

**АКУСТИЧНО МОДЕЛИРАНЕ НА ГЛАСНИТЕ
С ОТЧИТАНЕ НА ПЕРЦЕПТИВНАТА ИНТЕГРАЦИЯ
НА ФОРМАНТНИТЕ ЗОНИ**

Руси Николов, Маджед Хусеин***

**Пловдивски университет „Паусий Хилендарски“*

***Медицински университет – Пловдив*

**ACOUSTIC MODELING OF VOWELS BY TAKING INTO
ACCOUNT
THE PERCEPTUAL INTEGRATION OF THE FORMANT AREAS**

Rousi Nikolov, Madzhed Husein***

** Paisii Hilendarski University of Plovdiv*

***Medical University of Plovdiv*

A new approach is presented here, as well as a new software tool for plotting vowels – a technique as simple as the F2/F1 vowel space mapping but perceptually more adequate. The formants taken into account in acoustic analysis of vowels are more often reduced to their two first distinctive frequency components. However, the role of F3 is known to be significant in the perception of certain vowel contrasts, and this role is particularly important when it comes to contrastive analysis of Bulgarian and French vowels. The new technique was applied to account for the interactions between prosthetic treatment and speech function in Bulgarian-speaking patients.

***Key words:** Acoustic Analysis, Third Vowel Formant, NLP Software Tools, Prosthetic Dentistry*

В настоящата статия е представен нов акустико-функционален модел на гласните звукове. Моделът е акустичен, защото гласните се представят въз основа на акустични данни, той е и функционален, защото отчита функционалната, в частност – фонологичната стойност на различните акустични параметри на гласните. Фонологичната стойност на един акустичен параметър се определя от неговата перцептивна тежест в процеса на перцептивен анализ на звука и идентифициране на съответната фонема.

Необходимостта от подобряване на методите за акустичен анализ на гласните е обоснована експериментално в статията *Значение на третия формант в сравнителния анализ на българските и френските гласни* (Николов 2016: 42 – 47). Тук ще бъдат обосновани, концептуално и експериментално, адекватността и целесъобразността на предлагания модел, като се оценяват експерименталните резултати (т.е. гласните, представени с помощта на този модел) в светлината на следните два критерия:

Критерий 1. Един акустичен модел на гласните може да бъде определен като акустико-фонетичен (т.е. функционален акустичен модел с фонологична стойност), при условие че **резултатите от приложението на модела отговарят в достатъчна степен на резултатите от съответните перцептивни оценки**. Когато при един твърде опростен модел се стига на практика до явни парадокси в това отношение, моделът е просто неадекватен (каквото се оказва понякога двуформантният модел); обратно, когато моделът е сложен (в граничния случай – спектърът на звука) и е трудно или невъзможно да се интерпретира от функционална гледна точка, моделът остава само акустичен, или по-общо казано – физичен.

Този критерий (или изискване) не отменя практическата стойност на акустико-фонетичния анализ – със своите бързи, лесни и обективни резултати той е по-достъпен и по-продуктивен от съответните перцептивни тестове.

Критерий 2. Кодовете за устна комуникация (каквито са звуковите подсистеми на естествените езици) са най-използваните и са претърпели най-продължителна оптимизация в процеса на тяхната еволюция, което ги прави надеждни и ефективни. Тази ефективност може да бъде оценена количествено с помощта на два показателя: 1) разстоянието (контрастивността) между звуковите единици: по-голямото разстояние осигурява по-надеждна разграничимост на елементите на кода; 2) вариативността на звуковите единици по отношение на различните говорители: езиково релевантните звукови характеристики са толкова порелевантни, колкото по-близки са тези характеристики за различните говорители. Тези съображения ни позволяват да дефинираме критерий 2 като комбинация от следните два подкритерия: 2.1. максимална фонетична разлика между различните фонемите и 2.2. минимална фонетична разлика за всяка отделна фонема в реализациите на различните говорители. Ако за един модел и двата подкритерия (или с други думи – принципа за максимална контрастивност и минимална вариативност) са изпълнени в по-голяма степен спрямо друг модел, ще приемем това като

допълнителен показател, че първият модел е по-адекватен във функционално отношение от втория модел.

