

Visualizing Russian: Teaching Vocabulary at the Intersection of Frequency, Grammar, and Communication

October 28, 2021 — 1:30 pm - 2:30 pm

Steven Clancy

Harvard University

Visualizing Russian (<https://visualizingrussian.fas.harvard.edu>) is a suite of tools benefiting language learners, teachers, and researchers and enabling each user-group to access the complex system of the Russian language through visualization methods in order to leverage the powers of compression and expansion of a massive dataset. Tools developed so far analyze texts for relative difficulty with regard to vocabulary content, create frequency lists and display word forms used for each lemma in a text, compare target vocabulary to coverage in a text, identify the most commonly used word forms for a lemma, compare the relative frequency of near-synonyms or other items in a semantic domain, show the usage of nouns with various cases or verbs with various person/number/tense/aspect combinations, and identify the case governance or preposition usage of Russian verbs. Additional tools break down word-formation for prefixes-roots-suffixes in words and identify the field of morphologically-related or semantically-related words for any target word. These tools have contributed to the creation of a new Russian textbook series, Foundations of Russian (Clancy, Egorova, Green, Willis, forthcoming from Routledge Publishers), which presents a 4000-word vocabulary based on the most frequently occurring and communicatively necessary words in Russian.

Visualizing Russian project (Clancy 2014-2021)

Main site: <http://visualizingrussian.fas.harvard.edu/>

Tools in development: visualizingrussian-dev.fas.harvard.edu

General thoughts

- for a long time, we've had a desire to use authentic materials in language teaching materials and in the classroom
- yet such materials — whether simply written or audio-visual materials created by and for native speakers for communicative purposes or raw sentence examples taken from language corpora — remain difficult, particularly in languages such as Russian with vocabulary largely unrelated to English and with a high degree of morphological complexity
- notions like Stephen Krashen's "i+1" and comprehensible input suggest that created or adapted materials are useful for learners
- so we have a move from raw corpus materials to **corpus-based materials**

e.g., examples in the SMARTool, dialogues/text in *Min russiske reise*, dialogues, texts, examples in *Foundations of Russian*

- how can we use the quantitative tools and data sources that are familiar to linguists in the language classroom and in our creation of teaching materials?
- the tools we create should have a balance with regards to audience students—teachers—researchers

Visualizing Russian

Original Database and Development

- origins in hand-typed compiled database of various textbook vocabulary lists (*Live from Moscow* > *Live from Russia*, *Making Progress in Russian*, *Leaping into Russian*/Словарь в карьер, Начало, etc.) begun around 2009
- combined with lexeme and frequency information from Serge Sharoff's frequency lists based on the RNC (lemma frequency list, form frequency list)

Before the publication of the dictionary, Sharoff had made available various frequency lists based on the Russian National Corpus on his website:

<http://www.comp.leeds.ac.uk/sSharoff/frqlist/frqlist-en.html>

<http://corpus.leeds.ac.uk/serge/frqlist/>

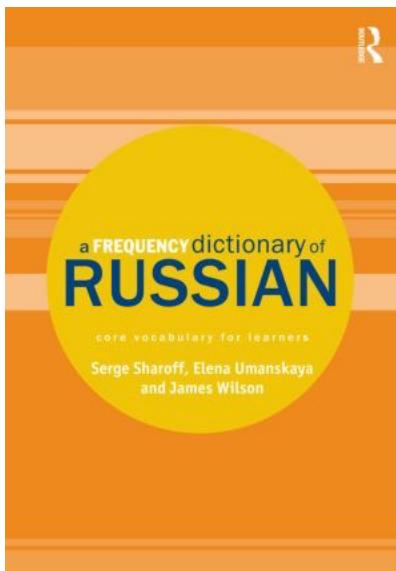
The exact files no longer seem to be available.

32617 lemmas with frequency, part of speech information

69307 word forms with frequency information

Our database was informed by the frequency information in these lists, but is an amalgam of many sources of information about Russian at present.

- Sharoff later published part of this data in 2013 as *A Frequency Dictionary of Russian* (but with only 5000 entries); provides **static lemma and form frequency information** in our database



A Frequency Dictionary of Russian

core vocabulary for learners

By **Serge Sharoff, Elena Umanskaya, James Wilson**
2013

A Frequency Dictionary of Russian is an invaluable tool for all learners of Russian, providing a list of the 5,000 most frequently used words in the language and the 300 most frequent multiword constructions.

A CD version is available to purchase separately. Designed for use by corpus and computational linguists it provides the full text in a format that researchers can process and turn into suitable lists for their own research purposes.

<https://www.routledge.com/products/9780415521420>

- scraped morphological forms from ru.wiktionary.org (not complete, but very thorough in forms incorporated c. 2017)
- now working on ways to incorporate more **dynamic lemma and form frequency information** directly from the full RNC and/or the downloadable portion of the RNC
- the vocabulary database of 33,000+ words (lexemes|lemmas) and phrases (set multi-word expressions and some common collocations (e.g., день рождения, домашнее задание))

core — targets ~1500 lexemes (1884), most frequent vocabulary for the most part
foundations — builds another ~2500 lexemes (3364), again informed by frequency, but also by pedagogical/communicative goals

expansions — next most frequent 15K words in the database

specializations — least frequent 12.5K words in the database

fifth level — items not in the database — mostly proper nouns, neologisms, slang, recent borrowings

- What do do with a database of (currently) 33225 lexeme entries (lemmas and some multi-word expressions, collocations) and more than 300 fields for each entry? How do make that quantitative information available, useful, and relevant

for **students, teachers, researchers**

Early inspirations from...

- developments with new HTML-5, JavaScript, CSS
even almost a decade later, this is a very impressive and inspiring piece technologically:
<https://www.nytimes.com/projects/2012/snow-fall/index.html#/?part=tunnel-creek>
D3 visualization library
<https://observablehq.com/@mbostock>
<https://observablehq.com/@d3/gallery>
- online interlinear readers and dictionaries
<https://nodictionaries.com>
<https://interlinearbooks.com>
- concordancing tools such as Laurence Anthony's AntWordProfiler
<http://www.laurenceanthony.net/software/antwordprofiler/>
- interactive readers with community commentary
<https://www.coerll.utexas.edu/ecomma/demo>
- the work and play of “internet polyglot entrepreneurs”
Steve Kaufman of YouTube and Lingq.com
<https://www.youtube.com/c/TheLinguist/featured>
<https://www.lingq.com/>
[great for the Krashen-inspired “just listen and read” approach to language acquisition]
[but not great for morphologically complex languages, e.g. Эбода, Эболу, Эболы all treated as separate words rather than separate **word forms** of a single **lemma/dictionary entry**]

Visualizing Russian Tools (overview)

- suite of tools developed and in development
 - Visible Vocabulary
 - Mini-Story Creator
 - Quick Lemma Tool
 - Word Formation Tool
 - Case Distribution Tool
 - Verb Form Distribution Tool (in development)
<https://eelegiap.github.io/verbradarcharts/>
 - Wordburst: Word Formation Tool (словообразование)
<https://eelegiap.github.io/wordburst/>
 - Similarity Tool
 - 3-D graphs of word embeddings from vectorization of navec data
Natasha (navec) developed from Librusec corpus (<https://librusec.pro>)
<https://github.com/natasha/navec>
чай: 25
https://eelegiap.github.io/three_d/three_d/chai25.html
and 41
https://eelegiap.github.io/three_d/three_d/chai41.html
- male vs female: https://eelegiap.github.io/three_d/three_d/index.html
- countries and capitals: https://eelegiap.github.io/three_d/three_d/countrycapital.html
- numbers: https://eelegiap.github.io/three_d/three_d/numbers.html
- utensils: https://eelegiap.github.io/three_d/three_d/utensils.html
- similarity network based on чай
<https://eelegiap.github.io/similarity-network/>
- hierarchical word tree for литература
<https://eelegiap.github.io/wordtreedemo/>
- other tools in the works!

In development:

- Verb histograms with surrounding case and prep+case information
<https://eelegiap.github.io/verbhistograms/>

Examples

analyze texts for relative difficulty with regard to vocabulary content, create frequency lists and display word forms used for each lemma in a text, compare target vocabulary to coverage in a text, identify the most commonly used word forms for a lemma, compare the relative frequency of near-synonyms or other items in a semantic domain, show the usage of nouns with various cases or verbs with various person/number/tense/aspect combinations, and identify the case governance or preposition usage of Russian verbs. Additional tools break down word-formation for prefixes-roots-suffixes in words and identify the field of morphologically-related or semantically-related words for any target word.

Visualizing Russian tools

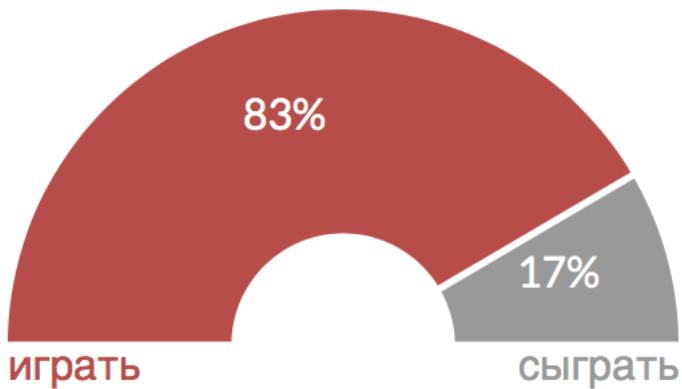
Visible Vocabulary text parser/analyzer

- cut-and-paste any text, receive an analysis by word level as conceived for the *Foundations of Russian* textbook series
- four-levels of vocabulary difficulty, percentage of words at each level, clickable information on words with bonus features (e.g., stress patterns in nouns, aspectual ratios with verbs, and more such features to come)
- can read online or cut-and-paste to Word or other programs
- amazing job of analyzing any text; catches almost everything; can do large amounts of text comparable to about 1/4 of *The Brothers Karamazov* at one time

students	teachers	researchers
<ul style="list-style-type: none"> • target vocabulary learning (early levels green and blue words; later levels tend to be the domain-specific words, technical jargon, etc. of various fields) • know what to pay attention to and what to ignore at various levels of study without a native speaking teacher/tutor to help • know how appropriate or difficult a particular reading will be • do more reading and spend more time with the target language 	<ul style="list-style-type: none"> • great for making handouts (what should students know at particular levels of study; what words need to be glossed) • excellent tool for creating and choosing written materials for use in class or your courses • keeps you focused on what students need at particular levels of study • great for curating and selecting texts for students at various levels 	<ul style="list-style-type: none"> • great for text analysis (difficult levels, content) • converts texts from word forms to lemmas (good for frequency lists)

