

Определяне фонологичните варианти на фонемите в съвременния български език с оглед компютърна обработка на текст

Десислава Баева, Димитрина Игнатова-Цонева

Defining the Phonological Variants of Contemporary Bulgarian Phonemes with Respect to the Computer Processing of Text: The aim of the present paper is to develop a framework for the number of phonological variants of contemporary Bulgarian vowel and consonant phonemes with respect to the computer processing of speech.

Key words: *allophones, distribution of consonants and vowels, segment level*

Фонемата като най-малката сегментна единица на речта служи като различител на морфемите и словоформи. В речта тя се реализира под формата на различни варианти, например: като произнасяме думите *лапа, лале, лук, лък, лик, кална, кълвач, дол, рало, жал, Лиляна*, имаме впечатлението, че във всяка от тези думи казваме все едно и също [л], но при звукозапис или при малка пауза пред следващия звук и внимателно и търпеливо вслушване, можем да установим, че в действителност произнасяме поредица от фонетично подобни звукове: [л₁, л₂, л₃, л₄, л₅, л₆, л₇, л₈ и т.н.]

Същото се наблюдава и с гласната фонема [а] в думата *разплàкала* :

[а] **под ударение** има най-голяма дължина и сила, височината му е еднаква с [а] пред и след ударение и равна мелодия с низходящи модуляции;

[а] **пред ударение** е второ по дължина след удареното [а], еднакво по сила с него и със слабионизходяща мелодия. Долавя се и известна зависимост между учленението на съгласната [р] и силата на предудареното [а], тъй като съгласната [р] е сонорна и засилва изговора на следващата гласна;

[а] **след ударение** има най-малка дължина, по сила е по-слабо от предходните две [а]-та, по височина е еднакво с удареното [а] и има равна мелодия;

Крайното [а] има по-голяма дължина от следудареното [а] и по-малка от останалите две, силата и височината му са еднакви със следудареното [а] и е с леко възходяща мелодия. [Георгиева 1968: 567].

Излиза, че в думата *разплàкала* има четири различни нюанса на фонемата [а].

Видовете (нюансите) звукове, които можем да установим при внимателно наблюдение, наричаме **фони**.

Когато се правят изследвания на някой говорим език, най-напред се определят отделните фони. След това се установява в каква позиция и в какво звуково обкръжение се срещат, каква е тяхната дистрибуция (разпределение). За съвременния български език например установяваме, че съществуват фони [**В₁**, **В₂**, **В₃**, **В₄**, **В₅**,], които се срещат в определени позиции, например: [**В₁**]. - пред задни (твърди) гласни в началото и в средата на думата – *ваза, вар, вода, вуйчо, върна*; [**В₂**]. - пред предни (меки) гласни – *вея, верен, вихър, вир*; [**В₃**]. - пред сонорни съгласни – *време, внедря, вместо, влюбя*; [**В₄**]. - пред [**й +а, ъ, о, у**] – *вяра, в юли*. Тези фони взаимно се изключват в една и съща позиция, т.е. там, където се среща [**В₁**], не се среща [**В₂**], нито [**В₃**] или [**В₄**]. Те са **комплементарни (разпределени)** – разпределят така позициите си, че там, където се среща единият фон, другият или другите не се срещат, и там, където се среща другият или другите фони – първият фон не се появява. Въпреки това те са фонетично подобни и принадлежат към една и съща фонема – фонемата [**в**].

Фоните, които са фонетично подобни и са комплементарни, се наричат **алофони (комбинаторни варианти, позиционно обусловени варианти, позиционни варианти)** [Манголд 1988: 91].

Броят на фоните на една фонема зависи от характера на изследването и от точността, с която те се отчитат. В действителност те могат да бъдат стотици.

Най-общо многообразието от конкретни варианти, с които се реализират фонемите в речта, може да бъде сведено до два основни типа – **фонологични и екстрафонологични варианти на фонемата**.

