

Russian Linguistics Research

28 Feb 2021

Morfologija Web Scraping 2021 Power Point Presentation

Perry Koob

Irina V. Ivliyeva

Missouri University of Science and Technology, ivliyeva@mst.edu

Follow this and additional works at: https://scholarsmine.mst.edu/russian_linguistics_research

 Part of the [Russian Linguistics Commons](#)

Recommended Citation

Koob, Perry and Ivliyeva, Irina V., "Morfologija Web Scraping 2021 Power Point Presentation" (2021).
Russian Linguistics Research. 2.
https://scholarsmine.mst.edu/russian_linguistics_research/2

This Data is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Russian Linguistics Research by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

The A.A. Zalizniak grammatical
dictionary of the Russian language

Web Scraping Project 2020-2021

February 26, 2021

The research team

Perry B. Koob, koobp@mst.edu

Database Administrator/System Administrator
Academic Technology Support Team
Missouri S&T Information Technology

Dr. Irina V. Ivliyeva, ivliyeva@mst.edu

Professor of [Russian](#), Arts, Languages, and Philosophy Department
Missouri University of Science and Technology

What is this presentation about?

We will review the general outline of steps for the **The A.A. Zalizniak grammatical dictionary of the Russian language (Morfologija.ru) Project.**

For inquiries on the theoretical principles, research methodology, vocabulary inventory, and technical specifics, go to:

[Koob and Ivliyeva, 2020](#)

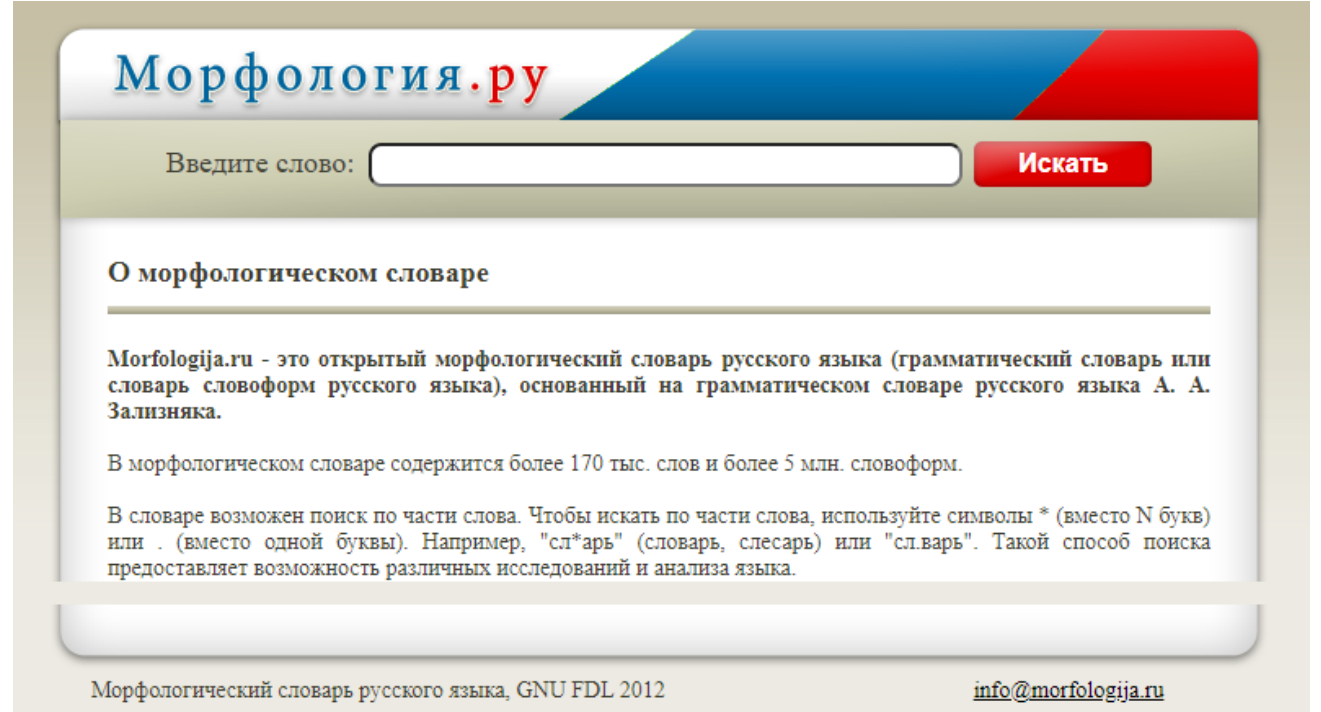
[Ivliyeva, I. 2020](#)

Morfologija.ru

Morfologija.ru is a grammatical dictionary of the Russian language. It contains more than 170,000 words and over 5,000,000 wordforms.

For our research purposes we are interested only in the search data from the future tense, imperatives, and gerund forms of Russian infinitives.

