BALTISTICA XLVIII(2) 2013 301–312

Juris GRIGORJEVS
Latvian Language Institute of the University of Latvia

PROBLEMS USING THE TRADITIONAL ACOUSTIC CUES FOR THE PHONOLOGICAL INTERPRETATION OF VOWELS

Introduction

Working on the Phonetics chapter of the new academic edition of "Grammar of the Latvian Language" the traditional description and classification of speech units had to be revised. Since the main part in speech communication is the acoustic signal, the description of acoustic targets of speech sounds has to be included into the description of a language inventory. This is important also due to the fact that it is easier to perform an acoustic analysis of some speech fragment than to perform an articulatory analysis on it. Knowing the relation between certain articulatory and acoustic parameters the phonological analysis and classification of sounds could be based on the acoustically defined distinctive features instead of rather arbitrary assigned articulatory and/or auditory ones. The aim of the present study was to check the relations between traditional distinctive features and the acoustic or psycho-acoustic features suggested by different authors, as well as to check the acoustic parameters associated with certain features. The Cardinal Vowels and the IPA vowel system based on them is an idealized representation of the vowels that a human is able to produce. These vowel systems are often used as a reference describing vowels of any language. The task for the present study has been set to check if the acoustic parameters suggested by different authors and distinctive features based on them allow a correct phonological interpretation of IPA and Cardinal Vowels. The hypothesis was made that the acoustic parameters and the distinctive features based on them (compactness, graveness and flatness) would allow classifying vowels as good as the articulatory features of height, backness and lip rounding.

Background (acoustic parameters of monophthong features)

It is not a secret for phoneticians that correspondence of the vowel articulation to its placement in the vowel quadrilateral is rather arbitrary (Wood 1975). Despite an extensive criticism of the tongue-arching model advocated by Daniel Jones, it is still used in IPA vowel quadrilateral. The placement of vowels in the quadrilateral is supposed to reflect the position of the highest point of the tongue and therefore is misleading. Luciano Canepari suggests to interpret the vowel position in the vocograms (including IPA vowel quadrilateral) as placement of center of the mediumdorsum (absolute center of the back of the tongue) in relation to various fixed points on the palatal vault ranging from palatal to velar (Canepari 2007, 112–119). This is an acceptable compromise in order to use traditional vowel description by height, backness and lip rounding.

It is more difficult to find the acoustic characteristics bound with these articulatory dimensions, because the acoustic signal produced by speaker depends mainly on the size and location of the constriction in the vocal tract. The system of acoustically based features was suggested by Roman Jakobson, Gunnar Fant and Morris Halle in "Preliminaries to Speech Analysis" (Jakobson et al. 1963).

The articulatory description of vowel height, traditionally expressed by the opposition open vs. close, in "Preliminaries" is replaced by the opposition of **compact** vs. **diffuse**, where "open vowels are the most compact, while close vowels are the most diffuse" (Jakobson et al. 1963, 27), and is bound to the value of F1.

> In the case of vowels this feature manifests itself primarily by the position of the first formant: when the latter is higher (i. e. closer to the third and higher formants), the phoneme is more compact. The closer the first formant is to the upper formants, the higher will be the intensity level of the region above the first formant, especially level between peaks.
>
> The essential articulatory difference between the compact and diffuse phonemes lies in the relation between the volume of the resonating cavities in front of the narrowest stricture and those behind this stricture. The ratio of the former to the latter is higher for the compact than for the corresponding diffuse phonemes. (Jakobson et al. 1963, 27)

The articulatory description of backness, traditionally expressed by the opposition back vs. front, is replaced by the opposition of **grave** vs. **acute**, where feature is based on the predominance of one side of the significant part of the spectrum over the other – "when the lower side of the spectrum predominates, the phoneme is labeled grave; when the upper side predominates, we term the phoneme acute" (Jakobson et al. 1963, 29).

> The position of the second formant in relation to that of the other formants in the spectrum is the most characteristic index of this feature: when it is closer to the

first formant the phoneme is grave; when it is closer to the third and higher formants it is acute.

The gravity of a consonant or vowel is generated by a larger and less comparted mouth cavity, while acuteness originates in a smaller and more divided cavity.

