Intelligent Tutor to Support Teaching and Learning of Tatar

Alsu Zakirova,[◇] Jue Hou,[†⋆] Anisia Katinskaia,[†⋆] Anh-Duc Vu,[†⋆] Roman Yangarber[⋆]
[◇]Moscow State University, Russia
[†]Department of Computer Science, University of Helsinki, Finland
[⋆]Department of Digital Humanities, University of Helsinki, Finland
zakirova.alsu112@student.msu.ru first.last@helsinki.fi

Abstract

This paper presents our work on tools to support the Tatar language, using Revita, a web-based Intelligent Tutoring System for language teaching and learning. The system allows the users—teachers and learners—to upload arbitrary authentic texts, and automatically creates exercises based on these texts that engage the learners in active production of language. It provides graduated feedback when they make mistakes, and performs continuous assessment, based on which the system selects exercises for the learners at the appropriate level. The assessment also helps the students maintain their learning pace, and helps the teachers to monitor their progress. The paper describes the functionality currently implemented for Tatar, which enables learners—who possess basic proficiency beyond the beginner level—to improve their competency, using texts of their choice as learning content. Support for Tatar is being developed to increase public interest in learning the language of this important regional minority, as well as to to provide tools for improving fluency to "heritage speakers"—those who have substantial passive competency, but lack active fluency and need support for regular practice.

1 Introduction

Tatar is a minority language spoken in the Russian Federation and by the Tatar diaspora worldwide. Although Tatar is an important Turkic language with over seven million speakers, it remains a low-resource language from the technological perspective, with little language technology to support its wider use online. This reduced online presence, in turn, limits and diminishes the overall vitality of the language.

Interest in second-language (L2) learning is continually increasing, with a growing number of resources available for learners at various proficiency levels. However, most of these resources either provide only an elementary introduction to the basics of the language, or try to increase proficiency by memorizing advanced vocabulary or complex grammatical structures, such as verb tenses. Despite this variety, it is difficult to find tools that make the learning process interactive and *personalized*—engaging the learners' interests and adapting to their level. The Revita approach to language learning and teaching[1] is founded on allowing the users themselves—students or teachers—to select any authentic material as learning content. The system then automatically generates exercises based on the chosen content, monitors the learner's performance on these exercises to assess the learner's proficiency in multiple dimensions, and adjusts the difficulty of the exercises according to the learner's current level. Currently, there is no similar online service for teaching Tatar to non-beginner students, using text material chosen by the students themselves. Implementing this plan will enable anyone to learn Tatar, using the latest methods from artificial intelligence and language technology.

Creating opportunities for learning Tatar and promoting its use within speaker communities is of great importance to supporting the language. Tatars form the largest linguistic minority in Russia, with diasporas in many other countries. It is crucial to stimulate interest in learning this language to preserve its heritage and expand its use geographically.

The need to create and maintain learning platforms such as Revita, as well as the importance of supporting the study of the Tatar language, underlies the relevance of this work. Intelligent support for language learning is a rapidly evolving and complex area of research. The problem becomes especially challenging in the case of *low-resource* languages: on one hand, the need is more urgent,

[1]revita.cs.helsinki.fi

Proceedings of the First Workshop on Natural Language Processing for Turkic Languages (SIGTURK 2024), pages 92–101
August 15, 2024

since many of the low-resource languages are endangered, and their speaker communities urgently require support. On the other hand, building intelligent tools for such languages is much more difficult due to the paucity of foundational tools and resources.

Some work is being done in this direction, e.g., by Apertium (Mirzakhalov et al., 2021; Forcada et al., 2011; Khanna et al., 2021). However, the availability of natural language processing (NLP) resources for Turkic languages, particularly the endangered ones, still lags far behind that of, e.g., the major European languages.

This paper introduces and describes the work on Tatar in the Revita system. Section 2 outlines the broad principles and capabilities of Revita, describes the work on the system, and explains the notion of "construct" within the framework. Section 3 details all the constructs implemented in the Revita platform and provides examples of exercises that can be created based on these constructs. Section 4 summarizes the results achieved during the adaptation of Revita to the Tatar language, and outlines the next steps for future work.

2 Features of the Revita system

2.1 System Capabilities

Tools for natural language processing (NLP) and automatic text analysis are understood to be central in the creation of platforms for language teaching and learning (Slavuj et al., 2015). Such platforms do not aim to *replace* the teacher, but rather aim to serve as an effective intelligent *assistant* to the teacher (Al Emran and Shaalan, 2014). Thanks to such systems, students can continue learning the language beyond school hours, and practice on their own time while tracking their progress independently. Developments focused on teaching rare languages are particularly valuable because it is more difficult to find teachers who speak these languages at the proper level.

