

ФОНЕТИКО-АКУСТИЧЕСКАЯ БАЗА ДАННЫХ ДЛЯ МНОГОЯЗЫЧНОГО СИНТЕЗА РЕЧИ ПО ТЕКСТУ НА СЛАВЯНСКИХ ЯЗЫКАХ¹

PHONETIC-ACOUSTICAL DATABASE FOR MULTILANGUAGE SLAVONIC TEXT-TO-SPEECH SYNTHESIS

Б. М. Лобанов (lobanov@newman.bas-net.by)

Л. И. Цирульник (liliya_tsirulnik@ssrlab.com)

Объединённый институт проблем информатики НАН Беларуси, Минск, Беларусь

Б. Пьорковска (boncia@wp.pl)

Я. Рафалко (jrafalko@wp.pl)

Э. Шпилевский (edwshp@hotmail.com)

Институт информатики Белостокского университета, Белосток, Польша

Проводится анализ особенностей фонетических систем белорусского, польского и русского языков, на основе которого предлагается общий подход к созданию единой фонетико-акустической БД для многоязычного синтеза речи по тексту. Описываются принципы создания и обработки текстовых и речевых корпусов для каждого из языков.

Введение

В настоящее время для некоторых славянских языков, таких как русский, чешский, польский, украинский, уже существует практически используемые или экспериментальные образцы синтезаторов речи по тексту (СРТ) [1]. Не имеется, однако, никаких сведений о создании многоязычных СРТ для славянских языков, а также СРТ для белорусского языка. Данная работа является продолжением проводимых ранее исследований, базирующихся на аллофонно-волновом представлении речевого сигнала [2] и направленных на разработку многоголосых [3] и многоязычных [4] СРТ. Синтез речи по тексту на славянских языках - белорусском, польском, русском - предполагает создание фонетико-акустической базы данных, построенной на единых принципах, отражающих внутри- и межъязыковую специфику фонетических систем и позиционно-комбинаторных явлений, порождающих аллофонию речи. Для этого необходимо определить принципы создания и обработки текстовых и речевых корпусов для каждого из языков и особенности формирования на их основе БД аллофонов и мультифонов. Решению именно этих вопросов посвящена данная работа.

1. Особенности фонетических систем белорусского, польского и русского языков

Фонетические системы языков, относящихся к группе славянских, имеют между собой значительное сходство, однако каждый из них обладает также специфическими особенностями, иногда значительными. Исследуемые фонетические системы белорусского, польского и русского языков являются относительно близкими, особенно русского и белорусского. В белорусском языке насчитывается 41 фонема, из них 6 гласных и 35 согласных, а в русском всего - 42, гласных - 6 и согласных - 36. Польский язык фонетически более разнообразен. В нём насчитывается 51 фонема, из них 8 гласных и 43 согласных. В таблице 1 представлена обобщённая информация о фонемном составе 3-х языков и об их различии по способу и месту образования. В каждой ячейке таблицы представлены имена фонем, характеризующихся определённым способом и местом образования, для белорусского, польского и русского языков порядке «сверху - вниз». Для обозначения фонем используются традиционные для каждого языка буквы алфавита.

В таблице 1 затемнены ячейки, фонетическое качество звуков которых имеет практически полное сходство для каждого из языков. Как видно из таблицы, количество таких ячеек в процентном отношении ко всем используемым ячейкам довольно значительно - 66%. Отличительные особенности фонетических систем белорусского и русского языков заключаются в следующем. В белорусском языке отсутствуют следующие фонемы:

- мягкие согласные *Т, Д, Ш, Ч, Р;*
- мягкая и твёрдая *Г.*

¹ Работа выполнена при поддержке европейского фонда INTAS в рамках проекта «Разработка многоголосовой и многоязыковой системы синтеза и распознавания речи (языки: белорусский, польский, русский)» в соответствии с грантом INTAS № 04-77-7404.

В белорусском языке имеется ряд специфических фонем, отсутствующих в русском:

- плавная *Ў*;
- мягкая *Ц* и твёрдая *Ч*;
- мягкая аффриката *Дз* и твёрдая *Дж*;
- мягкая и твёрдая щелевая *Гх*.

