

# Алгоритмы преобразования «Фонема— Аллофон» двуязычного синтезатора речи

*Гецевич Ю.С., кандидат технических наук*

*Лобанов Б.М., доктор технических наук*

*Покладок Д.А., аспирант*

В статье рассмотрены особенности аллофонного набора для русского и белорусского языков. Предложен метод преобразования «Фонема—Аллофон» на основе экспертных правил. Реализован языконезависимый алгоритм преобразования «Фонема — Аллофон», общий для белорусского и русского языков, используемый в двуязычном синтезаторе речи по тексту.

- *фонема* ● *аллофон* ● *русский язык* ● *белорусский язык*
- *синтез речи по тексту*

---

The peculiarities of allophones set for Russian and Belarusian languages are described. A method of phoneme-allophone conversion based on expert rules is proposed. Language independent phoneme-allophone conversion algorithm common to the Belarusian and Russian languages used in a bilingual text to speech synthesizer is implemented.

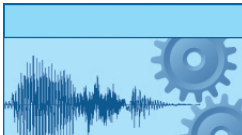
- *phoneme* ● *allophone* ● *Russian language* ● *Belarusian language*

## **Введение**

В потоке речи в зависимости от конкретных условий акустико-артикуляторные характеристики фонем изменяются, что приводит к появлению оттенков фонем, или аллофонов. Аллофоны подразделяются на позиционные и комбинаторные [1].

Позиционные аллофоны определяются положением фонемы по отношению к ударному слогу в слове. Ударные слоги в качественном и количественном отношении реализуются наиболее полно, в то время как безударные слоги, а в особенности гласные в заударных слогах могут редуцироваться в очень сильной степени.

Комбинаторные аллофоны определяются ближайшим контекстом фонемы и обусловлены наложением в потоке речи артикуляции одного звука на другой в результате действия эффектов коартикуляции и ассимиляции. В результате эффекта коартикуляции согласная фонема может в значительной степени приобрести окраску последующей гласной фонемы (например, огубление согласной /x/ в слове



«сухое — *сухойэ*» в сравнении с «сухая — *сухайа*»). Гласная же фонема в значительной степени приобретает окраску предшествующей согласной (например, смягчение гласной /а/ в слове «тяпки — *т'апк'и*» в сравнении с «*тапки* — *тапк'и*»).

Эффекты ассимиляции в наибольшей степени проявляются при стечении двух или более согласных (например, смягчение и потеря взрывного сегмента согласной /т/ в слове «отнять — *ат'н'ат'*»).

Эффекты редукции, коартикуляции и ассимиляции проявляются в основном внутри слоговых комплексов, в меньшей степени на стыках слогов и в ещё меньшей степени, но всё ещё заметной, на стыках слов внутри синтагмы. Описанные эффекты практически полностью прекращают своё действие только на стыках синтагм и фраз.

Описание фонем в потоке речи в виде последовательности аллофонов может быть осуществлено с различной степенью подробности. В общем случае невозможно дать точную оценку количества аллофонов, так как она напрямую зависит от степени детализации учёта влияния позиционных и комбинаторных факторов. При этом общее количество получаемых аллофонов может колебаться от нескольких сотен до нескольких десятков тысяч [2, 3]. В данной работе описывается практически опробованный и хорошо зарекомендовавший себя вариант аллофонного набора гласных, обеспечивающий достаточно подробное описание их оттенков в потоке речи и подтверждённый достаточно высокими результатами тестирования синтезированной речи [4].

*Таблица 1*

**Позиционные аллофоны гласных**

Фонема	Аллофон							
	Полноударный (0-й степени редукции)		Частично-ударный (1-й степени редукции)		Безударный (2-й степени редукции)		Безударный (3-й степени редукции)	
Язык	Бел.	Рус.	Бел.	Рус.	Бел.	Рус.	Бел.	Рус.
А	A <sub>0xx</sub>	A <sub>0xx</sub>	A <sub>1xx</sub>	A <sub>1xx</sub>	A <sub>2xx</sub>	A <sub>2xx</sub>	A <sub>3xx</sub>	A <sub>3xx</sub>
Е	E <sub>0xx</sub>	E <sub>0xx</sub>	E <sub>1xx</sub>	E <sub>1xx</sub>	-	E <sub>2xx</sub>	-	E <sub>3xx</sub>
І	I <sub>0xx</sub>	I <sub>0xx</sub>	I <sub>1xx</sub>	I <sub>1xx</sub>	I <sub>2xx</sub>	I <sub>2xx</sub>	I <sub>3xx</sub>	I <sub>3xx</sub>
О	O <sub>0xx</sub>	O <sub>0xx</sub>	O <sub>1xx</sub>	O <sub>1xx</sub>	-	-	-	-
У	U <sub>0xx</sub>	U <sub>0xx</sub>	U <sub>1xx</sub>	U <sub>1xx</sub>	U <sub>2xx</sub>	U <sub>2xx</sub>	U <sub>3xx</sub>	U <sub>3xx</sub>
Ү	Y <sub>0xx</sub>	Y <sub>0xx</sub>	Y <sub>1xx</sub>	Y <sub>1xx</sub>	Y <sub>2xx</sub>	Y <sub>2xx</sub>	Y <sub>3xx</sub>	Y <sub>3xx</sub>
Количество	6	6	6	6	4	5	4	5
У с л о в и е ПФА	Перед знаком ударения (+)		Перед знаком ударения (=)		В 1-й предударной позиции, в начале или в конце слова		В остальных случаях	

*Примечание к таблице 1:*

«**ж**» — *неопределяемые данной таблицей 2-е и 3-и индексы комбинаторных аллофонов.*

## 1. Аллофонный набор для гласных фонем русского и белорусского языков

В описываемом наборе каждый аллофон сохраняет имя соответствующей гласной фонемы и снабжается 3-мя цифровыми индексами: 1-й индекс обозначает степень её позиционной редукции, 2-й индекс — тип комбинаторного влияния левого контекста, а 3-й индекс — правого.

В таблице 1 представлен набор позиционных аллофонов, их обозначения, количество и условия преобразования «Фонема-Аллофон» (ПФА).

Как видно из таблицы, особенностью белорусского языка является отсутствие безударных аллофонов 2-й и 3-й степени редукции у фонемы «Е». Характерным для обоих языков является отсутствие безударных аллофонов 2-й и 3-й степени редукции у фонемы «О».

В таблицах 2 и 3 представлены наборы комбинаторных аллофонов гласных, их обозначения, количество и условия ПФА, определяемые, соответственно, левым и правым контекстом.

Таблица 2

### Комбинаторные аллофоны гласных, обусловленные левым контекстом

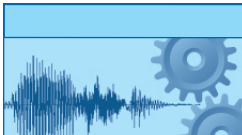
Условия ПФА (левый контекст)		Комбинаторные аллофоны левого контекста					
1	Пауза (начало синтагмы)	A <sub>x0x</sub>	E <sub>x0x</sub>	I <sub>x0x</sub>	O <sub>x0x</sub>	U <sub>x0x</sub>	—
2	Губные твёрдые согласные: В, F, M, V, L*, W, гласные U, O (внутри слова, на стыке слов)	A <sub>x1x</sub>	E <sub>x1x</sub>	I <sub>x1x</sub>	O <sub>x1x</sub>	U <sub>x1x</sub>	Y <sub>x1x</sub>
3	Переднеязычные твёрдые согласные: С, D, N, R, S, SH, Z, T, ZH, L**, CH, DZH, гласные А, Е (внутри слова, на стыке слов)	A <sub>x2x</sub>	E <sub>x2x</sub>	I <sub>x2x</sub>	O <sub>x2x</sub>	U <sub>x2x</sub>	Y <sub>x2x</sub>
4	Заднеязычные твёрдые согласные: G, H, K, GH, гласная Y (внутри слова, на стыке слов)	A <sub>x3x</sub>	E <sub>x3x</sub>	I <sub>x3x</sub>	O <sub>x3x</sub>	U <sub>x3x</sub>	—
5	Мягкие согласные: M', N', L', R', J', V', Z', F', S', SH', H', B', D', G', P', T', K', CH', GH', гласная i (внутри слова)	A <sub>x4x</sub>	E <sub>x4x</sub>	I <sub>x4x</sub>	O <sub>x4x</sub>	U <sub>x4x</sub>	—
6	Мягкие согласные: M', N', L', R', J', V', Z', F', S', SH', H', B', D', G', P', T', K', CH', GH', гласная i (на стыке слов)	A <sub>x5x</sub>	E <sub>x5x</sub>	I <sub>x5x</sub>	O <sub>x5x</sub>	U <sub>x5x</sub>	—
Количество левых контекстов		6	6	6	6	6	2

#### Примечания к таблице 2:

«x» — неопределяемые данной таблицей 1-е и 3-и индексы;

L\* — перед гласными A, E, I, Y;

L\*\* — перед гласными O, U.



