

# Преобразование между орфографическим и фонемным текстами для моделирования спонтанного произношения

*Робейко В.В., научный сотрудник*

*Сажок Н.Н., кандидат технических наук,  
старший научный сотрудник*

В статье рассматривается актуальный вопрос порождения последовательностей фонем по орфографическому тексту (транскриптор) и обратное преобразование. Основная цель работы — разграничение оперативного кода (команд) и данных, что позволит превратить транскриптор в гибкий и удобный инструмент исследователя. Предложена модель, в которой заложены закономерности связи между фонетическими и орфографическими символами. Многозначные преобразования, осуществляемые согласно модели, эквивалентны построению направленного графа. При переходе вдоль стрелок графа осуществляется генерирование последовательностей фонетических символов, которые соответствуют исходному тексту. Параметры модели задаются таблично в виде контекстно-зависимых правил, которые формирует эксперт. В модели предусмотрено введение дополнительных уровней транскрибирования, что позволяет эксперту строить сложные преобразования, работая с относительно простыми правилами. Разработана система, на вход которой подаётся орфографический текст с обозначениями ударений, а на выходе получаем последовательности фонем, соответствующих различным вариантам произношения исходного текста. Практическое применение системы для автоматического распознавания спонтанной украинской речи позволяет увеличить надёжность распознавания для ряда случаев. В заключение обсуждаются сопутствующие задачи и перспективные исследования.

• графема • фонема • преобразование • транскрибирование • спонтанная речь.

In this paper we consider the actual problem of grapheme-to-phoneme (GTP) conversion for forward and backward directions. The main purpose is to separate the operational code (instructions) and the data that is the way to obtain the flexible and convenient GTP tool for the researcher. We propose a model describing regularities of relations between the phonetic spelling and symbols. Multi-decision transformations carried out according to the model, are equivalent to building a directed graph. Moving along the arrows we can generate multiple phoneme sequences corresponding to the input text. The model parameters are specified in tabular form as a set of context-dependent rules formed by the expert. We introduce additional model levels allowing the expert to build complex transformations, still working with relatively simple set of rules. The developed system takes the orthographic text with the pointed accentuation and produces phoneme sequences corresponding to the different pronunciation ways and manners of the input text. Practical applications of the system for automatic recognition of Ukrainian spontaneous speech showed increasing accuracy for certain cases. Finally, we discuss related tasks and further research.

• grapheme • phoneme • conversion • transcription • spontaneous speech.

Порождения последовательностей фонем по орфографическому тексту и последовательностей слов по фонемному тексту — актуальная проблема в области речевой информатики. Графемно-фонемные преобразования нужны для транскрибирования текстов в системах озвучивания информации, для формирования словарей произношения при оценке параметров акустической модели в различных схемах декодирования речевого сигнала и т.п. [1–4].

Преобразование орфографического текста (графем) в фонемный текст невозможно без изучения закономерностей связи между фонетическими и орфографическими символами.

Долгое время графемно-фонемное преобразование реализовывалось в виде определённого алгоритма — транскриптора, воспроизводящего в программном коде правила чтения из учебника или справочника [6].

Нужно отметить, что, в отличие от орфографического, фонемный текст для каждого определённого слова не является устойчивым как по субъективным, так и по объективным причинам.

Во-первых, до сих пор среди исследователей в области фонетики и фонологии продолжают дискуссии относительно алфавитов фонем для языков и диалектов (несмотря на принятый стандарт Международной фонетической ассоциации — IPA). Во-вторых, алфавит фонем для системы распознавания или синтеза речи зависит не только от свойств языка, но и от того, какое фонетическое явление (ассимиляция, редукция, палатализация и др.) учитывается на уровне символов, а какое — на уровне акустической модели путём введения контекстной зависимости (CD-фонемы, фонемы-трифоны), использование смеси гаусоидов (GMM) или изменения топологии акустической модели HMM. В-третьих, при разработке речевых технологий должны учитываться индивидуальные, ситуативные особенности произношения, произношение слов в потоке речи, словарные и смысловые ударения. Это влечёт введение многозначности при переходе к фонемному тексту.

Отсутствие разграничения между оперативным кодом (командами) и данными — серьёзное препятствие для модификации и усовершенствования транскриптора и делает невозможным его превращение в гибкий инструмент исследователя.

