TIMBRE – VERBAL DESCRIPTION, SPECTRAL SOURCES AND CONTEXT DEPENDENCE

Jan Štěpánek

Musical Acoustics Research Centre (MARC) Prague
Sound Studio of Music Faculty, Academy of Performing Arts in Prague,
Malostranské. nám. 13,11800 Praha 1, Czech Republic
jan.stepanek@hamu.cz   www.hamu.cz/sound

Abstract: Basic methodological rules of psychoacoustic research of sound timbre are specified. The concept of complex acoustic taxonomy is introduced as methodology for the characterization and structuring of sets of objects (sound sources, sounds and sound percepts), for the labeling of methods of investigation and classification (instrumentally deterministic classification of sources, physically statistical taxonomy of sounds and subjectively statistical taxonomy of percepts). Steps of psychoacoustic experiment and levels of research are described and documented by examples of nonlistening tests focused on verbal description of sound timbre and listening tests with violin and organ pipe sounds. The main result of nonlistening experiments based on the collection of musicians opinions was the finding of three basic independent bipolar dimensions of timbre: 1. gloomy, dark – clear, bright, 2. harsh, rough – delicate, 3. full, wide – narrow. The spectral characteristics mainly influencing timbre perception were searched in experiments with listening tests of violin and organ sound contexts and compared. These characteristics consisted in levels of the first harmonic, in levels of high frequency harmonics and in levels of nonstationary sound components (noise). Frequency positions and mutual relations of these perceptual resources of timbre qualification and differences are dependent on pitch and on type of sound source. The verbal timbre dimensions are compared with verbal descriptions of violin and organ pipe sounds and related to spectral characteristics.

1. Introduction

Timbre is one of the basic attributes of the sound perception. Whereas perception of loudness and pitch have relatively unambiguous physical causality in sound intensity (SPL) and frequency, resources of timbre perception consist in the properties of the whole sound spectrum and its time variation. This physical complexity is the source of difficulties connected with the timbre definition itself (see for example ANSI norm [1]: "Timbre is that attribute of auditory sensation which enables a listener to judge that two sounds, similarly presented and having the same loudness and pitch, are dissimilar"). ANSI definition is the operational one, enabling postulation of conditions for the experimental study of the timbre but not its description using physical quantities [2]). Nevertheless multidimensional nature of timbre was undoubtedly accepted by majority of researchers.

The first attempt for timbre description going from acoustic properties of harmonic sound signal was given by Helmholtz [3]. He postulated two theories, based on relative or absolute relations between spectrum and timbre. Stumpf [4] on the other side used psychoacoustic view on timbre, he used semantic analysis of verbal attributes and without any experiment (in contemporary sense) has obtained three dimensions:

1. *dunkel – hell*
   (dark – bright in English)
2. *stumpf / weich – scharf / rauh*
   (unpointed / soft – sharp / rough)
3. *voll / breit – leer / dünn*
   (full / wide – empty / narrow)

Contemporary experimental approach to the study of timbre is based on psychological experiment [5] and was determined in seventies of the last century (see for example [6, 7]). This approach attempts to search for timbre dimensions, their verbal description and/or spectral interpretation. Skepticism raised from disapproving results of different experiments and difficulty in creation of common timbre theory led to the postulation of timbre research aim by Grey [6]: "A major aim of research in timbre perception is the development of a theory for the salient dimensions or features of classes of sounds". This postulation reflects the fact of deep dependence of timbre perception on the context of listened sounds.

This contribution describes used approaches and reached experimental results in the timbre study in Sound studio of Music Faculty Prague during more then last decade.

2. Methods

The main goal of the contemporary Musical Acoustic Research Centre (MARC) project is the research of timbre perception of musical instruments and its relationship to sound quality. This includes search for common timbre structure (and individual dimensions of timbre), their verbal description and causality in musical signal (sound or more specific tone) or signal source conditions (musical instrument construction and playing).

2.1. Methodical resources

For the description of basic approaches we introduce the concept of **complex acoustic taxonomy** as methodology for the characterization, description and structuring of sets of objects and for the labeling of methods of investigation and classification.

We distinguish three sets of objects: **space of sound sources**, its objects – ***conditions of tone creation*** – are characterized by physical quantities going from material properties, geometry and construction of sound source up to playing technique characteristics. Classification method of this space we call ***instrumentally deterministic classification*** which includes musical instruments taxonomy, classification of playing technique and tone determination (tone, dynamic, etc.).

Space of sounds with ***tones*** as objects and characterized by acoustic quantities. ***Physically statistical taxonomy*** as the method of investigation and classification of tones includes musical sounds classification and data reduction which preserves taxonomy attributes.

Space of sound percepts has ***subjective perception of tones*** as objects, which are characterized by psychoacoustic quantities. The method of ***subjectively statistical taxonomy*** is based on the classification of musical sounds perception and on psychological tests and their statistical evaluation.

In common there are three levels of scientific knowledge [8]:

1) Description, classification, categorization.
2) Correlation, prediction.
3) Causality.

The causality is the highest level of knowledge but at the same time it is the most difficult to reach it. The most of previous timbre research results fall into first two levels.

The main MARC project aim is the description of these three spaces and the search for their causal relations going from properties of musical instrument through musical sound to the perception of this sound.

2.2. Psychoacoustic experiment

2.2.1. Basic steps

Psychoacoustic as a research discipline is based on psychological experiment [5] but has some specific aspects. Following is our postulation of basic steps of psychoacoustic experiment:

1) ***Research aim*** specification.
2) Formulating the ***goal of the experiment*** (questions in the listening test).
3) Sound ***stimuli*** preparation (properties, presentation manner, number).
4) Assembling the group of ***judges*** (requirements on judges, number).
5) Choice of ***test method*** (psychological measurement type).
6) ***Experiment plan*** (technique, time schedule).
7) ***Listening test*** (according to experiment plan).
8) Statistical ***evaluation*** of test results.
9) ***Interpretation*** of results.
10) ***Formulation of hypothesis*** on stimuli (implied from 8 and 9).
11) Completion, selection, or ***adjustment of stimuli*** (according to 10).
12) ***Listening test for hypothesis verification*** (steps 4 – 9).
13) ***Decision about hypothesis*** (accept or reject).

Steps of psychoacoustic experiment number 1) – 9) are used in the most of published timbre studies but they are sufficient only up to correlation or prediction level of conclusions. Steps 10) – 13) were added to reach and prove causality between sounds and their perception.

2.2.2. Used methods

The previous MARC research was focused on **timbre of stationary sounds** and on the verbal description of timbre in common. Mostly used experimental test method in realized studies was dissimilarity pair test [5] and modification of Verbal Attribute Magnitude Estimation (VAME) [2] which we call Verbal Attribute Ranking and Rating (VARR) [9]. For the collection of verbal attributes describing the timbre properties of listened sounds the Spontaneous Verbal Description (SVD) was used [10 – 12].

Dissimilarity test results (set of individual dissimilarity matrices) were evaluated using MDS latent class approach (program CLASCAL) [13, 14]. The program solves two optimization tasks:

1) To create distances of Euclidean space of appropriate low dimension to match the stimuli dissimilarities (***construction of perceptual space*** of stimuli); this Euclidean space constitutes perceptual space.

2) To establish the appropriate number of (latent) classes of test respondents and to assign each individual respondent to one of these classes (***a posteriori grouping*** of respondents). Perceptual space of each individual class has the same dimensions with only different weights (dimension scales).

The results of VARR test (set of variables describing listened sounds) were evaluated using Factor Analysis (FA) [15]. This method simplifies the description of sound context properties (variables) by small number of factors. Space of factor loadings or factor scores we may consider for perceptual space.

