



2022

# Empirically Determined Strategic Input and Gamification in Mastering Russian Word Forms

Laura A. Janda

Jan Ove Nikolai Almendingen

Linn Thea Kaldager Josefsen

Follow this and additional works at: <https://scholarsarchive.byu.edu/rlj>



Part of the [Slavic Languages and Societies Commons](#)

---

### Recommended Citation

Janda, Laura A.; Almendingen, Jan Ove Nikolai; and Kaldager Josefsen, Linn Thea (2022) "Empirically Determined Strategic Input and Gamification in Mastering Russian Word Forms," *Russian Language Journal*: Vol. 72: Iss. 1, Article 3.

Available at: <https://scholarsarchive.byu.edu/rlj/vol72/iss1/3>

This Article is brought to you for free and open access by the Journals at BYU ScholarsArchive. It has been accepted for inclusion in Russian Language Journal by an authorized editor of BYU ScholarsArchive. For more information, please contact [ellen\\_amatangelo@byu.edu](mailto:ellen_amatangelo@byu.edu).

## **Empirically Determined Strategic Input and Gamification in Mastering Russian Word Forms**

**LAURA A. JANDA, JAN OVE NIKOLAI ALMENDINGEN  
LINN THEA KALDAGER JOSEFSEN**

### **1. Introduction**

We propose two designs to gamify second language (L2) learning of Russian inflectional morphology: Treasure Hunt and Story Time. The goal of these designs is to focus learning on high-frequency word forms that are most strategic and effective for L2 acquisition in a way that stimulates engagement and builds lifelong learning skills.

These two gamification designs emerged from a student focus group that was convened to propose implementations for the SMARTool (see Section 3). After an initial brainstorming session, the ideas were further developed by the instructor, honed by the students, and tested in class. Students have also contributed to and commented on the contents of this article.

In Section 2 we briefly identify the problem, namely, the enormous number of paradigm forms potentially present in Russian paradigms and their skewed distribution. We cite research showing that inflectional morphology is a major hurdle for L2 learners but not for native speakers, who use only a fraction of the potential forms and can easily understand and produce forms that they have never encountered. Furthermore, evidence demonstrates that learning can be enhanced by strategically concentrating on the highest-frequency forms. Access to the highest-frequency forms of over 3,000 lexemes is provided by the SMARTool described in Section 3, but that resource is relatively static, meaning that more guidance is needed on how to implement this tool in the classroom and in self-study. Our two proposed designs are presented in Section 4 (Treasure Hunt) and Section 5 (Story Time). Conclusions are offered in Section 6.

### **2. The challenge of inflectional morphology**

Russian is just one of many languages that have *rich inflectional morphology*, meaning that words can have many different forms to signal grammatical

categories such as case, number, person, tense, and so forth. Each Russian noun has at least a dozen forms, each adjective about 30 forms, and each verb several dozen forms (20 verb forms, plus inflections of participles). The full paradigms for even a modest vocabulary of a few thousand words constitute an array of over 100,000 word forms. However, the frequency of word forms follows a highly skewed Zipfian distribution (Karlsson, 1986; Zipf, 1949), meaning that only a handful of the potential forms of any given word occur frequently, while the remainder are rare (many vanishingly so).

In other words, while Russian morphology can produce a huge volume of word forms, only a small fraction of word forms are commonly used. Only three word forms are needed to account for the majority of uses of an average high-frequency inflected Russian word (Janda & Tyers, 2021). For many words, including all lower-frequency words as well as words that are closely associated with a given grammatical construction, over 90% of uses involve only one inflected form. For example, *протяжение* [expanse] is a fairly high-frequency word (with over 31 occurrences per million words in the Russian National Corpus,<sup>1</sup> henceforth “RNC”). Although this word is attested in all 12 paradigm forms in the RNC, the locative singular *протяжении* accounts for 92% of the uses of this word. The dative plural *протяжениям* and instrumental plural *протяжениями* forms are attested only once each, and these in the 19th and 18th centuries, respectively. Given that the RNC contains over 337 million words, a quantity roughly equivalent to the lifetime exposure of a human being between 40 and 70 years old to their native language, this fact indicates that many native speakers have probably never encountered these word forms. However, all native speakers of Russian can be expected to readily understand and produce these forms in appropriate contexts, as evidenced by rare occurrences that turn up in Google searches. Janda and Tyers (2021) showed that less than one tenth of 1% of Russian nouns are attested in all 12 paradigm forms, regardless of the size of the corpus. The skewed distribution of forms is much more pronounced with lower-frequency words, which typically occur in only one inflected form (note that given the Zipfian distribution, approximately one half of unique lexemes are

---

<sup>1</sup> See <https://ruscorpora.ru>.

very rare, and these are known as *hapaxes*). This means that Russian native speakers are exposed only to partial paradigms for the vast majority of words that they know.

