Acoustical correlates of the main features of violin timbre perception

Stepanek Jan, Musical Acoustics Research Centre, Academy of Performing Arts Prague, Czech Republic
stepanek@hamu.cz www.hamu.cz/zvuk

Otcenasek Zdenek, Musical Acoustics Research Centre, Academy of Performing Arts Prague, Czech Republic
otcenasek@hamu.cz www.hamu.cz/zvuk

Proceedings of the Conference on Interdisciplinary Musicology (CIM05)
Actes du Colloque interdisciplinaire de musicologie (CIM05)
Montréal (Québec) Canada, 10-12/03/2005

Abstract

An experimental approach to the study of timbre is described. Series of listening tests were performed separately with recordings of violin tones B3, F#4, C5, G5 and D6, played on various quality violins. Two main approaches to the discovering of acoustical correlates of most important perceptual features are described. The first approach is directly focused on interpretation of perceptual spaces using acoustic characteristics. The second one uses selected verbal attributes used by listeners to the description of timbre differences and tries to found their explanation using acoustic characteristics. Method of immersion of acoustic characteristic into perceptual space revealed the importance of the first harmonic, higher harmonics and spectral centre of gravity in all studied tones. The change of their relations (similarity, contradiction, or independence) with changing pitch is also discussed. Mutual relations of verbal attributes *sharp*, *dark*, *clear* and *narrow* in sounds of five studied tones are discovered; acoustical correlates of the main attributes *sharp* and *narrow* are found, compared and discussed. Moreover acoustical correlates of attribute *rustle* in high violin tones are mentioned. Findings of both approaches to the study of timbre are compared and discussed.

Introduction

This contribution provides an example of experimental approach to the study of timbre. Contemporary view on timbre, based also on many experimental results, consist in conviction of its multidimensional nature. The understanding of timbre is sometimes associated with the identification of sound source or with verbal description. Both these perceptual phenomena are based on two modes of perception (Handel 1995): 1) source mode is focused on physical invariants connected with the nature of sound source (vibration mechanism structure and manner of its excitation); 2) interpretative mode is focused on acoustic properties of sound which is created in specific conditions, connection between these properties and the source is learned by experience.

Traditional definition of timbre is operational (see for example ANSI S3.20-1973) enabling postulation of conditions for the creation of experimental sound contexts (usually a set of sounds of equal pitch, loudness and duration) and for the postulation of criteria for the judges decision making (for example rating or description of differences in pairs of sounds); these aspects are frequently used in timbre studies (Grey 1977; Bismarck 1974).

Actually used approach is based on psychological experiment methods used in psychoacoustics (Guilford 1954), on description of signals by means of acoustic characteristics obtained by signal analysis used in acoustics and based also on physiology of hearing (Moore 1995), on statistical methods enabling evaluation of experiment results (for example Borg & Groenen 1997; Harman 1976), and on computational methods used for experiment results interpretation.

Used approach to timbre study lies in search of basic perceptual features (dimensions), or their verbal description used by judges for the description of timbre differences on specific sound contexts (musical signals), in search of main acoustic characteristics correlated with these perceptual features, in verification of causality between acoustic characteristics and perceptual features (or verbal attributes), in comparison of above-mentioned findings for different sound contexts and their generalization.

The study of timbre of stationary parts of violin tones is described in this contribution.

Methods

Series of listening tests were performed separately with violin tones B3 (played on string G), F#4 (string D), C5 (string A), G5 and D6 (string E), played on various quality violins in the prescribed manner (detache, non vibrato, bow position naturale, mezzo forte), recorded in an anechoic room and manipulated to weaken the influence of transients on perception (sounds of the same pitch – accuracy of fundamental frequency was ±5 cents, loudness – levels from 75 to 80 dB, and duration 1135 ms with fade in 115 ms and fade out 170 ms). Judges in listening tests were professors and students of violin play from the Faculty of Music, Academy of Performing Arts in Prague.

Perceptual spaces constructed from the results of listening tests, verbal description of timbre of violin sounds and acoustical characteristics of these sounds were used for the tests evaluation. Some results of these studies were published on previous acoustic conferences, the main findings will be summarized in following paragraphs together with results of newly carried out analyses.