I. Фонологични варианти на фонемата. Имат задължителна реализация, защото са предопределени от автоматизираните учленителни движения, съставляващи правоговорната норма на съответния език. [Вътов 2002: 169]. Те са два вида: **основен вариант и специфични варианти**.

1. **Основният вариант** се осъществява в най-благоприятни произносителни условия. Може да се произнесе и изолирано.

За гласните фонеме в съвременния български език основни варианти са **в позиция под ударение**, например: *лак, лък, век, вик, учен, очен и др.*

За съгласните основните варианти са пред задните (твърдите) гласни в началото или в средата на думата, например: *ла̀ра, ба̀ра, фа̀р, ва̀р, та̀м, да̀м, съ̀н, зъ̀н, цо̀л, жа̀л, ша̀л, джа̀с, ча̀с, ко̀ст, го̀ст, ку̀ха, ку̀ка, ма̀й, на̀й, лу̀к, сро̀к* и т.н.

2. **Специфичните варианти** съдържат и някои временни признаци, комбинаторно или позиционно обусловени. Те се реализират в произносителни условия, при които съседните фонеме влияят върху дадената фонема или позицията им спрямо ударението ги променя в известна степен.

Специфични алофони са гласните пред или след ударение или в края на думата: *оста̀-уста̀, перо̀н-пиро̀н, преда̀ва-прида̀ва, веко̀ве-вико̀ве, пи́ле-пи́пи* и т.н.

Специфични варианти за съгласната [в] са например: *ви́е, ведро̀, вре́ме, вял, пра́вя, ви́я, врэ́е, прэ́е* и т.н.

II. Екстрафонологични варианти на фонемата. Различни фактори повлияват за появата на извънфонологичните варианти – възрастови особености, диалектно или чуждоезиково влияние, анатомични особености при артикулирането и др.

Екстрафонологичните варианти се делят на четири групи [Вътов 2002: 171]:

1. **Факултативни варианти** – възникват от въздействието на различни фактори – свързани са с изговорните стилове или са резултат от диалектно влияние в книжовния изговор, напр: [дитету́] вм.[детето́], [зна́йми] вм. [зна́ем]; [мо́йте, тво́йте] вм. [мо́ите, тво́ите], [Сòвийа] вм. Сòфия
2. **Индивидуални варианти** – свързани са със специфична артикулация или някои други логопатни проблеми, напр. [Уайна] или [Айна] вм. [Райна];
3. **Акцентни варианти** – получават се при интерференция, когато родният (изходният) език оказва влияние на втория език, [мётивация́] вм. [мотивация́] и т.н. при рускоговорещите.
4. **Стилистични варианти** – дължат се на емоционалната нагласа на говорещия, който умишлено натоварва една или друга фонема с по-особена квантитетно – динамична особеност, например удължаване на гласните за изразяване на възхищение или упрек: Оооох! Виикам, виикам напрасно! Хууубава работа!

Проверката дали в дадена позиция функционира една и съща фонема или две различни фонеме става чрез **метода на комутацията**, т.е. на взаимната замяна на звуковите сегменти. Ако от взаимната замяна се е получила нова словоформа, налице са две различни фонеме и обратно – ако словоформата остава една и съща, налице са фонема и вариант. Така например в *глас* и *гласо̀ве* ударената [а] в

първата дума и неударената гласна [а] във втората дума във фонетично отношение са твърде различни – гласежът на неударената [а] е по-близък до гласежа на гласната [ъ], отколкото до гласежа на ударената гласна [а]. Но тяхната взаимна замяна не води до образуване на нови, различни словоформи. Дали ще изговорим гласната [а] в думата *гласове* с гласежа ѝ под ударение или без ударение (редуцирано или нередуцирано), няма никакво значение, защото и в двата случая словоформата си остава една и съща. Потъмнелият гласеж на гласната [а], близък много до гласежа на [ъ], е в резултат на предударената ѝ позиция. В различна степен тази ситуация се среща и при останалите гласни в българския език. [Тилков, Бояджиев 1977: 38]

Доколко и дали трябва да се отчитат всички варианти на фонемите и да се установява съответният брой алофони, зависи от характера на описанието.