- works with brute-force database of all morphological forms we could assemble
- text parsing deals with...
 - hyphenated items as separate words (француженка-гувернантка) except for high-frequency items (кто-то, что-нибудь, etc.)
 - includes an increasing number of multi-word expressions (MWEs), потому что or потому, что после того, как or после того как как раз, только что, на всякий случай
- **these are (or will be!) all considered single items as fixed MWEs**
- we plan to be able to deal with common collocations with inflected parts

- день рождения (with all inflected forms of день)
- домашнее задание (with all inflected forms of the adj-noun phrase)
- disambiguation is a function of frequency (homonyms are sometimes found)
 - МГУ is seen as Asg of мга,
 - Маша, Машу, М уни gets picked up as verbal adv of махать, машу as 1sg, М уни as imperative
 - Путин as Gpl of путына
 - Ломоносов, Ломоносове as ломонос 'clematis'
 - Юля as verbal adverb of юлить 'fuss, fidget, play up to'
- the features in this tool were developed first and form the backbone of related tools and spin-off ideas
- elements of what are separate tools at present will be combined over time into single tools that perform in multiple areas



Taking into account aspectual usage

Lemma	Occurrences in RNC	Documents in RNC	Aspectual Usage
играть	89401	18045	82%
сыграть	16436	7063	15%
поиграть	3318	1871	3%
петь	47846	10067	90%
спеть	5140	2380	10%
попеть	463	350	1%
танцевать	13280	4054	94%
станцевать	292	215	2%
потанцевать	537	406	4%

Mini-Story Creator

- idea was to create a tool for writing a new text or analyzing an existing text that compares that text to a set of target vocabulary
- tool for creating mini-stories (small contextualized examples of language in use) from a target set of lexical items
- also serves to convert any text into a frequency list generator
- shows the specific word forms used and their counts for the lexemes in the text

students	teachers	researchers
<ul style="list-style-type: none"> • write essays on specific prompts with target vocabulary • keep track of what vocabulary you've used 	<ul style="list-style-type: none"> • create pedagogical texts and dialogues for students while keeping focused on vocabulary targets • keep track of used/unused vocabulary • check existing materials for how well they cover target vocabulary 	<ul style="list-style-type: none"> • convert texts to frequency lists • examine contents of texts by word forms and counts for each lemma in a text

Quick Lemma Tool

- quickly access form and lemma frequency information in the database and in the RNC
- compare frequencies for various lemmas (especially in particular semantic domains, near-synonyms, etc.)
- generate and compare frequencies of word forms for a lemma
- Ex. читать (especially with all the forms we shouldn't expect to see, but do find in the RNC)
- additional comparison of forms in Google Books n-gram viewer as an embedded tool
<https://books.google.com/ngrams>

students	teachers	researchers
<ul style="list-style-type: none"> • generate the forms of inflected words • compare their frequency and see the most important forms of the word for learners 	<ul style="list-style-type: none"> • identify most important forms for particular lemmas • compare frequencies for items in particular semantic domains • compare frequencies for near-synonyms 	<ul style="list-style-type: none"> • compare different lemmas by frequency and use with respect to RNC data and Google Books n-gram viewer • see the most frequent forms for a given lemma

<i>lexical sets</i>
понедельник среда четверг пятница суббота воскресенье
январь февраль март апрель май июнь июль август сентябрь октябрь ноябрь декабрь
футбól хоккéй баскетбól бейсбól тéннис гóльф волейбól бадминтón ракетбól réгби лакróсс сквóш софтбól фрýзби шáхматы шáшки брайдж билья́рд бóулинг дárтс крокéт
гитáра скрипка áльт виолончéль контрабáс валтóрна трубá тромбóн тýба гобóй кларнéт саксофон фагóт флéйта форте́йно рояль пианино барабáн
<i>near-synonyms</i>
офис кабинет контора
штаны брюки джинсы
фабрика завод комбинат
<i>aspectual comparisons</i>
играть сыграть поиграть петь спеть попеть танцевать станцевать потанцевать
<i>domains—FEAR</i>
страх, страшный, страшно, опасаться, пугаться, пугать, бояться, напугать, испугать, пугнуть, испугаться, ужас, ужасный, ужасно, спугивать, спугнуть, запугивать, запугать, испуг

Noun Case Distribution Tool

- spider chart/radar chart view of a noun's case usage profile
- based on downloadable, morphologically-tagged version of the RNC
- some words are better than others based on counts and coverage in the RNC
- attempts to automate based on forms alone, but syncretism can't yet be accounted for

Verb Form Distribution Tool (in development)

- spider chart/radar chart view of a verb's tense usage profile
- includes aspect ratio fuel gauge graph (impf/pf frequency)
- basic view (5):
infinitive, conjugated forms, imperative, verbal adv/adj, past
- detailed view (13):
infinitive, 1sg, 2sg, 3sg, 1pl, 2pl, 3pl, imperative, verbal adv/adj, past m f n pl
- based on downloadable, morphologically-tagged version of the RNC
- collapse/expand to show basic or extended verbal categories
- both tools provide handy visualizations of the semantic space of the inflected items
- nouns overlap with sg and pl maps
- verbs overlap with impf and pf maps

students	teachers	researchers
<ul style="list-style-type: none"> • students see which case or tense/aspect categories are most relevant for individual words • what are the most important forms to know and why 	<ul style="list-style-type: none"> • can take into account the case and tense/aspect profiles of vocabulary in teaching and creation of materials • can show these items to students in class 	<ul style="list-style-type: none"> • can compare the case and tense/aspect profiles of various words • can look for semantic similarity in similar profiles • can see what forms are most prominent for each lexical item

In development:

Verb histograms with surrounding case and prep+case information

<https://eelegiap.github.io/verbhistograms/>

Word Formation Tool

- view an interactive word network surrounding individual Russian roots
- view by level or by POS
- interactive D3 visualization lets you move words around
- frequency indicated by distance from the central node (the root)

students	teachers	researchers
<ul style="list-style-type: none"> • learn about roots and word formation in Russian • learn related words to words you already know • control the volume of words through level parameters and POS settings • see translations for green and blue words 	<ul style="list-style-type: none"> • view roots and related word forms • increase relevance for students by controlling for level or POS • consider vocabulary for deepening and expanding your students' knowledge of Russian vocabulary and word formation 	<ul style="list-style-type: none"> • easily view related Russian words by roots • see the relative frequency of the various words

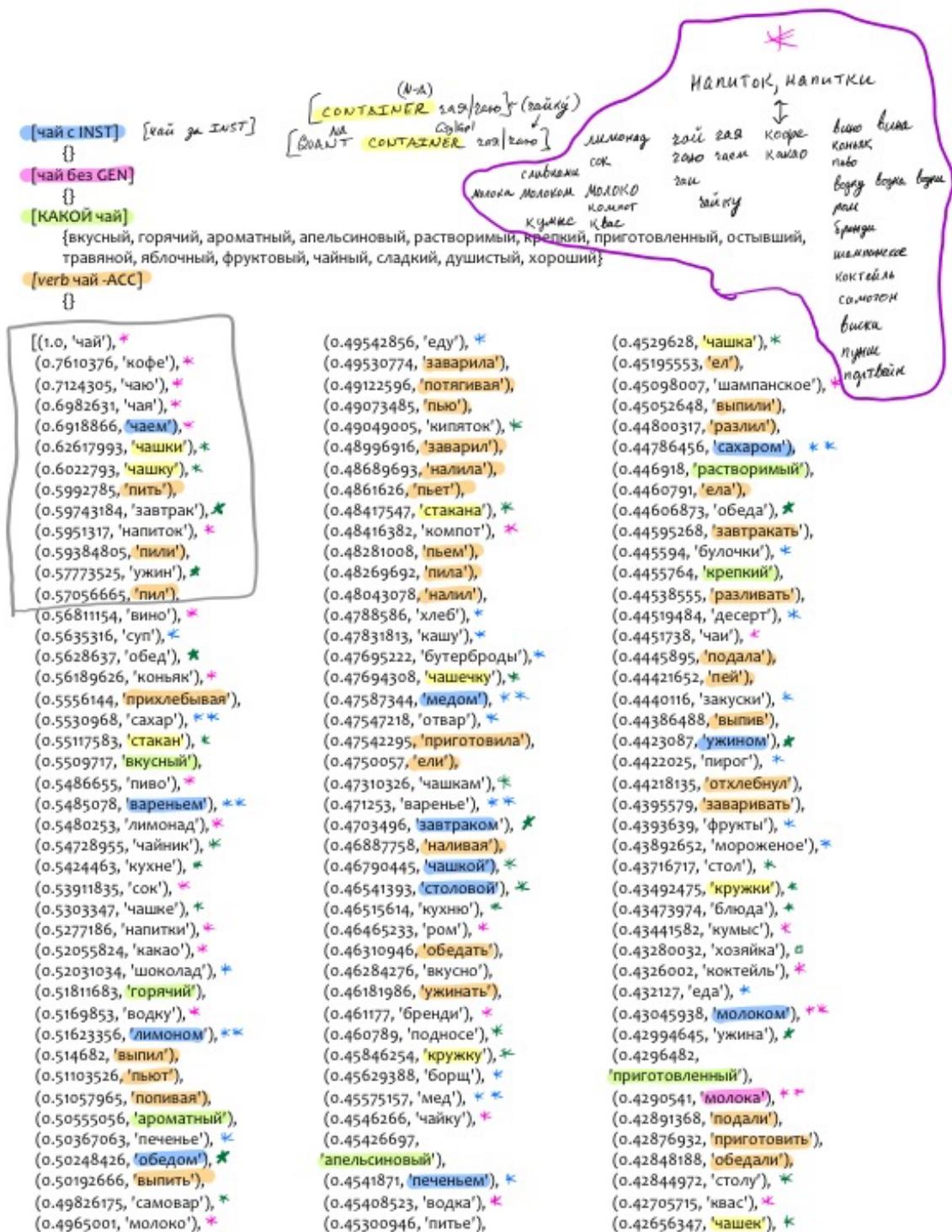
Wordburst: Word Formation Tool (словообразование) — under development

- enter a word and the tool splits the word up into prefix(es)—root(s)—suffix(es)
- engage students in word formation and analysis