Способ образования		Согласные															
		Глухие			Звонкие			Сонорные									
		Взрывные	Аффрикаты	Щелевые	Взрывные	Аффрикаты	Щелевые	Дрожжацие	Носовые	Боковые	Плавные	Гласные	Передняя	Высокая	Огубленная	Назальная	
Задне-язычные	Мягкие	к', k', к'	~	х', h', х'	г', g', г'	~	жх', ~ ~	~	~	~	~	й j й	у и у	0	1	1	0
	Твёрдые	к k к	~	х h х	г g г	~	жх ~ ~	~	~	~	~	~	о о о	0	0	1	0
Средне-язычные	Мягкие	~	~	~	~	~	~	~	~	~	~	а а а	0	0	0	0	
	Твёрдые	~	ч cz ~	ш sz ш	~	дж dż ~	ж ż ж	р r р	~	~	~	э е э	1	0	0	0	
Передне-язычные	Мягкие	~	ц', c', т'	с', s', с'	д', d', д'	дз', ~ ~	з', z', з'	~	н', n', н'	л', l', л'	~	ы у ы	0	1	0	0	
	Твёрдые	т t т	ц c ц	с s с	д d д	дз dz ~	з z з	~	н n н	л l л	~	и i и	1	1	0	0	
Губные	Мягкие	п', p', п'	~	ф', f', ф'	б', b', б'	~	в', w', в'	~	м', m', м'	~	~	~ q ~	0	0	1	1	
	Твёрдые	п p п	~	ф f ф	б b б	~	в w в	~	м m м	~	ў ł ~	~ ε ~	1	0	0	1	

Таблица 1. Фонетические системы белорусского, польского и русского языков

Сравнивая фонетическую систему польского языка с русским, отметим некоторые её особенности. В польском языке присутствуют все фонемы, характерные для русского языка, однако произношение мягких фонем *Ш* и *Ч* отличается от польских мягких *Ś* и *Ć*, артикуляторный уклад которых промежуточный между мягкими русскими *С*, *Ш* и *Ц*, *Ч* соответственно. Кроме того, в польском языке имеется ряд специфических фонем, отсутствующих в русском:

- плавная *Ł*;
- мягкие *С*, *Ć* и твёрдая *Сz*;
- мягкая аффриката *Dź* и твёрдые *Dż* и *Dz*;
- назализованные гласные *А* и *Ę*.

Если сравнить фонетические системы всех рассматриваемых языков, а также каждую из пар языков, подсчитывая количество совпадений в ячейках таблицы 1, то получим следующие значения в процентах к общему количеству используемых ими ячеек:

- «русский – белорусский – польский» - 66%
- «русский – белорусский» - 71%
- «русский – польский» - 78%
- «польский – белорусский» - 69%.

Как это ни удивительно на первый взгляд, но белорусский язык по фонетическому составу отличается почти в равной степени как от польского, так и от русского. Сказанное, конечно, не учитывает статистику

употребления тех или иных фонем в различных языках. Так, хорошо известно, что схожие по звучанию русские и польские фонемы /t'/, /d'/, /s'/, /z'/, /l/, употребляемые в русском языке очень часто, в польском встречаются гораздо реже. В близких по звучанию словах вместо них используются, соответственно, специфические польские фонемы - /ć/, /dź/, /ś/, /ź/, /ł/.

2. Мини- и макси-наборы аллофонов для синтеза белорусской, польской и русской речи

Как известно, в речевом потоке фонемы реализуются в виде аллофонов, или иначе, в виде позиционных и комбинаторных оттенков фонем. Позиционный фактор учитывает позицию данной фонемы относительно словесного, акцентно-группового, синтагматического и фразового ударения. Комбинаторный фактор учитывает ближайшее фонемное окружение. В общем случае невозможно дать точную оценку количества аллофонов, т.к. она напрямую зависит от степени детализации учёта влияния позиционных и комбинаторных факторов. Однако качество синтезированной речи напрямую зависит от степени детализации. Стремление к большей детализации может привести к огромному количеству аллофонов (несколько сот тысяч), что делает задачу создания БД аллофонов неразрешимой. Опыт создания русскоязычных СРТ [2] показал, что синтезированная речь достаточно высокого качества может быть достигнута при некоторых определённых условиях генерации позиционных и комбинаторных аллофонов. Были исследованы 2 типа аллофонных наборов: так называемые макси- и мини-наборы.