Целта на настоящото изследване е да бъде изградена концепция за броя на фонологичните варианти при гласните и при съгласните фонемите в съвременния български език с оглед компютърно моделиране на реч.

I. Фонологични варианти при гласните фонемите

Най-важната особеност на българския вокализъм е реализирането и функционирането на гласните фонемите в зависимост от ударението. Под ударение системата действа с максималния си брой фонемите, а извън него в резултат на неутрализацията на признака компактност – с намален брой. Освен това извън ударението гласните се поддават не само на количествена, но и на качествена промяна, изразена в различните степени на звучене на неударения вариант в зависимост от позицията, която той заема спрямо ударението. [Тилков 1975: 3]

В артикулационно и акустично отношение българското ударение е силно централизиращо, поради което се създава силен контраст между ударени и неударени срички. Също така то е свободно и подвижно. С тези негови особености е свързана и редукцията на неударените български гласни. Тя се проявява преди всичко в движението на отворените гласни [а, о, е] към корелативните [ъ, у, и].

Редукцията е на принципа на прогресивното затихване: колкото неударената гласна е по-близо до ударението, толкова редукцията е по-слаба, и обратно – колкото гласната е по-далече от ударението, толкова по-силно се проявява редукцията.

В книжовния изговор на гласните редукцията засяга само [а] и [о], които в неударена позиция се приближават до [ъ] и [у], без обаче да се покриват изцяло с тях: *камбаната* [кѐмбѐнѐтѐ], *сладкарница* [слѐдкѐрницѐ], *закачалка* [зѐкѐчѐлкѐ] и др.

В зависимост от позицията си спрямо ударението гласната [о] се реализира под формата на два алофона: в първа предударена позиция – с по-слаба затвореност, а във втора предударена и в следударена позиция – със сравнително по-голяма степен на затваряне, например: *кожух* [кѐжѹх], *колело* [кѐлѐлѐ], *мѐлко* [мѐлкѐ] и т.н.

Само гласната [е] остава незасегната. Тя дори в неударена позиция запазва своите темброви качества, например: *вечѐрен*, *пѐсенен*, *предѐше* и др.

Както се вижда от таблицата, гласните [е], [и], [у] и [ъ] почти запазват качества си както в ударена, така и в наударена позиция, а [а] и [о] ги губят и понякога се изравняват с [ъ] и [у], за да се реализират под формата на **архифонемите** [Трубецкой 1960]. **Архифонемата (сврѣхфонемата)** е съвкупност от смислоразличителни признаци, които са общи за две фонемите. Неутрализиращите се опозиции пораждат архифонемите.

В своите изследвания М. Георгиева [Георгиева 1968: 574], като сравнява особеностите на всички гласни в българската литературна реч, посочва следните позиционни варианти на гласните:

I. Ударена гласна – има най-голяма дължина, сила и височина и равна или възходяща мелодия с възходящи или низходящи модуляции.

II. Гласна пред ударение – има по-малка дължина, сила и височина от ударената гласна и възходяща или равна мелодия с възходящи или низходящи модуляции.

III. Гласна след ударение – с най-малка дължина и сила и равна или низходяща мелодия, която започва с височината на ударената гласна или е равна на крайната гласна.

IV. Гласна в края на думата – най-често има дължина по-малка от ударената гласна (понякога равна с нея), сила – най-малка или равна на гласната след или пред ударението и низходяща, равна или възходяща мелодия, която започва с височината или под височината на ударената или следударената гласна, с възходящи или низходящи модуляции.