Following acoustic characteristics were calculated from time-average power spectra:

1) Levels of individual harmonics L_{Hi} (harmonic spectra).
2) Levels in critical bands L_{Bi} (Bark spectra).
3) The spectral centre of gravity f_{cg} (spectral centroid) as a characterisation of spectral energy distribution. The value of f_{cg} is defined by formula:

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

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 most important perceptual features are described in following paragraphs. The first approach is directly focused on interpretation of perceptual spaces using acoustic characteristics. The second one uses selected verbal attributes and tries to found their explanation using acoustic characteristics.

Perceptual spaces

Dissimilarity pair test

Timbre dissimilarity pair test with seventeen sounds for each studied tone was performed with twenty judges. Test results were evaluated using latent class approach (CLASCAL MDS, see Winsberg & De Soete 1993, Winsberg 2002) and perceptual spaces were constructed (Stepanek 2003, Stepanek & Otcenasek 2003). Optimal models are given in Table 1.

Table 1. Optimal latent class models: common perceptual spaces (CiDjSk, i ... number of respondent classes, j ... number of dimensions, k ... specificities: 0=No, 1=Yes).

Tone	B3	F#4	C5	G5	D6
Model	C2D3S0	C2D3S0	C2D2S1	C2D2S1	C2D2S1

Immersion of acoustic characteristics

Immersion of external variable into N dimensional perceptual space was defined as direction in perceptual space, based on the optimal fitting of external scale defined by values of external variable in individual sounds (Borg & Groenen 1997). External scale fitting is calculated using multiple regression formula (2) with external variable as dependent variable y and dimension coordinates as independent variables x_1, ..., x_N:

$$y = b_0 + \sum_{n=1}^{N} b_n x_n \quad (2)$$

Immersion is determined by direction cosines which are calculated using regression weights. Measure of success of immersion is given by multiple correlation coefficient; its value is equal to the Pearson correlation coefficient between external variable values and coordinates of the projection of sounds in perceptual space onto external variable immersion. So the successful immersion has significant positive correlation and reproduces well external variable, reproduction is expressed by coordinates of the projection.

Values of calculated acoustic characteristics were immersed into perceptual spaces. Only immersions with significant value of Pearson correlation coefficient on the level at least 5% were taken into account. The best immersed characteristics revealed significance levels 1% or even 0.1%, they are summarized in Table 2. The angles contained by selected attribute immersions for tones B3 and F#4 (three-dimensional perceptual spaces) are assigned in Table 3. The immersion directions for tones C5, G5 and D6 (two-dimensional perceptual spaces) are drawn in Figure 1.

Table 2. Acoustic characteristics most successfully immersed into perceptual spaces of individual tones.

Acoustical characteristic	Immersion 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

Table 3. Angles in grades (°) contained by immersion directions of selected acoustic characteristics in perceptual space of tones B3 and F#4 (three-dimensional perceptual space solution).

B3	L_{H1}	L_{H4}	L_{H5}	L_{B17}	L_{B22}	L_{B23}	L_{B24}	f_{cg}
L_{H1}	–							
L_{H4}	92	–						
L_{H5}	94	28	–					
L_{B17}	129	137	135	–				
L_{B22}	98	75	103	84	–			
L_{B23}	103	77	104	80	5	–		
L_{B24}	95	79	107	83	5	8	–	
f_{cg}	119	97	119	54	30	25	30	–

F#4	L_{H1}	L_{H2}	L_{B22}	L_{B23}	L_{B24}	f_{cg}
L_{H1}	–					
L_{H2}	97	–				
L_{B22}	122	64	–			
L_{B23}	131	51	19	–		
L_{B24}	132	57	15	6	–	
f_{cg}	142	118	67	71	67	–

