



Что мы слышим и как мы это пишем: **использование корпусов устной речи при** **обучении аудированию и письму**



Елена Игоревна Риехакайнен
СПбГУ



Несколько слов о лекторе😊

кандидат филологических наук,
доцент кафедры общего языкознания им. Л.А. Вербицкой
Санкт-Петербургского государственного университета



Несколько слов о лекторе 😊

Научные интересы: психолингвистика, фонетика, восприятие речи, русский язык как иностранный



<http://genlingnw.ru/person/Riekhakaynen.htm>

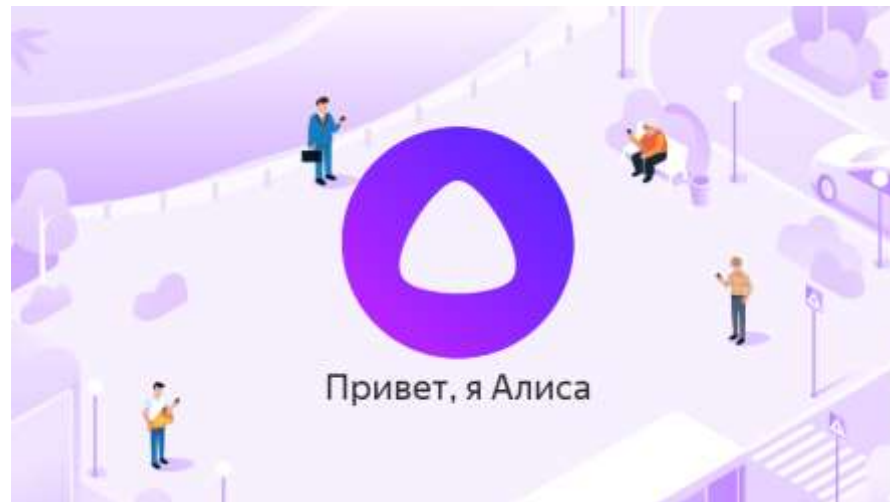
Онлайн-курс СПбГУ «Психолингвистика»