С оглед на подготовката на фонетичен материал за създаване на програмни средства за анализ на компютърен български текст (за целите на преобразуването текст – звук) в настоящото изследване се предлагат следните **съществени позиции за реализация на българските гласни фонemi**:

I. **Позиция под ударение;**

II. **Предударена позиция;**

III. **Следударена позиция.**

Те важат за всички фонемни противопоставяния в българската вокална система. Позициите за реализация на трите двойки фонемни опозиции [о :: у], [а :: ъ] и [е :: и] могат да бъдат представени в следната таблица:

Таблица 1. Позиционна реализация на вокалните фонemi в българския език

Фонemi	[а]	[ъ]	[о]	[у]	[е]	[и]
I-ва позиция (под ударение)	+	+	+	+	+	+
II-ра позиция (пред ударена сричка)	-	+	+	+	+	+
III-та позиция (след ударена сричка)	-	+		+	+	+

От таблицата се вижда, че елементарните звукови сегменти (алофони), които съставят българската вокална система, са 15.

A. Примери за реализация на гласните фонemi в опозиционната двойка [а::ъ]:

I позиция: пак – пѣк, лак – лѣк, прав – прѣв, плат – плѣт;

B. Примери за реализация на гласните фонemi в опозиционната двойка [о::у]:

I позиция: оѣен – ѹѣен, тѣѣен – тѹѣен, мѣѣен – мѹѣен, тоѣ – тѹѣ;

II позиция: остѣ – устѣ, копѣ – купѣ;

B. Примери за реализация на гласните фонemi в опозиционната двойка [е::и]:

I позиция: пѣѣ – пѹѣ, фѣѣрма – фѹѣрма, нѣѣѣ – нѹѣѣѣ, ѣѣѣст – ѣѹѣѣст;

II позиция: перѣѣ – пирѣѣ, веѣѣѣ – виѣѣѣѣ, предѣѣѣ – придѣѣѣ;

III позиция: покѣѣнѣх – покѹѣѣнѣх, отвѣѣѣѣрѣх – отвѹѣѣѣѣрѣх, пѣѣѣ – пѹѣѣѣ.

II. Фонологични варианти при съгласните фонемите

Реализацията на съгласните значително се отличава от реализацията на гласните фонемите. В потока на речта съгласните не могат да се изговарят поотделно и изолирано една от друга. Те взаимно си влияят, като променят своите учленителни и акустични особености, без да променят фонологическите си характеристики. При един прецизен анализ могат да се отделят хиляди фони на съгласните фонемите, но като се имат предвид съчетанията на съгласните със звуковете, които са след тях, могат да се установят две позиции на по-значителни комбинаторни промени в рамките на думата [Тилков, Бояджиев 1977: 139]: **Съчетание на съгласна + гласна** и **съчетание на съгласна + съгласна**.

I. Съчетание на съгласна + гласна

1. Както бе посочено, за съгласните **основните варианти са пред задните (твърдите) гласни [а, ъ, о, у]** в началото или в средата на думата.

Интересно явление се наблюдава пред лабиалните гласни [о, у]. Участието на устните придава закръглен лабиален изговор, т.е. изтеглянето и закръглянето на устния отвор се извършва предварително, още при учленението на съгласните, които стоят пред тях, например: *аккумулятор, гугутка, кука* и др. Този процес се нарича **коартикуляция**. Лабиализацията на съгласните се извършва само в определени фонетични позиции и фонологически е несъществена.

2. В позиция пред [е, и] съгласните се **палатализират**. Разбира се, мекостта на консонантите пред предните вокали в книжовния изговор е слаба. Алофоните в тези позиции се определят като полумеки: *месеци, седеше, песните* и др.

II. Съчетание на съгласна + съгласна - асимилативно взаимно приспособяване между консонантите по място и начин на учленение.

1. **В позиция пред сонорните съгласни [л, м, р, н]:** *разпрѐ-разбрѐ, брѐви-прѐви, брѐме-врѐме, брак-фрак, влѐга-флѐга, враг-фраг*.