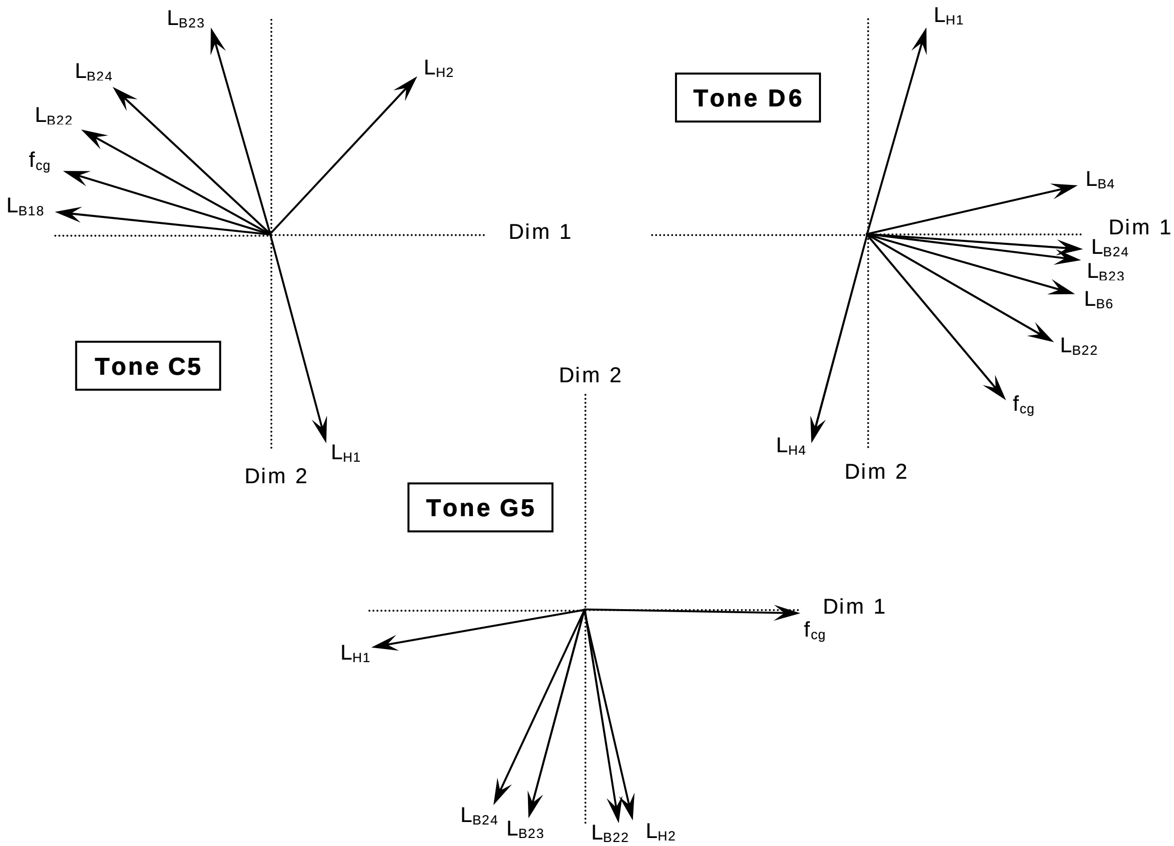

Figure 1. The most successfully immersed acoustic characteristics into perceptual space, immersions are indicated by arrows, direction of arrow agrees with growing of immersed characteristic. Drawings for tones C5, G5 and D6 with two-dimensional perceptual space solution are presented.

Discussion

Only acoustic characteristics successfully immersed into perceptual space are the candidates for the substantial influence on the perception on given sound context (Table 2). In all five studied tones these candidates revealed to be levels of fundamental (H1), levels of high frequency components (bark bands B22 – B24) and spectral energy distribution represented by values of spectral centre of gravity (centroid f_{cg}).

The discussion of relations of immersed acoustic characteristics in individual tones was based on the search of (nearly) orthogonal or (nearly) opposite directions, system of *n* orthogonal immersions in n-dimensional perceptual space was considered as its successful interpretation. Orthogonality of immersions means independent influence of immersed characteristics on perception, opposite directions means contradictory influence. Other criteria for the interpretation of immersions were the relation to the violin sound spectral shape and frequency position of the characteristic.

Tone B3 (fundamental frequency 247 Hz). High frequency components revealed similar influence on perception (see very small angles between immersions of levels of barks B22 – B24, Table 3). This group of directions originated (nearly) orthogonal system with the immersion of H1 and with immersions of the fourth (H4) and fifth (H5) harmonics. Immersion of B17 is not so different from orthogonality with B22 – B24 group. Immersion of spectral centroid is not fully defined by high frequency components B22 – B24 but could be also influenced by violin formant (B17).

Tone F#4 (fundamental frequency 370 Hz). The immersion of the group of B22 – B24 is not as compact as in tone B3 (Table 3). Only immersions of H1 and H2 are nearly orthogonal. The centroid is much more diverting from the immersion of the group B22 – B24.