At present, many platforms and applications exist for learners to get acquainted with a *new* language, and learn the basic structures using limited, pre-fabricated material. However, as students gradually acquire language skills, they often face a shortage of authentic and interesting material to practice more complex constructions at the intermediate to advanced level.

Further, learning a language is an ongoing process that requires a significant investment of time on the part of the learner. To achieve mastery, it is crucial to have a sufficient supply of practice material and exercises. Therefore, the automatic, intelligent generation of exercises, based on an unrestricted amount of text, can meet the needs of students aiming to reach advanced competency. Revita is designed to fulfill these requirements (Katinskaia et al., 2017, 2018).

Initially, the purpose of Revita, developed at the University of Helsinki, was to revitalize and support endangered Finno-Ugric languages (Katinskaia and Yangarber, 2018). More recently, this approach has been applied to language teaching and learning more generally, including for the "majority" languages. Currently, the Finnish and Russian languages are the most developed in terms of the richness of the kinds of exercises the system is able to generate, and the number of various grammatical concepts that it covers. It is used by teachers and students of Finnish and Russian at several universities. Other languages are under development, including major European languages (e.g., Italian, German, Swedish) and minority languages (e.g., Udmurt, Northern Sami). Revita has also been partially adapted for the endangered Turkic language Sakha (Ivanova et al., 2019).

An important aspect of Revita is that the approach is not intended for beginners, which distinguishes it from many other existing learning approaches and platforms. The approach assumes that the learner already knows some basic vocabulary (500–1000 words) and is familiar with elementary grammar. Revita tries to provide a "starter" library of texts for each language, but the main principle is to teach the language using materials that interest the learner. Students can independently choose the texts, based on which the exercises are generated, making use of the structure and vocabulary of the chosen material.

The system aims to act as a teacher's assistant: supporting continuous and effective learning, maintaining the students' motivation, and keeping their attention on the study objectives. To promote motivation and provide a variety of exercise modes, Revita employs various gamification features, as introduced in (Hou et al., 2022).

Before starting to work with texts, the student can take an adaptive test, usually consisting of 50–60 questions. This test estimates the level of language proficiency using several types of tasks: identifying a word by its meaning, choosing the correct structure, testing knowledge of phraseol-