Thus we face a linguistic conundrum termed the *Paradigm Cell Filling Problem* (Ackerman et al., 2009): the fact that native speakers of languages with rich inflectional morphology routinely recognize and produce forms that they have never been exposed to. It is not entirely clear how first (native) language (L1) learners acquire and navigate rich morphologies. Evidence shows that native speakers are sensitive to frequency (Goldberg, 2006, Chapter 5). Janda and Tyers (2021) suggested that native speakers acquire many partial paradigms. Since the frequency distribution of forms is unique for each word, the partial paradigms overlap, and collectively they cover the entire set of paradigm cells for each declension and conjugation class. This makes it possible for native speakers to triangulate from words whose given form is very common to words whose form is very rare. Returning to the preceding example of *протяжение* [expanse] that has no RNC attestations of dative and instrumental plural in the 20th or 21st centuries, native speakers have recourse to other words in the same declension class that have very frequent dative and instrumental plural forms, such as *упражнение* [exercise] (see the common phrases *ответы к упражнениям* [answer key] and *тетрадь с упражнениями* [book of exercises]).

Empowering L2 learners to navigate rich inflectional morphology in a native-like manner is a significant challenge (Hopp, 2010). Morphology is considered to be both essential to L2 acquisition and a “bottleneck,” as well as more difficult than both syntax and semantics, and multiple studies have shown that learning the myriad word forms in inflected languages is more difficult than learning other aspects of language (Jensen et al., 2019; Slabakova, 2009, 2014). L2 learners largely lack the resources of language experience that native speakers can fall back on when manipulating morphological forms. L2 learner acquisition is necessarily compressed because L2 learners do not have the tens of thousands of hours of language exposure that enables native speakers to build up their extensive reservoirs of overlapping partial paradigms. Can we find shortcuts to help L2 learners acquire a native-like fluency with inflectional morphology?

Janda and Tyers (2021) conducted a machine learning experiment using data from the SynTagRus corpus. The experiment compared two training simulations: one that involved learning the full paradigms for Russian nouns, verbs, and adjectives (henceforth “full model”) and one in which the computer learned only the single most-frequent inflected form of each lexeme (henceforth “single form model”). In the testing phase, the task for both models was the same: to predict given forms for new (previously unseen, not included in the training) lexemes. In other words, the training would include (among others) the lexeme *книга* [book], for which the full model was trained on all forms for all case and number combinations, but the single form model was trained only on the most common form, namely the accusative singular *книгу*. The testing phase then asked each of the two models to predict the accusative singular form of a new lexeme, for example *рыба* [fish], that was not included in the training phase. The training phase for both models began with 100 lexemes and was repeatedly increased in round after round to 200, 300, and so on, up to 5,400 lexemes. In each round, the two models were tested on the prediction of given inflected forms for 100 new lexemes. From the 11th round (after training on 1,100 lexemes) through all subsequent rounds, the single form model consistently outperformed the full model. The full model never scored above 80% correct predictions, whereas the single form model scored above 80% on most rounds, and above 90% on some rounds. Analysis of errors showed that even when the single form model made incorrect predictions, its errors were less serious (measured in *Levenshtein distance*, the number of letters needed to be changed to achieve the correct answer; see Levenshtein, 1965/1966). In other words, at least for a computer, learning Russian inflectional morphology is more effective when focusing on the most frequent forms instead of memorizing entire paradigms.