Подкритерият 2.2. за максимална парадигматична контрастивност при фонемите е логичен от функционална гледна точка. Не е случайно например, че кардиналните гласни в международната фонетична азбука – тези, които очертават акустичния трапец, са екстремните гласни във френския език – език с особено богата вокална система. По същата логика обаче трябва да очакваме, че даден фонетичен контраст може да се неутрализира, или по-общо казано – ще зависи от производителността на съответната фонологична опозиция. Прилагайки новия модел, ще установим например, че разстоянието между френските гласни /i/ и /e/ е значително по-голямо от това между /e/ и /ɛ/; обяснението е просто, ако се позоваваме на почти пълната неутрализация на опозицията между полузатворената и полуотворената гласна /E/ в съвременния френски език. Всички двойки френски гласни от тези две категории (/e/ и /ɛ/, /ø/ и /œ/, /o/ и /ɔ/) функционират в голяма степен като позиционни варианти, което позволява свиване на перцептивното пространство в тяхната зона.

Приложението на горните критерии за оценка на резултатите от традиционните методи за акустично представяне на гласните разкрива съществени слабости, специфични за съответните модели и касаещи тяхната функционалната адекватност (т.е. фонетичната им релевантност) и тяхната практическа стойност за непосредствена оценка на гласните (т.е. тяхната информативност). Съображенията по-долу обосновават това твърдение.

Акустико-фонетичните модели на гласните се основават на тяхната т.нар. формантна структура, но на практика те отчитат единствено пиковите стойности на формантите (включително моделите с отчитане на перцептивната интеграция на формантите). При този опростен подход се пренебрегва разпределението (най-често асиметрично) на акустичната енергия в рамките на формантните зони. Липсата на прецизност се задълбочава още повече поради несъвършенството на алгоритмите за автоматично определяне на пиковите стойности на формантите. От друга страна, при директна оценка или съпоставка на спектрални обвивки акустичният образ е максимално изчерпателен (спектърът съдържа наистина пълна информация за звука в даден момент от времето), но той подлежи трудно на интерпретация както на артикулационно, така и на перцептивно равнище. Същото е валидно и за всеки модел с отчитане на повече от два форманта.

Класическият двуформантен модел на гласните (Фант 1960) е едновременно прост и високо информативен, което е фундаментално важна характеристика на всеки фонологично релевантен акустичен модел. Той е прост, защото съдържа единствено необходимата и достатъчна информация за идентифициране на гласните; моделът е високо информативен, защото позволява директно да се „превеждат“ акустичните резултати на езика на артикулационната фонетика благодарение на аналогията между акустичната точкова графика в равнината F2/F1 (т.нар. акустичен трапец или триъгълник) и артикулационната точкова графика, изобразяваща зоните на учленение на гласните във вертикалната надлъжна равнина през гласовия канал. Добре известна слабост на модела е обаче неговата неприложимост за някои гласни в някои езици.

