LISTENER COMMON AND GROUP PERCEPTUAL DIMENSIONS IN VIOLIN TIMBRE

Jan Štěpánek, Zdeněk Otčenášek

Sound Studio of the Faculty of Music, Academy of Performing Arts Prague
stepanek@hamu.cz, otcenasek@hamu.cz

ABSTRACT

Results of listening tests of violin tones are studied with respect to listeners and their perceptual models. Five sets of violin tone recordings (pitch B3, F#4, C5, G5, D6) were used. Twenty experienced listeners – violin players – assessed dissimilarities in timbre in pairs of tones.

The results of five listening tests (individual dissimilarity matrixes) were separately processed using latent class approach (CLASCAL). This approach yields to perceptual spaces of common dimensions shared with all listeners and defines listener classes (groups). The perceptual models of CLASCAL groups were also studied. The results revealed that the stability is higher in group perceptual models then in common models of all listeners. Affinity of dimensions and models across groups showed potential existence of additional perceptual dimensions in some class models.

1. INTRODUCTION

1.1 Timbre studies

Usually the aim of study of musical sound timbre is a detection of distinctive features of specific sound context and the search for their acoustic correlates. More technically, a (multidimensional) perceptual space of sound stimuli is constructed and attempts are made to interpret its individual dimensions using adequate spectral characteristics. Subjective data are collected in listening tests using different methods that allow the use of Factor analysis [1] or Multidimensional scaling of different kind [2, 3] for the evaluation of results. Some studies have been focused only on stationary parts of the sounds [1, 2] or on both transient and stationary parts [3]. As for listeners, one group or more groups of judges (e.g. musicians and nonmusicians in [1]) are apriori chosen to participate in listening tests.

1.2. Listener classes

A question arises whether the listeners are in higher or lower concordance in their judgements or even whether there exist groups of listeners with different perceptual models. Latent class approach (CLASCAL MDS) [4, 5] results in a perceptual space of common dimensions shared by all listeners, but it also determines classes of listeners based on the similarity of individual perceptual models (aposteriori defined groups of listeners); the group models differ in weights of common dimensions.

In this contribution the study of violin timbre focused on stationary part of sounds is described. The results of listening tests (individual dissimilarity matrixes) were processed using latent class approach applied on weighted Euclidean model (CLASCAL) and extended CLASCAL model [4] (overview of models see in [6]). The first results of these analyses – common perceptual space dimensionality and main source of the differentiation of listeners into classes – were described in [7]. Further the classical Euclidean model and extended Euclidean model [6] were applied separately on a set of individual dissimilarity matrixes of each class (group) of listeners. Common and group perceptual spaces and dimensions were further compared.

2. METHOD

Five sets of violin tone recordings (pitches B3, F#4, C5, G5, D6) were used in the study [8]. Attack and decay transients were unified to weaken their influence on judgements. Seventeen tones for each pitch were listened in headphones and judged. Twenty experienced listeners – violin players (four Academy professors and sixteen students) assessed dissimilarities in timbre in pairs of violin tones with marks 0, 0.5, 1, ..., 4.5, 5.

2.1. Latent class approach

Five sets of individual dissimilarity matrixes were separately processed using latent class approach applied on weighted Euclidean model (CLASCAL) and extended CLASCAL. Latent class approach (CLASCAL MDS) solves two optimization tasks:

1. To fit stimuli dissimilarities into distances of Euclidean space of appropriate low dimension.
2. To establish the appropriate number of (latent) classes of listeners and to add each individual listener to one of these classes (aposteriori grouping).

Optimal model selection is an iterative process in which the solution of the first and the second optimization task is alternately improved.

The new version of the CLASCAL program with a bootstrap procedure [5] was used. It enables an application of latent class approach with both weighted Euclidean model (each stimulus is described by its coordinates in common dimensions; each class of listeners has its own weight of every dimension) and extended weighted Euclidean model (each stimulus is described moreover by its specificity value, which indicates the existence of a stimulus feature not shared with other stimuli; set of all specificities is weighted separately for each class of listeners).