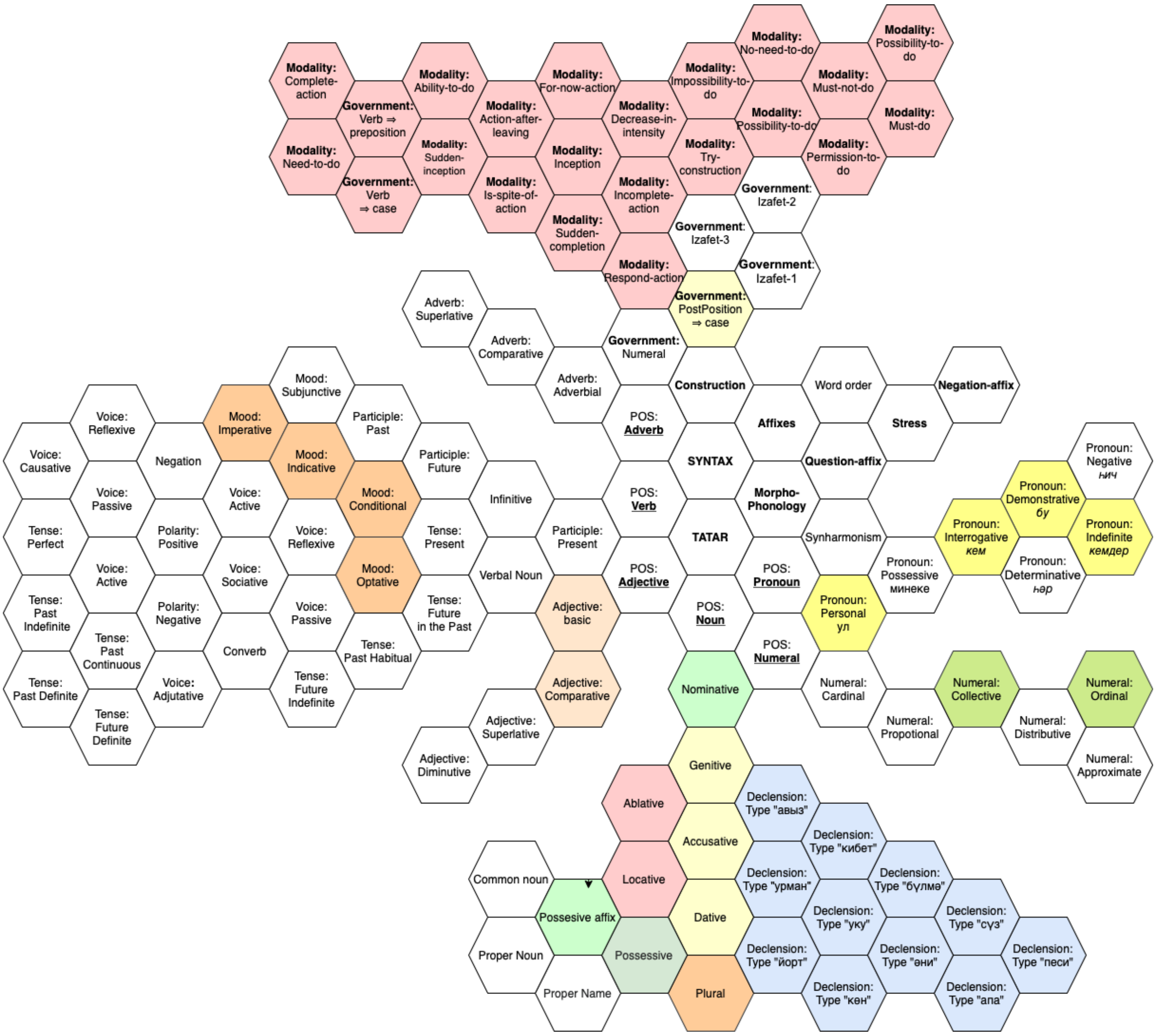

Figure 1: Heatmap of constructs for Tatar in Revita.

ogy and expressions, orthography, grammatical forms of words, etc.

To begin practicing within Revita, the student selects a text. Texts can be uploaded directly by the learner or by the teacher. Once the text has been analyzed and exercises have been created, the Preview mode allows the students to familiarize themselves with the grammatical structures (and vocabulary) present in the text. Since the platform adapts to the student's level over time, the learner can immediately start completing exercises proposed by the system. If the number of erroneous answers is high, the system will generate tasks based on easier grammatical topics.

Revita currently creates three types of exercises. "*Cloze*" (fill-in-the-gap) exercises require the student to produce the correct grammatical form based on the context of the word; the hint given to the student is the lemma (base form of the word). In multiple-choice (MC) tasks, the learner is asked to select the correct option from a drop-down list of answers. The challenge in generating MC questions is automatically finding appropriate "distractors"—options which are not suitable for the context, and yet not obviously incorrect (which would make the exercise too easy and uninteresting). Listening exercises are aimed at training auditory perception of spoken language. In auditory comprehension exercises, the student needs to enter the word pronounced by a speech generator. The system provides a set of settings to adjust the difficulty level of exercises. The student can select the type of exercises as desired.

Personalized feedback is a central aspect of the practice mode in Revita. The system analyzes learner errors and presents hints that help the student find the correct answer independently—rather than giving away the correct answer in case

