

Die physikalischen Korrelate von Instrumentalklangfarben

Saleh Siddiq¹, Christoph Reuter¹, Isabella Czedik-Eysenberg¹, Denis Knauf²

¹ *Institut für Musikwissenschaft, Universität Wien, E-Mail:saleh.siddiq@univie.ac.at*

² *Technische Universität Wien*

Hintergrund

In der Klangfarbenforschung werden häufig Verfahren eingesetzt, die perzeptiven Klangfarbenunterschiede als Entfernungen in virtuellen, n-dimensionalen Räumen darstellen [1][2][3][4][5][6]. Die Ähnlichkeitsverhältnisse der Klangfarben werden in Hörversuchen empirisch ermittelt. Die für die Forscher zu beantwortende Frage ist, welche physikalischen Klangparameter die Hauptachsen (Dimensionen) des virtuellen Raums inhaltlich sinnvoll beschreiben. Diese Verbindung war zumeist das Ergebnis eines „educated guess“, also der (subjektiven) Interpretation des Datenmaterials durch die Forscher.

In einer Vielzahl von Studien wurden zahlreiche Parameter ermittelt, um die Dimensionen dieses Raums zu erklären. Für die ersten beiden Dimensionen haben sich in großer Übereinstimmung eine Skala der spektralen sowie der zeitlichen Hüllkurve als quasi-Standard etabliert [7]. Erstere wird meist als Schärfe- oder Helligkeitsachse bezeichnet und beschreibt die Energieverteilung im Klangspektrum. Letzterer liegen Parameter zugrunde, die das Einschwingverhalten der Klänge beschreiben. Eine eventuell vorhandene dritte Dimension konnte nicht so eindeutig durch eine bestimmte Klangeigenschaft beschrieben werden [7].

Diese bisherigen Ergebnisse sind aus zwei Gründen problematisch: 1. Gemäß der Definition des *American National Standards Institute* (ANSI) ist Klangfarbe dasjenige Attribut, anhand dessen ein Hörer zwei Klänge unterscheiden kann, die bei gleicher Lautheit und Tonhöhe zeitgleich präsentiert werden [8]. Obwohl diese Definition stark kritisiert wurde [9][10][11], führte sie dazu, dass Instrumente meist nur auf einer Tonhöhe verglichen und damit nur durch einen einzigen Ton repräsentiert wurden.

2. In praktisch allen Studien wird nach Unterschieds- bzw. Identitätsmerkmalen der Klangfarben von Musikinstrumenten gefragt. Musikinstrumente können jedoch auf vielfältige Weise angeregt werden und folglich eine große Bandbreite an Klangfarben erzeugen. Instrumente auf einen Ton zu reduzieren hat zur Folge, dass effektiv nicht Musikinstrumente, sondern nur einzelne, von verschiedenen Instrumenten produzierte Klänge erforscht wurden. Forschungsfrage und Methode greifen letztlich nicht in gewünschter Weise ineinander.

Fragestellung

Tatsächlich lassen Studien mit mehr als einem Ton pro Instrument vermuten, dass der virtuelle Raum in diesem Fall durch andere Klangparameter erklärt werden muss [12][13]. Die Frage der hier vorgestellten Studie lautet folglich: Welche physikalischen Klangparameter sind geeignet, die perzeptive Klangverteilung im virtuellen Raum zu erklären?

Darüber hinaus stellt sich die Frage, welche dieser Klangparameter zur Identifikation bzw. Diskrimination von Musikinstrumenten geeignet sind.

Methode

Im Hörversuch wurden die subjektiven Unterschiede der verwendeten Klänge ermittelt. Diese wurden in Unähnlichkeitsmatrizen gespeichert, welche mittels Multidimensionaler Skalierung (MDS) in eine räumliche Darstellung übertragen wurden. Zudem wurden die Klänge mithilfe der Möglichkeiten des Music Information Retrieval (MIR) auf ihre physikalischen Eigenschaften hin analysiert. Dabei wurden die Ausprägungen der verschiedenen Eigenschaften der Klänge (Deskriptoren) in numerischen Werten ausgedrückt. Um Zusammenhänge zwischen den Dimensionen des perzeptiven Raumes und den Deskriptoren offenzulegen, wurden die Korrelationen zwischen den Koordinatenvektoren der Raumdimensionen und den Wertetabellen der Deskriptoren berechnet.