При использовании макси-набора аллофонов для синтеза русской речи создаются следующие позиционные аллофоны гласных: ударный - (0), частично ударный - (1), первый предупредительный - (2), не первый предупредительный - (3), заударный - (4). Всего 5 позиций. С учётом левого контекста создаются следующие комбинаторные аллофоны гласных: после синтагматической паузы - (0), после большинства губных - (1), переднеязычных - (2) и заднеязычных - (3) твёрдых, после /Л/- (4), /Р/- (5), /М/- (6), /Н/- (7), после большинства мягких - (8), после /Р'/ - (9), /М'/ - (10), /Н'/ - (11), после гласных /У/- (12), /О/- (13), /А/- (14), /Э/- (15), /Ы/- (16), /И/- (17). Всего 18 левых контекстов. Для учёта правого контекста создаются следующие комбинаторные аллофоны гласных: перед синтагматической паузой - (0), перед переднеязычными и заднеязычными твёрдыми согласными и гласными /У/, /О/, /А/, /Э/, /Ы/- (1), перед губными твёрдыми - (2), перед губными мягкими - (3) перед не губными мягкими согласными и гласным /И/- (4). Всего 5 правых контекстов. Итого, для 6-ти гласных создаются $N_v = 5 \cdot 18 \cdot 5 \cdot 6 = 2700$ аллофонов.

Позиционные аллофоны согласных для макси-набора включают два положения: в ударном слоге - (0) и в безударном слоге - (1). Левый контекст согласных включает следующие группы: после паузы - (0), после глухих - (1) и звонких - (2) согласных, после гласных - (3). Правый контекст: перед паузой - (0), перед глухими - (1) и звонкими - (2) согласными, перед безударными - (3) и ударными - (4) гласными. Итого, для всех 36-ти согласных создаются $N_c = 2 \cdot 4 \cdot 5 \cdot 36 = 1440$ аллофонов. Всего создаётся: $2700 + 1440 = 4140$ аллофонов русской речи.

При использовании мини-набора для синтеза русской речи создаётся только 2 типа позиционных аллофонов гласных: ударный - (0), безударный - (1). С учётом левого контекста создаются следующие комбинаторные аллофоны гласных: после синтагматической паузы - (0), после твёрдых губных - (1), передне- и среднеязычных - (2), после твёрдых заднеязычных и гласных - (3) и после мягких - (4). Всего 5 левых контекстов. С учётом правого контекста создаются следующие комбинаторные аллофоны гласных: перед синтагматической паузой - (0), перед переднеязычными и заднеязычными твёрдыми согласными и гласными /У/, /О/, /А/, /Э/, /Ы/- (1), перед губными согласными - (2), перед мягкими согласными и гласной /И/- (3). Итого, для 6-ти гласных создаются $N_v = 2 \cdot 5 \cdot 4 \cdot 6 = 240$ аллофонов. Аллофоны согласных создаются только с учётом правого контекста: перед паузой - (0), перед глухими - (1) и звонкими - (2) согласными, перед безударными - (3) и ударными - (4) гласными. Итого, для всех 36-ти согласных создаются $N_c = 5 \cdot 36 = 180$ аллофонов. Всего создаётся: $240 + 180 = 420$ аллофонов русской речи.

Полученные оценки количества аллофонов, рассчитанные теоретически, являются сильно завышенными из-за того, что, во-первых, очень многие позиционные и комбинаторные ситуации вообще не встречаются в речи и, во-вторых, для многих аллофонов акустические различия настолько невелики, что ими можно пренебречь. В результате, как показывает практика, используемое количество аллофонов в макси-наборе оказывается более чем в 2 раза, а в мини-наборе в 1,5 раза меньшим.

Результаты подсчёта теоретического и практически используемого количества аллофонов для каждого из 3-х языков приведены в таблице 2.

Язык	Белорусский				Польский				Русский			
	Теоретическое		Практич. используемое		Теоретическое		Практич. используемое		Теоретическое		Практич. используемое	
Тип набора	Макси	Мини	Макси	Мини	Макси	Мини	Макси	Мини	Макси	Мини	Макси	Мини
Гласных	2520	240	1480	170	3600	320	2050	224	2700	240	1550	175
Согласных	720	180	217	76	860	215	279	113	720	180	209	81
Всего	3240	420	1697	246	4460	535	2329	337	3420	420	1759	256

Таблица 2. Количество аллофонов

Для обозначения имён аллофонов при синтезе речи используется имена соответствующих фонем (латинские буквы), а также 3 цифровых индекса. При этом 1-й индекс обозначает позицию фонемы относительно полноударного гласного, 2-й индекс – левый контекст, а 3-й индекс – правый контекст. В таблице 3 приведены единые обозначения аллофонов, используемых для синтеза речи на трёх славянских языках.