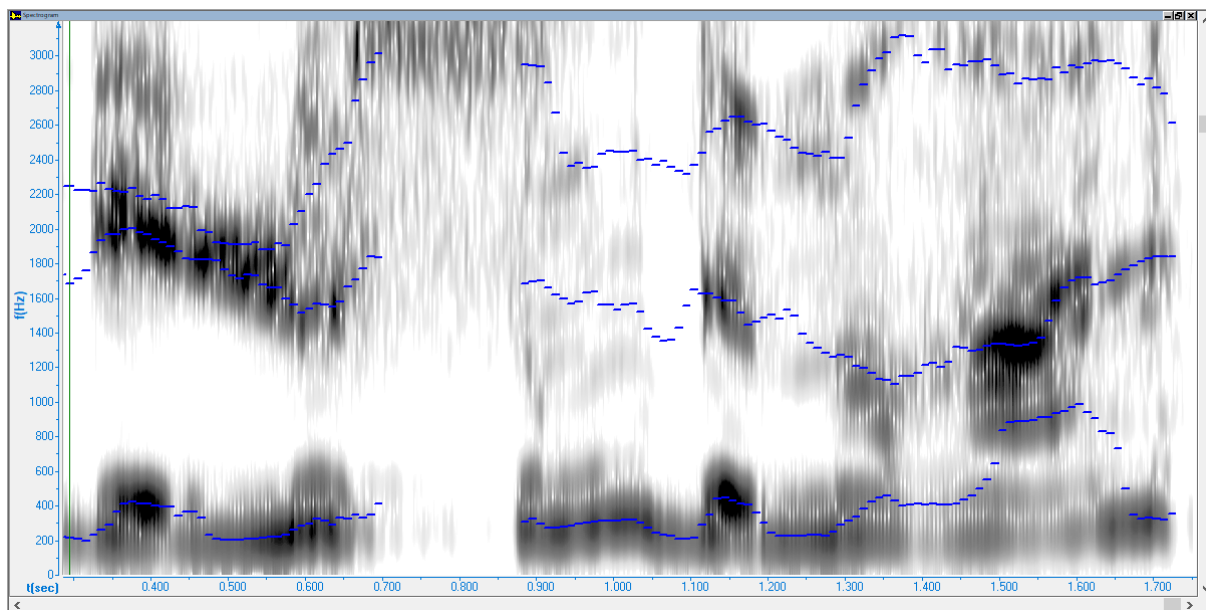
Новият модел запазва предимствата на класическия двуформантен модел: образът на всяка гласна е отново точка в равнинна координатна система, като конфигурацията на гласните за всеки език с три или повече гласни е аналогична на добре познатия акустичен трапец/триъгълник в координатната система F2/F1. Разликата е в начина на определяне на стойността на втория формант, а целта е да се избегне 1) недостатъкът при отчитане на пиковите стойности на формантите и 2) специфичният за класическия двуформантен модел недостатък – закономерната вероятност за неадекватност на резултатите. Тази слабост се проявява единствено при някои гласни в някои езици, но универсалността (широката приложимост) е също така фундаментално изискване за всеки модел, каквото е изискването моделите да бъдат едновременно прости и информативни. Неадекватността се проявява в липсата на съответствие между акустични съпоставителни резултати, от една страна, и перцептивни съпоставителни резултати, от друга страна, което може да се определи като парадокс. Парадокс е например честото акустично съвпадение на френска закръглена предна гласна и българска предна (незакръглена) гласна. Перцептивно същите гласни се оценяват лесно и недвусмислено като различни както от говорители французи, така и от говорители българи.

Решението на този парадокс се свежда до отчитане на допълнителни акустични параметри с перцептивна тежест, по-специално на третия формант – именно третият формант се променя значително в резултат на закръглянето на устните при предните гласни. Оказва се, че отчитането на третия формант невинаги е достатъчно само по себе си, защото различни формантни структури могат да имат идентичен перцептивен ефект в резултат на перцептивната интеграция на съсед-

