

**ТЕМБЪРЪТ НА ГЛАСА
КАТО ФНОСТИЛИСТИЧЕН ПАРАМЕТЪР**

Руси Николов
Пловдивски университет „Паусий Хилендарски“

**VOICE TIMBRE: QUANTIFYING THE PHONOSTYLISTIC
VARIATION**

Rusi Nikolov
Paissii Hilendarski University of Plovdiv

Timbre is the perceptual correlate of the sound spectrum (as, in some ways, the color is the eye's response of the light spectrum). In the speech stream, timbre involves two basic types of sets of invariant neural properties, since it is what identifies the segments as well as the speakers with their anatomical peculiarities and individual-specific articulatory habits. When it comes to evaluate the speakers' styles of pronunciation or their physical and mental tone, in contrast with prosodic parameters, the auditory range (if any) of timbre is insignificant, while the corresponding acoustic variation, as we will show in this paper, seems to be relevant in an interesting and easily understandable manner.

Key words: experimental phonetics and phonology, phonostylistics, natural language processing, vowels

Въз основа на съпоставяне на акустични данни – актуални и референтни – за един и същ говорител могат да се извлекат повече или по-малко надеждни показатели за неговия физически и психически тонус, по-специално за относителната степен на умора, мотивираност, емоционалност и т.н. При такъв сравнителен анализ е необходимо да се елиминира позиционно (или контекстуално) обусловената в речта произносителна вариативност, което може да се постигне статистически, т.е. чрез изследване на достатъчно голям корпус, или чрез разглеждане на едни и същи речеви форми, реализирани по различно време.

По отношение на тембъра тази възможност може да бъде конкретизирана със следния пример. Всяко от изреченията: 1) *Три, две, едно,*

нула, *старт!* и 2) *Още утре, или да, или не!* – съдържа всички гласни в българския език в ударена позиция (с изключение на централната гласна [ъ]). Всяка от гласните за всяко от изреченията ще бъде произнесена вероятно с различен тембър от един и същ говорител по различно време и в различни условия (така както постиженията при едно и също физическо упражнение в различни моменти на деня ще бъдат различни). Не можем обаче за оценка на тонуса на говорителя да сравняваме директно съответните гласни в двете изречения – тембърът на гласната [е] например значително се различава в думите *не* и *две* под влияние на предходната съгласна (гласната е по-предна в думата *не*).

Прозодията на речта – основен тон, интензитет и времетраене – е по-непосредствен, по-леснодостъпен показател от спектралните ѝ характеристики за състоянието на говорителя – в смисъл, че мелодията (монотонна или раздвижена), интензитетът (висок или нисък) и темпото (бързо или бавно) могат директно да се оценят слухово и сравнително лесно да се интерпретират с оглед на всяка подобна задача за екстралингвистична оценка от този вид на речта и говорителя.

От друга страна, нормалната интериндивидуална вариативност на спектралните характеристики на звуковете в речта трудно може да се оцени слухово поради категориалния характер на перцепцията¹. Съвременните изследвания в областта на функционалната невроанатомия на перцепцията на речта, касаещи физиологичната репрезентация на формантната структура на звука в човешкия мозък (Кац и кол. 2015: 533 – 534), показват, че активираните зони в мозъчната кора в резултат на звуков стимул с определена формантна структура не съответстват на установената с прости звукове мозъчна честотна „карта“². Тези резултати потвърждават установеното в перцептивната фонетика явление, наречено *ефект на „центъра на тежестта“* (Чистович и Люблинская 1979: 185 – 195), който се изразява в следното: ако два форманта са достатъчно близки един до друг по психоакустичните си честотни характеристики (макс. разлика 3 – 3,5 Bark), те могат да бъдат заместени с един-единствен формант с еквивалентен перцептивен ефект и междинна стойност на честотата му (която е по-близка до тази на форманта с по-голям интензитет). Освен всичко това именно спектърът дефинира фонологично звука и поради това неговата психо-

¹ „[...]discrimination of stimuli from within the same phonemic category tended to be poor“ (Джасчик и Лус 2002: 497).

