Воронежский государственный университет

Проектная работа

на тему:

Понятие информационных технологий в лингвистике. Автоматизированное рабочее место лингвиста. Понятие естественного и искусственного языка. Виды искусственных языков.

Подготовила студентка ВГУ

1 курс

Направление «Лингвистика»

Профиль «ТиМПИЯиК»

Онучина Анна Вениаминовна

Воронеж- 2018

Мотивация

Современное информационное пространство обязывает человека, изначально обладающего определенными знаниями, опытом и компетенциями, уметь обучаться: находить нужную информацию, уметь пользоваться различными источниками. Постоянное приобретение дополнительных знаний помогает решать профессиональные и личные задачи.

Ресурсы, в которых можно найти информацию о необходимости и эффективности применения информационных технологий в процессе лингвистической деятельности:

<https://qwizz.ru/информационные-технологии-лингвис/>

Теория.

1.Понятие информационных технологий в лингвистике.

В настоящее время лингвистическая наука находится в тесной связи с информационными технологиями. С помощью технологий можно успешно решать всевозможные задачи лингвистики. В то же время сама эта наука оказывает значительное воздействие на формирование современных инноваций.

Информационные технологии в лингвистике являются совокупностью законов, средств и методов, позволяющих получать, хранить, передавать, распространять и преобразовывать информацию о языке и принципах его функционирования при помощи современных компьютерных программ.

<https://qwizz.ru/информационные-технологии-лингвис/> Перейдя по этой ссылке, вы сможете более подробно ознакомиться с понятием информационных технологий в лингвистике.



2.Автоматизированное рабочее место лингвиста.

**Общие сведения:**

Автоматизированное рабочее место (АРМ) — это рабочее место специалиста, оснащенное персональным компьютером, программным обеспечением и совокупностью информационных ре­сурсов индивидуального или коллективного пользования, которые позволяют ему вести обработку данных с целью получения информации, обеспечивающей поддержку принимаемых им решений при выполнении профессиональных функций.  
  
Что касается лингвистики, рабочее место специалиста данной сферы должно быть оснащено совокупностью аппаратных, программных и лингвистических средств, необходимых для автоматическое обработки лингвистических данных.

**Программах и ресурсы, необходимые лингвисту:**

**Электронные словари:**

ABBYY Lingvo

Babylon-Pro

АльфаЛекс

Pop-up Dictionary

SKIIN

Albion Explanatory Dictionary

Rdic

EDictionary

VU-Dictionary

Parley Box

Context

MultiLex

**Контекстные словари:**

Artefact Dictionary

TranslateIt!

TransLite

MyDic

QDictionary

 LearnWords

ABBY LINGVO

**Контекстный словарь** представляет собой словарь каждая запись которого содержит контекст употребления одного определенного значения слова. Контекстный словарь создается с целью дальнейшего, более эффективного изучения слов в смысловом окружении, контексте.

<http://www.contdict.ru> – это словарь, работающий в онлайн-режиме, который позволяет пользователю работать более чем с десятком языков, а так же имеет собственный каталог обучающих видео- <http://www.contdict.ru/videos>

**Лингвистические банки данных:**

Подлингвистическиим банками данныхпонимаются представленные в электронной форме языковые источники (корпусы текстов) и лингвистические описания.

<http://www.ruscorpora.ru> - На этом сайте помещен корпус современного русского языка общим объемом более 600 млн слов. Корпус русского языка — это информационно-справочная система, основанная на собрании русских текстов в электронной форме. Корпус предназначен для всех, кто интересуется самыми разными вопросами, связанными с русским языком: профессиональных лингвистов, преподавателей языка, школьников и студентов, иностранцев, изучающих русский язык.

<https://corpus.byu.edu/bnc> - это [корпус текстов](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D1%80%D0%BF%D1%83%D1%81_%D1%82%D0%B5%D0%BA%D1%81%D1%82%D0%BE%D0%B2) из 100 миллионов слов, содержащий образцы письменного и разговорного [британского английского](https://ru.wikipedia.org/wiki/%D0%91%D1%80%D0%B8%D1%82%D0%B0%D0%BD%D1%81%D0%BA%D0%B8%D0%B9_%D0%B2%D0%B0%D1%80%D0%B8%D0%B0%D0%BD%D1%82_%D0%B0%D0%BD%D0%B3%D0%BB%D0%B8%D0%B9%D1%81%D0%BA%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0) языка из широкого круга источников. Корпус охватывает британский английский конца XX в., представленный широким разнообразием [жанров](https://ru.wikipedia.org/wiki/%D0%96%D0%B0%D0%BD%D1%80), и задуман как образец типичного разговорного и письменного британского английского языка того времени.

Lancaster/Oslo-Bergen Corpus - имеет в своем составе два подкорпуса - Leeds-Lancaster Treebank и Lancaster Parsed Corpus, - снабженных синтаксической разметкой.

