Analysis of Musical Timbre Semantics through Metric and Non-Metric Data Reduction Techniques

Asterios Zacharakis,[*1] Konstantinos Pastiadis,[#2] Joshua D. Reiss[*3], George Papadelis[#]

[*]*Queen Mary University of London, Centre for Digital Music, London, U.K.*

[#]*Aristotle University of Thessaloniki, Department of Music Studies, Thessaloniki, Greece*

[1]asteriosz@eecs.qmul.ac.uk, [2]pastiadi@mus.auth.gr, [3]josh.reiss@eecs.qmul.ac.uk

ABSTRACT

This study investigated the underlying structure of musical timbre semantic description. Forty one musically trained subjects participated in a verbal attribute magnitude estimation listening test. The objective of the test was to rate the perceptual attributes of 23 musical tones using a predefined vocabulary of 30 English adjectives. The perceptual variables (i.e. adjectives) were then analyzed through Cluster and Factor Analysis techniques in order to achieve data reduction and to identify the salient semantic dimensions of timbre. The commonly employed metric approach was accompanied by a non-metric counterpart in order to relax the assumption of linear relationships between variables and to account for the presence of monotonic nonlinearities. This rank transformation into an ordinal scale has offered a more compact representation of the data and thus confirmed the existence of nonlinearities. Three salient, relatively independent perceptual dimensions were identified for both approaches which can be categorized under the general conceptual labels: *luminance*, *texture* and *mass*.

I. INTRODUCTION

Musical timbre has been a subject of scientific research since the end of the 19th century. Von Helmholtz (1877) was the first to investigate verbally expressed perceptual attributes of musical timbre and their acoustic correlates. The description of timbre through the use of semantic scales was further explored by other researchers (e.g. Lichte, 1941; von Bismarck, 1974a, b; Kendall and Carterette, 1993a, b). In these approaches, sound objects were represented by a feature vector of semantic attributes which are usually elicited in the form of descriptive adjectives. Additionally, verbal terms have been also used for the description of musical instrument properties (Disley and Howard, 2004; Fritz et al., 2008; Barthet et al., 2010), description of polyphonic timbre (Alluri and Toiviainen, 2010) and acoustic assessment of concert halls (Lokki et al., 2011).

The most widely applied methods for obtaining semantic descriptions of timbre are semantic differential (Osgood et al., 1957) and its variation verbal attribute magnitude estimation (VAME) (e.g. Kendall and Carterette, 1993a,b). In semantic differentiation the rating of the different sounds is made on scales whose extremes are labeled by two opposing verbal attributes such as 'bright-dull', whereas in VAME the scales are labeled by an attribute and its negation ('not harsh-harsh'). Dimension reduction techniques such as Principal Components Analysis (PCA) (e.g. von Bismarck, 1974b; Lokki et al., 2011) or Factor Analysis (e.g. Alluri and Toiviainen, 2010) and classification techniques such as Cluster Analysis (e.g. Kendall and Carterette, 1993a; Disley et al., 2006) are usually applied in order to reduce high dimensionality and increase interpretability of the final solution.

In a highly cited work, von Bismarck (1974a,b) performed a semantic differential listening test where participants were asked to rate the perceptual attributes of 35 steady state synthetic tones using 30 bipolar verbal scales. This study identified four orthogonal perceptual dimensions: fullness (full-empty), luminance and texture (dull-sharp), color (colorful-colorless) and density (compact-diffused).

Other similar studies have also identified three or four perceptual axes. Pratt and Doak (1976), working with simple synthetic tones, proposed a 3-D space featuring dimensions of luminance (bright-dull), temperature (warm-cold) and richness (rich-pure). Stepanek's study (2006) revealed dimensions associated with sight (gloomy-clear), texture (harsh-delicate), fullness (full-narrow) and hearing (noisy/rustle-?). Moravec's work (2003) also resulted in the proposition of four perceptual axes related to sight/luminance (bright/clear-gloomy/dark), texture (hard/sharp-delicate/soft), width (wide-narrow) and temperature (hot/hearty-?). Finally, Disley's study (2006) uncovered four salient dimensions labeled by the terms: bright/thin/harsh - dull/warm/gentle, pure/percussive - nasal, metallic-wooden and evolving.