Usually, however, a notable auxiliary factor in the formation of grave phonemes (back vowels ..) is a contraction of the back orifice of the mouth resonator, through a narrowing of the pharynx, whereas the corresponding acute phonemes (.. front vowels) are produced with a widened pharynx. (Jakobson et al. 1963, 30)

As can be seen from the definitions above the main parameter characterizing graveness or acuteness is location F2 in respect to F1 and F3.

The articulatory description of lip rounding is replaced by the opposition of **flat** vs. **plain**, where "flattening manifests itself by a downward shift of a set of formants or even of all the formants in the spectrum" (Jakobson et al. 1963, 31).

Flattening is chiefly generated by a reduction of the lip orifice (rounding) with a concomitant increase in the length of the lip constriction. (Jakobson et al. 1963, 31)

Summing up everything said so far, it can be concluded that:

1) vowel height acoustically is bound to the value of F1;

2) backness is determined by the acoustic distance between F1 and F2, which for back vowels is smaller than distance between F2 and F3;

3) lip rounding is expressed acoustically by lowering values of all or certain set of formants.

Mona Lindau (1975) suggested using frequency values of the first formant (F_1) as a physical correlate of vowel height, and the difference between the frequency values of the second and the first formant (F_2-F_1) as a correlate of backness, but lip rounding should be associated with the decrease of frequency values of the second and the third formant (thus also the value of F_2-F_1).

Testing the acoustic parameters on the Cardinal and IPA vowel data

To see if the dimensions suggested by M. Lindau allow transforming the acoustic data into the shape similar to traditional vowel quadrilateral the formant data of IPA vowels (Hayes 2011) were used (Figure 1). The physical formant center values (measured in Hz) before plotting were transformed to perceptual, i. e., psycho-physical units (measured in barks (z)) according to formula by Hartmut Traunmüller (1988, 97). This was done to achieve a more even spacing of vowels along the horizontal and vertical axis taking into account the logarithmic nature of perception.

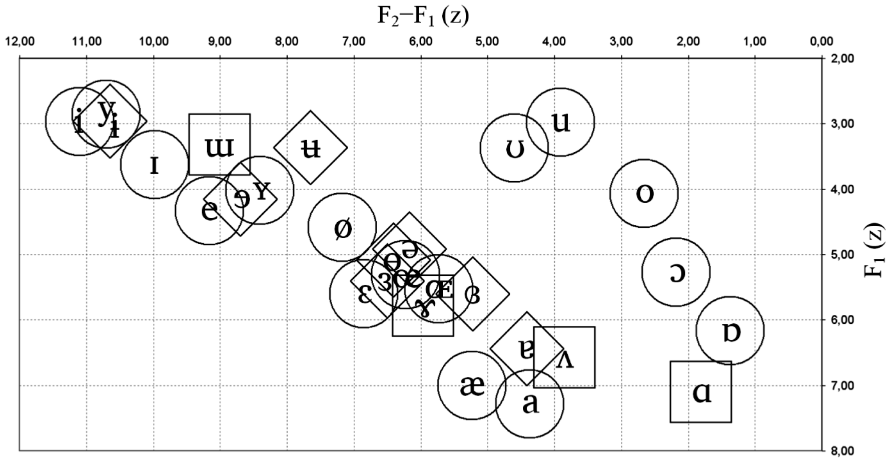

Figure 1. **The placement of IPA monophthongs in psycho-physical vowel plane using dimensions suggested by M. Lindau** (all front and rounded back monophthongs are shown by circles, unrounded back monophthongs – by squares, and central monophthongs – by rhombs)

It can be observed (Figure 1) that the canonical distinction between front, central and back vowels is lost - central and unrounded back monophthongs (except [ɑ]) can not be separated either mutually or from front monophthongs on the basis of the value F_2-F_1. M. Lindau pointed to a similar problem observed in the descriptions of different vowel systems by different authors, but ascribed it to the inconsistency of authors.

> There is a problem with assessing systems with reported central or back unrounded vowels. Linguists do not consistently use the same symbols for these vowel classes. As it turns out it may be a pseudoproblem: these two vowel classes never contrast for non-low vowels. (Lindau 1975, 19)

The examples of vowel systems provided by M. Lindau (1975, 12–20) suggest a conclusion that there are no languages contrasting central (unrounded or rounded) and unrounded back vowels of the same height because of the vague perceptual contrast.

