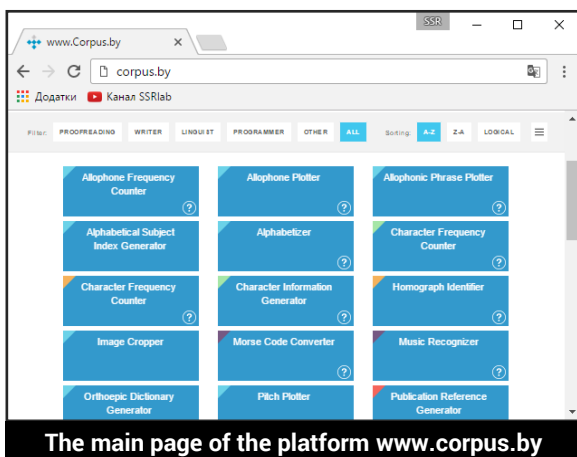


Corpus.by is a computational text&speech platform that includes about 50 services for electronic text and sound data processing.



- free access
- free use
- different languages
- 5 different solutions
- open source

Visit corpus.by!

All input data in this resource will be used only in scientific purposes for algorithms improvement.



contacts

Speech Synthesis and Recognition Laboratory

Surhanava St., 6

Minsk 220012

BELARUS



8 (017) 284-27-73
8 (017) 331-84-03 (fax)



www.ssrlab.by
ww.corpus.by



ssrlab221@gmail.com



We will be happy to receive your offers
to corpus.by@gmail.com!

© SSRLab, UIIP NAS Belarus, 2012-2018



CORPUS.BY

COMPUTATIONAL TEXT&SPEECH PLATFORM

SPEECH SYNTHESIS AND RECOGNITION LABORATORY

THE UNITED INSTITUTE OF INFORMATICS PROBLEMS

NATIONAL ACADEMY OF SCIENCES OF BELARUS

Text-to-Speech Synthesizer

Please input a stressed text

Primary stressed vowel must be marked by ^ˈ or ^ˆ, a secondary stressed vowel – by ^ˑ or ^ː.
To mark two words as one phonetic word use ^ˈ or ^ˆ.

or example: Паўночна-заходні вятрыск садыму-^ˈбы ўсе лісьце на^ˑчэсны, але^ˆп'ятым калісьці.
or: Паўночна-заходні вятрыск садыму-^ˈбы ўсе лісьце на^ˑчэсны, але^ˆп'ятым калісьці.

Лісьц.
вядуць.
два.
тры.
чатыры.
пяць.
шасць.

Belarusian BorisBel Show log information

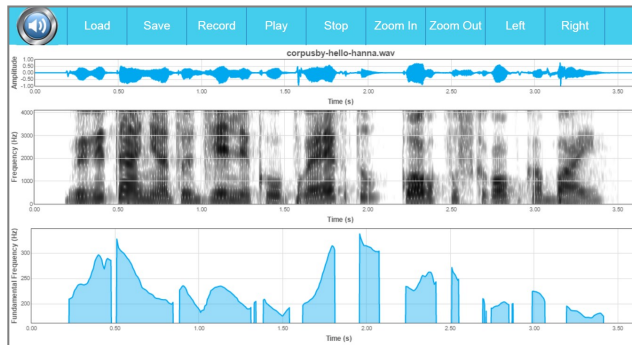
Listen to the generated speech:

0:00 / 4.45

Treated during 2 sec. language be.

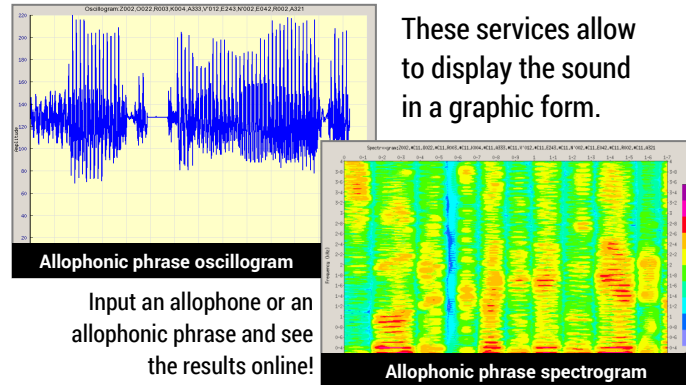
Processes a text automatically and forms an audio file that a user can listen to, download and save.

Pitch Plotter

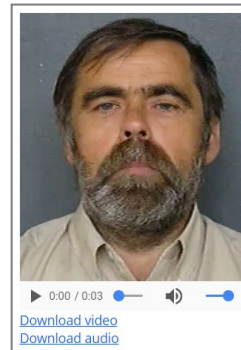


Allows to get a graphical representation of voice phrase online.