Janda and Tyers’s (2021) experimental results suggest that learning should be focused on the most-frequent inflected forms rather than on whole paradigms. Language instructors have probably always tried to emphasize the word forms that seem most common, but thanks to the existence of large corpora like the RNC, it is possible to scientifically determine exactly what forms are the most common. However, this is not a trivial task because each lexeme has a unique distribution of inflected forms. For example, the top three most-frequent forms of *церковь* [church] are genitive singular, nominative singular, and instrumental singular,

but for the near-synonym *храм* [temple, house of worship], the top three most frequent forms are accusative singular, accusative plural, and dative singular. Furthermore, just knowing the most-frequent forms gives us an incomplete picture. The grammatical constructions and collocations that motivate the same forms for different lexemes can be very different. As we saw previously, *протяжение* [expanse] occurs predominantly in the locative singular form, a fact that is motivated by its prominent role in a grammatical construction meaning “during” that consists of the preposition *на* [on] followed by *протяжении*, in turn followed by a noun phrase in the genitive case that refers to a time period. There are many other nouns that have a preference for the locative singular, and for each lexeme there is a specific motive, involving different prepositions, meanings, and collocations.

To reduce the burden of memorizing inflected forms for L2 learners of Russian and boost their morphological accuracy, we need to focus on the most frequent word forms. The selection of high-frequency forms can be informed by corpus data. However, each and every lexeme presents a unique set of motives for its highest-frequency forms, requiring investigation of the grammatical and lexical contexts that are most typical for each word.

Evidence shows that the majority of language produced by native speakers of any language consists of stringing together prefabricated units (*chunks*) such as “read a book” or “I’m trying to” in English. Estimates vary, but perhaps over 80% of language is the recombination of well-rehearsed chunks (see Dąbrowska, 2004, p. 19 for an overview of scholarly literature). It therefore makes sense to steer L2 learners’ attention to the word forms and contexts that predominate in Russian discourse.

### **3. Strategic stratification for learning inflectional morphology**

Linguistic corpora are not in themselves new, but there has to date been little substantial implementation of corpora in language teaching. The oldest language corpora were founded in the 1970s, and large digital collections of language samples with hundreds of millions of words have existed for over a decade. However, with some notable exceptions (Hopp, 2010), corpus resources have been aimed primarily at linguists, not L2 learners, and it has been difficult to find ways to connect L2 learners to the powerful benefits of using corpus language data.

The research described in Section 2 has inspired the development of the SMARTool<sup>2</sup> (Strategic Mastery of Russian Tool; Janda, 2019). The aim of the SMARTool is to give learners and instructors access to Russian word forms stratified by frequency, with the focus restricted to the word forms and contexts that are most strategic for learners to acquire.

The SMARTool is a free, publicly available resource that does not require a password, is accessible across a multitude of devices, and requires nothing more than a stable internet connection. The SMARTool was built using open-source code stored on GitHub and was deliberately designed to facilitate portability to other languages. Over 3,000 nouns, verbs, and adjectives are represented in the SMARTool, spanning Common European Frame of Reference (CEFR) proficiency Levels A1 through B2, representing a basic minimal vocabulary for each level. Corpus data<sup>3</sup> has been used to determine the most-frequent inflected forms of each lexeme. For most words, the three most-frequent forms are included, but if only one or two word forms account for over 90% of attestations of a given lexeme, then only those forms are included. There are therefore about 9,000 word forms represented in the SMARTool, less than 10% of the total number of potential word forms associated with the vocabulary. The collocational preferences and typical grammatical contexts of every single word form have been identified on the basis of corpus data, and all word forms are presented in their characteristic contexts, namely, in a corpus-inspired example sentence. Audio versions of all sentences are available at the click of a button, as are English translations, so learners can check both pronunciations and meanings.

For example, if a user looks up the noun *вопрос* [question] in the SMARTool, they receive the following sentences (including the translations, if the user has checked the box to request them):

---

<sup>2</sup> See the SMARTool at <https://smartool.github.io/smartool-rus-eng/>.

<sup>3</sup> The SynTagRus ([https://github.com/UniversalDependencies/UD\\_Russian-SynTagRus](https://github.com/UniversalDependencies/UD_Russian-SynTagRus)) corpus was used to determine the most-frequent inflected forms of each lexeme. In addition, the Russian National Corpus (see Section 2) and the Collocations, Colligations and Corpora resource (CoCoCo; <https://cococo.cosyco.ru/download.html>) were consulted to determine collocational preferences and typical grammatical contexts. Example sentences are inspired by these corpus resources, meaning that they have been simplified to focus on the given word forms and their immediate contexts. It is not feasible to use unedited corpus examples in beginning and intermediate L2 Russian instruction because (a) individual sentences extracted from a corpus are often hard to understand even for native speakers without more context, and (b) corpus sentences tend to be long, containing extraneous information that distracts from the learning goals.

