

грамматическая валентность

Шаг 2. Типы описания значений:

Энциклопедическое

Толковое

языковое

контекстное

Этимологическое

Отсылочное

к гиперониму

к производящему слову

морфемная связь

безморфемная связь

метафора

метонимия

синекдоха

к эквиваленту в языке-источнике

Синонимическое

полные синонимы

частичные синонимы

контекстуальные синонимы

стилистические синонимы

Антонимическое

языковые антонимы

контекстуальные антонимы

стилистические антонимы

Шаг 3. Дополнительные значения

Стиль

Коннотации

### Список литературы

- Баранов 2003 — Баранов, В. А. Формирование определительных категорий в истории русского языка / В.А.Баранов. — Казань: Изд-во Казанского гос. ун-та, 2003. — 390 с.
- Баранов 2006 — Баранов, В. А. От оцифрованных коллекций средневековых рукописей к электронным многофункциональным интернет-библиотекам / В. А. Баранов // [данный сборник] — С. 3–9.
- Никитин 1996 — Никитин, М. В. Курс лингвистической семантики : учеб. пос. для студ., асп. и препод. лингв. дисциплин в школах, лицеях, колледжах и вузах / М. В. Никитин. — СПб. : Науч. центр проблем диалога, 1996. — 760 с.

## Технологии и проблемы кодирования транскрипции в русской фонетике на основе различных алфавитных систем

О. В. Овчинников

Ижевский государственный технический университет, Россия

*Drawing up transcribitional binary latin-cyrillic system on the basis of coding Unicode according to the requirements of IPA (International Phonetic Alphabet): 1) revealing of signed unit of the transcribitional systems designated above alphabets; 2) comparison of the revealed pairs marks to coding Unicode; 3) recognition and completion of missing transcribitional pairs token from system IPA with adequate (exact or close) value; 4) "binding" of the revealed pairs transcribitional symbols under coding Unicode to one code value.*

В настоящее время для записи звучащей речи используется несколько систем знаков на основе разных алфавитов. Для русского языка — это кириллица и латиница. Однако до сих пор не решен вопрос адекватной кодировки/декодировки речи в компьютерных системах. Кроме того, желание исследователей применить наборную транскрипцию для широкого круга задач с использованием информационных систем приводит к постоянному изменению и/или дополнению существующих перечней знаков. Поэтому задача *адекватного* перевода знаков одной транскрипционной системы в другую полностью не решена.

Целью нашей работы является составление транскрипционной бинарной латинско-кириллической системы на основе кодировки Unicode в соответствии с требованиями IPA (International Phonetic Alphabet).

*Задачи:* 1) выявление единства транскрипционных систем на основе латинского и кирилловского алфавитов; 2) сопоставление выявленных пар знаков с кодировкой Unicode; 3) выявление недостающих транскрипционных знакопар в системе IPA с адекватным (точным либо близким) значением; 4) «привязка» выявлен-

ных пар транскрипционных символов Unicode к одному кодовому значению.

Основной механизм работы состоит из двух этапов: 1) сопоставление транскрипционных знаков с аналогичным значением (например, ‘с’ *кир.* и ‘s’ *лат.* репрезентируют фонему ‘шумный щелевой глухой переднеязычный зубной согласный’); 2) установление связи типа «аналог».

Одной из проблем является то, что внешне схожие знаки в разных транскрипционных системах имеют разное значение. К примеру, знак ‘Λ’ *кир.* обозначает ‘безударный гласный первой степени редукции средне-нижнего подъема средне-заднего ряда’, а ‘Λ’ *лат.* — ‘неогубленный гласный средне-нижнего подъема заднего ряда’.

В тех случаях, когда не удастся напрямую подобрать пару адекватных транскрипционных знаков, обращаемся к IPA. Сопоставление значений знаков в разных транскрипционных системах позволяет подобрать необходимый недостающий знак в одном из диапазонов Unicode. К примеру, такие транскрипционные знаки, как ‘ts’, ‘tʃ’ и ‘tɕ’, репрезентирующие аффрикаты, в кириллической (русской) традиции обозначаются знаками ‘ц’, ‘ч’ (= [tʃ]) и ‘ч’ (= [tʃʲ]) соответственно.

Однако подбор знаков с помощью IPA не всегда результативен, так как несколько знаков этой системы могут обозначать звуки, близкие по характеристикам, различающиеся лишь тонкими артикуляционными характеристиками. Например, символы, обозначающие гласные [ɔ] и [ɛ], имеют совпадающие характеристики ‘гласный заднего ряда нижнего подъема’, но различаются тем, что первый звук произносится без огубления, второй — лабиализован. На основе этого принимается решение о вводе в одну из транскрипционных систем дополнительного знака, если это диктуется фонетическими характеристиками звука или фонемы.

При отсутствии знака в IPA с необходимым для нас значением должна быть решена следующая задача — необходимость создания комбинированных знаков на основе уже существующих знаков в кодировке Unicode.