² „[...]these studies provide evidence that the cortex represents the formant structure of vowels in a manner that is unrelated to the tonotopic map“ (Кац и кол. 2015: 533).

акустична обработка и функционална идентификация е силно категориална; затова например една малко по-предна или по-задна, по-затворена или по-отворена гласна ще се идентифицира като една и съща функционална единица. От артикулационна гледна точка обаче тази „скрита“ (от гледна точка на слушателя) произносителна вариативност е свързана с упражняване на различно по големина биомеханично усилие. Следователно вариативността на спектралните характеристики на звуковете в речта, установена инструментално, може да се използва като отражение на промените в психо-физическото състояние на говорителя. Функцията на тази зависимост е обаче 1) глобална и 2) с малка „чувствителност“, което означава съответно, че: 1) състоянието на говорителя или използваният от него стил на говорене може да се оценят само приблизително и като цяло, без да се познават отделните му компоненти, и 2) значими промени в психо-физическото състояние на говорителя могат да се отразят само незначително на спектралните характеристики на звуковете в речта му. В този смисъл са неприемливи заключения като следното – при това в резултат на субективна слухова оценка на тембъра на гласа при публичните изяви на известни политически личности: Дейвид Камерън, министър-председател на Великобритания (10 август 2011 г.), и президента на Латвия (28 май 2011 г.): „Тембърът на гласа на ораторите изразява достойнство, съпричастност и лична ангажираност“³ (Бреде 2013: 4).

Изискването за научна строгост в лингвистиката не подлежи на коментар, но във фоностилистиката то има още по-голяма тежест, защото фонетиката е точна експериментална дисциплина, от една страна, а от друга страна, когато се дават личностни оценки посредством речта, отговорността е още по-голяма. Фактът, че авторката на цитираното изследване анализира тембъра в статия, посветена на интонацията, дава допълнително основание да се уточни понятието тембър.

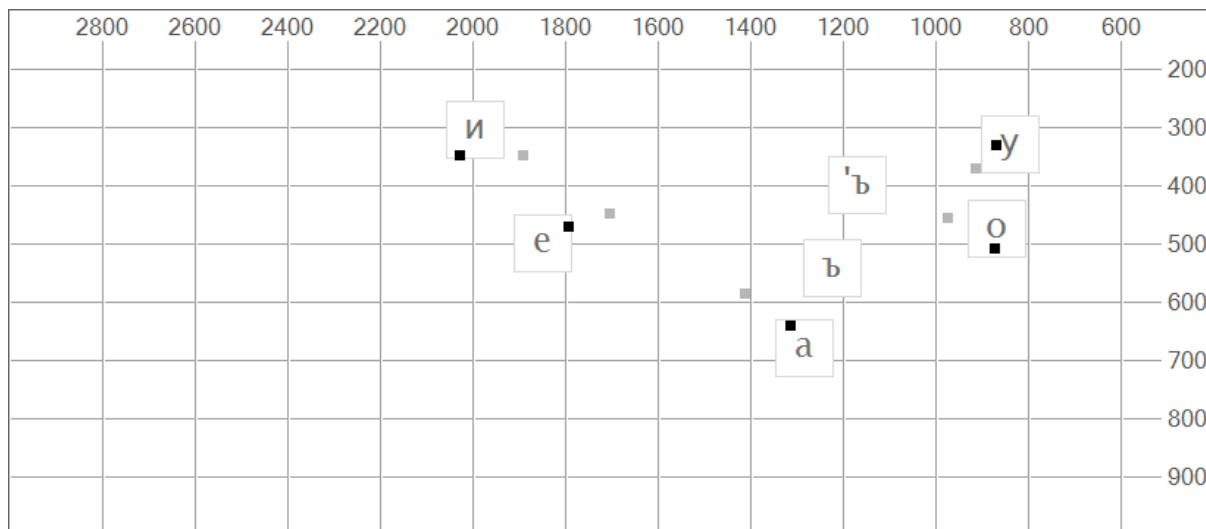
Тембърът на говорните звукове е перцептивно понятие, което отговаря от акустична гледна точка на техните спектрални характеристики, които от своя страна зависят от геометрията на гласовия канал и от динамиката на областта, в която се образува първичният звук. Част от геометричните характеристики на гласовия канал с акустична стойност са постоянни в речта на даден говорител и зависят единствено от анатомичните му особености. Именно тези характеристики идентифицират в най-голяма степен говорителя и подлежат