Similarity Tool

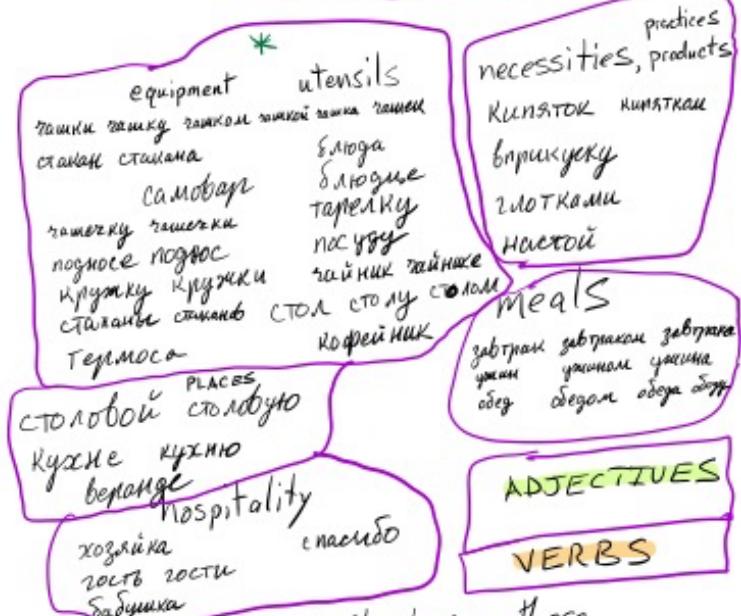
- 3-D graphs of word embeddings from vectorization of navec data
Natasha (navec) developed from Librusec corpus (<https://librusec.pro>)
<https://github.com/natasha/navec>
- explorations reveal a set of related words with very interesting semantic connections
- чай

manual analysis (on following pages)



(0.4265178, 'пирожные'), *
 (0.42636424, 'кушать'),
 (0.42629275, 'водки'), *
 (0.42494076, 'поднос'), *
 (0.4243299, 'посуду'), *
 (0.42381668, 'разливая'),
 (0.4235259, 'вприкуску'), *
 (0.42299774, 'глотками'), *
 (0.4226973, 'допил'),
 (0.42258814, 'заваривает'),
 (0.42243925, 'готовить'),
 (0.42241824, 'кофейник'), *
 (0.4205355, 'оставший'),
 (0.4193354, 'стаканы'), *
 (0.41854712, 'каша'), *
 (0.4178444, 'потягивал'),
 (0.417272, 'гости'), □
 (0.41660762, 'бабушка'), □
 (0.4160004, 'травяной'),
 (0.4157949, 'подавали'),
 (0.41503167, 'самогон'), *
 (0.414775, 'приготовлю'),
 (0.41391462, 'виски'), *
 (0.41354406, 'термоса'), *
 (0.41318315, 'выпила'),
 (0.413109, 'сливками'), **
 (0.41262156, 'салат'), *
 (0.4124164, 'стаканов'), *
 (0.4114, 'полить'),
 (0.41113812, 'яблочный'),
 (0.41071177, 'принесла'),
 (0.41067833, 'едят'),
 (0.41065747, 'едой'), *
 (0.410259, 'пунш'), *
 (0.4102473, 'приготовил'),
 (0.4099795, 'конфеты'), *
 (0.4099773, 'фруктовый'),
 (0.40991524, 'чайный'),
 (0.40915298, 'сыр'), *
 (0.40817428, 'сахара'), **
 (0.40785566, 'столовую'), *
 (0.4070255, 'веранде'), *
 (0.40693328, 'настой'), *
 (0.40690854, 'портвейн'), *
 (0.4067746, 'гости'), □
 (0.40667805, 'заваривала'),
 (0.40613306, 'спасибо'), □
 (0.4060068, 'тарелку'), *
 (0.40587488, 'пьете'),
 (0.40560496, 'масло'), *
 (0.4055304, 'завтрака'), *
 (0.4044205, 'пирожными'), *
 (0.40429533, 'пирожки'), *

(0.4040651, 'вины'), *
 (0.40403584, 'ест'),
 (0.40345475, 'сладкий'),
 (0.403358, 'кинжалом'), *
 (0.4032641, 'прихлебывал'),
 (0.40273315, 'душитый'),
 (0.4021364, 'столом'), *
 (0.40211138, 'обеду'), *
 (0.40133402, 'заказал'),
 (0.4013239, 'чайнике'), *
 (0.40091068, 'хороший'),
 (0.40074712, 'чашечки'), *
 (0.40044048, 'блюдце')] *



* one misses the opportunity to see these words in context

* we really only have CXX in context and those words merely suggest those CXX and contexts; we'd still want confirmation from the corpus

* there isn't really anything here that isn't relevant (with the exception of a few oddities, probably due to the corpus)
 (кумыс, брюксука)

- * I wonder what the relationship here is between cosine similarity and frequency?
- * for instance, the brown boxed words greater than .57 are all great and seem like exactly what verbs you'd need, but not much further and you get нужно зубар 'sipping' which will be rare as a verb lemma and rare as a verb
form but maybe the only time it was used in the corpus was with rain, so a very strong correlation

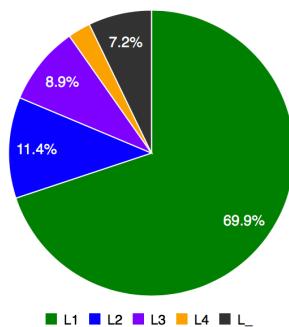
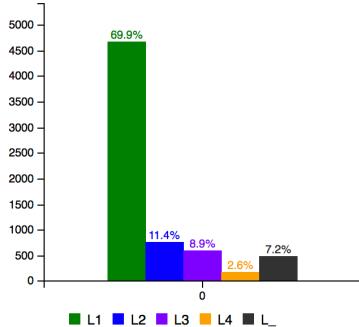
- 3-D charts of чай
чай: 25
https://eelegiap.github.io/three_d/three_d/chai25.html
and 41
https://eelegiap.github.io/three_d/three_d/chai41.html
- other examples
male vs female: https://eelegiap.github.io/three_d/three_d/index.html
countries and capitals: https://eelegiap.github.io/three_d/three_d/countrycapital.html
numbers: https://eelegiap.github.io/three_d/three_d/numbers.html
utensils: https://eelegiap.github.io/three_d/three_d/utensils.html
- similarity network based on чай
<https://eelegiap.github.io/similarity-network/>
- hierarchical word tree for литература
<https://eelegiap.github.io/wordtreedemo/>
- other tools in the works!

Visualizing Russian Tools in Action**Visible Vocabulary in Action**

I

Все счастливые семьи похожи друг на друга, каждая несчастливая семья несчастлива по-своему.
 Все смешалось в доме Облонских. Жена узнала, что муж был в связи с бывшую в их доме француженкой-гувернанткой, и объявила мужу, что не может жить с ним в одном доме. Положение это продолжалось уже третий день и мучительно чувствовалось и самими супругами, и всеми членами семьи, и домочадцами. Все члены семьи и домочадцы чувствовали,

Parse Text **Clear**



Word Count: 6671

Unparsed Count: 478

L1 Count: 4663

L2 Count: 762

L3 Count: 596

L4 Count: 172

Combine L1 and L2**Parsed Text****U Lemmas****Word Info**[Лев Николаевич Толстой](#)[Анна Каренина](#)[Мне отмщение, и аз воздам](#)[ЧАСТЬ ПЕРВАЯ](#)

Figure 1: Analysis of Tolstoy's Anna Karenina reveals a text with 81% comprehensible vocabulary to a student who has completed our intermediate Russian level with *Foundations of Russian*.

Parsed Text

U Lemmas

Word Info

Лев Николаевич Толстой

Анна Каренина

Мне отмщение, и аз воздам

ЧАСТЬ ПЕРВАЯ

I

Все счастливые семьи похожи друг на друга, каждая несчастливая семья несчастлива по-своему.

Все смешалось в доме Облонских. Жена узнала, что муж был в связи с бывшую в их доме

француженкою-губернанткой, и объявила мужу, что не может жить с ним в одном доме. Положение это продолжалось уже третий день и мучительно чувствовалось и самими супругами, и всеми членами семьи, и домочадцами. Все члены семьи и домочадцы чувствовали, что нет смысла в их сожительстве и что на каждом постоялом дворе случайно сошедшиеся люди более связаны между собой, чем они, члены семьи и домочадцы Облонских. Жена не выходила из своих комнат, мужа третий день не было дома. Дети бегали по всему дому, как потерянные; англичанка поссорилась с экономкой и написала записку приятельнице, прося приискать ей новое место; повар ушел еще вчера со двора, во время обеда; черная кухарка и кучер просили расчета.

На третий день после ссоры князь Степан Аркадьевич Облонский -- Стива, как его звали в свете, -- в обычный час, то есть в восемь часов утра, проснулся не в спальне жены, а в своем кабинете, на сафьяновом диване... Он повернул свое полное, выхоленное тело на пружинах дивана, как бы желая опять заснуть

Объявила

Lemma: **объявить**
 Parts of Speech: **verb**
 Aspect: **perfective**
 Levels: **21**
 Inflections: **Fpast**
 Translation: **announce**
 Stress Pattern: **n/a**
 Frequency:



Figure 2: Text may be exported to retain color-coding and printed or used electronically. Text in the online reader is clickable for further information such as a verb tool showing the prevalence of imperfective vs. perfective aspect for a given verb.

"Ах, ах, ах! Аах!.." -- замычал он, вспоминая все, что было. И его воображению представились опять все подробности ссоры с женой, вся безвыходность его положения и мучительнее всего собственная вина его.

"Да! она не простит и не может простить. И всего ужаснее то, что виной всему я, виной я, а не виноват. В этом-то вся драма, -- думал он. -- Ах, ах, ах!" -- приговаривал он с отчаянием, вспоминая самые тяжелые для себя впечатления из этой ссоры.

Неприятнее всего была та первая минута, когда он, вернувшись из театра, веселый и довольный, с огромной грушей для жены в руке, не нашел жены в гостиной; к удивлению, не нашел ее и в кабинете и, наконец, увидел ее в спальне с несчастною, открывшою все, запиской в руке.

Она, эта вечно озабоченная, и хлопотливая, и недалекая, какою он считал ее, Долли, неподвижно сидела с запиской в руке и с выражением ужаса, отчаяния и гнева смотрела на него.

-- Что это? это? -- спрашивала она, указывая на записку.

И при этом воспоминании, как это часто бывает, мучала Степана Аркадьевича не столько самое событие, сколько то, как он ответил на эти слова жены.