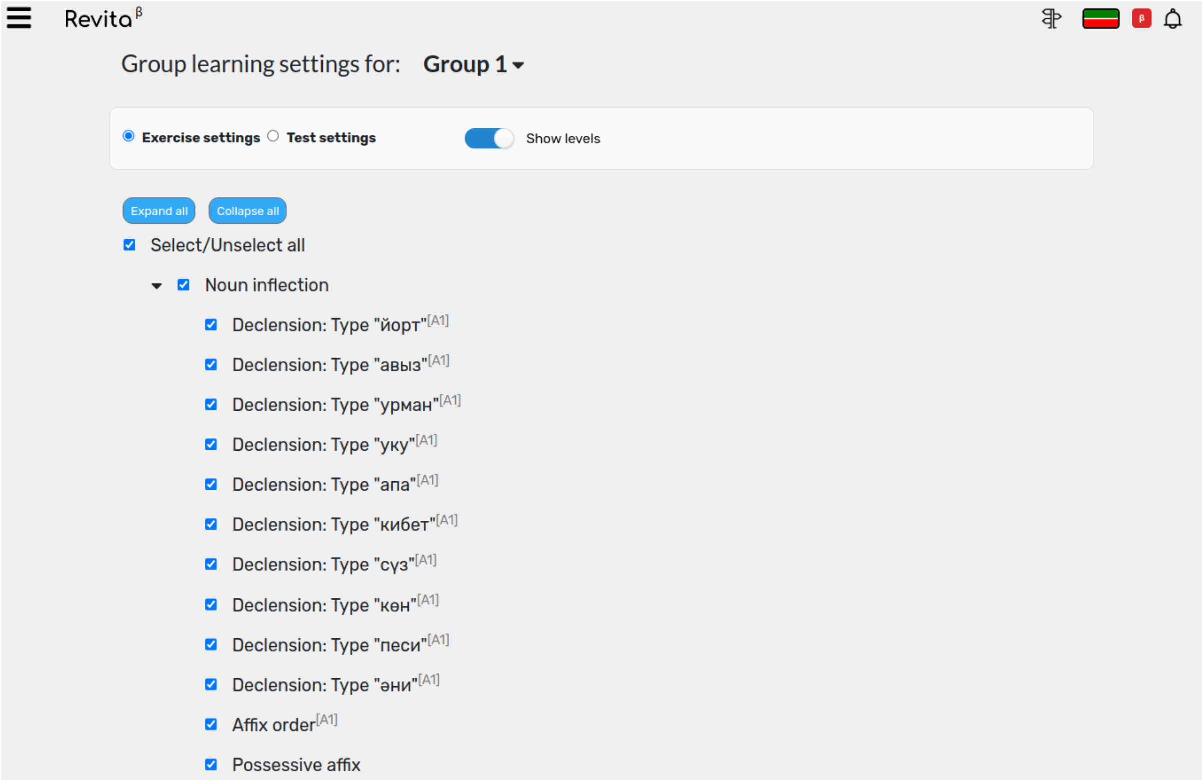

Figure 2: Selection of (top 12) constructs for learning during exercise sessions.

of a mistake. This way, the student not only discovers whether the answer is correct, but also understands which grammatical features need to be changed to complete the task correctly. After several incorrect attempts, the system will show the correct answer, with a detailed explanation of the correct form of the word or phrase.

2.2 Construct-centred Learning

The Revita approach treats *constructs* as the central unit of L2 teaching and learning, as introduced in prior research (Boas, 2022; Katinskaia et al., 2023). Language constructs may describe individual word forms, phrases, or clauses, and encompass topics on various linguistic levels—grammatical, lexical, orthographic, morphological, etc.—for each language (Katinskaia and Yangarber, 2018).

When developing a new language in Revita, we use the notion of a "chunk"—in linguistics, a chunk is a collocation or construction, which is regulated by certain rules. Chunks have a main word that controls and dependent ones, see, e.g., Figure 4. In this example, the analysis of the sentence Кояшка таба әйлән! (*"Turn towards the sun"*) is presented—the analyzer identifies the post-positional construction in the sentence, and highlights the components of the chunk: the postposition—таба ("towards"), which governs the noun—кояш=ка (*"sun"*)—which must be in the required (dative) case.

All constructs implemented for a given language can be graphically seen on the *heatmap*, which can be examined in the learner's profile. This allows both the student and the teacher to track progress, Figure 1. The size of the cell indicate how many times this construct has been practiced (relative to other constructs), and its color shows how well the construct is mastered over the selected period of time (which can be selected by the user). Using the heatmap, the teacher can visually assess the level of mastery of the lesson's topic (construct) by an individual student or by the group as a whole. As the number of constructs increases, the map also expands.

Constructs are linked to specific language proficiency levels. If a student believes her level is, e.g., B1, she can choose to study all constructs related to this level and below. In the system settings, it is possible to select constructs, which determines which exercises will be generated. In case a learner evaluates her level of language proficiency as insufficient for training a particular construct, she can disable exercises for this construct.