**Также существует отдельный сайт, который содержит каталог лингвистических программ и ресурсов в сети -** [**https://rvb.ru/soft/catalogue/catalogue.html#Resources**](https://rvb.ru/soft/catalogue/catalogue.html#Resources)

3.Понятие естественного и искусственного языка. Виды искусственных языков.

«Естественный» и «искусственный» - это деление языков по происхождению.

**Естественный язык**- в лингвистике и философии языка язык, используемый для общения людей и не созданный искусственно (в отличие от искусственных языков)

**Искусственные языки** - специальные языки, которые, в отличие от естественных, сконструированы целенаправленно. Они могут конструироваться с помощью естественного языка или ранее построенного искусственного языка.

**Различают следующие виды искусственных языков:**

· Языки программирования и компьютерные языки - языки для автоматической обработки информации с помощью ЭВМ.

· Информационные языки - языки, используемые в различных системах обработки информации.

· Формализованные языки науки - языки, предназначенные для символической записи научных фактов и теорий математики, логики, химии и других наук.

· Языки несуществующих народов, созданные в беллетристических или развлекательных целях, например: эльфийский язык, придуманный, Дж. Толкином клингонский язык, придуманный Марком Окрандом для фантастического сериала *«Star Trek»*, язык На'ви, созданный для фильма «Аватар».

· Международные вспомогательные языки - языки, создаваемые из элементов естественных языков и предлагаемые в качестве вспомогательного средства межнационального общения.

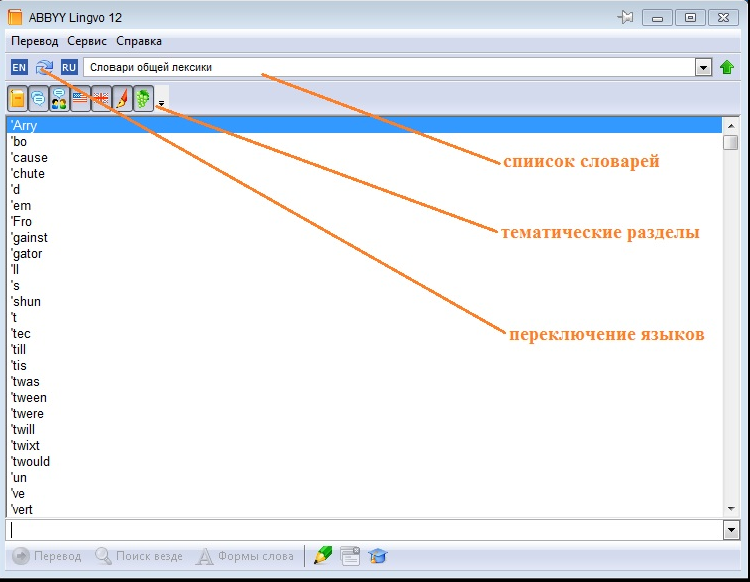
Практика.

Ознакомление с ABBYY Lingvo 12 Multilingual Edition

Это электронный словарь, который содержит современную лексику из разных тематических областей, как универсальную, так и специальную.



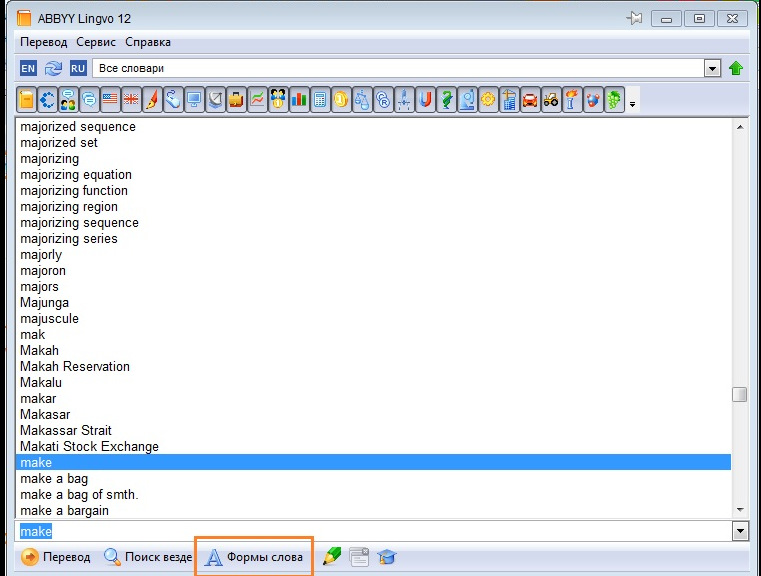
Качаем этот словарь по ссылке – <https://soft.sibnet.ru/soft/12491-abbyy-lingvo-12-ml/download> и устанавливаем.