The application of PCA or exploratory Factor Analysis techniques on the data does not account for potential nonlinear relationships between the measured variables (i.e. verbal descriptors). On the contrary, a non-metric analysis has been shown to relax the more strict assumption of linear relationships between variables allowing for the investigation of monotonic nonlinearities (Woodward, 1976). Building on previous work (Zacharakis et al., 2011), this study investigates the contribution of a rank ordinal transformation towards a better modeling of the relationships between variables.

The following section will describe the method of the listening test as well as the analytic techniques applied on the acquired data. Next, the comparison between the original and the rank transformed data representations will be presented and discussed. The paper concludes by summarizing the findings of this work.

II. METHOD

A listening test using the verbal attribute magnitude estimation (VAME) method was designed and conducted. The subjects were provided with a vocabulary of 30 adjectives and were asked to describe the timbral attributes of 23 sound stimuli by choosing the most appropriate descriptors for each stimulus. Once a subject chose a descriptor he was asked to estimate its relevance on a scale anchored by the extreme of the verbal attribute and its negation, such as 'not brilliant-very brilliant'. This rating was input using a horizontal slider with a hidden continuous scale ranging from 0 to 100. The verbal descriptors provided were intended for the description of sound impression (Wake and Asahi, 1998) and were selected among adjectives that are commonly found in musical timbre perception literature (Ethington and Punch, 1994; von Bismarck, 1974b, a; Faure et al., 1996; Disley et al., 2006). The collection of the terms along with basic statistics for each one is presented in Table 1. However, it has been pointed out that verbal descriptors within a predefined set may not correspond to descriptors chosen spontaneously by the participants (Donnadieu, 2007). In order to address this criticism we allowed our subjects to freely insert up to three additional adjectives of their own choice for describing each stimulus in case they felt that the provided terms were inadequate.

Table 1. Mean and maximum values for each verbal descriptor.

Descriptor	Max	Mean	Descriptor	Max	Mean
Brilliant	33.90	9.20	Rich	31.97	12.90
Hollow	28.04	8.61	Bright	33.53	14.50
Clear	30.02	10.99	Dense	18.70	7.50
Rough	42.61	10.95	Full	37.19	12.02
Metallic	58.82	19.10	Nasal	30.12	10.60
Warm	45.44	13.06	Soft	34.43	8.07
Smooth	22.09	7.14	Dark	29.09	9.35
Thick	33.58	10.74	Compact	16.51	5.55
Rounded	47.34	14.46	Dirty	44.70	8.85
Harsh	39.58	12.32	Empty	12.95	4.75
Dull	21.92	8.18	Messy	27.41	5.55
Thin	30.40	9.37	Light	29.36	5.53
Shrill	24.17	7.17	Dry	24.19	8.13
Cold	14.51	6.61	Distinct	26.48	10.66
Sharp	27.39	8.18	Deep	59.73	8.84

A. Stimuli and Apparatus

Aiming to promote ecological validity, 23 musical timbres drawn from acoustic instruments, electric instruments and synthesizers were employed. The following 14 instrument tones come from the MUMS (McGill University Master Samples) library: violin, sitar, trumpet, clarinet, piano at A3 (220 Hz), Les Paul Gibson guitar, baritone saxophone B flat at A2 (110 Hz), double bass pizzicato at A1 (55 Hz), oboe at A4 (440 Hz), Gibson guitar, pipe organ, marimba, harpsichord at G3 (196 Hz) and french horn at A3# (233 Hz). A flute recording at A4 was also used along with a set of 8 synthesizer sounds: Acid, Hammond, Moog, Rhodes piano at A2, electric piano (rhodes), Wurltitzer, Farfisa at A3 and Bowedpad at A4.

The samples were loudness equalized in an informal listening test within the research team. The RMS playback level was set between 65 and 75 dB SPL (A-weighted). 93% of the subjects found that level comfortable and 85% reported that loudness was perceived as being constant across stimuli.

The listening test was conducted under controlled conditions in acoustically isolated listening rooms. Sound stimuli were presented through the use of a laptop computer (MacBook Pro, 2.4 GHz Intel Core 2 Duo, 4 GB Ram, Mac OSX 10.6.8), with an M-Audio (Fast Track Pro USB) external audio interface, and a pair of Sennheiser HD60 ovation circumaural headphones. The interface of the experiment was built in Max/MSP.