For a phonological classification of vowels G. Fant (1983) suggested using values of the first formant and the effective second formant[1] (F_1 and F_2'). For

[1] The effective second formant is defined as the formant with frequency value of the second formant in two-formant best-match synthetic replica of a human vowel (Fant 1983, 7).

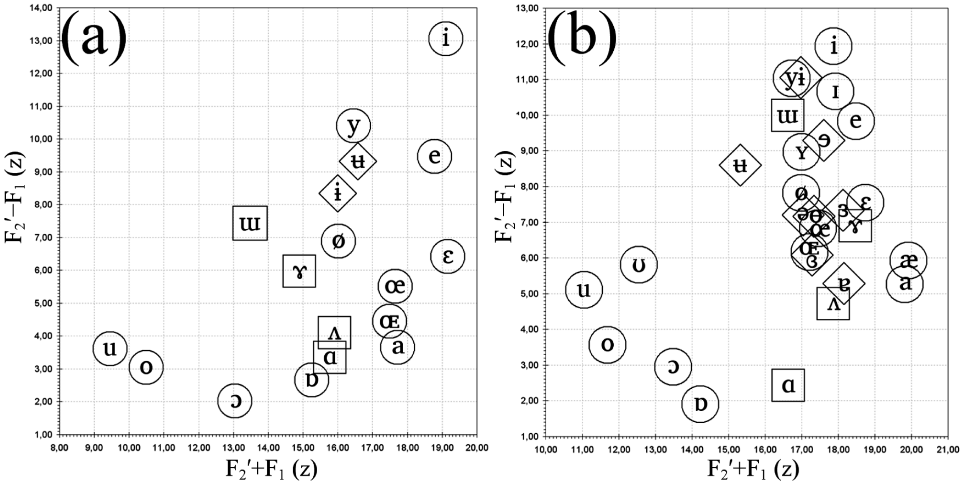

Figure 2. **The placement of Cardinal (a) and IPA (b) monophthongs in psycho-physical vowel plane using dimensions suggested by G. Fant** (all front and back rounded monophthongs are shown by circles, back unrounded monophthongs – by squares, and central monophthongs – by rhombs)

practical reasons G. Fant conceded to use equations for calculating F_2' from measured F_1, F_2, F_3 and F_4 (Bladon, Fant 1978, 3) instead of F_2' obtained in matching experiments. To assign values to the binary features [±grave], [±flat], [±extra flat], [±diffuse] and [±sharp] used for classification of the Swedish vowels, G. Fant suggested using dimensions of spectral spread (F_2'–F_1) and spectral flatness (–(F_2'+F_1)) plotting vowels in acoustic auditory spread vs. flat plane employing transform of frequency values from hertz (Hz) to bark (z).

Using formant data of Cardinal (Bladon, Fant 1978, 4) and IPA (Hayes 2011) vowels two plots were made (Figure 2). Since Bruce Hayes had provided values only for F1, F2 and F3, the F_2' of IPA vowels were calculated using another formula (Carlson et al. 1970, 19). The formant values were transformed from hertz (Hz) to bark (z) units using H. Traunmüller's formula, and the dimension of spectral spread (F_2'–F_1) was marked on the vertical axis and spectral flatness (–(F_2'+F_1)) – on the horizontal axis. It can be observed that the parameters based on the acoustic characteristics suggested by G. Fant (1973, 186–188; 1983, 13–14) for phonological classification of vowels fail to make a distinction between front and central vowels, as well as between rounded front, central and unrounded non-low back vowels.

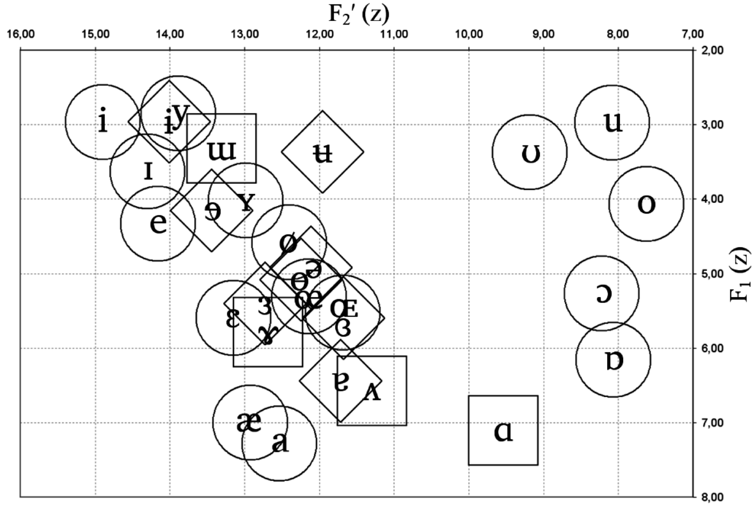

Figure 3. **The placement of IPA monophthongs in psycho-physical F2′/F1 vowel plane** (all front and back rounded monophthongs are shown by circles, back unrounded monophthongs – by squares, and central monophthongs – by rhombs)