Morfologija.ru



The screenshot shows the homepage of the Morfologija.ru website. At the top, the site's name "Морфология.ру" is displayed in a blue and red font. Below the name is a search bar with the placeholder text "Введите слово:" and a red button labeled "Искать". The main content area is titled "О морфологическом словаре" and contains the following text:

Мorfologija.ru - это открытый морфологический словарь русского языка (грамматический словарь или словарь словоформ русского языка), основанный на грамматическом словаре русского языка А. А. Зализняка.

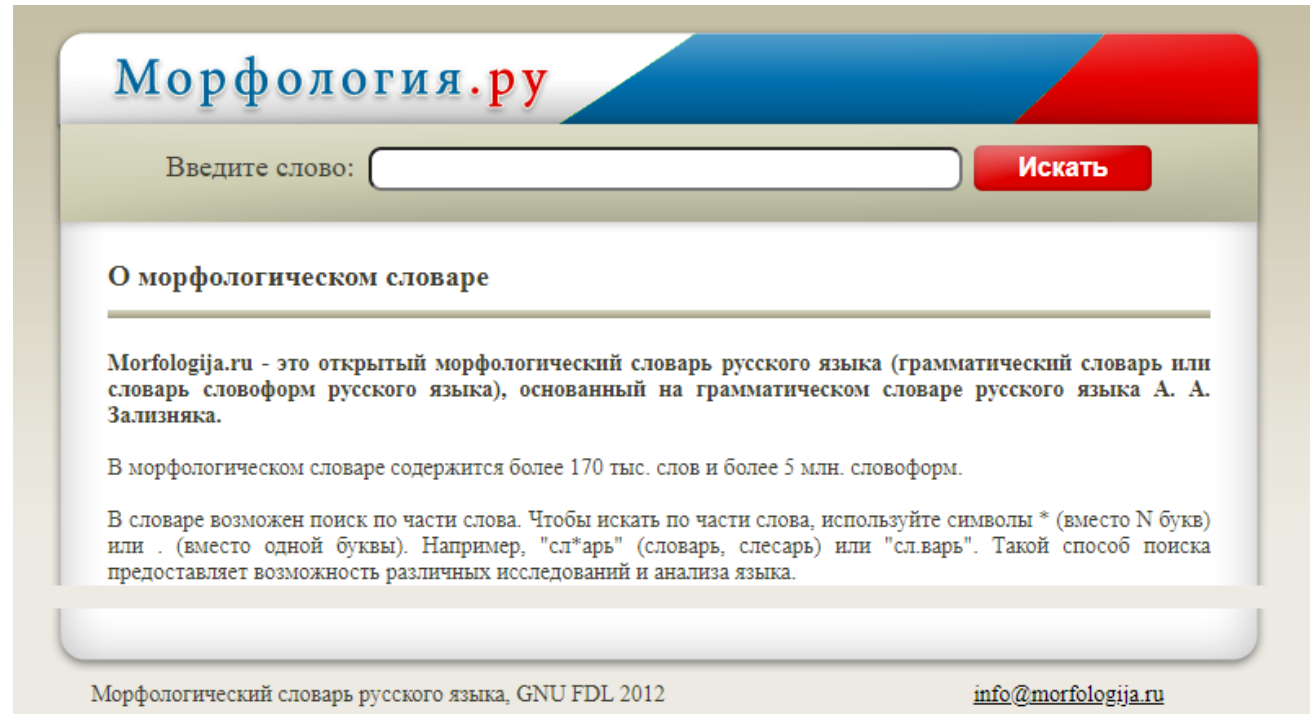
В морфологическом словаре содержится более 170 тыс. слов и более 5 млн. словоформ.

В словаре возможен поиск по части слова. Чтобы искать по части слова, используйте символы * (вместо N букв) или . (вместо одной буквы). Например, "сл*арь" (словарь, слесарь) или "сл.варь". Такой способ поиска предоставляет возможность различных исследований и анализа языка.

At the bottom of the page, the text "Морфологический словарь русского языка, GNU FDL 2012" is visible on the left, and the email address "info@morfologija.ru" is on the right.

URL Hacking

We start by submitting a search through the website and decoding the URL to determine how we can automate the generation of URL requests to Morfologija.ru.



The screenshot shows the homepage of Morfologija.ru. At the top, the site's name "Морфология.ру" is displayed in a blue and red gradient. Below the header is a search bar with the placeholder text "Введите слово:" and a red "Искать" button. The main content area features a section titled "О морфологическом словаре" (About the morphological dictionary). The text in this section describes the site as an open morphological dictionary of the Russian language, based on the dictionary by A. A. Zaliznjak. It mentions that the dictionary contains over 170,000 words and over 5 million word forms. It also explains the search syntax, such as using asterisks for wildcards and dots for single-letter replacements. At the bottom of the page, there is a footer with the text "Морфологический словарь русского языка, GNU FDL 2012" and the email address "info@morfologija.ru".