The CLASCAL MDS procedure generally yields to a model, which best fits dissimilarity data. Every discussed model will be denoted here as $C_iD_jS_k$: *i* classes (C_i), *j* dimensions (D_j) and with or without specificities (S_1 or S_0). The model parameters are as follows:

- The number of listener classes; each listener belongs to a certain class
- The number of common dimensions

- The stimulus coordinates according to the dimensions, defining its position in the common perceptual space
- The stimulus specificity value (optional), one for each stimulus
- The weight, one for each common dimension (eventually for a set of all specificities) and each class

The application of CLASCAL procedure on the studied data is described in [7]. The main results may be summarized in all five pitches as follows:

a) Two classes revealed as the best solution ($C_2D_jS_k$); division of the listeners into classes is very stable.
b) The weights of all dimensions/specificities of the first class are less then those of the second class.
c) One of the differences between classes is the measure of the exploitation of dissimilarity scale by individual listener; the mean dissimilarity value of the first class is significantly less then that of the second class.

2.2. Comparison of models

Outstanding listeners were excluded from each class before calculating individual class models to obtain more homogeneous classes. The exclusion criteria consisted in instability in class belonging with respect to models with different number of dimensions, in bootstrap instability (see Figure 2 in the next paragraph), in extreme value of the mean dissimilarity with respect to the other class members.

Euclidean and extended Euclidean models [6] applied on every clarified listener class resulted into models denoted here as $Class_iD_jS_k$. The most appropriate model for each class was established.

As a model development we will denote the sequence of models with increasing number of dimensions (number of classes and type of specificities S_0 or S_1 is not changed). Two types of model developments will be used here: $C_2D_jS_k$ = {$C_2D_1S_k$, $C_2D_2S_k$, ..., $C_2D_6S_k$} for group of all listeners and $Class_iD_jS_k$ = {$Class_iD_1S_k$, $Class_iD_2S_k$, ..., $Class_iD_6S_k$} for group consisted of listener class *i*. The Pearson correlation coefficient between stimuli coordinates of individual dimensions in two different models was used to compare them:

1. Inside model development – correlations in consecutive models exhibit stability of an individual dimension and (qualified over all dimensions) stability of the listener group model.
2. Across model developments – correlations exhibit affinity of dimensions between groups and (qualified over all dimensions) affinity of models of different groups of listeners.

3. RESULTS

3.1. Clarification of classes

The optimal model selection procedure is described in [7] where also results for all studied pitches are in details and they are summarized here in the paragraph 2.1. Two classes were found as the most appropriate in all five pitches, in the tones B3 and F#4 it was the model $C_2D_3S_0$ and in tones C5, G5 and D6 it was the model $C_2D_2S_1$.

The outstanding listeners were excluded from each class to obtain more homogeneous classes; the exclusion criteria are listed in the paragraph 2.2. For an excluded listener, several criteria were often met. Exclusion of outstanding listeners is marked in Figure 1; bootstrap instability is illustrated by a cluster tree in Figure 2. The number of listeners in clarified classes is summarized in Table 1.

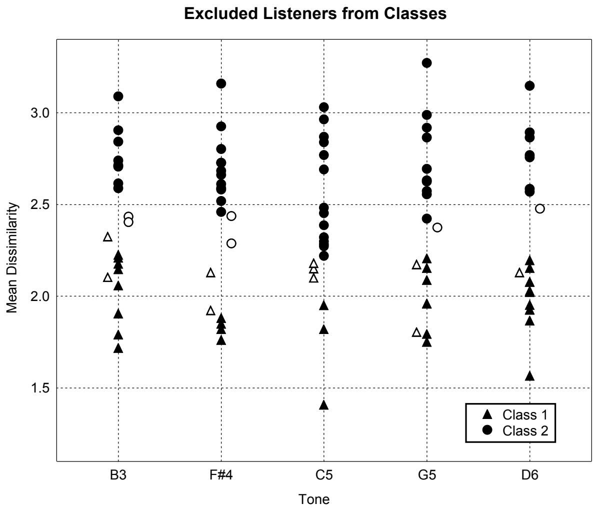

Figure 1: *Mean dissimilarity values of individual listeners belonging to class 1 or class 2. Excluded outstanding listeners are marked by not filled point patterns.*