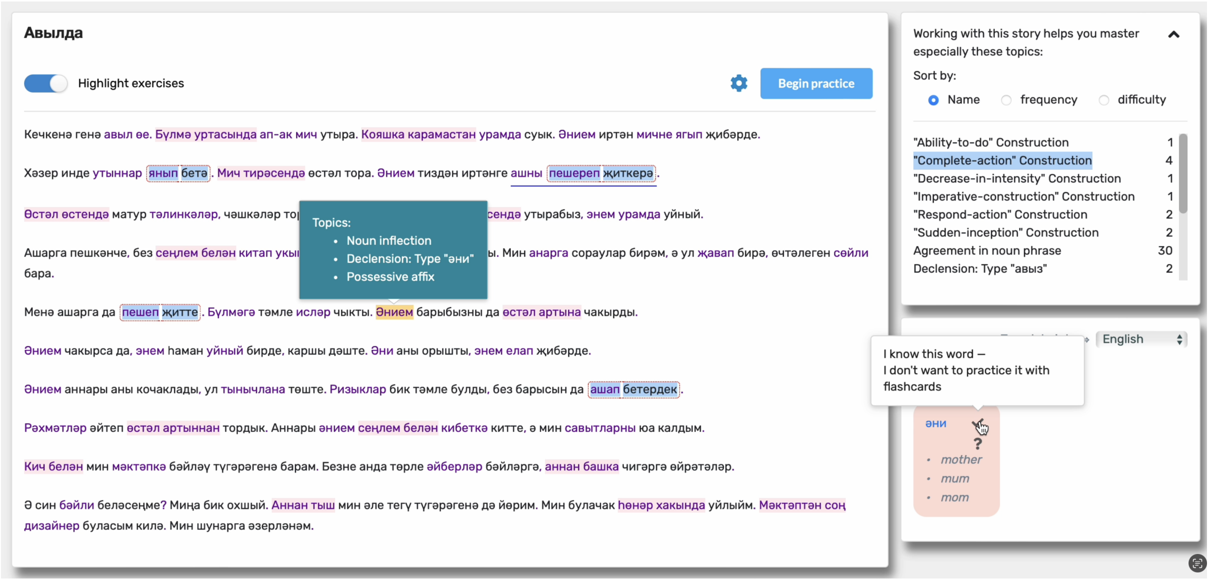

Figure 3: Preview Mode.

Conversely, the learner can request tasks to work with more complex material.

Exercises on finding the correct grammatical form of a word in its context (and on auditory perception) are tied to language constructs, allowing for the adjustment of task complexity. This is convenient for teachers—when working with a group of students, they can choose constructs relevant to the lesson and focus on training only those.

The system displays a tree structure of constructs, with each item attached to a CEFR level[2] corresponding to the construct, see Figure 2, and *topics* in upper-right box in Figure 3. The latter tells the learner which grammatical concepts will be presented during practice with this text.

To view all constructs contained in the chosen text, the system offers the Preview Mode, Figure 3. Hovering over each word brings up the list of constructs attached to it by Revita. The student can preview a selected text *before* practice, view the highlighted constructions and the of constructs identified for each word. Additionally, when the student clicks on a word, Revita displays its translation into a preferred language.

2.3 Technical Implementation of Constructs

To adapt the system to a new language (here: Tatar) we need to specify and implement the constructions of the language—e.g., rules for syntactic agreement, government, and many more. The system performs *chunking* based on these rules, and uses the chunks when creating exercises.

Syntactic government is a particularly important area in L2 teaching and learning. A government bank is a collection of declarative rules—each rule describes an essential pattern of interaction between, e.g., a head verb and its syntactic dependents which it governs—nouns, pre- or post-positional phrases, etc. These banks describing the government of adpositions, nouns, adjectives and verbs—as well as banks of more complex constructions—are, of course, language-specific, they must be created for each language separately. Currently, government banks in Revita are constructed manually, though our recent work attempts to extend government banks automatically by probing pre-trained large language models (Hou et al., 2024; Klyshinsky et al., 2023).

According to the defined rules, the system finds constructions in the text, as shown in Figure 3. The identified constrictions are then used to generate exercises. For example, the lemma of the governed noun can be used as the hint in the exercise (so the student must inflect it in the correct case), or the governed adposition (post-position in Tatar) can be replaced by a list of options from which the student must choose the correct one.

The distractors for multiple-choice exercises are also created using language-specific rules. This is a hard problem, since the tutor must avoid both those options that are a. *obviously incorrect* in the given context (which would make the exercise too easy), and b. those that could be *also correct* in the

[2] Common European Framework Reference

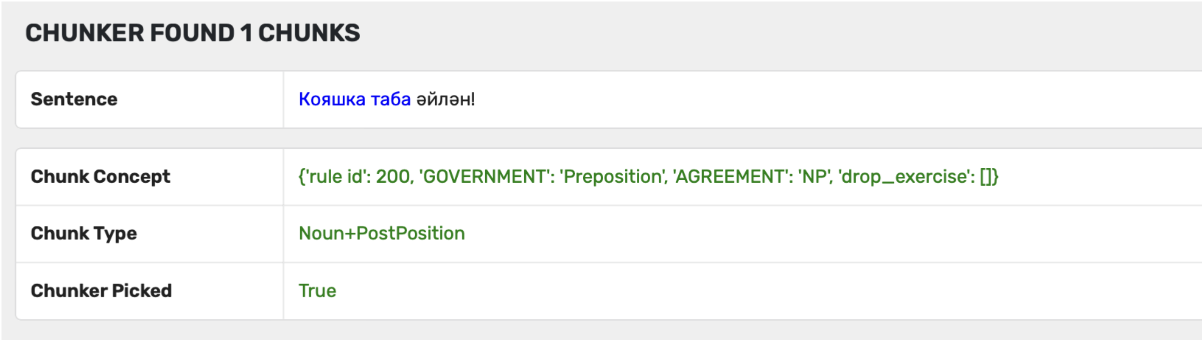

Figure 4: Example of post-positional construction (in developer's user interface for testing constructions.

context (making the exercise too difficult or impossible to solve). Deciding which distractors are suitable in a given context is an important problem that we are actively researching (Katinskaia et al., 2019; Katinskaia and Yangarber, 2021, 2023).