С ним случилось в эту минуту то, что случается с людьми, когда они неожиданно уличены в чем-нибудь слишком постыдном. Он не сумел подготовить свое лицо к тому положению, в которое он становился перед женой после открытия его вины. Вместо того чтобы оскорбиться, отрекаться, оправдываться, просить прощения, оставаться даже равнодушным -- все было бы лучше того, что он сделал! -- его лицо совершенно невольно ("рефлексы головного мозга", -- подумал Степан Аркадьевич, который любил физиологию), совершенно невольно вдруг улыбнулось привычно, доброю и потому глупою улыбкой.

Эту глупую улыбку он не мог простить себе. Увидав эту улыбку, Долли вздрогнула, как от физической боли, разразилась, со свойственною ей горячностью, потоком жестоких слов и выбежала из комнаты. С тех пор она не хотела видеть мужа.

"Всему виной эта глупая улыбка", -- думал Степан Аркадьевич.

руке

Lemma: **рука**

Parts of Speech: **noun**

Levels: **1E**

Inflections: **LOCsg, DATsg**

Translation: **hand; arm**

Stress Pattern: **UM**

[UM]

N	—	’	N	—	’	—
A	—	’	A	—	’	—
G	—	’	G	—	’	—
L	—	’	L	—	’	—
D	—	’	D	—	’	—
I	—	’	I	—	’	—

Figure 3: Noun analyses make available information on stress patterns along with other grammatical details about the noun.

Mini-Story Creator in action**Enter Mini-story:**

Ли́зе бо́льше всего́ нра́вится ходи́ть в теа́тр. В Мáлом драмати́ческом теа́tre идёт но́вый спектáкль, но очень тру́дно на́йти билéты. Она кáждый день ходи́ла в театрálную кáссу, но билéтов не́ было. Вдруг в сре́ду подрúга сказáла ей, что у неё есть ли́шний билéт.

В спорту́зле есть бассéйн. Я очень люблю́ плáвать и там можна́ занимáться плáванием в любо́е вре́мя гóда, даже зи́мой. Наша семья́ туда ча́сто ходит. Я серьёзно плáваю пол ча́са, потóм купáюсь с детьми. Мы ли́бум проводи́ть свобóдное вре́мя вме́сте в спорту́зле.

Generate word list Visualize levels

Check Story Vocabulary

Vocabulary:

число | числа цифра считать | сосчитать
числительное | числительные чётное
(число) нечётное (число) много (GENpl)
многий, многие немногий, немногие оба
| обе полтора |

Story lemmas (248):

Lemma	Forms (in order of appearance)	Count	Level
в	в(x22), во	23	1
и	и(x16)	16	1
на	на(x14)	14	1
она	она(x6), ей, неё, её	9	1
он	него, он(x5), ему(x2), нему	9	1
они	они(x7), них, их	9	1
что	что(x8)	8	1
с	с(x7)	7	1
не	не, не(x5)	6	1
быть	было(x2), была, был, будет, быть	6	1
играть	играли, играть(x3), играет(x2)	6	1
ру́сский	ру́сским(x3), ру́сская, ру́сский, ру́сской	6	1

Vocabulary usage (314):

Word	Used?	Matching Lemma(s)
американский	<input checked="" type="checkbox"/>	американский
бар	<input checked="" type="checkbox"/>	барин
в	<input checked="" type="checkbox"/>	во
вечер	<input checked="" type="checkbox"/>	вечер
вечера	<input checked="" type="checkbox"/>	вечер
вечером	<input checked="" type="checkbox"/>	вечер
во	<input checked="" type="checkbox"/>	в
вторник	<input checked="" type="checkbox"/>	вторник
вторникам	<input checked="" type="checkbox"/>	вторник
вчера	<input checked="" type="checkbox"/>	вчера
вчерашний	<input checked="" type="checkbox"/>	вчерашний

Figure 4: Mini-Story Creator for frequency lists.

как	как(x556)	556	1
князь	князь(x382), князей(x2), князю(x54), князя(x85), князем(x9), князья(x2), князе(x2)	536	3
по	по(x239)	239	1
генерал	генерала(x36), генералом(x6), генерал(x173), генералу(x13), генералов	229	3
смешной	смешно(x3), смешной, смешное(x2), смешными, смешные, смешон(x2), смешному, смешным(x3), смешная	15	1
прочесть	прочтите(x2), прочли(x3), прочесть(x5), прочел(x3), прочтено, прочту	15	3
сойти	сойдет, сошел(x2), сошла(x10), сошли, сойдешь	15	2
святотатство	святотатство(x4)	4	4
потихоньку	потихоньку(x4)	4	3
раздражить	раздражил, раздражена, раздражить, раздражен	4	2
соображать	соображая, соображал(x3)	4	3
княжна	княжна(x2), княжны(x2)	4	3
присутствовать	присутствует, присутствующих, присутствовали(x2)	4	2
карета	карету(x3), карета	4	4
отличаться	отличались(x2), отличалась(x2)	4	2
никуда	никуда(x4)	4	1
выбрать	выбрал, выберите, выбрали, выбран	4	1

Figure 5: Frequency Lists and Color Levels indicate words to target for learners at different stages and for different purposes (text sample from beginning chapters of Dostoevsky's Идиот).

Quick Lemma Tool in Action**Quick Lemma Tool**

офис
контора
кабинет

Lemma: e.g. Compare русский vs. российский

Inflection: e.g. Compare the forms масло, масла, маслом

Get forms from lemma: e.g. Input улица → Output улица, улице, улицу, улицы, улицах...

Get RNC Counts: on forms in the [Russian National Corpus \(RNC\)](#)

[Get lemma frequencies](#) [Get form frequencies](#) [Get forms from lemma](#) [Get RNC counts](#)

Query data is from the Russian National Corpus, which has 321,712,061 words and 124,565 documents in total.

Word	Occurrences in RNC	Documents in RNC
кабинет	4642	16444
контора	1043	1676
офис	889	1622

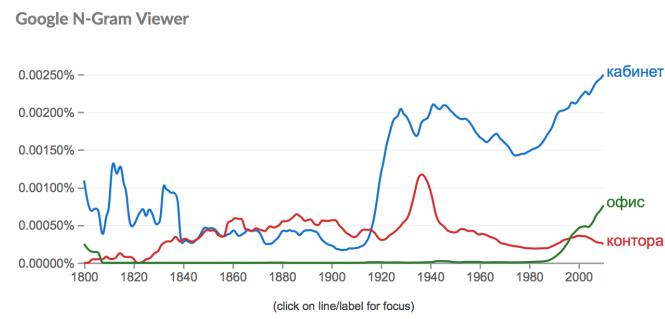


Figure 6: The Quick Lemma Tool allows for comparison of any set of words by frequency from the Russian National Corpus along with input from the Google Books N-Gram Viewer. For any noun, adjective, verb, you can also generate and analyze all forms of the paradigm.

Query data is from the Russian National Corpus, which has 321,712,061 words and 124,565 documents in total.

Word	Occurrences in RNC	Documents in RNC	Google N-Gram Viewer
нравится	6772	19247	
нравилось	2744	6010	
нравилась	1743	3093	
нравился	1725	3031	
нравятся	1960	2995	
нравились	1476	2386	
нравиться	1391	2164	
нравлюсь	377	515	
нравишься	293	396	
нравитесь	289	358	
нравимся	48	58	
нравься	None found	None found	
нравьтесь	None found	None found	
нравящееся	27	27	
нравящихся	25	26	
нравившиеся	21	23	

Figure 7: Form frequencies generated by the tool and referenced to the database or to the Russian National Corpus reveal the grammatical/semantic profile of the verb, e.g. here that нравиться is a strongly 3rd-person verb.

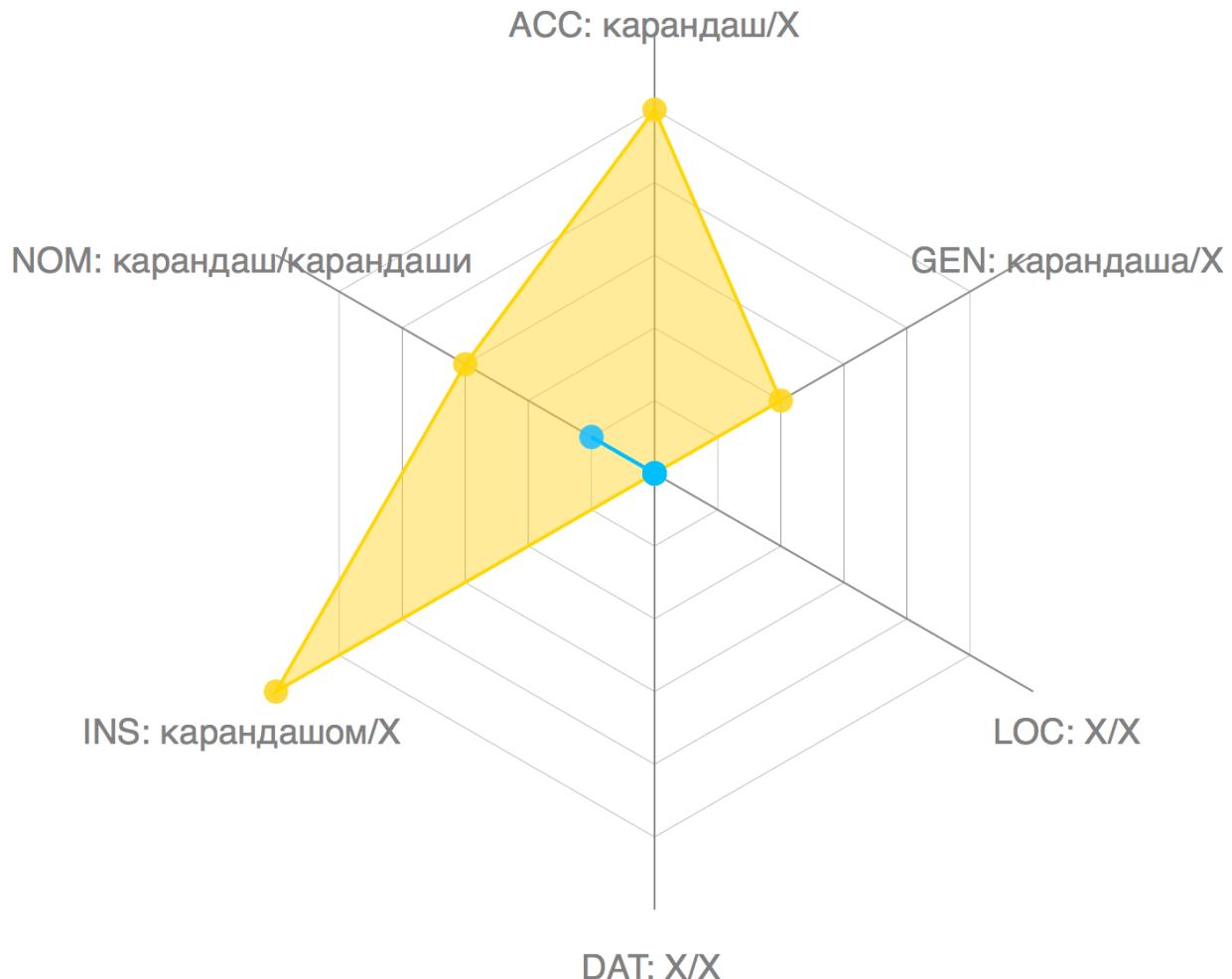
Case Distribution Tool in action

Figure 8: Detailed Sample of a Spider/Radar chart visualization of case semantic space. Here we see that карандаш ‘pencil’ has a strong INST association and a secondary association with the ACC case.

