Linguistische Berichte

Heft 230

Herausgegeben von
Günther Grewendorf
und Arnim von Stechow

BUSKE

Linguistische Berichte

Redaktion

Günther Grewendorf
Eric Fuß

Johann W. Goethe-Universität
Fachbereich Neuere Philologien
Institut für Linguistik
Grüneburgplatz 1
D-60629 Frankfurt am Main
Tel. +49-69-798-323 98
Fax +49-69-798-323 99
E-Mail: lb@lingua.uni-frankfurt.de

Die *Linguistischen Berichte* sind in Bezug auf Gegenstände und Methoden der Linguistik auf maximale Offenheit hin ausgerichtet, halten im Hinblick auf die zugrunde gelegten wissenschaftlichen Standards aber an einem hohen theoretischen und empirischen Anspruch fest. Eingereichte Manuskripte werden von anonymen Gutachtern beurteilt (»Peer Review«).

Eine Auswertung der *Linguistischen Berichte* erfolgt in: BLLDB (Bibliographie Linguistischer Literatur Datenbank), CSA Arts & Humanities, IBR (Internationale Bibliographie der Rezensionen geistes- und sozialwissenschaftlicher Zeitschriftenliteratur), IBZ (Internationale Bibliographie der geistes- und sozialwissenschaftlichen Zeitschriftenliteratur)

Jährlich erscheinen vier Hefte (Februar, Mai, August und November) mit einem Umfang von je etwa 128 Seiten. Zudem kann jährlich ein Sonderheft erscheinen, das den Abonnenten mit einem Nachlass von 15% auf den jeweiligen Ladenpreis geliefert wird.

Das Institutsabonnement (Print- und Onlineausgabe) kostet € 138,- pro Jahr, das Privatabonnement (Print- und Onlineausgabe) € 86,- (jeweils zuzüglich Versandspesen: Inland € 7,- / Ausland € 16,-). Der Preis für ein Einzelheft beträgt € 42,-. Kündigungsfrist: 6 Wochen zum Jahresende.

Neue Abonnements nehmen der Helmut Buske Verlag, Richardstraße 47, 22081 Hamburg, Tel. 040 / 299 95 80, Fax 040 / 299 36 14, E-Mail: info@buske.de sowie jede Buchhandlung entgegen.

ISSN 0024-3930

Werkdruckpapier: alterungsbeständig nach ANSI-Norm resp. DIN-ISO 9706, hergestellt aus 100% chlorfrei gebleichtem Zellstoff. Printed in Germany.

www.buske.de

Inhalt

LB 230/2012

Beiträge aus Forschung und Anwendung

Romanische Syntax

Phonologie

Grammatiktheorie

Rezensionen

Informationen und Hinweise

Phonologie

Acoustic and Auditory Characteristics of the Latvian Monophthong System

Juris Grigorjevs

Abstract

This paper is intended to show the attempt of defining the acoustic and auditory targets of the Latvian monophthong system. To define the acoustic targets vowels produced in isolation are analyzed since such a production is the closest to the mental prototype. Different methods of interspeaker normalization are utilized in order to make productions by informants of different gender comparable. The auditory targets are registered in perception experiments using synthesized two-formant stimuli. The deviation of experimental results for the front vowels from the formant values predicted by formulas is explainable by the lack of rounded front vowels in Latvian. Certain acoustic characteristics of the Latvian monophthongs are employed for the phonological classification, and the use of particular IPA symbols is suggested comparing the acoustic data of Latvian monophthongs with those of Cardinal vowels.

1 Introduction

When considering the practical realization of a phoneme (an abstract unit of speech) the complex phenomenon of a sound target (Rosner & Pickering 1994: 281–285) should be taken into account. Since people learn to speak imitating pronunciation of others, it is plausible that they store in memory only some kind of ideal auditory representation (auditory target) of each speech sound and the information about the deviations from it based upon different language-inherent rules, instead of storing all the possible productions of the same sound by different speakers. This ideal auditory representation is associated with the acoustic signal of certain quality (acoustic target). The acoustic signal is a mediator between the speaker and the listener, and conveys the message from one to the other. It is known that nearly the same acoustic signal can often be generated by several different articulatory postures. On the one hand it is good, since it enables such phenomenon as articulatory compensation. On the other hand it prevents associating certain acoustic characteristics of a sound with the only possible corresponding articulatory posture. Since the recognition of the utterance is based mainly on the acoustic signal and articulatory gestures usually provide only extra cues that are helpful, but not essential, the information about the acoustic and auditory targets is very important for the description of the sound system of any language. The evolution of speech technology is also tightly connected with the available information on the physical characteristics of speech.

Speech is the oldest form of language, and is still today the main steppingstone for human communication. The result of speaking is a chain of deliberately produced sounds, which can be described by the same physical characteristics as any sound in the world. Vowels due to their sonority and relative consistency form the basis for speech communication.

Properties of the Latvian vowels have been mainly described from the point of view of their articulation. The most extensive acoustic research has been done on vowel quantity and tonal modulation (Liepa 1979), although some research of the spectral quality of vowels has been done, too (Stelle 1971, Markus 1983, Sarkanis 1993, Bond 1994 etc.). The data on the vowel quality acquired in these investigations could not be used for the characterization of the vowel targets of Latvian. There are different reasons for it, e. a. sub dialect, phonetic environment, research methods etc., therefore it was necessary to undertake a new acoustic and auditory research of vowels. This research and description of vowel targets covers only the monophthongs of Standard Latvian. The system of Latvian diphthongs being a subject of equal importance is worth a thorough investigation in its own right.

2 Acoustic Targets of the Latvian Monophthongs

A wholesome speech communication is impossible without the mediation of an acoustic signal between the speaker and the listener. The aim of the speaker is to encode the idea in a string of such acoustic signals, that the hearing system of the listener would be able to decode it, comparing the acoustic signals with the auditory targets stored in his/her memory. It is likely, that in the consciousness of every language speaker there exists an auditory target for every speech sound used to distinguish the meaning. This auditory target should be shaped on the basis of an acoustic signal of a certain quality. To draw up a classification of sounds in any language and to describe the phonological rules operating in it, it is important to clarify – what kind of an acoustic signal (acoustic target) corresponds to an auditory target and how it changes under the influence of the phonetic and phonological processes occurring in language. Hence the necessity of investigating acoustic targets. For the material of the investigation of the acoustic vowel targets isolated monophthongs /#V#/ were chosen[1], because when

---

[1] For the sake of convenience, in this article the Latvian low mid vowel is transcribed using IPA symbol /a/, the mid back vowel – symbol /o/, but the length of all long vowels is marked using a colon (:) instead of the IPA length diacritic (ː).

producing an isolated vowel the speaker usually tries to imitate the mental image of a vowel (auditory target) as close as possible, and the quality of the vowel is not influenced by the surrounding sounds forming the phonetic environment.

For the present research recordings of 6 informants were used. Every vowel was recorded pronounced 3 times by each informant. The 6 informants (3 males and 3 females) were chosen as having faultless articulation of a larger group of Standard Latvian speakers on the basis of prior investigation data (Grigorjevs 1995), where the pronunciation of vowels in a word context in carrier sentences by 20 males and 5 females was investigated. The recordings were analyzed by the computer program for the acoustic analysis CSRE 4.5, registering the duration of each vowel, its fundamental frequency (F0) and the frequency values of the first four formants (F1, F2, F3 and F4). From the acquired data for each vowel statistical means and standard deviations were calculated (Table 1 and 2) using the computer program SPSS for Windows 7.0.

Vowel	**F0**		**F1**		**F2**		**F3**		**F4**		**Duration**	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
[i:]	117	11	286	42	2189	76	2947	155	3734	286	487	156
[i]	128	7	312	52	2182	73	2894	196	3700	275	207	34
[e:]	114	14	426	26	1975	92	2644	117	3469	75	481	159
[e]	118	9	458	37	1921	66	2630	129	3483	127	208	40
[æ:]	113	16	681	58	1532	85	2488	51	3442	125	493	118
[æ]	111	9	675	60	1564	81	2485	44	3475	225	219	33
[a:]	110	12	657	50	1005	52	2668	70	3452	109	501	114
[a]	108	9	690	63	1044	69	2546	144	3458	168	229	40
[o:]	112	11	468	31	731	55	2503	167	3371	116	488	127
[o]	112	10	509	34	787	53	2474	81	3334	102	229	44
[u:]	120	10	325	41	589	51	2239	104	3210	88	436	110
[u]	120	8	323	54	642	79	2350	158	3121	60	192	25

Table 1. The values of the fundamental frequency, the first four formants (in Hz) and the duration (in ms) for short and long isolated monophthongs of the Standard Latvian. The value of statistical mean (Mean) and standard deviation (SD) is calculated on the basis of pronunciation data, when every monophthong was articulated three times by three **male** informants.

On the basis of analysis of the acoustic measurement data the following conclusion was drawn – the quality of long and short monophthongs produced in isolation differs very little, therefore it is reasonable to single out 6 vowel pairs of equal quality, where the difference between the members of each pair is determined by the quantity ratio (on the average the ratio of the short vs. long vowel is 1:2.3).

Vowel	F0		F1		F2		F3		F4		Duration	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
[i:]	223	29	360	30	2788	107	3336	182	4367	236	458	88
[i]	226	25	352	41	2767	54	3287	147	4353	208	224	68
[e:]	217	26	604	42	2411	75	3038	159	4359	173	492	74
[e]	221	27	634	33	2411	58	3054	184	4329	186	254	59
[æ:]	215	27	881	109	1863	206	2986	218	4299	117	473	71
[æ]	218	32	916	90	1993	137	3000	265	4299	150	245	64
[a:]	211	25	892	103	1313	91	2924	167	3993	168	486	54
[a]	218	30	938	120	1369	141	2951	253	4071	151	250	68
[o:]	214	26	602	91	960	52	2780	254	3884	86	490	75
[o]	216	33	670	44	1038	60	2897	219	3960	109	253	63
[u:]	213	28	374	44	778	40	3073	279	4014	118	428	83
[u]	215	27	412	38	773	47	3000	267	4066	146	197	37

Table 2. The values of the fundamental frequency, the first four formants (in Hz) and the duration (in ms) for short and long isolated monophthongs of the Standard Latvian. The value of statistical mean (Mean) and standard deviation (SD) is calculated on the basis of pronunciation data, when every monophthong was articulated three times by three *female* informants.