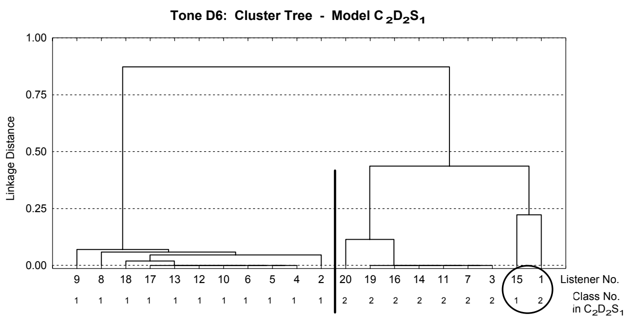

Figure 2: *Example of listeners cluster tree for tone D6. Listener distances are based on bootstrap procedure – see [5, 7]. Excluded listeners are marked by a circle.*

Tone	B3	F#4	C5	G5	D6
Class 1	**8**=10-2	**4**=6-2	**3**=6-3	**7**=9-2	**11**=12-1
Class 2	**8**=10-2	**12**=14-2	**14**=14-0	**10**=11-1	**7**=8-1

Table 1: *Number of listeners in clarified classes (bold). Original number of listeners and number of excluded listeners are on the right side of the equation.*

3.2. Class models

Euclidean models ($Class_iD_jS_0$) and extended Euclidean models ($Class_iD_jS_1$) [6] were calculated for each clarified listener class. The most appropriate models for each class both without and with specificities are summarized in Table 2 together with the most appropriate models found previously for group of all listeners.

Tone	B3	F#4	C5	G5	D6
All (S_0)	**$C_2D_3S_0$**	**$C_2D_3S_0$**	$C_2D_4S_0$	$C_2D_4S_0$	$C_2D_4S_0$
All (S_1)	$C_2D_2S_1$	$C_2D_2S_1$	**$C_2D_2S_1$**	**$C_2D_2S_1$**	**$C_2D_2S_1$**
$Class_1(S_0)$	D_3S_0	D_3S_0	D_3S_0	**D_2S_0**	D_2S_0
$Class_1(S_1)$	**D_2S_1**	**D_2S_1**	**D_2S_1**	D_2S_1	**D_2S_1**
$Class_2(S_0)$	D_4S_0	**D_6S_0**	**D_5S_0**	D_5S_0	D_2S_0
$Class_2(S_1)$	**D_3S_1**	D_3S_1	D_2S_1	**D_2S_1**	**D_2S_1**

Table 2: *The most appropriate models for group of all listeners (All) and for clarified classes; S_0: without specificities, S_1: with specificities. Optimal models irrespective to specificities are bold.*

3.3. Stability of dimensions inside model development

A stability of dimensions inside the model development was observed using the Pearson correlation coefficient between the stimuli coordinates with increasing number of dimensions. An example of a graphical view of significant correlations (α=5%) between dimensions in the model development is in Figure 3.