Заметим, что в мире для преобразования графем в фонемы распространены статистические методы, которые анализируют созданный экспертами словарь произношений [5]. Это оправдано для языков, в которых орфографическое написание слов существенно отличается от их произношения (например, для английского или французского). К недостаткам такого подхода можно отнести игнорирование возможной коартикуляции на стыках слов и субъективизм экспертов.

Задача обратного перехода от фонемного текста к орфографическому является относительно новой. Она приобрела актуальность с развитием многозначной многоуровневой модели понимания речевого сигнала [4]. Реализованному в то время алгоритму подавалось на вход менее 30 правил преобразования «фонема — графема» для украинского языка, в результате чего генерировались гипотетические варианты орфографического текста для многозначного ответа распознавания свободного порядка следования фонем. Такое незначительное количество правил объясняется не только свойствами украинской орфографии, в которой правила написания главным образом основываются на фонетическом принципе (пишем то, что слышим). На выходе фонемно-графемный преобразователь дополняется

лексическим фильтром, который отбрасывает недопустимые гипотезы орфографического написания.

Дальнейшие исследования алгоритма заключались в его использовании при обратном процессе — преобразовании орфографического текста в фонемный. При этом ставилась цель сохранения прозрачности и относительной простоты правил для эксперта и одновременного расширения подхода на языки, у которых отличие написания от произношения весьма существенно, и таким образом получить универсальный транскриптор, который можно сравнительно легко модифицировать на уровне структур данных без дополнительной компиляции программного кода.

### Модель многозначного преобразования последовательностей символов

Пусть задана некоторая конечная последовательность символов

$$(a_1, a_2, \dots, a_n, \dots, a_N) \equiv a_1^N, a_n \in \mathbf{A}, \quad (1)$$

где  $\mathbf{A}$  — алфавит входящих символов. Сконструируем отображение этой последовательности на множество последовательностей выходящих символов из некоторого иного алфавита  $\mathbf{B}$ .

Рассмотрим функцию  $f$ , отображающую последовательность  $a_1^N$ , начиная с её  $n$ -го символа, в символ алфавита  $\mathbf{B}$  или пустое множество:

$$f : a_n^N \rightarrow b, b \in \mathbf{B} \cup \emptyset, 1 \leq n \leq N. \quad (2)$$

Заметим, что (2) справедливо лишь в случае, когда входящая последовательность принадлежит области определения  $f$ , т.е.  $a_n^N \in \text{Def}(f)$ . Множество последовательных применений таких функций переводит  $a_n^N$  в последовательности символов из алфавита  $\mathbf{B}$ , и таким образом мы конструируем мультифункцию:

$$F(a_n^N) = \{f_1^k(a_n^N), f_2^k(a_n^N), \dots, f_{L_k}^k(a_n^N)\} \in \mathbf{B}^{L_k} \cup \emptyset, 1 \leq k \leq K_F, \quad (3)$$

где  $L_k$  — длина  $k$ -й выходящей последовательности, общее количество которых  $K_F$ , своё для каждой  $F \in F$ .

Определим аналог прямого произведения над множествами, полученными вследствие действий мультифункций из  $\mathbf{F}$ , как перебор всех вариантов объединения конечных последовательностей символов алфавита  $\mathbf{B}$ , т.е. опуская аргументы мультифункций:

$$F \otimes G = \{f_1^u, f_2^u, \dots, f_{L_u}^u, g_1^v, g_2^v, \dots, g_{L_v}^v\}, 1 \leq u \leq K_F, 1 \leq v \leq K_G. \quad (4)$$

Допускаем по определению, что если результат действия  $F$  или  $G$  является пустым множеством, то результатом их произведения будет пустое множество. В отличие от декартового произведения для определённого нами аналога выполняется свойство ассоциативности.

Рассмотрим упорядоченное множество  $\tilde{\mathbf{F}}$  мультифункций  $F \in F$ , которые сопроводим дополнительными параметрами:

$$\tilde{\mathbf{F}} = (F_{i,d_i,\delta_i}), 1 \leq i \leq |\tilde{\mathbf{F}}|, d_i > 0, \delta_i = \{0,1\}, \quad (5)$$

где  $i$  — индекс мультифункции в упорядоченном множестве  $\tilde{\mathbf{F}}$ ; параметр  $d_i$  — ширина шага анализа,  $\delta_i$  — условие исключительности. Через эти параметры формулируем ограничения при вычислении произведения

$$\bigotimes_{i,n} F_{i,d_i,\delta_i}(a_n^N), 1 \leq i \leq |\tilde{\mathbf{F}}|, 1 \leq n \leq N. \quad (6)$$