Морфология.ру

Введите слово: [Искать](#)

О морфологическом словаре

Мorfologija.ru - это открытый морфологический словарь русского языка (грамматический словарь или словарь словоформ русского языка), основанный на грамматическом словаре русского языка А. А. Зализняка.

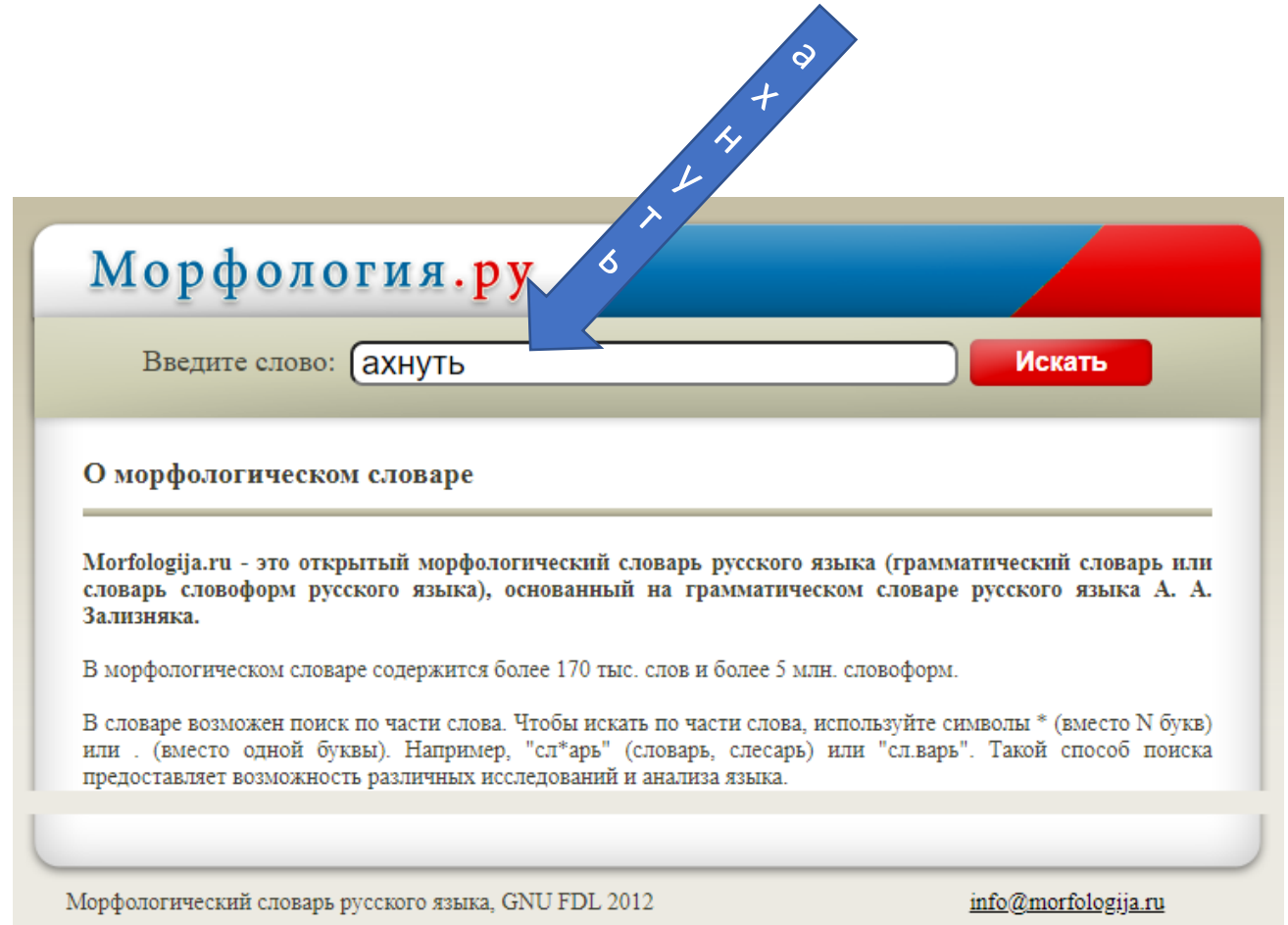
В морфологическом словаре содержится более 170 тыс. слов и более 5 млн. словоформ.

В словаре возможен поиск по части слова. Чтобы искать по части слова, используйте символы * (вместо N букв) или . (вместо одной буквы). Например, "сл*арь" (словарь, слесарь) или "сл.варь". Такой способ поиска предоставляет возможность различных исследований и анализа языка.