Губные согласные					Переднеязычные согласные					Среднеязычные согласные					Заднеязычные согласные и гласные				
№	Бел	Пол	Рус	Имя	№	Бел	Пол	Рус	Имя	№	Бел	Пол	Рус	Имя	№	Бел	Пол	Рус	Имя
1	<i>n</i>	<i>p</i>	<i>n</i>	<i>P_{ijk}</i>	16	<i>m</i>	<i>t</i>	<i>m</i>	<i>T_{ijk}</i>	31	<i>ч</i>	<i>cz</i>	-	<i>Ch_{ijk}</i>	46	<i>к</i>	<i>k</i>	<i>к</i>	<i>K_{ijk}</i>
2	<i>ф</i>	<i>f</i>	<i>ф</i>	<i>F_{ijk}</i>	17	<i>ц</i>	<i>c</i>	<i>ц</i>	<i>C_{ijk}</i>	32	<i>ш</i>	<i>sz</i>	<i>ш</i>	<i>Sh_{ijk}</i>	47	<i>х</i>	<i>h</i>	<i>х</i>	<i>H_{ijk}</i>
3	<i>б</i>	<i>b</i>	<i>б</i>	<i>B_{ijk}</i>	18	<i>с</i>	<i>s</i>	<i>с</i>	<i>S_{ijk}</i>	33	<i>дж</i>	<i>dž</i>	-	<i>Dh_{ijk}</i>	48	<i>гх</i>	<i>g</i>	<i>г</i>	<i>G_{ijk}</i>
4	<i>в</i>	<i>w</i>	<i>в</i>	<i>V_{ijk}</i>	19	<i>д</i>	<i>d</i>	<i>д</i>	<i>D_{ijk}</i>	34	<i>ж</i>	<i>ž</i>	<i>ж</i>	<i>Zh_{ijk}</i>	49	<i>к'</i>	<i>k'</i>	<i>к'</i>	<i>K'_{ijk}</i>
5	<i>м</i>	<i>m</i>	<i>м</i>	<i>M_{ijk}</i>	20	-	<i>dz</i>	-	<i>Dz_{ijk}</i>	35	<i>р</i>	<i>r</i>	<i>р</i>	<i>R_{ijk}</i>	50	<i>х'</i>	<i>h'</i>	<i>х'</i>	<i>H'_{ijk}</i>
6	<i>й</i>	<i>l</i>	-	<i>W_{ijk}</i>	21	<i>з</i>	<i>z</i>	<i>з</i>	<i>Z_{ijk}</i>	36	-	<i>é</i>	<i>ч'</i>	<i>Ch'_{ijk}</i>	51	<i>гх'</i>	<i>g'</i>	<i>г'</i>	<i>G'_{ijk}</i>
7	<i>н'</i>	<i>p'</i>	<i>н'</i>	<i>P'_{ijk}</i>	22	<i>н</i>	<i>n</i>	<i>н</i>	<i>N_{ijk}</i>	37	-	<i>ś</i>	<i>ш'</i>	<i>Sh'_{ijk}</i>	52	<i>й</i>	<i>j</i>	<i>й</i>	<i>J_{ijk}</i>
8	<i>ф'</i>	<i>f'</i>	<i>ф'</i>	<i>F'_{ijk}</i>	23	<i>л</i>	<i>l</i>	<i>л</i>	<i>L_{ijk}</i>	38	-	<i>dž</i>	-	<i>Dh'_{ijk}</i>	53	<i>у</i>	<i>u</i>	<i>у</i>	<i>U_{ijk}</i>
9	<i>б'</i>	<i>b'</i>	<i>б'</i>	<i>B'_{ijk}</i>	24	-	<i>t'</i>	<i>м'</i>	<i>T'_{ijk}</i>	39	-	<i>ź</i>	-	<i>Zh'_{ijk}</i>	54	<i>о</i>	<i>o</i>	<i>о</i>	<i>O_{ijk}</i>
10	<i>в'</i>	<i>w'</i>	<i>в'</i>	<i>V'_{ijk}</i>	25	<i>ц'</i>	<i>с'</i>	-	<i>C'_{ijk}</i>	40	-	<i>r'</i>	<i>р'</i>	<i>R'_{ijk}</i>	55	<i>а</i>	<i>a</i>	<i>а</i>	<i>A_{ijk}</i>
11	<i>м'</i>	<i>m'</i>	<i>м'</i>	<i>M'_{ijk}</i>	26	<i>с'</i>	<i>s'</i>	<i>с'</i>	<i>S'_{ijk}</i>	41	-	-	-	-	56	<i>э</i>	<i>e</i>	<i>э</i>	<i>E_{ijk}</i>
12	-	-	-	-	27	<i>дз'</i>	<i>d'</i>	<i>д'</i>	<i>D'_{ijk}</i>	42	-	-	-	-	57	<i>ы</i>	<i>y</i>	<i>ы</i>	<i>Y_{ijk}</i>
13	-	-	-	-	28	<i>з'</i>	<i>z'</i>	<i>з'</i>	<i>Z'_{ijk}</i>	43	-	-	-	-	58	<i>и</i>	<i>i</i>	<i>и</i>	<i>I_{ijk}</i>
14	-	-	-	-	29	<i>н'</i>	<i>n'</i>	<i>н'</i>	<i>N'_{ijk}</i>	44	-	-	-	-	59	-	<i>q</i>	-	<i>O'_{ijk}</i>
15	-	-	-	-	30	<i>л'</i>	<i>l'</i>	<i>л'</i>	<i>L'_{ijk}</i>	45	-	-	-	-	60	-	<i>ę</i>	-	<i>E'_{ijk}</i>