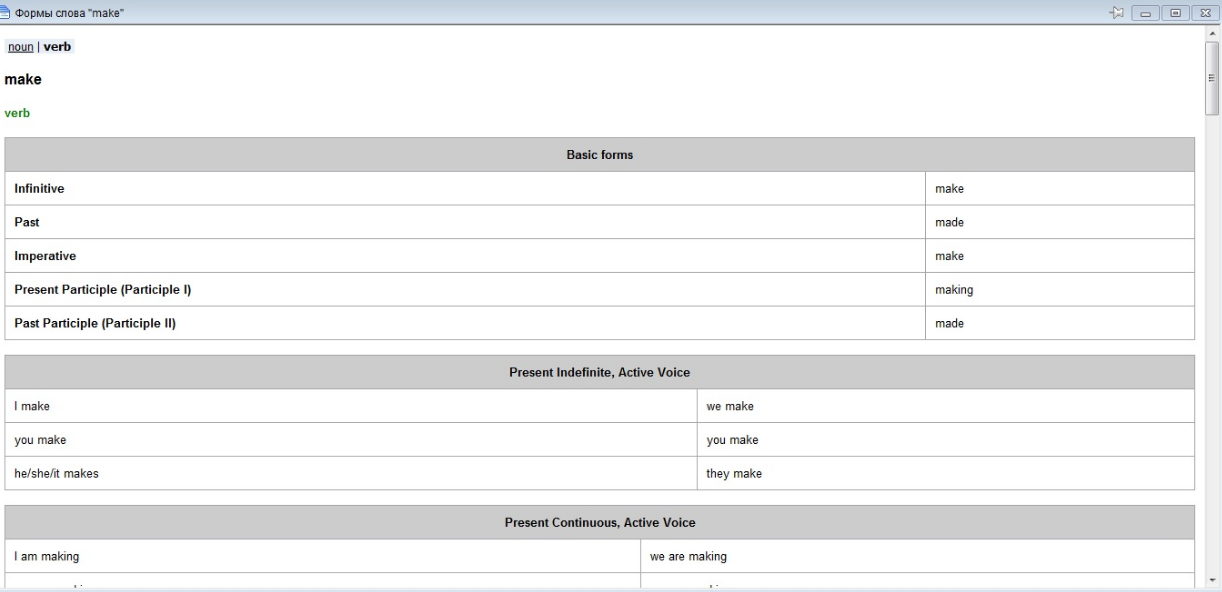


В списке словарей по умолчанию доступно 5 групп:

* словарь общей лексики;
* словари по естествознанию;
* словари по информатике;
* словари по экономике и праву;
* технические словари.

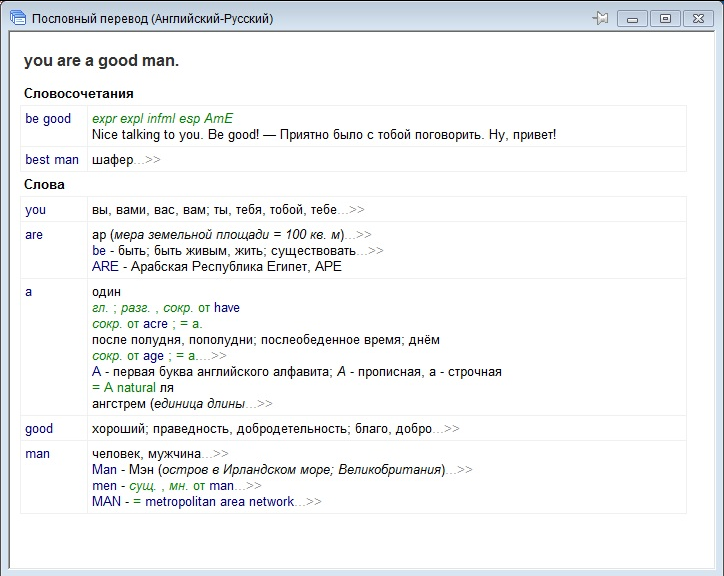
Каждая группа состоит из определенных тематических разделов. Например, выбрав словарь общей лексики, видим, что он состоит из следующих тематических ниш: универсальный (LingvoUniversal), идиомы (Idioms), неформальное общение (Informal), американский английский (American), британский английский (GreatBritain), виноделие (Wine).

Если кнопка тематического раздела нажата то, при переводе, система будет осуществлять поиск слова в данной теме. Для удобства можно в списке словарей сразу выбрать «Все словари» и работать с ними всегда. 

Перед нами откроется окно, в котором можно посмотреть всевозможные формы например слова «make», для каждой части речи, к которой оно может принадлежать.