Stimuli

Es wurden fünf Instrumente (Fagott, Flöte, Klarinette, Posaune, Cello) aus dem Orchesterrepertoire europäischer Kunstmusik in jeweils drei Tonhöhen und drei Dynamikstufen (pp, mf, ff), also mit jeweils neun Tönen verglichen. Eine detaillierte Beschreibung der Stimuliauswahl findet sich in Siddiq et al. 2017 [14].

Versuchspersonen

Insgesamt wurden 43 Versuchspersonen (Vpn) im Alter von 13–75 Jahren ($\bar{O} = 32,86$, $SD = 14,93$, weibl. = 20, männl. = 22, andere/k.A. = 1) getestet. Davon waren 32 aktuell bzw. acht ehemals musikalisch aktiv (2–60 Jahre Erfahrung, $\bar{O} = 20,66$, $SD = 14,08$). Zwei Personen gaben an, nie musikalisch aktiv gewesen zu sein.

Hörversuch

Die Vpn sollten die Klangfarbenunterschiede aller Klänge im Paarvergleich subjektiv bewerten. Die Bewertung erfolgte mittels einer siebenstufigen Skala, wobei 1 der geringste Unterschied (größte Ähnlichkeit) und 7 der größte Unterschied war. Detailliertere Angaben zum Ablauf des Hörversuchs finden sich in Siddiq et al. 2017 [14].

Auswertung

Die subjektiven Daten wurden in einer Gesamtmatrix zusammengefasst und in R einer MDS (Package „smacof“, ordinal) unterzogen. Das Ergebnis ist eine 4-dimensionale Konfiguration ($Stress-1$ n. Kruskal = 0,0320). Die Eingangsdaten (Proximitäten) werden durch die MDS-Distanzen offenbar gut abgebildet ($r = 0,97$, $p < 0,00001$). Die MIR-Analyse wurde in MATLAB mittels der MIR-Toolbox [15] durchgeführt.

Ergebnisse

Ein Blick auf den 2D-Plot der Dimensionen 1 und 2 offenbart die Ordnung der Klänge nach Tonhöhe entlang Dimension 1. Entlang von Dimension 2 deuten die Lage der Celloklänge, die Tonhöhenverteilung innerhalb der Holzbläser und Dynamik innerhalb der Posaune eine auf spektraler Energieverteilung basierende Ordnung an.

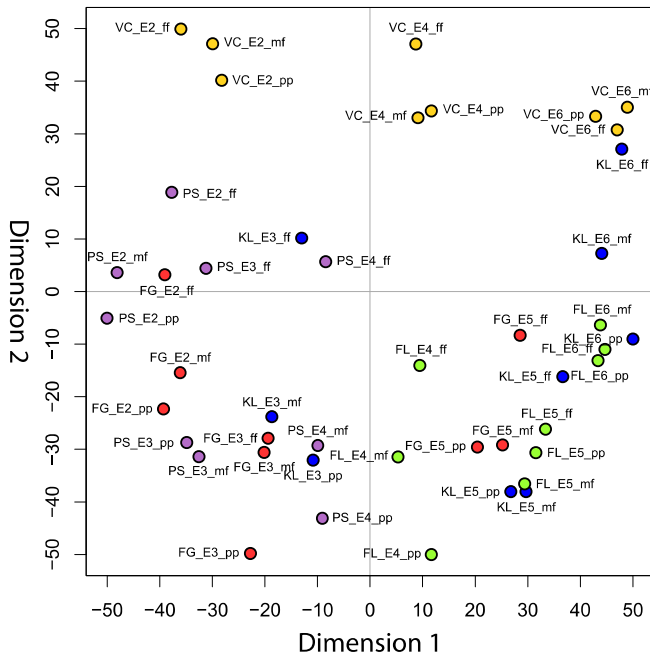


Abbildung 1: Dim1 (X-Achse) und Dim2 (Y-Achse) der MDS-Konfiguration. FG = Fagott (rot), FL = Flöte (grün), KL = Klarinette (blau), PS = Posaune (lila), VC = Cello (gelb). E2 bis E6 = Tonhöhe in wissenschaftl. Notation; pp, mf, ff = Dynamikstufen.