External interpretation of any perceptual space can be made using external scales, here verbal attributes, spectral characteristics or musical instrument properties, describing the same objects – test stimuli. These external scales – features of all three complex acoustic taxonomy spaces may led to the correlation or prediction level of knowledge. The method of embedding (optimal fitting of the external scale into perceptual space) [16] was applied [17]. Only successfully embedded external scales (represented by significant reproduction of external scale values by embedding) were selected for further analysis and interpretation. The angles between the embedding vectors were calculated [17, 18] and the relations between the embeddings were observed and discussed. A nearly orthogonal system of external scales completely describing perceptual space was searched. Also the correlation analysis of test results and features of object spaces was used.

The results of MARC research are documented by examples of nonlistening tests focused on verbal description of sound timbre in common and by listening tests with violin and organ pipe sounds. In specific situation of violin sounds additional tests were provided to prove causality between spectral characteristics and selected timbre properties (represented for example by attributes *sharp* and *narrow*) [19, 20].

3. Verbal description

There were made two different sets of experiments according to verbal description of timbre. The first one was based on the collection of musician opinions and experiences (and provided without listening to any sound). These experiments let to the construction of common perceptual space. The second one was provided using standard listening experiments with violin and organ sounds, SVD method was used for the collection of verbal attributes during dissimilarity test. Embedding of collected verbal attributes led to the verbal interpretation of perceptual spaces in both instruments.

3.1. Musicians opinions

The collection of musicians opinions was based on nonlistening experiments: collection of verbal attributes describing timbre – ***questionnaire survey***, and ***test with selected verbal attributes in dissimilarity*** according to timbre individual understanding (internal image of timbre of individual test respondent), to obtain attribute relations [21, 22].

The 25 most frequently used verbal attributes from 120 collected questionnaires (Table 1) were used as stimuli, 43 music professionals took part in the test (five professional groups: players of string, wind and key instruments, composers & conductors and sound designers) [21, 22].

The dissimilarity test results were evaluated using CLASCAL [13, 14] and common perceptual space of verbal attributes was constructed.

The angles between attribute positions in the coordinate system of common perceptual space were calculated and the relations between them were qualified as follows [17]:

a) small angles ($\alpha \leq 20^o$) => similar,

b) nearly orthogonal positions ($70^o \leq \alpha \leq 110^o$) => independent,

c) nearly opposite positions ($\alpha \geq 160^o$) => opposite

meaning according to timbre.

The aim of the interpretation was to search for a system of verbal attributes, which completely describes the perceptual space and constitutes nearly orthogonal system. The **three-dimensional common perceptual space** as the optimal model for dissimilarity data and following nearly orthogonal system of verbal attributes were identified (see also Table 2 and Figure 1):

1. ***temný*** / *tmavý* – ***jasný*** / *světlý* — ***gloomy*** / *dark* – ***clear*** / *bright*
2. ***drsný*** / *hrubý* – ***jemný*** — ***harsh*** / *rough* – ***delicate***
3. ***plný*** / *široký* – ***úzký*** — ***full*** / *wide* – ***narrow***

Table 1. List of 25 the most frequently used verbal attributes with their absolute (f_{abs}) and relative (f_{rel}) frequencies.

Verbal attribute		Frequency	
Czech	English	f_{abs}	f_{rel}
ostrý	sharp	94	78.3
temný	gloomy	79	65.8
měkký	soft	78	65.0
jasný	clear	75	62.5
sametový	velvety	61	50.8
jemný	delicate	58	48.3
kulatý	round	58	48.3
tupý	unpointed	55	45.8
drsný	harsh	54	45.0
světlý	bright	54	45.0
tvrdý	hard	54	45.0
sladký	sweet	53	44.2
plný	full	51	42.5
hrubý	rough	46	38.3
tmavý	dark	46	38.3
teplý	warm	43	35.8
zářivý	radiant	42	35.0
čistý	pure	40	33.3
vřelý	hearty	40	33.3
barevný	colored	38	31.7
zvonivý	ringing	38	31.7
chladný	cool	36	30.0
průzračný	lucid	36	30.0
široký	wide	36	30.0
úzký	narrow	36	30.0

Table 2. Angles between selected verbal attributes representing dimensions of common perceptual space.

Angle [°]	gloomy	clear	harsh	delicate	full	narrow
gloomy	–		84	91	71	109
clear	176	–	97	89	106	75
harsh			–		92	77
delicate			159	–	106	85
full					–	
narrow					169	–

The division of musicians into a posteriori classes does not reflect their professional specialization. On the other side a priori professional groups possess at least the same first two dimensions (as optimal MDS model), three-dimensional solutions are in good agreement with common one [22]. It means found common perceptual space is shared with all musicians.

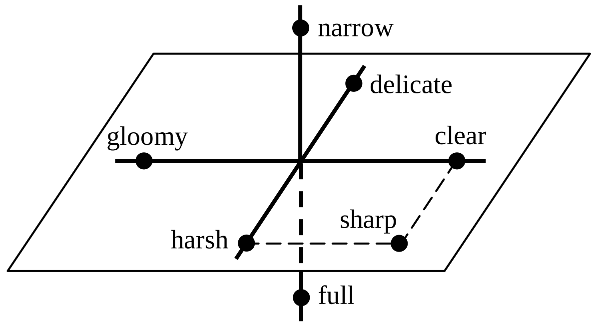

Figure 1. Nearly orthogonal system of verbal attributes in three-dimensional common perceptual space of timbre.

3.2. Violin sounds

Violin sounds of five different pitches were used in experiments: B3 (fundamental frequency 247 Hz, played on G string), F#4 (370 Hz, D), C5 (523 Hz, A), G5 (784 Hz, E), and D6 (1175 Hz, E). The instruments of various quality, going from factory to master, were played using the same technique: *détaché*, *naturale*, non-vibrato, and *mezzoforte*. Recordings were made in anechoic room and uniformly adapted to reduce the influence of transients on the perception (uniform fade in and fade out). The results of dissimilarity tests with twenty judges and seventeen sounds in each pitch led to the construction of perceptual spaces [10], their dimensionality is in Table 3.

These spaces were interpreted using verbal attributes collected in SVD. The attributes were embedded into perceptual spaces, the number of use of verbal attribute in each sound defined external scale values (Table 4, Figure 2 and 3).

Table 3. Number of perceptual space dimensions of optimal MDS solution and number of verbal attributes which have number of use at least 10 and embedding correlation significance at least α.

Tone		B3	F#4	C5	G5	D6
No. of dimensions		3	3	2	2	2
No. of attributes	$f_{abs} \geq 10$	65	64	58	64	65
	$\alpha \leq 5\%$	45	41	32	31	49
	$\alpha \leq 1\%$	27	28	22	12	31
	$\alpha \leq 0.1\%$	13	13	12	4	15

Table 4. The best embedded verbal attributes with number of use at least 30 and embedding significance $\alpha \leq 0.1$ %, sorted according to embedding correlation R_{embed}.