If the acoustic data of the vowels are grouped according to the gender of the speakers (male and female), it can be observed that the data of vowels articulated several times by different informants of the same gender form a rather compact, non-overlapping vowel quality zone in the F2/F1 plane (Figure 1 and 2). Thus it is possible to register distinctly separated vowel zones, which correspond to the traditional description of the Latvian monophthongs by the horizontal and vertical position of the highest point of the tongue body.

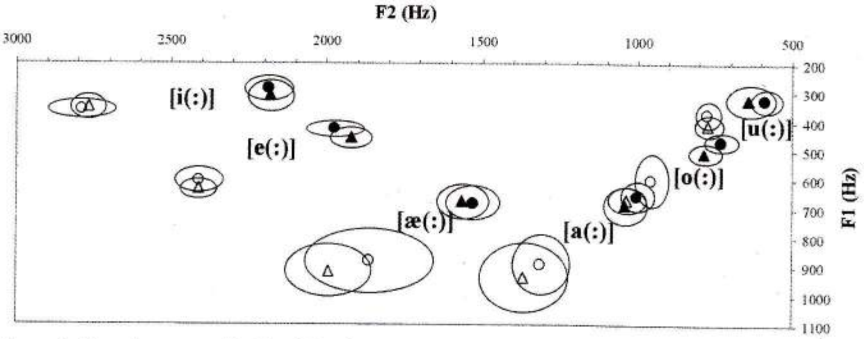

Figure 1. The placement of isolated Latvian monophthongs in the acoustic F2/F1 plane on the basis of the statistical data (Table 1 and 2). The black symbols depict the male, but the white ones – the female data. The circles are used for the long, but the triangles for the short monophthongs; those are based on the values of the statistical means. The ellipses are based on the values of standard deviations and they depict the possible vowel zones.

The arrangement of the female data in the acoustic F2/F1 plane form a vowel space which resembles the male vowel space but is stretched in the dimensions of F1 and F2, as well as moved away from the zero crossing of the axis. It can be explained from the acoustical point of view in terms of female resonant cavities being smaller than those of the males, thus resonating at higher frequencies.

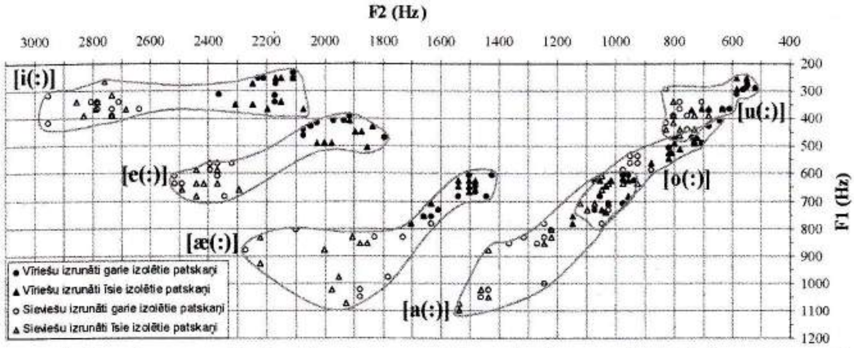

Figure 2. The placement of isolated Latvian monophthongs in the acoustic F2/F1 plane. The coordinates of the vowel points are defined by the values of F1 and F2 acquired in the acoustic measurements. The symbols are used as in Figure 1. The data points belonging to the same quality category are encircled with a grey line forming a dispersion zone of each vowel. The female vowel points in zone [o(:)] overlap the male points in zone [a(:)], and the female points in zone [u(:)]slightly overlap the male points in zone [o(:)].

One has to take the size factor into account if one is going to compare vowel qualities of male and female groups. To compare the female pronunciation data to the male data the procedure of interspeaker normalization has to be performed first.

There are three interspeaker normalization methods widely utilized depending on the goal of an investigation and the use of its results. The simplest of them is the method of uniform normalization, which takes into account only the length differences between the acoustic resonators formed by male and female supralaryngeal cavities (Mol 1970, Nordström & Lindblom 1975). Instead of calculating the normalization coefficient from the vocal tract length measures, it is often calculated from the acoustic data of specific formants, which are related to the vocal tract as a whole or to a particular cavity (Fant 1975).

$$k = \left(\frac{F_{2if}}{F_{2im}} - 1 \right) \cdot 100\%$$

Formula 1. In this formula ***k*** stands for the calculable normalization coefficient, F_{2i} – the frequency value of the second formant of the vowel [i(:)], ***m*** marks the male pronunciation, and ***f*** – the female pronunciation.

In the course of present research different methods for the calculation of the normalization coefficient were tested. The best result in the uniform normalization of the Latvian vowel data (Figure 3 and 4) was achieved using coefficient $k = 27\%$ ($k = k_{2i}$), that was calculated from F2 data of the vowel [i(:)] without the optimization with G. Fant's universal coefficients (Fant 1975: 8) according to the formula 1:

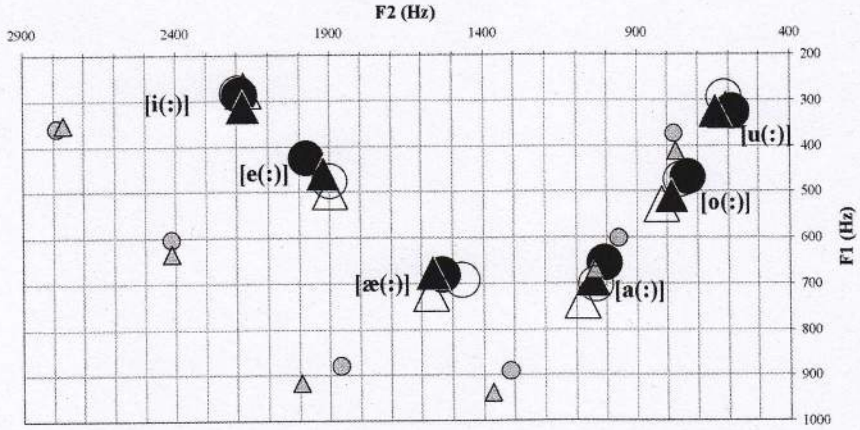

Figure 3. The placement of the female vowel data points in the acoustic F2/F1 plane before (the grey symbols) and after (the white symbols) the uniform normalization with coefficient k=27%. The male data (the black symbols) are used as a reference. The coordinates of the data points are defined by the mean values in Table 1 and 2.

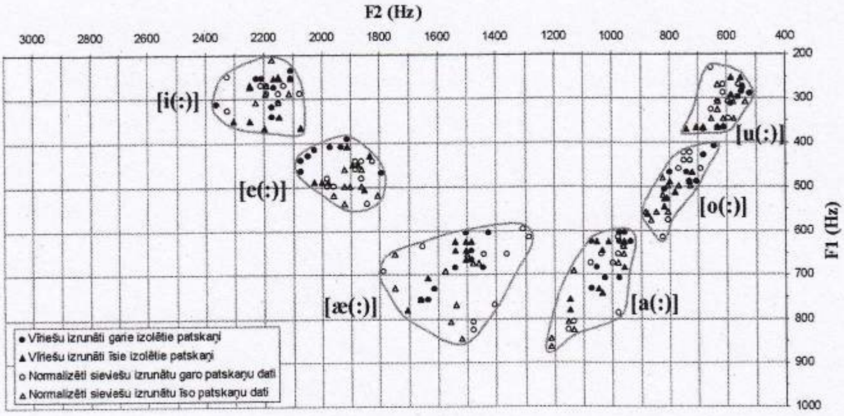

Figure 4. The placement of the female vowel data points (the white symbols) in the acoustic F2/F1 plane after the uniform normalization with coefficient k=27%. The male data (the black symbols) are used as a reference. The coordinates of the male data points are defined by F1 and F2 values obtained in the acoustic measurements, but the female values are normalized with the coefficient. The grey line encircles the data points belonging to the same vowel category.

The same numeric value of the coefficient (k = 27%) can be obtained if the coefficient for the uniform normalization is calculated from the F2 data of all long and short vowels of Latvian; however preference is given to the calculation from the F2 data of the vowel [i(:)], because it is theoretically bound with the length ratio of male and female pharyngeal cavities (Fant 1975).

G. Fant has indicated that after the uniform normalization differences between male and female data are reduced significantly, however not fully. The remaining differences can be reduced by half if the non-uniform normalization procedure is used instead of the uniform data normalization (Fant 1975: 13).

G. Fant concludes that “the non-uniform method removes a greater part of the language universal trends than the uniform method thus sharpening up any dialectal differences” (Fant 1975: 12). The advantage of the non-uniform normalization is in the fact that it is category specific and takes into account not only differences in resonator total lengths, but also the length ratios of the mouth resonator to the pharyngeal cavity. This is achieved by calculating separate normalization coefficients for each formant of each vowel. The relative stability of the mouth to pharyngeal cavity ratio ensures the perceptual contrast and invariance of vowels. This stable cavity ratio can be achieved by modifying, if necessary, the articulation of specific sounds by one group of speakers.

Coefficients for the non-uniform normalization were calculated for each formant of each vowel according to the following formula:

$$k_n = \left(\frac{F_{nf}}{F_{nm}} - 1 \right) \cdot 100\%$$

Formula 2. In this formula ***k*** stands for the calculable normalization coefficient, ***n*** – for the number of formant and the corresponding normalization coefficient, ***m*** marks the male pronunciation, and ***f*** – the female pronunciation.

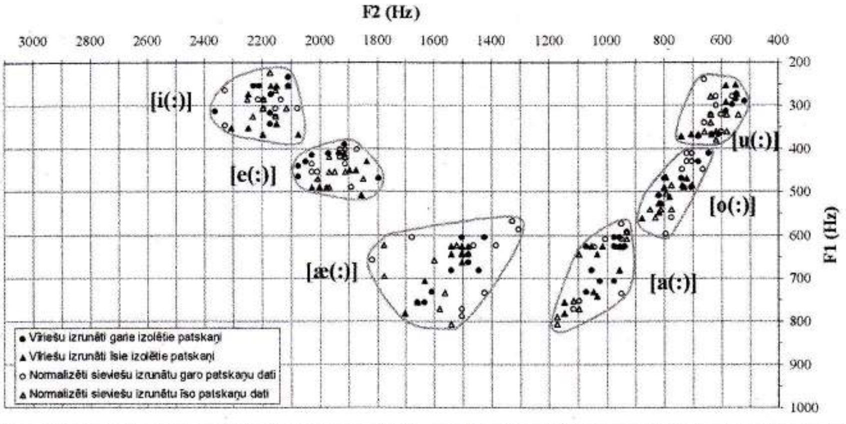

Figure 5. The placement of the female vowel data points (the white symbols) in the acoustic F2/F1 plane after the non-uniform normalization with the coefficients from Table 3. The male data (the black symbols) are used as a reference. The coordinates of the male data points are defined by F1 and F2 values obtained in the acoustic measurements, but the female values are normalized with the corresponding coefficients. The grey line encircles the data points belonging to the same vowel category.