Similar problem is encountered if the same IPA vowels are plotted in psycho-physical F_2'/F_1 vowel plane (Figure 3), where more traditional value of F2 is replaced by the value of F2′ to take into account acoustic effect of formants higher than F2. As it can be observed in Figure 1 and Figure 3, the values of F1 can be used to judge the vowel height, although H. Traunmüller has found in his investigations (Traunmüller 1981) that the distance between frequencies of fundamental and F1 (F_1-F_0) is a better acoustic correlate of vowel openness than the value of F1 alone. Lip rounding causes reduction of the frequency values of F2 and higher formants, and therefore affects the values of F_2-F_1 (Figure 1), as well as of F_2' (Figure 2 and 3), thus making noticeable difference with unrounded counterpart in rounded-unrounded vowel pair.

The only feature expressed ambiguously by the acoustic parameters is backness, if one should consider central and unrounded back vowels. If the monophthongs of a given language had to be labeled front, central and back on the basis of definition of grave and acute phonemes by R. Jakobson, G. Fant and M. Halle in "Preliminaries to Speech Analysis" (Jakobson et al. 1963, 29–30), acute phonemes should be labeled *front*, grave phonemes – *back*, but those that are neither acute nor grave – *central*. As it was described earlier, this should be done mainly on the basis of the proximity of F2 either to F1 or to F3.

The research on spectral integration (Chistovich, Lublinskaya 1979) has shown that formants being less than 3.5 z apart form a "center of gravity" in the vowel spectrum, and this spectral dominance is one of the main characteristics of the perceptual vowel quality. G. Fant uses the distance of 3.5 z to define Swedish vowels as being front, central or back (Fant 1983, 12), i. e., if the distance between F2 and F1 is less than 3.5 z, the vowel is *back*, if this distance is more than 3.5 z and the distance between F2 and F3 is less than 3.5 z, the vowel is *front*, but if the F2 is more than 3.5 z from both F1 and F3, the vowel is *central*.

To check if this approach can be used in all situations, the data of Cardinal (Bladon, Fant 1978, 4) and IPA (Hayes 2011) monophthongs were used to create plots (Figure 4) in psycho-physical F_2-F_1 vs. F_3-F_2 plane (values in barks – z). Judging by the Figure 4 the approach has failed for both Cardinal and IPA vowels. If the 3.5 z distance is employed for Cardinals (Figure 4a), unrounded back monophthongs (except [ɑ]), as well as [a] and [u] have to be classified as central, but [ɨ] and [ʉ] – as front. Applying the same rules to IPA vowels (Figure 4b) only [o], [ɔ], [ɒ] and [ɑ] can be labeled as back, [u] and [ʊ] happen to be central, but all the other monophthongs (front, central and non-low unrounded back) fit into category of front vowels. This suggests that the perception of backness is not based merely on the position of F2 in relation to F1 and F3.

Plotting IPA vowels in a three-dimensional vowel space (Figure 5a) using values of their F1, F2 and F3 as coordinates shows the solution to this problem. Most of the monophthongs overlapping in vowel planes (Figure 1 to Figure 4) are clearly separated in the vowel space on the basis of differences in values of at least one formant (F_1, F_2 or F_3). The same can be observed if a combined vowel plane is constructed where the values of F2 are plotted against the values of F1 and against the values of F3 (Figure 5b). For most monophthongs that have less than 1 bark distance to the neighboring ones in F_2 vs. F_1 plane the distance is increased in F_2 vs. F_3 plane. An insufficient increase in perceptual distance can be observed between the front and central monophthongs [y] and [ɨ], [ʏ] and [ɘ], [ø] and [ɵ], [ɛ] and [ɜ], as well as between central and back unrounded monophthongs [ɜ] and [ɤ], [ɐ] and [ʌ].