Tone C5 (fundamental frequency 523 Hz). The immersion of the group B22 – B24 is even more spread then in lower tones (Figure 1) but centroid immersion is between the group and end of violin formant (B18) immersions. Nearly orthogonal are immersions of H2 and group B22 – B24.

Tone G5 (fundamental frequency 784 Hz). The behaviour of immersions in this tone is very different from all others. The immersion of the group B22 – B24 is otherwise again more spread (Figure 1) but centroid immersion is nearly orthogonal to this group and nearly opposite to H1, which is nearly orthogonal to H2.

Tone D6 (fundamental frequency 1175 Hz). The immersions of B22 – B24 are again more spread (Figure 1), centroid is immersed near to this group. New group of immersions (B4, B6), representing frequencies below fundamental with basic resonant modes of violins A0, T1, C3 (modes designation see Moral & Jansson 1982), is mixed with the group B22 – B24. Opposite immersions of H1 and H2 are (nearly) orthogonal to both groups.

Verbal attributes

Spontaneous verbal description (SVD)

Spontaneous verbal description (SVD) of timbre differences in pairs of sounds was performed on representative subset of eleven sounds for each of five studied tones with ten respondents. The analysis of SVD results was based on frequencies of occurrence of attributes on individual sounds summed over all test respondents.

Correlation analysis results revealed relations between verbal attributes (similarity or contradiction) and offered also external data for the interpretation of perceptual spaces constructed from the results of dissimilarity pair test (Stepanek 2004 b). Comparison of results for all five studied tones led to the establishing of four possible dimensions of timbre of stationary violin sounds (Stepanek & Otcenasek & Melka 1999), see Table 4.

Table 4. Possible dimensions of timbre of stationary violin sounds.

Dim.	Dimensional attributes: in Czech (in English)
I.	*měkký* (*soft*) – *ostrý* (*sharp*)
II.	*jasný* (*clear*) – *zastřený* (*damped*)
III.	*tmavý* (*dark*) – *světlý* (*bright*)
IV.	*úzký* (*narrow*)

Verbal attribute ranking and rating (VARR)

The four verbal attributes of violin timbre: *sharp*, *clear, dark* and *narrow* were selected for following experiments. The method of Verbal attribute ranking and rating (VARR) (Stepanek 2002 a; 2002 b) was developed, implemented on PC and applied on above mentioned attributes and also for the assessment of perceived sound quality. The same sounds of all five tones as in SVD test were used with eleven respondents. Resulted rates were evaluated separately for each studied tone using factor analysis (Harman 1976). The scheme of the development of the relations between attributes according to changing pitch is in Figure 2. Successful prediction of perceived sound quality from these attributes was also verified for all five tones (Stepanek 2002 a).

The following selections for subsequent investigation were made based on results demonstrated in Figure 2:

a) Two main and representative verbal attributes of violin timbre *sharp* and *narrow*.
b) Sounds of tones B3, C5 and D6.

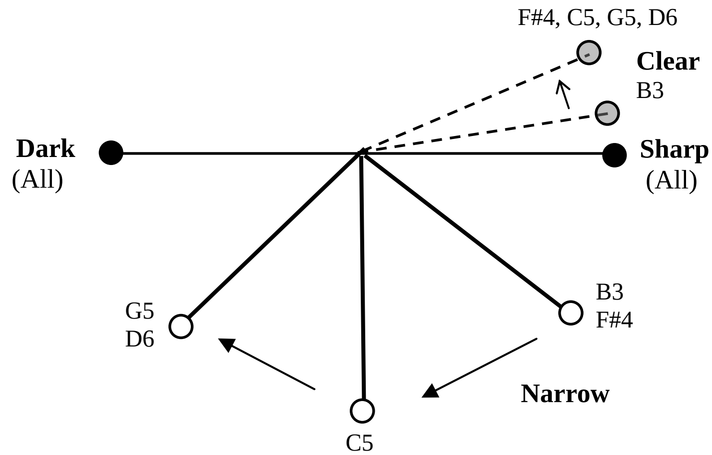

Figure 2. The scheme of two-dimensional factor analysis solution demonstrating relations among verbal attributes and their changes with pitch.