Entsprechend korreliert Dimension 1 hoch mit Deskriptoren der Tonhöhe. Der Zusammenhang ist sogar noch stärker, wenn nicht die Werte, sondern die Ränge betrachtet werden. Diese grundsätzliche Feststellung gilt auch für alle weiteren r-Werte, die jeweils für die Ränge angegeben sind. Dimension 1 korreliert besonders stark und signifikant mit der Grundfrequenz (f_0 ; $r = 0,97$; $p < 0,000001$), Roughness (nach Aures [16]; $r = -0,86$; $p < 0,000001$) und der Zero Crossing Rate ($r = 0,76$; $p < 0,000001$).

Dimension 2 korreliert tatsächlich mit Deskriptoren, welche die Energieverteilung des harmonischen Spektrums beschreiben: Brightness ($r = 0,86$; $p < 0,000001$), Spectral Rolloff ($r = 0,72$; $p < 0,000001$), Spectral Centroid ($r = 0,62$; $p < 0,000001$) und überdies mit den nominellen Dynamikstufen der Klänge ($r = 0,56$; $p < 0,0001$).

Für die Dimensionen 3 und 4 sind die Zusammenhänge weniger deutlich. Dimension 3 korreliert negativ schwach aber signifikant mit dem Deskriptor First Attack Time, der die Einschwingzeit angibt ($r = -0,32$; $p < 0,05$). Für Dimension 4 ergaben sich keine aussagekräftigen Korrelationen.

Betrachtet man die Spektrogramme der Klänge entlang der einzelnen Dimensionen, werden nicht nur die rechnerischen

Korrelationen, sondern mitunter darüber hinausgehende Zusammenhänge sichtbar, die nicht durch geeignete Deskriptoren erfasst wurden (siehe Abb. 3). Auch hier ist Dimension 1 eindeutig als Dimension der Tonhöhe und Dimension 2 als Achse der spektralen Energieverteilung identifizierbar.

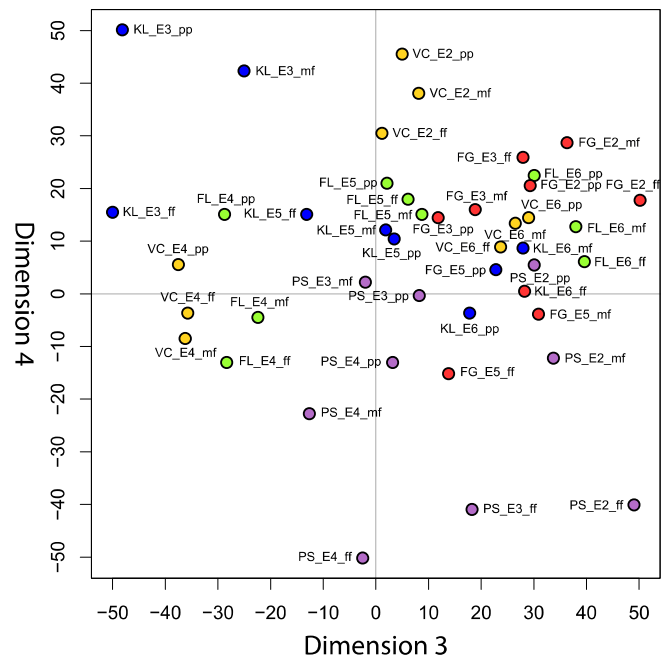


Abbildung 2: Dim3 (X-Achse) und Dim4 (Y-Achse) der MDS-Konfiguration. Label siehe Abb. 1.

Die negative Korrelation von Dimension 3 und Einschwingzeit wird bei Betrachtung der Wellenform deutlich, wo rechts tendenziell steilere Flanken erkennbar sind als links. Der optische Zusammenhang ist deutlicher als der statistische, da die First Attack Time nur den Anstieg bis zum ersten Peak misst und somit Unregelmäßigkeiten im Einschwingen als dessen Ende versteht (vgl. z.B. VC_E4 (alle) oder KL_E5 ff und pp in Abb. 3).

Dimension 4 korreliert mit den verwendeten Deskriptoren nur schwach und insignifikant. Beim gründlichen Studium der Spektrogramme, aber v.a. beim Abhören der Klänge entsteht der Eindruck, dass die begleitenden Nebengeräusche, also v.a. Blas- und Streichgeräusch, entlang Dimension 4 dominanter werden.