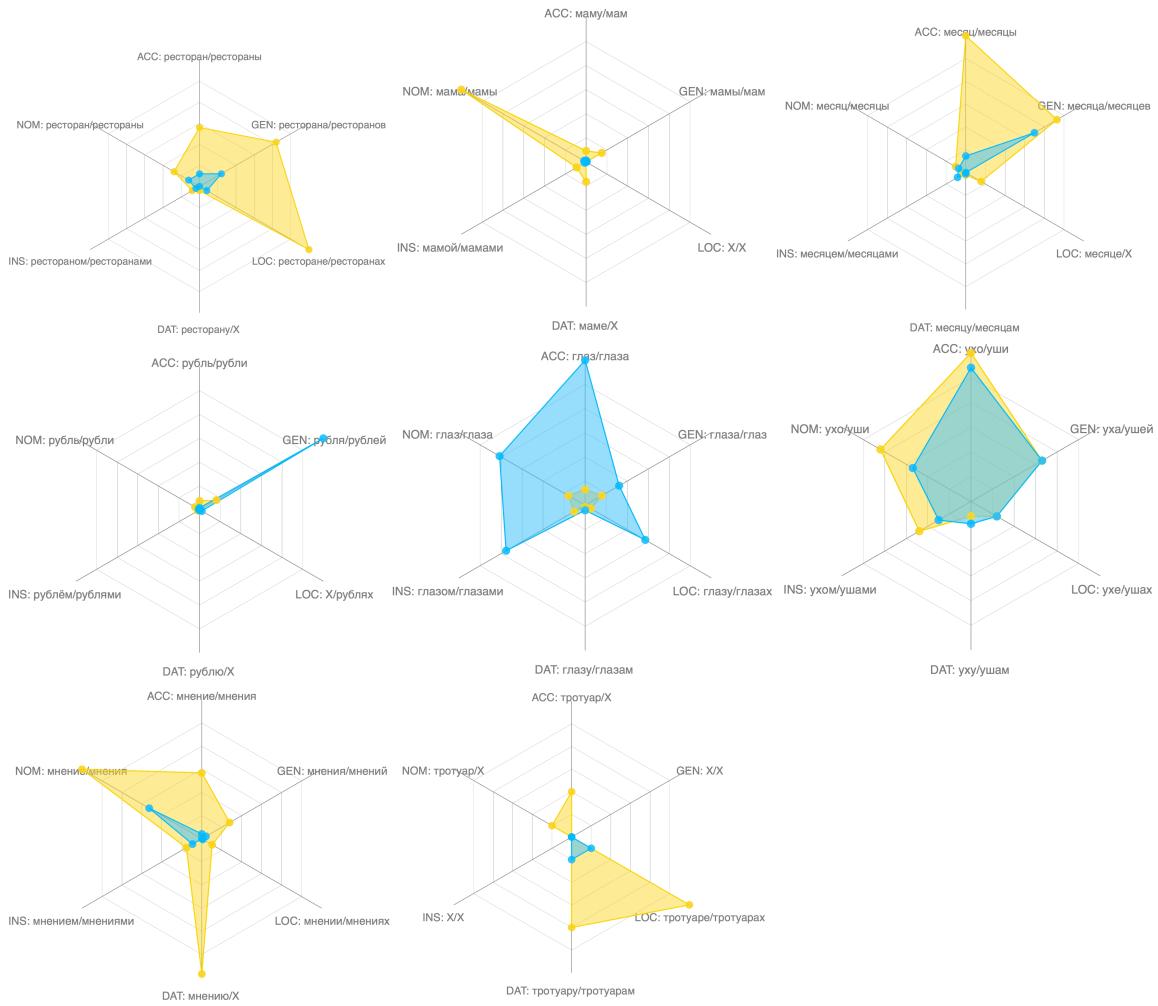


Figure 9: Several shots of spider/radar semantic case maps of words showing traits such as strong LOC association (ресторан ‘restaurant’), strong NOMsg subject association (мама ‘mom’), strong ACC/GEN association (time expression and counted noun месяц ‘month’), strong GENpl association (money рубль ‘ruble’), strong plural association (глаз|глаза ‘eye(s)’), balanced sg/pl profile (ухо|ухи ‘ear(s)’), extreme DAT association (мнение ‘opinion’), strong DAT/LOC association (тротуар ‘sidewalk’).

Compression and Expansion

The Russian language as a complex, but largely rule-governed linguistic system, presents an overwhelming variety of grammatical forms playing out along multiple semantic dimensions and presenting a seemingly insurmountable task for the language learner. The focus on most frequent forms in the SMARTool allow students to compress rather than expand the paradigms and morphological complexity of Russian.

<https://smartool.github.io/smartool-rus-eng/>

As both Janda and Tyers (2018) and Janda (2019) detail, Russian is highly-inflected, having 12 case/number slots for nouns, 28 case/number/gender slots for adjectives, and a very high number of tense/aspect categories inflected in verbs. See below for more details on word density in Russian vs. English.

Although we teachers of Russian only manage occasionally to offer exercises or classroom activities that focus solely on nominal declension or verb conjugation (at least in my experience), beginning learners especially occupy themselves with learning declension and conjugation patterns. As the demands on students merely to learn the meanings of new lexical items increase, the time available to focus on declensions and conjugations decreases. Students clearly need a way to incorporate both the morphology and the semantics of their increasing exposure to the Russian grammatical—lexical continuum.

The SMARTool offers a focused way to see the most important forms for any inflected item and to offer a path to mastery of declension and conjugation patterns through the acquisition of these key forms.

students	teachers	researchers
<ul style="list-style-type: none"> learn how to get the most mileage out of paradigms they are learning can focus on the most important forms of new inflected words 	<ul style="list-style-type: none"> can focus on most important forms to teach students can include useful and level-appropriate examples in their teaching materials and class activities 	<ul style="list-style-type: none"> can see what the most frequent uses are for a representative sample of 3000 words at various levels of proficiency can gain a sense of constructions used across those lexemes

Word form density in a morphologically simple vs. complex language

For the **most frequent 100 words**, for instance, **Russian has 5 times as many related word forms than does English** and this most frequent layer of vocabulary in any language consists of forms such as prepositions, pronouns, particles and function words that tend to have one (*the, if, would; esli ‘if’, by ‘would’*) or only a few forms (*a|an; v|vo ‘in’, o|ob|obo ‘about’*). Once you get out into the wider swathes of verbs, adjectives, and nouns, consider that **an English verb has only 4-5 forms**, whereas **a Russian verb has 31-33 (15-17 forms across 2 aspects)** and that is low-ballng the actual number of forms, **neglecting all the case forms of the participles (adjectives), the imperfective budu-future (6 additional forms), conditionals (a particle added to the past tense forms), etc.**

Verb forms in English [4-5 forms per verb]

*walk, walks, walking, walked
take, takes, taking, took, taken*

Verb forms in Russian [>32-33 forms per verb]

čitat'	pročitat' 'read'
čitaju	pročitaju
čitaješ'	pročitaješ'
čitajet	pročitajet
čitajem	pročitajem
čitajete	pročitajete
čitajut	pročitajut
čital, čitala, čitalo	pročital, pročitala, pročitalo
čitali	pročitali
čitaj	pročitaj
čitajte	pročitajte
čitaja	pročitav
čitajuščij	
čitavšij	pročitavšij
čitajemyj	pročitanneyj

An English adjective has maybe three forms if we count comparative and superlative versions, whereas a Russian adjective has 12 distinct forms plus comparative and superlative constructions and that considers only the distinct forms, not all the separate gender and case forms we could analyze in any given usage (6 case forms for masculine, neuter, feminine, and plural plus four short-form adjectives). Unlike in some inflected languages (such as Latin), adjectives and nouns in Russian belong to quite different declensional patterns, adding to the morphological complexity of a phrase or sentence. **For the verb pair given above, čitat'|pročitat', for the imperfective verb alone we find 71 distinct verb forms attested in the Russian National Corpus (RNC).** A single verbal notion could have 140 verb forms without even considering compound forms (with an auxiliary verb in the imperfective future, or the

conditional particle). Although we routinely tell students that Russian only has two conjugations and about 5 irregular verbs, this masks the fact that there are 25 regular verb types belonging to these two conjugations and that there are really more than 5 irregular verbs.

So when we look at the neighbors of *walk*, *walks*, *walking*, *walked*, we only have 4 forms to somehow relate together under the lemma of *walk*, whereas Russian might have 70 forms for each verb. Here, we've stumbled onto another thorny area in Russian. Verbs of motion, where each basic type of motion (by foot/near distance, by vehicle/farther distance, carrying, leading, conveying by vehicle, flying, swimming/sailing, running, crawling, climbing, etc.) is represented by 3 basic verbs and 2 prefixable aspect forms that can combine with 15 separate motion prefixes.