³ „*The public speakers' voice timbre is dignified, concerned and personally involved*“

евентуално на естетическа или друга оценка (без да се пренебрегват и придобитите индивидуални артикулационни навици). Дистинктивната стойност на тези характеристики очевидно няма лингвистичен характер и в частност – фоностилистичен. Затова в по-тесен, лингвистичен аспект тембърът на говорния звук може да се определи като глобален перцептивен корелат на системата от езиково релевантни спектрални характеристики на звука.

Езиково релевантните спектрални характеристики на звуковете са лесно измерими и същевременно достатъчно показателни за целите на това изследване, ако се сведат до формантната структура на гласните. За разглежданите езици – български и френски – е достатъчно като първо приближение да се отчитат първите два форманта. Въпреки тези опростяващи предпоставки спектралните отклонения в гласа не могат да се интерпретират еднозначно, защото всеки елементарен компонент на тези отклонения има специфична тежест в зависимост от 1) говорителя, 2) гласната и 3) вида на компонента – първи или втори формант. Тези специфични тежести могат да бъдат определени емпирично за всеки отделен говорител.

За български език е използвано изречението *Три, две, едно, нула, старт!*, повторено многократно. Това кратко изречение съдържа всички български гласни в ударена позиция с изключение на централната [ъ]. Интересна е констатацията, че граничните реализации – началната и крайната – са по-грижливи, по-прецизни от артикулационна гледна точка. Този резултат не е съществен по отношение на конкретната цел, но е интересен с оглед на аналогията с експерименталните резултати в психологията при запаметяване на списъци от семантично несвързани елементи – най-добре се запомнят началните и крайните елементи (Либерман 2012: 311), докато в нашия случай най-добре се артикулира в началото и в края на задачата. Обяснението за различния мускулен тонус, с който артикулираме в началото, средата и края на задачата, може да се свърже условно със следните три етапа в състоянието на говорещия: 1) ентузиазъм от начинанието, 2) отегчение и умора от „безсмислените“ повторения и 3) тонизиращо облекчение от мисълта за предстоящия край на задачата. На Фиг. 1 по-долу е показан акустичен триъгълник на българските гласни (маркерите с етикети), реализацията на ударените гласни в изречението *Три, две, едно, нула, старт!* (с черни маркери), които са максимално близки до модела (говорителят е един и същ), както и гласните в същото изречение, произнесено в позиция от повторенията, съответстваща на понижен тонус (маркерите в сиво).



