

УДК 004.056.5

СИНТЕЗ РЕЧЕПОДОБНЫХ СИГНАЛОВ НА БЕЛОРУССКОМ ЯЗЫКЕ

Г.В. ДАВЫДОВ, В.А. ПОПОВ, А.В. ПОТАПОВИЧ, Е.Н. СЕЙТКУЛОВ*, И.В. САВЧЕНКО

*Белорусский государственный университет информатики и радиоэлектроники
П. Бровки, 6, Минск, 220013, Беларусь*

**Евразийский национальный университет им. Л.Н. Гумилева
Мирзояна, 2, Астана, Казахстан*

Поступила в редакцию 24 марта 2015

Проанализированы методы синтеза речеподобных сигналов по базе аллофонов. Обоснована актуальность разработки синтеза речеподобных сигналов на белорусском языке. Изложена методология формирования базы аллофонов на белорусском языке с учетом фонетических особенностей белорусского языка.

Ключевые слова: защита информации, маскирующие сигналы, речевая информация, речеподобные сигналы, база аллофонов, белорусский язык.

Введение

Речеподобные сигналы широко используются в телефонетрии для тестирования качества передачи речевой информации по линиям связи и в системах защиты речевой информации, в качестве одной из компонент комбинированных маскирующих речевых сигналов для защиты ее от утечки по акустическим каналам [1–5]. В соответствии с [1–3] для тестирования линий связи основные требования к речеподобным сигналам заключаются в наличии трех компонент: звонких вокализованных сигналов с имитацией свойств голоса; сигналов с постоянной спектральной плотностью мощности в частотной области; наличие паузы для обеспечения свойств, присущих сигналам с амплитудной модуляцией.

Речеподобные сигналы в последнее время широко применяются в качестве маскирующих сигналов для активных методов защиты речевой информации от утечки по акустическим каналам. При этом чаще всего используются комбинированные маскирующие сигналы, включающие «белый» шум и речеподобные сигналы. Соотношение между «белым» шумом и речеподобными сигналами устанавливается по уровням среднеквадратичного значения напряжения и лежит обычно в пределах 3–12 дБ [4].

Важным требованием к маскирующим сигналам является требование, чтобы они имели случайный характер, чтобы «белый» шум формировался за счет тепловых шумов полупроводниковых приборов или других видов физических шумов, а речеподобные сигналы формировались с использованием генератора случайных чисел, построенного на тепловых шумах. Эти требования обусловлены необходимостью исключения какой-либо возможности шумоочистки перехваченных акустических сигналов от маскирующих сигналов.

К речеподобным сигналам для систем защиты речевой информации дополнительно предъявляются требования, обусловленные характерными лингвистическими и фонетическими свойствами определенного языка и акустическими свойствами голоса определенного диктора. Лингвистические требования включают статистические данные о длине предложений, длине синтагм, длине слов, а также вероятности появления определенных фонем для определенного языка. Фонетические требования обусловлены необходимостью формирования речеподобных сигналов со спектральными характеристиками, не отличающимися от спектральных характеристик голоса диктора, речь которого необходимо защитить.

Методы синтеза речеподобных сигналов

Синтез речеподобных сигналов так же, как и синтез речи, может быть выполнен двумя основными методами. Первый метод – это синтез речи с использованием фонемного синтезатора, суть которого заключается в генерации фонем и дальнейшей компиляции из них слов и фраз.

Фонемный синтез содержит три основных этапа формирования речевого сигнала. На первом этапе выполняется перекодировка последовательности орфографических символов в последовательность кодов фонем; на втором рассчитывается набор акустических параметров формантных частот и полос частот, частоты основного тона и амплитуда вокализации. На третьем этапе формируется выходной речевой сигнал. Фонемный синтезатор не позволяет сформировать речевой сигнал с окраской и тембром, свойственным определенному диктору. Поэтому применение фонемного синтезатора для формирования речеподобных сигналов свойственных определенному диктору, не является эффективным с точки зрения защиты речевой информации маскирующими сигналами. Это обусловлено тем, что остаются различия в окраске и тембре речеподобных сигналов, сформированных фонемным синтезатором и информационным речевым сигналом определенного диктора.