Adjective forms in English (1)

interesting [more interesting, most interesting]
good [better, best]

Adjective forms in Russian (12)

(fewer than the $6 \times 4 = 24$ forms we might encounter)

interesnyj, interesnogo, interesnom, interesnomu, interesnym, interesnoe,
interesnaja, interesnuju, interesnoj, interesnye, interesnyx, interesnymi
[plus short forms *interesen, interesna, interesno, interesny*, comparative *interesnee*,
superlative compound forms in *samyj + ADJ*, and 12 forms of *interesnejšij* as another
superlative construction]

xorošij, xorošego, xorošem, xorošemu, xorošim, xorošee, xorošaja, xorošuju, xorošej, xorošie,
xorošix, xorošimi [*lučše, lučšij, samyj + ADJ*]

Noun forms in English (2-4) vs. Russian (7-10)

(fewer at least than the $6 \times 2 = 12$ forms we might expect to encounter)

<i>table; tables</i>	<i>stol, stola, stole, stolu, stolom;</i>
<i>(table's; tables')</i>	<i>stoly, stolov, stolax, stolam, stolami</i>
<i>dictionary, dictionaries</i>	<i>slovar', slovarja, slovare, slovarju, slovarjem;</i>
<i>(dictionary's, dictionaries')</i>	<i>slovary, slovarej, slovarjax, slovarjam, slovarjami</i>
<i>word, words</i>	<i>slovo, slova, slove, slovu, slovom;</i>
<i>(word's, words')</i>	<i>slov, slovax, slovam, slovami</i>
<i>field, fields</i>	<i>pole, polja, polju, polem;</i>
<i>(field's, fields')</i>	<i>polej, poljax, poljam, poljami</i>
<i>dog, dogs</i>	<i>sobaka, sobaku, sobaki, sobake, sobakoj; sobak,</i>
<i>(dog's, dogs')</i>	<i>sobakax, sobakam, sobakami</i>
<i>week, weeks</i>	<i>nedelja, nedelju, nedeli, nedele, nedelej; nedel',</i>
<i>(week's, weeks')</i>	<i>nedeljax, nedeljam, nedeljami</i>

*mouse, mice
(mouse's, mice's)
name, names
(name's, names')*

*myš', myši, myš'ju;
myšej, myšax, myšam, myšami
imja, imeni, imenem, imena, imen, imenax,
imenam, imenami*

*radio, radios
(radio's, radios')*

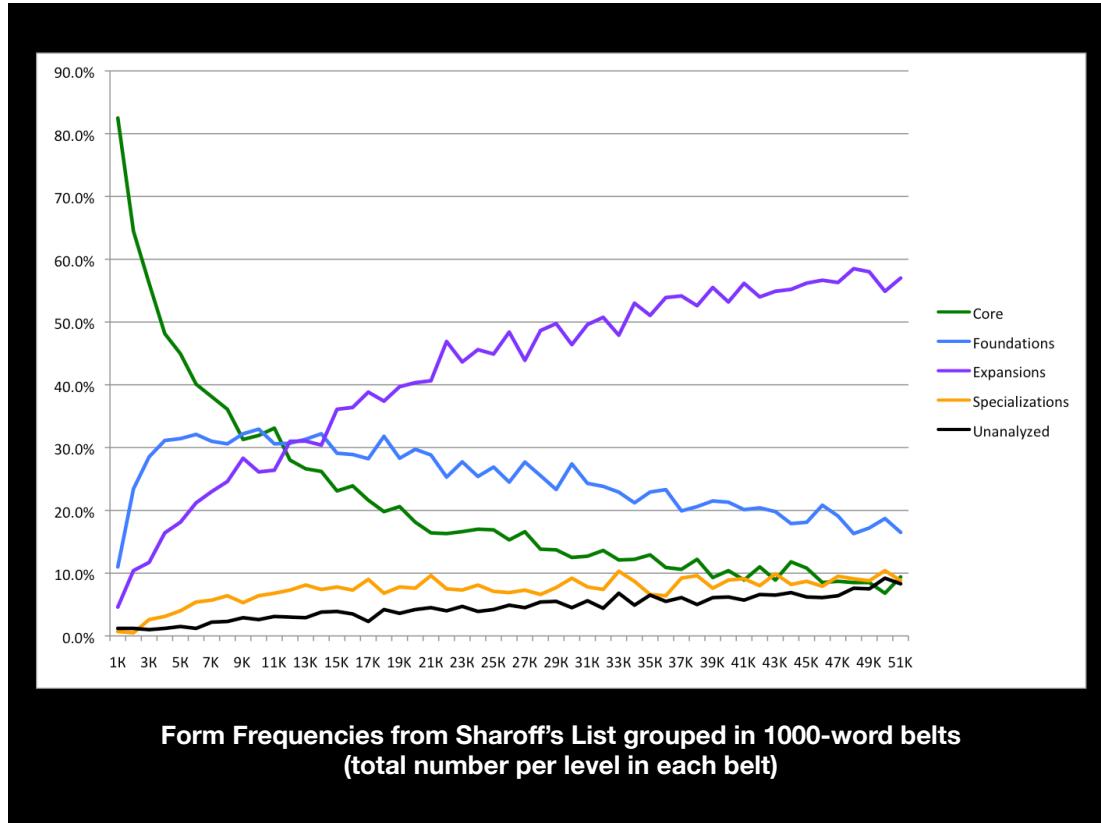
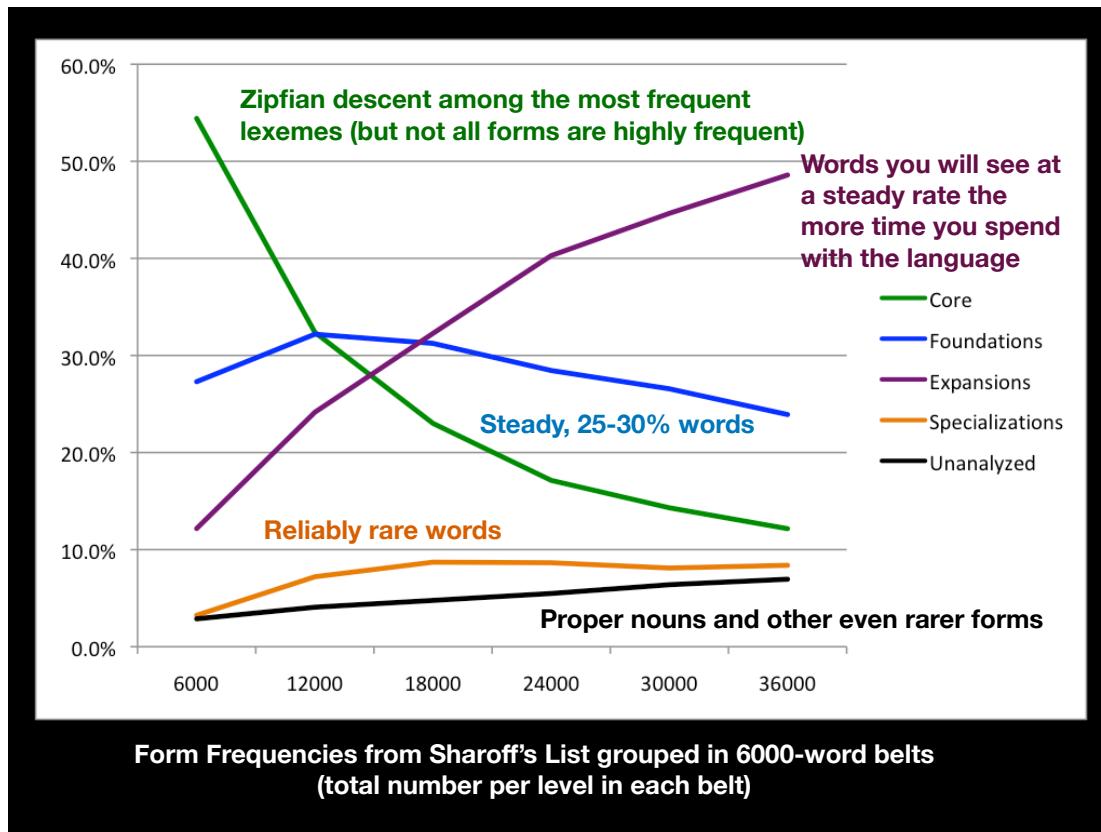
*...and there are some indeclinable nouns...
radio, kafe, kofe, kenguru, šimpanze, Kentuki,
etc.*

We can also consider that in addition to simplification by syncretism, there are also effects to be considered of incomplete paradigms, at least from a statistical perspective of usage. For any given word, there will be a profile of more intense usage among all the possible forms of the set (paradigm). See Janda and Tyers (2018).

Janda, Laura A., and Francis M. Tyers. 2018. “Less is more: why all paradigms are defective, and why that is a good thing”. *Corpus Linguistics and Linguistic Theory*, 14(2): 109-141. doi.org/10.1515/cllt-2018-0031.

Janda, Laura A. 2019. “Businessmen and Ballerinas Take Different Forms: A Strategic Resource for Acquiring Russian Vocabulary and Morphology”. *Russian Language Journal*, 69: 175-193.

NB: Incidentally, it would be interesting to look at the nature of the items in the 100-word belts used in the training and testing of models by Janda and Tyers (2018) in light of what I found below when looking at the four vocabulary levels we use in Visualizing Russian when compared to 6000-form or 1000-form belts of forms and which levels their lemmas belong to (see p. 29 below).



Russian Constructions and Word Formation

- compositional vs. non-compositional

sometimes the seemingly non-compositional is compositional (to a native speaker or an advanced student); sometimes the compositional is better left unanalyzed

Только что

как раз

после того, как

пешком

босиком

sometimes we can analyze and relate but the results are still not predictable

во время GEN ‘during’

вовремя ‘on time’

sometimes we do want our students to see the relationships between roots and related forms to the benefit of their learning

ухо|уши ‘ear’ > наушники ‘headphones’, подушка ‘pillow’

Visualizing Russian and Foundations of Russian textbook

Foundations of Russian: A cognitive and constructional approach to teaching Russian enriched by frequency data (working title)

Steven Clancy, Harvard University, <sclancy@fas.harvard.edu>

Daniel Green, Cambridge University, <dlcg3@cam.ac.uk>

Veronika Egorova, Harvard University, <egorova@fas.harvard.edu>

Oksana Willis, Bucknell University, <owoo3@bucknell.edu>

Foundations of Russian features ambitious exposure to the most frequent and communicatively necessary words in the Russian language. Based on **frequency of use, communicative needs of students, the consensus of textbooks** over the years, and our own **teaching experience**, we have identified four levels of vocabulary: **Core, Foundations, Expansions, and Specializations**.

The **Core** words comprise around 1500 lexical entries corresponding to our targeted vocabulary for introductory Russian. The **Foundations** level presents another estimated 2500 words and greatly expands and deepens the student's vocabulary.

The database behind our vocabulary levels as well as the online tools and website for the books currently contains 1893 items in the **Core** level and 3350 items in the **Foundations** level. We consider regular adverbs as part of the lemmas for adjectives (**хороший|хорошо́, бы́стрый|бы́стро, интересный|интересно**, etc.) and consider verbs as part of aspectual pairs (**чита́ть|прочитáть, запи́сывать|запи́сáть**) or triplets (**ходи́ть, идти́|пойти́, носи́ть, нести́|понести́**), thus bringing that number down to our target of approximately 4000 words. A student with mastery of these two levels will have a robust vocabulary of 4000 Russian words, or 5243 lexical items if we separate adjectives and adverbs and consider imperfective and perfective verbs separately.

