

Zwischen psychoakustischen Ähnlichkeiten und mentalen Konzepten: Ein Experiment zur Gruppierung von Instrumentalklängen

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

Studien zur Klangfarbenwahrnehmung zeigen, dass allein auf klanglichen Merkmalen basierende Ähnlichkeitsurteile keine Instrumentencluster produzieren, sondern einzig auf klanglicher Ebene sortieren [1][2][3]. Aus akustischer Sicht ist dies nicht verwunderlich, denn obwohl die akustischen Charakteristika von Instrumentenklängen eine direkte Folge der Klangentstehung am Instrument sind, werden sie doch durch spieltechnische Aspekte und andere Begebenheiten mehr oder weniger stark überformt. Empfundene Klangähnlichkeiten sind folglich nicht zwingendermaßen ein Hinweis auf eine Entstehung am gleichen Instrument [3][4]. Jüngere Publikationen [5][6] legen jedoch nahe, dass die kategoriale Zuordnung von Instrumentenklängen nicht allein aufgrund klanglicher Eigenschaften geschieht, sondern durch ein semantisches Konzept von Musikinstrumenten als klangliche Entitäten geleitet werden (bereits 1926 bei [7]). Das wirft die Frage auf, inwieweit die Empfindung der Zusammengehörigkeit von Instrumentalklängen von psychoakustischen Faktoren und inwieweit von semantischen Konzepten beeinflusst wird.

Fragestellung

In dieser Studie soll daher zum einen der Frage nachgegangen werden, inwieweit die Entstehung der perzeptiven Identität von Musikinstrumenten durch psychoakustische Merkmale gebildet bzw. behindert wird. Zum anderen soll festgestellt werden, inwiefern die Präzision des Konzepts „Musikinstrument“ eine erlernte und überdies (musik-)kompetenzabhängige Fähigkeit ist.

Methode

In einer Gruppierungsaufgabe sollten Klänge verschiedener Orchesterinstrumente gemäß ihrer Instrumentenzugehörigkeit sortiert werden. Aus den Daten wurden individuelle Grauwertematrizen und daraus wiederum eine Ähnlichkeitsmatrix erstellt, die als Grundlage für die weitere Auswertung diente. Die Datenstruktur wurde mittels Clusteranalyse und multidimensionaler Skalierung (MDS) sowie deskriptiver Statistik analysiert.

Stimuli

Es wurden sechs Musikinstrumente aus dem Orchesterrepertoire europäischer Kunstmusik mit jeweils sieben Klängen in verschiedenen Tonhöhen untersucht. Die Instrumente waren Fagott (FG), Eb-Klarinette (KL), Querflöte (FL), Tenorposaune (PS), Violoncello (VC) und Waldhorn (HN). Die Tonhöhen wurden nach folgenden Kriterien gewählt: (1.) Die Klänge eines Instruments sollten möglichst selten das gleiche Tonchroma haben, um diesbezügliche Ähnlichkeiten zu vermeiden. (2.) jede Tonhöhe sollte nur einmal vorkommen, um die erwiesenermaßen starken Ähnlichkeitseffekte zu vermeiden [2][8]. (3.) sollten die Intervalle innerhalb eines Instruments möglichst äquidistant sein, um Clustereffekte zu vermeiden, jedoch minimal eine kleine Terz und maximal eine große Sext. Sämtliche Stimuli wurden aus der *Vienna Symphonic Library (VSL)* bezogen.

Versuchspersonen

Aus insgesamt 107 Teilnahmen konnten 41 verwendbare Datensätze gewonnen werden. Von diesen 41 Versuchspersonen (Vpn) im Alter von 19–78 Jahren ($\bar{O} = 42,7$; $SD = 16,9$; w/m/d = 21/16/4) waren 26 Personen aktuell bzw. 7 ehemals musikalisch aktiv (Musiker(innen) = M; 3–50 Jahre Erfahrung; $\bar{O} = 20,1$; $SD = 11,5$). Lediglich 8 Personen gaben an, nie musikalisch aktiv gewesen zu sein (Nichtmusiker(innen) = NM). Der Versuch wurde über ein eigens dafür entwickeltes Interface durchgeführt. Die Art des Abspielgeräts (Kopfhörer, Boxen, Ohrhörer, integrierte Lautsprecher) wurde erhoben, hatte allerdings keinerlei erkennbaren Einfluss auf die Daten.