Gruppierung von Musikinstrumenten

Die Klänge der Instrumente streuen in allen Dimensionen. Lediglich Dimension 2 korreliert mit der Kategorie „Instrument“ ($r = 0,68$; $p < 0,001$). Entlang der Regressionsgerade der Kategorie Instrument ist die Verteilung der Instrumente noch klarer ($r = 0,74$; $p < 0,001$).

Aus Abbildung 4 ist ersichtlich, dass sich das Cello klar absetzt. Fagott und Flöte haben je einen Ausreißer aus ihrer regionalen Begrenzung. Klarinette und Posaune streuen dagegen stärker. Alles in allem muss man festhalten, dass die gefundenen Parameter nicht geeignet sind, ganze Musikinstrumente zu unterscheiden.

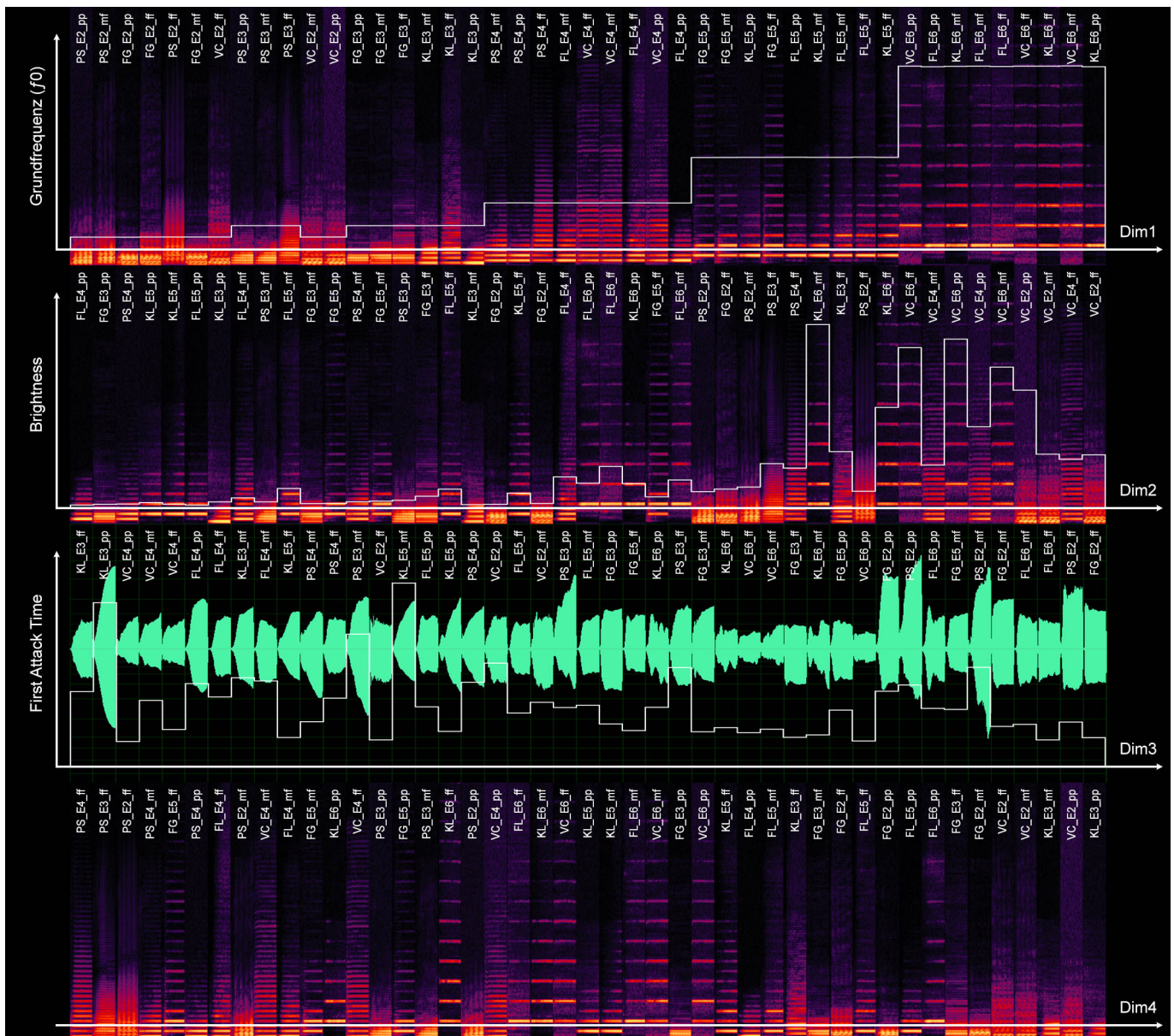


Abbildung 3: Spektrogramme bzw. Wellenform der Klänge entlang der Dimensionen (X-Achse). Darüber Plots der jeweils stärksten Korrelate. Die X-Achse der Plots entspricht den Spektrogrammen, so dass die Darstellungen ineinander passen. Die Y-Achse wurde für eine anschaulichere Darstellung jeweils skaliert. Für Dimension 4 ergaben sich keine aussagekräftigen Korrelate.

Zusammenfassung

Es lässt sich festhalten: Dimension 1 ist eindeutig eine Dimension der Tonhöhe. Dimension 2 korreliert stark mit Deskriptoren des harmonischen Spektrums. Folglich ließen sich die Dimensionen 1 und 2 zur Klangfarbe im engeren Sinn [17] zusammenfassen. Dimension 3 hängt offenbar mit der zeitlichen Klangstruktur, sprich dem Einschwingvorgang zusammen. Obwohl statistisch insignifikant, liegt der Schluss nahe, dass Dimension 4 durch die begleitenden Nebengeräusche bestimmt ist. Damit ließen sich die Dimensionen 3 und 4 der Klangfarbe im weiteren Sinn [17] zuordnen. Demzufolge lassen sich die vorliegenden Ergebnisse als empirischer Hinweis auf Carl Stumpfs Klangfarbenkonzept [17] interpretieren.