*Быть или не быть, вот в чём **вопрос**.* (Nom.Sing)

“To be, or not to be, that is the question.”

*Никто не может ответить на мои **вопросы**.* (Acc.Plur)

“Nobody can answer my questions.”

*У нас много **вопросов** к президенту.* (Gen.Plur)

“We have many questions for the president.”

This information indicates that the following three case and number combinations are most common for this word, in descending order: nominative singular, accusative plural, and genitive plural (a tab at the top of the page directs the user to the list of abbreviations if needed). Furthermore, we learn several crucial constructions that go with these three most-frequent word forms, namely (*вот*) *в чём вопрос* [that is the question/what is the question], *ответить на вопросы* [answer questions], *у + genitive вопрос/много вопросов* [somebody has a question/many questions], and *вопрос(ы) к + dative* [question(s) for somebody]. Learners are thus equipped with enough information to successfully interpret and use the word *вопрос* [question] in the very contexts they are most likely to encounter.

The SMARTool’s filters make it possible for users to select content according to CEFR Levels, Topics (in 18 categories such as *время* [time] and *еда* [food]), Analysis (combinations of grammatical categories, such as locative singular), and Dictionary (permitting the user to both type in part of a word and scroll through the entire inventory). The filters are designed to encourage learner experimentation and autonomy. Search by analysis makes it possible to reverse the perspective of learning inflectional morphology: instead of showing what word forms are most associated with each lexeme, searching by grammatical categories shows which lexemes are most associated with given grammatical categories. Especially when learning challenging verb forms (like gerunds and participles), it can be useful to find out what words actually occur frequently in those forms. For instance, examples are provided for high-frequency perfective gerunds such as *оглянувшись* [after taking a look around]. In addition, the SMARTool A1 vocabulary serves as the learner dictionary for *Min rusiske reise* [*My Russian Journey*], an online beginner



course in Russian.<sup>4</sup> A scaled-down version of the SMARTool (sourced from the same data set) that can be filtered for the 35 lessons in that course is also available.<sup>5</sup>

Representation of the Russian language in the SMARTool is limited somewhat by the available data, which is itself skewed, particularly in terms of gender. The Russian language expresses gender in all singular past-tense verb forms (e.g., “was”: *был* [masculine singular], *была* [feminine singular], *было* [neuter singular]), as well as singular forms of adjectives and participles, and for many ethnonyms and professional titles there are distinct male and female forms (e.g., “an American”: *американец* [masculine], *американка* [feminine]). Kuznetsova (2015) showed that in corpus data, past-tense forms of verbs associated with human subjects typically have three times more attestations of masculine forms than of feminine forms. For example, the RNC contains 407,823 attestations of *сказал* [he said] but only 119,855 attestations of *сказала* [she said], a ratio of over 3.4:1. Skewed data of this type is not particular to Russian or to language corpora. As Criado-Perez (2019) and D’Ignazio and Klein (2020) have shown, underrepresentation of women is endemic across all kinds of data. We aim to correct for the skew in data by taking appropriate steps to improve the gender balance in our gamification exercises.

The SMARTool is an important step forward in using corpus data to make a real difference in the experience of L2 learners of Russian. The tool gives learners and instructors access to the most strategic inflected forms and usage contexts for a basic vocabulary of nouns, verbs, and adjectives. It is interactive in the sense that users can search according to topics, grammar, lexicon, and proficiency level, with both translations and audio on demand. However, beyond this, the SMARTool is a static resource and risks being underutilized, like a reference book that merely collects dust on a shelf. Users need instructions on how to use the SMARTool and a motive to do so. We offer two designs for engaging users in such a way that they will learn by doing, and in so doing acquire lifelong learning skills that they can apply beyond the tasks at hand. While gamification is meant to add some fun to the business of acquiring inflectional morphology, it is also more than that. Gamification encourages learners to transition from passive reception to

---

<sup>4</sup> See <https://mooc.uit.no/courses/course-v1:UiT+C001+2020/about>.