During the investigation it was discovered that for the non-uniform normalization of the Latvian vowel data it is advisable to use coefficients that are calculated from the data obtained by measuring vowel pronunciation in equal or very similar phonetic environments.

The best results for the normalization of isolated vowels (Figure 5) were achieved using coefficients calculated from the data of isolated vowels (Table 3).

Vowel	k1	k2
[i]	20%	27%
[e]	40%	24%
[æ]	33%	25%
[a]	36%	31%
[o]	31%	32%
[u]	22%	26%

Table 3. Normalization coefficients k1 and k2 common for the long and short Latvian vowels calculated separately for F1 and F2 of each vowel from the mean data of long and short vowels in Table 1 & 2.

These coefficients yield good results not only for normalizing mean data of female pronunciation, but also for every separate female pronunciation data point obtained in the measurements (Figure 5).

A hypothesis was made that, to find a universal set of normalization coefficients which would be equally efficient in eliminating sex determined interspeaker differences in all phonetic environments, these coefficients have to be calculated from the data of both isolated vowels and vowels in all possible phonetic environments.

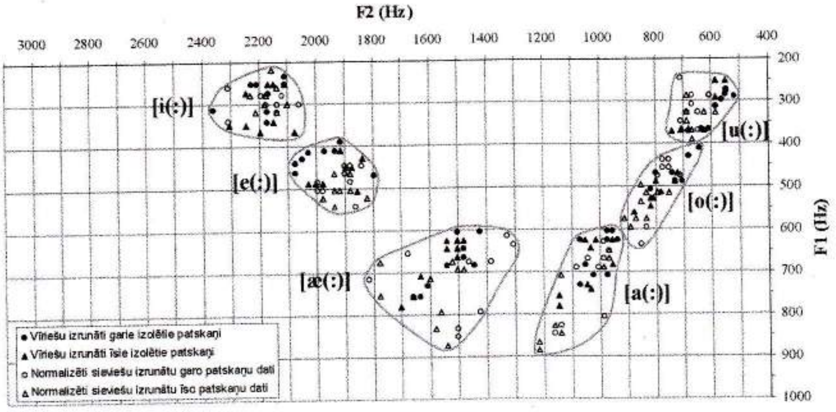

Figure 6. The placement of the female vowel data points (the white symbols) in the acoustic F2/F1 plane after the non-uniform normalization with the universal coefficients from Table 4. The male data (the black symbols) are used as a reference. The coordinates of the male data points are defined by F1 and F2 values obtained in the acoustic measurements, but the female values are normalized with the corresponding coefficients. The grey line encircles the data points belonging to the same vowel category.

Since the author of the present research had an access only to data of isolated vowels and vowels pronounced in the phonetic environment of /t/ (Grigorjevs 1995), the normalization coefficients were calculated on the basis of these data (Table 4). The coefficients are universal only for vowels in zero and /t/ phonetic environments. In all the other environments they should be regarded as quasi-universal. As stated before, the calculation of genuinely universal normalization coefficients should involve the use of vowel data from all possible phonetic environments.

Vowel	k1	k2
[i]	22%	28%
[e]	26%	26%
[æ]	23%	25%
[a]	24%	26%
[o]	23%	23%
[u]	19%	17%

Table 4. The universal normalization coefficients k1 and k2 common for the long and short Latvian vowels calculated separately for F1 and F2 of each vowel from the mean data of the long and short isolated vowels (Table 1 and 2) and vowels in the phonetic environment of /t/ (Grigorjevs 1995).

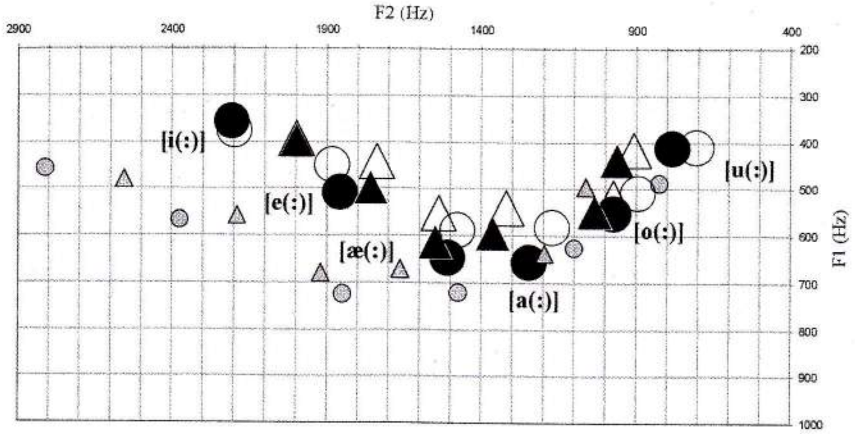

Figure 7. The placement of the female vowel data points in the acoustic F2/F1 plane before (the grey symbols) and after (the white symbols) the non-uniform normalization with the universal coefficients from Table 4. The male data (the black symbols) are used as a reference. The coordinates of the data points are defined by the mean values for vowels in the phonetic environment of /t/ (Grigorjevs 1995).

The non-uniform normalization with the universal coefficients calculated from the data of isolated vowels and vowels pronounced in the phonetic environment of /t/ gave satisfactory results reducing interspeaker differences in the data of isolated vowels (Figure 6) and the data of vowels in the phonetic environment of /t/ (Figure 7).

Both methods described are associated with the production phase in the process of speech communication, because they are normalizing differences in the acoustic data determined by the articulatory mechanisms. The normalization of the acoustic data is possible also from the viewpoint of speech perception, taking into account the anatomy and functions of the hearing system. To perform such normalization the physical units of acoustic measurements have to be changed to psychophysical ones reflecting the subjective perception of sound. Since the perception of pitch and spectral composition of a sound, as well as the perception of a sound's energy, i. e., loudness is of logarithmic nature (especially in frequencies over 1000 Hz), all the most popular psychophysical scales are also logarithmic – at least to some extent. The transformation of the acoustic data into psychoacoustic alone does not remove the dissimilarity of male and female pronunciation data, but it balances the spread of the vowel space in the horizontal and vertical dimensions. In hearing research it has been found that the frequency analyzer located in the inner ear reacts to sound as to the matrix of energy.

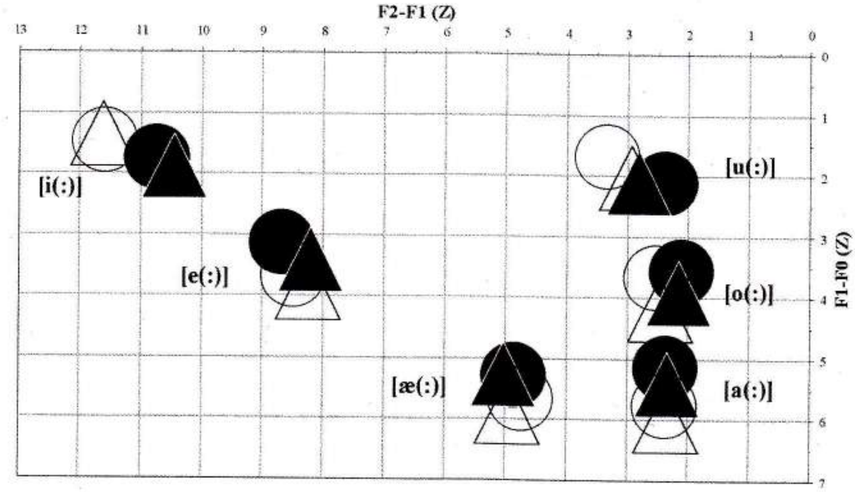

Figure 8. The placement of the Latvian vowel data points in the psychophysical plane. The coordinates of the data points are defined by the tonotopic distances F1–F0 and F2–F1 in Barks (Z), estimated from the mean values of the acoustic data (Table 1 and 2) transformed to psychoacoustic units using Traunmüller's formulas. The female data are shown by the white symbols, the male data – by the black symbols. The diameter of the symbols is 1 Z that relatively corresponds to the size of vowel zones.

The pattern of energy distribution in the spectrum of a sound and the corresponding excitation pattern along the Basilar membrane of the inner ear are more important in perception of the sound than precise characteristics of each peak of spectral energy or displacement of Basilar membrane (Traunmüller 1981). According to the Space-pattern theory, every sound can be characterized by tonotopic distances between the peaks of the spectral energy generated by resonances of the vocal tract. Tonotopic distances between the fundamental frequency and the first formant (F1–F0) and between the first and the second formant (F2–F1) were chosen for characterization of the vowel system of Standard Latvian, because the first of these distances is closely related to the openness of the vowel, and the second to its frontness or backness. Characterizing vowels with these tonotopic distances, the optimum results of interspeaker normalization (Figure 8 and 9) can be achieved using the Bark scale (Miller 1989: 2119).