B. Listening Panel

Forty one English speaking subjects (aged 17-61, mean age 29.6, 13 female) participated in the listening test. None of them reported any hearing loss or synaesthesia and all participants were critical listeners and had been practicing music for 18.8 years on average (ranging from 4 to 45). Subjects were researchers from the Centre for Digital Music at Queen Mary University of London, students of the Royal College of Music and of the Music Department of Middlesex University in London.

C. Procedure

Initially the listeners were presented with a preliminary stage which consisted of the random presentation of the entire stimuli set in order for them to become familiar with the timbral range of the experiment. For the main part of the experiment the playback of each sound was allowed as many times as needed prior to submitting a rating. The sounds were presented in random order for each listener. Subjects were advised to use as many of the terms as they felt were necessary for an accurate description of each different timbre, and also to take a break in case they felt signs of fatigue. They were also free to withdraw at any point. The overall procedure, including instructions and breaks, lasted around 45 minutes for most of the subjects. The majority of participants rated the above procedure as easy to follow, clear and meaningful.

D. Cluster Analysis, Exploratory Factor Analysis and rank ordering transformation

Two statistical analysis techniques were applied to the data in order to reach conclusions regarding the salient perceptual dimensions of timbre. Cluster Analysis is a technique that seeks to identify homogeneous subgroups within a larger set of observations (Romesburg, 2004) and has been used in order to indicate groups of semantically related verbal descriptors. Factor Analysis is a multivariate statistical technique that is used to uncover the latent structure of a set of inter-correlated variables (Harman, 1976).

Given the objective of this study, Exploratory Factor Analysis (EFA) is more suitable than the frequently used Principal Components Analysis (PCA). This is because EFA aims at modeling the structure of correlations among the original variables (i.e. adjectives) rather than achieving a simple data reduction (see Fabrigar et al., 1999). The difference between the two techniques lies in the fact that EFA treats each measured variable as a linear combination of one or more common factors and one unique factor while PCA does not differentiate between common and unique variance. Therefore, PCA merely represents the variance of a set of variables in contrast to EFA which targets at explaining the relationships among variables within a set. Thus, EFA was preferred for identifying the underlying perceptual dimensions of timbre.

In order to better account for possible nonlinear relationships between the variables (i.e. adjectives) we applied a simple rank ordering transformation to the data. We then performed the same analytic approach for both original and transformed variables so as to test the hypothesis of potential existing nonlinearities by examining the effect of the transformation on the final solution.

III. ANALYSIS

The quantity estimations on each verbal descriptor and each musical timbre were averaged over the 41 subjects. 66% of the subjects used at least one extra term, thus providing 131 additional verbal descriptors. 35 of these terms were inserted more than once and 26 were used by more than one participant.

The analytic strategy used in order to reduce the large number of variables (30) was structured upon four basic steps. In the initial step, a centroid Hierarchical Cluster Analysis based on squared Euclidean distances was employed in order to reveal the major clusters and outliers among the adjectives. The outliers were the adjectives that could not be grouped with other adjectives as they appeared to have many instances of low inter-correlation coefficients. Such variables were identified and discarded by observation of the dendrograms.

A preliminary factor analysis (FA) with a non-orthogonal *Oblimin rotation*[1] of the extracted factors was then performed within each of the clusters in order to identify its salient adjectives. The original variables featured an extreme positive skewness while the transformed variables were uniformly distributed. Thus, the FA technique used was the Principal Axis Factoring that makes no distributional assumptions. The adjectives with extracted *communalities*[2] < 0.6 were then discarded. This criterion ensures that only the verbal descriptors that are explained adequately by the model for each cluster were maintained.

Subsequently, an inspection of the correlation matrix led to the removal of multicollinear verbal descriptors.

A final FA, again with a non-orthogonal rotation of the factors, applied on this reduced set of salient adjectives resulted in the major factors (i.e. perceptual dimensions). The descriptors featuring communalities < 0.6 were again discarded and the remaining set of descriptors was subjected to a final FA. The final data reduction step used *factor loadings* as a criterion. Factor loadings are the regression coefficients between variables and factors. Their values indicate the relative contribution that a variable makes to a factor and are crucial for the labelling and interpretation of the factors. Only the descriptors with factor loadings > 0.75 were considered significant in this work.