2. **Пред [в]:** *дворѐц-творѐц, пѐзва-пѐсва, звѐнѐц-свинѐц, цвят-свят, цвѐтна-свѐтна*. „По-особеното въздействие на съгласната [в] върху съседните шумови съгласни се дължи на исторически причини и може да се обясни с някогашното ѝ билабиално сонорно учленение. И днес съгласната се отличава от останалите шумови съгласни с изразения си сонорен характер – спектърът ѝ има добре очертана формантна структура, подобно на сонорните съгласни. Едва след като се променя начинът на учленяването ѝ, тя преминава към шумовите съгласни, а по-късно се развива и беззвучният ѝ съответник – фонемата [ф].- в структурата на българските думи...» [Вътов 2002: 202].

3. **В абсолютен край на думата.** В краесловието звучните фонемите се обеззвучават, т.е. **реализират се с архифонемите:** [грат Русе], [Вътоф], [корал], [гняф] и др. Обеззвучането става дори ако в края на думата стоят две звучни съгласни: [гросст или грост] вм. *грозд*, [вошт или вош] вм. *вожд*

4. В настоящото изследване приемаме положението на твърда съгласна пред глайд за още една **позиция за реализация на съгласните – пред [й], последван от твърдите вокали [а, ъ, о, у]**, напр.: *гювеч [гйувеч], люляк [лйуляк], мярка [мйарка], няма [нйама]* и др.

5. **Пред беззвучна или звучна съгласна.** Когато в началото или в средисловието застанат една до друга различни по звучност шумови съгласни, те се изравняват под въздействието на регресивна асимилация:

А) Пред беззвучна съгласна: *ножка* → [ношкѐ],

Б) Пред звучна съгласна: *сграда* → [зграда], *сватба* → [свадба]

В действителност позицията на шумовите съгласни пред шумови съгласни с противоположна звучност е функционално слаба позиция. В тази позиция те или се

озвучават, или губят звучността си, като се реализират в архифонемни, т.е. във варианти, общи за съответните им корелати по качеството звучност. [Вътов 2002: 169].:

Таблица 2. Позиционна реализация на съгласните в българския език

Фонемни	б	п	в	ф	д	т	з	с	дз	ц	ж	ш	дж	ч	г	к	х	м	л	р	н	й
I позиция	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
- пред [а, ъ, о, у]																						
II позиция	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
- пред [е, и]																						
III позиция	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
- пред сонорни съгл.																						
IV позиция	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
- пред [в]																						
V позиция	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	+	+	+	+	+	+
- в края на думата																						
VI позиция	+	+	+	+	+	+	+	+	-	+	-	-	-	-	+	+	-	+	+	+	+	-
-пред [й + а, ъ, о, у]																						
VII позиция	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	+	+	+	+
-пред беззв. съгласни																						
VIII позиция	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	+	+	+	+	+
-пред звучни съгласни																						

Изводи:

1. Към петте съществени позиции за реализация на съгласните фонемни:

I. Пред задни гласни;

II. Пред предни (меки) гласни;

III. Пред сонорните съгласни [л, р, м, н];

IV. Пред съгласната [в];

V. В края на думата

трябва да се добавят още три:

VI. – пред глайд + вокал . [а, ъ, о, у],

VII. Шумова съгласна пред беззвучна съгласна

VIII. Шумова съгласна пред звучна съгласна в началото или в средисловието

2. Независимо че в традиционното българско езикознание и във всички съвременни учебници и учебни помагала се посочва, че българската фонемна система се състои от 6 гласни и 39 съгласни, като приемем факта, че в съвременния

български език само [й] е запазила своята палаталност, налага се броят им да бъде прецизиран – **на 6 гласни и 22 съгласни**. Това е особено необходимо не само за обучението по фонология, за писмената практика, за работата по формиране на фонематичен слух, но и за информационните технологии и по-точно за разработването на програми за компютърно разпознаване на българска реч.