Фиг. 1. Силна и слаба реализация на гласните в силна позиция

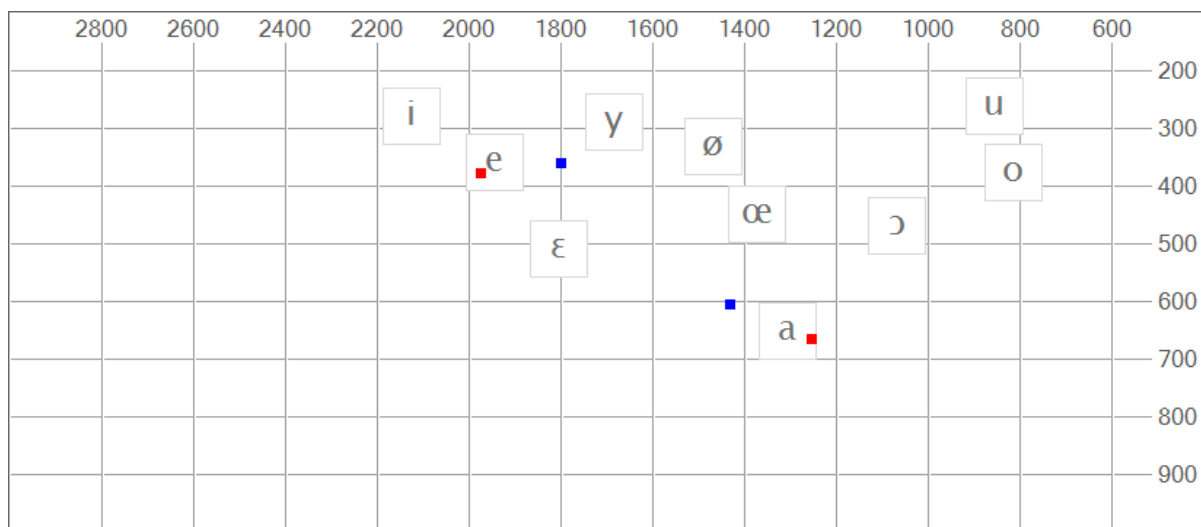
Този резултат е очакван и не е изненадващ нито с количествени-те, нито с качествените си параметри, но онагледяването му на Фиг. 1 подсказва, че може да се използва площта на фигурата, която се получава като обвивката на всяка отделна точкова диаграма, като глобална количествена оценка за „грижливостта“, с която са артикулирани гласните. Тази оценка е възможен фоностилистичен показател за психо-физическото състояние на говорителя (мотивираност, умора и т.н.). Оценка е ненадежден показател в случаите на близко съвпадение между актуални и референтни стойности, защото тя е глобална и може да е резултат на неутрализиране на противодействащи фактори (напр. силна мотивираност, компенсираща ефекта на умората); обратно, едно значително отклонение от индивидуалната норма в една или друга посока може да се интерпретира с голяма степен на достоверност. Методът е лесно приложим в практиката с помощта на свободно достъпната компютърна програма *TREFL*.

Освен това този резултат повдига един важен принципен въпрос за възможните сравнения и интерпретации. Тук ще обособим три случая:

1. Различни състояния на един и същ говорител, отразени в различието на акустичните данни.
2. Различни произносителни стилове при различни говорители.
3. Различно произношение на двама различни говорители при съвпадащи и очевидно непълни акустични данни.

Първият случай беше разгледан и илюстриран на Фиг. 1. Вторият случай е интересен с това, че в множество сравнителни анализи на

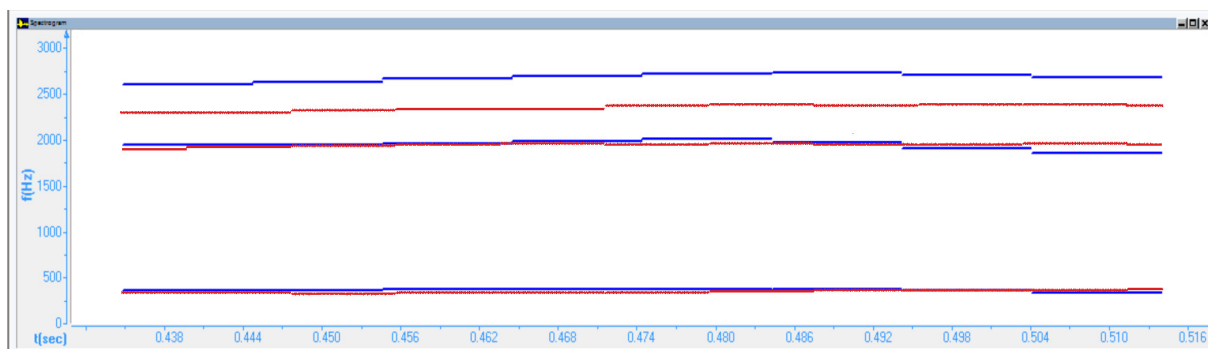
произношението на водещи журналисти и техните гости (напр. Б.В. и Р.П., Николов 2014: 246) се оказва, че резултатите възпроизвеждат показаните на Фиг. 1. По-конкретно обвивката на точкова диаграма на интервюирания се намира по-често във вътрешността на тази на журналиста. Без претенции за широкообхватни и достоверни резултати, наблюденията показват, че тенденцията се обръща „в полза“ на политците при съпоставяне на стила на журналисти и политици по френските медии, което всъщност би могло да се тълкува отново в полза на професионализма на журналиста, който остава по този начин „в сянка“. Изводът е, че придобитите артикулационни навици на професионално ангажираните с гласовата комуникация говорители включват характеристиката по-силно изразена екстремност на т.нар. кардинални гласни. Един такъв експериментален сравнителен анализ на базата на интервю по френското радио (*France Info*) от 14 октомври 2014 г. е илюстриран на Фиг. 2, където резултатите потвърждават констатираната тенденция: речта на журналиста (Jean-François Kelly) е забързана, докато министърът на финансите (Michel Sapin) придава тежест на всяка своя дума, което се отразява на тембъра на гласните.