Морфологический словарь русского языка, GNU FDL 2012 info@morfologija.ru

Search for the infinitive “ахнуть”

We will use the infinitive “ахнуть” as our test word to figure out, document, and illustrate the procedure



The screenshot shows the website Morfologija.ru. At the top, the site name is displayed in blue and red. Below it, there is a search bar with the text "Введите слово:" followed by a text input field containing "ахнуть". To the right of the input field is a red button labeled "Искать". A blue arrow points from the top right towards the search bar, with the Cyrillic letters "Б Т У Н Х Ф" written along its shaft. Below the search bar, there is a section titled "О морфологическом словаре" followed by a horizontal line. The text below the line describes the website as an open morphological dictionary of the Russian language, based on the dictionary of A. A. Zaliznjak. It mentions that the dictionary contains over 170,000 words and over 5 million word forms. It also explains that the dictionary allows for searching by parts of a word using symbols like asterisks and dots.

Морфология.ру

Введите слово:

О морфологическом словаре

Morfologija.ru - это открытый морфологический словарь русского языка (грамматический словарь или словарь словоформ русского языка), основанный на грамматическом словаре русского языка А. А. Зализняка.

В морфологическом словаре содержится более 170 тыс. слов и более 5 млн. словоформ.

В словаре возможен поиск по части слова. Чтобы искать по части слова, используйте символы * (вместо N букв) или . (вместо одной буквы). Например, "сл*арь" (словарь, слесарь) или "сл.варь". Такой способ поиска предоставляет возможность различных исследований и анализа языка.

Морфологический словарь русского языка, GNU FDL 2012 info@morfologija.ru

Results

A standard set of results can be seen here.

Infinitive

Present, past, future, Imperative

Active Present Participles

Active Past Participles

Passive Present Participles

Passive Past Participles

morfologija.ru/словоформа/ахнуть

Морфология.ру

Введите слово:

Словоформы слова "ахнуть"

ахнуть → **ахнуть** – инфинитив, действ. зал.

Часть речи: инфинитив - **ахнуть**

Часть речи: глагол

	Настоящее время	Прошедшее время	Будущее время	Повелительное наклонение
Я		ахнул; ахнула; ахнуло	ахну	
Ты		ахнул; ахнула; ахнуло	ахнешь	ахни
Он/она		ахнул; ахнула; ахнуло	ахнет	
Мы		ахнули	ахнем	ахнем; ахнемте
Вы		ахнули	ахнете	ахните
Они		ахнули	ахнут	

Часть речи: деепричастие

Несовершенный вид		Совершенный вид	
		ахнув;	ахнувши

Часть речи: причастие

Действительное причастие:

	Настоящее время			
	Единственное число		Средний род	Множественное число
Мужской род	Женский род			
Им.				
Рд.				
Дт.				
Вн.				
Тв.				
Пр.				

	Прошедшее время			
	Единственное число		Средний род	Множественное число
Мужской род	Женский род			
Им.	ахнувший	ахнувшая	ахнувшее	ахнувшие
Рд.	ахнувшего	ахнувшей	ахнувшего	ахнувших
Дт.	ахнувшему	ахнувшей	ахнувшему	ахнувшим
Вн.	ахнувшего; ахнувший	ахнувшую	ахнувшее	ахнувших; ахнувшие
Тв.	ахнувшим	ахнувшей; ахнувшею	ахнувшим	ахнувшими
Пр.	ахнувшим	ахнувшей	ахнувшим	ахнувших

Страдательное причастие:

	Настоящее время			
	Единственное число		Средний род	Множественное число
Мужской род	Женский род			
Им.				
Рд.				
Дт.				
Вн.				
Тв.				
Пр.				

	Прошедшее время			
	Единственное число		Средний род	Множественное число
Мужской род	Женский род			
Им.	ахнутый	ахнутая	ахнутое	ахнутые
Рд.	ахнутого	ахнутой	ахнутого	ахнутых
Дт.	ахнутому	ахнутой	ахнутому	ахнутым
Вн.	ахнутого; ахнутый	ахнутую	ахнутое	ахнутых; ахнутые
Тв.	ахнутым	ахнутой; ахнутою	ахнутым	ахнутыми
Пр.	ахнутом	ахнутой	ахнутом	ахнутых

Часть речи: кр. причастие

Decoding the URL

Looking at the URL we can see it is composed of the site name, followed by a slash,

the Russian word for “word form”, словоформа, followed by a slash,

then the word we are searching for: ахнуть.

We can craft URLs for our automated Web Scraping by just replacing last word with the word we want to search for. The examples below refer to the prefixed derivatives as well as the base-words, e.g. ахать – ахнуть – взахать; отбахать.