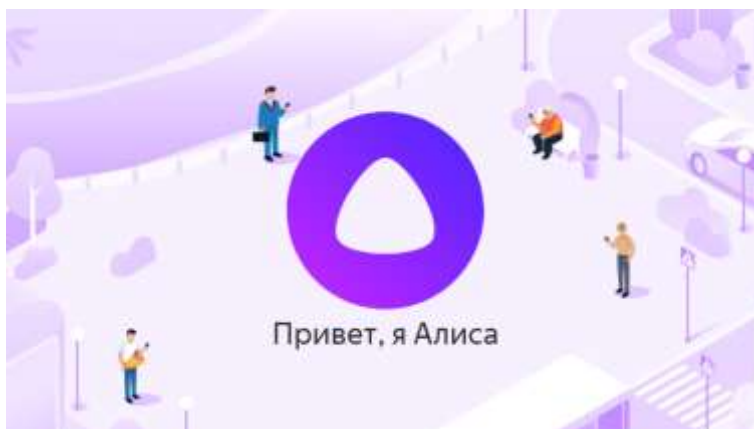


<https://openedu.ru/course/spbu/PSYLING/>

<https://ru.coursera.org/learn/psikholingvistika>







Мне очень нравится Туапсе

Мне очень рад ситуации

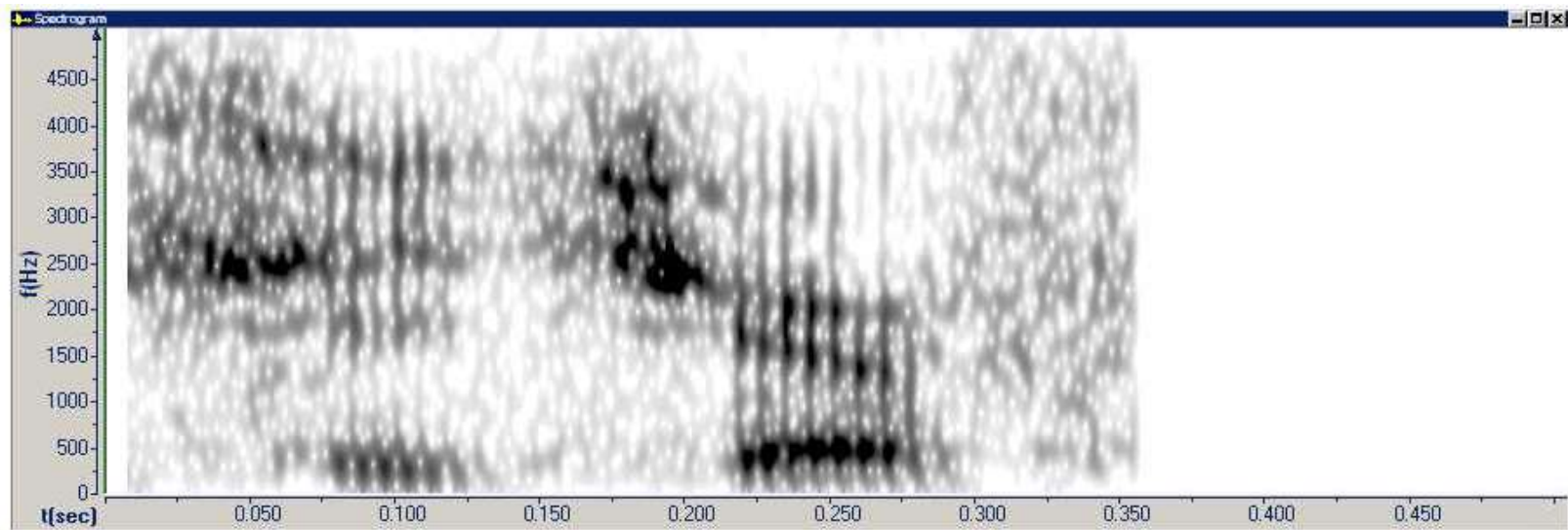
Яркие особенности русской устной речи

1) Редукция словоформ

Яркие особенности русской устной речи

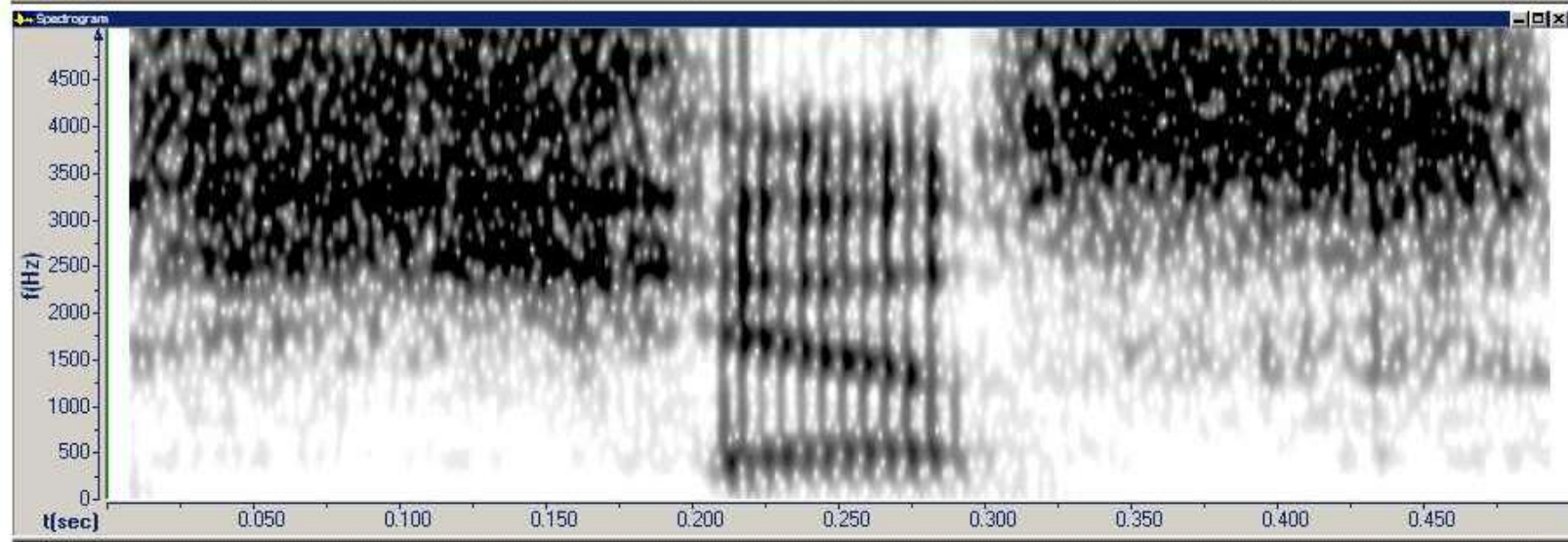
1) Редукция словоформ:

- типичные редуцированные варианты



сейчас

[sʲit͡ɕas]



[ɕas]

Яркие особенности русской устной речи

1) Редукция словоформ:

- типичные редуцированные варианты (*щас*);

- редукция в типичном фонетическом контексте:

когда, тогда, всегда, никогда, иногда

только, сколько, несколько, столько, насколько, настолько

Яркие особенности русской устной речи

1) Редукция словоформ:

- типичные редуцированные варианты (*щас*);

- редукция в типичном фонетическом контексте:

когда, тогда, всегда, никогда, иногда

только, сколько, несколько, столько, насколько, настолько;

- нетипичные редуцированные варианты частотных слов;

Яркие особенности русской устной речи

1) Редукция словоформ:

- типичные редуцированные варианты (*щас*);

- редукция в типичном фонетическом контексте:

когда, тогда, всегда, никогда, иногда

только, сколько, несколько, столько, насколько, настолько

- нетипичные редуцированные варианты частотных слов;

- большое количество вариантов произнесения одного и того же слова.