It can be noticed that in dimension of F1 quasi equidistance between close, close-mid, open-mid and open vowels appears only for unrounded

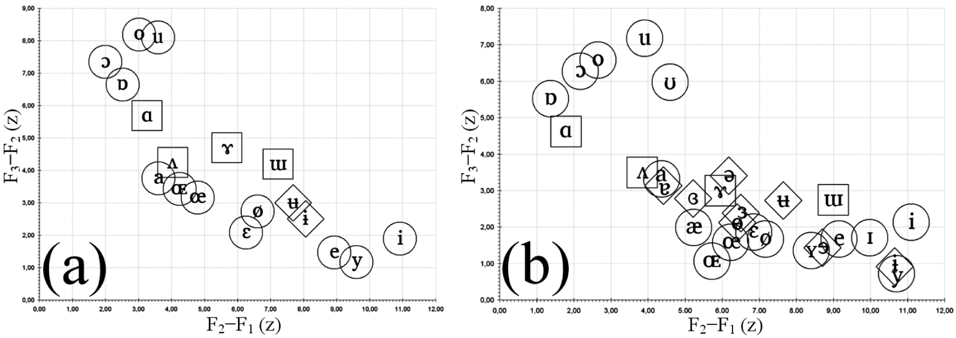

Figure 4. **The placement of Cardinal (a) and IPA (b) monophthongs in psycho-acoustic vowel plane using dimensions characterizing backness and frontness** (all front and back rounded monophthongs are shown by circles, back unrounded monophthongs – by squares, and central monophthongs – by rhombs)

front ([i]-[e]-[ɛ]-[a]) and rounded back ([u]-[o]-[ɔ]-[ɒ]) monophthongs, and more or less for unrounded central ([ɨ]-[ɘ]-[ɜ]-[ɐ]) monophthongs, too. For rounded front ([y]-[ø]-[œ]-[ɶ]) and central ([ʉ]-[ɵ]-[ɞ]) monophthongs and unrounded back ([ɯ]-[ɤ]-[ʌ]-[ɑ]) monophthongs the distance between close and close-mid vowels is much larger than the distances between close-mid, open-mid and open vowels.

It can be observed (Figure 5b) that for vowels of the same height and lip rounding category the values of F2 decrease moving in direction from front to back, e. g., [i]-[ɨ]-[ɯ] and [y]-[ʉ]-[u]. Rounded counterparts of unrounded vowels are characterized mainly by decrease in the value of F2 (while in the pair [a]-[ɶ] the value of F1 is reduced more than that of F2). For distinguishing rounded and unrounded counterparts of front vowels the decrease in values of F3 is even more significant than the changes in values of F2. For close and close-mid vowels the F_2 of rounded front monophthongs are lower than those for unrounded central, as well as the F_2 of rounded central monophthongs are lower than those of unrounded back. So, there is clearly the pattern in the IPA vowel system that is lost if these vowels are viewed in any of simple vowel planes (Figure 1 to Figure 4 and other similar).

Discussion

It seems that there are very few languages in the World contrasting rounded and unrounded central vowels (Lindau 1975, 17), as well as contrasting central and back unrounded non-low vowels (Lindau 1975, 19; Ladefoged,