Предположим, что мы уже вычислили выражение (6) на некоторых упорядоченных индексных множествах  $J$  и  $M$  и получили некоторое непустое множество

$$G_{J,M} = \bigotimes_{u \in J, v \in M} F_{u,d_u,\delta_u}(a_v^N). \quad (7)$$

Пусть  $j$  и  $t$  являются последними элементами индексных множеств  $J$  и  $M$  соответственно. Тогда при рассмотрении следующего компонента произведения,  $F_{i,d_i,\delta_i}(a_n^N)$ , проводим вычисления согласно с определением (4), если выполняются такие условия:

$$\begin{cases} m + d_j = n; \\ \delta_r \neq 1, 1 \leq r < i; \\ \bigotimes_{u \in J, v \in M} F_{u,d_u,\delta_u}(a_n^N) \otimes F_{r,d_r,\delta_r}(a_n^N) \neq \emptyset, 1 \leq r < i, \text{ если } \delta_i = 1. \end{cases} \quad (8)$$

В противном случае, при поступлении следующего компонента произведения получаем пустое множество.

Выражением (6) порождаются последовательности выходящих символов по некоторой последовательности входящих символов. Если исходный алфавит совпадает с алфавитом букв определённого языка, а выходящий алфавит состоит из фонем, то получаем многозначный транскриптор орфографического текста. И наоборот, если на входе — фонемный алфавит, а на выходе — алфавит букв, то получаем многозначное преобразование из фонемного текста в орфографический. Возможны промежуточные варианты.

Пример порождения реализаций фонемного текста по орфографическому тексту приведён на рис. 1. Рассматривается орфографический текст слова «снег». С целью обобщения стандартный алфавит дополнен символом «\_», который разделяет слова. Соответствующие ударным гласным буквы переводятся в верхний регистр, все остальные буквы — в нижний. Позиция ударения определяется по орфоэпическому словарю с учётом омографии [10] или же учитываются все допустимые позиции ударения.

Таким образом, имеем на входе последовательность из шести символов  $a_1^N = (\langle \_ \rangle, \langle c \rangle, \langle n \rangle, \langle E \rangle, \langle z \rangle, \langle \_ \rangle)$ ,  $N = 6$ . На графе отображены все допустимые мультифункции  $F_{i,d_i,\delta_i}(a_n^N)$ ,  $1 \leq n \leq N$ . Производя переход по стрелкам, получаем произведения вида (6), генерирующие четыре последовательности фонем или фонемных текста:

$\langle \_ c n' E z \_ \rangle$ ;  $\langle \_ c n' E k \_ \rangle$ ;  $\langle \_ c' n' E z \_ \rangle$ ;  $\langle \_ c' n' E k \_ \rangle$ .

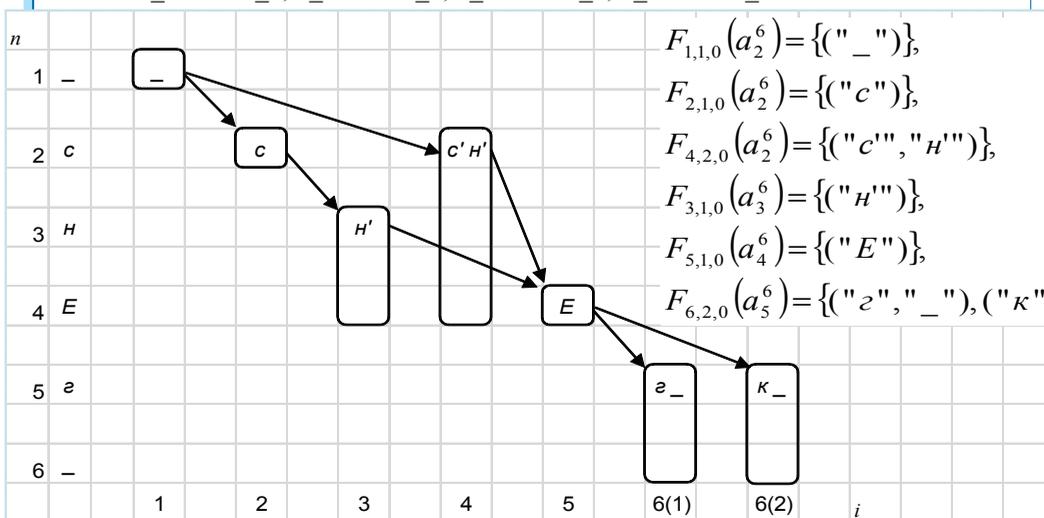


Рис. 1. Граф многозначного преобразования орфографического слова «снег» в фонемную транскрипцию