IV. RESULTS

The strategy described in the analysis section was applied to both original and rank transformed data. The application of centroid Hierarchical Cluster Analysis revealed the clusters and the outliers. Figures 1 and 2 show the dendrograms of the original and rank transformed adjectives respectively. The comparison between the two dendrograms shows that the transformation results in a higher organization of the data. This is more evident in the clusters *dirty-messy-rough-harsh-nasal* and *dark-deep-thick-rich-full-dense* which become significantly tighter. Furthermore, the cluster *empty-light-hollow* is dissolved and its members are grouped into the other existing clusters. Cluster *smooth-soft-warm-rounded-dull* also becomes tighter except for *dull* which is more separated and *hollow* which is added. Finally, the largest cluster becomes only slightly looser and with one additional member in the transformed case.

Overall, a dendrogram featuring one less cluster and two others significantly tighter supports that the application of the rank ordering transformation leads to a more compact representation of the data. This fact is a first indication of existing nonlinear relationships between the original variables that are accounted for by the applied transformation.

The data reduction strategy that was described in the analysis section resulted in a 3 factor solution for both the original and the transformed variables. Table 2 shows the percentage of total and factorial variance explained prior the non-orthogonal rotation. It is evident that there is both a small increase of the total explained variance and a significantly higher concentration of the accounted variance (additional 7.6%) in the first two factors for the transformed variables. This suggests the presence of higher correlations between the transformed variables and is also in agreement with the more compact representation of the data that was evident from the centroid Hierarchical Cluster Analysis. Thus, even a simple rank ordering transformation of the verbal attributes seems to account for the presence of nonlinearities and results in a more consistent solution compared to the untransformed variables.

Table 2. Total and factorial variance explained prior the non-orthogonal rotation for the original and rank transformed variables .

	1st Factor	2nd Factor	3rd Factor	Total
Original	42.8%	24.5%	9.8%	77.1%
Rank Transf.	46.4%	28.5%	5.2%	80.1%

Two goodness-of-fit indexes of the model that were produced from the final FA of our data reduction strategy are presented below. The Kaiser-Meyer-Olkin (KMO) criterion[3] can be calculated for individual and multiple variables and varies between 0 and 1. In our analysis KMO was 0.645 for the original and 0.686 for the transformed case both of which are regarded as ‘mediocre’ but acceptable (Hutcheson and Sofroniou, 1999) and the Bartlett's test of sphericity also showed statistical significance[4].

Table 3 shows the correlation coefficients and angles between the non-orthogonally rotated factors for both original and transformed variables.

[1] A non-orthogonal rotation of the initial orthogonal solution that allows factors to be correlated is preferred in order to increase interpretability of the perceptual dimensions.

[2] Communalities measure the percent of variance in a given variable explained by all the factors jointly.

[3] KMO assesses the sample size (i.e. cases/variables) and predicts if data are likely to factor well based on correlation and partial correlation.

[4] Bartlett's test examines the hypothesis that the correlation matrix under study is significantly different from the identity matrix. Significance on this test confirms this hypothesis.

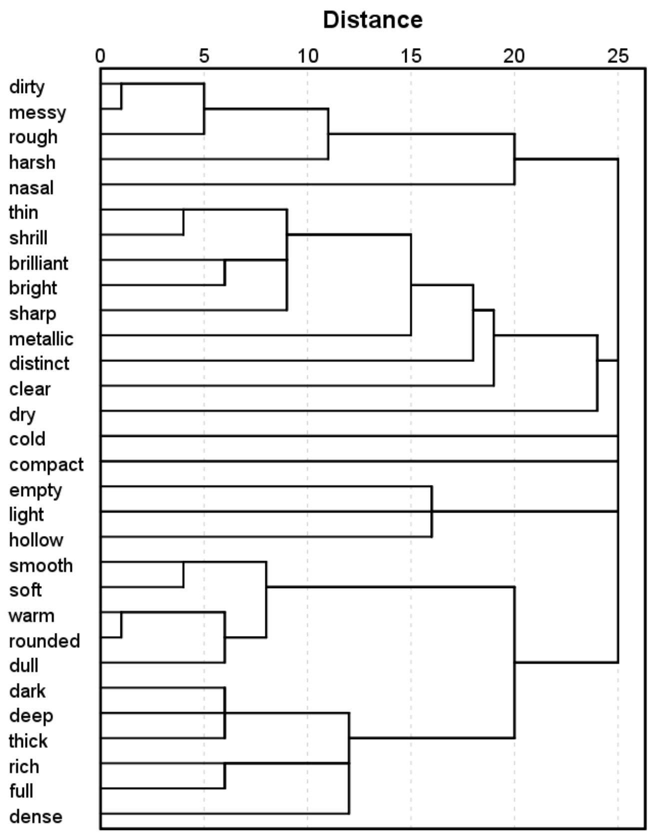

Figure 1. Dendrogram of the original adjectives.