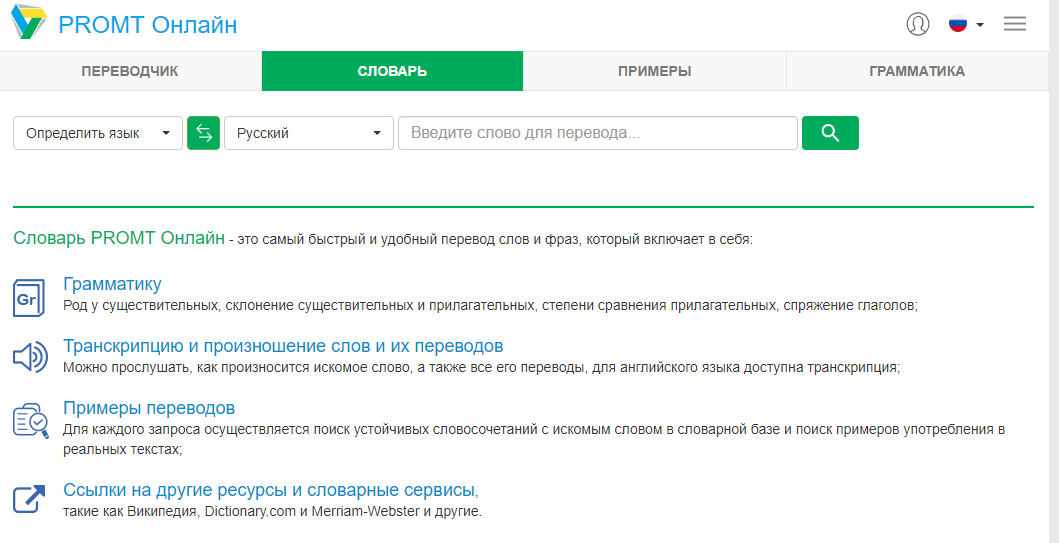
Следующей крайне полезной функцией программы является перевод слов при помощи их выделения и нажатия комбинации клавиш Ctrl+C+C. Допустим, вы читаете одну из моих статей на английском языке. Попробуйте выделить незнакомое слово и, зажав Ctrl, два раза нажать клавишу С. Вот, что из этого получится:

Однако данный словарь не предназначен для перевода случайных фраз и предложений. Программа выдаст перевод фразы лишь в том случае, если данное выражение является устойчивым, идиомой или фразовым глаголом. Если же вам требуется перевести какое-то случайное предложение, то вам будет предложен перевод каждого слова в отдельности. Например, давайте переведен фразу «You are a good man». Вводим выражение в строку перевода и нажимаем Enter.