Tone B3				
Verbal attribute		no. of use	rank	R_{embed}
Czech	English			
temný	gloomy	144	3	0.97
hladký	smooth	45	17.5	0.95
úzký	narrow	105	6	0.94
mečivý	bleaty	87	9	0.93
tmavý	dark	208	1	0.92
kovový	metallic	99	7	0.90
ostrý	sharp	176	2	0.87
jemný	delicate	45	17.5	0.87

Tone F#4				
Verbal attribute		no. of use	rank	R_{embed}
Czech	English			
kulatý	round	73	8	0.98
tmavý	dark	157	2	0.96
sametový	velvety	48	14	0.95
kovový	metallic	72	9.5	0.95
pronikavý	penetrating	41	20	0.90
měkký	soft	96	4	0.89
temný	gloomy	78	6	0.88
ostrý	sharp	185	1	0.86
šustivý	rustle	51	13	0.85
úzký	narrow	107	3	0.85

Tone C5				
Verbal attribute		no. of use	rank	R_{embed}
Czech	English			
měkký	soft	121	6	0.98
jemný	delicate	108	7	0.97
kovový	metallic	76	9	0.96
ostrý	sharp	208	1	0.96
světlý	bright	143	3	0.94
kulatý	round	123	5	0.93
vysoký	high	45	16.5	0.92
tmavý	dark	152	2	0.92
pronikavý	penetrating	35	21	0.91

Tone G5				
Verbal attribute		no. of use	rank	R_{embed}
Czech	English			
tenký	thin	33	24	0.95
měkký	soft	75	7.5	0.93
úzký	narrow	151	2	0.87

Tone D6				
Verbal attribute		no. of use	rank	R_{embed}
Czech	English			
měkký	soft	134	2	0.98
ostrý	sharp	249	1	0.95
kulatý	round	92	7	0.94
šustí	rustle	95	6	0.92
hladký	smooth	59	12	0.92
tmavý	dark	60	11	0.90
pronikavý	penetrating	57	14	0.90
temný	gloomy	71	8	0.88
přidušený	damped	40	18	0.87
jemný	delicate	62	10	0.87
pískový	sandy	53	15	0.87
vysoký	high	38	19	0.86

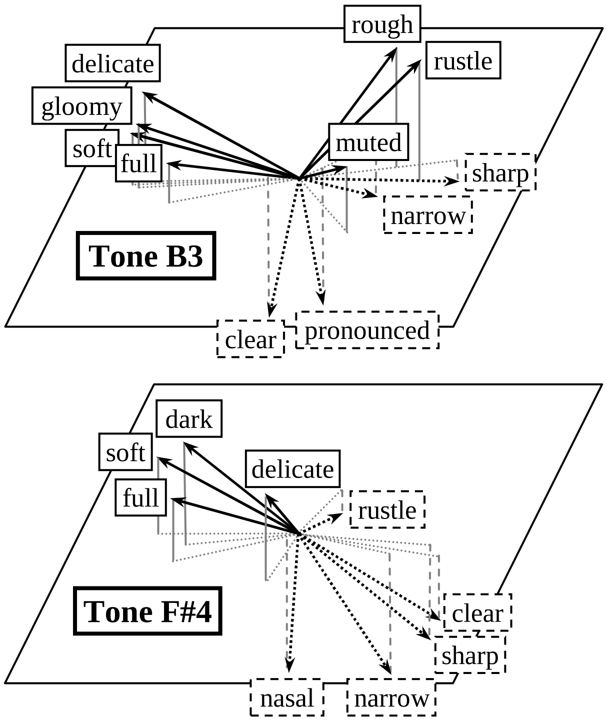

Figure 2. Embeddings of selected verbal attributes into three-dimensional perceptual spaces (tones B3, F#4) with R_{embed} at least 5%. Embeddings are indicated by arrows, direction of arrow agrees with growing of embedded attribute number of use. The appropriate rotation of perceptual spaces was done to reach bigger mutual similarity.

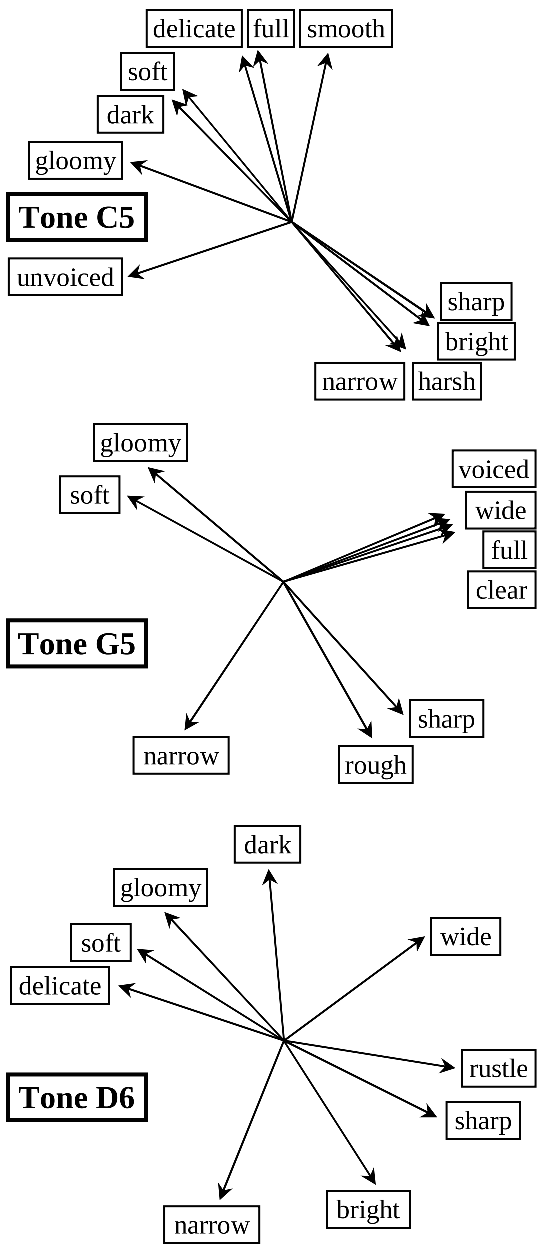

Figure 3. Embeddings of selected verbal attributes into two-dimensional perceptual spaces (tones C5, G5, D6) with R_{embed} at least 5%. Embeddings are indicated by arrows, direction of arrow agrees with growing of embedded attribute number of use. The appropriate rotation of perceptual spaces was done to reach bigger mutual similarity.

Comparison of results led to the assessment of four representative attributes for the description of violin timbre: *sharp*, *clear*, *dark*, and *narrow* [23]. Another experiment with eleven judges and eleven representative sounds in each pitch used VARR method to quantify these four attributes as well as perceived sound quality. FA model (principal components, varimax rotation) led to the two-dimensional perceptual space of verbal attributes and quality in all five pitches (Figure 4).

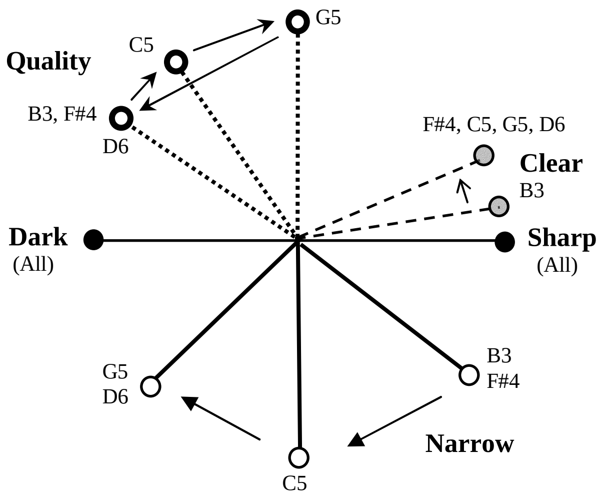

Figure 4. Violin timbre: schematic factor space demonstrating relations among verbal attributes and also perceived sound quality and their changes with pitch.

The extremely often use of the attribute *rustle* in the highest violin tone D6 was the reason for it further investigation [24, 25], the description of it spectral causality see in the paragraph 4.1.

Subjective evaluation using SVD method of one violin sounds recorded in different directions of radiation led to the identification of two specific verbal attributes: *buzzing* and *glossy* [26]. Their spectral causality is described in the paragraph 4.1.