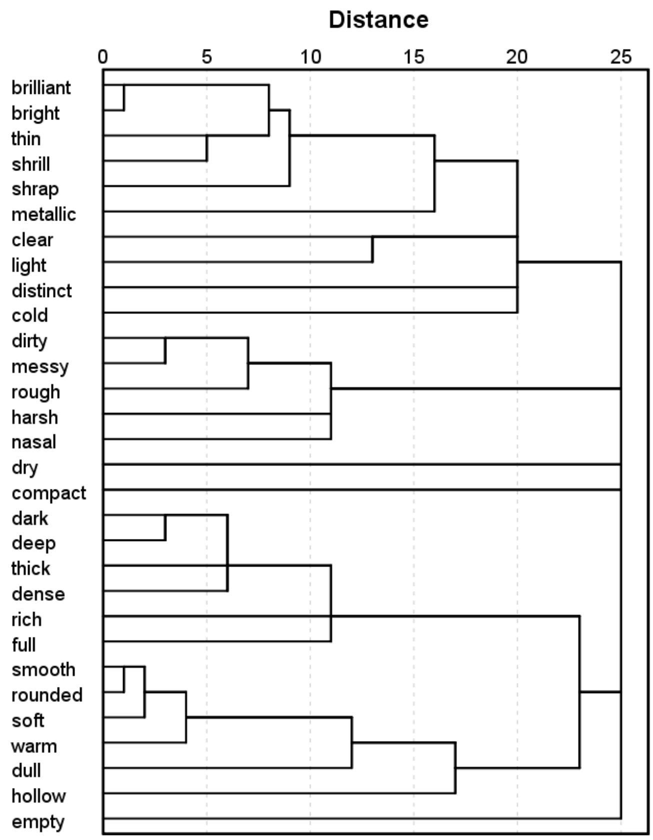

Figure 2. Dendrogram of the transformed adjectives.

The angles are calculated by the $cos^{-1}(corr.coeff)$. Table 3 shows that the rotated solution for the transformed variables exhibits higher inter-dimensional correlations than the solution for the original ones.

Table 3. Inter-dimension correlations and angles.

Correlation coefficient	Orig. variables	Transf. variables
r12	-0.217 (77.5^o)	-0.343 (69.9^o)
r23	0.033 (88.1^o)	0.251 (75.46^o)
r31	0.320 (71.33^o)	0.432 (64.4^o)

Tables 4 and 5 show the pattern matrices for the original and transformed variables respectively. Only the adjectives with factor loadings > 0.75 are depicted and will be used for factor interpretation. For the original variables, the factors could be labeled as Factor 1: *Brilliance/Sharpness*, Factor 2: *Roughness/Dirtiness* and Factor 3: *Thickness/Fullness*. Similarly, for the transformed variables Factor 1: *Lightness/Thinness vs. Thickness/Density*, Factor 2: *Roughness vs. Clarity* and Factor 3: *Brilliance/Sharpness vs. Warmth*.

V. DISCUSSION

As highlighted in the result section, the rank ordering transformation of the variables (semantic descriptors) contributed towards a more compact representation of the data something that is demonstrated by the clearer formulation of clusters in the dendrogram.

Furthermore, the transformed Factor Analysis solution explained a slightly higher percentage of total variance and significantly increased the concentration of the accounted variance in the first two factors. This fact is an indication that a non-metric transformation of perceptual variables can account for existing nonlinear relationships between them and provide a potentially more consistent factor configuration. Future work should consider the application of more elaborate nonlinear transformations on the data in order to achieve an even more robust modeling of nonlinearities.