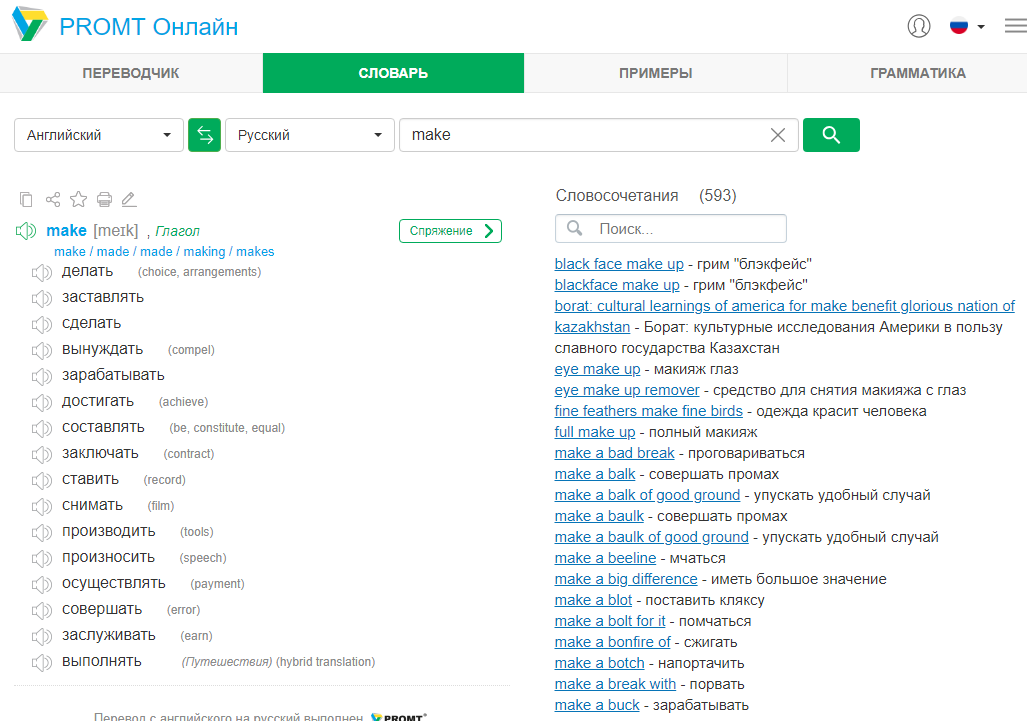
****

**Ознакомление с Translate.ru**

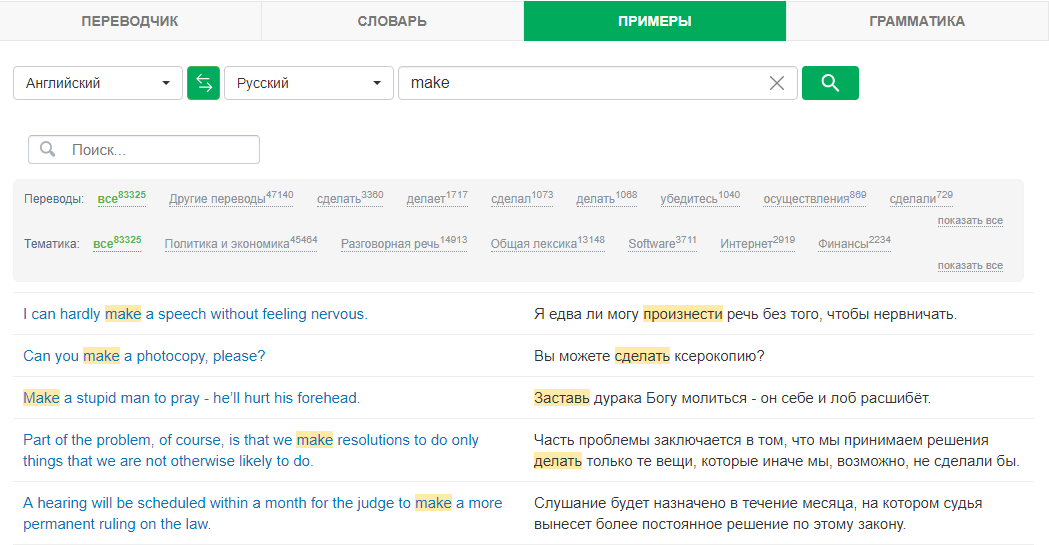
По этой ссылке мы переходим на сайт словаря <https://www.translate.ru>

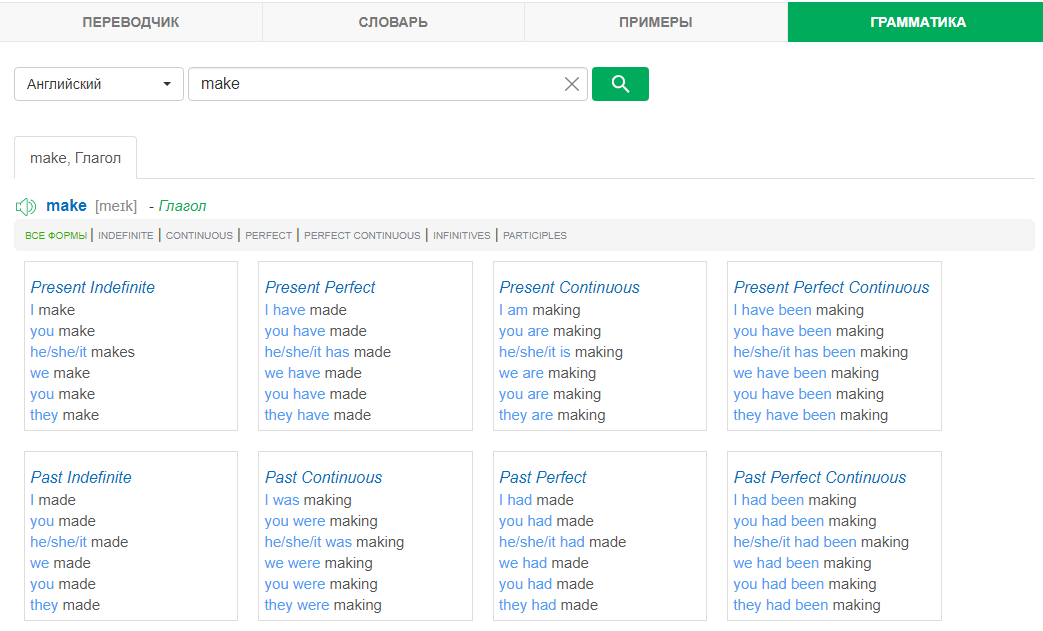


На данном изображении перечислены основные достоинства и возможности данного словаря.

В поиковую строку мы вводим слово, которое требует перевода, предварительно выбрав исходный язык( однако программа способна сама определить язык по заданному слову)

Словарь показывает нам всевозможные переводы данного слова, а так же словосочетания с этим словом.

Затем в меню сайта мы выбираем графу «Примеры» и программы выдает нам всевозможные примеры использования данного слово в предложениях и текстах с различными вариантами контекстного перевода, а так же предлагает выбрать тематику употребления данного слова.