Заметим, что на практике используется не вся подпоследовательность вида  $a_n^N$ , а лишь её часть,  $a_n^{n-1+T}$ , где положительный параметр  $T$  зависит от конкретной мультифункции и определяет ширину контекста, в котором рассматриваются входящие символы. На рис. 1 высота прямоугольников в узлах графа соответствует ширине контекста.

Результат преобразования может быть многократно подвергнут описанным преобразованиям с теми же или другими параметрами. Таким образом мы можем вводить дополнительные уровни транскрибирования, постепенно переходя к выходным символам. Это позволяет существенно упростить спецификацию параметров, что важно для языков, в которых написание во многом отличается от произношения. Также появляется возможность применить идентичный алгоритм с другим набором правил к (многозначной) расшифровке чисел и сокращений и транслитерации.

### Система многоуровневого многозначного транскрибирования орфографических текстов

Предлагается система, на вход которой подаётся орфографический текст, содержащий только символы из алфавита букв, включая символы границы между словами и морфемами и обозначения ударения. Такой текст получаем вследствие автоматической обработки произвольного текста или списка слов, которая заключается в вербализации (расшифровке) символов, чисел и сокращений, расстановке ударений и разделения на синтагмы, слова и морфемы [6]. На выходе системы получаем последовательности фонем, соответствующие различным вариантам произношения входного текста.

Эта система транскрибирования используется для распознавания украинской речи [8, 11].

Разработанная система многозначного транскрибирования орфографических текстов использует модель (1) — (8), которая предусматривает возможность таблично задавать контекстно-зависимые правила преобразований одних обобщённых последовательностей символов на другие. При этом в каждом правиле задаётся ширина шага, по которому происходит переход к следующей последовательности символов. Применение многих правил позволяет генерировать сразу несколько вариантов транскрипции одного и того же слова или генерировать нужный вариант из нескольких возможных, например, описывая спонтанную речь говорящего или группы дикторов (таблица 1).

Таблица 1

#### Примеры построения правил транскриптора для украинской речи

Входящая обобщённая подпоследовательность символов	Выходящие символы	Ширина шага	Пояснения
[зсц] [жшч]	[жшч]	1	<b>з, с, ц</b> перед <b>ж, ш, ч</b> переходят соответственно в <b>ж, ш, ч</b>
[тс][дтзснц][ісюяіЄЮЯ]	т'	1	<b>т и с</b> перед мягкими <b>д, т, з, с, н, ц</b> смягчаются
с т [лн]	с	2	<b>т</b> между <b>с и л</b> или <b>н</b> выпадает

В целом для транскрибирования украинской речи (для литературного варианта произношения) достаточно ввести чуть более 30 подобных правил. Для литературного варианта русской речи уже потребовалось вводить уровни, постепенно отходя от орфографического написания к фонемному тексту. Предложено 5 уровней и около 120 правил.

Последующее развитие системы транскрибирования для украинского языка происходило на основе анализа спонтанной речи нескольких сотен дикторов [8,11]. Для всех дикто-



ров был создан общий вариант транскрибирования на основе литературного произношения. Кроме этого, все дикторы были распределены на группы, для которых разработаны свои правила индивидуализированного транскрибирования, дополняющие или заменяющие основной вариант [9]. Также для одной из групп дикторов был разработан общий вариант транскрибирования спонтанной слитной речи [12].

Результаты изучения спонтанного вещания многих дикторов свидетельствуют о том, что никто из них не соблюдает орфоэпических правил произношения в полном объёме. Прежде всего, это касается запрещённых литературной нормой регрессивной ассимиляции по глухости в паре фонем «звонкая + глухая» и оглушение согласных перед паузой (**тобто** → **т о п т о**; **підтримати** → **п' і т т р и м а т и**; **робив** → **р о б и ф**). Дикторы с такими особенностями произношения были собраны в отдельную группу. Было также установлено много других характерных черт произношения разных дикторов: редукция окончаний некоторых слов (прилагательных, глаголов) в слитной речи (**шановний** → **ш а н о в н и**; **доброго** → **д о б р о**), «аканье» (**робити** → **р а б и т и**), твёрдое произношение мягких согласных (**синього** → **с и н о г о**) и пр.