Условия левого контекста в таблице 2 включают 6 различных условий (наборов фонем), встречающихся перед гласной внутри или на стыке слов. Состав наборов фонем выбран экспериментально, исходя из особенностей и степени их влияния на последующую гласную, и группируются, в основном, по признакам места образования фонемы. Выбранное количество из 6-ти различных левых контекстов является достаточным для достижения приемлемого покрытия всего многообразия оттенков гласных фонем, обусловленных комбинациями с предшествующей фонемой.

На этой и последующих таблицах **жирным прямым шрифтом** помечены согласные фонемы, существующие только в русском языке, а **жирным курсивом** помечены согласные фонемы, существующие только в белорусском языке.

В таблице 3 представлен набор комбинаторных аллофонов гласных, определяемых правым контекстом фонемы. Экспериментальные исследования показывают, что правый контекст по сравнению с левым оказывает меньшее воздействие на фонетическое качество гласной фонемы, так что условия правого контекста включают всего 4 различных набора фонем. Их количество является достаточным для достижения приемлемого покрытия всего многообразия оттенков гласных фонем, обусловленных комбинациями с последующей фонемой.

Таблица 3

**Комбинаторные аллофоны гласных, обусловленные правым контекстом**

Условия ПФА (правый контекст)		Комбинаторные аллофоны правого контекста					
1	Пауза(конец синтагмы)	A <sub>xx0</sub>	E <sub>xx0</sub>	I <sub>xx0</sub>	O <sub>xx0</sub>	U <sub>xx0</sub>	Y <sub>xx0</sub>
2	Губные твёрдые согласные: В, F, M, V, L, <b>W</b> , гласные U, O	A <sub>xx1</sub>	E <sub>xx1</sub>	I <sub>xx1</sub>	O <sub>xx1</sub>	U <sub>xx1</sub>	Y <sub>xx1</sub>
3	Переднеязычные и заднеязычные твёрдые согласные: C, D, N, R, S, SH, Z, T, ZH, <b>CH, DZH, G</b> , H, K, <b>GH</b> , гласные A, E, Y	A <sub>xx2</sub>	E <sub>xx2</sub>	I <sub>xx2</sub>	O <sub>xx2</sub>	U <sub>xx2</sub>	Y <sub>xx2</sub>
4	Мягкие согласные: M', N', L', <b>R'</b> , J', V', Z', F', S', <b>SH'</b> , H', B', <b>D'</b> , <b>G'</b> , P', T', K', <b>CH'</b> , <b>GH'</b> , гласный I	A <sub>xx3</sub>	E <sub>xx3</sub>	I <sub>xx3</sub>	O <sub>xx3</sub>	U <sub>xx3</sub>	Y <sub>xx3</sub>
Количество правых контекстов		4	4	4	4	4	4

Примечание к таблице 3:

«x» — неопределяемые данной таблицей 1-е и 2-е индексы.

Количество аллофонов для каждой из гласных фонем подсчитывается как произведение количеств позиционных и комбинаторных аллофонов, приведённых в таблицах 1–3.

Для русского языка:

$$N_a=4*6*4=96; N_e=4*6*4=96; N_i=4*6*4=96; N_o=2*6*4=48; N_u=4*6*4=96; N_y=4*2*4=32.$$

Для белорусского языка:

$$N_a=4*6*4=96; N_e=2*6*4=48; N_i=4*6*4=9; N_o=2*6*4=48; N_u=4*6*4=96; N_y=4*2*4=32.$$

Общее количество аллофонов гласных фонем:

для русского языка  $N_{all}^v = 464$ ;

для белорусского языка  $N_{all}^v = 416$ .

На *рис. 1, 2* приведены примеры спектрограмм, иллюстрирующих комбинаторное влияние левого и правого контекстов на гласную «А» 0-й и 2-й степени редукции.

## 2. Аллофонный набор для согласных фонем русского и белорусского языков

В сравнении с гласными фонемами, согласные значительно более устойчивы к позиционным факторам изменчивости и в меньшей степени к комбинаторным. Причём влияние комбинаторных факторов различно для различных подмножеств согласных фонем. По степени и типу комбинаторной изменчивости согласные фонемы целесообразно разделить на следующие подгруппы.

Глухие: Р, Р', Т, Т', К', С, **CH'**, F, F', S, S', SH, **SH'**, Н', **C'**, **CH**, **SCH**.

Звонкие: В, В', Д, Д', **G'**, Z, Z', ЗН, L, L', М, М', N, N'**DZ'**, **DZH**, **GH'**.

3. Твёрдые заднеязычные: К, **G**, Н, **GH**.

4. Плавные V, V', J', **W**.

5. Дрожжацие: R, **R'**.

В отличие от гласных фонем, для которых решающее влияние на акустическое качество звука оказывает левый контекст, согласные фонемы наиболее сильно подвержены влиянию правого контекста. Влияние левого контекста практически отсутствует на

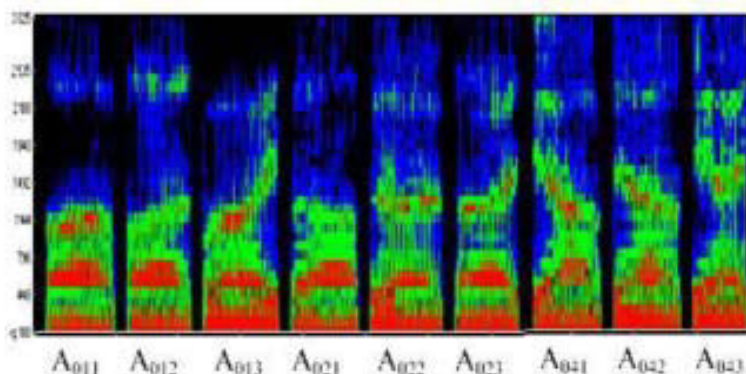


Рис. 1. Спектрограммы комбинаторных аллофонов полноударной гласной «А»

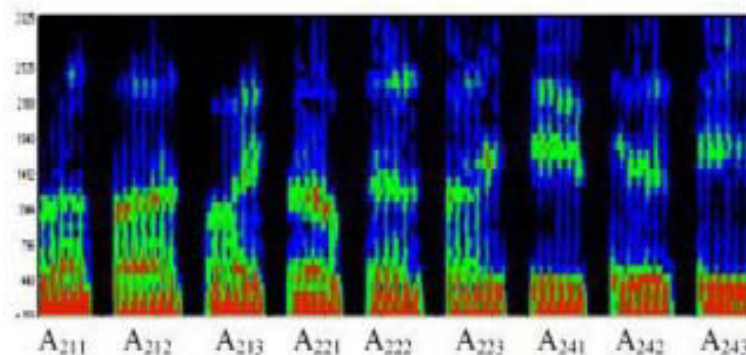


Рис. 2. Спектрограммы комбинаторных аллофонов гласной «А» 2-й степени редукции



согласные 1-й, 2-й и 3-й подгрупп, и проявляется только в 4-й и 5-й подгруппах. Позиционная изменчивость согласных также менее значительна, чем у гласных, для каждой из перечисленных подгрупп.

В таблице 4 представлен набор комбинаторных аллофонов согласных 1-й подгруппы (глухие), их обозначения и условия ПФА, определяемые правым контекстом фонемы, а также количество аллофонов для русского и белорусского языков.