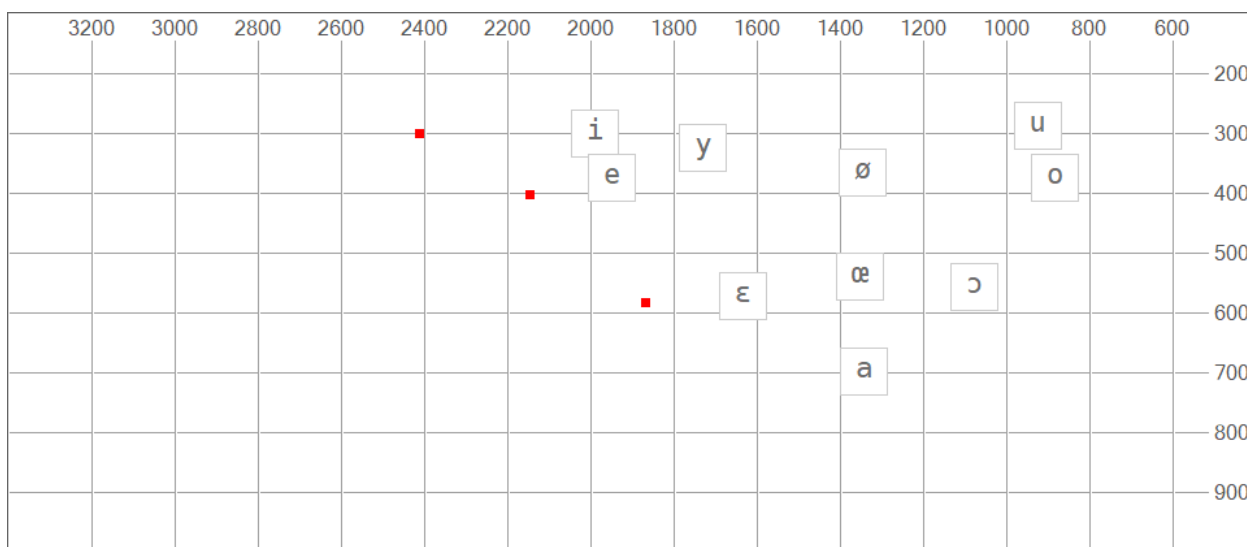
ни форманти (Чистович и Люблинская 1979: 185). Оказва се още, че именно в проблемните случаи при акустико-фонетичното моделиране на гласните разстоянието между F2 и F3, измерено в психоакустичната мерна единица *bark*, е такова, че позволява, и по-точно – налага да се отчете перцептивната интеграция на формантите ($\Delta < 3.5 \text{ bark}$). Идеята за необходимостта от отчитане на перцептивната интеграция на формантите не е нова, но конкретната реализацията тук на тази идея под формата на софтуерно осигурен метод за практическо моделиране на гласните се основава на по-различна концепция: перцептивната интеграция се изразява не чрез определяне на *претеглената средна стойност* на двата съседни и достатъчно близки форманта (т.е. на техните пикови стойности), а чрез определяне на *претеглената средна стойност* на двете съседни формантни зони, т.е. на всички хармонични честоти, които попадат 1) в зоната на формантите и 2) в зоната между тях. Този подход осигурява едновременно *по-точно* и *по-надеждно* отчитане на спектралното (т.е. на честотното) разпределение на акустичната енергия. Отчитането е по-точно, защото така получената средна стойност се изчислява въз основа на изчерпателна спектрална информация; то е и по-надеждно, защото не се използват пиковите стойности на формантите, чието определяне е ненадеждно, когато се разчита на автоматичното им изчисляване с помощта на съществуващите алгоритми (вж. Фиг. 1), особено при някои гласни (напр. френските носови гласни), при глас с висок основен тон (чест случай при говорител жена или дете), както и при неспазване на сравнително високите изисквания при дигитализацията на звука.

Фиг. 2 илюстрира вариативността при някои гласни на техните *най-релевантни* фонетични характеристики, получени с използване на класическия двуформантен модел. Тази вариативност е прекомерно голяма и поставя под съмнение релевантността на използваните данни (F1 и F2) при този модел. Фиг. 3 потвърждава, че и добавянето на третия формант в акустичното описание на една гласна не е достатъчно само по себе си, за да се представят акустично две реализации на една гласна като идентични, каквито са гласните на перцептивно равнище. Дефинирането на формантната структура на една гласна като подредена двойка или тройка от числа – пиковите стойности на формантите – води до парадокса, илюстриран на Фиг. 3. Ако обаче при определяне на формантната структура се отчете перцептивната интеграция на някои от тях (в съответствие с условията, дефинирани от Чистович и Люблинская 1979), се оказва, че между различни реализации на една