Яркие особенности русской устной речи

- 1) Редукция словоформ: только в разговорной речи и только в частотных словах? **НЕТ**

Яркие особенности русской устной речи

1) Редукция словоформ: только в разговорной речи и только в частотных словах? **НЕТ**

Факторы, которые влияют на возникновение редукции:

- частотность;
- количество слогов;
- является ли слово прилагательным, вводным словом, предлогом.



(Riekhakaynen, Dayter 2020)

Яркие особенности русской устной речи

- 1) Редукция словоформ
- 2) Стяжения на стыке слов



Яркие особенности русской устной речи

- 1) Редукция словоформ
- 2) Стяжения на стыке слов
- 3) Неожиданные паузы или отсутствие пауз там, где они должны быть

Паузация в спонтанной речи

- **Каноническое членение паузами:**

*(пауза 505 мс) э-э хотя как известно (пауза 441 мс) в нашей стране
разбираться значительно труднее (пауза 423 мс)*

- **Неканоническое членение паузами**

Паузация в спонтанной речи

- **Каноническое членение паузами:**

(пауза 505 мс) э-э хотя как известно (пауза 441 мс) в нашей стране разбираться значительно труднее (пауза 423 мс)

- **Неканоническое членение паузами**

1) единое смысловое целое разрывается одной или несколькими паузами:

потом он стал (вдох 477 мс) двадцать пятой коррекционной (пауза 418 мс) э-э-э школой

Паузация в спонтанной речи

- **Каноническое членение паузами:**

(пауза 505 мс) э-э хотя как известно (пауза 441 мс) в нашей стране разбираться значительно труднее (пауза 423 мс)

- **Неканоническое членение паузами**

1) единое смысловое целое разрывается одной или несколькими паузами:

потом он стал (вдох 477 мс) двадцать пятой коррекционной (пауза 418 мс) э-э-э школой

2) смысловые единства следуют друг за другом без каких-либо пауз:

(вдох 220 мс) чтобы у наших детей было понятие где они находятся ради чего они живут (вдох 239 мс)

ПРЕДЛОЖЕНИЯ ПО ОЗНАКОМЛЕНИЮ СТУДЕНТОВ, ИЗУЧАЮЩИХ РУССКИЙ ЯЗЫК КАК ИНОСТРАННЫЙ, С РЕДУЦИРОВАННЫМИ ЕДИНИЦАМИ

1. Начальный этап изучения языка:

- указание на то, что некоторые слова могут произноситься сокращенно;
- примеры наиболее частотных вариантов реализации форм вежливости (*здрасти, пасиба*);
- упоминание возможных редуцированных вариантов местоимений, наречий и т.п. по мере знакомства студентов с соответствующим грамматическим материалом.

ПРЕДЛОЖЕНИЯ ПО ОЗНАКОМЛЕНИЮ СТУДЕНТОВ, ИЗУЧАЮЩИХ РУССКИЙ ЯЗЫК КАК ИНОСТРАННЫЙ, С РЕДУЦИРОВАННЫМИ ЕДИНИЦАМИ

1. Начальный этап изучения языка:

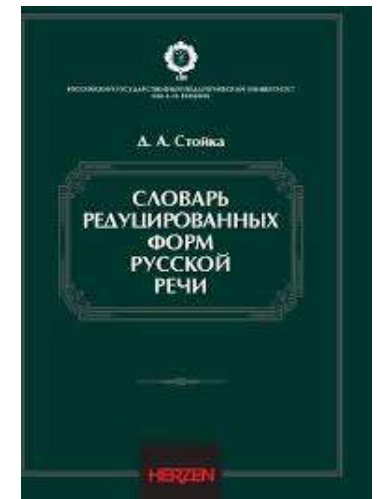
- указание на то, что некоторые слова могут произноситься сокращенно;
- примеры наиболее частотных вариантов реализации форм вежливости (*здрасти, пасиба*);
- упоминание возможных редуцированных вариантов местоимений, наречий и т.п. по мере знакомства студентов с соответствующим грамматическим материалом.

Откуда брать информацию?

ПРЕДЛОЖЕНИЯ ПО ОЗНАКОМЛЕНИЮ СТУДЕНТОВ, ИЗУЧАЮЩИХ РУССКИЙ ЯЗЫК КАК ИНОСТРАННЫЙ, С РЕДУЦИРОВАННЫМИ ЕДИНИЦАМИ

Откуда брать информацию?