For example:

<https://www.morfologija.ru/словоформа/ахнуть>

<https://www.morfologija.ru/словоформа/взахать>

<https://www.morfologija.ru/словоформа/отбахать>

The screenshot shows the website [morfologija.ru/словоформа/ахнуть](https://www.morfologija.ru/словоформа/ахнуть). The page title is "Морфология.ру". The search bar contains the word "ахнуть" and a red "Искать" button. The results section is titled "Словоформы слова 'ахнуть'". It shows the word "ахнуть" is an infinitive (инфинитив, действ. зал.). The part of speech is "инфинитив - ахнуть". Below this, it lists the part of speech as "глагол" and provides a conjugation table for the present, past, and future tenses, as well as the imperative mood. The table has five columns: "Настоящее время", "Прошедшее время", "Будущее время", and "Повелительное наклонение". The rows represent the pronouns "Я", "Ты", "Он/она", "Мы", "Вы", and "Они".

	Настоящее время	Прошедшее время	Будущее время	Повелительное наклонение
Я		ахнул; ахнула; ахнуло	ахну	
Ты		ахнул; ахнула; ахнуло	ахнешь	ахни
Он/ она		ахнул; ахнула; ахнуло	ахнет	
Мы		ахнули	ахнем	ахнем; ахнемте
Вы		ахнули	ахнете	ахните
Они		ахнули	ахнут	

Below the table, it lists the part of speech as "деепричастие" and provides a table for the imperfect and perfect forms. The imperfect form is empty, and the perfect form is "ахнув; ахнувши".

Несовершенный вид	Совершенный вид
	ахнув; ахнувши

Accessing the html

Using your web browser, you can look at the code that is used to generate the html page.

We can use our knowledge of html to determine how to extract the data we are looking at.

The image shows a side-by-side comparison of a web page and its source code. On the left is the rendered page from 'Морфология.ру' for the word 'gasp'. On the right is the raw HTML code that generated the page.

Rendered Page Content:

Enter a word:

Word forms of the word "gasp"

gasp → **gasp** · infinitive, action. Hall.

Part of speech: infinitive - **gasp**

Part of speech: verb

	Present time	Past tense	Future tense	Imperative mood
I		gaped ; gasped ; gasped	ahnu	
You		gaped ; gasped ; gasped	gasp	ahni
He		gaped ; gasped ; gasped	gasp	
she		gaped ; gasped ; gasped	gasp	
we		gaped	ahem	ahnem ; ahnemte
You		gaped	gasp	gasp
They		gaped	gasp	

Part of speech: verbal participle

Imperfect species	Perfect view
	gaspig ; gasping

Part of speech: participle

Active participle:

	Present time			
	Masculine gender	Feminine	Neuter gender	Plural
Them.				
Rd.				
Dt.				
Ext.				
TV.				
Etc.				