At present, the Apertium morphological analyzer is used for Tatar. The analyzer can recognize the grammatical form of a word, and generate required forms based on a lemma, which will be used for constructing future exercises (in particular, distractors in MC questions). The analyzer is still under development, but it partially meets the needs for analyzing forms, recognizing structures and creating exercises.

As the quality of morphological analysis improves in the future, we can expect to be able to make enhancements to the quality of the exercises. As mentioned in the introduction, the lack of foundational resources for Turkic languages, and Tatar in particular, is a major bottleneck, ultimately limiting the quality of downstream applications.

3 Tatar Constructs Implemented in Revita

3.1 Basic Constructs

The list of constructs that are currently implemented can be viewed both in the heatmap and in the system's settings. The selectable constructs are shown in these views. Not all of them are fully recognized by the system at present. Basic constructs include, e.g., declension of nouns, degrees of comparison of adjectives, tenses of verbs, etc. Tatar constructs are listed based on the inventory in the *Guide to the Tatar language and Tatar grammar* (Guzev, 2015; Mansurova, 2018; Nigmatullina, 2011; Nurmukhametova, 2008; Sharafutdinova, 2018).

Basic constructs are formulated at the beginning of development stage, since they form the foundation for further development. On the heatmap, basic constructs are displayed as branching out from the main parts of speech. At the top of the heatmap, more complex constructs/constructions are highlighted in light red (in the Figure).

3.2 Constructions

The more complex constructions are developed based on their descriptions in the works listed above, Constructions implemented so far include:

- post-positional constructions;
- verbal constructions;
- and modality constructs.

At present, the system recognizes 67 post-positional constructions, 344 verbal constructions, and 18 modality constructs. A detailed analysis for these constructions can be explored in the "Grammar Tester" interface, as shown in the example for a post-positional construction, Figure 4.

This interface is intended exclusively *for developers*. It provides detailed insights into the functionality and performance of the system's analyzers, and helps the developers tune the recognition algorithms. The system highlights post-positional constructions (Galiyeva, 2020) in blue, showing a detailed analysis of these constructions.

The structure and analysis of verbs can also be examined using this tool, see the example in Figure 5. This demonstrates the analysis of the sentence containing a verb construction: Тәртипкә гадәтләнергә кирәк! (*"One must get used to the order!"*). The main word of the construction—the verb гадәтләнергә (*"to get used to"*)—and the noun dependent on it тәртип=кә (*"[to the] order"*), where case=Dative. Note, in the Preview Mode, for verb constructions that require post-positional control (Gatiatullin, 2012), both the entire verb-government construction and the

Found 1 PATTERN/CONSTRUCTION

Sentence	Тәртипкә гадәтләнергә кирәк!

name	Тәртипкә+гадәтләнергә
pattern_type	Verb
rule id	None
source	government
target	гадәтлән

GOVERNMENT HEAD:
{'base': 'гадәтлән'}

ARGUMENT:
{'HEAD': 'Noun', 'CASE': 'Dative'}

Figure 5: Verbal Construction.

post-positional construction inside it will be highlighted.

Modality constructions (Gatiatullin, 2012; Tatevosov, 2018) are of particular interest, as their correct use can indicate that language proficiency has progressed to a higher level. Each of these constructions has a logical, concise name, which helps the student remember its meaning, see Figure 6. This example shows the structure and characteristics of a modal construction in the sentence Мин кибеткә барам, ә син укый тор. (*"I will go to the store, and you stay here and study"*). The modal meaning is conveyed through the serial verb construction: the 3SG.PRS verb укый (*"study"*) followed by the imperative verb тор (*"stay"*).

3.2.1 Examples of Exercises

In this section, we present examples of exercises that can be generated using these constructs. Various types of exercises implemented in the system will be described, along with the various kinds of feedback that the student can receive.

When working with the selected text, the student may be asked to inflect a noun or pronoun into the correct form:

- Original text: Мәктәптән соң кунакка барам. (*"After school I will go for a visit"*).
- Task: [мәктәп] соң ("after [school]")—the *lemma* of the noun (мәктәп) is presented as the hint for this cloze exercise—which must be inflected correctly by the learner.
- Correct answer: мәктәптән (*"school"*), Case=ablative.
- If the learner answers incorrectly, the system will offer *graduated feedback*—a sequence of increasingly more specific hints on each attempt:
  1. Pay attention to the part of speech.
  2. Choose the correct case.
  3. Inflect the noun мәктәп into the correct case, as required by the following post-position.
  4. Inflect the noun мәктәп into the ablative case.