From examination of tables 4 and 5 it is evident that despite the variations of the exact labeling terms, the two solutions retain a conceptual similarity among the identified dimensions. The three perceptual dimensions could be conceptualized as related to *luminance*, *texture* and *mass*. This finding is of particular interest as it is in agreement with the results of an identical study that we have performed in the Greek language (see Zacharakis et al., 2011). This provides evidence of inter-linguistic similarities regarding musical timbre semantic description. It also seems to support the findings of one of the first studies in the field of musical timbre semantics conducted by Lichte (1941) and recent findings regarding polyphonic timbre verbalization by Alluri and Toiviainen (2010).

It is apparent that especially for the transformed variables case, the borders between percepts are not always absolute. As an example, the *mass* dimension is also represented by terms such as: *dark* which belongs to the conceptual category of *luminance*.

Table 4. Factor loadings of the original variables. Only the loadings > 0.75 are presented.

	Factor		
	1	**2**	**3**
Brilliant	-0.945	…	…
Sharp	-0.814	…	…
Bright	-0.779	…	…
Thick	…	…	0.828
Dense	…	…	0.803
Rough	…	0.964	…
Harsh	…	0.750	…
Dirty	…	0.877	…
Full	…	…	0.794

Table 5. Factor loadings of the transformed variables. Only the loadings > 0.75 are presented.

	Factor		
	1	**2**	**3**
Brilliant	…	…	0.804
Sharp	…	…	0.809
Clear	…	-0.874	…
Warm	…	…	-0.819
Thick	-0.872	…	…
Dense	-0.814	…	…
Thin	0.758	…	…
Rough	…	0.870	…
Harsh	…	0.750	…
Dirty	…	0.787	…
Soft	…	…	-0.753
Light	0.792	…	…
Dark	-0.766	…	…
Deep	-0.751	…	…

Furthermore, the *luminance* dimension is also represented by the terms *sharp* and *soft* which better fit in the *texture* dimension.

As noted previously, the dimensions that resulted from the transformation of the variables are more inter-correlated compared to the original solution. *Luminance* and *mass* are the stronger correlated ones for both cases ($r = 0.320$ and $r = 0.432$). *Texture* and *mass* is also weakly correlated ($r = -0.343$) for the transformed case while totally independent for the original case ($r = 0.033$). Finally, *luminance* and *texture* exhibit an equally weak correlation for both cases ($r = -0.217$ and $r = 0.251$).

As a final remark, it is evident that the non-metric transformation did not affect the qualitative interpretation of the perceptual dimensions. However, the value of this approach lies in the yield of more accurate representation of sound stimuli positions within the identified perceptual timbre space. This is particularly significant in the search for acoustic correlates of the perceptual dimensions.

VI. CONCLUSION

This study not only investigated the latent dimensions of a verbally described perceptual space of musical timbre, but also addressed potential nonlinear relationships between the perceptual variables.

The data of a verbal attribute magnitude estimation (VAME) listening test have been processed by dimension reduction techniques. A metric and a non-metric approach have been carried out and compared. Results showed that a simple rank ordering transformation of the data improved their representation and explained a larger amount of variance with fewer dimensions compared to the untransformed case. This supports the hypothesis of existing nonlinearities among the perceptual variables that have been more efficiently modeled by the non-metric approach.

The three identified perceptual dimensions explained around 80% of the total variance and shared common qualitative characteristics for both metric and non-metric approach. They can be categorized as relating to the description of: *luminance*, *texture* and *mass* of a sound object. This finding is in agreement with previous studies on musical timbre verbalization and can be further exploited for the development of a semantic framework for musical timbre description.

ACKNOWLEDGMENT

The authors would like to thank Dr. Robin Hart, Susan Sturrock and Dr. John Dack who have helped with organizing the listening tests at the Royal College of Music and Middlesex University. We would also like to thank all the participants of the listening tests.

REFERENCES

Alluri, V. & Toiviainen, P. (2010). Exploring perceptual and acoustical correlates of polyphonic timbre. *Music Perception, 27*, 223-242.

Barthet, M., Guillemain, P., Kronland-Martinet, R., & Ystad, S (2010). From clarinet control to timbre perception. *Acta Acustica united with Acustica, 96*, 678-689.

Disley, A. C., & Howard, D. M. (2004).Spectral correlates of timbral semantics relating to the pipe organ. *TMH-QPSR, 46*, 25-40.