	Past tense		
	Singular	Plural	

HTML Code Snippets:

```
<input type="text" value="аХНУТЬ" />
<input type="button" value="Search for" />
Word forms of the word "gasp"
gasp → gasp · infinitive, action. Hall.
Part of speech: infinitive - gasp
Part of speech: verb


|      | Present time | Past tense              | Future tense | Imperative mood |
|------|--------------|-------------------------|--------------|-----------------|
| I    |              | gaped ; gasped ; gasped | ahnu         |                 |
| You  |              | gaped ; gasped ; gasped | gasp         | ahni            |
| He   |              | gaped ; gasped ; gasped | gasp         |                 |
| she  |              | gaped ; gasped ; gasped | gasp         |                 |
| we   |              | gaped                   | ahem         | ahnem ; ahnemte |
| You  |              | gaped                   | gasp         | gasp            |
| They |              | gaped                   | gasp         |                 |


Part of speech: verbal participle


| Imperfect species | Perfect view     |
|-------------------|------------------|
|                   | gaspig ; gasping |


Part of speech: participle
Active participle:


|       | Present time     |          |               |        |
|-------|------------------|----------|---------------|--------|
|       | Masculine gender | Feminine | Neuter gender | Plural |
| Them. |                  |          |               |        |
| Rd.   |                  |          |               |        |
| Dt.   |                  |          |               |        |
| Ext.  |                  |          |               |        |
| TV.   |                  |          |               |        |
| Etc.  |                  |          |               |        |



|  | Past tense |        |  |
|--|------------|--------|--|
|  | Singular   | Plural |  |
|  |            |        |  |


<table border="1" class="grameasTable">
<tr>
<td colspan="2">gaspig ; gasping</td>
</tr>
</table>
<table border="1" class="grameasTable">
<thead>
<tr>
<th rowspan="2"></th>
<th colspan="4">Present time</th>
</tr>
<tr>
<th>Masculine gender</th>
<th>Feminine</th>
<th>Neuter gender</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Them.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rd.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dt.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ext.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TV.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Etc.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
<table border="1" class="grameasTable">
<thead>
<tr>
<th rowspan="2"></th>
<th colspan="3">Past tense</th>
</tr>
<tr>
<th>Singular</th>
<th>Plural</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Accessing the html

Your experience may be different, with different browsers, browser versions, and different themes, but it should be similar.

Here I use Google Chrome version 88.0.43124.190.

To see the html, select the configuration menu, it looks like ellipses running vertically in the upper right-hand corner, then select “More tools” menu option, and then “Developer tools”.

Also, I always use a dark theme that has a black background to work better in low light. It will also help distinguish between results and html.

The screenshot shows the Morphology.ru website interface. At the top, there is a search bar with the word "ахнуть" entered and a "Search for" button. Below the search bar, the word forms of "gasp" are displayed. The part of speech is identified as "verb". A table shows the present, past, and future tenses for various subjects. Below this, the verbal participle is shown in two views: Imperfect species and Perfect view. The active participle is also shown in a table with columns for Present time and Past tense, and rows for Masculine gender, Feminine, Neuter gender, and Plural.

	Present time	Past tense	Future tense	Imperative mood
I		gasped ; gasped ; gasped	ahnu	
You		gasped ; gasped ; gasped	gasp	ahni
He		gasped ; gasped ; gasped	gasp	
she		gasped ; gasped ; gasped	gasp	
we		gasped	ahem	ahnem ; ahnemte
You		gasped	gasp	gasp
They		gasped	gasp	

Part of speech: verbal participle

Imperfect species	Perfect view
	gaspig ; gasping

Part of speech: participle

Active participle:

	Present time			
	Masculine gender	Feminine	Neuter gender	Plural
Them.				
Rd.				
Dt.				
Ext.				
TV.				
Etc.				

Part of speech: participle

	Past tense			
	Masculine gender	Feminine	Neuter gender	Plural
Them.				
Rd.				
Dt.				
Ext.				
TV.				
Etc.				

The right side of the image shows the developer tools interface with a dark theme. The HTML code is visible, showing the structure of the page, including the search bar, the word forms section, and the tables. The code is rendered in a dark background with light text, making it easy to read in low light.

Web Scraping

Using the “Developers tools” you can match the html with the sections of the results page.

Here we find that the results we are interested in are in a bunch of tables with a class name of gramemasTable.

To get our results we need to extract the html tables with the class name of gramemasTable

The screenshot shows the website 'Морфология.ру' with a search bar containing 'ахнуть'. The search results are displayed on the left, and the corresponding HTML code is shown on the right. A blue arrow points from the search bar to the HTML code, and several orange arrows point from specific HTML elements to their corresponding content on the page.

Word forms of the word "gasp"

gasp → **gasp** - infinitive, action. Hall.

Part of speech: infinitive - **gasp**

Part of speech: verb

	Present time	Past tense	Future tense	Imperative mood
I		gasped ; gasped ; gasped	ahnu	
You		gasped ; gasped ; gasped	gasp	ahni
He		gasped ; gasped ; gasped	gasp	
she		gasped ; gasped ; gasped	gasp	
we		gasped	ahem	ahnem ; ahnemte
You		gasped	gasp	gasp
They		gasped	gasp	

Part of speech: verbal participle

Imperfect species	Perfect view
	gasping ; gasping

Part of speech: participle

Active participle:

	Present time			
	Singular			Plural
	Masculine gender	Feminine	Neuter gender	
Them.				
Rd.				
Dt.				
Ext.				
TV.				
Etc.				

Past tense

	Singular			Plural
	Masculine gender	Feminine	Neuter gender	
Them.				
Rd.				
Dt.				
Ext.				
TV.				
Etc.				

```
<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" class="translated-ltr">
<head>
</head>
<body marginheight="0" marginwidth="0" leftmargin="0" rightmargin="0" topmargin="0" bottommargin="0">
<div id="fb-root" class="fb_reset">
<script>
</script>
<div id="wrap">
<div style="margin-left:10px;">
<div id="top">
<div id="middle">
<div id="content">
<div id="printIt">
<table class="gramemas">
<div class="hr">
<b>
<font style="vertical-align: inherit;">
<a href="списокформ/ахнуть" class="w yellow" title="action Hall.">
<div class="hr">
<b>
<font style="vertical-align: inherit;">
<table class="gramemasTable">
<div class="hr">
<b>
<font style="vertical-align: inherit;">
<table class="gramemasTable">
<div class="hr">
<b>
<font style="vertical-align: inherit;">
<br>
<br>
<table class="gramemasTable">
<table class="gramemasTable">
<br>
<br>
<table class="gramemasTable">
<table class="gramemasTable">
<div class="hr">
<b>
<font style="vertical-align: inherit;">
<b>
<br>
<table class="gramemasTable">
<table class="gramemasTable">
</div>
</div>
</div>
<div id="bottom">
</div>
<script type="text/javascript">
<div id="goog-gt-tt" class="skiptranslate" dir="ltr">
<div class="ac_results" style="display: none; position: absolute;">
<div class="goog-te-spinner-pos">
</body>
</html>
```

Tables

This is a closer look. You can see all the gramemasTable classes.