Второй метод синтеза речеподобных сигналов – это компиляционный синтез, который основан на формировании речевого сигнала путем последовательного акустического воспроизведения единиц речевого сигнала, которые подготовлены заранее и сохраняются в памяти. Расстановка в определенной последовательности единиц речевого сигнала и их акустического воспроизведения являются основными этапами компиляционного синтеза речи. Метод является наиболее простым решением для достижения натуральности и естественности звучания речи [7]. К структурным единицам речи относятся аллофоны, дифоны, трифоны, полифоны, слоги, отдельные слова и словосочетания. Выбор структурной единицы речи для синтеза речи, с одной стороны, более просто вести по коротким сегментам с общим незначительным объемом памяти. Однако при этом имеет место большое количество переходов от одного фрагмента речи к другому, что может сказаться на качестве синтезируемой речи, если не применять сплайны. С другой стороны, при выборе в качестве структурных единиц речи более длинных по звучанию фрагментов, речь становится более естественной, однако необходимы при этом большие объемы памяти и большие базы структурных единиц речи, создание которых является трудоемким процессом. Поэтому для синтеза речеподобных сигналов предлагается использовать аллофоны в качестве структурной единицы речи, а по речевой базе аллофонов можно будет формировать речеподобные сигналы голосом определенного диктора.

Статистические характеристики белорусского языка

Синтез речеподобных сигналов на белорусском языке необходимо выполнять с учетом вероятностей длины предложений и длины слов в белорусском языке, а также вероятностей появления определенных аллофонов.

Распределение вероятностей длины предложений (числа слов в предложении) для белорусского языка является не определяющим параметром при синтезе речеподобных сигналов. Следует использовать при синтезе речеподобных сигналов длину синтагмы, на которые делится предложение (фраза) и количество фраз в фоноабзаце. Среднее число слов в предложении для белорусского языка составляет 10. Однако эти характеристики для каждого диктора могут быть свои. Поэтому следует использовать усредненные показатели. В табл. 1 представлены статистические характеристики белорусского языка.

Таблица 1. Статистические характеристики количества слов в синтагме для белорусского языка

Число слов в синтагме	1	2	3	4	5	6	7
Вероятность	0,05	0,10	0,30	0,25	0,15	0,10	0,05

Распределение вероятностей количества букв в слове для белорусского языка представлено в табл. 2.

Таблица 2. Распределение вероятностей количества букв в слове

Число букв в слове	1	2	3	4	5	6	7	8
Вероятность	0,103	0,126	0,103	0,108	0,130	0,128	0,115	0,077
Число букв в слове	9	10	11	12	13	14	15	
Вероятность	0,051	0,031	0,016	0,007	0,003	0,001	0,001	

Распределение вероятностей появления печатных знаков в тексте на белорусском языке представлено в табл. 3.

Таблица 3. Распределение вероятностей появления печатных знаков в тексте на белорусском языке

Аллофон	Вероятность	Аллофон	Вероятность	Аллофон	Вероятность
а	0,1642	м	0,0274	ч	0,0158
б	0,0211	н	0,0571	ш	0,0143
в	0,0287	о	0,0413	ы	0,0405
г	0,0198	п	0,0289	ь	0,0149
д	0,0339	р	0,0393	э	0,0076
е	0,0395	с	0,0404	ю	0,0055
ж	0,0085	т	0,0336	я	0,0376
з	0,0303	у	0,0338	ё	0,0065
й	0,0092	ф	0,0005	і	0,0508
к	0,0412	х	0,0125	ў	0,0257
л	0,0436	ц	0,0260		

Для формирования речеподобных последовательностей используется генератор псевдослучайных чисел, выполненный программно. По значениям, полученным от генератора псевдослучайных чисел, и с учетом распределения вероятностей длины предложений определяется длина первого предложения, т.е. число слов, входящих в состав предложения (или синтагмы). Далее с учетом значения следующего псевдослучайного числа определяется длина первого слова в предложении, т.е. число аллофонов в первом слове предложения. Это процесс повторяется, пока не будут получены значения длин слов для первого предложения. Далее по значениям последующего псевдослучайного числа и с учетом таблиц вероятностей появления аллофонов в русской речи находится соответствующий первый аллофон первого слова для первого предложения. После этого выполняется процесс нахождения всех аллофонов для первого предложения. Таким образом, формируется текст, который воспроизводится системой преобразования текста в речь по базе аллофонов и передается на звуковую карту компьютера.

При формировании текста для речеподобных последовательностей вводится ряд ограничений, таких как: не может быть в слове последовательно расположенных трех гласных и четырех согласных, не может быть слова, состоящего из двух и более только гласных и двух и более только согласных.

Кроме вероятностей появления аллофонов необходимо учитывать вероятности встречаемости звуков в начале слова. Вероятности встречаемости фонем и их сочетаний в начале слова не исследовались.