Disley, A. C., & Howard, D. M., Hunt A. D. (2006). Timbral description of musical instruments. In M. Baroni, A. R. Addessi, R. Caterina, M. Costa (Eds.), *Proceedings of the 9th International Conference on Music Perception and Cognition (ICMPC 9)* (pp. 61-68). Italy: University of Bologna.

Donnadieu, S. (2007). Mental representation of the timbre of complex sounds. In J. W. Beauchamp (Ed.), *Analysis, synthesis, and perception of musical sounds* (pp. 272-320). Springer Science+Business Media.

Ethington, R., & Punch, B. (1994). Seawave: A system for musical timbre description. *Computer Music Journal, 18*, 30-39.

Fabrigar, L. R., Wegener, D. T., McCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*, 272-299.

Faure, A., McAdams, S., & Nosulenko, V. (1996). Verbal correlates of perceptual dimensions of timbre. In B. Pennycook, E. Costa-Giomi (Eds.), *Proceedings of the 4th International Conference on Music Perception and Cognition (ICMPC 4)* (pp. 79-84). Canada: McGill University, Montreal.

Fritz, C., Blackwell, A., Cross, I., Moore, B., & Woodhouse, J. (2008). Investigating English violin timbre descriptors. In K. Miyazaki, Y. Hiraga, M. Adachi, Y. Nakajima, & M. Tsuzaki (Eds.), *Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC 10)* (pp. 638-639). Japan: Hokkaido University.

Harman, H. H. (1976). *Modern factor analysis*. 3rd edition, Chicago, IL: University of Chicago Press.

Hutcheson, G. & Sofroniou, G. (1999). *The multivariate social scientist.* London: Sage.

Kendall, R. A., & Carterette, E. C. (1993a). Verbal attributes of simultaneous wind instruments timbres: I. von Bismarck's adjectives. *Music Perception, 10*, 445-468.

Kendall, R. A., & Carterette, E. C. (1993b). Perceptual scaling of simultaneous wind instruments timbres: II. Adjectives induced from Piston's "Orchestration". *Music Perception, 10,* 469-502.

Lichte, W. (1941). Attributes of complex tones. *Journal of Experimental Psychology: 28*, 455-480.

Lokki, T., Patynen, J., Kuusinen, A., Vertanen, H., & Tervo, S. (2011). Concert hall acoustic assessment with individually elicited attributes. *Journal of the Acoustical Society of America, 130*, 835-849.

Moravec, O., & Stepanek, J. (2003). Collection of verbal descriptions of musical sound timbre in Czech language. In V. Verfaille (Ed.), *Proceedings of the 7th International Colloquium "ACOUSTICS '03"* (pp. 23-26). Slovakia: Zvolen , Sachticky.

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning*. Urbana, IL: University of Illinois Press.

Pratt, R. & Doak, P. (1976). A subjective rating scale for timbre. *Journal of Sound and Vibration, 45*, 317-328.

Romesburg, S. (2004). *Cluster Analysis for researchers*. North Carolina: Lulu Press.

Stepanek, J. (2006). Musical sound timbre: Verbal descriptions and dimensions. In V. Verfaille (Ed.), *Proceedings of the 9th International Conference on Digital Audio Effects (DAFx-06)* (pp. 121-126). Canada: McGill University, Montreal, Quebec.

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

von Bismarck, G. (1974b). Sharpness as an attribute of the timbre of steady sounds. *Acustica, 30*, 159-192.

von Helmholtz, H.L.F. (1887). *On the sensations of tone as a physiological basis for the theory of music*. 4th edition, translation by A.J. Ellis, New York: Dover (1954).

Wake, S. Asahi, T. (1998). Sound retrieval with intuitive verbal expressions. In *International Conference of Auditory Display (ICAD'98).* U.K.: University of Glasgow.

Woodward, J. A., & Overall, J. E. (1976). Factor analysis of rank-ordered data: An old approach revisited. *Psychological Bulletin, 83*, 864-867.

Zacharakis, A., Pastiadis, K., Papadelis, G., & Reiss, J. D. (2011). An investigation of musical timbre: Uncovering salient semantic descriptors and perceptual dimensions. In A. Klapuri, C. Leider (Eds.), *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 12)* (pp. 807-812). USA: Miami, FL.