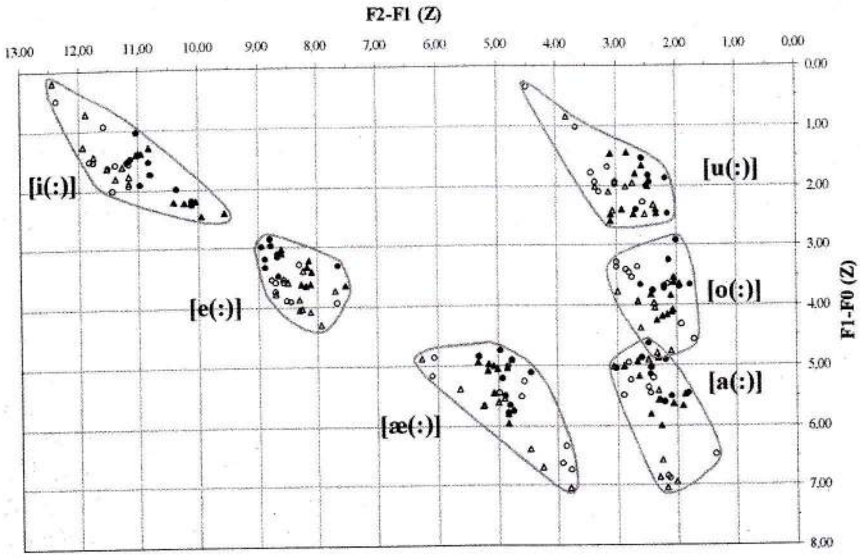

Figure 9. The placement of the Latvian vowel data points in the psychophysical plane. The coordinates of the data points are defined by the tonotopic distances F1–F0 and F2–F1 in Barks (Z), estimated from the values of the acoustic data obtained in measurements and transformed to psychoacoustic units using Traunmüller's formulas. The grey line encircles the data points belonging to the same vowel category.

To calculate the perceptually important psychophysical units Barks, which correspond to the placement or sequential number of the auditory filter along the Basilar membrane, from the data of vowel formant centre frequencies in Hertz the formula (Formula 3) suggested by H. Traunmüller (Traunmüller 1988) was used:

$$z = \left(26.81 \cdot \frac{f}{1960 + f}\right) - 0.53$$

Formula 3. In this formula z is the value of Critical Bands in Barks, and f – frequency in Hertz.

If $z < 2.0$ Barks, to account for the differences in perception of low frequencies the other Traunmüller's formula (Formula 4) was used:

$$z' = z + 0.15 \cdot (2 - z)$$

Formula 4. In this formula z is the value of Critical Bands in Barks, which is calculated by the previous formula and is under 2 Barks ($z < 2.0$), and z' – the final value of Critical Bands in Barks.

It can be observed that after the transform to psychophysical units and using values of tonotopic distances instead of formant centre frequencies, the difference between the placement of the male and female data points has been reduced to a great extent (cf. Figure 2 and Figure 9). In most cases, vowel points are placed less than 1 Bark apart; therefore not only short and long vowels but also vowels produced by the male and female informants are perceived as belonging to the same quality category. The Bark scale is also useful because it allows the transformed acoustic data to be used for phonological classification of vowels (Miller 1989; Fant 1983; Iivonen 1987).

During the investigation, it became clear that any method of interspeaker normalization independently of its type significantly reduced the differences in the acoustic data of vowels pronounced by the male and female informants. The use of a particular method should be considered on each occasion according to the goal of investigation and application field of the obtained results.

3 Auditory Targets of the Latvian Monophthongs

In the prior research carried out by phoneticians of different countries it has been proved that in the listener's perception, the formant frequencies of auditory targets of any vowel seldom corresponded to the formant frequencies that could be observed in the pronunciation of the same vowel by any informant in normal speech conditions (Eek & Meister 1994, Johnson et al. 1993b). It is explained by a more even distribution of the vowel targets and the variants of their possible realization in the psychophysical sphere of perception in comparison with the acoustic sphere of production. The acoustic properties of the articulatory vowel targets are usually associated with the pronunciation of isolated vowels, when an informant consciously tries to pronounce the vowel according to his/her comprehension of the ideal sound of the particular vowel. For the fixation of the perceptual, i. e. auditory vowel target, several variants of the vowel are played to the listener, from which the listener is to choose the variant that is the closest to his/her concept of the ideal sound.

For a long time in acoustic phonetics, a view was maintained that the acoustic quality of a vowel is best characterized by the frequency values (F_1 and F_2) of the first two vowel formants. The value of F1 centre frequency is usually associated with the vertical position of the tongue or mouth opening (meaning the opening between the highest point of the tongue and the palate, not the lip opening), and the value of F2 centre frequency – with the position of the highest point of the tongue in the mouth in the front-back dimension and also with labialization. In fact, the formant frequencies of a given vowel are related to the cavity sizes of the resonator determined by the postures of the speech organs, and to the size ratio of these cavities. This is why it would be advisable to characterize vowels by the place of vocal tract's maximum constriction and not the place of the highest point of a tongue (Rosner & Pickering 1994).

The research on spectral integration (Chistovich & Lublinskaya 1979; Chistovich et al. 1979; Johnson et al. 1993a) suggests that formants, closely located in the spectrum of a sound, in perception form a single peak of spectral energy, the so called "centre of gravity", thus estimating in perception the quality of a vowel which substitutes the information provided by centre frequencies of separate formants. This agrees with the hypothesis that the value of F_1–F_0 (F_1 – frequency value of the first formant, F_0 – frequency value of the fundamental) is more important in determining "closeness" or "openness" of a vowel than the frequency value of F1 (Traunmüller 1981; Traunmüller & Lacerda 1987). The increased value of the effective second formant[2] of synthesized two-formant front vowels in comparison with the value of the second formant of the corresponding four-formant vowels can also be explained by the influence of the "centre of gravity" (Carlson et al. 1970; Bladon & Fant 1978). As the theory of the perceptual domain of vowels is based to a great extent on a research using two-formant vowel stimuli, synthesized two-formant vowels were chosen also for the estimation of the auditory targets of the Latvian vowels. By using this method the results obtained in this investigation are comparable with the results of other similar investigations. Moreover, the use of two-formant vowels in the parametric synthesis of speech allows reducing the load on computer memory resources.

To detect fixed auditory vowel targets in the listeners' memory, phonological experiments were needed. The stimuli would be arranged in such a way that the sequence started with a synthesized four-formant vowel as a reference, followed by two-formant stimuli with gradually changing frequencies of the first or the second formant. In such an arrangement, the reference stimulus provides the listeners information about the target vowel to which the following stimuli have to be compared in order to find the best match. Since the reference stimulus is heard only once in the beginning of the stimuli sequence and the pauses between the stimuli are sufficiently long so that the respondent (listener) looking for the best matching two-formant stimulus is guided not by the quality of the reference stimulus, but by the quality of the auditory vowel target stored in his/her memory. By this procedure it is possible to define both the coordinates of each auditory vowel target and the approximate boundaries of vowel zones in the perceptual space. To begin with, two experiments were carried out with the Latvian respondents using two-formant stimuli with changing F2.

For the first, i. e. pilot experiment synthesized two-formant vowel stimuli were generated[3] in Kay CSL 4300 using LPC synthesis (software version ASL

---

[2] F2' – the effective second formant in synthesized two-formant vowels, where F2' substitutes all formants higher than F1, i. e., F2, F3, F4 etc.

[3] Formants higher than F2 were deleted. The amplitude values of the base vowel as well as the bandwidths of F1 and F2 and the linkage to F0 were preserved. The F2 frequency of the two-formant stimulus was altered in 0.33 Bark increments. Two sequences of two-formant stimuli were created for each vowel (/iː, eː, æː, aː, oː, uː/). The first sequence was organized in a way that it started with the stimulus in which F2 was well below the frequency of a naturally produced base vowel. The F2 frequency of each following stimulus was increased by 0.33 Bark until it reached the frequency that was well above the frequency of a naturally produced base vowel. In the second sequence the same stimuli were arranged in the opposite order. The minimum and the maximum F2 frequency in the sequence was chosen so that it fell, if possible, in the frequency range of another vowel of approximately the same openness. The aim of this was to check if the boundary between two vowels of approximately the same openness, e. g. [eː] and [oː], was perceived with the help of some F2 value falling between the frequencies typical to [eː] and to [oː], or if each of these vowels had a certain range of frequencies (making up the quality zone in F2 dimension) where it was recognized as a realization of a certain phoneme. Outside this range the sound by its auditory quality

4304) proceeding from the LPC spectrum of the base vowels pronounced by one of the informants. The pilot experiment was carried out at the University of Latvia with 16 listeners as experimental subjects: 3 phoneticians, 4 linguists and 9 fourth year students of Linguistics. After instructions, each sequence of stimuli was played to subjects 3 times. The responses of all subjects were summarized and a statistical analysis of the acquired data was carried out. The results showed that every vowel had an F2 zone, inside which the changes of F2 were perceived as quality changes corresponding to allophonic realizations of the same phoneme. If the value of F2 was outside such a zone, the vowel was perceived as "not belonging to the Latvian vowel system". The F2 zones corresponding to Latvian vowels were mutually isolated, i. e. they did not form a common boundary for two vowels of similar openness.

To synthesize stimuli for the second experiment, the synthesis software SENSYN (produced by *Sensimetrics Corp.*) was used. The four-formant reference stimuli were synthesized using the formant data of naturally pronounced vowels. The unvarying F1 frequency of two-formant stimuli also agreed with the F1 frequency of the corresponding vowel, but the F2 frequency was altered in 0.33 Bark increments. The duration of all the stimuli was 300 ms. The fundamental frequency was made equal for all vowels – in the beginning of the vowel it was 90 Hz, during the first 100 ms it gradually increased to 100 Hz, then gradually decreased reaching the value of 80 Hz in the end of the vowel. The amplitudes of spectral harmonics were computed automatically on the basis of the location of formants in the spectrum. The formant bandwidths were also made equal for all vowels: bandwidth of F1 was 60 Hz, of F2 – 90 Hz (for four-formant reference stimuli: F3 – 150 Hz and F4 – 200 Hz). Like in the first experiment, the stimuli were arranged in sequences. There were two stimuli sequences for each vowel: the first with an increasing, the second with a decreasing frequency of F2. Every sequence started with a synthesized four-formant reference vowel. There was a 3-second pause between all stimuli in a sequence, and the pause between two sequences starting with the same reference vowel was 5 second long. Since the results of the pilot experiment pointed to the existence of mutually isolated F2 zones conforming to the quality of each vowel, the number of stimuli could be reduced, as there was no need for the highest or lowest F2 values falling into the zone of a neighboring vowel. The two-formant stimuli in sequences were generated so that the minimum value of F2 was lower and the maximum higher than in the naturally pronounced vowel. The extreme values of F2 were chosen well below and well above the marginal values of the F2 zones determined in the first experiment.