3.3. Organ sounds

The sound recordings of twelve pipe organs from four European countries (six Baroque, two Romantic and four contemporary) measured in situ (list of instruments is in [27]) were subjectively evaluated. Individual sound signals were manipulated to decrease the influence of transients on the perception (uniform fade in and fade out). Two listening tests with sounds of tones C2 (fundamental frequency 65 Hz), C3, C4, C5 and C6 of Principal 8' were provided. During the first test (dissimilarity in pairs) with twelve judges (nine organ players, three organ builders) also collection of verbal attributes (SVD) was realized [28].

The three-dimensional perceptual spaces (constructed from dissimilarities) were appropriate for interpretation in all C-tones. The verbal attributes going from SVD were embedded into perceptual spaces, the number of use of verbal attribute in each sound defined external scale values. Selected verbal attributes use and embedding success are in Table 5. Schematic perceptual space with embeddings of selected verbal attributes is in Figure 5.

Table 5. Selected verbal attributes describing Principal 8' sound timbre (embedding significance $\alpha \leq 5$ %, values of R_{embed} with $\alpha \leq 0.1$ % are bold).

Verbal attribute	German / English	eng / narrow	rauschig / noisy	rund / round	scharf / sharp	weich / soft
C2	no. of use	15	6	10	6	15
	rank	1.5	12.5	5	12.5	1.5
	R_{embed}	**0.86**	–	**0.85**	0.61	0.78
C3	no. of use	9	4	6	4	9
	rank	3.5	14	6.5	14	3.5
	R_{embed}	0.62	0.63	0.76	0.67	**0.94**
C4	no. of use	18	12	10	5	8
	rank	2	4.5	6.5	14	8.5
	R_{embed}	**0.95**	0.64	0.81	0.66	0.70
C5	no. of use	19	6	7	19	19
	rank	3.5	12.5	10	3.5	3.5
	R_{embed}	0.65	0.66	0.58	**0.86**	0.72
C6	no. of use	9	1	4	38	12
	rank	5.5	58.5	14	1	4
	R_{embed}	0.81	–	0.71	**0.98**	0.58

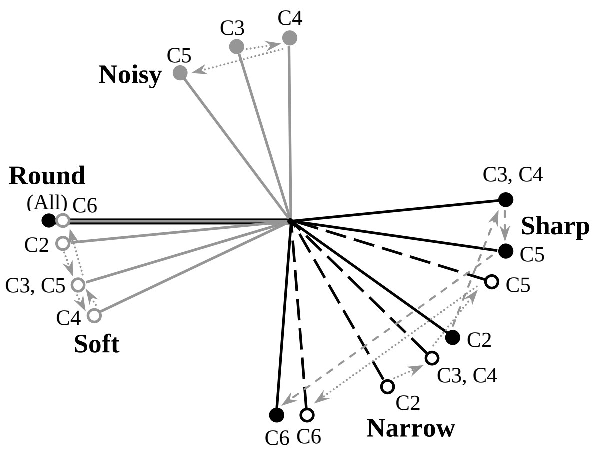

Figure 5. Organ Principal 8' timbre: schematic perceptual space demonstrating relations among selected verbal attributes and their changes with pitch. Attributes positions are relative to the fixed position of attribute *round*.

In the second listening test only the most contrasting pairs in timbre were selected and judged by eight organ builders. The listeners ranked spontaneously expressed verbal attributes into four prescribed categories according to their meaning describing pipe scaling, pipe voicing, prominent or insufficient partials, and timbre. The verbal attributes collected in this test are in German and are summarized in Table 6.

Table 6. The number of use of verbal attributes from contrasting pairs test in prescribed categories: **S**caling, **V**oicing, **P**artials, **T**imbre. The most frequently used attributes and attributes with ambiguous rating were selected [28].

verbal attribute	all	S	V	P	T
Mensur eng	71	71			
obertönig	42			42	
rauschig	35		9	9	17
hell	33				33
laut	32		2		30
Lautstärke	27				27
Oktave	27			27	
streichend	27		9		18
Mensur weit	26	26			
stark	26		4		22
weich	26		2		24
unruhig	25		4	12	9
quintig	23			21	2
rund	23				23
Aufschnitt niedrig	18		17		1
bläst	18			15	3
flötig	15	2			13
Mensur normal	14	14			
dunkel	11			1	10
gut	11		8	2	1
ö	11				11
Aufschnitt hoch	10		10		
Mensur mittel	10	10			
ä	9			1	8
grundtönig	9			9	
prinzipalig	9	2			7
singend	7			1	6
u	7				7
Winddruck höher	7		7		
gerausch	6		4	1	1
Kernspalte weit	6		6		
kräftig	6				6
wenig Teiltöne	6			6	
zu viel Wind	6		6		
blasend	5			1	4
dumpf	5				5
mittle Teiltönen	5			5	
nebengerausch	5		4	1	
o	5				5
klar	4				4
schwacher	4				4
ü	4			1	3
viele Teiltöne	4			4	
direkte Ansprache	3		3		
instabil	3		1	1	1
scharf	2				2
Terz	2			2	

3.4. Discussion

Three-dimensional common perceptual space originated from the **musicians opinions** is similar to the Stumpf conclusions made without any psychoacoustic experiment [4]. More precisely the dimensions one and three are nearly identical, only second dimension is expressed with slightly different attributes but with similar semantic meaning. Only attribute *sharp* in common perceptual space is not dimensional one but lies in the plane of the first two dimensions.

The situation in **violin** (Figure 2 and 3) demonstrated the complexity of timbre perception and its dependence on sound context. The dimensional attributes from common perceptual space are not always opposite (even in second dimension is sometimes one of the attributes not present). In the contrary attributes *sharp* – *dark*/*gloomy* are nearly opposite in all five studied pitches. The position of attribute *narrow* moves from the proximity of *sharp* towards *dark*/*gloomy* (see also [12]). Opposition of *sharp* and *dark* as well as the movement of *narrow* was corroborated in VARR experiment results (Figure 4) [9].

In **organ** sound (Figure 5) the attributes *round* and *soft* are close each other, *sharp* and *narrow* show similar movement according to *round*. The *noisy* attribute is nearly independent (orthogonal) to *round*/*soft* and constitutes thus new independent dimension.

The different use of verbal attributes could have to be influenced also by the fact that violin experiments were made with Czech speaking musicians and organ ones with German speaking organists / organ builders.

The results of listening tests of violin and organ sounds revealed that dimensional attributes from common perceptual space belong to important and very often spontaneously used also by listening to specific sounds, but their mutual relations are strictly dependent on the sound context (type of musical instrument and pitch).

4. Spectral sources

The time-average power spectra were calculated from quasi-stationary parts of time-flow of listened violin and organ sounds, then following spectral characteristics were calculated:

1) Levels of individual harmonics L_{Hi} (harmonic spectra).
2) Levels in critical bands L_{Bi} (Bark spectra) or in thirds of octaves L_{Ti} (third-octave spectra). Appropriate frequency band limits are for example in [29], numbering of critical bands is standard, numbering of third-octaves starts from $f_{T1} \in \langle 18, 22.4 \rangle$ Hz.
3) The spectral centre of gravity f_{cg} (spectral centroid) as a characterization of spectral energy distribution, defined by formula:

$$f_{cg} = f_1 \frac{\sum_{k=1}^{N} kA_k}{\sum_{k=1}^{N} A_k} \quad [\text{Hz}] \qquad (1)$$

where N is the number of harmonics, f_1 is the fundamental frequency, and A_k is the amplitude of k-th harmonic.

Two main approaches to the discovering of acoustical correlates of the most important perceptual features are described in this paragraph. The first one uses embedding of spectral characteristics into perceptual space. The second one uses selected verbal attributes and tries to found their explanation using spectral characteristics due to correlation analysis.