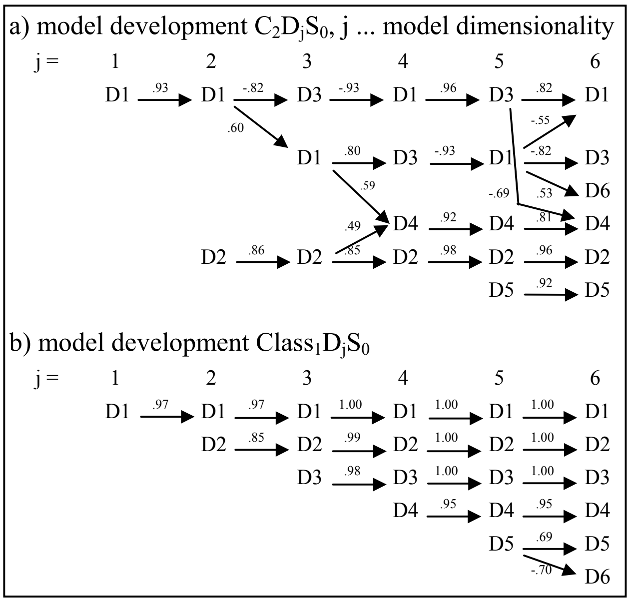

Figure 3: *Significant correlations in model development for tone F#4, a) group of all listeners, b) clarified class 1. Individual dimensions are denoted as D1, D2, ..., D6. Numbers around arrows are values of significant correlation coefficient between dimension coordinates.*

The model developments revealed similar fashion for the group of all listeners and for both classes in all five pitches.

Stability of the dimension coordinates in consecutive models is documented in Figure 4.

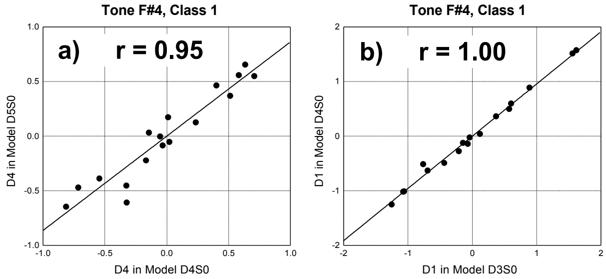

Figure 4: *Stability of dimension coordinates in model development for tone F#4, class 1: a) dimension D4 from models D_4S_0 and D_5S_0, Pearson correlation r=0.95, b) dimension D1 from models D_3S_0 and D_4S_0, r=1.00.*

3.4. Affinity of dimensions across model developments

An affinity of the dimensions across the model developments was observed using the Pearson correlation coefficient between stimuli coordinates of models of different groups (All listeners, Class1, and Class2). Dimensions from most appropriate models (see Table 2) or dimensions from other models having the same number of dimensions were correlated. The examples of significant correlations are in Figure 5.

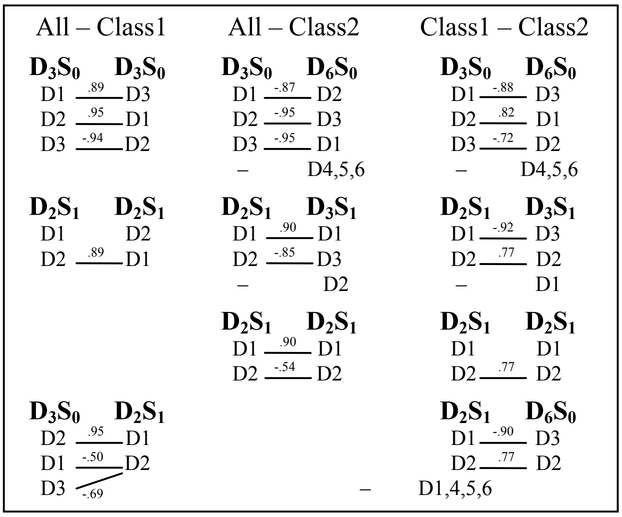

Figure 5: *Examples of significant correlations between dimensions across model developments for tone F#4. Numbers around lines are values of significant correlation coefficient between dimension coordinates.*

From the correlations across the models, it was possible to recognize the dimensions shared across groups. The dimensions were designated by letters according to their stability and stable order in class model developments. For the tone F#4, the dimensions A, B, and C are shared with the groups (this is in accordance to the CLASCAL common dimensions) but they differ in order and sign according to groups.

Shared dimensions and other dimensions used by individual groups for all five studied tones are indicated in Table 3.