The **Core** and **Foundations** levels are not exclusively based on frequency, but were chosen and are hand-curated by our team of authors. Although we do occasionally change our minds about what vocabulary belongs to the **Core** and **Foundations** levels, the items in these levels are stable and represent the 1500 most important words for beginning students to learn first and the most important 2500 words for continuing students to settle into as they move from intermediate to advanced students of Russian. A thorough mastery of these two levels puts the highest frequency words in Russian at the learner's disposal, providing them with the vocabulary necessary for critical communicative functions in any area of language use.

The remaining words students will encounter in Russian are considered **Expansions** (14.9K words) and **Specializations** (12.6K words), categorized largely according to decreasing frequency of occurrence.

The [Expansions](#) level comprises more nuanced general vocabulary and the domain specific terms and jargon used in various fields of human endeavor. These are the words a continuing learner will be exposed to over many years or a lifetime of exposure to the Russian language and well exceed our expectations for the organized curriculum a student might pass through in high school or university study. On a more positive note, many words in the [Expansions](#) level are cognates with English or are internationalisms across languages.

Decreasing frequency of occurrence continues into the [Specializations](#) level, but here we find words that are very low frequency.

A **fifth level** of vocabulary emerges in our online tools from words not contained in the database at all, revealed to be mostly proper nouns (place names, people, characters), neologisms, and slang. Although such words will be encountered in authentic materials that accompany this book, they are not the focus of this level and students can use the color categorization to focus their attention on more vital words.

The textbook strikes a balance between thorough grammatical coverage and treatment of thematic and communicative material. Grammar is neither ignored nor the principal motivating factor in the organization of the book.

Book 1

- 1 Мы ráды сно́ва вíдеть вас! [знакомство, учеба, университет]
- 2 Кудá ушлó лéто? [отпуск, каникулы, отдых, хобби, развлечения, досуг]
- 3 Чúвствуйте себя́ как дóма! [гостеприимство, еда, напитки, тосты]
- 4 Кто на свéте всéх милéе! [одежда, внешность; части тела]
- 5 Без труда́ нет добра́ [работа, характер]
- 6 Приятного аппетита! [приготовление еды, partitive GEN, verbs of SENSING]
- 7 В гостя́х хорошó, а дóма лúчше! [дом, квартира, сад, огород, покупки, услуги]
- 8 —Счастли́вого пúти! —Счастли́во оставáться! [туризм, летать, собирать чемодан]
- 9 С празднико́м! [праздники]

Book 2

- 10 У приро́ды нет плохой погóды [погода, природа, окружающая среда]
- 11 Порá в пúть-дорóгу! [путешествие по всему миру, страны мира, достопримечательности, иммиграция, история]
- 12 Бýдьте здорóвы! [здравье, медицина, лечение, лекарство, лíчная гигиéна]
- 13 @ — дру́г человéка [технология (интернét, компьютер), экономика, сообщéние]
- 14 Семéйное счастье [семья́, бráк, отношéния]
- 15 Граждани́н и госудárство [государство, правительство, власть, государственный строй, преступление и наказание, право и закон, правила]
- 16 Войнá и мíр [войнá и мíр, дипломáтия, догово́ры, согла́сие]
- 17 «Жíзнь коротká, иску́сство вéчно» [иску́сство и мúзыка]
- 18 Нáше всё [литератúра, аргумéнты и фáкты]

Visualizing Russian and other textbooks

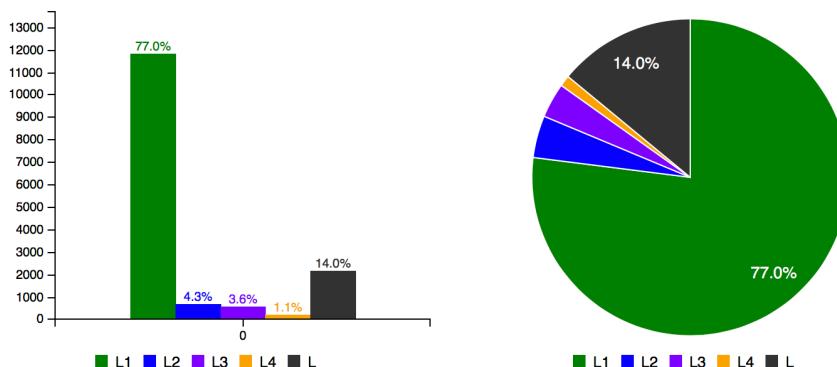
The Visualizing Russian tools can be applied to any textbooks. Eventually tools could be added so that the vocabulary levels could be tweaked to correspond to the target vocabulary in other textbooks or to the learned/unknown/target vocabulary of individual learners as in the Hedera Project by my colleague Dr. Ivy Livingston (Harvard University):

Hedera: Reading that fits

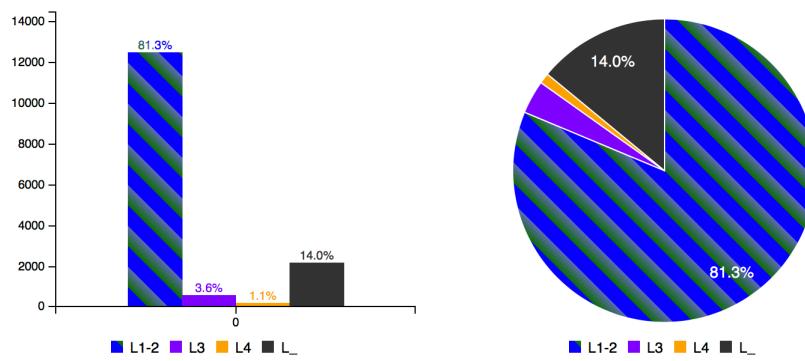
<https://atg.fas.harvard.edu/hedera>

<https://hedera.fas.harvard.edu>

I looked into the contents of the new elementary/beginning Russian textbook in development in Tromsø, *Min russiske reise*, for analysis in Visible Vocabulary. The book contains a compelling running narrative through prose texts and dialogues totalling 15K words, well-suited to the elementary level based on analysis and comparison with *Foundations of Russian vocabulary* level: 77% level one (green) words.



Word Count: 15305
Unparsed Count: 2147
L1 Count: 11785
L2 Count: 657
L3 Count: 544
L4 Count: 172



Word Count: 15305
Unparsed Count: 2147
L1 Count: 11785
L2 Count: 657
L3 Count: 544
L4 Count: 172

Special thanks to the Visualizing Russian development team, the various funding support and to my coauthors on *Foundations of Russian*.

Development of Visualizing Russian (2014-2021)

Development Team

Steven Clancy, Slavic Languages and Literatures, Harvard University

with colleagues at Harvard University

Arthur Barrett, (2017-present), Senior Software Engineer, Academic Technologies for the Faculty of Arts and Sciences

Paige Lee (2020-present), student research assistant

Christopher Morse (2014-2017), formerly of Arts and Humanities Research Computing

Cole Crawford, research computing specialist with Arts and Humanities Research Computing

Bill Barthelmy, Senior Technical Architect with Academic Technologies for the Faculty of Arts and Sciences

Funding Support

The Visualizing Russian project has benefitted from the financial support of the **Barajas Dean's Innovation Fund for Digital Arts and Humanities** (Dean of Arts and Humanities, Harvard University) (2018-present) and additional support from the **Foreign Language Advisory Group** (FLAG, Harvard University), the **Davis Center for Russian and Eurasian Studies** (Harvard University), and from the **Academic Technologies for the Faculty of Arts and Sciences Group** (Harvard University).

Extras

frequency isn't enough of a guide on its own

- typical domains included in textbooks are present in most frequent words, but not fully fleshed out as semantic groups
- **MONTHS** appear by lemma frequency from ranks 1938-3587, **DAYS OF THE WEEK** from 1547-8834, **NATIONALITY** (country, citizens (M-F-PI), adj, language) sets are dispersed throughout and would never be encountered in a purely frequency-based selection of lexical items

Франция француз француженка французский по-французски
Франция, француз, француженка, французский, (по-французски)
Россия, русский, русская, русские, российский, россиянин, россияне, (по-русски)
Англия, английский, англичанин, англичанка, англичане, (по-английски)

- lexemes that cross domains (средá, трубá, шáшка out-punch their weight in a set or at least skew the results)
- leveling is complex (how to decide what lexemes are appropriate to each level, how many levels, etc.), but also order in one's textbook, time is required for development (blue words are on the intermediate-level, but that level is traversed over 18 chapters in our book)
- grammar/lexicon/communicative themes and functions all demand their own approaches; can be tricky to combine
- just like the measured frequencies of various words will change depending on the content of and updates to our corpus, the items in our green/blue/purple/orange/black levels are subject to change, but are relatively stable and have their own characteristics apart from mere frequency

Teaching target vocabulary

What we do as teachers...

Teach students lots of useful, difficult, hard-won vocabulary...

выбирáть|выíбрать
входíть|войти
засыпáть|заснúть
|сказáть
варíться|сварíть

Lots about VOM, but then you still need things like:

добра́ться до GEN
попа́сть в ACC
пройти к DAT

and then they go and find things to read or listen to, etc.

and encounter instead...

подбира́ть подобráть	выбирайтъ выíбрать
уснúть	засыпáть заснúть
заходíты зайтý	входíть войти
подскажáть	сказáть
завáривать заварíть чай	варíться сварíть чай
напугáть	пугáть испугáть

«Каждое слово имеет значение, – говорит она. – Убеждать – значит правильно подбира́ть выражения.

Стрóя домá из кўбиков маленький Боря скучал, зевал, а потом и совсем уснул.

Вы не подскáжете, как добра́ться до Эрмитáжа?

Какими масками можно напугáть людей на Хеллоуин?

Take-away: Students need to develop a solid foundation in Russian vocabulary, but also learn when prefixes **make a big difference** and when they simply indicate near-synonymous expressions.

How to *best* represent four + one categories in text?