Также сайт дает возможность ознакомиться с грамматикой данного слова. Для этого нам нужно прейти в раздел «Грамматика» в меню сайта.

**Ознакомление с Корпусами**

<https://corpus.byu.edu/bnc> - это [корпус текстов](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D1%80%D0%BF%D1%83%D1%81_%D1%82%D0%B5%D0%BA%D1%81%D1%82%D0%BE%D0%B2) из 100 миллионов слов, содержащий образцы письменного и разговорного [британского английского](https://ru.wikipedia.org/wiki/%D0%91%D1%80%D0%B8%D1%82%D0%B0%D0%BD%D1%81%D0%BA%D0%B8%D0%B9_%D0%B2%D0%B0%D1%80%D0%B8%D0%B0%D0%BD%D1%82_%D0%B0%D0%BD%D0%B3%D0%BB%D0%B8%D0%B9%D1%81%D0%BA%D0%BE%D0%B3%D0%BE_%D1%8F%D0%B7%D1%8B%D0%BA%D0%B0) языка из широкого круга источников. Корпус охватывает британский английский конца XX в., представленный широким разнообразием [жанров](https://ru.wikipedia.org/wiki/%D0%96%D0%B0%D0%BD%D1%80), и задуман как образец типичного разговорного и письменного британского английского языка того времени.

На главной странице мы видим строку поиска, в которую необходимо вставить интересующее слово. В результате откроется страница с примерами предложений, в которых оно встречается.

Например, проверим, а используется ли глагол «recommend» с инфинитивом, т.е., а можно ли сказать «recommend to do».

Вставляем в строку поиска «[recommend to](http://bnc.bl.uk/saraWeb.php?qy=recommend+to&mysubmit=Go" \t "_blank" \o "Recommend to )»

Корпус показывает нам, что нельзя.

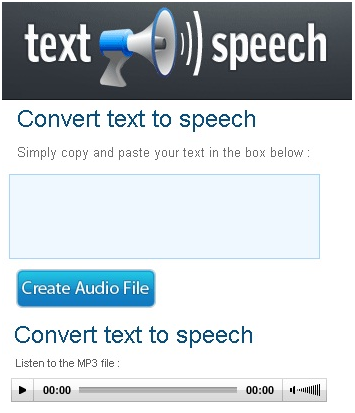
Вот [еще одна база данных](http://corpus.leeds.ac.uk/protected/query.html) <http://corpus.leeds.ac.uk/protected/query.html> , которая состоит из разных корпусов, в том числе и BNC.

Здесь поиск того же «[recommend to](http://smlc08.leeds.ac.uk/cgi-bin/cqp.pl?q=recommend+to&c=REUTERS&searchtype=conc&contextsize=60c&sort1=word&sort2=right&terminate=100&llstat=on&cleft=0&cright=1&cfilter=" \t "_blank" \o "Recommend to do)» дал другие результаты, и выяснилось, что хоть и редко, но «recommend to do» все-таки встречается.

**Ознакомление с синтезаторами речи**

Программы-синтезаторы речи с каждым годом всё больше входят в нашу жизнь. Они позволяют нам более досконально учить иностранные языки, переводят тексты в удобный аудиоформат, используются в функционале различных служебных программ и многое другое. И когда у некоторых из нас возникает потребность воспроизвести онлайн какой-либо текст в аудиформате, тогда многие из нас обращаются к различным сервисам и программам по синтезу речи, способным помочь нам в трансформации нужного там текста. Данные программы могут во многом облегчить труд лингвиста.

Чтобы перевести текст в речь онлайн можно воспользоваться сервисом fromtexttospeech <http://www.fromtexttospeech.com> . Он работает по принципу конвертации текста в аудиофайл формата mp3, который затем можно скачать себе на компьютер.  Сервис поддерживает конвертацию текста величиной в 50 тыс. символов, что является достаточно значительным объёмом.