Стойка, Д.А., Богданова-Бегларян Н.В. Словарь редуцированных форм русской речи. СПб.: Издательство РГПУ им. А. И. Герцена, 2019.



ПРЕДЛОЖЕНИЯ ПО ОЗНАКОМЛЕНИЮ СТУДЕНТОВ, ИЗУЧАЮЩИХ РУССКИЙ ЯЗЫК КАК ИНОСТРАННЫЙ, С РЕДУЦИРОВАННЫМИ ЕДИНИЦАМИ

2. Продвину́тый этап: работа с записями естественной речи – прослушивание и анализ:

- 1) фрагментов телевизионных передач;
- 2) фрагментов радиопередач;
- 3) фрагментов реалити-шоу / бытовых диалогов.

Откуда брать информацию?



Корпусы русской устной речи

Корпусы русской устной речи

- 1) **Национальный корпус русского языка** (<http://ruscorpora.ru/new/>):
 - устный подкорпус (13 001 271 словоупотребление);

1. [Беседа участников реалити-шоу «Дом-2», ТНТ \(2005\) // практиканты](#) [омонимия снята]
2. [Лекция \(отрывок\), Москва \(2005\) // Данилкина | практиканты](#) [омонимия снята]
3. [Праздный разговор молодых людей, Московская область \(2005\) // Лобзина | практиканты](#) [омонимия снята]
4. [Разговор в магазине \(2005\) // практиканты](#) [омонимия снята]
5. [Разговор на кухне, деревня в Челябинской области \(2005\) // Воротникова | практиканты](#) [омонимия снята]
6. [Разговор на лекции, Московская область \(2005\) // Данилкина | практиканты](#) [омонимия снята]
7. [Разговор на улице между женщиной и женщиной \(2005\) // практиканты](#) [омонимия снята]
8. [Разговор о будущем, Москва \(2005\) // Воротникова](#) [омонимия снята]
9. [Разговор при выходе из дома, Москва \(2005\) // практиканты](#) [омонимия снята]
10. [Совещание альпинистов, Бишкек \(2000-2005\) // Интернет](#) [омонимия снята]
11. [В. Жириновский. Выступление В. Жириновского в программе «Свобода слова», НТВ \(2004\) // Интернет](#) [омонимия снята]
12. [Беседа в Новосибирске \(2004\) // Фонд «Общественное мнение»](#) [омонимия снята]

3. Разговор на кухне, деревня в Челябинской области (2005) // Воротникова | практиканты [омонимия снята] [Все примеры](#)

[45, торговый работник, Н, жен] Н. Да ничего **я** не спросила. Л. А бабушка-то как? [Разговор на кухне, деревня в Челябинской области (2005) //

Воротникова | практиканты] [омонимия снята] ←...→

[19, студент, S, жен] Ааа/ ну слава Богу / а то **я** уже напугалась... [Разговор на кухне, деревня в Челябинской области (2005) // Воротникова | практиканты]

[омонимия снята] ←...→

[40, домохозяйка, Л, жен] Ну это-то ещё он заработал / как **я** поняла / в армш?... [Разговор на кухне, деревня в Челябинской области (2005) //

Воротникова | практиканты] [омонимия снята] ←...→

[13, школьник, К, жен] Ешь / ешь... К. **Я** ещё хочу мороженое... Пойду Клаву покормлю... Л. [Разговор на кухне, деревня в Челябинской области

(2005) // Воротникова | практиканты] [омонимия снята] ←...→

[45, торговый работник, Н, жен] **Я** ему полис отдала / пусть там на работе сходит... [Разговор на кухне, деревня в Челябинской области (2005) //

Воротникова | практиканты] [омонимия снята] ←...→

[S] С. Нет как-то... **Я** не ходила туда... Л. Такое сочетание у него там/ зелёное с синим / пол оранжевый... [Разговор на кухне, деревня в

Челябинской области (2005) // Воротникова | практиканты] [омонимия снята] ←...→

[40, домохозяйка, Л, жен] Нам надо сдавать назад / а он там лежит.... **Я** говорю / "Джувльбик / отойди". Не видно же его / он же маленький.

