

Тананайко Светлана Олеговна, Садуртинова Ксения Радиевна, Васильева Людмила Анатольевна
**СОВРЕМЕННЫЕ ИТ-ТЕХНОЛОГИИ ДЛЯ ИССЛЕДОВАНИЯ ФОНЕТИЧЕСКИХ ВАРИАНТОВ
ЗНАЧИМЫХ ЕДИНИЦ ЯЗЫКОВОЙ СИСТЕМЫ**

В статье рассматриваются инновационные методы изучения акустических вариантов морфологических единиц. Эта тема привлекает пристальное внимание лингвистов, поскольку в связи с высокой фонетической вариативностью морфологических единиц существует острая необходимость описания этой вариативности и представления данных в удобном для анализа виде. В статье описана программа для создания словаря акустических вариантов морфологических единиц в виде базы данных.

Адрес статьи: www.gramota.net/materials/2/2013/8-2/50.html

Источник

Филологические науки. Вопросы теории и практики

Тамбов: Грамота, 2013. № 8 (26): в 2-х ч. Ч. II. С. 182-184. ISSN 1997-2911.

Адрес журнала: www.gramota.net/editions/2.html

Содержание данного номера журнала: www.gramota.net/materials/2/2013/8-2/

© Издательство "Грамота"

Информация о возможности публикации статей в журнале размещена на Интернет сайте издательства: www.gramota.net

Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: voprosy_phil@gramota.net

ORGANIZATION OF CONFERENCE WEEK AS MEANS OF INCREASING TECHNICAL HIGHER EDUCATION ESTABLISHMENT STUDENTS' MOTIVATION TO STUDY FOREIGN LANGUAGE**Sumtsova Ol'ga Vital'evna***National Research Tomsk Polytechnic University**olgasumtsova0205@mail.ru*

The author considers the problem of increasing nonlinguistic specialties students' motivation in the study of a foreign language by involving them into participation in conference weeks, analyzes the main types of motivation, describes the experience of conducting conference weeks at National Research Tomsk Polytechnic University, contributing to its increase, and pays particular attention to activities organized within the framework of conference weeks.

Key words and phrases: conference week; extrinsic motivation; intrinsic motivation; communicative competence; foreign language communication.

УДК 81'34

Филологические науки

В статье рассматриваются инновационные методы изучения акустических вариантов морфологических единиц. Эта тема привлекает пристальное внимание лингвистов, поскольку в связи с высокой фонетической вариативностью морфологических единиц существует острая необходимость описания этой вариативности и представления данных в удобном для анализа виде. В статье описана программа для создания словаря акустических вариантов морфологических единиц в виде базы данных.

Ключевые слова и фразы: фонетическая вариативность; морфемы; уровни языковой системы; акустические варианты значимых единиц; база данных; словарь морфем; русский язык.

Тананайко Светлана Олеговна, к. филол. н.**Садуртинова Ксения Радиевна****Васильева Людмила Анатольевна***Санкт-Петербургский государственный университет**stananaiko@mail.ru; ksenia.sadurtinova@gmail.com; lavasil@mail.ru***СОВРЕМЕННЫЕ ИТ-ТЕХНОЛОГИИ ДЛЯ ИССЛЕДОВАНИЯ ФОНЕТИЧЕСКИХ ВАРИАНТОВ ЗНАЧИМЫХ ЕДИНИЦ ЯЗЫКОВОЙ СИСТЕМЫ[©]**

Статья посвящена теме, в равной степени актуальной как для практических лингвистически ориентированных разработок, так и для теоретической фонетики русского языка. Речь идет о выявлении и систематизации фонетических вариантов одних и тех же значимых единиц языковой системы. Появление таких вариантов неизбежно в звучащей речи, и это может объясняться самыми разными причинами: уровнем речевой культуры диктора, темпом речи, стилем произнесения, видом устной речи (чтение, подготовленная речь, спонтанный диалог и т.д.). На первом этапе обширного исследования данной темы, проводимого в настоящее время на кафедре фонетики СПбГУ, было принято решение сосредоточиться на фонетических вариантах морфем как единиц уровня языковой системы, непосредственно следующего за фонетическим и как самого нижнего уровня, знаковая природа единиц которого не оспаривается никем из лингвистов (что касается фонетического уровня, то признание его элементов, фонем, знаков зависит от того, какой точки зрения на знак придерживается исследователь, – унилатеральной (односторонней) или билатеральной (двусторонней)) [2].