Расстановка ударений в словах выполняется с учетом номера слога и числа слогов в слове по правилу, представленному в табл. 4.

Таблица 4. Правило расстановки ударного слога в слове

Число слогов в слове	1	2	3	4	5	6	7
№ ударного слога	1	2	2	3	3	4	5

Распределение вероятностей встречаемости гласных ударных фонем в середине слова представлены в табл. 5.

Таблица 5. Распределение вероятностей встречаемости гласных ударных фонем в середине слова

Фонема	А	Е	Ё	И	О	У	Ы	Э	Ю	Я
Вероятность	0,325	0,196	0,013	0,143	0,146	0,070	0,061	0,012	0,006	0,028

Фонетические особенности белорусского языка

Основной особенностью белорусского языка, в отличие от русского, является его напевность, более продолжительное звучание гласных и согласных фонем, отсутствие напряженности при звукообразовании, сильное смягчение. В белорусском языке отсутствует мягкая и твердая фонема **Г**, вместо нее используется мягкая и твердая фонема **Гх**. В белорусском языке отсутствуют мягкие согласные фонемы **Д, Р, Т, Ш, Щ, Ч**. В отличие от русского языка, в белорусском языке имеются фонемы **Ў** (у короткое), мягкая **Ц**, твердая **Ч**. Отличительной особенностью белорусского языка является свободный, без напряжения, проход воздуха через артикуляционный аппарат при произношении гласных звуков. В фонетическом аспекте ударный слог характеризуется более продолжительным звучанием и более высоким уровнем создаваемого звукового давления по сравнению с таким же безударным слогом. Поэтому при формировании базы аллофонов необходимо использовать связный текст, из которого путем сегментации речи на аллофоны создается акустическая база.

Формирование базы аллофонов белорусского языка

Аллофоны необходимо классифицировать по фонетическим особенностям таким образом, чтобы учитывались фонетические особенности предыдущего и последующего аллофонов, т.е. его окружение. Это связано с тем, что звучание аллофонов определяется переходом формы речевого аппарата из одного положения в другое. Состояние речевого аппарата при произношении изменяется и это является динамическим процессом. Поэтому весьма важно при формировании базы аллофонов классифицировать их с учетом изменения формы речевого аппарата, что принято называть артикуляцией.

Аллофоны белорусского языка можно разделить на вокализованные (тоновые), шумовые и тоново-шумовые. Тоновые аллофоны это гласные **А, Е, О, У, Ы, І**. Они характеризуются открытым положением речевого аппарата при их произношении, а спектр имеет ярко выраженные частоты основного тона и форманты. Согласные звуки **Б, П, Т, Д, К, Г** образуются при коротком движении воздуха и их нельзя протяжно произнести, не прибавив к ним гласный. Кроме того, при произношении гласных фонем необходимо в первую очередь учитывать артикуляцию губ и языка и степень открытия рта.

Аллофоны согласных фонем делятся на твердые и мягкие, а по акустическим характеристикам и спектру делятся на шумные **Б, Г, Д, З, Ж, К, П, С, Ц, Х, Т, Ш, Ч, А** и сонорные **В, Ў, Л, М, Н, Ё**, в которых вокализованность фонемы превышает шумность.

Создание баз аллофонов для славянских языков рассматривается в работе [8], в которой указывается, что теоретически минимальная база аллофонов для русского и белорусского языков составляет 420, а для польского 535. На практике эти данные являются завышенными из-за того, что очень многие позиционные и комбинаторные ситуации вообще не встречаются в речи, и для ряда аллофонов акустические различия настолько невелики, что этими различиями можно пренебречь.

Для синтеза речеподобных сигналов на белорусском языке предложена база из 476 аллофонов, обозначение которых по принципу согласуется с обозначениями, предложенными Б.М. Лобановым [8]. Обозначение аллофонов состоит из названия аллофона и цифрового кода из трех позиций, например **a101**. Это означает, что в wav-файле записан аллофон **а**, безударный, окружение слева отсутствует, а справа перед фонемами **г, г', д, з, з', ж, ж', к, к', л, л', н, н', р, с, с', ц, ц', х, х', т, ш, й, ч, дз, дз', дж, дж'**. Верхний индекс [**'**] означает, что согласный мягкий. Если первый цифровой индекс после названия гласного аллофона **0**, то это означает, что гласный аллофон является безударным. Если же первый цифровой индекс после названия гласного аллофона **1**, то это означает, что гласный аллофон ударный. Для согласных аллофонов первый цифровой индекс **1** означает, что согласный аллофон твердый. Если первый цифровой индекс для согласного аллофона **2**, то это означает, что согласный аллофон мягкий.