4.1. Violin sounds

The spectral characteristics described above were calculated for the stimuli signals of violins used in the experiments described in the paragraph 3.2 and their values were embedded into appropriate perceptual space. The most successfully embedded spectral characteristics in individual tones are in Table 7. The embeddings into perceptual spaces are in Figure 6 and 7.

Verbal attributes *sharp* and *narrow*, rated in VARR experiment were evaluated by spectral characteristics using correlation analysis. The causal influence of correlating spectral characteristics was verified by additional listening to appropriately manipulated signals [20, 30, 31]. The most significant correlations are in Figure 8.

Buzzing and *glossy* spectral sources were found using correlation analysis, their causality in groups of high harmonic components was verified by additional listening to manipulated signals [26]. The necessity of an existence of the group of neighboring harmonics in the band 5.3 – 7.0 kHz for the perception of *buzzing* (Figure 9) and in the band 7.3 – 8.5 kHz for the perception of *glossy* was proved.

Table 7. Spectral characteristics of stimuli the most successfully embedded into perceptual spaces of individual violin tones.

Acoustic characteristic	Embedding significance level [%]				
	B3	F#4	C5	G5	D6
L_{H1}	0.1	0.1	0.1	0.1	0.1
L_{H2}	–	0.1	5	1	–
L_{H4}	1	–	–	–	1
L_{H5}	1	–	–	–	–
L_{B4}	–	–	–	–	0.1
L_{B6}	–	–	–	–	1
L_{B17}	0.1	–	–	–	–
L_{B18}	–	–	1	–	–
L_{B22}	0.1	1	5	5	0.1
L_{B23}	0.1	1	0.1	1	0.1
L_{B24}	1	0.1	5	1	1
f_{cg}	0.1	0.1	0.1	1	0.1

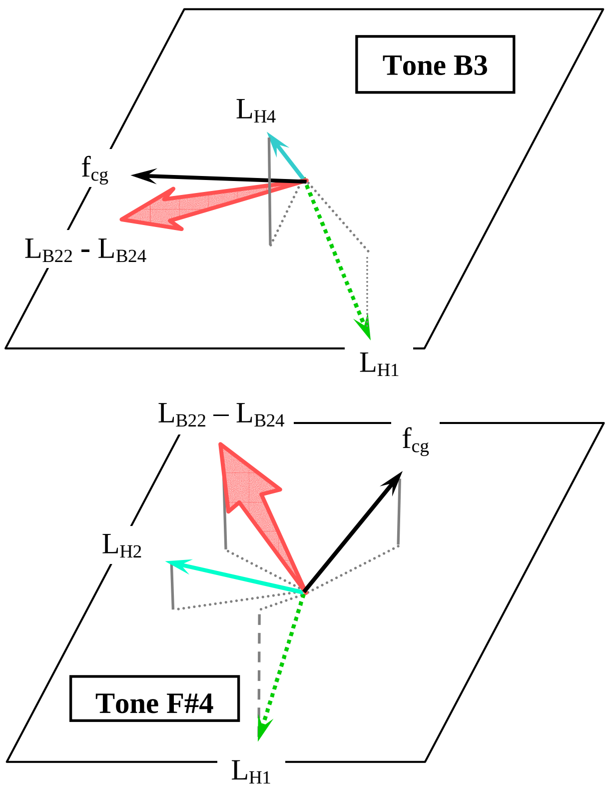

Figure 6. The most successfully embedded spectral characteristics into perceptual space, embeddings are indicated by arrows, direction of arrow agrees with growing of embedded characteristic. Drawings for tones B3 and F#4 (three-dimensional perceptual space solutions). The appropriate rotation of perceptual spaces was done to reach bigger mutual similarity.

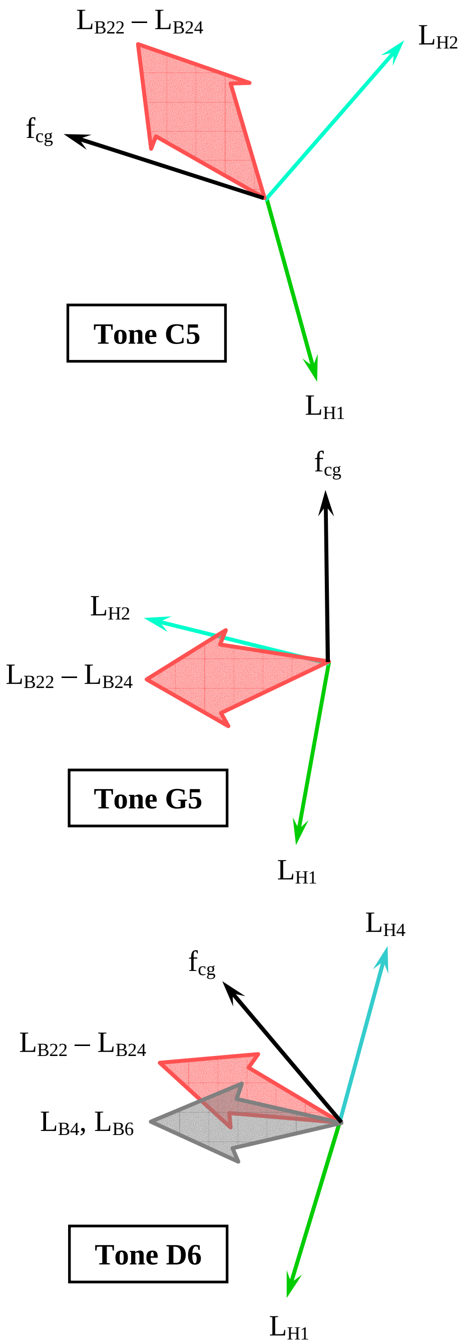

Figure 7. The most successfully embedded spectral characteristics into perceptual space, embeddings are indicated by arrows, direction of arrow agrees with growing of embedded characteristic. Drawings for tones C5, G5 and D6 (two-dimensional perceptual space solutions). The appropriate rotation of perceptual spaces was done to reach bigger mutual similarity.

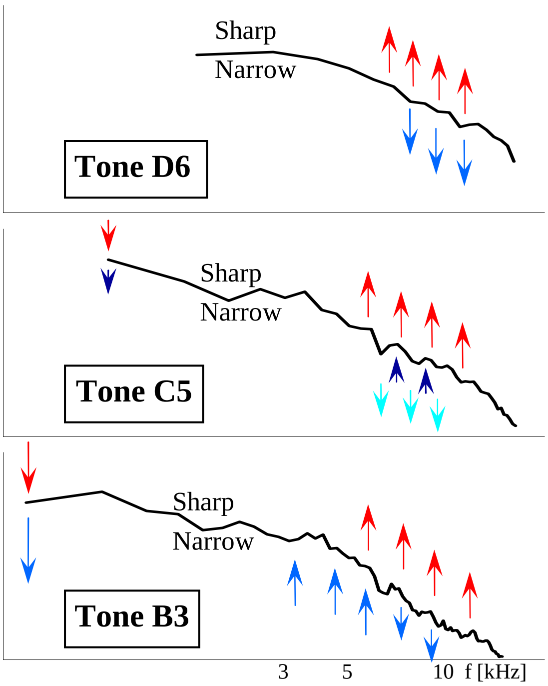

Figure 8. Increase of the attributes *sharp* and *narrow* perception caused by spectrum level change in arrow direction (length corresponding to the size of increase). There are two different strategies of judges for tone C5 and *narrow* assessment (the firs one like in tone B3, the second one like in tone D6).