---

should not correspond to any vowel phoneme of Latvian. In the beginning of each sequence there was a synthesized four-formant stimulus (in the process of LPC synthesis all spectral information above F4 in a naturally produced vowel was deleted) that was used as a reference for estimation of the ideal two-formant vowel. The duration of all stimuli in the sequence was equal – 300 ms. There was a 2 second pause between all stimuli in a sequence, and the pause between two sequences starting by the same reference vowel was 5 second long.

7 listeners participated in the second experiment: 1 phonetician and 6 persons whose native language is Latvian and whose pronunciation does not bear traits of dialectal accents. At the time of the experiment, none of the subjects had any speech or hearing disorder that could affect the results. After instructions, each sequence of stimuli was played to the subjects 3 times. The responses of all subjects were summarized and a statistical analysis of the acquired data was carried out. First the mean value of the preferred F2 was calculated for each sequence, and then the value of $F2^i$ was calculated using the mean values of two sequences of each vowel.

Vowel	**Measured F2**		**$F2^i$ (2nd exp.)**		**Calculated F2'**	
	Hz	**Z**	**Hz**	**Z**	**Hz**	**Z**
i:	2177	13,58	2597	14,75	3058	15,81
e:	2048	13,17	2196	13,64	2296	13,98
æ:	1639	11,68	1763	12,17	1704	11,94
a:	1147	9,37	1128	9,26	1152	9,39
o:	868	7,70	872	7,73	869	7,71
u:	699	6,52	688	6,44	702	6,54

Table 5. Values in Hertz and Bark of the second formant in sustained isolated pronunciation by one informant (Measured F2), values of the effective second formant ($F2^i$) obtained in the experiment, values of the predicted effective second formant (Calculated F2').

As expected, the data for back vowels were in close agreement with the theoretically predicted[4] values (Table 5). The results of the second experiment were taken into account in preparing the material for the third experiment.

The stimuli for the third experiment were generated using the same equipment and the same methodology as in the second experiment. Four-formant reference vowels were synthesized on the basis of formant frequency data of naturally produced vowels. The unvarying F2 frequency of two-formant stimuli for each vowel agreed with the F2 frequency of the stimulus, which during the second experiment was most frequently marked as the best match to the reference stimulus of the corresponding vowel (this frequency value was not the same as calculated value of $F2^i$!). The varying F1 frequency was altered in 0.33 Bark increments. All the other synthesis parameters and principles for the stimuli presentation were the same as in the second experiment. The two-formant stimuli sequences were generated so that the minimum value of F1 was lower and the maximum higher than in the naturally pronounced isolated vowel. F1 values for vowels of the same tongue advancement (i. e., front or back vowels) were overlapping – e. g., the highest F1 values of close vowels were chosen in such a way that they corresponded to the F1 values of mid vowels.

---

[4] The predicted values of the effective second formant F2' were calculated from the pronunciation data using formula suggested by Bladon and Fant (Blandon and Fant 1978: 3).

12 subjects (6 of them participated in the second experiment) took part in the third experiment: 1 phonetician and 11 persons whose native language is Latvian and in whose pronunciation there are no traits of dialectal accents. At the time of the experiment none of the subjects had any speech or hearing disorder that could affect the results. After instructions, each sequence of stimuli was played to subjects 3 times. The responses of all subjects were summarized and a statistical analysis of the acquired data was carried out. First the mean value of the preferred F1 was calculated for each sequence, and then the value of $F1^i$ was calculated using the mean values of two sequences of each vowel.

Vowel	Measured F1		$F1^i$ (3rd exp.)		Calculated F1'	
	Hz	Z	Hz	Z	Hz	Z
i:	314	3,17	383	3,85	226	2,26
e:	458	4,55	504	4,95	295	2,91
æ:	681	6,38	730	6,75	–	–
a:	768	7,02	773	7,05	–	–
o:	529	5,17	531	5,19	–	–
u:	328	3,31	314	3,17	233	2,33

Table 6. Values in Hertz and Bark of the first formant in sustained isolated pronunciation by one informant (Measured F1), values of the effective first formant ($F1^i$) obtained in the experiment, values of the predicted[5] effective first formant (Calculated F1').

It can be concluded on the basis of the results obtained for close vowels in the third experiment that there has been neither spectral integration between the fundamental frequency and F1, nor has the frequency value of F1 decreased due to the effect of "perceptual hypersphere" (Table 6). On the contrary, the increase in the preferred value of F1 frequency of open vowels conforms to the assumption that in the perceptual sphere, open vowels become even more open (Johnson et al. 1993b).

To see whether auditory vowel targets differ from acoustic targets significantly, a set of complimentary experiments was prepared and carried out. The aim of these experiments was to establish two-formant stimuli, which were the best match to the synthesized four-formant reference stimuli (not the auditory vowel targets). To achieve this goal, the stimuli in sequences have to be organized in pairs, where one member of the pair is a reference stimulus, and the other – a two-formant stimulus with altered F1 or F2 frequency. To make the comparison of the two stimuli easier, pauses between the members of each pair had to be shorter than or equal with the duration of the stimuli. Such pairs of stimuli were prepared using the material generated for the second and the third experiment. Using this material two complimentary experiments ("matching experiments") were carried out. The same subjects who took part in the second and the third experiment (conventionally referred to as "phonological experi-

---

[5] The predicted effective first formant was calculated as a midpoint between values of F1 and fundamental for vowels where the tonotopic distance between F0 and F1 was less than 3.5 Bark.

ments") participated in these experiments. The responses of all subjects were summarized, and a statistical analysis of the acquired data was carried out corresponding to the principles described above in relation to the phonological experiments.

VOWEL		$F1^i$ ↑ (Hz)	$F2^i$ ↑ (Hz)	$F1^i$ ↓ (Hz)	$F2^i$ ↓ (Hz)	$F1^s$ ↑ (Hz)	$F2^s$ ↑ (Hz)	$F1^s$ ↓ (Hz)	$F2^s$ ↓ (Hz)
aː	**Mean**	757	1141	781	1120	734	1141	781	1101
	SD	36	54	39	30	24	38	50	48
æː	**Mean**	696	1742	720	1811	689	1654	697	1773
	SD	36	162	64	127	36	132	47	203
eː	**Mean**	478	2127	523	2210	447	2084	472	2198
	SD	28	116	58	156	30	248	48	239
iː	**Mean**	331	2587	343	2617	330	2530	336	2546
	SD	55	133	66	235	41	277	45	227
oː	**Mean**	516	890	535	838	522	880	529	863
	SD	29	79	32	42	20	27	33	32
uː	**Mean**	314	701	303	672	314	701	325	675
	SD	27	45	32	54	34	16	25	37

Table 7. The mean values of F1 and F2 and their standard deviations in Hertz calculated from the results of the phonological and matching experiments with the Latvian respondents – arrows pointing up mark the results for corresponding stimuli sequences with the increasing formant frequency value, but arrows pointing down – with decreasing formant frequency value.

The values $F2^s$ and $F1^s$ obtained in the matching experiments are closer to the F2 and F1 values of naturally pronounced isolated vowels than the values $F2^i$ and $F1^i$ obtained in the phonological experiments, but in general they reveal the same tendencies (Table 7). To carry out the analysis and interpretation of the auditory experiments easier, all the obtained data were plotted in the psychophysical F2/F1 plane (Figure 10). The large white dots correspond to the registered values of naturally pronounced isolated vowels, the coordinates of the large grey dots are determined by the theoretically predicted (Bladon & Fant 1978) values of F2' (only front vowels are shown, because the calculated F2' of back vowels are equal to F2 values registered in the acoustic measurements). The results of the phonological experiments are depicted by the small black dots and black ellipses, and the results of the matching experiments – by the small grey dots and grey ellipses. The size of the measured and the predicted data points is chosen as 1 Z in order to mark in the psychophysical F2/F1 plane the zones belonging to a certain quality category and to make the comparison with vowel zones (marked by ellipses) established in auditory experiments easier.

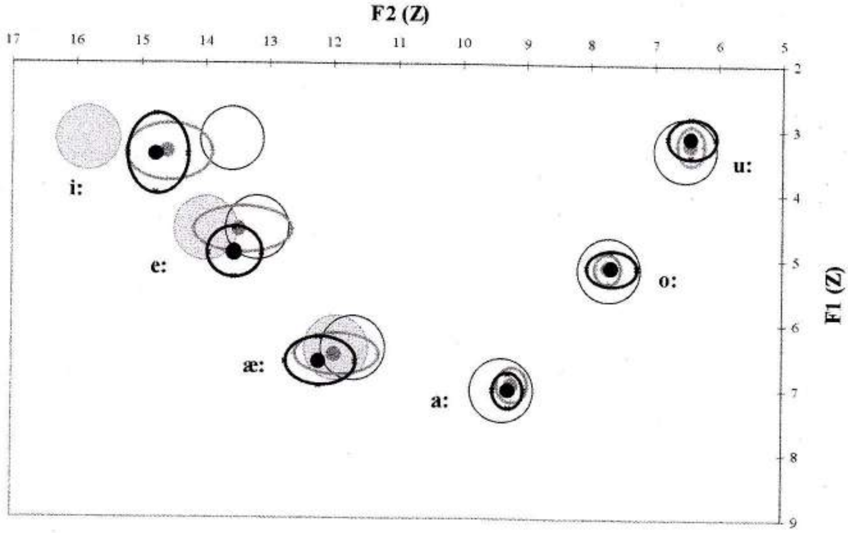

Figure 10. The placement of the Latvian vowels in the psychophysical F2/F1 (Z) plane according to: a) the values registered from the pronunciation data (the large white dots), b) the values of the calculated F2' (the large grey dots), c) the values registered in the phonological experiments (the black dots and black ellipses), d) the values registered in the matching experiments (the little grey dots and grey ellipses).

Comparing the results of the phonological and the matching experiments with the registered and theoretically predicted data (Figure 10) it can be observed that:

- data points of the two-formant stimuli for the vowels [a:], [o:] and [u:] perceived as the best matching to reference in both the phonological and the matching experiments, which are allocated by mean values of F1 and F2, as well as the zones determined by standard deviations of these values and marked by ellipses overlap with the points of the registered pronunciation data;
- the values of the front vowels obtained in the phonological experiments differ more from the values of the pronunciation data than those obtained in the matching experiments;
- the F2 values of the vowels [i:] and [e:] registered in the experiments are amid the F2 values registered in measurements of pronunciation and the F2' values calculated using the formula by Bladon and Fant, but the F2 values of the vowel [æ:] are even higher than the calculated F2'.