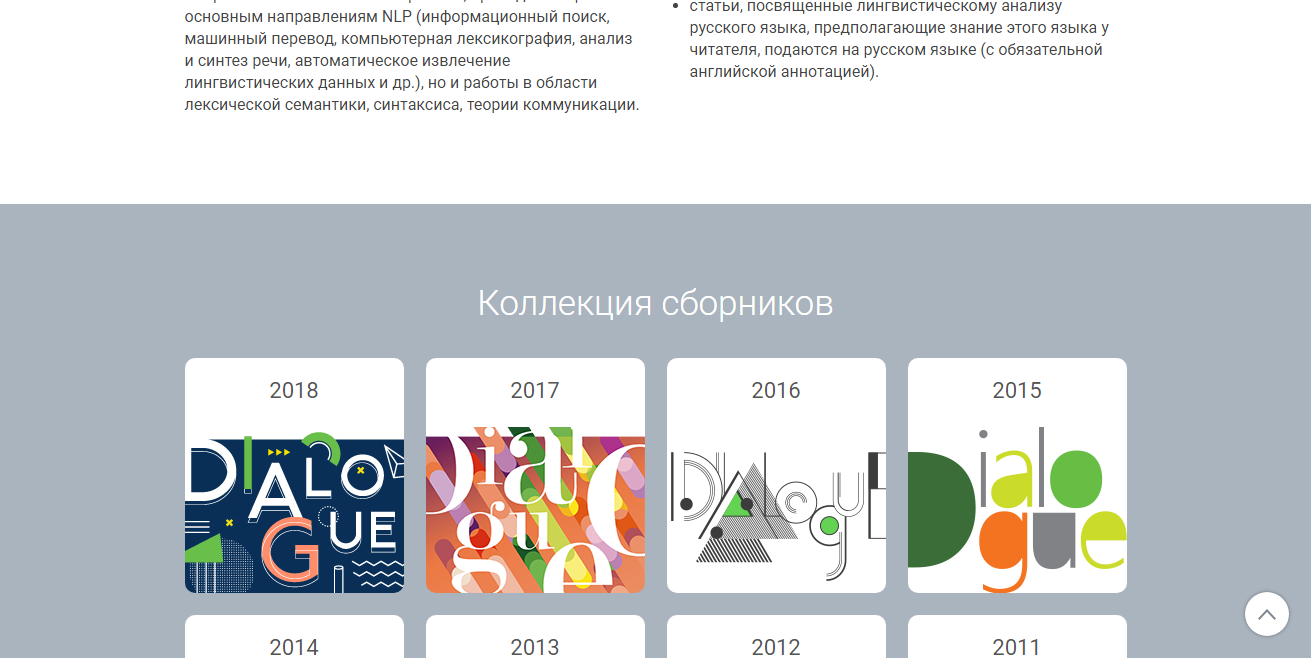
  
Для работы с сервисом fromtexttospeech перейдите на него, в опции «Select Language»  выберите «Russian» .

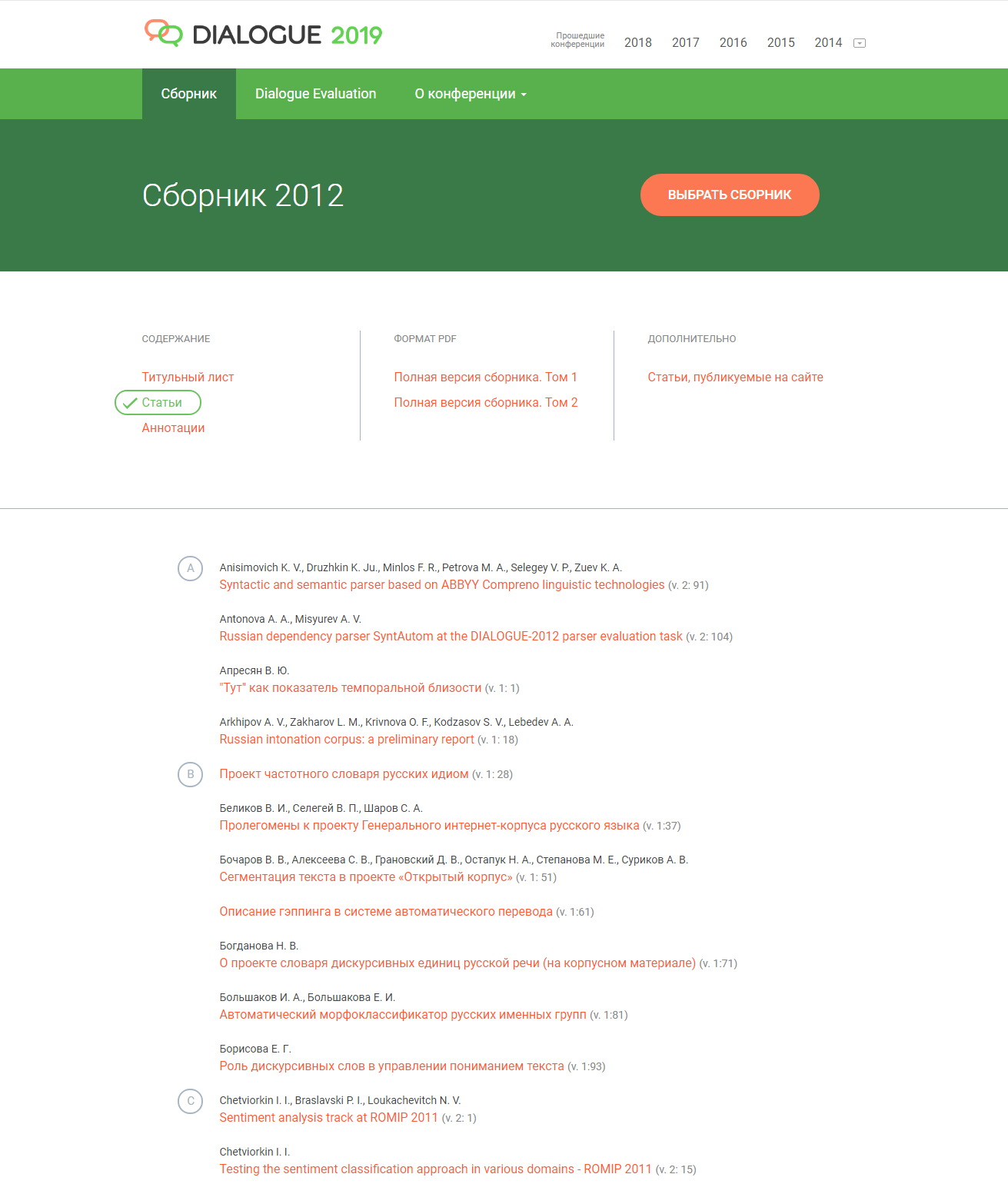
В большом окне введите (вставьте) нужный для озвучки текст, затем нажмите на кнопку «Create Audio File». Текст будет обработан, затем вы сможете послушать полученный результат, а потом и скачать его себе на ПК.