Tone	Order	Class 1	Class 2	All listeners
B3	1	A	A	A
	2	B	- [(B) & (C)]	B
	3	C	**D**	C
	4		(B)	
F#4	1	- C	A	- B
	2	A	B	- C
	3	- B	C	- A
	4-6		**D**, **E**, **F**	
C5	1	- A	A	A
	2	- **C**	B	B & (-A)
	3	(**D**)	**C**	
	4	(B)	**D**	
G5	1	A	A	A
	2	B	B	B
	3	C	- C	C & D
	4	D	D	D & C
D6	1	A	A	A
	2	B	- B	B & (-A)

Table 3: *Dimensions used by groups. Extra dimensions added to CLASCAL common dimensions are bold, dimensions correlated only with significance 1%<α<5% are in parenthesis, with α<1% as normal characters.*

4. DISCUSSION

The results of analyses of the clarified CLASCAL classes exhibited some interesting properties in comparison to the usual CLASCAL approach dealing with the common dimensions.

The splitting of the dimensions and the change of their order (see example for the tone F#4 in Figure 3) is frequent in the model development in the group of all listeners. The class models showed stable dimensions (with very high significant correlations), new dimensions are added by dimension increase (dimension splitting is exceptional) and the order of the dimensions is stable.

It was shown previously [7] that the mean dissimilarity values are significantly higher in Class2 (due to use of higher marks in dissimilarity judgements) than in Class1 in all five pitches. As for most appropriate models, Class1 optimal models revealed lower number of dimensions and Class2 the higher ones than models for all listeners (Table 2). Even if there is a good concordance among common dimensions and lower dimensions of individual classes (Table 3), it is not evident whether additional dimensions in Class2 manifest a different perceptual model (use of more dimensions), or only the use of more levels in judgement decreases information "noise" allowing to discern next existing dimensions.

Usually used criteria for establishing the number of dimensions [4, 5] are information criteria (AIC, BIC) or criteria based on randomizing (Monte Carlo, bootstrap). In our approach we prefer interpretability as a criterion for the acceptation of a perceptual dimension. This is the reason we permit also models with the higher number of dimensions, to keep any potential dimension (tones B3, F#4, C5 in Table 3).

Specificity values decreased with increasing model dimensionalities, so models D_jS_0 and D_jS_1 became more similar with increasing dimensionality. The specificity feature of any signal must be judged individually by listening.

5. CONCLUSIONS

Latent class approach (CLASCAL) defines classes of listeners; clarified classes revealed better stability of perceptual dimensions and models than the group of all listeners. Affinity of dimensions and models across groups showed potential existence of additional perceptual dimensions in models of listeners using the full dissimilarity scale. The correlation with some acoustic parameters of signals will be the criterion for the acceptance of a perceptual dimension. A preliminary calculation revealed importance of spectral energy distribution (the center of gravity) and of the levels of fundamental and lower harmonics. The importance of these spectral features for violin timbre must be verified in further listening tests.

6. ACKNOWLEDGEMENTS

The research was supported partly by the Ministry of Education and Youth of the Czech Republic (Project No. 511100001) and by the Grant Agency of the Czech Republic (Project No. 202/02/1370).

The authors are gratefully acknowledged to Suzanne Winsberg for the offering of the new version of the CLASCAL program with bootstrap and for guidelines and beneficial discussions for its use.

7. REFERENCES

[1] Bismarck, G. von, "Timbre of steady sounds: A factorial investigation of its verbal attributes", *Acustica*, 30, 146-159, 1974.

[2] Plomp, R., Aspects of tone sensation, London, Academic Press, 1976.

[3] Grey, J. M., "Multidimensional perceptual scaling of musical timbres", J. Acoust. Soc. Am., 61 (5), 1270-1277, 1977.

[4] Winsberg, S., De Soete, G., "A latent class approach to fitting the weighted Euclidean model, CLASCAL", Psychometrika, 58, 315-330, 1993.

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

[6] McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., Krimphoff, J., "Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes", Psychological Research, 58, 177-192, 1995.

[7] Štěpánek, J., "Latent listener classes and class models in violin timbre", DAGA 2003, Aachen, in print.

[8] Štěpánek, J., Otčenášek, Z., Melka, A., Syrový, V., "Violin Tones Spectra and their Relationship to Perceived Sound Quality" in Proceedings of the Institute of Acoustics ISMA´97, 19 (5), Edinburgh, 125-130, 1997.