Второй индекс аллофона характеризует его окружение слева. Если второй индекс **0**, то это означает, что окружение слева отсутствует, т.е. с этого аллофона начинается новое слово. Если второй индекс **1**, то слева ему предшествует гласная фонема. Если второй индекс **2**, то слева ему предшествует согласная фонема.

Третий индекс аллофона характеризует его окружение справа. Если третий индекс в обозначении аллофона **0**, то это означает, что окружение аллофона справа отсутствует, т.е. этим аллофоном заканчивается слово. Если третий индекс **1**, то справа за ним следует гласная фонема. Если третий индекс **2**, то справа за ним следует согласная фонема.

При синтезе белорусской речи по базе аллофонов некоторые трудности возникают при появлении в тексте мягкого знака. Для исключения этой трудности было предложено ввести в базу аллофонов согласные смягченные мягким знаком, такие как **Ць, Ль, Нь, Зь, Сь, Дзь**.

В соответствии с принятым обозначением перечень аллофонов для формирования речеподобных сигналов на белорусском языке приведен в табл. 6.

Таблица 6. Перечень аллофонов белорусского языка

a000	в212	з220	я020	ю010	л221	о111	у000	ц101	ш120	дж110
a001	в220	з221	я021	ю011	л222	о112	у001	ц102	ш121	дж111
a002	в221	з222	я022	ю012	м101	о120	у002	ц110	ш122	дж112
a010	в222	ж101	я101	ю020	м102	о121	у010	ц111	ч101	дж120
a011	ў100	ж102	я102	ю021	м110	о122	у011	ц112	ч102	дж121
a012	ў110	ж110	я110	ю022	м111	п101	у012	ц120	ч110	дж122
a020	ў112	ж111	я111	ю101	м112	п102	у020	ц121	ч111	дж201
a021	г101	ж112	я112	ю102	м120	п110	у021	ц122	ч112	дж202
a022	г102	ж120	я120	ю110	м121	п111	у022	ц201	ч120	дж210
a101	г110	ж121	я121	ю111	м122	п112	у101	ц202	ч121	дж211
a102	г111	ж122	я122	ю112	м201	п120	у102	ц210	ч122	дж212
a110	г112	ж201	е001	ю120	м202	п121	у110	ц211	ф101	дж220
a111	г120	ж202	е002	ю121	м210	п122	у111	ц212	ф102	дж221
a112	г121	ж210	е010	ю122	м211	п201	у112	ц220	ф110	дж222
a120	г122	ж211	е011	к100	м212	п202	у120	ц221	ф111	ць210
a121	г201	ж212	е012	к101	м220	п210	у121	ц222	ф112	ць212
a122	г202	ж220	е020	к102	м221	п211	у122	х101	ф120	ль210
б101	г210	ж221	е021	к110	м222	п212	э001	х102	ф121	ль212
б102	г211	ж222	е022	к111	н101	п220	э002	х110	ф122	нь210
б110	г212	і000	е101	к112	н102	п221	э010	х111	ф201	нь212
б111	г220	і001	е102	к120	н110	п222	э011	х112	ф202	зь210
б112	г221	і002	е110	к121	н111	р101	э012	х120	ф210	зь212
б120	г222	і010	е111	к122	н112	р102	э020	х121	ф211	сь210
б121	д101	і011	е112	к201	н120	р110	э021	х122	ф212	сь212
б122	д102	і012	е120	к202	н121	р111	э022	х201	ф220	дзь210
б201	д110	і020	е121	к210	н122	р112	э101	х202	ф221	дзь212
б202	д111	і021	е122	к211	н201	р120	э102	х210	ф222	
б210	д112	і022	ё001	к212	н202	р121	э110	х211	дз101	
б211	д120	і101	ё002	к220	н210	р122	э111	х212	дз102	
б212	д121	і102	ё010	к221	н211	с101	э112	х220	дз110	
б220	д122	і110	ё011	к222	н212	с102	э120	х221	дз111	
б221	з100	і111	ё012	л101	н220	с110	э121	х222	дз112	
б222	з101	і112	ё020	л102	н221	с111	э122	т101	дз120	
в101	з102	і120	ё021	л110	н222	с112	ы010	т102	дз121	
в102	з110	і121	ё022	л111	о001	с120	ы011	т110	дз122	
в110	з111	і122	ё101	л112	о002	с121	ы012	т111	дз201	
в111	з112	й120	ё102	л120	о010	с122	ы020	т112	дз202	
в112	з120	й121	ё110	л121	о011	с201	ы021	т120	дз210	
в120	з121	й122	ё111	л122	о012	с202	ы022	т121	дз211	
в121	з122	я000	ё112	л201	о020	с210	ы110	т122	дз212	
в122	з201	я001	ё120	л202	о021	с211	ы111	ш101	дз220	
в201	з202	я002	ё121	л210	о022	с212	ы112	ш102	дз221	
в202	з210	я010	ё122	л211	о101	с220	ы120	ш110	дз222	
в210	з211	я011	ю001	л212	о102	с221	ы121	ш111	дж101	
в211	з212	я012	ю002	л220	о110	с222	ы122	ш112	дж102	