Sharp and narrow

Acoustic characteristic values were correlated with rates of verbal attributes *sharp* and *narrow*. Possible spectral sources for the perception of these verbal attributes were recognized. Originally no correlation between attribute *narrow* and any acoustic characteristic was found in tone C5, but the division of respondents into two groups was satisfactory. Subsequent listening tests using appropriate manipulation of selected original sounds (Stepanek 2004a; Stepanek & Otcenasek 2002; 2004) supported verification of spectral sources for *sharp* and *narrow*, which are presented in Figure 3.

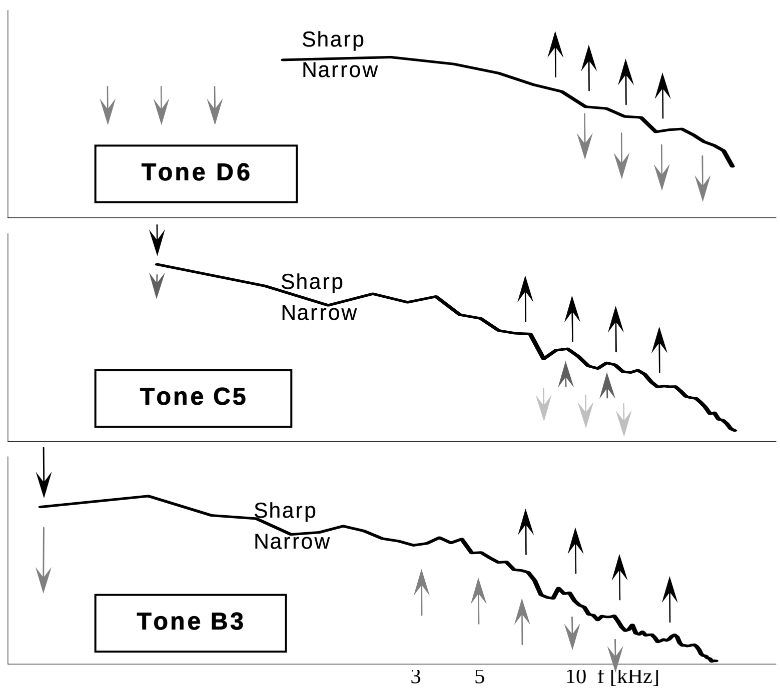

Figure 3. Increase of the *sharp* and *narrow* attributes caused by spectrum level change in arrow direction (the length corresponds to the size of increase). There are two different strategies for C5 and *narrow*. The arrows in low frequency region for *narrow* in tone D6 are for noise components below the fundamental. Spectral envelopes are mean values from eleven signals used in SVD and VARR experiments.

Rustle

Substantial increase of use of verbal attribute *šustivý* (*rustle*) in highest studied tone D6 was the main reason for the search for spectral sources of perception of this attribute (Stepanek & Otcenasek 1999). Significant correlations with acoustic characteristics were found in two frequency regions displayed in Figure 4. Subsequent listening tests using appropriate manipulation of selected original sounds supported predominant origin of *rustle* perception in spectral components below the fundamental (Figure 4).

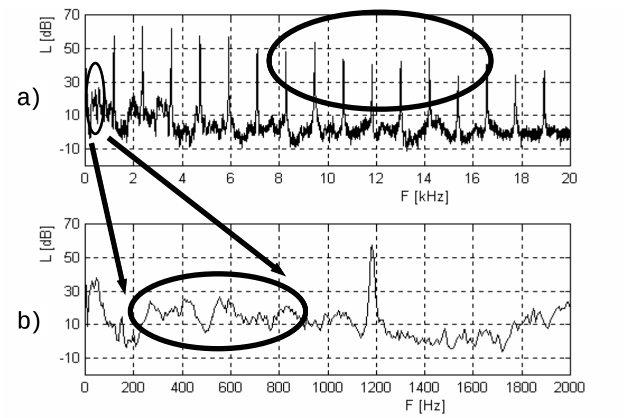

Figure 4. The tone D6 potential sources of *rustle* (marked with ellipses); a) the power spectrum, b) detail of its lower part. Predominant source of rustle was identified in spectral components below the fundamental.