[Разговор на кухне, деревня в Челябинской области (2005) // Воротникова | практиканты] [омонимия снята] ←...→

[40, домохозяйка, Л, жен] Не видно же его / он же маленький. **Я** говорю / "Тебя ж задавят / отойди..." Серёже говорю / пока не / не это / не

сдавай / **я** сейчас Джувльбика уберу... [Разговор на кухне, деревня в Челябинской области (2005) // Воротникова | практиканты] [омонимия снята] ←...→

[40, домохозяйка, Л, жен] Серёже говорю / пока не / не это / не сдавай / **я** сейчас Джувльбика уберу... [Разговор на кухне, деревня в Челябинской

области (2005) // Воротникова | практиканты] [омонимия снята] ←...→

[40, домохозяйка, Л, жен] Серёже говорю / пока не / не это / не сдавай / **я** сейчас Джувльбика уберу... **Я** говорю / "Иди сюда!" . Дверца

машины открыта... **Я** сижу на сиденье. [Разговор на кухне, деревня в Челябинской области (2005) // Воротникова | практиканты] [омонимия снята] ←...→

4. Разговор на лекции, Московская область (2005) // Данилкина | практиканты [омонимия снята] [Все примеры](#)

[19, студент, № 3, жен] Давай быстрее. № 3. Если Катя пойдёт / **я** тоже пойду. А че мы скажем / когда будем выходить? [Разговор на лекции,

Московская область (2005) // Данилкина | практиканты] [омонимия снята] ←...→

Корпусы русской устной речи

- 1) **Национальный корпус русского языка (<http://ruscorpora.ru/new/>):**
 - устный подкорпус (13 001 271 словоупотребление);
 - мультимедийный подкорпус (4 540 588 словоупотреблений).


[Скачать](#)

Александр Рогожкин. Операция С Новым Годом, к/ф (1996)



[Сергей Олегович (Сергей Русскин), муж, 41, 1955] Я на компьютере работаю.
[Поздняков (Александр Поздняков), муж, 44, 1952] Тогда раньше.
[Сергей Олегович (Сергей Русскин), муж, 41, 1955] Ага. Угу.
[Поздняков (Александр Поздняков), муж, 44, 1952] Чёрт / я же ещё подарки на Новый Год не купил!
[Сергей Олегович (Сергей Русскин), муж, 41, 1955] Книга / лучший подарок!
[Поздняков (Александр Поздняков), муж, 44, 1952] да.
[Александр Рогожкин. Операция С Новым Годом, к/ф (1996)] [омонимия не снята]

Жесты:

Имя	Пол	Активный орган	Название жеста	Значение жеста
Александр_Поляков	мужской	рука	чесать_за_ухом	задуматься
Александр_Поляков	мужской	голова	посмотреть_на_часы	оценка_времени
Александр_Поздняков	мужской	рука	отвести_руку_в_сторону	что_тут_поделаешь

[Скачать](#)

Александр Рогожкин. Операция С Новым Годом, к/ф (1996)

Корпусы русской устной речи

1) Национальный корпус русского языка (<http://ruscorpora.ru/new/>):

- устный подкорпус (13 001 271 словоупотребление);
- мультимедийный подкорпус (4 540 588 словоупотреблений);

2) Рассказы о сновидениях и другие корпуса звучащей речи

(<http://spokencorpora.ru/>)



[Рассказы о сновидениях](#)

Корпус состоит из 129 рассказов детей и подростков от 7 до 17 лет об увиденном ими во сне. Рассказы записывались непосредственно после пробуждения. Общая длительность звучания — около 2 часов; объем корпуса — около 14 тысяч словоупотреблений.

[Подробнее](#)

[Рассказы сибиряков о жизни](#)

Корпус состоит из 17 рассказов взрослых жителей Новосибирска (от 19 до 70 лет) о каких-либо ярких событиях в их жизни. Общая длительность — около 40 минут звучания; объем — приблизительно 5 тысяч словоупотреблений.

[Подробнее](#)

[Весёлые истории из жизни](#)

Корпус состоит из 40 пар рассказов взрослых людей (от 18 до 60 лет) о смешных происшествиях в

[К списку корпусов](#)

NDS_001-f-z 7 Ж

[К списку записей
корпуса](#)

[Показать описание записи](#)

[Скачать
файл
ELAN](#)

Проигрыватель
файла целиком
[Старт](#)

Комментарии

[Скрыть](#)

Вид транскрипции

[Полная](#)

[Упрощённая](#)

[Минимальная](#)

Проигрывание
фрагмента
(1.84–2.97)

[Старт](#)

2.83 / 87.46

Транскрипция

Время	№	ЭДЕ	Комментарии
0.00	1.	Пото= ==	
0.23	2.	А /потом ··(0.08) про \зай <u>ч</u> ика сон.	
1.84	3.	··(0.46) А /зай <u>ч</u> ик <u>в</u> от,	В начальной паузе шумный вдох. Слово <u>вот</u> — скрипучим голосом.
2.97	4.	··(0.53) он был в \лес <u>у</u> .	
4.56	5.	···(1.20) А /я была у \р-р <u>е</u> чки.	
7.19	6.	··(0.43) {ЧМОКАНЬЕ 0.09} ··(0.46) А /потом ··(0.17) ког= ··(0.37) <u>в</u> от я /пош <u>л</u> а,	
10.41	7.	··(0.19) к \дом <u>о</u> й,	
11.05	8.	··(0.22) а /зай <u>ч</u> ик за \мн <u>о</u> й,	