Таблица 4

**Комбинаторные аллофоны согласных 1-й подгруппы:**  
 P, P', T, T', K', C, CH', F, F', S, S', SH, SH', H', C', CH, SCH

Условия ПФА правого контекста						
№ пп	Одиночная фонема			Удвоенная фонема		
	Ударная гласная <sup>1</sup>	Безударная гласная или любая согласная <sup>2</sup>	Пауза (конец синтагмы)	Ударная	Безударная	Пауза
	1	2	3	4	5	6
1	P <sub>001</sub>	P <sub>002</sub>	P <sub>000</sub>	P <sub>101</sub>	P <sub>102</sub>	-
2	P' <sub>001</sub>	P' <sub>002</sub>	P' <sub>000</sub>	P' <sub>101</sub>	P' <sub>102</sub>	-
3	T <sub>001</sub>	T <sub>002</sub>	T <sub>000</sub>	T <sub>101</sub>	T <sub>102</sub>	T <sub>100</sub>
4	T' <sub>001</sub>	T' <sub>002</sub>	T' <sub>000</sub>	T' <sub>101</sub>	T' <sub>102</sub>	-
5	K' <sub>001</sub>	K' <sub>002</sub>	-	K' <sub>101</sub>	K' <sub>102</sub>	-
6	C <sub>001</sub>	C <sub>002</sub>	C <sub>000</sub>	C <sub>101</sub>	C <sub>102</sub>	-
7	CH' <sub>001</sub>	CH' <sub>002</sub>	CH' <sub>000</sub>	CH' <sub>101</sub>	CH' <sub>102</sub>	-
8	F <sub>001</sub>	F <sub>002</sub>	F <sub>000</sub>	F <sub>101</sub>	F <sub>102</sub>	F <sub>100</sub>
9	F' <sub>001</sub>	F' <sub>002</sub>	F' <sub>000</sub>	F' <sub>101</sub>	F' <sub>102</sub>	-
10	S <sub>001</sub>	S <sub>002</sub>	S <sub>000</sub>	S <sub>101</sub>	S <sub>102</sub>	S <sub>100</sub>
11	S' <sub>001</sub>	S' <sub>002</sub>	S' <sub>000</sub>	S' <sub>101</sub>	S' <sub>102</sub>	-
12	SH <sub>001</sub>	SH <sub>002</sub>	SH <sub>000</sub>	SH <sub>101</sub>	SH <sub>102</sub>	-
13	SH' <sub>001</sub>	SH' <sub>002</sub>	SH' <sub>000</sub>	SH' <sub>101</sub>	SH' <sub>102</sub>	-
14	H' <sub>001</sub>	H' <sub>002</sub>	-	H' <sub>101</sub>	H' <sub>102</sub>	-
15	C' <sub>001</sub>	C' <sub>002</sub>	C' <sub>000</sub>	C' <sub>101</sub>	C' <sub>102</sub>	-
16	CH <sub>001</sub>	CH <sub>002</sub>	CH <sub>000</sub>	CH <sub>101</sub>	CH <sub>102</sub>	-
Количество аллофонов в русском языке						71
Количество аллофонов в белорусском языке						66

Примечание к таблице 4:

<sup>1</sup> — полноударные гласные: A<sub>0</sub>, E<sub>0</sub>, O<sub>0</sub>, U<sub>0</sub>, I<sub>0</sub>, Y<sub>0</sub> ;

<sup>2</sup> — гласные: A<sub>1</sub>, E<sub>1</sub>, O<sub>1</sub>, U<sub>1</sub>, I<sub>1</sub>, Y<sub>1</sub>, A<sub>2</sub>, E<sub>2</sub>, U<sub>2</sub>, I<sub>2</sub>, Y<sub>2</sub>, A<sub>3</sub>, E<sub>3</sub>, U<sub>3</sub>, I<sub>3</sub>, Y<sub>3</sub> ;

согласные: P, P', T, T', K', C, CH', F, F', S, S', SH, SH', H', C', CH, SCH, K, G, H, GH, B, B', D, D', G', Z, Z', ZH, L, L', M, M', N, N', DZ', DZH, GH', R, R', V, V', J, W.

Приведём несколько примеров акустической реализации комбинаторных аллофонов согласных этой подгруппы. На рис. 3 представлена осциллограмма фразы: «Он сумел сам», в которой реализованы два типа аллофонов одиночной фонемы «S» в позициях перед безударной гласной в слове «сумел» и в позиции перед полноударной гласной в слове «сам». Как видно из



рисунка, аллофоны S002 и S001 незначительно отличаются по спектру, а основной отличительной характеристикой является их длительность: 145мс и 185мс, соответственно. Так как реальная длительность аллофонов зависит от темпа речи, то для их характеристики используют относительные значения длительности в сравнении, например, с длительностью аллофона в позиции перед полноударной гласной. Относительная длительность рассматриваемых аллофонов равна соответственно 0,75 и 1,0.

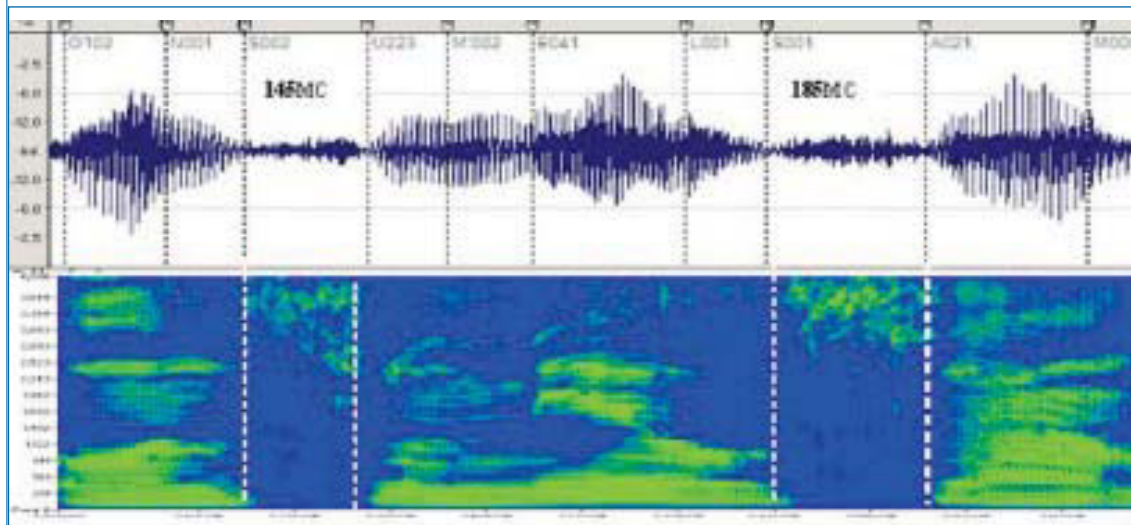


Рис. 3. Спектрограмма (внизу) и осциллограмма фразы: «Он сумел сам»

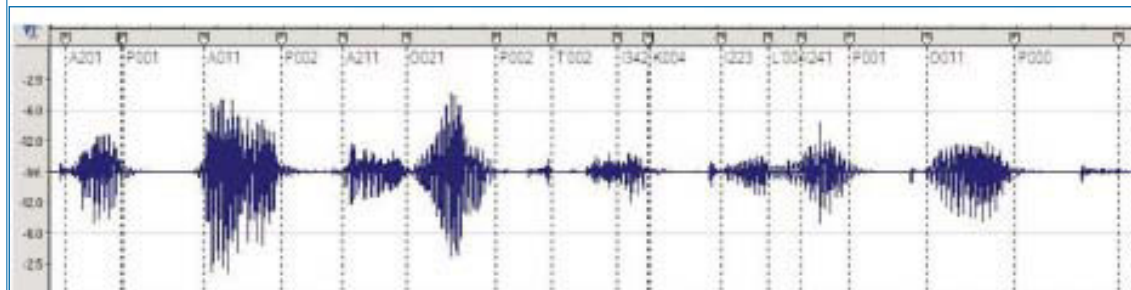


Рис. 4. Осциллограмма фразы: «А папа оптик или поп?»

На рис. 4 представлена осциллограмма фразы: «А папа оптик или поп?», в которой реализованы три типа аллофонов одиночной фонемы «Р», а именно:

P001 — в позициях перед ударными гласными в словах «папа» и «поп» с длительностями 122 мс и 117 мс (среднее значение — 120 мс);

P002 — в позициях перед безударной гласной в слове «папа» и в позиции перед согласной в слове «оптик» с длительностями 95 мс и 85 мс (среднее значение — 90 мс);

P000 — в позиции «конец синтагмы» в слове «поп» с длительностью 155 мс.