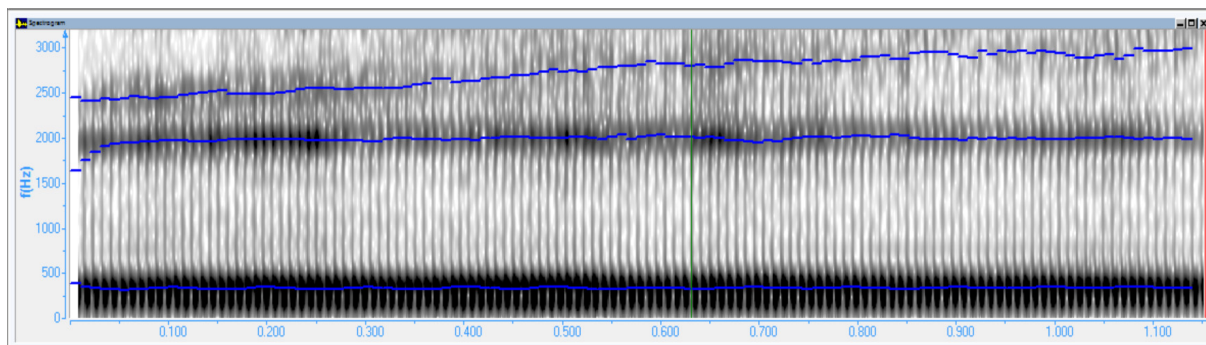
Фиг. 2. Гласните [e] и [a] в речта на френски журналист и политик

Третият случай е особено важен едновременно от теоретична и приложна гледна точка – например за чуждоезиковото обучение. В координатната система F2 vs. F1 (като тази на Фиг. 1 и 2) се случват неочаквани съвпадения, които, ако не се дължат на неточност в отчетите, имат само едно обяснение – несъвършенство на модела, според който първите два форманта са достатъчни за описанието на гласните. Възможно е например българската гласна (и) да съвпадне точно с френс-

ката гласна (e), докато в същото време двете гласни ясно се оценяват слухово като различни. На Фиг. 3 по-долу и двете гласни – българското [и] и френското затворено [e] – са в крайна, отворена, ударена сричка, т.е. в силна позиция; първите два форманта съвпадат, докато третият формант на [и] е със значително по-ниска стойност от тази на [e], което е свързано с по-малкия отвор на гласовия канал при устните за българската гласна. Обяснението произтича от физичните свойства на акустичните резонатори и за случая на речеобразуването е илюстрирано на Фиг. 4 с прехода от българското [и] към френското затворено [e]. (Непосредствено потвърждение може да получим, разбира се, и чрез наблюдение на артикулационните жестове на устните в отчетлива реч на френски.) Това ни води до един терминологичен парадокс. Ако върху двойната графика на Фиг. 3 насложим още една – очертанията на първите три форманта на френското [i], няма да е изненадващо, ако първите два форманта отново съвпаднат, докато третият, също напълно закономерно, е с по-висока стойност от тази за [e] (и естествено – за [и]), което осигурява перцептивната идентичност на гласната. (Констатацията е направена за журналиста Жан-Франсоа Кели и думите *continuer* и *dire* – и двете под ударение и от един и същ кратък текст – интервюто по радио *Франс-Енфо* с Мишел Сапен от 20 октомври 2014 г.) Парадоксът се състои в това, че гласната [i] може да се окаже реално по-отворена от [e] (като се отчита различният отвор на гласовия канал при устните и еднаквата геометрия на гърлената и на устната кухина), без да се отразява тази артикулационна особеност при съпоставката на двете гласни на тяхната перцептивна оценка. Дори и в общия случай, при усреднени стойности за френския език, разстоянието между двете гласни е незначително в координатната система F2 vs. F1. Тази близост обаче е подвеждаща, защото моделът на двуформантната структура не отчита ролята на третия формант, който увеличава допълнително контраста между гласните.