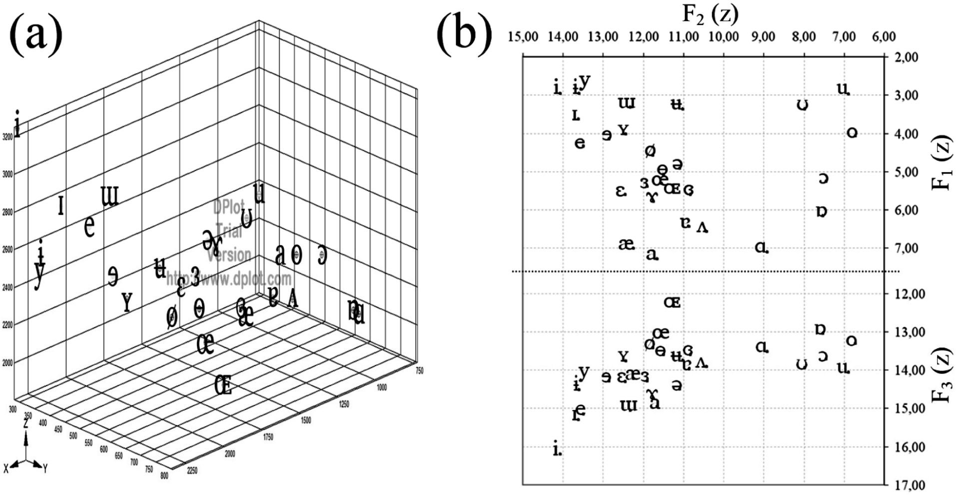

Figure 5. **The placement of IPA monophthongs based on their formant data** (Hayes 2011) in **(a)** the three-dimensional vowel space created using DPlot software (X axis shows the values of F2, Y – values of F1 and Z – values of F3 in Hz) and in **(b)** the combined plane of values of F_2 vs. F_1 and F_2 vs. F_3 in barks (z)

Maddieson 1998, 290–292) therefore monophthongs of most vowel systems are usually well separated in simple vowel planes. Judging by the vocograms included into L. Canepari's book (Canepari 2007, 307–438) and by the description of vowel systems in University of Victoria Phonetic Database (UVPDb, 1999) the World's languages tend to maintain as large perceptual contrast as possible between the entities of their vowel systems. More than 2/3 of vowel systems (43) out of 57 in University of Victoria Phonetic Database employ vowels resembling those from the set of the Primary Cardinals alone (Latvian is among them; Lithuanian is not included in this database, but it would belong to this group, too) or in combination with some central or unrounded back vowels. If central and unrounded back vowels are used in the same vowel system they usually differ either by height or by lip rounding or both. The same statement is true for the remaining 14 vowel systems additionally including front rounded vowels. The perceptual distance between close constituents (Figure 5) of such systems, e. g., in case when /y/ and /ɨ/ are used, is increased by some modification of the production of one or both constituents. As a result – inside one vowel system vowels of different quality usually are well separated even in simple vowel planes. The use of vowel plots in simple vowel planes based on the formant values is popular among phoneticians not only because it is convenient for the illustration purposes, but also because these plots resemble monophthong mapping in the IPA vowel quadrilateral. Nevertheless the reader of such plot should keep in mind that although the formant values depend on the configuration of the vocal tract the acoustic or psycho-physical dimensions used in the plot do not exactly correspond either to the articulatory or perceptual dimensions.

Conclusions

1. It is possible to establish the relation between the articulatory and acoustic distinctive features so that [±low] corresponds to [±compact], [±back] to [±grave], [±rounded] to [±flat], but it is impossible to define one universal set of acoustic parameters that could differentiate all the vowels by height/compactness, backness/graveness and lip rounding/flatness.

2. The value of F1 is related to the feature of the tongue height/compactness, but it is impossible to define one numeric value that would allow drawing borders between, e. g., close and close mid or open mid and open vowels, because these values will differ for front, central and back vowels, as well as for the rounded and unrounded vowels of the same degree of backness. If the backness/gravity is associated with the value of F2, of F2′ or of the distance between F2 (or F2′) and F1 it is impossible to set the border between front, central and back vowels of different degree of height or lip rounding.

3. The tonotopic distance between values of F2 and F3 can not be used to separate front vowels from central and unrounded back vowels, because it is smaller than 3.5 z for all of them. The tonotopic distance between values of F1 and F2 can not be used to separate back vowels from central and front vowels, because it is larger than 3.5 z for unrounded back vowels except [ɑ] (in case of the IPA vowels – even for rounded [u] and [ʊ]).

4. Since both living Baltic languages, i. e., Latvian and Lithuanian, have monophthongs corresponding to a modified set of the Primary Cardinals it is possible to classify their monophthongs on the basis of acoustic features and parameters suggested by G. Fant and other authors bearing in mind that neither acoustic nor articulatory characteristics correspond exactly to the idealized representation of vowels in the IPA vowel quadrilateral.

KLASIKINIŲ AKUSTINIŲ POŽYMIŲ PASIRINKIMO PROBLEMOS ATLIEKANT FONOLOGINĘ BALSIŲ INTERPRETACIJĄ

Santrauka

Šio straipsnio tikslas – aptarti klasikinius akustinius artikuliacinių požymių koreliatus ir nustatyti, kiek jie yra universalūs aprašant įvairių kalbų fonemų diferencinius požymius. Taip pat svarbu panagrinėti, kurie akustinių ir psichofizinių ypatybių grafinio vaizdavimo būdai padėtų patikslinti fonologinę vienos ar kitos kalbos garsų interpretaciją.