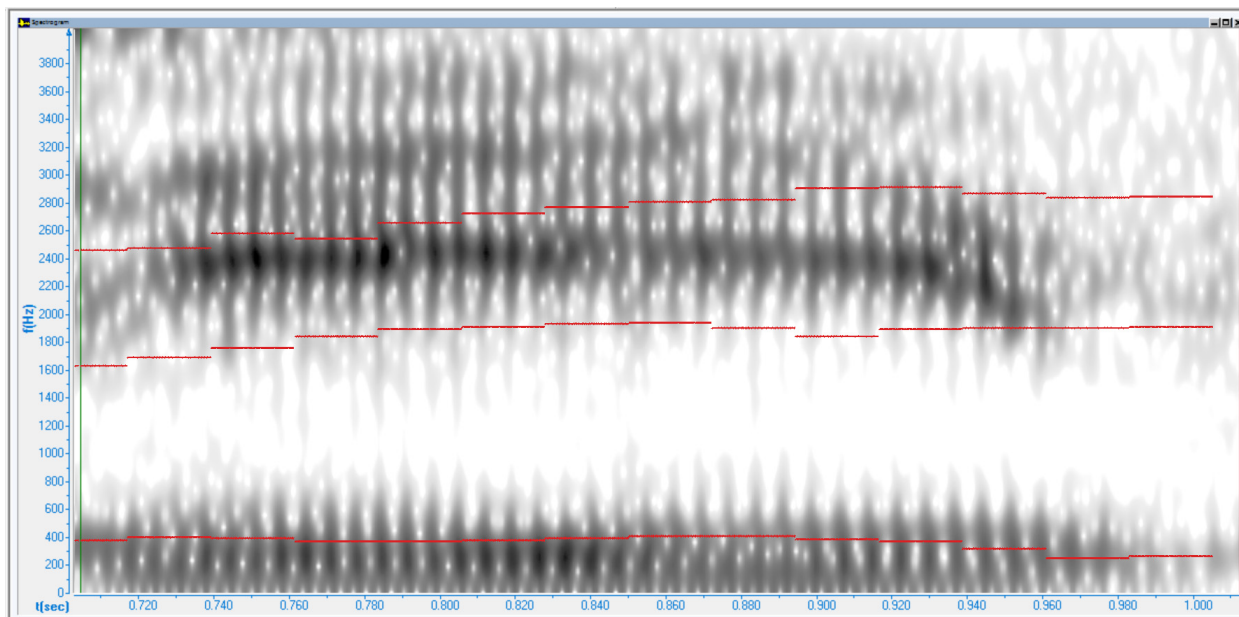
гласна има добро съвпадение на акустично равнище тогава и само тогава, когато има съвпадение на перцептивно равнище (Фиг. 4).



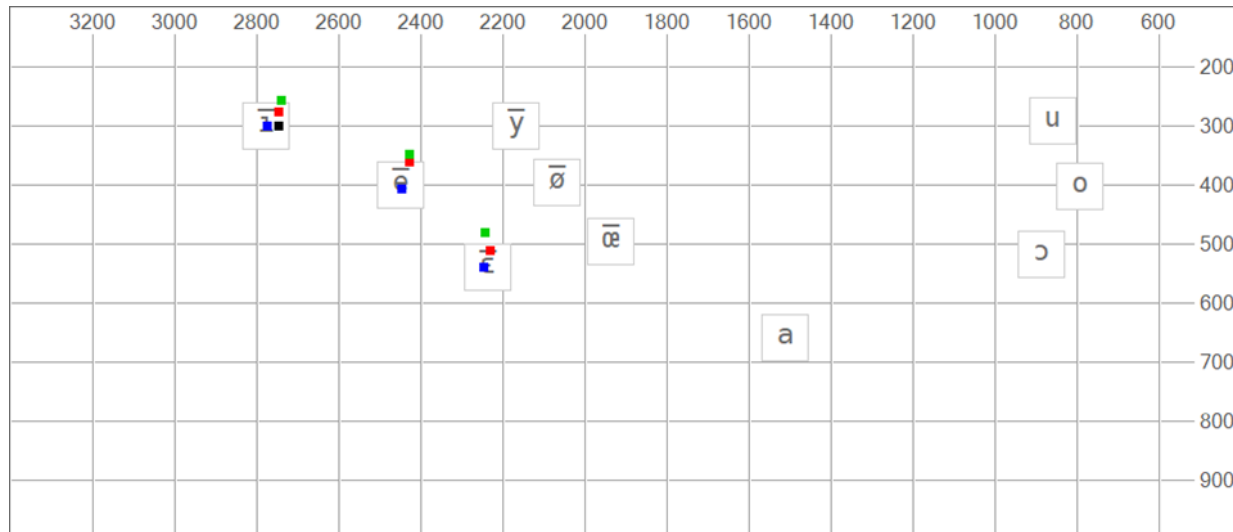
Фигура 1. Спектрограма (*des leçons de morale*; студиен запис; говорител: *Jean-Marie Le Pen*) – илюстрация на липсата на надеждност при автоматично определяне на формантите