Существенная проблема заключается и в том, где разместить такой комбинированный знак в кодировочной таблице. Это зависит, в частности, от того, какая из алфавитных систем принимает

ся в качестве основной. Основной мы определи для себя транскрипцию на базе латиницы, поэтому именно в латинском диапазоне Unicode мы разместили глифы латинского и русского алфавитов.

В русском языке существуют также звуки, для которых ни в одной из имеющихся кодировок нет соответствующих символов. Для таких звуков созданы символы, содержащие в своем составе дополнительные значки, которые передают артикуляционные особенности этих звуков.

Таким образом, разработана система знаков, с помощью которой можно подготовить транскрипцию как на латинского и кирилловского алфавитов с возможностью транслитерации без потерь информации.

Этот набор транскрипционных символов используется для создания фонетического атласа русских говоров Удмуртии в лингвогеографической информационной системе «Диалект», созданной для систематизации и структурирования лексических и фонетикофонологических особенностей говоров.

### Список литературы

- Аванесов 1968 — Аванесов, Р. И. Русское литературное произношение : учеб. пособие для студ. пед. ин-тов / Р. И. Аванесов. — М. : Просвещение, 1968. — 287 с.
- Баранов и др. 1998 — Баранов, В. А. Методические указания к диалектологической практике (для студентов филологического факультета) / В. А. Баранов, С. А. Никифорова, Н. Е. Якименко. — Ижевск : Удм. госун-т, 1998. — 82 с.
- Суслов 1999 — Суслов, И. П. Вариации на фонетическую тему // Лингвистический вестник. — Вып. I. — Ижевск : Изд-во УдГУ, 1999. — С. 13 — 19.
- Моисеев 1980 — Моисеев, А. И. Русский язык: Фонетика. Морфология. Орфография : пособие для учителей / А. И. Моисеев. — М. : Просвещение, 1980. — 255 с.
- Осипов 1984 — Осипов, Б. И. Справочные таблицы к практическим занятиям по фонетике современного русского языка / Б. И. Осипов. — Ижевск : Изд-во УдГУ, 1984. — 14 с.
- Руководство 1998 — Руководство по фонетической транскрипции : методические указания для студентов филологического факультета / сост. Б. И. Осипов. — 2-е изд., испр. и доп. — Омск : Омск. гос. ун-т, 1998. — 22 с.

- IUC 2006 — Internationalization & Unicode Conference [Электронный ресурс]. — 2006. — Режим доступа: <http://www.unicodeconference.org>, свободный. — Загл. с экрана.
- UCL 2006 — UCL Dept of Phonetic & Linguistics [Электронный ресурс]. — 2006. — Режим доступа: <http://www.phon.ucl.ac.uk>, свободный. — Загл. с экрана.
- Unicode 2005 — Unicode [Электронный ресурс]. — 2005. — Режим доступа: <http://www.unicode.org>, свободный. — Загл. с экрана.
- Wikipedia 2006 — Wikipedia The Free Encyclopedia [Электронный ресурс]. — 2006. — Режим доступа: [http://en.wikipedia.org/wiki/Main\\_Page](http://en.wikipedia.org/wiki/Main_Page), свободный. — Загл. с экрана.

## **Об электронной версии «Словаря современного русского города»**

Б. И. Осипов

Омский государственный университет, Россия

1. С 1988 г. словарная группа филологического факультета Омского университета начала работу над словарем народно-разговорной речи современного русского города, включающим все разновидности разговорного лексикона: разговорную литературную, просторечную, жаргонную лексику, а также ту часть диалектизмов, которая продолжает функционировать в речи горожан. В 2003 г. «Словарь современного русского города» в его печатной версии был опубликован в Москве издательством «Русские словари». Планируется и второе издание этого словаря.

2. Однако параллельно с печатной словарная группа создавала и электронную версию. Подготовка печатного издания на какое-то время задержала эту работу, но теперь она возобновляется.

3. Электронная версия отличается тем, что, кроме постоянного пополнения материала, позволяет автоматически находить нужную лексикологу информацию. Предусмотрен автоматический вывод слов, имеющих одинаковые акцентные пометы (тип акцентной парадигмы), а также пометы грамматические, стилистические и характеризующие сферу употребления слова. Предусматривается также автоматический вывод слов с заданным буквосочетанием, что должно облегчить поиск однокоренных образований при анализе словообразовательной системы городской речи.

4. Кроме того, предусмотрена автоматическая система вывода в прямом и обратном алфавитном порядке всех словоформ, встречающихся в текстовых примерах, приводимых при толковании слов. Это позволяет отчасти придать словарю характер полного, хотя в общем он проектируется как дифференциальный (отражающий только лексику обозначенных выше разрядов).

5. В перспективе планируется предоставление словаря пользователям машинного фонда русского языка при Институте русского языка Российской академии наук.