3. В човешката реч фонемата се реализира чрез позиционно обусловени варианти (алофони), които са подобни и се срещат комплементарно. Явяват се така, че там, където се среща единият алофон, другият или другите не се срещат и там, където се среща другият или другите алофони, първият алофон не се среща. Броят на алофоните, които изграждат една фонема, зависи от точността, с която те се описват фонетично. Доколко трябва да се отчитат всичките фини различия, зависи от целта на описанието. По наше мнение за компютърната обработка на българска реч трябва да се кодират основните позиционни варианти на гласните и съгласните фонемни и чрез набор от специално подбрани лексеми да се установи акустичната характеристика на всеки един алофон.

4. На сегментно равнище българските гласни и съгласни се реализират от общо 159 комплементарни алофона (**15 за гласните и 144 за съгласните**).

Както при формирането на „еталонни зони“ в паметта на човека за родните или чуждоезиковите звукове, така и при компютърното разпознаване на текст в паметта на програмата трябва да бъдат фиксирани моделите, еталоните на звуковете, които участват в речта на ползвателя. Това е необходимо, за да може програмата да приема всеки отделен алофон в потока на речта и да го съотнася към правилния еталонен образ, т.е. към съответната фонема. След като бъде разпознат и идентифициран „слуховият образ“, той трябва да бъде обозначен с подходяща графема.

Разбира се, при разработване на компютърни програми за разпознаване на реч вниманието трябва да бъде насочено не само към различаване, идентифициране и правилно съотнасяне на речевите звукове към конкретни фонемни на езика, но трябва да се вземат под внимание и суперсегментните елементи, за които е необходимо допълнително кодиране.

ЛИТЕРАТУРА

Вътов 2002: Вътов, В. Фонетика и фонология на българския език. В.Търново, ВТУ „Св. св.Кирил и Методий“, 2002.

Георгиева 1968: Георгиева, М. Позиционни варианти на гласните в българската литературна реч. – Изв. На Инст. За бълг.език., №16, 1967, С., 1968, с.565 – 576.

Манголд 1988: Манголд, М. Увод в езикознанието с оглед и на българския език. София, СУ „Кл.Охридски“, 1988.

Тилков 1975: Тилков, Д. Вокални фонемни варианти в книжовния български език –Български език, 1975, №1, с.3 – 8.

Тилков, Бояджиев 1977: Тилков, Д., Т. Бояджиев. Българска фонетика. Наука и изкуство, София, 1977.

Трубецкой 1960: Трубецкой, Н.С. Основы фонологии. М., 1960, с.130.

За контакти:

Доц. д-р Димитрина Цонева, Катедра *Български език, литература и изкуство*, Русенски университет „Ангел Кънчев“, тел.: 082-888 738, e-mail: dtzoneva@uni-rise.bg

Ас. Десислава Баева, докторант към Катедра *Информатика и информационни технологии*, Русенски университет „Ангел Кънчев“, тел.: 082-888 214, e-mail: dbaeva@ami.uni-ruse.bg

Докладът е рецензиран.

РУСЕНСКИ УНИВЕРСИТЕТ „АНГЕЛ КЪНЧЕВ“
UNIVERSITY OF RUSE „ANGEL KANCHEV“

ДИПЛОМА

Програмният комитет на
Научната конференция RU&SU'11
награждава с КРИСТАЛЕН ПРИЗ
“THE BEST PAPER”

ас. ДЕСИСЛАВА БАЕВА, докторант,
доц. д-р ДИМИТРИНА ЦОНЕВА,
автори на доклада

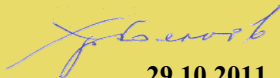
“Определяне фонологичните варианти на
фонемите в съвременния български език
с оглед компютърна обработка на текст”.

DIPLOMA

The Programme Committee of
the Scientific Conference RU&SU'11
awards the Crystal Prize "THE BEST PAPER"
to Assist. Prof. DESISLAVA BAEVA, PhD Student,
Assoc. Prof. DIMITRINA TSONEVA, PhD,
authors of the paper

“Defining the Phonological Variants of
Contemporary Bulgarian Phonemes with Respect
to the Computer Processing of Text”.

РЕКТОР
RECTOR



Проф. д.т.н. Христо Белоев
Prof. DSc Hristo Beloev

29.10.2011