Для некоторых слов (например, служебных частей речи, слов с разными ударениями) задаётся несколько вариантов транскрипций — с ударением на различных слогах или вообще без ударения: **коли** → **к о л и**; **к о л и**; **к о л и**.

Такие тенденции моделируются путём изменения правил перехода от одних последовательностей символов к другим и расширением действующих правил.

Все правила индивидуализированной модификации транскрипций можно разделить на несколько групп (на основе обработки звуковых записей).

К *позиционным*, зависящим от общих фонетических условий — позиции в составе / слове, ударности / безударности т.д. [7] нелитературным изменениям звуков в потоке речи относим:

- 1) ослабленное произношение *о* как *а* в безударной позиции, реже встречается редукция безударных гласных до полного исчезновения (**тепер** → **т и п е р**, **зозуля** → **з у з у л' а**, **боротьба** → **б а р а д' б А** или **б р а д' б А**);
- 2) оглушение звонких согласных перед паузой (**бріd** → **б р' і т**, **зараз** → **з А р а с**);
- 3) редукцию в терминальных частях слов в процессе произношения — исчезновения согласного звука в окончаниях — *ого*, — *их*, — *ич*, — *ий*, — *их*, — *ий*, — *иї*, — *ої*, — *еї*, — *ою*, — *ею*, — *ити* и подобных (**коротший** → **к о р о ч ш и**, **Іванович** → **і в А н о в и**, **синіх** → **с и н' і**, **безпеку** → **б е с п е к о у**); исчезновения конечного гласного звука в окончаниях — *ою*, — *ею*, — *ею* и подобных (**доброю** → **д о б р о й**, **землею** → **з е м л е й**) и пр.

К *комбинаторным* (качественные и количественные изменения соседних звуков [7]) нелитературным изменениям относим:

- 1) полную регрессивную ассимиляцию по глухости в сочетании „звонкий+глухой» на границе любых морфем в слове и на границе слов (**без причини** → **б е с п р и ч и н и**, **розсунути** → **р о с с у н у т и**, **книжка** → **к н и ш к а**, **сядьте** → **с' А т' т е**);
- 2) ассимиляцию по мягкости свистящих и шипящих согласных, губных и заднеязычных согласных (**злі** → **з' л' і**, **шлях** → **ш' л' А х**, **квітка** → **к' в' і т к а**);
- 3) произнесение удлинённого согласного звука как обычного неудлинённо-

го, произнесение двух гласных как одного звука (віддати → в' і д А т и, знання → з н а н' А, зоопарк → з о п А р к, аеропорт → а р о п О р т);

4) неполное упрощение в группах согласных, его отсутствие (чесний → ч Е с т н и й) и пр.

Для создания индивидуализированных словарей произношения все исследуемые дикторы были распределены на группы. Это происходило в несколько этапов: первый предусматривал выделение группы дикторов с литературной речью и дикторов с отклонениями от литературной нормы; следующим участком работы было выделение общих черт речи во второй группе и создание соответствующих транскрипций для каждой из подгрупп. Таким образом было сгенерировано 18 видов транскрипций словаря для тех или иных дикторов. Адекватность индивидуализированных транскрипций проверялась с помощью распознавания речи (улучшение / ухудшение результатов распознавания для конкретного диктора по сравнению с результатами распознавания для этого же диктора с использованием литературной транскрипции). Каждая из транскрипций была проверена на всех дикторах (таблица 2).

Таблица 2

### Примеры индивидуализированных словарей спонтанной речи

Литературная транскрипция \ Индивидуализированный словарь	devocal	end_cons	a	devocal_a
р о б И в	р о б И ф	р о б И	р а б И в	р а б И ф
в О р о г	в О р о х	в О р о	в О р а г	в О р а х

Для исследований произношения и для обучающей выборки при распознавании использовались записи заседаний Верховной Рады Украины продолжительностью 99 тыс. секунд, в которых встретилось более 210 тыс. слов. Всего было записано 208 дикторов. Дикторов с продолжительностью записи более 300 секунд оказалось 87 (запись длительно-стью менее 300 секунд является слишком короткой для объективных выводов в процес-се распознавания) [8].