На рис. 5 представлена осциллограмма фраз:

«А поп падал? И поп Пахом?», в которых реализованы два типа аллофонов удвоенной фонемы «П», а именно:

P101 — на стыке слитно произносимых слов «поп» и «падал» перед ударной гласной с длительностью 220 мс;



P102 — — на стыке слитно произносимых слов «поп» и «Пахом» перед безударной гласной с длительностью 175 мс;

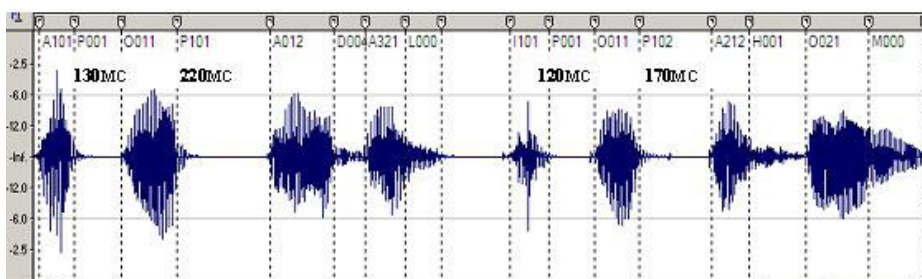


Рис. 5. Осциллограмма фраз: «А поп па+дал? И поп Пахо+м?»

В этих фразах дважды присутствует также аллофон P001 с длительностями 130мс и 120мс (среднее значение — 125мс).

Относительные длительности аллофонов согласной «Р», рассчитанные в соответствии с данными рис. 4, 5, приведены в таблице 5.

Таблица 5

**Относительная длительность аллофонов согласной «Р»**

Условия ПФА правого контекста	Одиночная фонема			Удвоенная фонема	
	Ударная гласная <sup>1</sup>	Безударная гласная или любая согласная <sup>2</sup>	Пауза (конец синтагмы)	Ударная	Безударная
	1	2	3	4	5
Имя аллофона	P <sub>001</sub>	P <sub>002</sub>	P <sub>000</sub>	P <sub>101</sub>	P <sub>102</sub>
Относительная длительность	1,00	0,75	1,30	1,75	1,40

Исследование относительной длительности остальных аллофонов этой подгруппы показали результаты, сходные с приведёнными в таблице 5.

В таблице 6 представлен набор комбинаторных аллофонов согласных 2-й подгруппы (звонкие), их обозначения и условия ПФА, определяемые правым контекстом фонемы, а также количество аллофонов для русского и белорусского языков. Как и для аллофонов 1-й подгруппы, основным различительным признаком аллофонов 2-й подгруппы является их относительная длительность, которая для соответствующих аллофонов 2-й подгруппы распределяется примерно так же, как показано в таблице 5.



Таблица 6

**Комбинаторные аллофоны согласных 2-й подгруппы:  
В, В', D, D', G', Z, Z', ZH, L, L', M, M', N, N'DZ', DZH, GH'**

№ п/п	Условия ПФА правого контекста					
	Одиночная фонема				Удвоенная фонема	
	Ударная гласная <sup>(1)</sup>	Безударная гласная <sup>(2)</sup>	Любая согласная <sup>(3)</sup>	Пауза (конец синтагмы)	Ударная	Безударная
	1	2	3	4	5	6
1	V <sub>002</sub>	V <sub>004</sub>	V <sub>001</sub>	-	V <sub>102</sub>	V <sub>104</sub>
2	V' <sub>002</sub>	V' <sub>004</sub>	V' <sub>001</sub>	-	V' <sub>102</sub>	V' <sub>104</sub>
3	D <sub>002</sub>	D <sub>004</sub>	D <sub>001</sub>	-	D <sub>102</sub>	D <sub>104</sub>
4	D' <sub>002</sub>	D' <sub>004</sub>	D' <sub>001</sub>	-	D' <sub>102</sub>	D' <sub>104</sub>
5	G <sub>002</sub>	G <sub>004</sub>	G <sub>001</sub>	-	G <sub>102</sub>	G <sub>104</sub>
6	Z <sub>002</sub>	Z <sub>004</sub>	Z <sub>001</sub>	-	Z <sub>102</sub>	Z <sub>104</sub>
7	Z' <sub>002</sub>	Z' <sub>004</sub>	Z' <sub>001</sub>	-	Z' <sub>102</sub>	Z' <sub>104</sub>
8	ZH <sub>002</sub>	ZH <sub>004</sub>	ZH <sub>001</sub>	-	ZH <sub>102</sub>	ZH <sub>104</sub>
9	L <sub>002</sub>	L <sub>004</sub>	L <sub>001</sub>	L <sub>000</sub>	L <sub>102</sub>	L <sub>104</sub>
10	L' <sub>002</sub>	L' <sub>004</sub>	L' <sub>001</sub>	L' <sub>000</sub>	L' <sub>102</sub>	L' <sub>104</sub>
11	M <sub>002</sub>	M <sub>004</sub>	M <sub>001</sub>	M <sub>000</sub>	M <sub>102</sub>	M <sub>104</sub>
12	M' <sub>002</sub>	M' <sub>004</sub>	M' <sub>001</sub>	M' <sub>000</sub>	M' <sub>102</sub>	M' <sub>104</sub>
13	N <sub>002</sub>	N <sub>004</sub>	N <sub>001</sub>	N <sub>000</sub>	N <sub>102</sub>	N <sub>104</sub>
14	N' <sub>002</sub>	N' <sub>004</sub>	N' <sub>001</sub>	N' <sub>000</sub>	N' <sub>102</sub>	N' <sub>104</sub>
15	DZ' <sub>002</sub>	DZ' <sub>004</sub>	DZ' <sub>001</sub>	-	DZ' <sub>102</sub>	DZ' <sub>104</sub>
16	DZH <sub>002</sub>	DZH <sub>004</sub>	DZH <sub>001</sub>	-	DZH <sub>102</sub>	DZH <sub>104</sub>
17	<b>GH'</b> <sub>002</sub>	<b>GH'</b> <sub>004</sub>	<b>GH'</b> <sub>001</sub>	-	<b>GH'</b> <sub>102</sub>	<b>GH'</b> <sub>104</sub>
Количество аллофонов в русском языке						76
Количество аллофонов в белорусском языке						86

**Примечание к таблице 6:**

<sup>(1)</sup> — полноударные гласные: A<sub>0</sub>, E<sub>0</sub>, O<sub>0</sub>, U<sub>0</sub>, I<sub>0</sub>, Y<sub>0</sub>;

<sup>(2)</sup> — гласные: A<sub>1</sub>, E<sub>1</sub>, O<sub>1</sub>, U<sub>1</sub>, I<sub>1</sub>, Y<sub>1</sub>, A<sub>2</sub>, E<sub>2</sub>, U<sub>2</sub>, I<sub>2</sub>, Y<sub>2</sub>, A<sub>3</sub>, E<sub>3</sub>, U<sub>3</sub>, I<sub>3</sub>, Y<sub>3</sub>;

<sup>(3)</sup> — согласные: P, P', T, T', K', C, CH', F, F', S, S', SH, SH', H', C', CH, SCH, K, G, H, GH, B, B', D, D', G', Z, Z', ZH, L, L', M, M', N, N', **DZ', DZH, GH', R, R', V, V', J', W.**

Особенностью аллофонов 2-й подгруппы является выделение отдельного условия правого контекста: «Любая согласная», в отличие от 1-й подгруппы — «Безударная гласная или любая согласная». Это связано с тем, что при указанном условии между звонкой согласной 2-й подгруппы и последующей согласной, как правило, появляется специфическая «гласная вставка». Это явление иллюстрируется рисунком 6, где представлены осциллограммы и спектрограммы слов: «Обня+л» и «Немно+го», в которых реализовано условие 3-го столбца таблицы 6. На осциллограмме и спектрограмме видно, что в слове «Обня+л» у аллофона V<sub>001</sub> появляется гласная вставка длительностью 55 мс, а в слове «Немно+го» — у аллофона M<sub>001</sub> длительностью 45 мс (гласные вставки обозначены значком (+) и отмечены на спектрограмме штриховыми линиями).

В таблице 7 представлен набор комбинаторных аллофонов согласных 3-й подгруппы, их обозначения и условия ПФА, определяемые правым контекстом фонемы, а также количество аллофонов для русского и белорусского языков.