Фигура 2. Предни незакръглени гласни в силна позиция (стандартен френски език, Париж, говорители: *Jean-Yves Dommergues* и *Paul Colonna*) – илюстрация на вариативността



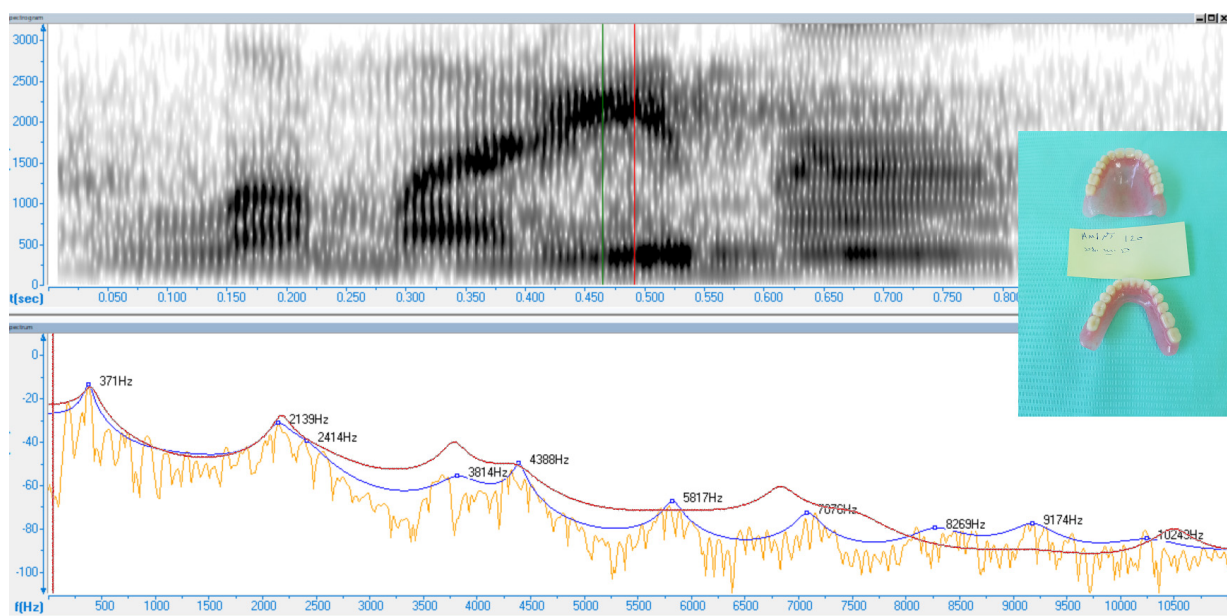
Фигура 3. Френската гласна [i] в силна позиция (стандартен френски език, Париж, говорители: Jean-Yves Dommergues и Paul Colonna) – перцептивна идентичност в рамките на произносителната норма и значителна разлика във формантната структура (F2 и F3)



Фигура 4. Интегрална формантна структура на френските гласни (Jean-Marie Le Pen, Jean-Yves Dommergues и Paul Colonna) – илюстрация на акустичната инвариантност като корелат на фонологичната идентичност на различните реализации

Методът с отчитане на перцептивната интеграция е лесно приложим с помощта на софтуерната програма *TREFL* (web.uniplovdiv.bg/rousni) и е използван от Маджед Хусеин (Хусеин и Нико-

лов, 2018: 104 – 108) за акустично-перцептивна оценка на влиянието на зъбните протези върху артикулацията на говорните звукове. Използването на този метод за контрол на говорната функция в областта на протетичната дентална медицина позволява да се оценят промените в тембъра на гласа, които биха могли да повлияят съществено на комуникативните възможности на пациентите, разграничавайки ги надеждно от други промени, които влияят единствено върху специфичните за индивида тембрални характеристики на гласа на пациента. Промените от втория тип се установяват също така лесно с помощта на софтуерната система *Speech Analyzer* (software.sil.org/speech-analyzer/) и *TREFL* благодарение на възможността на контрастна суперпозиция на различните акустични графики (Фиг. 5).