Schlussfolgerungen

Folgende Erkenntnisse lassen sich aus den Ergebnissen ableiten: 1. Die Tonhöhe ist ein Faktor der Klangfarbe. Mit steigender Tonhöhe nimmt die Helligkeit zu und die Rauigkeit ab [17]. Zudem ist die Tonhöhenabhängigkeit der Instrumentalregister bekannt.

2. Die (Denk)Kategorie „Instrument“ ist unbedeutend. Letztlich gruppieren sich Klänge physikalisch wie perceptiv gemäß ihrer Eigenschaften. Offensichtlich sind Klänge gleicher Tonhöhe usw. einander ähnlicher als verschieden hohe Töne des gleichen Instruments. Gleiche physikalische/perceptive Eigenschaften deuten umgekehrt auch nicht auf das gleiche Instrument hin.

Basierend auf diesen Erkenntnissen ist eine Präzisierung des Begriffs Klangfarbe wünschenswert. Um die verschiedenen Ebenen Ton, Klang, Instrument gezielt anzusprechen und zu

unterscheiden, soll folgende Sprachregelung vorgeschlagen werden:

Tabelle 1: Begriffsebenen des Konzepts „Klangfarbe“

Ebene	physikalisches Korrelat	Begriff
Ton	Sinuston	Tonfarbe
Klang	Einzelklang (isoliertes Schallereignis)	Klangfarbe
Instrument	Gesamtheit aller Klänge eines Instruments	Instrumentalfarbe

Schon Stumpf griff, zur Beschreibung der „Klangfarbe“ von Tönen, auf *Tonfarbe* zurück [17]. Der Begriff *Instrumentalfarbe* taucht ebenfalls bereits bei Stumpf auf [18]. Der Begriff *Klangfarbe* ließe sich folglich (relativ) missverständnisfrei auf Klänge und Geräusche anwenden.