```
...<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN" "http://www.w3.org/TR/html4/loose.dtd"> == $0
<html xmlns="http://www.w3.org/1999/xhtml" class="translated-ltr">
  <head>...</head>
  <body marginheight="0" marginwidth="0" leftmargin="0" rightmargin="0" topmargin="0" bottommargin="0">
    <div id="fb-root" class=" fb_reset">...</div>
    <script>...</script>
    <div id="wrap">
      <div style="margin-left:10px;">...</div>
      <div id="top">...</div>
      <div id="middle">
        <div id="content">
          <div id="printMe">
            <h1>...</h1>
            <table class="gramemas">...</table>
            <div class="hr"></div>
            <b>...</b>
            <font style="vertical-align: inherit;">...</font>
            <a href="сноваформа/аххуа" class="w yellow" title="action Hall.">...</a>
            <div class="hr"></div>
            <b>...</b>
            <font style="vertical-align: inherit;">...</font>
            <table class="gramemasTable">...</table>
            <div class="hr"></div>
            <b>...</b>
            <font style="vertical-align: inherit;">...</font>
            <table class="gramemasTable">...</table>
            <div class="hr"></div>
            <b>...</b>
            <font style="vertical-align: inherit;">...</font>
            <br>
            <br>
            <b>...</b>
            <br>
            <table class="gramemasTable">...</table>
            <table class="gramemasTable">...</table>
            <b>...</b>
            <br>
            <table class="gramemasTable">...</table>
            <table class="gramemasTable">...</table>
            <div class="hr"></div>
            <b>...</b>
            <font style="vertical-align: inherit;">...</font>
            <b>...</b>
            <br>
            <table class="gramemasTable">...</table>
            <table class="gramemasTable">...</table>
          </div>
        </div>
      </div>
    <div id="bottom">...</div>
  </div>
  <script type="text/javascript">...</script>
  <div id="goog-gt-tt" class="skiptranslate" dir="ltr">...</div>
  <div class="ac_results" style="display: none; position: absolute;"></div>
  <div class="goog-te-spinner-pos">...</div>
</body>
</html>
```

Морфология.ру

Введите слово:

Искать

Словоформы слова "ахнуть"

ахнуть → **ахнуть** – инфинитив, действ. зал.

Часть речи: инфинитив - **ахнуть**

Часть речи: глагол

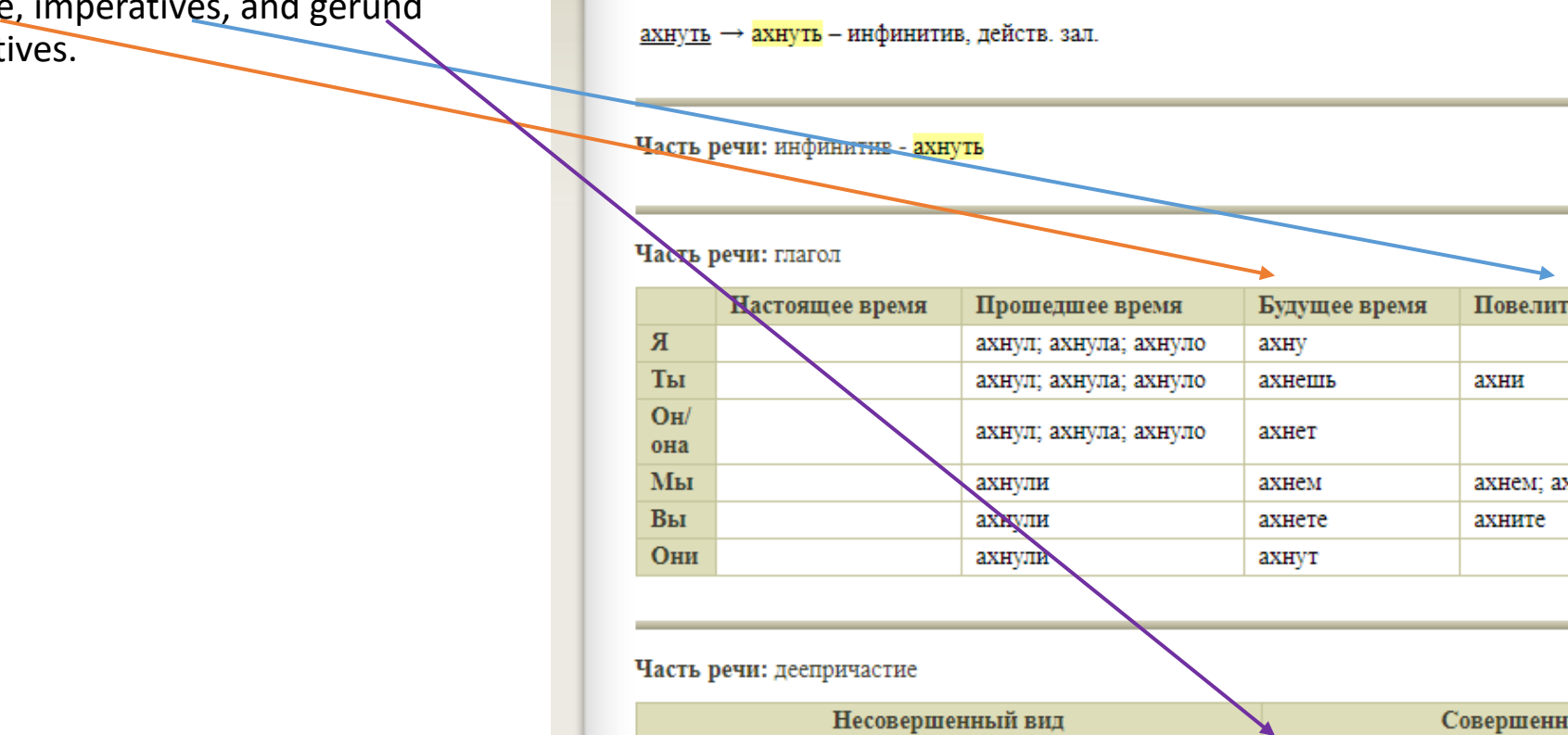
	Настоящее время	Прошедшее время	Будущее время	Повелительное наклонение
Я		ахнул; ахнула; ахнуло	ахну	
Ты		ахнул; ахнула; ахнуло	ахнешь	ахни
Он/ она		ахнул; ахнула; ахнуло	ахнет	
Мы		ахнули	ахнем	ахнем; ахнемте
Вы		ахнули	ахнете	ахните
Они		ахнули	ахнут	

Часть речи: деепричастие

Несовершенный вид	Совершенный вид
	ахнув; ахнувши

Results we want

We are not looking for all the results. Just the future tense, imperatives, and gerund forms for infinitives.



What we extract

We find that the future tense and imperatives are in the first gramemasTable table, and gerunds are in the second gramemasTable table.