Таблица 3. Перечень имён аллофонов, используемых для синтеза речи на белорусском, польском и русском языках

3. Текстовые и речевые корпуса для создания БД аллофонов

Процесс создания БД аллофонов включает следующие этапы:

- формирование представительного текстового корпуса (набора текстов) и соответствующих этим текстам фонограмм речи (речевой базы) диктора;
- обработка созданной речевой базы, включающая фонемную сегментацию речевого сигнала, аллофонную маркировку сегментов и сохранение полученного набора в аллофонно-волновой БД.

Текстовые корпуса созданы на основе специально подобранного набора слов в количестве, равном числу используемых в каждом из языков аллофонов. Каждое из слов отбиралось исходя из критерия наилучшей репрезентации данного аллофона в речи диктора. Речевые корпуса, соответствующие текстовым корпусам, создавались в студийных условиях специально проинструктированными профессиональными дикторами. Ниже, в таблицах 4 и 5, приведены фрагменты списка слов для создания («нарезки») БД для макси-набора аллофонов согласных и гласных польской речи, в таблицах 6 и 7 – для создания мини-набора слов для 3-х языков.

Правый контекст (третий индекс)	Пауза (0)	Глухой согласный (1)	Звонкий согласный (2)	Безударный гласный (3)	Ударный гласный (4)
Левый контекст (второй индекс)					
Пауза (0)	–	<i>Sztuka</i>	<i>Szmal</i>	<i>Szanować</i>	<i>Szybkość</i>

		(Sh ₀₀₁)	(Sh ₀₀₂)	(Sh ₁₀₃)	(Sh ₁₀₄)
Глухой согласный (1)	Wieprz (Sh ₀₁₀)	Kształtowanie (Sh ₁₁₁)	–	Przełomu (Sh ₁₁₃)	Przez (Sh ₀₁₄)
Звонкий согласный (2)	–	–	Spójrzmy (Sh ₁₂₂)	Rozszerzalność (Sh ₁₂₃)	Sfałszować (Sh ₀₂₄)
Гласный (3)	Również (Sh ₁₃₀)	Przemieszczenia (Sh ₀₃₁)	Wprzeszłość (Sh ₀₃₂)	Nasze (Sh ₁₃₃)	Naszego (Sh ₀₃₄)

Таблица 4. Фрагмент списка слов для «нарезки» аллофонов согласной /Sh/ польской речи (в скобках после каждого слова указан соответствующий аллофон)