В настоящей статье описывается опыт представления акустических вариантов морфологических единиц в компьютерном виде, удобном для дальнейшей компьютерной обработки и анализа, а именно в виде словаря акустических вариантов морфологических единиц, организованного как база данных, являющаяся наиболее удобным способом представления подобного словаря. Особая роль отводится описанию реляционного подхода к организации данных.

Говоря о теоретических основаниях проводимого исследования, необходимо упомянуть о свойствах языкового знака [4] и о двойном членении [3], т.е. возможности провести членение языковой последовательности, с одной стороны, на значимые единицы – морфемы, слова, синтагмы и т.д., а с другой стороны, на формальные единицы – фонемы, слоги, не имеющие самостоятельного значения, но обеспечивающие различие значимых единиц. Именно возможность рассмотрения единиц фонетического уровня речи как единиц формальных, в определенной степени произвольных и не связанных со значением, и позволяет столь широко использовать современные ИТ-технологии для их хранения в звуковой форме, их анализа и изучения.

Кроме того, при обсуждении результатов взаимодействия единиц разных уровней языка надо упомянуть о достижениях структурной, в частности дескриптивной, лингвистики. Структуралисты одними из первых

заговорили об уровне и иерархическом устройстве языковой системы, и именно структуралисты, и дескриптивисты в частности, проработали понятие уровня языка и представили язык как иерархическую систему таких уровней. В отечественной фонологии были приняты и получили свое развитие многие идеи структурной лингвистики. Так, Л. В. Бондарко [6, с. 81] говорит о следующих уровнях языковой системы: фонетический (или фонологический), морфологический, синтаксический и семантический. При этом про взаимодействие единиц каждого уровня Бондарко пишет следующее: «Анализ механизмов парадигматического противопоставления единиц каждого уровня показывает, что оно реализуется единицами низшего уровня; фонемы обеспечивают противопоставленность экспонентов морфем, словоформы противопоставлены с помощью морфем и т.д. Можно сказать, что переход с уровня на уровень осуществляет преобразование единиц низшего уровня в отношении единиц высшего уровня» [Там же, с. 82]. Однако при этом важно замечание о том, что единица высшего уровня не является суммой единиц более низкого уровня: «морфема не есть сумма фонем, а слово – сумма морфем, поскольку процесс функционирования языковой системы предполагает обязательное —приращение значений— на любом уровне... значение единиц более низкого уровня всегда является результатом анализа формальной структуры более высокого уровня» [Там же, с. 84].

Словарь акустических вариантов морфологических единиц создан на материале высококачественных звуковых записей, полученных при разработке корпуса *CORPRES* для синтеза речи на кафедре фонетики и методики преподавания иностранных языков филологического факультета СПбГУ в 2007-2009 годах. Общий объем материала, представляющего собой подготовленное чтение четырьмя профессиональными дикторами литературных текстов, составляет 35 000 словоупотреблений. В качестве основного морфемного словаря, послужившего основой представляемой базы данных, был использован «Словарь морфем русского языка» А. И. Кузнецовой и Т. Ф. Ефремовой [1].

Разработанный словарь морфологических единиц – это информационная система, основной задачей которой является хранение большого объема структурированных данных, таких как морфологические единицы и соответствующие им акустические и идеальные транскрипции. В словаре хранятся орфографические записи морфологических единиц, встречающихся в обработанном корпусе, с указанием класса морфологической единицы (приставка, корень, суффикс и т.д.). Кроме того, в словаре сохраняются взаимосвязи между морфологическими единицами и соответствующими акустическими и идеальными транскрипциями.