Further investigations revealed connection of these components with violin body radiation near resonant modes A0, T1, C3 (modes designation see Moral & Jansson 1982) induced by aperiodic events in bowed string motion (Stepanek & Otcenasek 1999). Comparison of spectra of sounds of all five tones of violin with high degree of *rustle* with frequency characteristic of the same instrument is in Figure 5.

Conditions for perceptual detection of *rustle* were studied (Stepanek & Otcenasek & Moravec 2000) and levels of spectral components sufficient for disturbance by *rustle* were specified. These levels are reached predominantly in low quality instruments.

Discussion

Additional tests revealed that from the hypothesis of possible violin timbre dimensions (Table 4) only two main dimensions were acknowledged and are sufficient for the prediction of perceived sound quality of violin sounds. Attributes *sharp* and *clear* were rated similarly but contrariwise to *dark*, second dimension was created by "moving" attribute *narrow*. There is no contradiction with common perceptual space of attributes found in tests made without sound context but based on respondent experiences and opinions (Stepanek & Moravec 2005), only the first dimension (*harsh* – *delicate*) is missing.

The position of attribute *narrow* is changing from positive correlation with *sharp* in lower tones B3 and F#4 through independence in C5 to positive correlation with *dark* in highest tones G5 and D6 (Figure 2). These findings were corroborated in subsequent tests verifying spectral sources of attributes *sharp* and *narrow* (Figure 3). The influence of the first harmonic on the perception of *sharp* and *narrow* is decreasing with increasing pitch (Figure 3). The existence of two groups of listeners in *narrow* judgement in tone C5 (one with "B3 strategy", second with "D6 strategy") exhibited that the speed of strategy change with increasing pitch is not the same for all respondents.

Noise components of violin sounds expressed by listeners as *rustle* can influence the perception in high violin tones (here in D6) with negative impact on perceived sound quality. Periodicities in bowed string motion provoke radiation from resonant modes A0, T1, C3 (in violin body in frequency band 200 – 700 Hz that is below the fundamental in tone D6).

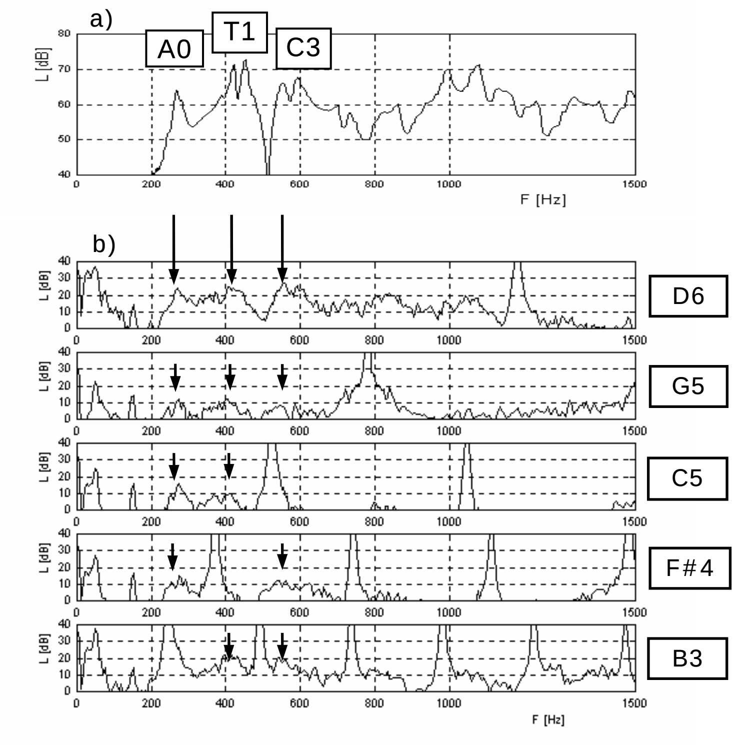

Figure 5. Comparison of the violin frequency characteristic (a) with time-averaged spectra of quasi-stationary part of tones (b) for the violin with high degree of *rustle*. Spectral signatures of violin modes A0, T1 and C3 in spectra are marked with arrows.

General discussion

The two approaches were used to the search for the most important perceptual features of timbre of stationary violin sounds and their acoustical correlates. The first approach was based on the assumption that most important perceptual features are contained in perceptual space constructed from dissimilarity judgements using multidimensional analysis. Acoustical correlates responsible for its creation were searched using immersion of acoustic characteristics. The first harmonic, higher harmonics and spectral energy distribution revealed to be most important in a span of fundamental frequencies going from 247 Hz to 1175 Hz, but their mutual relation changed with pitch.