Summarizing the observations presented in Figure 10, it can be concluded that the results of the phonological and the matching experiments do not differ significantly – the most noticeable differences can be observed for the vowel [e:] (in dimension of F1) and the vowel [æ:] (in dimension of F2), but expressed numerically they do not exceed 0.5 Bark. This means that the responses of subjects present a little evidence to the effect of perceptual hypersphere. If this effect is valid, the points representing the results of the phonological experiments ought to be displaced towards the periphery of F2/F1 plane much more than the points representing the results of the matching experiments. The data of front vowels obtained in the matching experiments confirm the phenomenon of spectral integration showing the influence of closely spaced higher formants on F2 – the F2 frequency of these vowels is increased in comparison to the registered pronunciation values. A conclusion can be drawn that for the vowels [e:] and [æ:] the spectral integration involves F2 and F3, and for the vowel [i:] – F2, F3 and F4. This could be predicted taking into account the tonotopic distances between the formants of these vowels. The fact that $F2^{i}$ and $F2^{s}$ values of the vowels [i:] and [e:] registered in the experiments do not agree with F2' calculated using formula suggested by Bladon and Fant (Bladon & Fant 1978), does not automatically point to the shortcomings of this formula. The difference can be explained by the lack of rounded front vowels in Latvian. Thus there is no need to mark the perceptual difference between rounded and unrounded front vowels, but the perceptual contrast with back vowels is sufficient even with F2 values being much lower. The data from auditory experiments with Latvian vowels confirm that the subjects have perceived as unrounded front vowels even those two-formant stimuli, whose F2 frequency has been equal to or lower than the F2 frequency in pronunciation of these vowels. In similar experiments with Estonian vowels (Eek & Meister 1994) it was discovered that in order to increase the contrast between rounded and unrounded vowels of the same openness Estonian subjects for unrounded front vowels had chosen two-formant stimuli with F2 frequency even higher than that predicted by formula. Thus the disagreement of the data obtained in the experiments with the Latvian vowels with the data theoretically predicted (calculated by the formula) or observed in languages that have an opposition of rounded and unrounded front vowels, can be explained by the conditions of providing sufficient perceptual contrast (Liljencrants & Lindblom 1972). The same phenomenon can be used to explain an increase in F1 frequency of the vowel [i:] – the increased F1 frequency does not affect the perceptual contrast between [i:] and [e:], which is strengthened by a very high F2 frequency of [i:]. Drawing a general conclusion about the values of F1, it can be inferred that the subjects have given preference to the two-formant stimuli, which have the tonotopic distance between F1 and F0 equal with or close to the distance observed in the pronunciation data. This hypothesis has to be tested by additional research.

4 Phonological Interpretation of the Latvian Monophthong System

The phonological classification of the Latvian vowels is a seemingly unsophisticated task, because the twelve Latvian vowels form six vowel pairs. In stressed positions members of a pair significantly differ only by their quantity. The quality differences between members of each vowel pair are within the scope of 1 Bark, and therefore are not significant for the categorization of these vowels.

Traditionally monophthongs of the Latvian language are classified according their articulation, characterizing them by their length, tenseness, labialization, the row[6] of articulation (depending on the horizontal position of the highest point of the tongue) and by the degree of the tongue elevation (associated with vowel openness). Difficulties start when an attempt to draw up a phonological classification upon objective physical data acquired in an acoustic research is undertaken. The role of speech production (articulation) is to supply sufficient encoding of the thought in order to ensure contrast and maximum distinctiveness in perception (Fant 1983: 1) and, consequently, correct decoding. The speech production process therefore should be considered as output-oriented, i. e. aimed at production of a certain acoustic signal. Not all the acoustic properties of vowels are significant in their perception. The main information for the characterization of vowels is given by the centre frequencies of the first four formants. Often the values of only the first three formants (F1, F2 and F3) are used, and vowels are characterized as points in a three-dimensional vowel space.

To make the acoustic classification closer to the traditional two dimensional articulatory description related to the sagittal cross-section of a head and showing the positions of the speech organs, the concept of an effective second formant (F2′) has been introduced. This system is used for the phonological classification of Latvian vowels, since it is the only relatively developed system the author of the present paper is familiar with, where the distinctive features are clearly defined by their relation to acoustic data. The same system has been used by A. Eek in drawing up the phonological classification of the Estonian vowel system (Eek & Meister 1994). The acoustic-auditory data used for the classification of Latvian vowels were calculated from the pronunciation data in the same way as the data used for the classification of the Estonian and Swedish vowels (Table 8). In one of his articles G. Fant has suggested the method for the phonological classification of vowels (Fant 1983: 13–14) and illustrated it on the material of Swedish. The back vowels are separated from the others using the parameter –(F2–F1), which corresponds to the feature [±grave] (where [+grave] = F2–F1<3.5 Z). These are the only vowels in Latvian that have opposition by lip rounding. Since the feature [flat] expressed by the parameter –(F2′+F1) is strongly related to the lip rounding and velarization it can be successfully used to separate [+grave] vowels on the basis of both of these properties.

To express three degrees of flatness[7] in binary oppositions two binary features are introduced – [±flat] and [±extra flat]. Since the vowel [a(:)] has a high F1 value related to a wide jaw opening and a low tongue position, it has the highest F2′+F1 value (Table 8) and the lowest degree of flatness, and it could be

---

[6] The term 'row' could be substituted with 'place', because it presumes a place of the highest point of the tongue ranging from the front of the hard palate for front vowels to the soft palate for back vowels.

[7] Correspond to three degrees of lip rounding in articulation – unrounded, rounded and labial.

described as [–flat]. The vowel [o(:)] has a medium F2′+F1 value and a medium degree of flatness associated with the tongue rising towards the soft palate (velarization) and a medium lip rounding, therefore it can be described as [+flat]. The vowel [u(:)] has the lowest F2′+F1 value and the highest degree of flatness determined by a very high tongue position (close to the soft palate) and an explicit lip rounding, protrusion, and reducing the area of lip opening, therefore it is described as [+extra flat], which automatically includes [+flat]. During the production of the vowel [o(:)] the lip rounding is significantly smaller, and although it is [+flat], it is therefore classified as [–extra flat].

	[u:]	*[u]*	*[o:]*	*[o]*	*[a:]*	*[a]*	*[æ:]*	*[æ]*	*[e:]*	*[e]*	*[i:]*	*[i]*
F1	3.3	3.3	4.6	5.0	6.2	6.4	6.4	6.3	4.3	4.5	2.9	3.1
F2′	5.7	6.1	6.8	7.2	8.6	8.9	11.6	11.7	14.0	13.8	15.2	15.0
F2′– F1	2.4	2.8	2.1	2.2	2.4	2.4	5.2	5.4	9.8	9.3	12.3	11.9
F2′+F1	8.9	9.3	11.4	12.2	14.8	15.3	17.9	18.1	18.3	18.4	18.0	18.2
F1 – F0	2.1	2.1	3.6	3.9	5.1	5.4	5.3	5.3	3.1	3.4	1.7	1.9
F2 – F1	2.4	2.8	2.1	2.2	2.4	2.3	4.9	5.0	8.7	8.2	10.7	10.5
F3 – F2	8.1	8.0	7.8	7.3	6.4	5.8	3.2	3.1	1.9	2.1	2.0	1.9
F4 – F2	10.5	9.9	9.7	9.2	8.0	7.8	5.3	5.2	3.7	3.9	3.4	3.4

Table 8. The acoustic-auditory data in Barks (Z) of the isolated vowels (male pronunciation) used for the phonological classification of the Latvian vowels. The transform into Barks was performed using formula suggested by H. Traunmüller.

Since there are no central vowels in Latvian [–grave] vowels can be only front vowels characterized by F2–F1>3.5 Z and F3–F2<3.5 Z (Table 8). Not a single vowel of the group has the lip rounding opposition (they are unrounded); therefore the feature [±flat] cannot be used for their separation. The characteristic feature of front vowels is spectral spread, but the feature [spread] has four degrees and cannot be used in binary oppositions. Instead G. Fant suggests the use of two binary features – [±diffuse] and [±sharp][8]. G. Fant describes [+diffuse] vowels as “less open front vowels” (Fant 1983: 16). The vowel [æ(:)] can be separated from the other front vowels by the feature [±diffuse], because it is pronounced with a wide jaw opening and a low tongue position, that results in high value of F1 and a lot of energy between spectral peaks and is therefore described as [+compact] or [–diffuse]. Following Fant’s suggestion the diffuseness is used to separate the vowels [i(:)] and [e(:)] as [+diffuse] from [æ(:)]. In this investigation the vowels are [+diffuse], if F2′–F1>7 Z (Table 8).

---

[8] It seems that G. Fant has mechanically divided four-degree feature [spread] into two binary features, since all these features are based on the value F2′–F1.

Parameter	Feature	u(:)	o(:)	a(:)	æ(:)	e(:)	i(:)
–(F2–F1)	*[±grave]*	+	+	+	–	–	–
–(F2′+F1)	*[±flat]*	*(+)*	+	–	–	–	–
–(F2′+F1)	*[±extra flat]*	+	–				
F2′– F1	*[±diffuse]*				–	+	*(+)*
F2′– F1	*[±sharp]*					–	+

Table 9. The distinctive features used for the phonological classification of the Latvian vowels and the corresponding acoustic-auditory parameters suggested by G. Fant (Fant 1983).

To separate the vowel [i(:)] from [e(:)] G. Fant used the feature [±sharp] explaining that this feature involves palatalization, i. e. a more advanced articulation. An advancement of articulation in palatal region results in frequency increase of F2′ and decrease of F1 thus increasing the spectral spread and the numerical value of F2′–F1. For the Latvian vowels the boundary of [+sharp] was estimated as F2′–F1>10.5 Z. Using this value, the vowel [i(:)] has been classified as [+sharp], but the vowel [e(:)] – as [–sharp] (Table 8). The results of this classification are summarized in Table 9. The same classification can be graphically represented as a distinctive feature branching tree for the Latvian vowels (Figure 11).