Корпусы русской устной речи

1) Национальный корпус русского языка (<http://ruscorpora.ru/new/>):

- устный подкорпус (13 001 271 словоупотребление);
- мультимедийный подкорпус (4 540 588 словоупотреблений);

2) Рассказы о свидениях и другие корпуса звучащей речи

(<http://spokencorpora.ru/>)

3) Корпус русской устной речи (<http://russpeech.spbu.ru/>)

Корпус русской устной речи

- [Наши ресурсы](#)
- [Новости](#)
- [Транскрипция](#)
- [Публикации](#)
- [Источники](#)
- [О проекте](#)
- [Коллектив](#)

«Корпус русской устной речи»

создается с 2009 г. при поддержке грантов

[Российского фонда фундаментальных исследований](#)

№№ 09-06-00244,

13-06-00374

и 19-012-00629.

Корпус русской устной речи создаётся в СПбГУ как собрание звучащих тестов, снабженных орфографической и акустико-фонетической транскрипцией. Потребность в таком корпусе возникла прежде всего в связи с исследованиями процессов восприятия речи носителями языка: собранный материал используется при тестировании элементов функциональной модели восприятия речи, а главное, позволил оценить, с каким именно сигналом приходится иметь дело носителю языка в условиях естественного речевого общения.

В настоящее время корпус содержит звучащие тексты общим объёмом более 22-х тыс. словоупотреблений, отражающие разные стили речи: профессиональное дикторское чтение, чтение рядовыми носителями языка, монологическая спонтанная речь и детская речь. При орфографическом аннотировании во всех текстах восстановлена в правах буква "ё". В корпус включаются тексты с начала 40-х гг. XX века до настоящего времени.

На базе корпуса создан частотный словарь орфографических словоформ с указанием всех встретившихся вариантов их акустической реализации объёмом более 12-и тыс. словоформ. Архивированная версия словаря доступна для скачивания на странице "Наши ресурсы".

На данный момент существует несколько общедоступных корпусов русской устной речи: Мультимедийный подкорпус Национального корпуса русского языка (http://ruscorp.ru/new/search_muroc.html), Всероссийский корпус устной речи и другие корпусы

Корпус русской устной речи

- [Наши ресурсы](#)
- [Новости](#)
- [Транскрипция](#)
- [Публикации](#)
- [Источники](#)
- [О проекте](#)
- [Коллектив](#)

«Корпус русской устной речи»

создается с 2009 г. при поддержке грантов

[Российского фонда фундаментальных исследований](#)

№№ 09-06-00244,
13-06-00374
и 19-012-00629.

Корпус русской устной речи создается в СПбГУ как собрание звучащих тестов, снабженных орфографической и акустико-фонетической транскрипцией. Потребность в таком корпусе возникла прежде всего в связи с исследованиями процессов восприятия речи носителями языка: собранный материал используется при тестировании элементов функциональной модели восприятия речи, а главное, позволил оценить, с каким именно сигналом приходится иметь дело носителю языка в условиях естественного речевого общения.

В настоящее время корпус содержит звучащие тексты общим объемом более 22-х тыс. словоупотреблений, отражающие разные стили речи: профессиональное дикторское чтение, чтение рядовыми носителями языка, монологическая спонтанная речь и детская речь. При орфографическом аннотировании во всех текстах восстановлена в правах буква "ё". В корпус включаются тексты с начала 40-х гг. XX века до настоящего времени.

На базе корпуса создан частотный словарь орфографических словоформ с указанием всех встретившихся вариантов их акустической реализации объемом более 12-и тыс. словоформ. Архивированная версия словаря доступна для скачивания на странице "Наши ресурсы".

На данный момент существует несколько общедоступных корпусов русской устной речи: Мультимедийный подкорпус Национального корпуса русского языка (http://ruscorp.ru/new/search_muroc.html), Всероссийский корпус и другие корпусы

Корпус русской устной речи

- [Наши ресурсы](#)
- [Новости](#)
- [Транскрипция](#)
- [Публикации](#)
- [Источники](#)
- [О проекте](#)
- [Коллектив](#)

«Корпус русской устной речи»

создается с 2009 г. при поддержке грантов

[Российского фонда фундаментальных исследований](#)

№№ 09-06-00244,
13-06-00374
и 19-012-00629.

Уважаемый пользователь!

В настоящее время Вы имеете возможность работать с корпусом звучащих текстов объемом около 20 тыс. словоупотреблений. Все тексты корпуса снабжены орфографической и акустико-фонетической транскрипцией, а также восстановлена в правах буква «ё».

По корпусу создан частотный словарь транскрибированных словоформ, насчитывающий более 12 тыс. единиц.