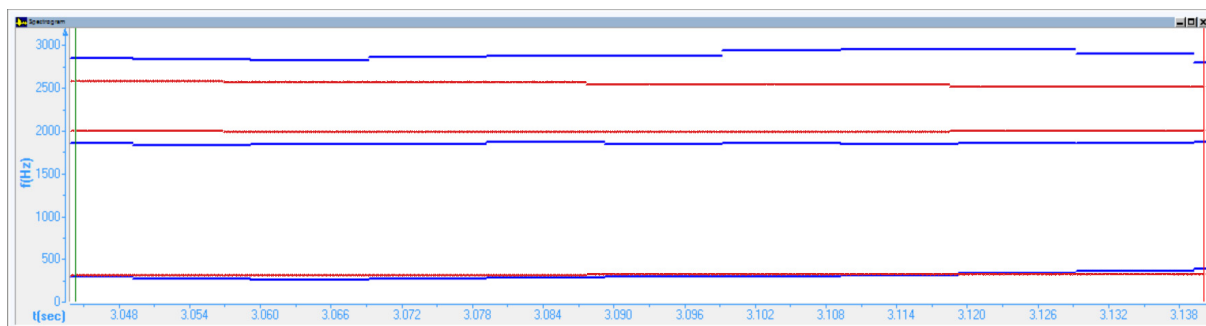


Фиг. 3. [u] vs. [e]



Фиг. 4. Преход от [u] към [e]

Следователно влиянието на третия формант, по-специално при предните затворени гласни, не може да се пренебрегне в нито един език с такива гласни, ако искаме акустичното описание на гласните да е функционално обвързано с тяхната перцептивна оценка. Това произтича директно от ефекта на „центъра на тежестта“ и се потвърждава от експерименталния факт, визуализиран на Фиг. 5: вторият формант на българското [и] е с по-висока стойност от тази за едно също така стандартно френско [i] (Жан-Франсоа Кели), което е необичайно, но този акустичен факт няма перцептивен корелат, защото ниската стойност на F2 се компенсира перцептивно с високата стойност на F3.



Фиг. 5. [u] vs. [i]

Необходимостта да се отчита и третият формант (F3) не е новост в теорията на гласните (референции), но практиката – както в научното описание на гласните, така и в чуждоезиковото обучение – значително изостава. Удобен и високопроизводителен инструмент за тези цели е най-новата версия на компютърната програма *TREFL*.

ЛИТЕРАТУРА

- Бреде 2013:** Brēde, Maija. Characteristics of the Publicistic Style of Intonation in English and Latvian. // *Baltic Journal of English Language, Literature and Culture*, Volume 3, University of Latvia, Riga 2013, 4 – 19.
- Джасчик и Лус 2002:** Peter W. Jusczyk and Paul A. Luce. Speech Perception. // *Stevens' handbook of experimental psychology Volume 1: Sensation and Perception*. Third Edition, John Wiley & Sons, Inc. New York, 2002.
- Кац и кол. 2015:** Katz, J., Chasin, M., English, K., Hood, L. J., Tillery, K. L. *Handbook of Clinical Audiology*. Wolters Kluwer Health, Philadelphia 2015.
- Либерман 2012:** Lieberman, David A. *Human Learning and Memory*, Cambridge University Press, 2012.
- Николов 2014:** Николов, Руси. Генериране и управление на езикови ресурси с многофункционалната програма TREFL. // *Езикови ресурси и технологии за български език*, София: Академично издателство „Проф. Марин Дринов“, 2014, 231 – 247.
- Чистович и Люблинская 1979:** Ludmilla A. Chistovich, Valentina V. Lublinskaya. 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.