Основные требования, которые предъявлялись к словарю в процессе его создания и которым он удовлетворяет, таковы:

1) удобство доступа к данным.

Словарь разрабатывался для наблюдения вариативности морфологических единиц. Поэтому информационная система спроектирована таким образом, чтобы существовала возможность простого и быстрого доступа к хранимой в ней информации. В частности, существует возможность по орфографической записи морфологической единицы найти в словаре соответствующие ей акустические и идеальные транскрипции;

2) удобство добавления и корректировки данных.

Имеется возможность добавления и изменения данных в словаре;

3) портативность.

Система может без особых трудностей переноситься между различными физическими устройствами;

4) надежность.

Словарь спроектирован с высокой степенью надежности с точки зрения защиты от потери информации;

5) отсутствие избыточности хранимой информации.

Избыточность (дублирование данных) потенциально может привести к логически ошибочным результатам поиска или изменения данных словаря.

В качестве предпочтительного варианта реализации словаря была выбрана система баз данных.

Словарь морфологических единиц как система баз данных требует дополнительных программных продуктов для реализации уровня программного обеспечения, а именно для реализации системы управления базами данных (СУБД). База данных словаря морфологических единиц предполагает наличие нескольких таблиц, часть из которых содержит информацию об объектах морфологических единиц, акустических и идеальных транскрипций, а другая часть – «связующие таблицы», содержат информацию о взаимосвязях между этими объектами.

Преимуществами такого подхода к реализации словаря морфологических единиц являются отсутствие избыточных данных, скорость поиска в словаре, простота редактирования уже внесенной в словарь информации, поддержка на уровне СУБД средств резервного копирования. Такой способ реализации при правильно спроектированной базе данных позволяет избежать дублирования информации. Каждому объекту, будь то морфологическая единица определенного типа, акустическая или идеальная транскрипция, сопоставляется одна запись в соответствующей таблице. Взаимосвязи между объектами реализуются с помощью «связующих» таблиц. Скорость поиска обеспечивается за счет использования индексов, поддерживаемых всеми современными СУБД. К преимуществам способа реализации можно отнести также защиту от дублирования информации на уровне СУБД. Большинство современных СУБД позволяет использовать индексы для контроля уникальности записи в пределах таблицы базы данных.

Высокая надежность словаря морфологических единиц при таком подходе к реализации достигается за счет использования систем резервного копирования, которые поддерживаются современными СУБД. Такие системы позволяют минимизировать риск необратимой потери данных. Основными недостатками такого

способа реализации являются необходимость установки дополнительных программных продуктов для использования словаря и, как следствие, сложность переноса программы между различными устройствами.

В качестве СУБД выбрана *Microsoft SQL Server 2008 sp1 express*. Причин такого выбора несколько. Во-первых, *Microsoft SQL Server express* – это СУБД, которая устанавливается по умолчанию вместе со свободно распространяемым программным продуктом *Microsoft Visual C# 2008*. Во-вторых, СУБД поддерживает технологию *ADO.NET Entity Framework*. *ADO.NET Entity Framework* – объектно-ориентированная технология доступа к данным для *.NET* от *Microsoft*. Эта технология в значительной мере упрощает взаимодействие с базой данных на этапе создания словаря, позволяет использовать средства *LINQ* в процессе работы с базой. В-третьих, СУБД поддерживает индексы и все необходимые операции над ними, а также связи между таблицами и обеспечение. Архитектура базы данных подробно описана в нашей работе [5].

Интерфейс рабочей версии позволяет выводить наборы идеальных и акустических транскрипций морфологической единицы. В результате программа предлагает пользователю соответствующие данные. Например, на Рисунках 1 и 2 приведен вывод данных для корня *-ход-* и приставки *пре-*.

Таким образом, благодаря созданному словарю акустическая вариативность морфологических единиц представлена в компьютерном виде, удобном для дальнейшей компьютерной обработки, а благодаря разработанному интерфейсу данные об акустической вариативности морфологических единиц удобны для анализа и для дальнейшего использования в прикладных и теоретических целях.