<sup>5</sup> See <https://smartool.github.io/min-russiske-reise/>.

active inquiry (see Harvey Arce & Cuadros Valdivia, 2020). Our goal is to spark curiosity and creative expression by challenging users to take advantage of the potential of the SMARTool. We plan to create apps like the SMARTool itself that can be used across various devices (laptop, tablet, smartphone).

Sections 4 and 5 present two gamification concepts that emerged from focus group meetings with undergraduate L2 learners of Russian in 2020 and 2021. The concepts were generated and initially developed in the focus group. The faculty member (Janda) further refined the ideas and worked out specific exercises that were vetted by focus group members, and in 2022 these exercises were piloted with a new cohort of undergraduate students. Two focus group members (Almendingen and Josefsen) were consulted in the writing and editing of this article.

#### **4. Treasure Hunt**

The Treasure Hunt design launches users on explorations into various corners of the Russian language. Explorations guide users to useful discoveries not only about inflectional morphology but also about phonology, semantics, syntax, derivational morphology, and even alternative ways of categorizing the human experience—all without needing to learn any linguistic terms. Treasure Hunt activities are stratified for proficiency level, and even A1 users (with a vocabulary of only a few hundred words) have ample opportunities to go on Treasure Hunts. Treasure Hunts can be undertaken in groups or individually, in the classroom or during self-study.

Each Treasure Hunt begins with a simple Prompt, an instruction on how to use a SMARTool search function to extract a target set of sentences, and a question to consider. After deducing an answer using the SMARTool, users can compare their answers with an Answer Key. A Take-Away Idea summarizes the result and what users can do with it.

Here we cite four examples of Treasure Hunts from the A1 level and describe some Treasure Hunts for more advanced levels.<sup>6</sup> Note that we do not cite the sentences that the SMARTool presents for these examples; the user will find these examples when they consult the resource. We also do not translate words here since users can find translations in the SMARTool.

---

<sup>6</sup> These and more Treasure Hunts are available at <https://smartool.github.io/exercises/>.

#### 4.1. Treasure Hunts for CEFR level A1

1.

Prompt:

Choose: Search by dictionary

Find all the words that begin with *a-* and *э-*.

What do these words have in common?

Answer Key:

All of the words that begin with *a-* and *э-* in the SMARTool dictionary are borrowed words in Russian.

Native Russian words do not begin with *a-*. The only exceptions are *ахать*, *ахнуть* [say ah!].

Native Russian words do not begin with *э-*. The only exceptions are *этом* [this/that] and other forms of this word (*эта*, *эти*, etc.).

In general, most Russian words begin with a consonant. This includes words that begin with *е-*, *ю-*, *я-*, which begin with the consonant *й-* (sounds like *y-* in English). If a Russian word begins with a vowel, it is one of these: *и-*, *о-*, or *у-*.

Take-Away Idea: If you encounter a long new word that begins with a vowel, it is probably a borrowed word. If you sound it out, you will probably recognize it. For example, *экономический* means “economic,” and you don’t need a dictionary to figure that out.

2.

Prompt:

Choose: Search by dictionary

Look up these words: *российский*, *русский*, *иностранный*. Look at the sentences.

What kinds of items can be *российский*, and what kinds can be *русский*? Can you compare this with the use of the word *иностранный*?

Answer Key:

We use *российский* to describe items connected to Russia as a state (*на-спорт*, *Федерация*).

We use *русский* to describe items connected to the Russian language, culture, and ethnic identity (*алфавит*, *литература*, *авангард*).

For many items, you can use both adjectives, depending on what you want to emphasize.

For example: *российские журналисты* are journalists from Russia, whereas *русские журналисты* are journalists who are Russian.

The word *иностранный* can be used to describe both geopolitical relationships and those of language and culture.

Take-Away Idea: *Русский* is about ethnic identity; *российский* is about a relationship to the Russian Federation.

3.

Prompt:

Choose: Search by dictionary

Look up these words: *кухня, ресторан*.

Read the sentences. There are two patterns that have to do with going to a place, being in a place, and going away from a place. Can you identify the two patterns?