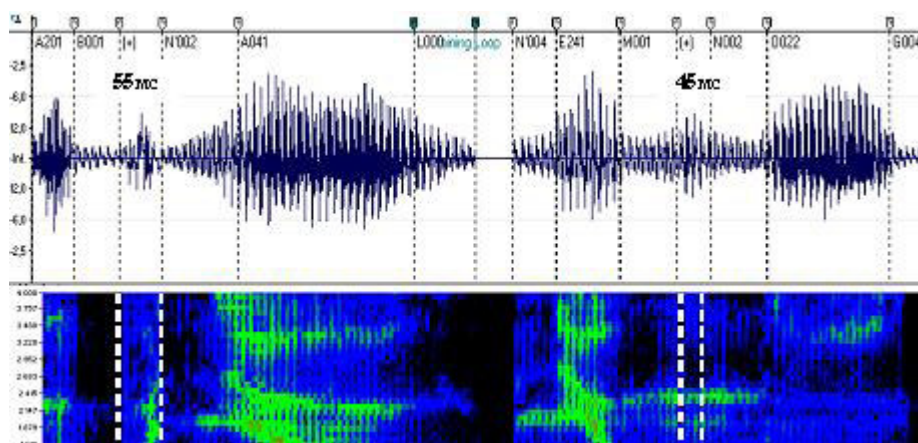


Рис. 6. Спектрограмма (внизу) и осциллограмма слов «Обня+л» и «Немно+го»

Таблица 7

**Комбинаторные аллофоны согласных 3-й подгруппы: К, Г, Н, ГН**

№ пп	Условия ПФА правого контекста				
	Ударная гласная U <sub>0</sub> или O <sub>0</sub>	Ударная гласная A <sub>0</sub> , E <sub>0</sub> , I <sub>0</sub> или Y <sub>0</sub>	Безударная гласная U <sub>1</sub> , U <sub>2</sub> , U <sub>3</sub> или O <sub>1</sub>	Безударная гласная A, E, I или Y, или любая согласная <sup>(1)</sup>	Пауза (конец синтагмы)
	1	2	3	4	5
1	K <sub>001</sub>	K <sub>002</sub>	K <sub>003</sub>	K <sub>004</sub>	K <sub>000</sub>
2	G <sub>001</sub>	G <sub>002</sub>	G <sub>003</sub>	G <sub>004</sub>	—
3	H <sub>001</sub>	H <sub>002</sub>	H <sub>003</sub>	H <sub>004</sub>	H <sub>000</sub>
4	<b>GH<sub>001</sub></b>	<b>GH<sub>002</sub></b>	<b>GH<sub>003</sub></b>	<b>GH<sub>004</sub></b>	—
Количество аллофонов в русском языке					14
Количество аллофонов в белорусском языке					14

Примечание к таблице 7:

<sup>(1)</sup> — согласные: P, P', T, T', K', C, CH', F, F', S, S', SH, SH', H', C', CH, SCH, K, G, H, GH, B, B', D, D', G', Z, Z', ZH, L, L', M, M', N, N', DZ', DZH, GH', R, R', V, V', J', W.

Особенностью аллофонов 3-й подгруппы (твёрдые заднеязычные) является выделение 2-х отдельных ударных и безударных условий правого контекста: «гласная U или O» (губные гласные фонемы) и «гласная A<sub>0</sub>, E<sub>0</sub>, I<sub>0</sub> или Y<sub>0</sub>» (остальные гласные). Это связано с тем, что акустические характеристики твёрдых заднеязычных согласных в значительной степени зависят от типа последующей гласной. Особенно контрастное изменение акустических характеристик этих согласных происходит при соседстве с погубленными и неогубленными гласными или согласными. Это явление иллюстрируется рисунком 7, где представлены осциллограммы и спектрограммы слов: «хутор — ху+тар», «хата — ха+та», «кто — хто+», в которых реализованы условиях 1-го, 2-го и 4-го столбца таблицы 7. На спектрограмме видно, что частотная локализация аспиративного шума у аллофона H<sub>001</sub> в слове «ху+тар» существенно отличается от частотной локализации шума у аллофона H<sub>002</sub> в словах «ха+та» и «хто+».

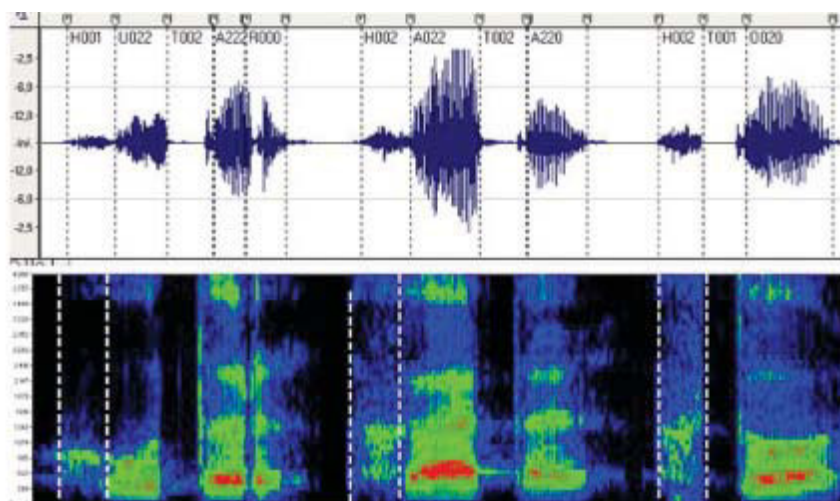


Рис. 7. Спектрограмма (внизу) и осциллограмма слов «Хутор», «Хата», «Кто»

В отличие от согласных фонем 1-й, 2-й и 3-й подгрупп, для которых решающее влияние на акустическое качество звука оказывает только правый контекст, согласные фонемы 4-й и 5-й подгрупп подвержены влиянию также и левого контекста. В таблице 8 представлен набор комбинаторных аллофонов согласных 4-й подгруппы, их обозначения и условия ПФА, определяемые левым и правым контекстом фонемы, а также количество аллофонов для русского и белорусского языков.

Таблица 8

**Комбинаторные аллофоны согласных 4-й подгруппы: V, V', J', W**

Правый контекст		Одиночная фонема				Удвоенная фонема	
		Ударная гласная (1)	Безударная гласная (2)	Любая согласная (3)	Пауза (конец синтагмы)	Ударная	Безударная
Любая согласная	1	V <sub>001</sub>	V <sub>002</sub>	V <sub>003</sub>	—	-	—
		V' <sub>001</sub>	V' <sub>002</sub>	—	—	-	—
		J' <sub>001</sub>	J' <sub>002</sub>	—	—	-	—
Любая гласная	2	V <sub>011</sub>	V <sub>012</sub>	V <sub>013</sub>	—	V <sub>111</sub>	V <sub>112</sub>
		V' <sub>011</sub>	V' <sub>012</sub>	V' <sub>013</sub>	—	V' <sub>111</sub>	V' <sub>112</sub>
		J' <sub>011</sub>	J' <sub>012</sub>	J' <sub>013</sub>	J' <sub>010</sub>	J' <sub>111</sub>	J' <sub>112</sub>
		W <sub>011</sub>	W <sub>012</sub>	W <sub>013</sub>	W <sub>010</sub>	—	—
Количество аллофонов в русском языке							24
Количество аллофонов в белорусском языке							28

Примечание к таблице 8:

- (1) — полноударные гласные: A<sub>0</sub>, E<sub>0</sub>, O<sub>0</sub>, U<sub>0</sub>, I<sub>0</sub>, Y<sub>0</sub>;
- (2) — гласные: A<sub>1</sub>, E<sub>1</sub>, O<sub>1</sub>, U<sub>1</sub>, I<sub>1</sub>, Y<sub>1</sub>, A<sub>2</sub>, E<sub>2</sub>, U<sub>2</sub>, I<sub>2</sub>, Y<sub>2</sub>, A<sub>3</sub>, E<sub>3</sub>, U<sub>3</sub>, I<sub>3</sub>, Y<sub>3</sub>;
- (3) — согласные: P, P', T, T', K', C, CH', F, F', S, S', SH, SH', H', C', CH, SCH, K, G, H, GH, B, B', D, D', G', Z, Z', ZH, L, L', M, M', N, N', DZ', DZH, GH', R, R', V, V', J', W.

Особенностью аллофонов 4-й подгруппы (плавные) является наличие 2-х условий левого контекста: «Любая согласная» и «Любая гласная». Акустические характеристики этих согласных в значительной степени зависят от указанных условий. Это иллюстрируется рисунком 8, где представлены осциллограммы и спектрограммы слов: «Майка», «Своя», «Пьяный», в которых реализованы условия 3-го столбца и 2-й строки таблицы 8 аллофона —  $J'_{013}$  в слове «Майка», 1-го столбца и 2-й строки аллофона —  $J'_{011}$  в слове «Своя», а также 1-го столбца и 1-й строки аллофона —  $J'_{001}$ , 4-го столбца и 2-й строки аллофона —  $J'_{010}$  в слове «Пьяный». На спектрограмме ясно видны их существенные отличия.