Verbal description of timbre was used as a starting point of the second approach. *Sharp* and *narrow* remain the most important verbal attributes for all studied tones, but their mutual relation changed with pitch like did their acoustic sources (again the first and higher harmonics). The method of hypotheses verification using appropriately manipulated signals revealed to be beneficial, just the correlation is not always satisfactory as was proved in *rustle* investigation. Attribute *rustle* importance in tone D6 and its acoustic sources are in good agreement with successful immersions of the same frequency bands (barks B4, B6) into perceptual space of tone D6.

The future investigation must apart from others answer following questions: Are found acoustic characteristics important really for violin sound or only specifically for studied sound contexts (set of selected instruments)? Do the acoustic invariants of here described verbal attributes (independent on pitch or on acoustic source – instrument type) exist? What are they? Or opposite: Why are the different kinds of acoustic characteristics denominated using the same word?

Acknowledgments. 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.

References

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

Bismarck, G. von. (1974). Timbre of steady sounds: A factorial investigation of its verbal attributes. *Acustica* 30: 146-159.

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

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

Guilford, J. P. (1954). *Psychometric methods*. McGraw Hill, New York.

Handel, S. (1995). Timbre Perception and Auditory Object Identification. In *Hearing, Handbook of Perception and Cognition*, (Editor Moore, B. C. J.) 2nd Edition, San Diego California, Academic Press: 425-461.

Harman, H. H. (1976). *Modern factor analysis*. Third Edition Revisited, The University of Chicago Press, London.

Moore, B. C. J. (Editor) (1995). *Hearing, Handbook of Perception and Cognition*, 2nd Edition, San Diego California, Academic Press.

Moral, A. J., and Jansson, E. (1982). Eigenmodes, input admittance, and the function of the violin, *Acustica* 50: 329-337.

Stepanek, J. (2002a). Evaluation of timbre of violin tones according to selected verbal attributes, In *Proceedings of 32nd International Acoustical Conference, European Acoustics Association (EAA) Symposium "ACOUSTICS BANSKÁ ŠTIAVNICA 2002"*: 129-132.

———. (2002b). The Study of Violin Timbre Using Spontaneous Verbal Description and Verbal Attribute Rating. In *Forum Acusticum Sevilla 2002*: MUS-06-008.

———. (2003). Latent listener classes and class models in violin timbre. In *Proceedings of DAGA 2003*, Aachen: 532-533.

———. (2004a). 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.

———. (2004b). Relations between perceptual space and verbal description in violin timbre. In *acústica 2004 Guimarães*, Portugal, CD ROM: AFP 077-S.

Stepanek, J., Moravec, O. (2005). Verbal description of musical sound timbre in Czech language and its relation to musicians profession and performance quality. *CIM05 Montreal*.

Stepanek, J., Otcenasek, Z. (1999). Rustle as an Attribute of Timbre of Stationary Violin Tones, *CASJ - Journal of the CATGUT Acoustical Society*, Vol. 3, No. 8 (Series II), November 1999: 32-38.

———. (2002). Spectral sources of selected features of violin timbre, In *Proceedings of 6 CFA*, Lille: 695-698.

———. (2003). Listener common and group perceptual dimensions in violin timbre. In *Proceedings of the Stockholm Music Acoustics Conference (SMAC 03)*, Stockholm, 663-666.

———. (2004). Interpretation of Violin Spectrum Using Psychoacoustic Experiments, In *CD of Proceedings of the International Symposium on Musical Acoustics (ISMA2004)*, Nara, Japan: 4-S2-1.

Stepanek, J., Otcenasek, Z., Melka, A. (1999). Comparison of five perceptual timbre spaces of violin tones of different pitches, In *CD-ROM of Joint Meeting 137th ASA, 2nd EAA Forum Acusticum 1999, 25th DAGA Berlin*, 5aMUb5.

Stepanek, J., Otcenasek, Z., Moravec, O. (2000). Analytical and Perceptual Detection of Rustle in Stationary Violin Tones, In *Proceedings of SFA / 5 CFA*, Lausanne 2000: 433-436.

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

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