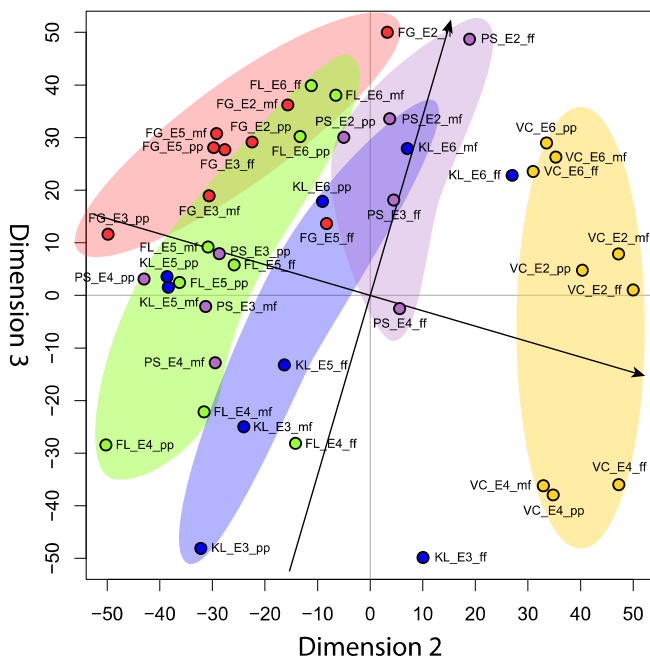


Abbildung 4: Dim2 (X-Achse) und Dim3 (Y-Achse) der MDS-Konfiguration. Die rotierte Dimension 2 entspricht der Regressionsgeraden der externen Skala „Instrument“. Die Farbfelder verdeutlichen die ungefähre Gruppierung der Instrumente. Label siehe Abb. 1.

Danksagungen

Diese Studie wurde gefördert durch den *Jubiläumsfonds der Österreichischen Nationalbank* (OeNB, Projektnr. 16473). Vielen Dank an die *Vienna Symphonic Library* für die Unterstützung in Form von zwei kostenfrei zur Verfügung gestellten *Vienna Super Package* Lizenzen.

Literatur

[1] Wedin, L.; Goude, G.: Dimension analysis of the perception of instrumental timbre. *Scandinavian Journal of Psychology*. 13 (1972), 228–240

[2] Grey, J.M.: An Exploration of Musical Timbre using computer-based Techniques for Analysis, Synthesis and perceptual Scaling. Stanford University, CCRMA Report No. STAN-M-2, 1975

[3] Wessel, D.L.: Timbre Space as a Musical Control Structure. *Computer Music Journal*. 3 (1979), 45–52

[4] Krumhansl C.: Why is Musical Timbre so hard to understand? Nielzen, S.; Olsson, O. (Eds.): *Structure and Perception of electroacoustic Sound and Music*. Elsevier, Amsterdam 1989, 43–53

[5] McAdams, S. et al.: Perceptual Scaling of synthesized Musical Timbres. *Psychological Research* 58 (1995), 177–192

[6] Lakatos, S.: A common Perceptual Space for harmonic and percussive Timbres. *Perception & Psychophysics* 62 (2000), 1426–1439

[7] Caclin, A.: Acoustic correlates of timbre space dimensions. A confirmatory study using synthetic tones. *JASA*. 118 (2005), 471–482

[8] American National Standards Institute: *American Standard. Acoustical Terminology (Including Mechanical Shock and Vibration)*, ASA, New York, 1960

[9] Schouten J.F. in: Plomp, R.: *Timbre as a multidimensional attribute of complex tones*. Plomp, R.; Smoorenburg, G.F. (Eds.). *Frequency analysis and periodicity detection in hearing*. Sijthoff, Leiden 1970, 397–414

[10] Bregman, A.S. *Auditory Scene Analysis. The Perceptual Organization of Sound*, MIT Press, Cambridge, 1990

[11] Patterson, R.D. et al.: *Reviewing the Definition of Timbre as it Pertains to the Perception of Speech and Musical Sounds*. Lopez-Poveda, E. (Eds.): *The Neurophysiological Bases of Auditory Perception*. Springer, New York 2010, 223–233

[12] Marozeau, J. et al.: The dependency of timbre on fundamental frequency. *Journal of the Acoustical Society of America* 114 (2003), 2946–2957

[13] Handel, S.; Erickson, M.L.: *Sound Source Identification: The Possible Role of Timbre Transformations*. *Music Perception*. 21 (2004), 587–610

[14] Siddiq, S. et al.: *Timbre Space reloaded: Tonhöhe und Dynamik als Teil der Klangfarbenempfindung*. *DAGA* 43 (2017), 194–197

[15] Lartillot, O. et al.: *A Matlab Toolbox for Music Information Retrieval*. *Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation E.V.* (2008), 261–268

[16] Sethares, W.A.: *Tuning, Timbre, Spectrum, Scale*. Springer, Berlin 1998

[17] Stumpf, C.: *Tonpsychologie Bd. 2*, Hirzel, Stuttgart, 1890

[18] Stumpf, C.: *Die Sprachlaute. Experimentell-phonetische Untersuchungen. Nebst einem Anhang über Instrumentalklänge*. Springer, Berlin 1926