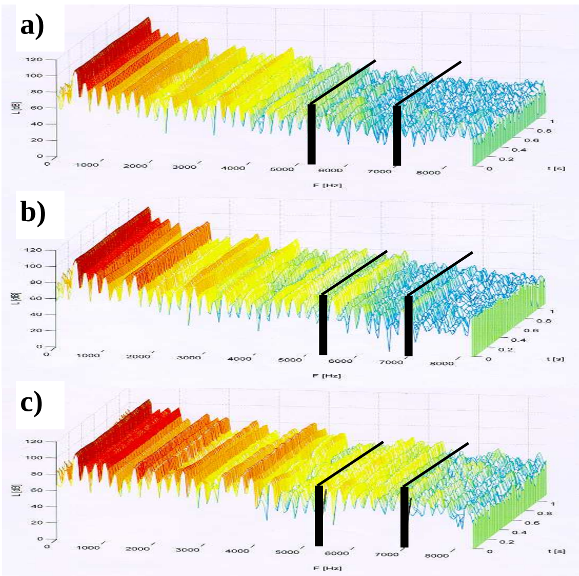

Figure 9. The spectrum waterfalls of radiation of one violin in different directions. Perception of *buzzing* feature is only in the sound corresponding to (c) (existence of the group of harmonics in frequency band 5.3 – 7.0 kHz) but not in (a) and (b) (no or only one harmonic in this band).

The numerous use of verbal attribute *rustle* together with attributes *sandy*, *hissy*, *scrubs*, *dusty* and *horse hair* in the SVD of the highest studied violin tone D6 was the reason for focus on this phenomenon [24, 25]. The correlation analysis and realization of additional experiments proved that the causal source for perception of attributes from the group *rustle* are time-varying levels (above 25 dB) in frequency bands 250 – 300 Hz (frequency position of Helmholtz resonance mode of violin corpus A0), 400 – 450 Hz (top plate mode T1) and 500 – 600 Hz (corpus mode C3), illustration see in Figure 10.

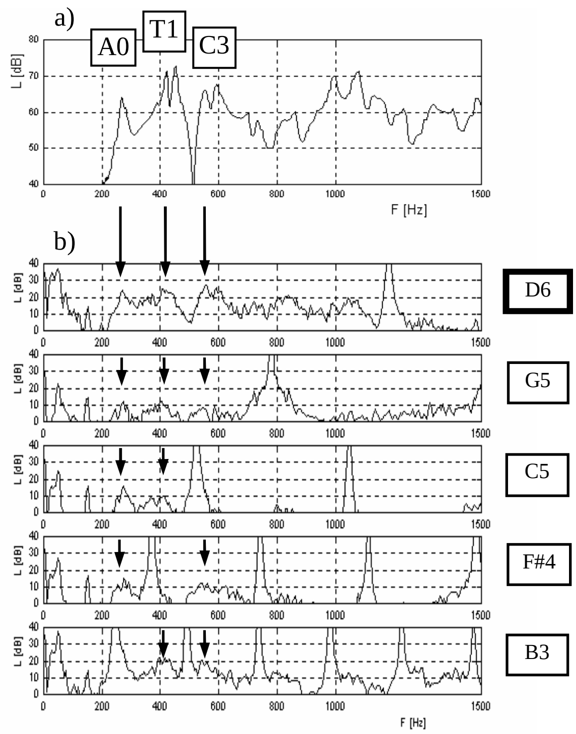

Figure 10. Comparison of the violin frequency characteristic (a) with time-averaged spectra of quasi-stationary tone parts (b) for the violin with high degree of *rustle* in tone D6. Spectral signatures of violin modes A1, T1 and C3 in spectra are marked with arrows.

4.2. Organ sounds

Successfully embedded spectral characteristics (levels in third of octaves) influencing the perception of organ C-tones of Principal 8' are marked in Figure 11 (embeddings into three-dimensional perceptual spaces) and Figure 12 (positions of the most important spectral components relative to the mean third-octave spectrum). Description and discussion of spectral characteristics see also in [27, 32, 33].

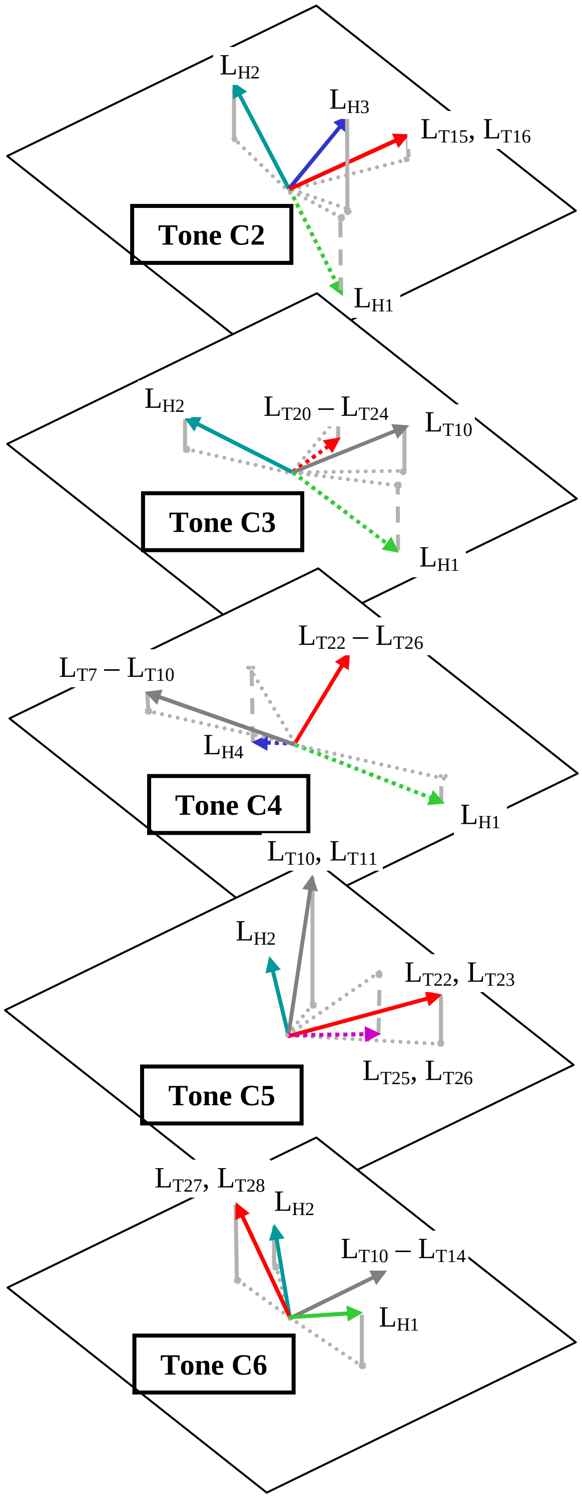

Figure 11. The most successfully embedded spectral characteristics into perceptual space of organ principal 8' sounds, embeddings are indicated by arrows, direction of arrow agrees with growing of embedded characteristic (three-dimensional perceptual space solutions). The third-octaves up to T_{14} represent noise components, from T_{15} higher harmonics.