Правый контекст (третий индекс)	Пауза (0)	Не губные твёрдые (1)	Губные твёрдые (2)	Не губные мягкие (3)	Губные мягкие (4)
Левый контекст (второй индекс)					
Пауза (0)	A (A ₀₀₀)	Adres (A ₀₀₁)	Amper (A ₀₀₂)	Ani (A ₀₀₃)	–
Губные твёрдые (1)	Ba (A ₀₀₁₀)	Najbardziej (A ₀₀₁₁)	Zaspawać (A ₀₀₁₂)	Projektowania (A ₀₀₁₃)	Pawie (A ₀₀₁₄)
Передне- и среднеязычные твёрдые (2)	Ta (A ₀₀₂₀)	Rozszerzalność (A ₀₀₂₁)	Samym (A ₀₀₂₂)	Przetwarzania (A ₀₀₂₃)	Zaletami (A ₀₀₂₄)
Нёбные твёрдые (3)	Ha (A ₀₀₃₀)	Bogaty (A ₀₀₃₁)	Gapa (A ₀₀₃₂)	Wspomagania (A ₀₀₃₃)	Zagapić (A ₀₀₃₄)
M (4)	Ma (A ₀₀₄₀)	Matlab (A ₀₀₄₁)	Mapa (A ₀₀₄₂)	Mazia (A ₀₀₄₃)	Zamawiać (A ₀₀₄₄)
N (5)	Na (A ₀₀₅₀)	Znacznie (A ₀₀₅₁)	Sygnalów (A ₀₀₅₂)	Naciąg (A ₀₀₅₃)	Nawiać (A ₀₀₅₄)
L (6)	La (A ₀₀₆₀)	Popularnym (A ₀₀₆₁)	Matłaba (A ₀₀₆₂)	Kolanie (A ₀₀₆₃)	Kalafior (A ₀₀₆₄)
R (7)	Ra (A ₀₀₇₀)	Pracy (A ₀₀₇₁)	Rama (A ₀₀₇₂)	Wyrzucić (A ₀₀₇₃)	Procedurami (A ₀₀₇₄)
Ł (8)	Pchła (A ₀₀₈₀)	Układów (A ₀₀₈₁)	Oklamać (A ₀₀₈₂)	Odlążyć (A ₀₀₈₃)	Odlawia (A ₀₀₈₄)
J (9)	Ja (A ₀₀₉₀)	Jakość (A ₀₀₉₁)	rozwijają się (A ₀₀₉₂)	Objaśnia (A ₀₀₉₃)	Zjawia (A ₀₀₉₄)
Не губные мягкие (10)	Pnia (A ₀₁₀₀)	Posiada (A ₀₁₀₁)	Rozdzigawa (A ₀₁₀₂)	Niania (A ₀₁₀₃)	Narzędziami (A ₀₁₀₄)
Губные мягкие (11)	Lwia (A ₀₁₁₀)	Świata (A ₀₁₁₁)	Biawar (A ₀₁₁₂)	Oglupianie (A ₀₁₁₃)	Kopiami (A ₀₁₁₄)

Таблица 5. Фрагмент списка слов для «нарезки» аллофонов гласной /A/ польской речи (в скобках после каждого слова указан соответствующий аллофон; второй индекс обозначен двумя цифрами)

Правый контекст (индекс аллофона)	Пауза (0)	Глухой согласный (1)	Звонкий согласный (2)	Безударный гласный (3)	Ударный гласный (4)
Язык					
Белорусский	Цяжар	Дзірка	Скарба	Сябраваць	Урад
Польский	Akr	Krtań	Grdyka	Środowisko	Program
Русский	Спор	Марка	Кордон	Караван	Парад

Таблица 6. Мини-набор аллофонов согласной /R/ для 3-х языков

Третий индекс, правый контекст (языки: белорусский, польский, русский)	0	1	2	3
пауза	пауза	n, ф, б, в, м, ў	t, ц, с, д, з, н, л, ч, ш, дж, ж, р, к, х, гх, у, о, а, э, ы	k', x', gх', й, ц', с', дз', з', н', л', n', ф', б', в', м', i
пауза	пауза	p, f, b, w, m, t	t, c, s, d, dz, z, n, l, cz, sz, dź, ź, r, k, h, g, u, o, a, e, e, y	k', h', g', j, ć, ś, dź, ź, r', t', c', s', d', z', n', l', p', f', b', w', m', i
Второй индекс, левый контекст (языки: белорусский, польский, русский)	пауза	n, ф, б, в, м	t, ц, с, д, з, н, л, ш, ж, р, к, х, г, у, о, а, э, ы	k', x', z', й, ч', ш', p', m', c', d', z', n', л', n', ф', б', в', м', u