Для проверки правильности транскрипций и для контрольной выборки при распознавании ис-пользовались записи продолжительностью 30 тыс. секунд, в которых встретилось более 68 тыс. слов. Эти записи были сделаны в отличные от обучающей выборки дни. Всего использовались записи 118 дикторов. Дикторов с продолжительностью записи более 300 секунд оказалось 37.

Словарь для генерирования транскрипций был создан из текстов стенограмм заседаний Верховной Рады Украины. С официального сайта Верховной Рады были загружены сте-нограммы заседаний, начиная с 1991 г. Все тексты стенограмм (более 15 млн. слов) были модифицированы для того, чтобы устранить лишнюю информацию (например, со-общения об аплодисментах или возгласы), записать числа словами, а также отделить украиноязычный текст от русскоязычного. В результате был получен словарь примерно из 150 тыс. слов, для которого вычислена частота употребления каждого слова. Был составлен словарь на 15 тыс. элементов из наиболее частотных слов (50 и более слово-употреблений), который и стал основой для анализа.

Чтобы выяснить, насколько влияет индивидуализация транскрипции на надёжность распоз-навания, был проведён ряд экспериментов. Для этого сделана оценка параметров аку-стической и лингвистической моделей [8] с использованием инструментария НТК [2]. Результаты экспериментов распознавания для некоторых транскрипций (улучшение / ухудшение результатов распознавания для конкретного диктора по сравнению с ре-зультатами распознавания для этого же диктора с использованием литературного тран-скрипции) приведены в таблице 3.

Таблица 3

Образцы результатов распознавания речи дикторов с использованием индивидуализированных транскрипций

Словарь Диктор	end_vow_rus	end_vow	end_cons_2	end_cons_1	a	duples
lit	0,16	0,14	-0,01	-0,01	-2,6	0,26
kir	1,37	1,82	0	-1,06	-3,3	0,91
dan	0,71	1,16	-0,09	-0,44	-5,4	0,18

После обработки результатов экспериментов каждому из дикторов приписываются те правила транскрибирования речи, которые повышают надёжность распознавания. Так формируется набор правил для генерирования индивидуализированных транскрипций того или иного диктора.

Как оказалось, часть правил транскрибирования может быть использована для большинства говорящих, например, редукция окончаний вследствие быстрого темпа и эмоциональности речи. Распространённое явление — оглушение, вызванное влиянием русского языка. Значительное количество дикторов избегает произношения удлинённых и удвоенных звуков. Часто встречается слабое произношение ударных гласных (ударение исчезает не только в односложных, двусложных словах, но и в трёхсложных). Другие правила могут быть использованы только для речи одного-двух дикторов.

Генерирование индивидуализированных транскрипций для распознавания речи позволило улучшить надёжность. В дальнейшем планируется увеличить количество индивидуализированных транскрипций за счёт детального анализа речи дикторов, а также автоматически относить дикторов к той или иной группе, учитывая все особенности их произношения.

Общий словарь транскрипций для спонтанной речи был создан для распознавания речи дикторов акустического корпуса украинского эфирного вещания (АКУЕМ) [12]. Он сочетает в себе как правила литературного произношения, так и индивидуальные черты произношения отдельных дикторов, а также очерчивает специфику спонтанной украинской речи (например, задаёт ударные и безударные формы для всех односложных слов и т.д.).

Для исследований произношения и для обучающей выборки при распознавании использовались аудиозаписи спонтанной речи из АКУЕМ продолжительностью около 190 тыс. секунд (более 2000 дикторов). Особое внимание обращалось на речь дикторов с продолжительностью записи более 600 секунд (около 150 дикторов).

Словарь для генерирования транскрипций был создан из текстов стенограмм АКУЕМ (около 50 тыс. слов).

Для проверки правильности транскрипций (и для контрольной выборки при распознавании) использовались записи продолжительностью 13,5 тыс. секунд, в которых встретилось 29 500 слов. Всего использовались записи 34 дикторов со средним и быстрым темпом речи.

Возможность генерировать сразу несколько вариантов транскрипции одного и того же слова позволяет продемонстрировать в словаре вариативность произношения наиболее частотных украинских слов, редуцирование и растяжение слов во время быстрого темпа речи, нечёткое произношение и подобные явления наряду с литературным вариантом произношения. Также система транскрибирования позволяет генерировать транскрипции для та-

ких специфических подсловарей, как словарь суржика, социальных и территориальных диалектов, аббревиатур и др.

Введение нескольких способов произнесения слов в словаре в целом улучшает надёжность распознавания спонтанной речи [11].