Для этого нажмите правой клавишей мыши на «Download audio file» и выберите в появившемся меню «Сохранить объект как».  
  
**Ознакомление с лингвистическими ресурсами**

**Изучим возможности лингвистических ресурсов на примере сайта** [**http://www.dialog-21.ru**](http://www.dialog-21.ru) **– на данном сайте представлены тексты всех докладов международной конференции «Диалог», посвященной проблемам и вопросам лингвистики и языкознания.**

**Данный сайт содержит коллекцию сборников с докладами, начиная с 2000г.**

****

Выбираем любой, например, сборник за 2012 год.

Выбрав раздел статьи, мы можем ознакомиться с текстом любого доклада, представленного на конференции за выбранный год.

**Задания для закрепления изученного материала:**

**1.Карточки Quizlet**

<https://quizlet.com/354914343/Закрепление-материала-по-изученной-теме-flash-cards>

**2.Подбор Quizlet**

<https://quizlet.com/354914343/match>

**3. Кроссворд**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  | **5** |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  | **3** |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  | **4** |  |  |  |  |
|  |  | **2** |  |  |  |  |  |  |  |  |  |  |  |
| **1** |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |

1.Как называется язык, создаваемый из элементов естественных языков и предлагаемый в качестве вспомогательного средства межнационального общения?

2.Как называется язык, используемый для общения людей и не созданный искусственно?

3. Как называется специальный язык, который сконструирован целенаправленно?

4. Как называется язык, используемый в различных системах обработки информации?

5. Как называется словарь каждая запись которого содержит контекст употребления одного определенного значения слова?

**4. Пройдите тест**

<https://docs.google.com/forms/d/1UDTjB5pzrVIuatSpk4MQYFq7MeSnsyama2BiMAtpPrs/edit?usp=sharing>

**5.Ответьте на вопросы в развернутом виде**

<https://docs.google.com/forms/d/1V7HpCLyH34U1mSE-KtWjDyi_dU-5Nq3Jzermefvc6lY/edit?usp=sharing>

Ресурсы, которыми я пользовалась для создания урока: html <http://biouroki.ru/workshop/crossgen.html>

<https://quizlet.com>

<http://www.google.md/intl/ru/forms/about>

Ответы на кроссворд

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  | и |  |  |  |  |  |  |
|  |  |  |  |  |  |  | н |  |  |  |  |  |  |
|  |  |  |  |  | **3** |  | ф |  |  |  |  |  |  |
|  |  |  |  |  | и |  | о |  |  |  |  |  |  |
|  |  |  |  |  | с |  | р |  | **4** |  |  |  |  |
|  |  | **2** |  |  | к |  | м |  | к |  |  |  |  |
| **1** | м | е | ж | д | у | н | а | р | о | д | н | ы | й |
|  |  | с |  |  | с |  | ц |  | н |  |  |  |  |
|  |  | т |  |  | с |  | и |  | т |  |  |  |  |
|  |  | е |  |  | т |  | о |  | е |  |  |  |  |
|  |  | с |  |  | в |  | н |  | к |  |  |  |  |
|  |  | т |  |  | е |  | н |  | с |  |  |  |  |
|  |  | в |  |  | н |  | ы |  | т |  |  |  |  |
|  |  | е |  |  | н |  | й |  | н |  |  |  |  |
|  |  | н |  |  | ы |  |  |  | ы |  |  |  |  |
|  |  | н |  |  | й |  |  |  | й |  |  |  |  |
|  |  | ы |  |  |  |  |  |  |  |  |  |  |  |
|  |  | й |  |  |  |  |  |  |  |  |  |  |  |