In multiple-choice exercises, the student may be offered several options as distractors: various post-positions or post-positional words, as well as different forms of a noun or a pronoun:

- Original text: Ул минем артыма яшеренде (*"He hid behind me"*).
- Task: [янына/артыма/хакында] яшеренде—multiple-choice menu of options.
- Correct answer: артыма (*"behind"*).
- Graduated feedback:
  1. Which post-position can follow the pronoun in the genitive case?
  2. Recall the meaning of the post-position.
  3. Translate *hid behind me (behind my back)*.

In an exercise with *modality constructions*, the learner may be asked to select the correct auxil-

CHUNKER FOUND 1 CHUNKS	
Sentence	Мин кибеткә барам, ә син укый тор
Chunk Concept	{'MODALITY': 'imperative-construction', 'rule id': 113, 'ANALYTIC': True, 'NUMBER': 'Singular', 'PERSON': '2', 'TENSE': 'Imperative', 'CHUNK_RULE': 'respond-action', 'drop_exercise': []}
Chunk Type	Verb+Verb
Chunker Picked	True

Figure 6: Example of Modality Construct—in the Grammar Tester interface.

iary verb, or the form of the "semantic" (i.e., the meaning-bearing) verb.

- Original text: Мин кибеткә барам, ә син укый тор. (*"I'll go to the store, and you stay here and read"*).
- Task 1: син укый [тор/башла/ал]—multiple-choice menu of options.
- Task 2: син укый [...]—a gap.
- Correct answer: [тор].
- Graduated feedback:
    1. Recall how the imperative construction is formed.
    2. Try to translate the phrase *"stay and read."*
    3. Use an auxiliary verb appropriate for this context.

Since the choice of an auxiliary verb may not always be unambiguous, the name of the construction (or its translation) can be used as a hint.

4 Conclusions

This paper presents the current state of our work on adapting, Revita, a system for L2 teaching and learning, to Tatar—a low-resource Turkic language, spoken by over 7M people. It describes in some detail the constructs implemented to date. To the best of our knowledge, this is the first work dedicated to the development of L2 teaching and learning tools specifically for Tatar at the intermediate to advanced levels, based on state-of-the-art technologies available at present.

Work is on-going on identifying constructs most useful for teaching, and classifying them by proficiency levels, in accordance with the scale of assessment of language competencies. The next major development phase is to extend the inventory of constructions to include more complex syntactic constructs, with a particular interest in synthetic subordinate clauses (Zakharova, 2016).

We are also working on expanding the selection of texts on various topics to create a more complete open library, to give learners who do not have a source of their own texts a wider choice from the public library. Having more texts will increase the amount of content for training AI models—for example, using transfer learning from Turkish—a related, higher-resource language—which may help address problems of low-resource languages. An essential system component, found to be extremely useful in the development of other languages, is a syntactic dependency parser. Such a parser (of reasonable quality) is not available for Tatar at present. When one becomes available, the quality of the analysis—and the variety and quality of the automatically generated exercises—will progress to the next level.

Further work will focus on developing the platform to support Tatar. The ultimate goal is to bring the Tatar Revita to a level of functionality that can be deployed for teaching and learning Tatar, e.g., in schools with teachers, as is done for other languages for which richer resources are available—currently Finnish and Russian. We also hope that this work will contribute to stimulating global interest in the study and development of Tatar—and other low-resource languages in need of support—using the latest NLP technologies and theories of L2 acquisition.

References

Mustafa Al Emran and Khaled Shaalan. 2014. A survey of intelligent language tutoring systems. In *Advances in Computing, Communications and Informatics*, pages 393–399. IEEE.

Hans C Boas. 2022. From construction grammar(s) to pedagogical construction grammar. *Directions for pedagogical construction grammar. Learning and teaching (with) constructions*, pages 3–43.

Mikel L Forcada, Mireia Ginestí-Rosell, Jacob Nordfalk, Jim O'Regan, Sergio Ortiz-Rojas, Juan Antonio Pérez-Ortiz, Felipe Sánchez-Martínez, Gema Ramírez-Sánchez, and Francis M Tyers. 2011. Apertium: a free, open-source platform for rule-based machine translation. *Journal of machine translation*, 25(2).

Gatiatullin Galiyeva. 2020. Semantics of constructions with the postposition 'belan' in the Tatar language: analysis of corpus data. *Philological Sciences. Questions of theory and practice*.