Синтез текста для речеподобных сигналов на белорусском языке и его лингвистический анализ

Для синтеза текста для речеподобных сигналов на белорусском языке необходимо учитывать статистические характеристики белорусского языка, приведенные в таблицах выше. При этом сформированный текст речеподобных сигналов должен быть подвергнут лингвистическому анализу.

Лингвистический анализ текста речеподобных сигналов проводится с целью исключения из текста элементов речи, не характерных для белорусского языка. Слова с одной буквой могут состоять только из букв **А, З, І, К, У, Ў, Я**. Слова из трех букв должны содержать хотя бы одну гласную. В белорусском языке отсутствуют сочетания букв **СЧ, ЧІ, ЧЕ, ЧЯ, ТІ, ТЕ, ТЯ ДІ, ДЯ, ДЕ, РІ, РЕ, РЯ, ШІ, ШЕ, ШЯ**. Слова не могут начинаться с буквы **Ы**. Гласный **Ы** пишется после твердых согласных. **Ў** может появляться в тексте только после гласной, как в слове, так и после окончания предыдущего слова, оканчивающегося на гласную. Если в сформированном тексте речеподобных сигналов после гласной следует **У**, то ее следует заменить на **Ў**. Гласный **І** пишется после мягких согласных и в начале слова.

Заклучение

Синтез речеподобных сигналов на белорусском языке основан на статистических характеристиках текстов (длительности слов, вероятностей появления букв в тексте) на белорусском языке и преобразовании его в речеподобные сигналы с использованием базы аллофонов.

Работа выполнена при финансовой поддержке КН МОН РК грант 0027/ПЦФ-14-ОТ и Белорусского республиканского фонда фундаментальных исследований проект Ф14КАЗ-020.

SYNTHESIS OF SPEECH-LIKE SIGNALS IN THE BELARUSIAN LANGUAGE

H.V. DAVYDAU, V.A. PAPOU, A.V. POTAPOVICH, Y.N. SEITKULOV, I.V. SAVCHENKO

Abstract

The methods of synthesis of speech-like signal on the basis of allophones are analyzed. The development urgency of speech-like signals synthesis in the Belarusian language is grounded. The methodology of forming the base of allophones in the Belarusian language with the phonetic features of the Belarusian language is described.

Список литературы

1. ITU-T P.501 Test signals for use in telephony/ Series P: Telephone Transmission Quality. Objective measuring apparatus. P.27.
2. ETSI 3rd Speech Quality Test Event. Anonymized Test Report «IP Phones». P. 81.
3. ITU-T P.50 Artificial voice/ Series P: Telephone Transmission Quality, Telephone installations, Local line networks. Objective measuring apparatus. P.14.
4. Воробьев В.И., Давыдов А.Г., Давыдов Г.В., Ивонин А.И., Лещенко Д.В., Лобанов Б.М., Лыньков Л.М., Попов В.А., Потапович А.В. Устройство защиты речевой информации от утечки по вибрационным и акустическим каналам / Патент РБ № 3053.
5. Воробьев В.И., Давыдов А.Г., Давыдов Г.В. // Докл. БГУИР. 2009. № 3 (41). С. 9–16.
6. Сейткулов Е.Н., Давыдов Г.В., Потапович А.В. // Вест. КазНТУ. 2014. № 2 (102).
7. Киселев В.В., Лобанов Б.М. // Докл. БГУИР. 2004. №4. С. 138–142.
8. Фонетико-акустическая база данных для многоязычного синтеза речи по тексту на славянских языках. [Электронный ресурс]. – Режим доступа: <http://www.dialog-21.ru/digests/dialog2006/materials/html/Lobanov.htm>. – Дата доступа: 24.03.2015.