Allophone Plotter and Allophone Phrase Plotter



Talking Head Synthesizer



Processes text automatically and forms a video file with animated head articulating the input phrase. A user can listen it to, download and save.

The main idea of this service is to help people with hearing disabilities.

Voiced Electronic Grammatical Dictionary

Word for searching:

Part-of-speech: Only initial forms

Tag (for "sbm1987"): For example, NN1*

Maximum results:

Dictionaries for searching:

- SBM1987 (навадзе публікацыяў "Словнік Беларускай мовы. Арфаграфіка, Арфапія, Акцэнтацыя, Словазлучэньне" / пад рэд. М.В. Брычкі. - Мінск, 1987)
- SBM2012initial (лексікаграфічны фармат публікацыі "Словнік Беларускай мовы" / навука, рэд. А.А. Лукашэвіч, В.П. Русак. - Мінск: Беларуская навука, 2012)
- NOUN2013 (навадзе публікацыяў "Граматычны слоўнік навуковай / навука, рэд. В.П. Русак. - Мінск: Беларуская навука, 2013")
- ASBM2017 (навадзе публікацыяў "Граматычны слоўнік Беларускай мовы" / навука, рэд. В.П. Русак, Ю.С. Геверыч, С.І. Лысы. - Мінск: Беларуская навука, 2017)
- ZALIZNIAK (навадзе публікацыяў "Граматычна-лексікаграфічны слоўнік Беларускай мовы: Словазлучэньне / А.А. Залізняк. - Москва: Русская мова, 1980. - 880 с.)
- CMU (навадзе публікацыяў "Cambridge English Pronunciation Dictionary")
- NEW (новыя словы для CMU)

The service provides information on the correct words spelling and pronunciation. It allows to see its transcription and find out its annotation.

Spell Checker

Please, enter a text for spell checking

It is a truth universally acknowledged, that a single man in possession of a good fortune must be in want of a wife. However little known the feelings or views of such a man may be on his first entering a neighbourhood, this truth is so well fixed in the minds of the surrounding families, that he is considered as the rightful property of some one or other of their daughters.

truth neighbourhood

Checks the correctness of words spelling. After processing a user receives a list of words that were not found in the built-in dictionaries. Also there is a possibility to create and add new dictionaries.

Service Demonstrator

Please, enter a text

Група цэляў апошняй год.
Усе галіны на, усе вайсковыя раскі, да апошняга прудка, былі ўсупакі буйным белым руковым шэцям. Яна кіпела, мілья і расквашвалася ў глыбым звоне, цягнула да сонца сталыя латы і расквашла ў яго званні малавыя, каляя палцы новых парасткаў. І была яна такая магутная і свежая, так утратана спрыяла ў яе руковым раі гнолы, што, здавалася, не будзе ёй зводу і не будзе ёй канца.

Select checkboxes:

- checkbox #1 (you can input a comment here)
- checkbox #2 (you can input a comment here)
- checkbox #3 (you can input a comment here)

Select radiobutton:

- radiobutton #1 (you can input a comment here)
- radiobutton #2 (you can input a comment here)
- radiobutton #3 (you can input a comment here)

Select option:

This is an open source which allows every interested developer to create his own service for www.corpus.by.

About 50 services for different solutions:

1. Allophone Frequency Counter
2. Allophone Plotter
3. Allophonic Phrase Plotter
4. Alphabetical Subject Index Generator
5. Alphabetizer
6. Character Frequency Counter
7. Character Information Generator
8. Emotion Speech Analyzer
9. Homograph Identifier
10. Image Cropper
11. Keyboard Layout Converter
12. Language Identifier
13. Language Tutor
14. Lemmatizer
15. Morse Code Converter
16. Music Recognizer
17. N-gram Frequency Counter
18. Orthoepic Dictionary Generator
19. Part-of-Speech Tagger
20. Pitch Plotter
21. Publication Reference Generator
22. RSS Reader
23. Romanizator
24. Service Demonstrator
25. Service DemonstratorPy
26. Service Demonstrator With Authorization

Proofreading · Writer · Linguist · Programmer · Other

27. Short U Spell Checker
28. Sound Recorder
29. Speaker Recognizer
30. Specialized Phonetic Dictionary
31. Speech Duration Predictor
32. Speech Recognition Trainer
33. Speech Segmentation Checker
34. Spell Checker
35. Syllabifier
36. Table Processor
37. Tag Identifier
38. Talking Head Synthesizer
39. Text-to-Speech Synthesizer
40. Thematic Lists Collector
41. Thematic Speech Recognizer
42. Tokenizer
43. Transcription Generator
44. UDC Decoder
45. Unknown Words Processor
46. Voiced Electronic Grammatical Dictionary
47. Web Page Content Saver
48. Word Frequency Counter
49. Word Paradigm Generator
50. ... Add a solution for your task!