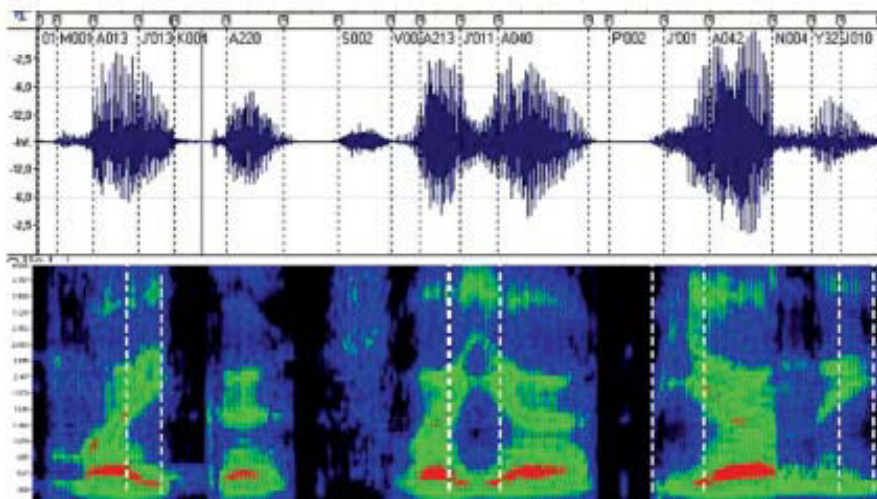


Рис. 8. Спектрограмма (внизу) и осциллограмма слов «Майка», «Своя», «Пьяный»

В таблице 9 представлен набор комбинаторных аллофонов согласных 5-й подгруппы (дрожащие), их обозначения и условия ПФА, определяемые левым и правым контекстом фонемы, а также количество аллофонов для русского и белорусского языков.

Таблица 9

**Комбинаторные аллофоны согласных 5-й подгруппы: R, R'**

Правый контекст		Одиночная фонема				Удвоенная фонема
		Глухие согласные <sup>(4)</sup>	Любая гласная <sup>(1)</sup>	Звонкие согласные <sup>(5)</sup>	Пауза (конец синтагмы)	
Левый контекст		1	2	3	4	5
Любая гласная <sup>(1)</sup>	1	$R_{003}$	$R_{002}$	$R_{001}$	$R_{000}$	$R_{102}$
		$R'_{003}$	$R'_{002}$	$R'_{001}$	$R'_{000}$	$R'_{102}$
Твёрдые губные согласные <sup>(2)</sup>	2	$R_{013}$	$R_{012}$	$R_{011}$	$R_{010}$	—
		$R'_{013}$	$R'_{012}$	$R'_{011}$	$R'_{010}$	—
Остальные согласные <sup>(3)</sup>	3	$R_{023}$	$R_{022}$	$R_{021}$	$R_{020}$	—
		$R'_{023}$	$R'_{022}$	$R'_{021}$	—	—
Количество аллофонов в русском языке						25
Количество аллофонов в белорусском языке						13



Примечание к таблице 9:

<sup>(1)</sup> — гласные: A<sub>0</sub>, E<sub>0</sub>, O<sub>0</sub>, U<sub>0</sub>, I<sub>0</sub>, Y<sub>0</sub>, A<sub>1</sub>, E<sub>1</sub>, O<sub>1</sub>, U<sub>1</sub>, I<sub>1</sub>, Y<sub>1</sub>, A<sub>2</sub>, E<sub>2</sub>, U<sub>2</sub>, I<sub>2</sub>, Y<sub>2</sub>, A<sub>3</sub>, E<sub>3</sub>, U<sub>3</sub>, I<sub>3</sub>, Y<sub>3</sub>;

<sup>(2)</sup> — согласные: B, P, F, M, V, W, L;

<sup>(3)</sup> — согласные: C, D, N, R, S, SH, Z, T, ZH, CH, DZH, SCH, H, G, K, GH, P', T', K', CH', F', S', SH', H', C', B', D', G', Z', L', M', N', DZ', GH', R', V', J';

<sup>(4)</sup> согласные: P, P', T, T', K', C, CH', F, F', S, S', SH, SH', H', C', CH, SCH, K, G, H, GH; <sup>(5)</sup> согласные: B, B', D, D', G', Z, Z', ZH, L, L', M, M', N, N', DZ', DZH, GH', R, R', V, V', J', W.

Особенностью аллофонов 5-й подгруппы (согласные R и R') является наличие 2-х дополнительных условий левого контекста: «Твёрдые губные согласные» и «Остальные согласные». Акустические характеристики согласных R и R' в значительной степени зависят от указанных условий. Это иллюстрируется рисунком 9, где представлены осциллограммы и спектрограммы слов: «Брал. Драл. Брил. Дрель», в которых реализованы условия 2-го столбца и 2-й строки таблицы 9 для аллофона — R<sub>012</sub> в слове «Брал», 2-го столбца и 3-й строки для аллофона — R<sub>022</sub> в слове «Драл», а также 2-го столбца и 2-й строки для аллофона — R'<sub>012</sub> в слове «Брил», 2-го столбца и 3-й строки для аллофона — R'<sub>022</sub> в слове «Дрель». На спектрограмме ясно видны их существенные отличия.

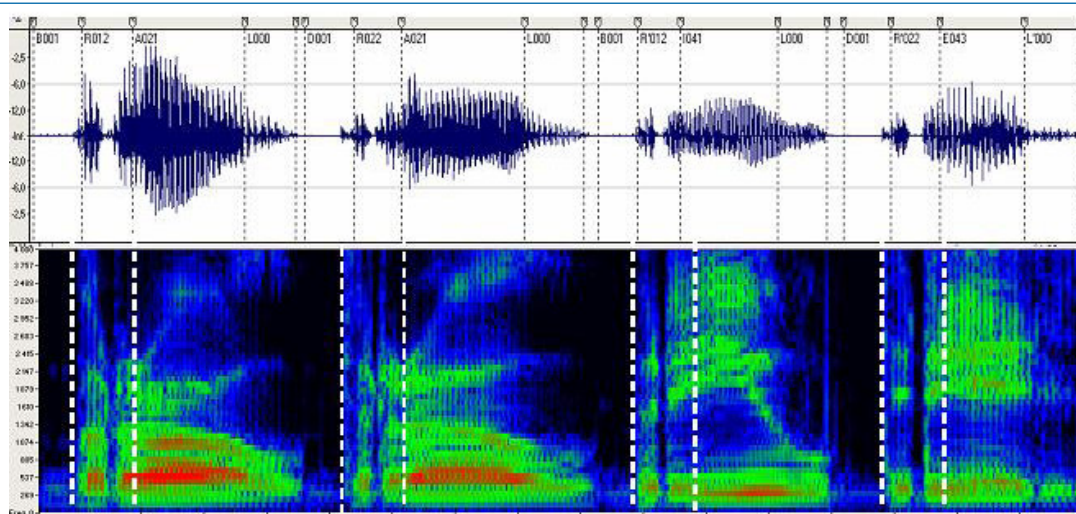


Рис. 9. Спектрограмма (внизу) и осциллограмма слов «Брал. Драл. Брил. Дрель»

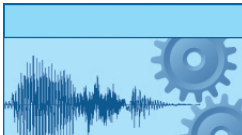
Общее количество аллофонов согласных фонем —  $N_{all}^c$ , подсчитываемое как их сумма аллофонов в каждой из 5-ти подгрупп (см. таблицы 1.10–1.14), равно:

- для русского языка  $N_{all}^c = 71 + 76 + 14 + 24 + 25 = 210$ .
- для белорусского языка  $N_{all}^c = 66 + 86 + 14 + 28 + 13 = 207$ .

Таким образом, несмотря на отличия числа аллофонов в разных подгруппах согласных русского и белорусского языков, их общее количество получилось примерно равным.

### 3. Алгоритм ПФА двуязычного синтезатора речи

Для записи описанных выше экспертных правил ПФА используется следующий способ: правила записаны в виде классификации групп центрального элемента, левого и правого



контекста, а также выходного значения центральных групп фонем для всех возможных вариаций контекстов. Для достижения оптимального соотношения производительности и потребления ресурсов экспертные правила преобразуются системой в машинные хэш-коды в процессе инициализации синтеза речи.