Once you have found the two patterns, look up and try to sort these words into two groups according to the two patterns:

*страна, мир, место, дом, школа, город, квартира, класс, свет, центр, улица, комната, район, театр, парк, музей, стадион, гостиница, остановка, вокзал, факультет, бассейн, общежитие, Россия, аэропорт, фабрика, столовая, аптека, номер, завод, университет, клуб, концерт, сад, площадь, здание, столица, кабинет, лекция, этаж*

Answer Key:

The two patterns are:

*на кухню* (accusative), *на кухне* (locative), *с кухни* (genitive)<sup>7</sup>

*в ресторан* (accusative), *в ресторане* (locative), *из ресторана* (genitive)

The pattern with *в* and *из* is used more than the one with *на* and *с*.

Take-Away Idea: The prepositions *на* and *с* are mostly used with large, open places (*стадион, фабрика, завод, остановка, свет, место, вокзал*), surfaces (*площадь, этаж*), and events (*концерт, лекция*). With other places, we use the prepositions *в* and *из*.

4.

Prompt:

Choose: Search by topic and choose *еда* [food].

Toggle through all the entries and look at the nouns. Notice what

---

<sup>7</sup> Note that the SMARTool represents patterns of highest frequency. It is also possible to say *в кухне*, but this phrase is much less common than *на кухне* in Russian.

words appear in singular and what words appear in plural. Can you make some generalizations?

Answer Key:

Only singular in the SMARTool: *вода, масло* (NB! both “butter” and “oil”), *сок, сыр, мороженое, сахар, картошка, колбаса, чай, пиво, хлеб, мясо, молоко, вино, еда*.<sup>8</sup>

Both singular and plural in the SMARTool: *продукт* (usually plural if referring to food), *салат* (plural refers to various kinds or portions of salad), *яйцо, суп* (plural refers to various kinds of soup), *соль* (if plural usually not about food but about chemicals), *курица* (plural *куры* is used for animals, not food), *рыба* (plural *рыбы* is used for animals, not food), *яблоко*.

Only plural in the SMARTool: *фрукт, овощ*.

Take-Away Idea: Many foods are primarily understood as substances in Russian, even if they come in fairly large pieces (potatoes, sausages, fish, chicken). These words tend to occur mostly or exclusively in the singular. Note that *фрукты, овощи, продукты* (when it means “groceries”) almost always occur in the plural, probably because they are not homogeneous (there are lots of kinds of fruits and vegetables and groceries). Food items that one tends to count (apples, eggs) are used in both singular and plural.

These and similar Treasure Hunts were piloted in a class with students that had just completed their first semester of study at UiT The Arctic University of Norway (A1 level) in January 2022. Students reported that this was a fun way to review vocabulary, that it was interesting to find differences between words and uses on their own, and that the Take-Away Ideas presented “cool facts.” All students reported that they had learned something useful and that they would recommend similar exercises to other students.

#### ***4.2. Treasure Hunts for more advanced levels***

At more advanced levels, Treasure Hunts target morphology (e.g., formation and use of short-form adjectives, comparatives), case usage (e.g., use of various cases with and without prepositions), and challenges

---

<sup>8</sup> Note that both *картошка* and *колбаса* can also appear in plural in Russian, though less often.

associated with nonfinite verb forms (all the various participles and gerunds). Motion verbs and aspect (including biaspectuals) can be addressed, along with prefixation (since there are separate perfective and imperfective entries for most verbs in the SMARTool). Advanced Treasure Hunts probe more nuanced questions, for example, the special meanings of the so-called “second genitive” and “second locative” as opposed to the genitive and locative, as in these SMARTool examples:

*Петя выпил два **стакана** компота.* (Gen.Sing)

“Petya drank two glasses of compote.”

*Хотите **компоту**?* (Gen.Sing)

“Do you want to drink some compote?”

*Я долго собиралась на работу, поэтому завтрак пришлось на **бегу**.* (Loc.Sing)

“It took me a long time to get ready for work, so I had to eat breakfast on the run.”

*В **беге** главное — правильная техника, иначе легко получить травму.* (Loc.Sing)

“The most important thing in running is the right technique, otherwise it’s easy to get injured.”

Treasure Hunts are designed to inspire linguistic curiosity and to encourage learners to gather data and deduce patterns and to incorporate these patterns into their own repertoires. Our plan is to devise a score system so that each student can work toward a personal goal at each proficiency level.