Pirmiausia straipsnyje apžvelgiami spektriniai tiriamųjų balsių parametrai, analizuojami įvairūs jų pateikimo modeliai. Pavaizdavus kardinalių ir TFA (an. *IPA*) balsių formančių ar jų skirtumų (pvz., F_2 ir F_1, F_2' ir F_1, F_2-F_1 ir F_1, F_2'-F_1 ir F_1 ir pan.) reikšmes (išreikštas hercais ir barkais) dvimatėje – akustinėje ar psichofizinėje – erdvėje, matyti, kad priešakiniai, tarpiniai ir užpakaliniai nelūpiniai balsiai persidengia. Dvimačiai fonologinių sistemų modeliai ne visai aiškiai perteikia garsų diferencinius požymius ir tuo atveju, kai pasirenkami Gunnaro Fanto pasiūlyti akustiniai artikuliacinių požymių koreliatai ‘bemolinis’ (an. *flat*) ir ‘ne žemasis’ (an. *spread*).

Be dvimačių, išanalizavus ir trimačius modelius paaiškėjo, kad geriausiai skiriami balsiai, jei jų požymiai pateikiami trimatėje erdvėje, tačiau svarbu nepamiršti, kad artikuliacinių ir akustinių ar psichofizinių garsų ypatybių negalima sieti tiesmukiškai, nes vienos artikuliacinės ypatybės pokytis paprastai lemia ne vieną, o kelis atitinkamų formančių reikšmių pokyčius.

REFERENCES

Bladon, R. A. (Anthony) W., Gunnar Fant 1978, A two-formant model and the cardinal vowels, in *Speech Transmission Laboratory – Quarterly Progress and Status Report* 1/1978, Stockholm: KTH, 1–8.

Canepari, Luciano 2007, *Natural phonetics and tonetics (Articulatory, auditory & functional)*, Muenchen: LINCOM GmbH.

Carlson, Rolf, Björn Granström, Gunnar Fant 1970, Some studies concerning perception of isolated vowels, in *Speech Transmission Laboratory – Quarterly Progress and Status Report* 2–3/1970, Stockholm: KTH, 19–35.

Chistovich, Ludmila, Valentina Lublinskaya 1979, The ‘Center of gravity’ effect in vowel spectra and critical distance between the formants: psychoacoustical study of the perception of vowel-like stimuli, in *Hearing Research* 1, ELSEVIER, 185–195.

Fant, Gunnar 1973, *Speech sounds and features*, Cambridge: The MIT Press.

Fant, Gunnar 1983, Feature analysis of Swedish vowels – a revisit, in *Speech Transmission Laboratory – Quarterly Progress and Status Report* 2–3/1983, Stockholm: KTH, 1–18.

Hayes, Bruce 2011, Vowel Chart with Sound Files, http://www.linguistics.ucla.edu/people/hayes/103/charts/VChart/ (2013 11 07).

Jakobson, Roman, Gunnar Fant, Morris Halle 1963, *Preliminaries to speech analysis*, 4th edition, Cambridge: The MIT Press.

Ladefoged, Peter, Ian Maddieson 1998, *The sounds of the world's languages*, Oxford: Blackwell Publishers Ltd.

Lindau, Mona 1975, Vowel features, in *Working Papers* 11, Lund: Phonetics Laboratory, Dpt. of General Linguistics, Lund University, 1–42.

Traunmüller, Hartmut 1981, Perceptual dimension of openness in vowels, *Journal of the Acoustical Society of America* 69, New York: American Institute of Physics, 1465–1475.

Traunmüller, Hartmut 1988, Analytical expressions for the tonotopic sensory scale, in *PERILUS* 8, Stockholm: University of Stockholm, 93–102.

UVPDb 1999 – University of Victoria Phonetic Database, Version 4, September, 1999, Victoria: Speech Technology Research Ltd.

Wood, Sidney 1975, The weaknesses of the tongue-arching model of vowel articulation, in *Working Papers* 11, Lund: Phonetics Laboratory, Dpt. of General Linguistics, Lund University, 55–107.

Juris GRIGORJEVS
LU Latviešu valodas institūts
Akadēmijas laukums 1
LV-1050 Rīga
Latvia
[*jugrig@latnet.lv*]