Ресурсы разработанного алгоритма ПФА имеют следующий вид:

**#Алфавит фонем**

A0,A1,A2,A3,E0,E1,E2,E3,I0,I1,I2,I3,O0,O1,Y0,Y1,Y2,Y3,U0,U1,U2,U3,B,B',D,D',G,G',  
 Z,Z',ZH,L,L',M,M',N,N',P,P',T,T',K,K',C,CH',F,F',S,S',H,H',SH,SH',V,V',J',R,R',C',  
 CH,SCH,GH,DZ',DZH,GH',W,#,

**#Центральные группы фонем**

C0: A0,A1,A2,A3,E0,E1,E2,E3,I0,I1,I2,I3,Y0,Y1,Y2,Y3,  
 C1: U0,U1,U2,U3,O0,O1,  
 C2: P,P',T,T',K',C,CH',F,F',S,S',SH,SH',H',C',CH,SCH,  
 ...  
 C6: R,R',

**#Группы фонем левого контекста**

L0: #,  
 L1: P,B,F,V,M,U0,U1,U2,U3,O0,O1,W,  
 L2: SH,ZH,R,T,C,S,D,Z,N,L,A0,A1,A2,A3,E0,E1,E2,E3,Y0,Y1,Y2,Y3,CH,SCH,DZH,  
 L3: K,G,H,GH,  
 ...  
 L11: B,B',D,D',G,G',Z,Z',ZH,L,L',M,M',N,N',P,P',T,T',K,K',C,CH',F,F',S,S',SH,SH',H,H',V,  
 V',J',R,R',C',CH,SCH,GH,DZ',DZH,GH',W,#,

**#Группы фонем правого контекста**

R0: #,  
 R1: P,B,F,V,M,L,W,O0,O1,U0,U1,U2,U3,  
 R2: SH,ZH,R,T,C,S,D,Z,N,K,G,H,A0,A1,A2,A3,E0,E1,E2,E3,Y0,Y1,Y2,Y3,CH,DZH,S,CH,  
 R3: P',T',K',CH',F',S',SH',H',G',B',D',Z',L',M',N',V',J',R',I0,I1,I2,I3,C',DZ',GH',  
 ...  
 R14: A0,A1,A2,A3,E0,E1,E2,E3,Y0,Y1,Y2,Y3,I0,I1,I2,I3,O0,O1,U0,U1,U2,U3,!

**#Правила преобразования для каждой центральной группы фонем**

C0	C1	C2
L0R0-00	L0R0-00	L10R0-00
L5R0-10	L1R0-10	L10R4-01
L6R0-20	L2R0-20	L10R5-02
...	...	...

*Алгоритм преобразования «Фонема–Аллофон» состоит из 5 этапов:*

1. Определение центральной, правой и левой фонем.
2. Определение первого индекса центральной фонемы.
3. Определение групп для правой контекстной фонемы.
4. Определение групп для левой контекстной фонемы.
5. Определение по группам для левой и правой фонем второго и третьего индекса центральной фонемы.



Для оптимизации вычисления входные ресурсы обрабатываются в процессе инициализации синтеза: контекстные группы записываются в хэш-таблицы с индексированными ключами названия групп, правила преобразования — в массивы с индексами левого и правого контекстов.

**Входные данные алгоритма:**

- 1) последовательность фонетических транскрипций синтагм входного текста Sph;
- 2) последовательность количества слогов в каждой синтагме Nsyl.

**Ресурсы алгоритма:**

1. Множество гласных фонем  $Vow = \{Vow_1, \dots, Vow_{NVow}\}$ , где  $NVow$  — количество гласных фонем; множество согласных фонем  $Con = \{Con_1, \dots, Con_{NCon}\}$ , где  $NCon$  — количество согласных фонем; множество разделителей  $D = \{D_1, \dots, D_{ND}\}$  (заметим, что  $D_1$  является символом начала или конца синтагмы), где  $ND$  — количество разделителей; множество обозначений словесных ударений  $Sstr = \{Sstr_1, Sstr_2\}$ ,  $Sstr_1$  — символ полноударности,  $Sstr_2$  — символ частичноударности.
2. Фонемные группы для центральных фонем  $C = \{\langle NameC_1, C_1 \rangle, \dots, \langle NameC_{NC}, C_{NC} \rangle\}$ ,  $m = 1, \dots, NC$ , где  $C_m = \{PhC_{m,r} \in \{VowUCon\} | r = 1, \dots, NC_m\}$ , где  $C_m$  — множество фонем группы,  $NameC_m$  — её уникальное название. Заметим, что фонемы  $PhC_{m,r}$  не могут использоваться в множестве  $C$  не больше одного раза.

Фонемные группы для левых фонем  $L = \{\langle NameL_1, L_1 \rangle, \dots, \langle NameL_n, L_{NL} \rangle\}$ ,  $n = 1, \dots, NL$ , где  $L_n = \{PhL_{n,s} \in \{VowUConUD\} | s = 1, \dots, NL_n\}$ , где  $L_n$  — множество фонем группы,  $NameL_n$  — её уникальное название.

Фонемные группы для правых фонем  $R = \{\langle NameR_1, R_1 \rangle, \dots, \langle NameR_{NR}, R_{NR} \rangle\}$ ,  $p = 1, \dots, NR$ , где  $R_p = \{PhR_{p,t} \in \{VowUConUD\} | t = 1, \dots, NR_p\}$ , где  $R_p$  — множество фонем группы,  $NameR_p$  — её уникальное название.

Заметим, что каждая группа фонем имеет уникальное название, например «C2», «L1», «R13». Первый элемент названия отражает принадлежность к месту расположения фонемы (C — группы для центральных фонем, L — группы для левых фонем, R — группы для правых фонем), второй — определённый состав фонем, входящих в данную группу.

3. Для каждой фонемной группы  $C_m$ ,  $m = 1, \dots, NC$  правила преобразования «Фонема—Аллофон»  $RA_m = \{\langle LR_1, Index_1 \rangle, \dots, \langle LR_{NRAm}, Index_{NRAm} \rangle\}$ ,  $NRAm$  — количество пар элементов правил  $LR = \{LR_1, \dots, LR_{NRAm}\}$  и  $Index = \{Index_1, \dots, Index_{NRAm}\}$ .

**Выходные данные:**

Последовательность аллофонных транскрипций синтагм SaI.

Функция  $F(LR_i, NameL_j, NameR_k)$  осуществляет проверку, равен ли первый элемент правил преобразования «Фонема—Аллофон»  $LR_i$  объединению названий групп для левых  $NameL_j$  и правых  $NameR_k$  фонем.

$$F(LR_i, NameL_j, NameR_k) = \begin{cases} 0, & LR_i \neq NameL_j \cup NameR_k; \\ 1, & LR_i = NameL_j \cup NameR_k. \end{cases}$$



**Структура алгоритма:**

**Шаг 1.** Полагаем, что  $Sph = \bigcup_{i=1}^{NSph} Sph_i$ ,  $Nsyl = \bigcup_{i=1}^{NSph} Nsyl_i$ ,  $NSph$  — количество синтагм с расписанными фонетическими транскрипциями.

**Шаг 2.** Для каждой синтагмы в фонемном виде  $Sph_i$ ,  $i=1, \dots, NSph$  (заметим, что  $Sph_i = \bigcup_{j=1}^{NPh} Ph_j$ ,  $NPh$  — количество фонем в синтагме) выполняем шаги 3–19. По завершении цикла переходим к шагу 20.

**Шаг 3.** Принимаем текущее количество слогов в синтагме  $Sph_i$  равным  $NsylT:=0$ ;

**Шаг 4.** Для каждой последовательности фонем  $Ph_{j-1}$ ,  $Ph_j$ ,  $Ph_{j+1}$  синтагмы  $Sph_i$  выполняем шаги 5–18, где  $j=NPh, \dots, 1$ . Причём должна выполняться следующая система условий:

$$\begin{cases} Ph_{j-1} := D_1, j-1 < 1; \\ Ph_{j+1} := D_1, j+1 > NPh. \end{cases}$$

Если фонемы в данной синтагме закончились, переходим к шагу 2.

**Шаг 5.** Если фонема  $Ph_j \in Vow$ , то переходим к шагу 6. Иначе переходим к шагу 10.

**Шаг 6.** Увеличиваем количество слогов в синтагме  $Sph_i$  на единицу  $NsylT:=NsylT+1$ .

**Шаг 7.** Если  $Ph_{j+1} = Sstr_1$  ( $Sstr_1=/+$ ), то фонема  $Ph_j$  является полноударной. Тогда первый индекс аллофона  $IndexA$  принимает значение 0, а именно  $IndexA:=0$ . Затем за фонему  $Ph_{j+1}$  принимаем фонему с индексом  $j+2$  с последовательности фонем  $Sph_i$ , а в последовательности фонем  $Sph_i$  удаляем фонему с индексом  $j+1$ . Общее количество фонем в последовательности  $Sph_i$  уменьшаем на 1,  $NPh:=NPh-1$ . Запоминаем, что текущая гласная  $Ph_j$  ударная через  $FlagSV:=1$ , и переходим к шагу 13.