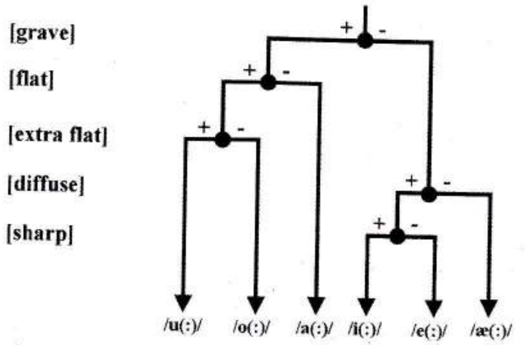

Figure 11. On acoustic-auditory parameters based distinctive feature branching tree for the Latvian vowels.

To test if the classification using features suggested by G. Fant is universal the author of the present paper has attempted to classify according to it the whole IPA vowel system. Since it was necessary to make a distinction between front, central and back vowels in IPA system, the feature [±acute] associated with frontness was used additionally to the feature [±grave]. In such a way back vowels are characterized as [+grave] and [–acute], front vowels – as [–grave] and [+acute], but central vowels – as [–grave] and [–acute]. To make a distinction between four degrees of vowel openness, the feature [±compact] was used along with features [±diffuse] and [±sharp]. The feature [±compact] is associated with a large articulatory opening. Using this feature it is possible to separate open vowels, which are [+compact] and [–diffuse] (automatically [–sharp]), from open-mid vowels, which are [–compact] and [–diffuse] (and [–sharp]).

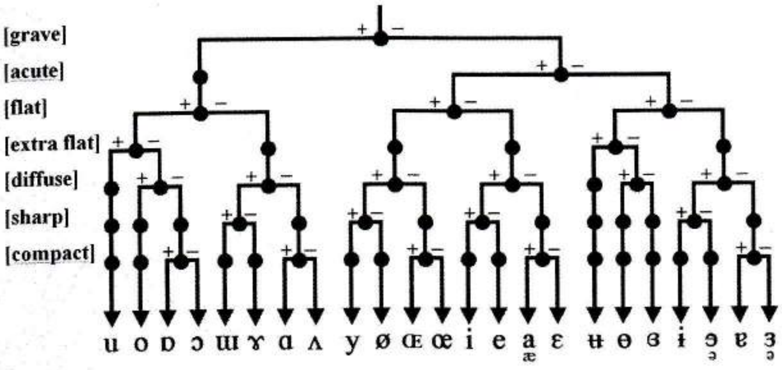

Figure 12. A version of distinctive feature branching tree for IPA vowel system.

According to this system close-mid vowels are characterized as [–compact], [+diffuse] and [–sharp]. To separate close vowels from close-mid vowels using the features suggested by G. Fant, the only solution was to use [+sharp] for all close vowels, despite the fact that this feature is related only to an advancement of front vowels. The graphical result of such a classification is shown in Figure 12. Even though such distinctive feature branching tree (Figure 12) includes nearly all IPA vowels, the classification has several shortcomings. First of all, the feature [±sharp] can be assigned only to front vowels because it is attributed to narrow palatal opening, which acoustically manifests as a high F2′–F1 value. Close central and back vowels do not have such values and therefore should be characterized as [–sharp]. Thus the use of [+sharp] for close central and back vowels is not accurate, however, it was introduced because of the need to separate close vowels from close-mid vowels using available distinctive features. The author of the present research would suggest substituting the inappropriate feature [±sharp] with the feature [±extra diffuse] whose acoustic correlate is a very low value of F1–F0 and therefore it could be related to all close vowels. Secondly, the use of the feature [±extra flat] is not justified, because there is no threefold opposition made possible by lip rounding for vowels of the same tongue position in the IPA vowel system. There is only a binary opposition *rounded* vs. *unrounded* that can be sufficiently marked by the feature [±flat]. Abandoning the feature [±extra flat] and replacing the feature [±sharp] with [±extra diffuse] would allow classifying most of the vowels of IPA system correctly. Other classification problems are related with the intermediate stages of openness: between close and close-mid (e. g., /ɪ, ʏ, ʊ/), close-mid and open-mid (e. g., /ə/), open-mid and open (e. g., /æ/ and /ɐ/, too). English phoneticians have used the feature [±tense] related to the duration and formant values of vowels to separate the vowel /i/ from /ɪ/ and the vowel /u/ from /ʊ/. It is possible that this feature can be used also to separate the vowel /y/ from /ʏ/, the vowel /a/ from /æ/ and the vowel /ɘ/ from /ə/ (if needed – also /ɐ/ form /ɜ/). This problem needs to be addressed in a separate investigation. Since the author of this article did not have access either to audio data of IPA vowels pronounced by several informants or to the valid data of the acoustic measurements, it was impossible to develop the robust, reliable scheme for the phonological classification based on acoustic facts.

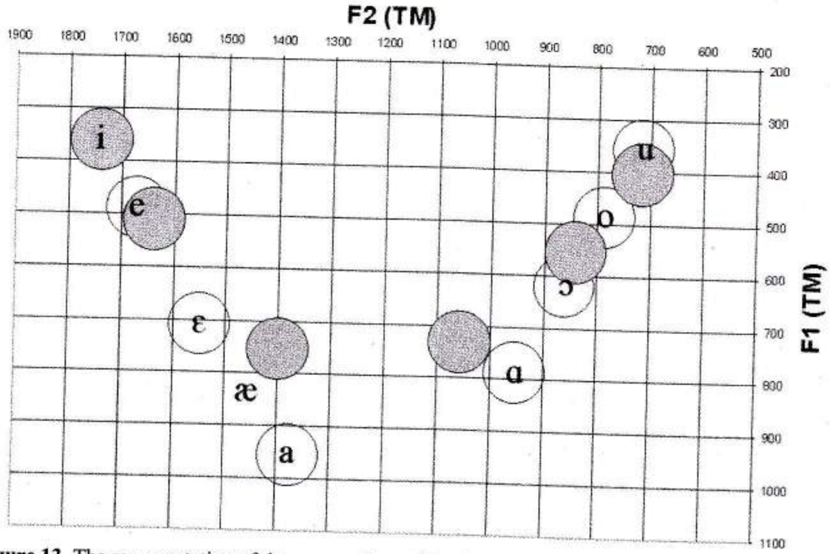

Figure 13. The representation of the mean values of the Cardinal vowels (the large white filled dots) and the Latvian vowels (the large grey filled dots) in the psychophysical F2/F1 plane. The units of measure are Technical Mel. The zone of the Latvian vowels is enlarged in proportion and displaced so that the points of vowels [i] and [u] coincide as much as possible with the corresponding Cardinal vowels

To choose the IPA transcription symbols for the Latvian vowels without such reliable classification, the acoustic data of Latvian vowels were compared with the data of Cardinal vowels, because the system of Cardinal vowels forms a base for the IPA vowel system. In the present research the placement of Latvian vowels in the acoustic (F2–F1)/F1 plane and the psychoacoustic F2′/F1 or F2/F1 plane was compared with the placement of Cardinal vowels as represented in the phonetic literature (Catford 1988; Jassem 1973; Ladefoged 1975). It was noted that the data of Cardinal vowels taken from different sources made a noteworthy dispersion in the acoustic and psychoacoustic plane.

A similar observation was made by the prominent American phonetician P. Ladefoged when he analyzed the Cardinal vowels pronounced by the creator of the system D. Jones and by 10 of his disciples (Ladefoged 1975). The recording of the material was supervised by D. Jones himself, and he also checked the recorded material for compliance to the auditory quality of the Cardinal vowels. P. Ladefoged (Ladefoged 1975: 97) concluded:

> […] the exact phonetic quality of a vowel sound does not depend on the absolute values of its formant frequencies, but on the relationship between the formant frequencies for that vowel and the formant frequencies of other vowels produced by that speaker.

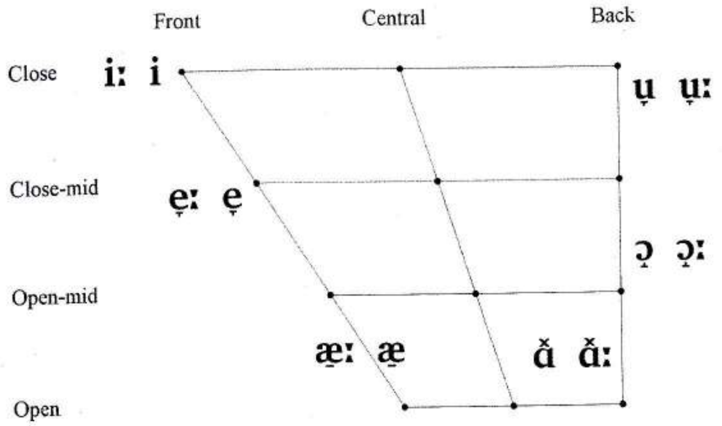

Figure 14. The representation of the Latvian vowel system in the IPA vowel quadrilateral, using IPA symbols corresponding to the acoustic-auditory quality of the Latvian vowels.

Ladefoged's conclusions are taken into account making comparison of the Latvian vowel system with the system of Cardinal vowels (Figure 13). In creating Figure 13 several important factors were taken into account. Firstly, since the goal of the investigation was to compare the pronunciation data of speakers representing certain systems (the vowel system of Standard Latvian and the system of Cardinal vowels), the mean values were used. The Latvian data reflect the male pronunciation of isolated vowels (Table 1), and the Cardinal vowel data reflect the pronunciation by 11 informants (Ladefoged 1975: 88–89). Secondly, to compare placement of vowels, the zone of the Latvian vowels was enlarged in proportion and displaced so that the point of the vowel [i] coincided with the corresponding Cardinal and the point of [u] overlapped the corresponding Cardinal as much as possible (Figure 13). The appropriate IPA symbols and the diacritic marks specifying the deviation in pronunciation of the Latvian vowels from the pronunciation of the Cardinal vowels (Figure 14) were chosen after the thorough analysis of Figure 13.

Since the data point of the Latvian back mid vowel almost equally overlaps with the points of the Cardinal vowels [o] and [ɔ] it is possible to choose between the two symbols for its notation. It has to be taken into account that the auditory quality of the Latvian vowel, judging by hearing, is closer to the quality of the Cardinal vowel /ɔ/ therefore the author of this article would suggest to transcribe it using the symbols ɔ̝ and ɔ̝ː. The author is aware that more extensive research on the articulation, acoustics and perception of Latvian vowels can provide more detailed information, which will affect the current placement of Latvian vowels in the IPA vowel quadrilateral and influence the phonological classification of these vowels and the choice of symbols for their transcription.