Аннотированный корпус:

- частотный словарь орфографических словоформ с указанием всех вариантов их акустической реализации (файлы в формате CP1251, упорядоченные по типам текстов и количеству вариантов акустической реализации каждой лексической единицы, собраны в доступный для скачивания архивный файл slovart.rar);
- создание конкорданса по текстам корпуса с возможностью сохранения соответствующих звуковых «цитат».

Пользователю предоставляется возможность создания конкорданса, т.е. набора примеров-цитат из текстов корпуса, содержащих заданное в запросе слово. На каждое

Корпус русской устной речи

- [Наши ресурсы](#)
- [Новости](#)
- [Транскрипция](#)
- [Публикации](#)
- [Источники](#)
- [О проекте](#)
- [Коллектив](#)

«Корпус русской устной речи»

создается с 2009 г. при поддержке грантов

[Российского фонда фундаментальных исследований](#)

№№ 09-06-00244,

13-06-00374

и 19-012-00629.

Уважаемый пользователь!

В настоящее время Вы имеете возможность работать с корпусом звучащих текстов объемом около 20 тыс. словоупотреблений. Все тексты корпуса снабжены орфографической и акустико-фонетической транскрипцией, а также восстановлена в правах буква «ё».

По корпусу создан частотный словарь транскрибированных словоформ, насчитывающий более 12 тыс. единиц.

Аннотированный корпус:

- частотный [словарь](#) орфографических словоформ с указанием всех вариантов их акустической реализации (файлы в формате CP1251, упорядоченные по типам текстов и количеству вариантов акустической реализации каждой лексической единицы, собраны в доступный для скачивания архивный файл slovart.rar);
- создание конкорданса по [текстам](#) корпуса с возможностью сохранения соответствующих звуковых «цитат».

Пользователю предоставляется возможность создания конкорданса, т.е. набора примеров-цитат из текстов корпуса, содержащих заданное в запросе слово. На каждое

Корпус русской устной речи

Параметры поиска

Словоформа (в орфографии):

когда

Поиск

Допускается использование минуса (как дефиса).

Без основного ударения;
Без вторичного ударения

Параметры поиска

- «Е» и «Ё» эквивалентны
- Искать в составных словах и при стяжениях

Всего найдено 49 вхождений в 8 файлах(ле)

В дикторском: 0

В прочитанном: 9

В монологическом: 40

В диалогах: 0

21 с Ю+рия Гео+ргиевича. прощу+. [s u:r'i g'o:rC praSu] 182.597 183.449
pause 183.449 183.801

когда+ меня+ пригласи+ли на э+ту переда+чу, я+ проводи+л таки+е ми+ни опро+сы обще+ственного мне+ния, [kada
m'@n'a: pr'QglaIs'il'o: n@ et p'ir'idaIC e: provaId'il talk'i m'in'i a:pros@ apSe:stno mn'e:n'a:] 183.801 188.131
включи+л всю+ семью+, все+х друзе+й. [fkl'u:Cil s'o: s'imiu: fs'e:v druIz'e:] 188.131 190.003
pause 190.003 190.245

[Сохранить звуковой файл](#)

Корпус русской устной речи

Параметры поиска

Словоформа (в орфографии):

Допускается использование минуса (как дефиса).

Без основного ударения;
Без вторичного ударения

Параметры поиска

- «Е» и «Ё» эквивалентны
- Искать в составных словах и при стяжениях

Всего найдено **5 вхождений** в **2 файлах(ле)**

В дикторском: **0**

В прочитанном: **1**

В монологическом: **4**

В диалогах: **0**

Сопоставление восприятия частотных слов русского языка носителями языка и иностранцами

Riekhakaynen E.I., Balanovskaia A. SLIPS OF THE EAR AS CLUES FOR SPEECH PROCESSING // Communications in Computer and Information Science (см. в книгах). 2019. Т. 943. С. 111-121.

Сопоставление восприятия частотных слов русского языка носителями языка и иностранцами

100 частотных слов: 5 списков

О. Н. Ляшевская, С. А. Шаров, Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). М.: Азбуковник, 2009. <http://dict.ruslang.ru/freq.php>

с_с	с_с^і	с^і_с	с^і_с^і
		і	
і			
а	а^і	іа	іа^і
о	о^і	іо	іо^і
и	и^і	іи	іи^і
е	е^і	іе	іе^і

Бондарко Л.В., Вербицкая Л.А., Гордина М.В. Основы общей фонетики. СПб.: Филологический факультет СПбГУ; М.: Издательский центр «Академия», 2004. С. 98.

СПИСОК 1:

человек, любить, готовый, фильм, сказать, женщина, вместе, порядок, первый, стихи, взять, хорошо, форма, деньги, далёкий, процесс, получить, русский, ещё, любой

СПИСОК 2:

берег, поэтому, каждый, дом, уметь, широкий, революция, включать, где, намёки, сфера, настоящий, брат, тихий, модель, хозяин, вид, писать, генерал, серьёзный

УЧАСТНИКИ:

- 30 носителей русского языка;
- 30 китайцев, изучающих русский язык как иностранный (B1 и B2).

