**Составление частотного словаря редуцированных форм слов для обучения аспектам русской разговорной речи на занятиях РКИ1**

***Мелкозерова В.И.***

*Студент, стажер-исследователь*

*Национальный исследовательский университет «Высшая школа экономики, Санкт-Петербургская школа гуманитарных наук и искусств, Лаборатории языковой конвергенции, Санкт-Петербург, Российская Федерация*

*E-mail:* *vimelkozerova@edu.hse.ru*

Главная цель обучения РКИ – овладение навыками аудиального восприятия русской звучащей речи и навыками формирования высказываний, соответствующих конкретной речевой ситуации. Известно также, что сверхчастотная лексика с наибольшей вероятностью подвержена редукции и звуковым изменениям, которые впоследствии могут укорениться и стать частью языковой системы (например, случай перехода глагола *молвил* в частицу-ксенопоказатель *мол*) [3]. Словоформы, которые «на сегментном уровне представлены меньшим количеством элементов, чем в полном, предусматриваемом нормами кодифицированного литературного языка варианте», называются *редуцированными формами слов* (далее — РФ) [1]. Владение иностранным языком кажется невозможным без представлений о реализации лексем частотного ядра в повседневной речи. Однако в современных учебных программах все еще наблюдается серьезный уклон в сторону изучения кодифицированного литературного языка и полного стиля произношения, из-за чего учащиеся сталкиваются с трудностями в условиях реального общения как при восприятии речи, так и при порождении.

Исследованиями восприятия РФ иностранными носителями языка занимается М. Эрнестус на материале нидерландского и французского [4]. В ходе экспериментов им было установлено, что РФ хорошо распознаются слушающими только в контексте, а без контекста или с ограниченным контекстом степень узнавания прямо зависит от степени редукции и частотности РФ. Более того, все участники (как носители, так и изучающие язык) распознавали высокочастотные, сильно редуцированные варианты быстрее, чем низкочастотные, слабо редуцированные варианты, поэтому знакомство с редуцированными вариантами на занятиях РКИ необходимо.

На материале русского языка проводились исследования порождения РФ китайскими студентами, изучающими русский язык не менее 4 лет с сертифицированным уровнем не ниже ТРКИ-1 [1]. Китайские студенты склонны редуцировать отдельные звуки также, как это делают носители русского языка, однако даже в высокочастотных словах наблюдается большое количество замен звуков. Этот эксперимент также подтверждает необходимость введения РФ в учебные материалы по РКИ.

Специализированные словари, в которых содержится информация об особенностях произношения и которые можно было бы использовать в практике преподавания РКИ, существуют: например, Д.А. Стойка составила наиболее полный дескриптивный словарь в ходе работы над диссертацией, которую можно считать фундаментальным исследованием редуцированных форм [3]. Словарь содержит 100 лексем, а также информацию об их частеречной принадлежности; устойчивые словосочетания; словоформы, подвергающиеся редукции; варианты написания; варианты произношения; примеры [2]. Но несмотря на то, что источниками для словаря были различные записи устной речи, в нем приводится информация о частотности РФ по Национальному корпусу русского языка, который содержит по большей части письменные тексты. Кроме того, количество примеров ограничено словарным форматом, а сами примеры зачастую относятся к источникам 19-20 столетий.

В настоящей работе был осуществлен автоматический поиск РФ на материале Корпуса устной речи студентов и молодежи, который состоит из наиболее современных записей [6]. Сленг молодежи эмоционален, в нем постоянно появляются новые формы выражения содержания. Это касается и употребления редуцированных форм: в группе людей до 30 лет сокращенные формы составляют 57,8% от общего количества сверхчастотных единиц [3].

Сбор материала в корпус ведется по непрерывной методике «Один речевой день», после чего записи проходят сегментацию на эпизоды, метаразметку, автоматическую акустическую расшифровку и экспертную редактуру [5].

В выборку вошли 207 расшифровок объемом почти 300 тысяч токенов, относящихся к 20 информантам. Важно отметить, что эпизоды относятся к различным коммуникативным ситуациям: диалоги происходят на работе, в университете, в магазине, дома и т.д. Тексты были токенизированы, после чего с помощью программы на Python осуществлялся поиск РФ. Для этого использовался список РФ Д.А. Стойка [3], который был дополнен РФ, найденными в акустических расшифровках. Была составлена таблица вида:

|  |  |  |  |
| --- | --- | --- | --- |
| РФ | Реплика, которая включает эту РФ | Номер эпизода | Таймкод реплики |
| те | А это случайно она не те? | escAF024\_18 | 00:13:15 |

Результаты были проверены вручную, затем был сформирован частотный словарь. Он включает 1417 вхождений, 24 типа. Наиболее частотные редуцированные формы и абсолютная частотность: *че*, 621 (что); *щас*, 330 (сейчас); *прям*, 283 (прямо); *тыщ\_*, 47 (лемма — тысяча); *ваще*, 34 (вообще); *ща*, 18(сейчас); *шо*, 15 (что); *седня*, 10 (сегодня); *те*, 10 (тебе); *скока*, 8 (сколько).

В этот словарь наиболее выразительные и частотные случаи редукции, которые зачастую отображаются и на письме, например, в интернет-дискурсе. В перспективе преподаватели РКИ могут обращаться непосредственно к данным о частотности конкретных РФ, репликам и диалогам, контекстам употребления РФ и соответствующим им записям звучащей речи.

Таким образом, были получены актуальные контекстные данные употребления РФ в устной речи и составлен частотный словарь РФ. Их уже сейчас можно использовать в качестве учебных материалов. В будущем целесообразно разработать эффективный метод ввода, усвоения и отработки РФ в иностранной аудитории.

**Литература**

1. Риехакайнен Е. И., Телова Ю. С. Фонетическая редукция словоформ в речи на русском языке как иностранном (в сопоставлении с устной речью на родном языке) // Социо- и психолингвистические исследования. 2019. №.7. С. 40-45.

2. Словарь редуцированных форм русской речи, Стойка Д.А. СПб, 2019.

3. Стойка Д.А. Редуцированные формы русской речи: лингвистический и экстралингвистический аспекты: дис. … канд. филол. наук СПб., 2017. 101 с.

4. Brand S. W. M., Ernestus M. T. C. Understanding reduced words. The relevance of reduction degree and frequency of occurrence. 2019.

5. Sherstinova T., Petrova I. ESC Corpus of Spoken Russian: Everyday Student Conversations Captured through Continuous Speech Recording in Natural Communicative Environments. 2024. P. 151–162.

6. Корпус устной речи студентов и молодежи: <https://esc-corpus.ru/>