5 Conclusions

(a) The hypothesis of the qualitative equivalence of the long and short Latvian vowels was tested using the data of the acoustic vowel targets. It was discovered that the difference in the acoustic characteristics of the long and short vowels is not significant, since their data points were overlapping in both the acoustic and psychophysical F2/F1 plane, and the differences in the placement of the long and short vowel zones did not exceed 1 Bark.

(b) Different normalization methods of the articulation data were evaluated. The correlation analysis was performed to determine whether there was a statistically significant difference between results obtained by the methods of normalization used. The analysis showed no statistically significant difference between the uniform and non-uniform normalization ($z = 0.00 - 0.236$, $p > 0.05$). It can be concluded from the results that for technical needs the uniform data normalization is advisable, because it is simple but sufficiently effective for the elimination of vowel data differences caused by gender and age of speakers. The best results for the normalization of the data of the Latvian acoustic vowel targets were achieved using coefficient $k = 27\%$, which was calculated from the ratio of the F2 centre frequencies of vowel [i(:)].

- For the non-uniform normalization of the Latvian vowel data it is advisable to use coefficients calculated from the data of vowels pronounced in equal or eventually close phonetic environment.
- To acquire as compact vowel zones as possible, before the normalization of differences determined by the gender, the data of speakers of the same gender should be normalized in respect to the data of a male or female prototype.

(c) From the theoretical literature and the experiments carried out for this research a conclusion is drawn that it is advisable to transform data from the acoustic units (Hz) to psychophysical units (Z) and to characterize the vowels by the tonotopic distances between the formants. In this way it is possible to relate the acoustic parameters of vowels with their perception, simultaneously almost completely eliminating vowel pronunciation data differences determined by the age and gender of speakers.

- Characterizing vowels by the tonotopic distances between their formants, 1 Bark interval suggested by A. Iivonen can be used determining qualitative resemblance of both members of the long and short vowel pair and data from male and female pronunciation.

(d) As a result of the analysis of the vowel targets in Latvian established by the auditory experiments carried out for the present research it was realized that:

- for back vowels the values of the effective second formant determined in experiments ($F2^i$ and $F2^s$) and calculated by Bladon and Fant's formula ($F2'$) agree with the values of F2 obtained in the acoustic measurements of the pronunciation data;

- for front vowels differences are observed between frequencies of the effective second formant obtained in the experiments and theoretically predicted, that can be explained as follows – since there are no rounded front vowels in Latvian, there is no necessity to increase the perceptual contrast of unrounded front vowels increasing the frequency of their effective second formant;
- since the frequencies of F1 and F2 of two-formant stimuli corresponding to the auditory vowel targets do not differ significantly from the frequencies obtained in the matching experiments, the conclusion can be drawn that as a result of phonological experiments the effect of perceptual hypersphere is not observed.

(e) The results of the auditory experiments point to the necessity of extensive further experiments varying the parameters of auditory stimuli to concretize the quality of auditory vowel targets and the shape of vowel zones in human perception, as well as to clarify, why the psychoacoustic data of the Latvian vowels described in this research differ from the data of other languages.

(f) The acoustic database of the Latvian vowels acquired as a result of the present research, the establishment of the acoustic and psychophysical parameters essential for the perception of vowels, as well as the tested methodology of research is a substantial contribution to the development of the acoustic phonetics of the Latvian language and in providing material for typologic research of world languages.

Literature

Bladon, R. & G. Fant (1978): „A two-formant model and the cardinal vowels". Speech Transmission Laboratory – Quarterly Progress and Status Report 1/1978, 1–8

Bond, Dz. (1994): „A note on the quality of Latvian vowels". Journal of Baltic Studies, XXV/1, 3–14

Carlson, R., B. Granström & G. Fant (1970): „Some studies concerning perception of isolated vowels". Speech Transmission Laboratory – Quarterly Progress and Status Report 2-3/1970, 19–35

Catford, J. (1988): A Practical Introduction to Phonetics. Oxford: Clarendon Press.

Chistovich, L. & V. Ljublinskaya (1979): „The 'centre of gravity' effect in vowel spectra and critical distance between the formants: psychoacoustical study of the perception of vowel-like stimuli". Hearing research 1, 185–195

Chistovich, L., R. Sheikin & V. Ljublinskaya (1979): „'Centres of gravity' and spectral peaks as the determinants of vowel quality". Frontiers of Speech Communication Research (Eds: B. Lindblom and S. Öhman). London: Academic Press.

Eek, A. & E. Meister (1994): „Acoustics and perception of Estonian vowel types". PERILUS 18, 55–90

Fant, G. (1975): „Non-uniform vowel normalization". Speech Transmission Laboratory – Quarterly Progress and Status Report 2-3/1975, 1–19

Fant, G. (1983): „Feature Analysis of Swedish Vowels – a revisit". Speech Transmission Laboratory – Quarterly Progress and Status Report 2-3/1983, 1–18

Grigorjevs, J. (1995): „Latviešu literārās valodas patskaņu akustisks raksturojums“. Baltu filoloģija 5, 81–83

Iivonen, A. (1987): „Kuulon kriittinen kaista mahdollisten vokaalien lukumäärän ja vokaalien psykoakustisten etäisyyksien selittäjänä“. Helsingin yliopiston fonetiikan laitoksen monisteita 12, Summary: „The critical band in the explanation of the number of possible vowels and psychoacoustical vowel distances“. Helsinki.

Jassem, W. (1973): Podstawy fonetyki akustycznej. Warszava: Państwowe Wydawnictwo Naukowe

Johnson, K., M. Fernandez, M. Henniger & J. Sandstrum (1993a): „Spectral integration in vowel perception: matching and discrimination studies“. UCLA Working Papers in Phonetics 83, 47–54

Johnson, K., E. Flemming & R. Wright (1993b): „The hyperspace effect: phonetic targets are hyperarticulated“. UCLA Working Papers in Phonetics 83, 55–73

Ladefoged, P. (1975): Three Areas of Experimental Phonetics (4th edn). London: Oxford University Press.

Liepa, E. (1979): Vokālisma un zilbju kvantitāte latviešu literārajā valodā. Rīga: Zvaigzne.

Liljencrants, J. & B. Lindblom (1972): „Numerical simulation of vowel quality systems: The role of perceptual contrast“. Language Vol. 48, No. 4, 839–862

Маркус, Д. (1983): Вокализм зиемерского говора (экспериментальное исследование). Дис. канд. филол. наук. Вильнюс: ВГУ.

Miller, J. (1989): „Auditory-perceptual interpretation of the vowel“. Journal of the Acoustical Society of America 85, 2114–2134

Mol, H. (1970): Fundamentals of Phonetics, Vol. 2. Mouton, The Hague.

Nordström, P.-E. & B. Lindblom (1975): „A normalization procedure for vowel formant data“. Proceedings of the 8th International Congress of Phonetic Sciences, Leeds, England.

Rosner, B. & J. Pickering (1994): Vowel perception and production. Oxford: Oxford University Press.

Sarkanis, A. (1993): Latviešu valodas Augšzemes sēlisko izloksņu prosodija un vokālisms: Eksperimentālie pētījumi. Filol. dokt. disert. Viļņa: VU.

Stelle, A. (1971): Latviešu literārās valodas uzsvērtā vokālisma akustisks skaidrojums. Filol. zin. kand. dis. Rīga: LVU.

Traunmüller, H. (1981): „Perceptual dimension of openness in vowels“. Journal of the Acoustical Society of America 69, 1465–1475

Traunmüller, H. (1988): „Analytical expressions for the tonotopic sensory scale“. PERILUS 8, 93–102

Traunmüller, H. & F. Lacerda (1987): „Perceptual relativity in identification of two-formant vowels“. Speech Communication 6, 143–157

Riga Juris Grigorjevs

Latvian Language Institute of the University of Latvia, Akadēmijas laukums 1, LV-1050 Riga
E-Mail: juris.grigorjevs@lu.lv

Hinweise für Autorinnen und Autoren

Alle *redaktionellen* Zuschriften und Sendungen erbitten wir an die Redaktion der *Linguistischen Berichte*:

Linguistische Berichte
z. Hd. Herrn Prof. Dr. Günther Grewendorf
Johann Wolfgang Goethe-Universität
Fachbereich Neuere Philologien, Institut für Linguistik
Grüneburgplatz 1
D-60629 Frankfurt am Main
E-Mail: lb@lingua.uni-frankfurt.de

- Wir bitten darum, jedes Manuskript in dreifacher Ausfertigung einzureichen; die Kopien sind für die Begutachtung erforderlich, der jedes Manuskriptangebot unterzogen wird.
- Bevorzugte Sprache der Beiträge ist Deutsch; englisch- und französischsprachige Beiträge werden akzeptiert, sofern sie stilistischen Standards genügen.
- Stellen Sie Ihrem Aufsatz eine kurze Zusammenfassung (abstract) in englischer Sprache voran.
- Namen und Vornamen aller (Ko-)Autor(inn)en sollen auf dem Manuskript vermerkt sein, einschließlich der Adresse (mit Telefonnummer), an die die Korrekturfahnen geschickt werden sollen.
- Die Autor(inn)en erhalten 20 Sonderdrucke ihres Beitrags.

Autor(inn)en, deren Beiträge zur Veröffentlichung vorgesehen sind, erhalten von der Redaktion ein Merkblatt mit *Richtlinien zur Manuskriptformatierung*, deren Berücksichtigung unbedingt erforderlich ist. Der Beitrag ist anschließend in elektronischer Form (E-Mail-Anhang) im Word-Format oder als rtf-Datei an die Redaktion zu senden. Zudem werden Autor(inn)en gebeten, der Redaktion einen inhaltsgleichen Ausdruck des Beitrags per Post zukommen zu lassen.

Alle *Besprechungsexemplare* von neu erschienenen Werken der Sprachwissenschaft und eng verwandter Disziplinen schicken Sie bitte an die Redaktion. Die Auswahl behält sich die Redaktion vor. Rücksendungen unverlangt eingesandter Bücher können nicht vorgenommen werden.

Mitteilungen für die *LB-Info* schicken Sie bitte an Klaus Müllner, Postfach 21 51, 65779 Kelkheim, kmuellne@uni-mainz.de.