**Шаг 8.** Если  $Ph_{j+1} = Sstr_2$  ( $Sstr_2=/=$ ), то фонема  $Ph_j$  является частичноударной. Тогда первый индекс аллофона  $IndexA$  принимает значение 1, а именно  $IndexA:=1$ . Затем за фонему  $Ph_{j+1}$  принимаем фонему с индексом  $j+2$  с последовательности фонем  $Sph_i$ , а в последовательности фонем  $Sph_i$  удаляем фонему с индексом  $j+1$ . Общее количество фонем в последовательности  $Sph_i$  уменьшаем на 1,  $NPh:=NPh-1$ . Запоминаем, что текущая гласная  $Ph_j$  ударная через  $FlagSV:=1$ , и переходим к шагу 13.

**Шаг 9.** Если  $FlagSV = 0$  и фонема  $Ph_j$  не находится в крайнем левом  $NsylT \neq 1$  или крайнем правом  $NsylT \neq Nsyl_i$  слоге, то данная фонема  $Ph_j$  является безударной гласной второй степени редукции, тогда первый индекс аллофона  $IndexA$  принимает значение 3, а именно  $IndexA:=3$ . Иначе, она является безударной гласной первой степени редукции, тогда первый индекс аллофона  $IndexA$  принимает значение 2, а именно  $IndexA:=2$ . Переходим к шагу 13.

**Шаг 10.** Если данная фонема согласная, т.е.  $Ph_j \in Con$ , то переходим к шагу 11. Иначе переходим к шагу 12.

**Шаг 11.** Если фонема  $Ph_j$  равна первой левой фонеме  $Ph_{j-1}$  относительно её ( $Ph_j = Ph_{j-1}$ ), то она считается удвоенной, т.е.  $IndexA:=1$ . Тогда за фонему

$Ph_{j-1}$  принимаем фонему с индексом  $j-2$  с последовательности фонем  $Sph_i$ , а в последовательности фонем  $Sph_i$  удаляем фонему с индексом  $j-1$ , общую нумерацию текущей фонемы уменьшаем на 1, т.е.  $j:=j-1$ , ибо общее количество фонем уменьшилось на 1 в последовательности фонем  $Sph_i$ ,  $NPh:=NPh-1$ .

Иначе, данную фонему считаем одинарной  $IndexA:=0$ .

Переходим к шагу 13.

**Шаг 12.** Если  $Ph_j$  является разделительным элементом, т.е.  $Ph_j \in D$ , , то запоминаем его во временной переменной  $TEMP:=Ph_j$  и переходим к шагу 18.

**Шаг 13.** Определяем множество фонемных групп, в которые входит центральный элемент  $Ph_j$ :

$$Cx = \{ \langle NameC_m, C_m \rangle | PhC_{m,r} \in C_m, C_m \in C, m=1, \dots, NC, r=1, \dots, NC_m, \text{ при условии } PhC_{m,r} = Ph_j \}$$

Учитывая формат правил преобразования «буква-фонема», множество  $Cx$  состоит из одного элемента. Запоминаем индекс  $m$ , при котором выполнилось условие  $PhC_{m,r} = Ph_j$ , т.е.  $iCx:=m$ .

По индексу  $iCx$  находим соответствующие этому множеству правила преобразования  $RA_{iCx}$ .

Определяем множество фонемных групп, в которые входит левый элемент  $Ph_{j-1}$ :

$$Lx = \{ \langle NameL_n, L_n \rangle | PhL_{n,s} \in L_n, L_n \in L, n=1, \dots, NL, s=1, \dots, NL_n, \text{ при условии } PhL_{n,s} = Ph_{j-1} \}$$

Определяем множество фонемных групп, в которые входит правый элемент  $Ph_{j+1}$ :

$$Rx = \{ \langle NameR_p, R_p \rangle | PhR_{p,t} \in R_p, R_p \in R, n=1, \dots, NR, t=1, \dots, NR_p, \text{ при условии } PhR_{p,t} = Ph_{j+1} \}$$

**Шаг 14.** Для каждой фонемной группы левого контекста  $Lx_{kj}$ ,  $Lx = \{Lx_1, \dots, Lx_{NLx}\}$ ,  $k1=1, \dots, NLx$ , где  $NLx$  — количество групп для левого контекста  $Ph_{j-1}$ , выполняем шаги 15–16. По завершении цикла переходим к шагу 17.

**Шаг 15.** Для каждой фонемной группы правого контекста  $Rx_{k2}$ ,  $Rx = \{Rx_1, \dots, Rx_{NRx}\}$ ,  $k2=1, \dots, NRx$ , где  $NRx$  — количество групп для правого контекста  $Ph_{j+1}$ , выполняем шаг 16.

**Шаг 16.** Для первого элемента  $LR_q$  каждого правила  $RA_{iCx}$ ,  $q=1, \dots, NRA_{iCx}$  проверяем значение функции  $F(LR_q, NameL_{k1}, NameR_{k2})$ . Если  $F(LR_q, NameL_{k1}, NameR_{k2})=1$ , то индекс  $q$  определяет соответствующий второй элемент правила  $RA_{iCx} - Index_q$ . Этот элемент  $Index_q$  является вторым и третьим индексом аллофона  $IndexB:=Index_q$ .

**Шаг 17.** Находим полное значение аллофона  $TEMP:=Ph_j \cup IndexA \cup IndexB$

**Шаг 18.** Добавляем  $TEMP$  к аллофонной транскрипции синтагмы  $Sal_i:=TEMP \cup Sal_i$  и переходим к шагу 4.

**Шаг 19.** Добавляем аллофонную транскрипцию синтагмы к выходному аллофонному тексту  $Sal:=Sal \cup Sal_i$ , и переходим к шагу 2.

**Шаг 20.** Конец алгоритма.

## Заключение

Сходство русского и белорусских языков позволило создать универсальный двуязычный алгоритм ПФА, в котором фонетические правила преобразования вынесены в отдельные ресурсы. Благодаря этому реализуется возможность достаточно простого изменения (уточнения) этих правил экспертами-лингвистами.



Для тестирования алгоритма ПФА были собраны корпуса для русского и белорусского языков из электронных словарей, в которых собраны слова со всеми возможными в данных языках 5-ти символьными комбинациями. Для русского языка корпус составил 101958 слов, для белорусского — 102169. Результаты тестирования для двух языков приведены в таблице 10.

Таблица 10

#### Результаты тестирования алгоритма

	Русский язык	Белорусский язык
Количество слов	101 958	102 169
Количество аллофонов	897 635	934 583
Количество правильных аллофонов	871 989	903 216
Процентное отношение корректно преобразованных аллофонов	97.143%	96.644%

Алгоритм на основе регулярных правил, реализованный в синтезаторе речи «МультиФон» и его мобильной реализации [5], позволил достичь достаточно высокой точности ПФА (около 97% правильно преобразованных фонем для русского и белорусского языков).

#### Литература

1. Матусевич М.И. Современный русский язык. Фонетика. М., 1976.
2. Лобанов Б.М., Пьорковска Б., Рафалко Я., Цирульник Л.И., Шпилевский Э. Фонетикоакустическая база данных для многоязычного синтеза речи по тексту на славянских языках // Компьютерная лингвистика и интеллектуальные технологии: труды междунар. конф. Диалог'2006, Бекасово, 31 мая — 4 июня 2006 г. / Институт проблем информатики РАН; отв. ред.: Н.И. Лауфер [и др.]. М.: Наука, 2006. С. 357–363.
3. Зиновьева Н.В., Кривнова О.Ф. Прикладные системы с использованием фонетических знаний // Проблемы фонетики: сб. статей. М.: Прометей, 1993. С. 288–300.
4. Лобанов Б.М., Цирульник Л.И. Компьютерный синтез и клонирование речи. Мн.: Белорусская наука, 2008.
5. Цирульник Л.И., Покладок Д.А. Система синтеза речи по тексту для мобильных телефонов // Речевые технологии. 2010. № 1. С. 83–91.