Результаты: носители русского языка

Ошибки в восприятии начального сегмента:

день → *тень*

больше → *Польша*

белый → *спелый*

цель → *щель*

Результаты: иностранцы

Ошибки в определении места ударного слога:

тѐтя → театр

пятница → пятнадцать

женщина → тишина

бывший → дыши

процесс → просит

готовый → годовой

Результаты: иностранцы

Ошибки в восприятии конечного сегмента:

тётя → театр

каждый → карты

начало → начать

первый → первого

Результаты: иностранцы

крикнуть →

регион →

фирма →

белый →

берег →

больше →

далёкий →

Результаты: иностранцы

крикнуть → кликнуть

регион → легион

фирма → фильма

белый → первый

берег → белок

больше → борщ

далёкий → дороги

Результаты: общие тенденции

Эффект прайминга:

- *готовый* → *готовить* (под влиянием предшествующего стимула *любить*);
- *общество* → *отчество* (под влиянием стимула *фамилия*).

Замена на более частотное слово:

фирма → *форма*, *сфера* → *вера*, *знать* → *значит* и др.

RLC

Русский учебный корпус

Русский язык в многоязычном мире

секция X Апрельской международной научной конференции НИУ ВШЭ

[Подробнее](#)



[Что такое Русский учебный корпус?](#)

[Метаданные](#)

[Партнёры](#)

[Языки](#)

[Поиск](#)

[Результаты поиска](#)

[Использование корпуса](#)

Что такое Русский учебный корпус?

В Русском учебном корпусе содержатся образцы устной и письменной речи двух категорий нестандартных говорящих на русском языке: изучающих русский язык как иностранный и так называемых зрительных говорящих. Для первой категории русский язык не является родным, представители же второй категории начали усваивать его в детстве как первый язык, но по разным причинам (в основном, это эмиграция) в качестве основного языка общения используют другой язык. Корпус позволяет производить поиск по лексико-грамматическим свойствам, а также по самым разным типам отклонений от стандартной русской речи - от орфографических ошибок до выбора лексических единиц и конструкций. Лингвистический анализ и разметка осуществляются членами Лаборатории по корпусным исследованиям НИУ ВШЭ под руководством Екатерины Рахилиной.

Основная часть текстов любезно предоставлена преподавателями русского как иностранного за рубежом, многие из которых специально работают и с зрительными говорящими. В Корпус входят как академические, так и неакадемические тексты. О представленных в Корпусе жанрах см. раздел [ПОМОЩЬ](#).

Частью Корпуса является лонгитюдный подкорпус письменной академической речи [РУДЕК](#), созданный Олесей Киселевой.

- [Метаданные](#)
- [Языки](#)
- [Типология ошибок](#)
- [Использование базы данных](#)

Что такое Электронная база речевых ошибок детей-инофонов?

База содержит речевые ошибки детей-инофонов, обучающихся в школах города Череповца Вологодской области. Дети-инофоны – это дети мигрантов, переехавших в Россию относительно недавно, владеющие русским языком на низком, пороговом уровне. Инофоны относятся к нестандартным носителям русского языка. Материалом послужили учебные письменные тексты: диктанты, сочинения, изложения, списывание. При проектировании базы учитывался опыт разработчиков учебных корпусов нестандартных носителей русского языка разных категорий.

Прежде всего это [Русский учебный корпус](#), разрабатываемый сотрудниками Лаборатории по корпусным исследованиям НИУ ВШЭ под руководством Екатерины Рахилиной.

Метаданные

В базе учтены следующие метаданные:

- имя ребенка-инофона
- возраст
- пол
- родной язык
- тип текста, в котором встретилась ошибка
- год создания текста

Языки

На данный момент в базе представлены тексты, созданные детьми-инофонами, для которых родными языками являются:

- армянский
- азербайджанский
- узбекский
- киргизский
- таджикский
- тагикский

Поиск в базе данных

Лексико-грамматический поиск

Словоформа Часть речи Тип текста Категория ошибки Дополнительно

я

Поиск

1

О. тип текста, часть речи, категория ошибки (тип ошибки), имя ребенка (возраст, родной язык)|год сбора материалов

словосочетание, содержащее **слово с ошибкой**

словосочетание с **исправленной ошибкой**

1. сочинение, местоимение-существительное, лексические (Лексические повторы в составе одного предложения или между несколькими предложениями), Тунар К. (10, азербайджанский)|2018

Я этим лето я подружился с новым другом.

- этим летом познакомился со своим новым другом.

Спасибо за внимание!



e.riehakajnen@spbu.ru