[4+1 levels in green—blue—purple—orange—black]

[4+1 levels in green—blue—purple—orange—black]

[4+1 levels in color blind readable categories green—blue—purple—orange—black]

[4+1 levels in categorically different font styles sans serif—serif—script—decorative—other]

[4 + 1 levels in somewhat more aesthetically pleasing font choices green—blue—purple—orange—black]

[4+1 levels in shades as green—blue—purple—orange—black]

[4+1 levels in point size green—blue—purple—orange—black]

[4+1 levels in styles green—blue—purple—orange—BLACK]

[4+1 levels in green—blue—purple—orange—black]

Великобритания может пойти на высылку российских дипломатов, объявить бойкот Чемпионату мира по футболу-2018, а также отозвать лицензию на телевещание у телеканала RT, опубликовать информацию о банковских счетах и активах российских политиков и бизнесменов, арестовать эти средства, отключить Россию от системы обмена банковской информацией SWIFT. Наконец, Великобритания может, апеллируя к 5-й статье устава НАТО, потребовать от других стран-участников Североатлантического альянса нарастить свое военное присутствие у границ России и включить РФ в список «стран-пособников терроризма». Скрипаль — далеко не первый перебежчик из России, который подвергся нападению на территории Великобритании, хотя найти неопровергимых доказательств причастности Москвы к подобным происшествиям Лондону так и не удалось. Однако никогда раньше, даже в случае летальных исходов, как, например, было с экс-сотрудником ФСБ Александром Литвиненко, таких серьезных угроз со стороны Лондона в адрес Москвы не было.

green	[15, 112, 1]
blue	[0, 0, 255]
purple	[107, 0, 255]
orange	[253, 148, 9]
black	[0, 0, 0]

[4+1 levels in green—blue—purple—orange—black]

Великобритания может пойти на высылку российских дипломатов, объявить бойкот Чемпионату мира по футболу-2018, а также отозвать лицензию на телевещание у телеканала RT, опубликовать информацию о банковских счетах и активах российских политиков и бизнесменов, арестовать эти средства, отключить Россию от системы обмена банковской информацией SWIFT. Наконец, Великобритания может, апеллируя к 5-й статье устава НАТО, потребовать от других стран участников Североатлантического альянса нарастить свое военное присутствие у границ России и включить РФ в список «стран-пособников терроризма». Скрипаль — далеко не первый перебежчик из России, который подвергся нападению на территории Великобритании, хотя найти неопровергимых доказательств причастности Москвы к подобным происшествиям Лондону так и не удалось. Однако никогда раньше, даже в случае летальных исходов, как, например, было с экс-сотрудником ФСБ Александром Литвиненко, таких серьезных угроз со стороны Лондона в адрес Москвы не было.

green	[15, 112, 1]
blue	[0, 0, 255]
purple	[107, 0, 255]
orange	[253, 148, 9]
black	[0, 0, 0]

[4+1 levels in color blind readable categories green—blue—purple—orange—black]

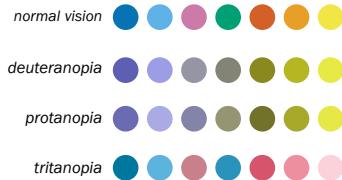
Великобритания может пойти на высылку российских дипломатов, объявить бойкот Чемпионату мира по футболу-2018, а также отозвать лицензию на телевещание у телеканала RT, опубликовать информацию о банковских счетах и активах российских политиков и бизнесменов, арестовать эти средства, отключить Россию от системы обмена банковской информацией SWIFT. Наконец, Великобритания может, апеллируя к 5-й статье устава НАТО, потребовать от других стран-участников Североатлантического альянса нарастить свое военное присутствие у границ России и включить РФ в список «стран-пособников терроризма». Скрипаль — далеко не первый перебежчик из России, который подвергся нападению на территории Великобритании, хотя найти неопровергимых доказательств причастности Москвы к подобным происшествиям Лондону так и не удалось. Однако никогда раньше, даже в случае летальных исходов, как, например, было с экс-сотрудником ФСБ Александром Литвиненко, таких серьезных угроз со стороны Лондона в адрес Москвы не было.

bluish green	[0, 158, 115]
blue	[0, 114, 178]
reddish purple	[204, 121, 167]
orange	[230, 159, 0]
black	[0, 0, 0]

conservative 7-color palette adapted for color blindness

Color	Color name	RGB (1–255)	CMYK (%)	P	D
Black	Black	0, 0, 0	0, 0, 0, 100	█	█
Orange	Orange	230, 159, 0	0, 50, 100, 0	█	█
Sky blue	Sky blue	86, 180, 233	80, 0, 0, 0	█	█
Bluish green	Bluish green	0, 158, 115	97, 0, 75, 0	█	█
Yellow	Yellow	240, 228, 66	10, 5, 90, 0	█	█
Blue	Blue	0, 114, 178	100, 50, 0, 0	█	█
Vermillion	Vermillion	213, 94, 0	0, 80, 100, 0	█	█
Reddish purple	Reddish purple	204, 121, 167	10, 70, 0, 0	█	█

SORTED BY SIMILARITY IN DEUTERANOPIA



Wong, B. (2011) Points of view: Color blindness. Nature Methods 8:441.

<http://mkweb.bcgsc.ca/colorblind>

[4+1 levels in categorically different font styles sans serif—serif—script—decorative—other]

Великобритания может пойти на высылку российских *дипломатов*, объявить *бойкот Чемпионату* мира по футболу-2018, а также *отозвать лицензию* на телевещание у телеканала RT, *опубликовать* информацию о *банковских* счетах и *активах* российских политиков и бизнесменов, арестовать эти средства, *отключить* Россию от системы обмена *банковской* информацией SWIFT. Наконец, Великобритания может, апеллируя к 5-й статье *устава НАТО*, потребовать от других стран-участников Североатлантического *альянса* *нарастить* свое *военное* присутствие у границ России и включить РФ в список «стран-пособников *терроризма*». Скрипаль – далеко не первый перебежчик из России, который подвергся нападению на территории Великобритании, хотя найти *неопровергнуемых* доказательств *причастности* Москвы к подобным происшествиям Лондону так и не удалось. Однако никогда раньше, даже в случае летальных *исходов*, как, например, было с экс-сотрудником ФСБ Александром Литвиненко, таких серьезных *ущербов* со стороны Лондона в адрес Москвы не было.

Gill Sans, Cambria, *Manuscript Cy*, Bradley Hand, Monaco

[4 + 1 levels in somewhat more aesthetically pleasing font choices green—blue—purple—orange—black]

Великобритания может пойти на высылку российских дипломатов, **объявить** бойкот Чемпионату мира по футболу-2018, а также отозвать лицензию на телевещание у телеканала RT, опубликовать информацию о банковских **счетах** и активах российских политиков и бизнесменов, **арестовать** эти **средства**, отключить Россию от системы **обмена** банковской информацией SWIFT. Наконец, Великобритания может, апеллируя к 5-й статье устава НАТО, **потребовать** от других стран-**участников** Североатлантического альянса нарастить свое военное **присутствие у границ** России и **включить** РФ в **список** «стран-пособников терроризма». Скрипаль — далеко не первый перебежчик из России, который подвергся **нападению** на **территории** Великобритании, хотя найти неопровергимых **доказательств** причастности Москвы к **подобным происшествиям** Лондону так и не **удалось**. Однако никогда раньше, даже в случае летальных исходов, как, например, было с **экструдником** ФСБ Александром Литвиненко, таких серьезных угроз со стороны Лондона в адрес Москвы не было.

Helvetica Neue Light—**Helvetica Neue Bold**—Cambria—Calibri—Georgia

[4+1 levels in shades as green—blue—purple—orange—black]

Великобритания может пойти на высылку российских дипломатов, объявить бойкот Чемпионату мира по футболу-2018, а также отозвать лицензию на телевещание у телеканала RT, опубликовать информацию о банковских счетах и активах российских политиков и бизнесменов, арестовать эти средства, отключить Россию от системы обмена банковской информацией SWIFT. Наконец, Великобритания может, апеллируя к 5-й статье устава НАТО, потребовать от других стран-участников Североатлантического альянса нарастить свое военное присутствие у границ России и включить РФ в список «стран-пособников терроризма». Скрипаль — далеко не первый перебежчик из России, который подвергся нападению на территории Великобритании, хотя найти неопровергимых доказательств причастности Москвы к подобным происшествиям Лондону так и не удалось. Однако никогда раньше, даже в случае летальных исходов, как, например, было с экс-сотрудником ФСБ Александром Литвиненко, таких серьезных угроз со стороны Лондона в адрес Москвы не было.

[4+1 levels in point size green—blue—purple—orange—black]

Великобритания **может** пойти на высылку **российских** дипломатов, объявить бойкот Чемпионату Мира по футболу-2018, а также отозвать лицензию на телевещание у телеканала RT, опубликовать информацию о банковских счетах и активах **российских** политиков и бизнесменов, арестовать эти средства, отключить Россию от системы обмена банковской информацией SWIFT. Наконец, Великобритания **может**, апеллируя к 5-й статье устава НАТО, потребовать от других стран-участников Североатлантического альянса нарастить свое военное присутствие у границ России и включить РФ в список «стран-последников терроризма». Скрипаль – далеко не первый перебежчик из России, **который** подвергся нападению на территории Великобритании, хотя найти неопровергимых доказательств причастности Москвы к подобным происшествиям Лондону так и не удалось. Однако никогда раньше, даже в случае летальных исходов, как, например, было с экс-сотрудником ФСБ Александром Литвиненко, **таких** серьезных угроз со стороны Лондона в адрес Москвы не было.

[4+1 levels in styles green—blue—purple—orange—BLACK]

ВЕЛИКОБРИТАНИЯ может пойти на высылку российских дипломатов, **объявить** бойкот Чемпионату мира по Футболу-2018, а также *отозвать лицензию* на ТЕЛЕВЕЩАНИЕ у ТЕЛЕКАНАЛА RT, *опубликовать* информацию о банковских **счетах** и активах российских политиков и бизнесменов, **арестовать** эти **средства**, *отключить* Россию от системы **обмена** банковской информацией SWIFT. Наконец, Великобритания может, апеллируя к 5-й статье устава НАТО, **потребовать** от других СТРАН-участников СЕВЕРОАТЛАНТИЧЕСКОГО альянса нарастить свое **военное присутствие** у границ России и **включить** РФ в **список** «СТРАН-пособников терроризма». Скрипаль — далеко не первый перебежчик из России, который подвергся **нападению** на **территории** Великобритании, хотя найти неопровергимых **доказательств** причастности Москвы к **подобным происшествиям** Лондону так и не удалось. Однако никогда раньше, даже в случае летальных исходов, как, например, было с ЭКС-сотрудником ФСБ Александром Литвиненко, таких серьезных угроз со стороны Лондона в адрес Москвы не было.