В дальнейшем планируется сбалансировать набор правил, которые порождают варианты произношения слов в потоке спонтанной речи, а также использовать индивидуализированные словари (для 18 групп дикторов) при создании единого словаря для распознавания спонтанной украинской речи.

## Выводы

Предложенная модель позволяет довольно компактно описать закономерности преобразований между графемами и фонемами в удобном для эксперта виде с учётом неоднозначности преобразований. Разработанная реализация модели также пригодна для расшивки чисел, символов и сокращений.

Введение в модель произвольного количества уровней транскрибирования помогает разложить правила, заданные экспертом, на более простые. Это открывает путь к применению предложенного метода для транскрибирования языков, в которых традиции написания не соответствуют произношению слов. В спонтанной украинской речи наблюдается подобное явление, но в обратном направлении: уход реальной живой речи от написания (и, соответственно, от правил канонического литературного произношения). В системе распознавания спонтанной речи нужно найти баланс между детальностью фонемной транскрипции, топологией акустических моделей фонем и количеством параметров, которые уточняют эти модели.

Результатом многозначного решения являются последовательности символов, но при этом не указывается, какие из них более вероятны. Оценка соответствующих вероятностей возможна на основе результатов распознавания, что является предметом дальнейших исследований.

## Литература

1. Винцюк Т.К. Анализ, распознавание и смысловая интерпретация речевых сигналов. Киев: Наукова думка, 1987.
2. Young S.J. et al. HTK Book, version 3.1, Cambridge University, 2002.
3. Lee, T. Kawahara and K. Shikano: Julius — an open source real-time large vocabulary recognition engine. In Proc. European Conference on Speech Communication and Technology (EUROSPEECH), 2001. P. 1691–1694.
4. Vintsiuk T., Sazhok M. Multi-Level Multi-Decision Models for ASR // Proceedings of the 10th Int. Conference on Speech and Computer — SpeCom'2005, Patras, 2005. P. 69–76.
5. Bisani M., Ney H. Joint-sequence models for grapheme-to-phoneme conversion // Journal Speech Communication, 50: 434–451, Elsevier, 2008.
6. Винцюк Т., Людовик Т., Сажок М., Селюх Р. Автоматичний озвучувач українських текстів на основі фонемно-трифонної моделі з використанням природного мовного сигналу // Праці 6-ї Всеукраїнської міжнародної конференції «Оброблення сигналів і зображень та розпізнавання образів» — УкрОбраз'2002, Київ, 2002.
7. Сучасна українська літературна мова. Фонетика: Навч. посібник для студентів-філологів. К.: Видавничо-поліграфічний центр «Київський університет», 2002. С. 60.
8. Пилипенко В.В., Робейко В.В. Автоматизированный стенограф украинской речи // Искусственный интеллект. № 4. 2008. С. 768–775.
9. Робейко В.В. Генерування індивідуалізованих транскрипцій для злитого мовлення // Мовні і концептуальні картини світу. Вип. 26. Ч. 3. Київ, 2009. С. 38–42.



10. *Робейко В.В.* Графічна омонімія як специфічна проблема синтезу мовлення за текстом. // Українське мовознавство. Вип. 39/1. Київ, 2009. С. 429–433.
11. *Людовик Т.В., Пилипенко В.В., Робейко В.В.* Автоматическое распознавание спонтанной украинской речи (на материале корпуса украинской эфирной речи). // Компьютерная лингвистика и интеллектуальные технологии: По материалам Международной конференции «Диалог». Вып. 10 (17). М.: РГГУ, 2011. С. 478–488.
12. *Васильєва Н.Б., В.В. Пилипенко, О.М. Радучький, В.В. Робейко, М.М. Сажок.* Створення акустичного корпусу українського ефірного мовлення. // Обробка сигналів і зображень та розпізнавання образів: Десята Всеукраїнська міжнародна конференція. Київ, 2010. С. 55–58.

### **Сведения об авторах**

***Робейко Валентина Васильевна —***

*научный сотрудник отдела распознавания и синтеза звуковых образов Международного научно-учебного центра информационных технологий и систем, Киев, Украина. valya.robeiko@gmail.com*

***Сажок Николай Николаевич —***

*кандидат технических наук, старший научный сотрудник отдела распознавания и синтеза звуковых образов Международного научно-учебного центра информационных технологий и систем, Киев, Украина. sazhok@gmail.com*