## 5. Story Time

The goal of Story Time is to build skills and confidence in productive communication in Russian. Story Time helps learners become confident writers, and, when used in the classroom, speakers. Story Time activities take advantage of the fact that all lexemes in the SMARTool are searchable according to topic, and many lexemes belong to more than one topic. Filtering lexemes by topic facilitates the targeting of word forms that learners can use to construct coherent narratives. Since there are 18 topics, and many groups of words can be sourced from each

combination of topic and proficiency (CEFR) level, Story Time provides ample opportunities for learners at all levels from A1 to B2. The task for learners is to use the models of word forms and their typical contexts presented in the SMARTool to build their own sentences and, ultimately, paragraphs.

Table 1 gives examples of how Story Time Prompts and expectations can be scaled up from A1 to B2. In the first example, an A1 user is asked to write one sentence based on the SMARTool model sentences for two words on a given topic. As an example, we show the topic *магазин* [shopping], which in level A1 includes, among others, the words *купить* [buy] and *одежда* [clothing] that have been selected for this prompt. These are just two of 40 words available for the combination of Level A1 and *магазин* [shopping]. Every combination of level and topic presents many lexemes in the SMARTool for many more Story Time prompts. The user receives only the information in the first four rows of Table 1 and begins their work from the prompt. When the user consults the SMARTool entries for *купить* [buy] and *одежда* [clothing], they find sentences that model these constructions and collocations: *красивая одежда* [beautiful clothing], *удобная одежда* [comfortable clothing], *постирать одежду* [laundry clothing], *я хочу купить + acc* [I want to buy something], and *он/она купил/купила (себе) + acc* [he/she bought (him/herself) something]. Based on these models, the learner can write a sentence like *Я хочу купить (себе) красивую одежду* [I want to buy (myself) beautiful clothing] (among many other good answers). As the learner advances through proficiency levels, the prompts involve more and more difficult vocabulary, along with greater expectations for length and coherence of narration. With the prompt for level B2, the learner can write a whole paragraph about international trade and economics.

Story Time can be a part of self-study, a homework assignment, or a classroom assignment. In a classroom setting, Story Time can be a competitive and/or group assignment in which students can perform their stories orally, and fellow students can also check each other's work to see whether the use of word forms and constructions matches the models in the SMARTool. An alternative classroom activity is the co-creation of a larger narrative by combining several prompts and having students or teams of students take turns adding to a story one sentence at a time. Since the SMARTool vocabulary is quite large, it can potentially

source tens of thousands of Story Time prompts.<sup>9</sup> If implemented as a regular part of a daily or weekly study routine, Story Time is an efficient way to hone communication skills, combining building up a repertoire of idiomatic phrases with students' creative expression of their own ideas. Ideally, Story Time will be linked to an analyzer specially designed to give feedback to L2 Russian learners on their writing errors (see Reynolds et al., 2022).

*Table 1: Examples of Prompts for Story Time Activities Across Proficiency Levels and Topics*

CEFR level	A1
Number of words in Prompt	2
Task	Write 1 sentence
Example of Topic and Prompt	Topic: магазин [shopping] Prompt: купить, одежда
Word forms, constructions, and collocations modeled in SMARTool	красивая одежда, удобная одежда, постирать одежду, я хочу купить + асс, он/она купил/купила (себе) + асс
CEFR level	A2
Number of words in Prompt	3
Task	Write 2 sentences
Example of Topic and Prompt	Topic: погода [weather] Prompt: юг, тёплый, лить
Word forms, constructions, and collocations modeled in SMARTool	на юг, на юге, с юга, на улице теплее, льёт дождь

<sup>9</sup> Some of these prompts are available at <https://smartool.github.io/exercises/>.



CEFR level	B1
Number of words in Prompt	4
Task	Write 2–3 connected sentences
Example of Topic and Prompt	Topic: здоровье [health] Prompt: принимать, операция, желудок, анализ
Word forms, constructions, and collocations modeled in SMARTool	принимать лекарство, принимать участие в + loc, операция на + loc, операция проводится под общим наркозом, у +gen болит желудок, боль в желудке, расстройство желудка, анализ крови, результаты анализа
CEFR level	B2
Number of words in Prompt	5+
Task	Write a paragraph of 3–5 sentences
Example of Topic and Prompt	Topic: учёба/работа [study/work] Prompt: вкладывать, безграничный, биржа, ввоз, бюджет
Word forms, constructions, and collocations modeled in SMARTool	вкладывать в бизнес/акции, вкладывать деньги/доходы, безграничные возможности, безграничный доступ, колебания биржи, на бирже, биржа труда, ввоз товаров/оружия, заниматься ввозом, попытка ввоза, федеральный бюджет, деньги в бюджете на + асс, внести поправки в бюджет