0	<i>пауза</i>	A000	<u>А</u>	A001	<u>Аўра</u>	A002	<u>Анджэй</u>	A003	<u>Альфа</u>
	<i>пауза</i>		<u>А</u>		<u>Ампер</u>		<u>Адрес</u>		<u>Апі</u>
	<i>пауза</i>		<u>А</u>		<u>Автор</u>		<u>Атом</u>		<u>Ася</u>
1	<i>n, ф, б, в, м, ў</i>	A010	<u>Барацьба</u>	A011	<u>Ва́бны</u>	A012	<u>Фарба</u>	A013	<u>Майстар</u>
	<i>p, f, b, w, m, ł</i>		<u>Ва</u>		<u>Zasprawać</u>		<u>Najbardziej</u>		<u>Равіе</u>
	<i>n, ф, б, в, м</i>		<u>Судьба</u>		<u>Баба</u>		<u>Вата</u>		<u>Батя</u>
2	<i>т, ц, с, д, з, н, л, ч, ш, дж, ж, р</i>	A020	<u>Кабала</u>	A021	<u>Згра́бны</u>	A022	<u>Цацка</u>	A023	<u>Талент</u>
	<i>t, c, s, d, dz, z, n, l, cz, sz, dź, ź, r</i>		<u>Та</u>		<u>Sąpym</u>		<u>Znacznie</u>		<u>Zaletami</u>
	<i>ш, ж, р, т, ц, с, д, з, н, л</i>		<u>Еда</u>		<u>Запад</u>		<u>Дата</u>		<u>Тяся</u>
3	<i>к, х, гх, у, о, а, э, ы</i>	A030	<u>Дачка</u>	A031	<u>Кава</u>	A032	<u>Казка</u>	A033	<u>Камень</u>
	<i>k, h, g, u, o, a, e, e, y</i>		<u>На</u>		<u>Гара</u>		<u>Vogaty</u>		<u>Zagarić</u>
	<i>к, х, г, у, о, а, э, ы</i>		<u>Нога</u>		<u>Гавкаць</u>		<u>Сказка</u>		<u>Галя</u>
4	<i>ц', с', дз', з', н', л', н', ф', б', в', м', к', х', гх', ў, і</i>	A040	<u>Мітусня</u>	A041	<u>Сябар</u>	A042	<u>Немаўляты</u>	A043	<u>Сядзеш</u>
	<i>t', c', s', d', z', n', l', é, ś, dź, ź, r', p', f', b', w', m', k', h', g', j, i</i>		<u>Рніг</u>		<u>Rozdziawa</u>		<u>Посіда</u>		<u>Коріамі</u>
	<i>т', с', д', з', н', л', ч', ш', р', н', ф', б', в', м', к', х', г', ў, и</i>		<u>Шутя</u>		<u>Тяпка</u>		<u>Тяга</u>		<u>Тянет</u>

Таблица 7. Мини-набор аллофонов ударной гласной /А/ для 3-х языков

4. Процедура создания БД звуковых волн аллофонов

Процедура обработки созданной речевой базы включает фонемную сегментацию речевого сигнала, аллофонную маркировку сегментов и сохранение полученного набора сегментов естественной речевой волны в аллофонно-волновой БД. Совершенно очевидно, что хотя использование для синтеза макси-набора обеспечит наивысшее качество речи, его создание «вручную» весьма затруднительно (порядка 2000 аллофонов!), если не невозможно. Создание «вручную» мини-набора (порядка 300 аллофонов) вполне реально. Мини-набор так же, как и макси-набор, обеспечивает синтез произвольного текста, хотя качество синтезированной речи при этом будет не столь высоким. Однако благодаря созданию мини-набора аллофонов становится возможным автоматизировать процесс «нарезки» макси-БД аллофонных волн, а при необходимости и более крупных единиц – мультифонов, реализующихся в виде последовательности аллофонов – диаллофонов, трифонов, слогов. Для автоматизации процесса создания БД аллофонных волн используется разработанная ранее технология клонирования персонального голоса и дикции [5,6].

Общая схема процедуры создания мини- и макси-БД аллофонных волн представлена на рис.1.

Заключение

Разработанные мини- и макси-наборы аллофонов для белорусского, польского и русского языков, а также созданные в соответствии с описанной технологией БД аллофонных волн для трёх языков используются в многоязычном и многоголосовом синтезаторе речи по тексту.

Кроме очевидного преимущества разработанной единой фонемно-аллофонной классификации – возможности создания многоязычного синтезатора – описанный подход позволяет также синтезировать речь с заданным акцентом, например, русскую речь с белорусским акцентом. Такое применение системы может понадобиться, в частности, при персонализированном синтезе речи по тексту для передачи индивидуальных фонетических особенностей дикции.