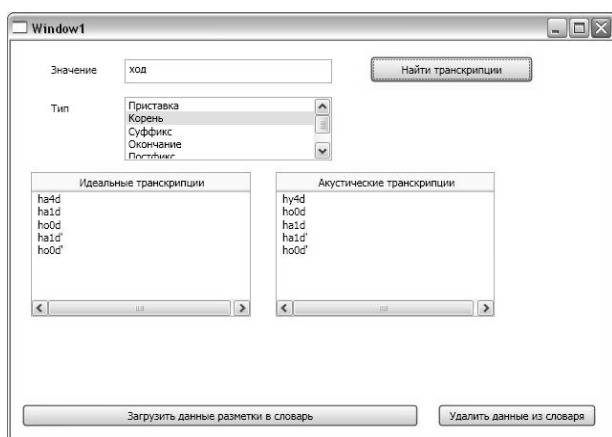


Рис. 1. Данные для корня *-ход-*

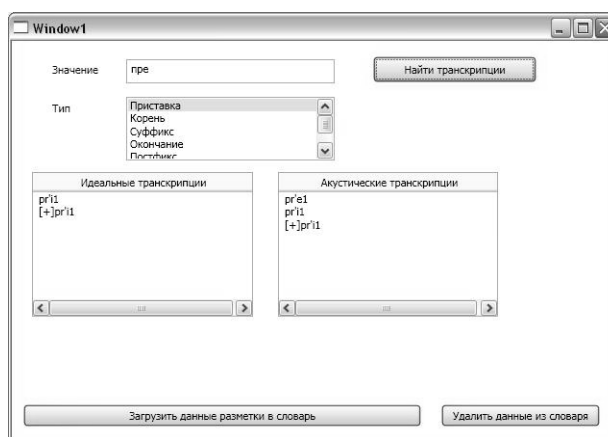


Рис. 2. Данные для приставки *пре-*

Список литературы

1. Кузнецова А. И., Ефремова Т. Ф. Словарь морфем русского языка. М.: Русский язык, 1986. 1136 с.
2. Кузнецова Н. Еще раз о понятии фонемы [Электронный ресурс]. URL: <http://www.dialog-21.ru/Archive/2005/KuznetsovaN/KuznetsovaN.pdf> (дата обращения: 02.07.2013).
3. Мартин А. Основы общей лингвистики. М.: Едиториал УРСС, 2004. 224 с.
4. Соссюр Ф. де. Курс общей лингвистики. М.: Едиториал УРСС, 2004. 256 с.
5. Тананайко С. О., Садуртинова К. Р. Фонетические варианты морфологических единиц: опыт создания словаря // Анализ разговорной русской речи (АР³-2012). СПб.: ГУАП, 2012. С. 40-45.
6. Уровни языка в речевой деятельности: к проблеме лингвистического обеспечения автоматического распознавания речи / отв. ред. Л. В. Бондарко. Л.: Изд-во ЛГУ, 1986. 260 с.

MODERN IT-TECHNOLOGIES FOR RESEARCH OF PHONETIC VARIANTS OF LANGUAGE SYSTEM NOTIONAL UNITS

Tananaiko Svetlana Olegovna, Ph. D. in Philology
Sadurtinova Kseniya Radievna
Vasil'eva Lyudmila Anatol'evna

St. Petersburg State University
stananaiko@mail.ru; ksenia.sadurtinova@gmail.com; lavasil@mail.ru

The authors consider the innovative methods for studying the acoustic variants of morphological units, tell that this topic draws attention of linguists, because due to the high phonetic variability of morphological units there is urgent need to describe this variability and data presentation in the form suitable for analysis; and describe the programme for creating the dictionary of the acoustic variants of morphological units in the form of a database.

Key words and phrases: phonetic variability; morphemes; language system levels; acoustic variants of notional units; database; dictionary of morphemes; the Russian language.