## 6. Conclusion

Our goal is to take the next step in realizing the potential benefits of the SMARTool by offering designs to engage L2 learners in constructing their own understanding of Russian vocabulary and grammar. Both Treasure Hunt and Story Time are student-centered activities that encourage users to make and implement their own discoveries. These designs present the SMARTool as a space for experimentation and development for learners who will continue to find new words and phrases about which to ask “How do you say that in Russian?” Treasure Hunt and Story Time provide guided prompts for open-ended learning experiences that can transfer to unguided lifelong learning skills. The variety of prompt levels in Treasure Hunt and Story Time facilitate use even in classrooms with students at different proficiency levels, a challenge we often encounter in L2 Russian instruction. The open-source architecture of the SMARTool invites the creation of parallel SMARTools for other languages, along with the exercise designs suggested here.

## References

- Ackerman, F., Blevins, J. P., & Malouf, R. (2009). Parts and wholes: Patterns of relatedness in complex morphological systems and why they matter. In J. P. Blevins & J. Blevins (Eds.), *Analogy in grammar: Form and acquisition* (pp. 54–82). Oxford University Press.
- Criado-Perez, C. (2019). *Invisible women*. Abrams Press.
- Dąbrowska, E. (2004). *Language, mind and brain*. Georgetown University Press.
- D’Ignazio, C., & Klein, L. F. (2020). *Data feminism*. MIT Press. <https://doi.org/10.7551/mitpress/11805.001.0001>
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalizations in language*. Oxford University Press.
- Harvey Arce, N., & Cuadros Valdivia, A. (2020). Adapting competitiveness and gamification to a digital platform for foreign language learning. *International Journal of Emerging Technologies in Learning (IJET)*, 15(20), 194–209. <https://www.learntechlib.org/p/218340/>
- Hopp, H. (2010). Ultimate attainment in L2 inflection: Performance similarities between non-native and native speakers. *Lingua*, 120(4), 901–931. <https://doi.org/10.1016/j.lingua.2009.06.004>

- Janda, L. A. (2019). Businessmen and ballerinas take different forms: A strategic resource for acquiring Russian vocabulary and morphology. *Russian Language Journal*, 69(1), 175–193.
- Janda, L. A., & Tyers, F. M. (2021). Less is more: Why all paradigms are defective, and why that is a good thing. *Corpus Linguistics and Linguistic Theory*, 14(2), 109–141. <https://doi.org/10.1515/cllt-2018-0031>
- Jensen, I. N., Slabakova, R., Westergaard, M., & Lundquist, B. (2019). The Bottleneck Hypothesis in L2 acquisition: L1 Norwegian learners' knowledge of syntax and morphology in L2 English. *Second Language Research*, 36(1), 3–29. <https://doi.org/10.1177/0267658318825067>
- Karlsson, F. (1986). Frequency considerations in morphology. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, 39(1), 19–28. <https://doi.org/10.1524/stuf.1986.39.14.19>
- Kuznetsova, J. (2015). Are verbs politically correct? A corpus study of gender in Russian verbs. *Gender and Language*, 9(3), 391–434. <https://doi.org/10.1558/genl.v9i3.16650>
- Levenshtein, V. I. (1965/1966). Dvojnye kody s ispravleniem vypadenij, vstavok i zameščenij simvolov. *Doklady Akademii Nauk SSSR*, 163(4), 845–848. Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, 10(8), 707–710.
- Reynolds, R. J., Janda, L. A., & Nessel, T. (2022). A cognitive linguistic approach to analysis and correction of orthographic errors. *Russian Journal of Linguistics*, 26(2), 390–407. <https://doi.org/10.22363/2687-0088-30122>
- Slabakova, R. (2009). How is inflectional morphology learned? *Eurosla Yearbook*, 9(1), 56–75. <https://doi.org/10.1075/eurosla.9.05sla>
- Slabakova, R. (2014). The bottleneck of second language acquisition. *Foreign Language Teaching and Research*, 46(4), 543–559.
- Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Addison-Wesley.