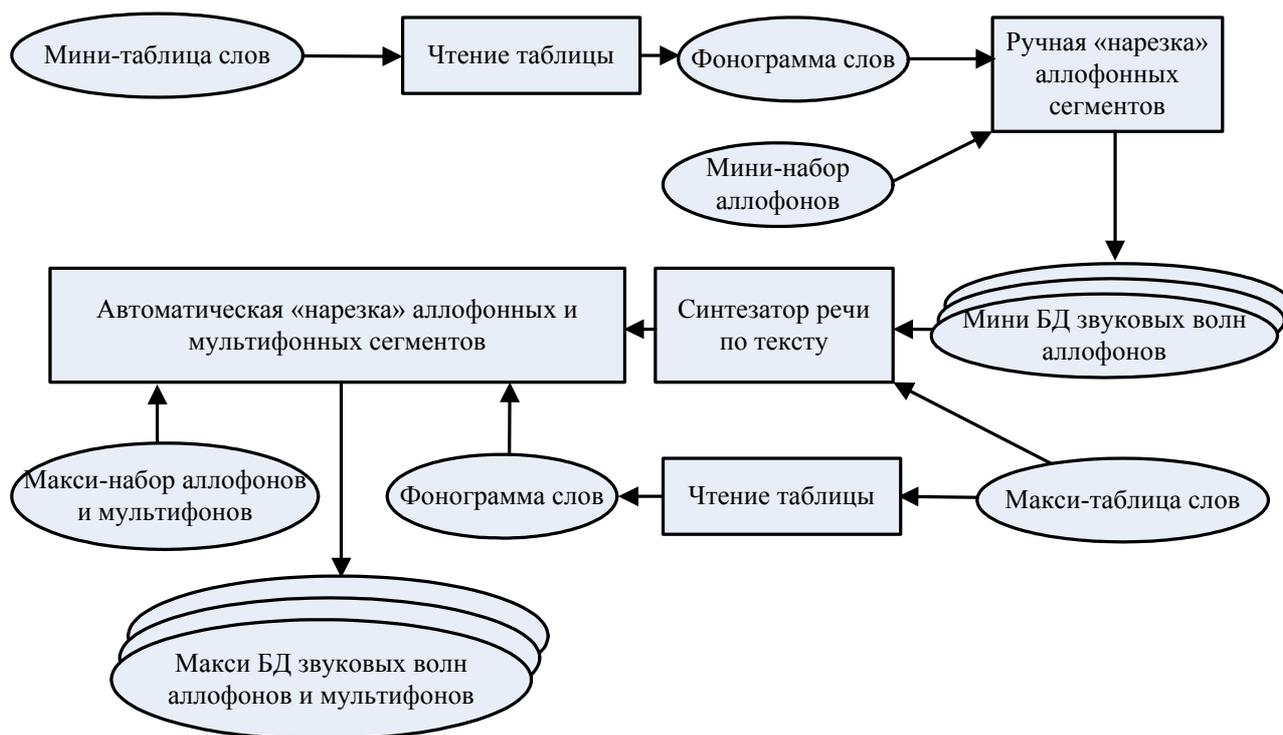


Рис. 1. Процедура создания мини- и макси-БД звуковых волн аллофонов

Список литературы

1. <http://www.speech.cs.cmu.edu/comp.speech/>.
2. Лобанов Б.М. Синтез речи по тексту // Четвёртая Международная летняя школа-семинар по искусственному интеллекту. Сб. науч. тр. Мн.:Изд. БГУ, 2000. С. 57-76.
3. Lobanov B.M., Tsurulnik L.I. Phonetic-Acoustical Problems of Personal Voice Cloning by TTS // Proc. of the International Conference «Speech and Computer» – SPECOM'2004, St.-Petersburg, 2004. P. 17 – 21.
4. Shpilewski E., Piurkowska B., Rafalko J., Lobanov B., Kiselov V., Tsurulnik. Polish TTS in Multi-Voice Slavonic Languages Speech Synthesis System. // Proc. of the International Conference «Speech and Computer» – SPECOM'2004, St.-Petersburg, 2004. P. 565 – 570.
5. Лобанов Б.М., Киселёв В.В. Автоматизация клонирования персонального голоса и дикции для систем синтеза речи по тексту // Международная конференция «Диалог-2003». Сб. науч. тр. М, 2003. С. 417-424.
6. Цирульник Л.И. Автоматизированная система клонирования фонетико-акустических характеристик речи // Информатика. № 1(9). Мн., 2006. С. 37-46.