Фигура 5. Спектрограма и спектър на българската гласна [u] в силна позиция – думата ламарина, изговорена от пациентка Ж. Д. 1) без и 2) с протеза, възстановяваща не само дъвкателната и естетическата, но и говорната ѝ функция

Наслагването на двете графики на спектъра на гласната показва, че влиянието на зъбната структура (нейната липса) се е отразило трайно и значително на тембралните характеристики на гласната, без обаче тази промяна да касае релевантните ѝ признаци (вижда се пълното съвпадение на първите три форманта). Това може да се обясни с постепенна адаптация след загубата на зъбите и използване на компенсаторни механизми от страна на пациентката за функционално възстановяване, макар и частично, на качеството на говорните звуко-

ве. Протезата възстановява това качество напълно, което е извод от слухова преценка (тембърът на гласа се оценява като подобрен и напълно естествен). За съжаление, липсват данни за спектралните характеристики на същата гласна при пациентката преди загубата на зъбите, които да потвърдят обективно този слухов анализ.

Усъвършенстването на методите за фонетичен анализ не може да отмени една принципна трудност при оценката на влиянието на зъбните протези върху говорната функция. Експерименталните резултати показват, че един и същ вид протеза при двама различни пациенти може да доведе първоначално до противоположни резултати – при единия пациент протезата възстановява веднага нормалната говорна функция (в съпоставка с езиковата норма), докато при другия я влошава (в съпоставка не само с нормата, но и със състоянието без протеза). Такова (временно) влошаване е неочаквано с оглед на факта, че протезата е възстановила в голяма степен геометрията на устната кухина, но може лесно да се обясни с по-голямата или по-малката адаптивност (съответно инертност) на артикулационните навици, които се променят, за да компенсират акустичните ефекти на структурните промени в гласовия канал.

Следователно една бърза и обективна оценка на влиянието на зъбните протези върху говорната функция може да придобие допълнителна практическа стойност, ако тя се приложи не като елемент от статистически данни, но и като конкретна „диагноза“ с възможност за насочване на компенсаторните артикулационни усилия на пациента.

ЛИТЕРАТУРА

- Николов 2016:** Николов, Р. Значение на третия формант в сравнителния анализ на българските и френските гласни. [Nikolov, R. Znacheniya na tretiya formant v sravnitelniya analiz na balgarskite i frenskite glasni.] // *Научни трудове на ПУ „Паисий Хилендарски“, Филология*. Пловдив: УИ „Паисий Хилендарски“, 2016, Том 54, кн. 1, сб. Б, 42 – 47.
- Хусейн, Николов, 2017:** Хусейн, М., Николов Р. Акустично-перцептивна оценка на влиянието на зъбните протези върху артикулацията на звуковете в арабския език. [Husein, M., Nikolov, R. Akustichno-pertseptivna otsenka na vliyanieto na zabnite protezi varhu artikulatsiyata na zvukovete v arabskiya ezik.] // *Научни трудове на Съ-*

юза на учените в България – Пловдив, Серия Г. Медицина, фармация и дентална медицина. Пловдив: Дом на учените, 2017, 104 – 108.

Чистович, Люблинская 1979: Chistovich, Ludmilla A., Lublinskaya, Valentina V. The ‘center of gravity’ effect in vowel spectra and critical distance between the formants: Psychoacoustical study of the perception of vowel-like stimuli. // *Hearing Research*, Volume 1, Issue 3, August 1979, 185 – 195, <<http://www.sciencedirect.com/science/article/pii/0378595579900121>> (12.11.2017).

Фант 1960: Fant, Gunnar. *Acoustic theory of speech production*. Mouton, the Hague, 2nd ed., 1970.