Suleimanov Gatiatullin. 2012. Model of verbose constructions of the Tatar language: analytical forms. *Kazanskaya Nauka*.

Guzev. 2015. *Theoretical Grammar of the Turkish Language*.

Jue Hou, Anisia Katinskaia, Lari Kotilainen, Sathianpong Trangcasanchai, Anh-Duc Vu, and Roman Yangarber. 2024. What do transformers know about government? In *Proceedings of the Joint Conference on Computational Linguistics and Language Resources and Evaluation*, Torino, Italy.

Jue Hou, Ilmari Kylliäinen, Anisia Katinskaia, Giacomo Furlan, and Roman Yangarber. 2022. Applying gamification incentives in the Revita language-learning system. In *Proceedings of the 9th Workshop on Games and Natural Language Processing within the 13th Language Resources and Evaluation Conference*, pages 7–16, Marseille, France. European Language Resources Association.

Sardana Ivanova, Anisia Katinskaia, and Roman Yangarber. 2019. Tools for supporting language learning for Sakha. In *22nd Nordic Conference on Computational Linguistics (NoDaLiDa)*.

Anisia Katinskaia, Jue Hou, Anh-duc Vu, and Roman Yangarber. 2023. Linguistic constructs represent the domain model in intelligent language tutoring. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, pages 136–144, Dubrovnik, Croatia. Association for Computational Linguistics.

Anisia Katinskaia, Sardana Ivanova, and Roman Yangarber. 2019. Multiple admissibility in language learning:: Judging grammaticality using unlabeled data. In *Workshop onBalto-Slavic Natural Language Processing*, pages 12–22. The Association for Computational Linguistics.

Anisia Katinskaia, Javad Nouri, and Roman Yangarber. 2017. Revita: a system for language learning and supporting endangered languages. In *6th Workshop on NLP for CALL and 2nd Workshop on NLP for Research on Language Acquisition, at NoDaLiDa*, Gothenburg, Sweden.

Anisia Katinskaia, Javad Nouri, and Roman Yangarber. 2018. Revita: a language-learning platform at the intersection of ITS and CALL. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

Anisia Katinskaia and Roman Yangarber. 2018. Digital cultural heritage and revitalization of endangered Finno-Ugric languages. In *Proceedings of the 3rd Conference on Digital Humanities in the Nordic Countries*, Helsinki, Finland.

Anisia Katinskaia and Roman Yangarber. 2021. Assessing grammatical correctness in language learning. In *Proceedings of the 16th Workshop on Innovative Use of NLP for Building Educational Applications*, pages 135–146.

Anisia Katinskaia and Roman Yangarber. 2023. Grammatical error correction for sentence-level assessment in language learning. In *Workshop on Innovative Use of NLP for Building Educational Applications*, pages 488–502. The Association for Computational Linguistics.

Tanmai Khanna, Jonathan N Washington, Francis M Tyers, Sevilay Bayatlı, Daniel G Swanson, Tommi A Pirinen, Irene Tang, and Hèctor Alòs i Font. 2021. Recent advances in Apertium, a free/open-source rule-based machine translation platform for low-resource languages. *Journal of Machine Translation*, pages 1–28.

Eduard Klyshinsky, Anna Bogdanova, and Mikhail Kopotev. 2023. Towards a corpus-based dictionary of verbal government for the Russian language. *Journal of Linguistics/Jazykovedný casopis*, 74(1):173–181.

Mansurova. 2018. Agreement of numerals in Arabic and Tatar languages. *Modern Muslim World*.

Jamshidbek Mirzakhalov, Anoop Babu, Duygu Ataman, Sherzod Kariev, Francis Tyers, Otabek Abduraufov, Mammad Hajili, Sardana Ivanova, Abror Khaytbaev, Antonio Laverghetta Jr, et al. 2021. A large-scale study of machine translation in the Turkic languages. *arXiv preprint arXiv:2109.04593*.

Nigmatullina. 2011. *Learning the Tatar language (rules and exercises)*.

Nurmukhametova. 2008. *Collection of rules on the Tatar language for Russian-speaking students*.

Sharafutdinova. 2018. Isafet constructions in the Tatar language. *Modern Muslim World*.

Vanja Slavuj, Božidar Kovačić, and Igor Jugo. 2015. Intelligent tutoring systems for language learning. In *Proceedings of the 38th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*. IEEE.

Tatevosov. 2018. On the subparadigm of analytical forms of the Tatar verb. *Ural-Altai Research*.

Zakharova. 2016. Concessive clauses in the Yakut, Tatar and Turkish languages: a comparative analysis of synthetic and synthetic-analytic constructions. *New Science: Current State and Ways of Development*.