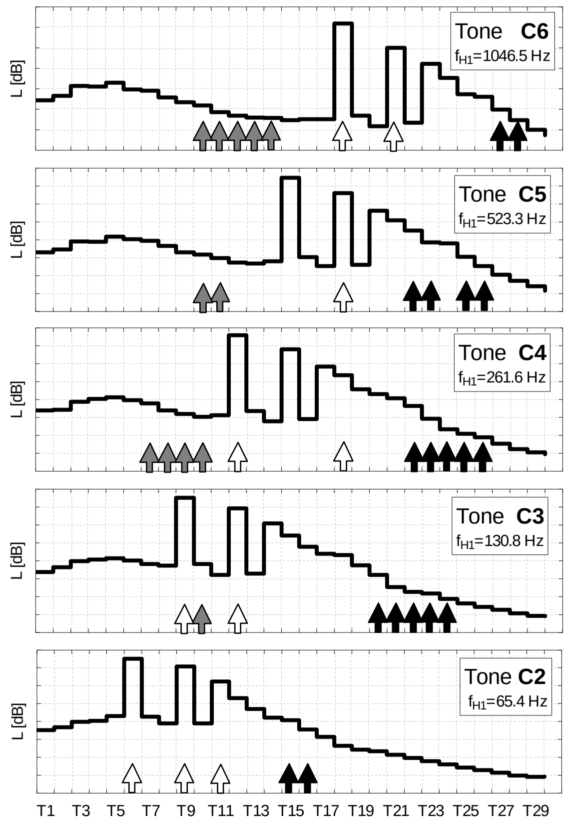

Figure 12. Frequency positions of the spectral characteristics mainly influencing perception in Principal 8' tones C2, C3, C4, C5 and C6. Drawings of mean third octave spectra over the tested set of sounds were used. The arrows have the following meaning: H1(H2|H3|H4), higher harmonics, noise.

4.3. Discussion

The substantial influence on the perception of **violin sound** in all studied pitches have levels of fundamental, levels of high frequency components in critical bands $B_{22} – B_{24}$ (frequency band 9 – 15.5 kHz) and the distribution of spectral energy (value of the spectral centroid f_{cg}), see Figures 6, 7.

The relations of the perceptual influence of individual spectral features revealed high degree of variability according to pitch. For example independent influence have fundamental and high frequency components in tones B3 and D6, high frequency components and spectral energy distribution in tones F#4 and G5. The influence of the spectral energy distribution is not fully defined by (not identical with) high frequency components. A nearly orthogonal system of spectral characteristics (its definition see in the paragraph 3.1) was found only in tone B3 (fundamental, fourth harmonic, high frequency components).

Spectral sources of specific verbal attributes describing violin timbre (*sharp*, *narrow*, *buzzing*, *glossy* and *rustle*) showed finer structure of human perception and causal influence of specific parts of sound spectrum.

There are three main spectral characteristics, which substantially influence the perception of **organ Principal 8' sound**: level of the fundamental (second harmonics in tone C5), levels of high frequency components, levels of noise (non-stationary) components. The second, third and fourth harmonic levels influence perception only in some pitches.

The frequency position of noise components is relatively stable or slightly increases with pitch and lies in frequency band about 70 – 450 Hz with the exception of tone C2, where the noise is probably masked by lower harmonics. The frequency position of high frequency components is increasing with an increasing pitch and changes from the band 450 – 710 Hz (tone C2) to 7100 – 11200 Hz (C6) (Figure 12).

The perceptual influence of fundamental (second harmonic in C5) is independent to the influence of high frequency components (angles varying from 82^o to 108^o). Also noise components (tones C3 – C6) influence is independent to high frequency components (angles varying from 73^o to 108^o). But nearly orthogonal system of spectral characteristics was found only in tones C5 and C6, in other tones was slightly violated. The relations of the influence of fundamental to second harmonic or to noise are changing with pitch.

Much higher stability in spectral characteristics influence on perception was found in organ Principal 8' than in violins. This documents the precisely defined role of Principal stop in organ sound on one side and extremely high variability of timbres in violins on the other side. Moreover the organ sound possess more noises than violin sound.

5. General discussion

The importance of dimensional verbal attributes from common perceptual space was corroborated in SVD collection of attributes and their successful embedding into perceptual spaces of violin and organ sounds. But their mutual relations (similarity, independence or opposition) in individual sound contexts were not always preserved. Also the listening tests respondents preferred other attributes very often used in the questionnaire of musicians opinions (*soft*, *round*) but positioned out of common perceptual space dimensional axes.

A new (fourth) dimension originated by attributes *rustle* in violins and *noisy* in organs could be also a perceptual importance. This 'noise' dimension has not occurred in musicians opinions apparently due to it negative meaning according to sound quality and aesthetic feeling.

The spectral features mainly influencing the perception and their relations were identified by embedding of spectral characteristics into sound context perceptual spaces. Importance of fundamental, high frequency components, spectral energy distribution and noise components (originated below fundamental) revealed in most of violin and organ sound contexts.

The relations between embeddings of verbal attributes and of spectral characteristics can be evaluated when we look on the angles they contain (Tables 8 and 9).

Table 8. Violin: The angles in degrees contained by embeddings of selected verbal attributes and perceptually the most important spectral characteristics. The angles satisfying conditions of similarity or opposition are bold.

$\mathbf{L_{H1}}$	B3	F#4	C5	G5	D6
sharp	136	145	139	129	100
dark/gloomy	39	38	29/54	60	21/59
full	32	58	**4**	61	–
narrow	**163**	**162**	154	156	**174**
soft	42	43	24	71	74
$\mathbf{f_{cg}}$					
sharp	**18**	**4**	**17**	40	24
dark/gloomy	148	159	151/**176**	131	145/**177**
full	150	138	118	108	–
narrow	45	21	32	35	62
soft	151	115	146	120	**162**
$\mathbf{L_{B22}\text{-}L_{B24}}$	mean value of $\alpha[^o]$				
sharp	40	65	**18**	58	**15**
dark/gloomy	137	130	**164**/152	130	109/147
full	128	136	150	**13**	–
narrow	68	62	**17**	131	98
soft	140	115	**162**	141	**162**

Verbal attribute *narrow* revealed especially stable relationships with the fundamental, but extremely unstable with high frequency components (B_{22}–B_{24}), where *full* is least stable. Other attributes span of values over all five pitches varies from 40 to 60 degrees. *Narrow* embeddings have very often nearly opposite position to fundamental, *sharp* similar position to spectral centre of gravity.

Table 9. Organ Principal 8': The angles in degrees contained by embeddings of selected verbal attributes and perceptually the most important spectral characteristics. Corresponding third-octave bands of high harmonics and noises for individual C-tones see in Figure 12. The angles satisfying conditions of similarity or opposition are bold.

L_{H1}	C2	C3	C4	C5 (L_{H2})	C6
narrow	105	**164**	144	81	127
noisy	–	41	35	108	–
round	72	45	67	148	45
sharp	116	102	91	84	118
soft	72	78	41	139	50
high harm.	mean value of α[°]			$T_{22,23}/T_{25,26}$	
narrow	63	87	66	23/67	
noisy	–	45	57	85/58	–
round	**162**	111	141	90/156	106
sharp	60	36	46	22/57	57
soft	**164**	120	121	115/92	92
noises	mean value of α[°]				
narrow	–	105	31	60	136
noisy	–	53	145	156	–
round	–	100	110	87	124
sharp	–	**11**	84	65	140
soft	–	133	140	135	119

Verbal attributes *round* and *soft* revealed especially stable relationships with noises, but extremely unstable with the fundamental. *Narrow* has unstable relationship with noises, *sharp* in opposite high stability with fundamental and high frequency components. The verbal attribute *noisy* (*rauschig* in German) has stable relationships with high frequency components, but extremely unstable and sometimes of high values with noises. This is an example of the situation which requires additional listening tests for causality verification. Other attributes span of values over all five pitches varies from 60 to 80 degrees.

In violins the causality between specific spectral features and attributes *sharp*, *narrow*, *glossy*, *buzzing* and *rustle* was verified in additional listening tests.

The stability of relationship of embedding of verbal attributes and spectral characteristics is higher in violin then in organ Principal 8' sounds, which can among others signify the higher quality of test procedures in violins.