```
▼<table class="gramemasTable">
  ▼<tbody>
    ▼<tr>
      <th></th>
      <th>Настоящее время</th>
      <th>Прошедшее время</th>
      <th>Будущее время</th>
      <th>Повелительное наклонение</th>
    </tr>
    ▼<tr>
      <th width="10">Я</th>
      <td></td>
      ▼<td>
        <a href="словоформа/ахнул" class="w" title="действ. зал., прош. вр., муж. р., ед. ч.">ахнул</a>
        "; "
        <a href="словоформа/ахнула" class="w" title="действ. зал., прош. вр., жен. р., ед. ч.">ахнула</a>
        "; "
        <a href="словоформа/ахнуло" class="w" title="действ. зал., прош. вр., ср. р., ед. ч.">ахнуло</a>
      </td>
      ▼<td>
        <a href="словоформа/ахну" class="w" title="действ. зал., 1 лицо, буд. вр., ед. ч.">ахну</a>
      </td>
      <td></td>
    </tr>
    ▼<tr>
      <th width="10">Ты</th>
      <td></td>
      ▼<td>
        <a href="словоформа/ахнул" class="w" title="действ. зал., прош. вр., муж. р., ед. ч.">ахнул</a>
        "; "
        <a href="словоформа/ахнула" class="w" title="действ. зал., прош. вр., жен. р., ед. ч.">ахнула</a>
        "; "
        <a href="словоформа/ахнуло" class="w" title="действ. зал., прош. вр., ср. р., ед. ч.">ахнуло</a>
      </td>
      ▼<td>
        <a href="словоформа/ахнешь" class="w" title="действ. зал., 2 лицо, буд. вр., ед. ч.">ахнешь</a>
      </td>
      ▼<td>
        <a href="словоформа/ахни" class="w" title="действ. зал., повелительное накл., 2 лицо, ед. ч.">ахни</a>
      </td>
    </tr>
  </tbody>
</table>
```


How we extract it

We use Web Scraping tools, such as the Python module Beautiful Soup, to extract the contents of the first and second gramemasTable, and save the contents into Excel workbooks.

Using a programming language, like Python, allows us to automate this process, and adapt to anomalies in the data.

References

- Koob, Perry B., and Ivliyeva, Irina V. "Russian National Corpus Web Scraping Project 2019-2020." [Electronic resource]. February 1, 2020. – Access mode: https://scholarsmine.mst.edu/artlan_phil_facwork/158/
- Ивлиева, И.В. Особенности семантики лакун-модификаций. К идее создания словаря лакун (на материале глаголов звучания) // Интерактивная наука. – 2020. – № 6 (52). – С. 57-70. – ISSN 2414-9411. Access mode: https://interactive-science.media/ru/article/551671/discussion_platform
- The grammatical dictionary of the Russian language A.A. Zalyzniak. Access mode: <https://morfologija.ru/>.