6. Conclusion

The dimensional attributes of three-dimensional common perceptual space together with 'sharpness' and 'noise' are satisfactory for the description of the main perceptual contrasts in the most studied violin and organ contexts focusing on quasi-stationary sounds.

The future research in timbre has huge tasks to broaden knowledge about more sound contexts (other pitches, dynamics, playing techniques) in individual musical instruments or their mixtures. The focus will be made on the search of perceptually important spectral differences among instruments as well as on the search of spectral invariants of the most important (dimensional) attributes on different instruments.

Introducing of the time characteristics of sounds into the listening test results interpretation and the study of the transients and complete sounds is further broadening of the timbre research.

All here enumerated tasks are the challenge for many researchers which will have to cooperate to improve the research effectiveness and portability of methods and results.

Acknowledgements

The research was supported by the Ministry of Education and Youth of the Czech Republic, partly by the Project No. MSM 511100001 and partly by the Project No. 1M6138498401, partly by the Grant Agency of the Czech Republic Project No. 202/02/1370 and partly by the EU CRAFT project G1ST-CT-2002-50267–DEMORGPIPE.

References

[1] - American national standard (1973): Psychoacoustical terminology. Timbre. ANSI S3.20-1973, New York, American National Standards Institute, 56.

[2] Hajda, J. M., Kendall, R. A., Carterette, E. C., Harshberger, M. L. (1997): Methodological issues in timbre research. In Deliege, I., Sloboda, J. (Ed.), Perception and Cognition of Music, *Psychology Press*, 253-286.

[3] Helmholtz, H. von (1863): Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik. VI Ausgabe Viewec, Braunschweig 1913.

[4] Stumpf, C. (1890): Tonpsychologie I-II. S. Hirzel Verlag, Leipzig 1883.

[5] Guilford, J. P. (1954): Psychometric methods. McGraw Hill, New York.

[6] Grey, J. M. (1977): Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America*, 61 (5), 1270-1277.

[7] Bismarck, G. von (1974): Sharpness as an Attribute of the Timbre of Steady Sounds, *Acustica*, 30, 159-172.

[8] Ferjenčík, J. (2000): Úvod do metodologie psychologického výzkumu, ISBN 80-7178-367-6, Portál, Praha (in Czech).

[9] Štěpánek, J. (2002): Evaluation of timbre of violin tones according to selected verbal attributes, 32nd International Acoustical Conference, European Acoustics Association (EAA) Symposium "ACOUSTICS BANSKÁ ŠTIAVNICA 2002" (ISBN 80-228-1159-9), 129-132.

[10] Štěpánek, J., Otčenášek, Z., Melka, A. (1997): Violin Sound Timbre: I. The Construction of a Perceptual Space from Results of Listening Tests, In *Proceedings of the 31st International Acoustical Conference (EAA Symposium),* Acoustics - High Tatras 97, Zvolen, 94-95.

[11] Štěpánek, J., Otčenášek, Z., Melka, A. (1997): Violin Sound Timbre: II. External Interpretation of Perceptual Space from Results of Listening Tests, In *Proceedings of the 31st International Acoustical Conference (EAA Symposium),* Acoustics - High Tatras 97, Zvolen, 96-97.

[12] Štěpánek, J. (2002): The Study of Violin Timbre Using Spontaneous Verbal Description and Verbal Attribute Rating, Forum Acusticum Sevilla 2002 (ISBN 84-87985-06-8), MUS-06-008.

[13] Winsberg, S., De Soete, G. (1993): A latent class approach to fitting the weighted Euclidean model, CLASCAL. *Psychometrika*, 58, 315-330.

[14] Winsberg, S., De Soete, G. (2002): A bootstrap procedure for mixture models: applied to multidimensional scaling latent class models. Applied Stochastic Models for Business and Industry, 18, 391-406.

[15] Harman, H. H. (1976): Modern factor analysis, Third Edition Revisited, The University of Chicago Press, London.

[16] Borg, I., Groenen, P. (1997): Modern Multidimensional Scaling, Theory and Applications, Springer-Verlag, New York.

[17] Štěpánek, J. (2005): Interpretation and comparison of perceptual spaces, Inter-Noise 2005, Rio de Janeiro, CD-ROM, \papers\doc_1627.pdf.

[18] Štěpánek, J. (2004): Relations between perceptual space and verbal description in violin timbre, acústica 2004, Guimarães, Portugal, CD ROM, AFP 077-S.

[19] Štěpánek, J., Otčenášek, Z. (2002): Spectral sources of selected features of violin timbre, 6 CFA, Lille, 695-698.

[20] Štěpánek, J. (2004): Perception of Sharpness and Narrowness in Violin Tones: Influence of Spectral Components with Changing Pitch, Proceedings of 8th International Colloquium ACOUSTICS'04, Zvolen, 49-51.

[21] Štěpánek, J., Moravec, O. (2005): Verbal description of musical sound timbre in Czech language and its relation to musicians profession and performance quality, CIM05 Montréal, Québec, Canada, Electronic proceedings, file: STEPANEK_J_CIM05_02.pdf.

[22] Štěpánek, J., Moravec, O. (2004): Percepční prostory barvy hudebního zvuku a jejich slovní popis, Závěrečná výzkumná zpráva projektu GA ČR 202/02/1370 (in Czech).

[23] Štěpánek, J., Otčenášek, Z., Melka, A. (1999): Comparison of five perceptual timbre spaces of violin tones of different pitches, CD-ROM of Joint Meeting 137th ASA, 2nd EAA Forum Acusticum 1999, 25th DAGA Berlin, 5aMUb5.

[24] Štěpánek, J., Otčenášek, Z. (1999): Rustle as an Attribute of Timbre of Stationary Violin Tones, CASJ - *Journal of the CATGUT Acoustical Society* (ISSN 0882-2212), Vol. 3, No. 8 (Series II), November 1999, 32-38.

[25] Štěpánek, J., Otčenášek, Z., Moravec, O. (2000): Analytical and Perceptual Detection of Rustle in Stationary Violin Tones, SFA / 5 CFA, Lausanne 2000, 433-436.

[26] Otčenášek, Z., Štěpánek, J. (2002): 'Glossy' and 'Buzzing' in timbre of violin sounds, 6 CFA, Lille, 691-694.

[27] Štěpánek, J., Otčenášek, Z., Syrový, V., Täsch, Ch., Angster, J. (2005): Spectral features influencing perception of pipe organ sounds, Forum Acusticum 2005, Budapest, 465-469.

[28] Štěpánek, J., Syrový, V., Otčenášek, Z., Täsch, Ch., Angster, J. (2006): Verbal description of organ Principal 8' sound, 32. DAGA, Braunschweig, in print.

[29] Rieländer, M. M. (1982): Reallexikon der Akustik, Verlag E. Bochinsky, Frankfurt am Main.

[30] Štěpánek, J., Otčenášek, Z. (2005): Acoustical correlates of the main features of violin timbre perception, CIM05 Montréal, Québec, Canada, Electronic proceedings, file: STEPANEK_J_CIM05_01.pdf.

[31] Štěpánek, J., Otčenášek, Z. (2004): Interpretation of Violin Spectrum Using Psychoacoustic Experiments, CD of Proceedings of the International Symposium on Musical Acoustics (ISMA2004), Nara, Japan, 4-S2-1.

[32] Otčenášek, Z., Štěpánek, J. , Täsch, C. (2005): Spectra of the steady state C tones of the Principal 8' of European pipe organs, 31. DAGA, München, 233-234.

[33] Štěpánek, J., Syrový, V., Otčenášek, Z., Moravec, O. (2005): Development of an Innovative Organ Pipe Design Method, EU project CRAFT–1999–72438 DEMORGPIPE, CRAFT Report, Partner B2.