Hörversuch

Die Vpn waren aufgefordert, die Klänge der untersuchten Instrumente so zu sortieren, dass jeweils die Klänge, die auf dem gleichen Instrument gespielt wurden (z. B. Fagott), zusammen gruppiert wurden. Abgesehen von der expliziten Vorgabe nach Instrument zu gruppieren, war die Zuordnung frei („free sorting task“). Auch zur Anzahl der formierten Gruppen sowie der gruppierten Klänge wurden keinerlei Vorgaben gemacht. Es wurden keinerlei Anhaltspunkte, wie z.B. spezielle Instrumentennamen verwendet; die Gruppierung erfolgte allein anhand auditiver Informationen. Im Anschluss an die Gruppierung sollten die gebildeten Instrumentengruppen frei benannt werden.

Auswertung

Die individuellen Resultate wurden in Kookurrenzmatrizen festgehalten und zunächst einer Clusteranalyse unterzogen um etwaige systematische Schichtungen der Stichprobe festzustellen (z.B: M/NM) und überdies Ausreißer zu identifizieren. Anschließend wurden die Matrizen zu Ähnlichkeitsmatrizen addiert (gesamt, M, NM). Mittels Clusteranalyse (complete-linkage, furthest neighbor) wurden die wahrnehmungsbasierten Cluster ermittelt. Zusätzlich wurde eine MDS (R-Package *smacof*, ordinal, 4D) errechnet. Mittels Audiosignalanalyse (*MiningSuite* 0.10 in Matlab) wurden die akustischen Klangmerkmale extrahiert und mit den Koordinatenvektoren der MDS-Dimensionen in Beziehung gesetzt, um so diejenigen Merkmale zu identifizieren, welche die Gruppierung der Klänge am ehesten beeinflusst haben.

Ergebnisse

Sortierung der Klänge

Insgesamt lässt sich festhalten, dass M im Vergleich zu NM feinere Unterteilungen vorgenommen haben. M haben zwischen 5 und 15 Instrumente angelegt ($\bar{O} = 8,1$, $SD = 2,6$), während NM zwischen 2 und 8 Instrumente ($\bar{O} = 3,6$, $SD = 1,99$) angelegt haben. Dabei ist allerdings zu bedenken, dass nur 7 NM den Test beendet haben. Heatmaps für M bzw. NM zeigen, welche Klänge am ehesten als zusammengehörig bewertet wurden. Implizit geht daraus hervor, welche Klänge am öftesten verwechselt wurden.

Insgesamt zeigen die Heatmaps für M bzw. NM eine hohe Übereinstimmung ($r = .886^{***}$). Beide Gruppen hatten offensichtlich keine Schwierigkeiten, das Cello von den anderen Instrumenten zu unterscheiden. Dagegen wurden die Klänge von Fagott, Horn und Posaune häufig zusammen gruppiert: Horn und Posaune $M = 42,8\%/NM = 63,56\%$, Fagott und Horn $27,13\%/38,19\%$ sowie Fagott und Posaune $28,39\%/39,94\%$.

	Cello	Fagott	Horn	Klarinette	Posaune	Flöte
Cello	77,84	3,79	4,96	6,12	4,96	5,54
	63,39					
Fagott	4,44	3,79	38,19	24,78	39,94	23,62
		42,74				
Horn	4,5	27,13	63,85	18,95	63,56	20,7
			48,02			
Klarinette	8,04	11,1	7,26	64,43	18,66	27,7
				56,42		
Posaune	4,74	28,39	42,8	8,16	68,51	21,87
					48,5	
Flöte	4,44	13,51	10,38	19,45	11,4	73,18
						59,3

Abbildung 1: Heatmap für M (untere Halbmatrix) und NM (obere Halbmatrix). Angaben in Prozent für alle möglichen Kombinationen.

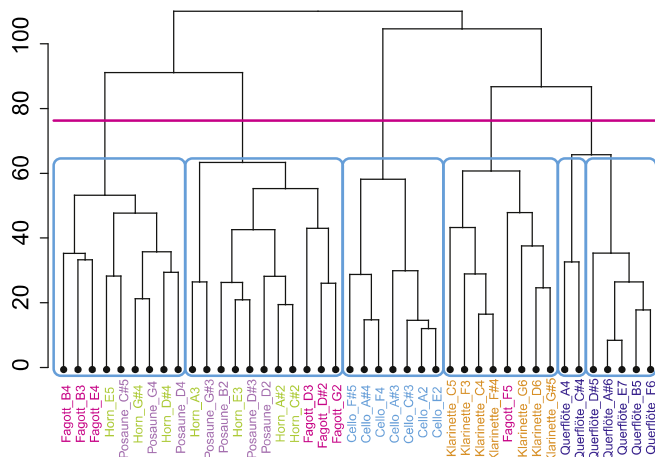


Abbildung 2: Dendrogramm der Clusteranalyse. Linie (Magenta) = errechnete Cluster (5), Rahmen (Cyan) = erzwungene Cluster (6).

Die anhand der Heatmaps bereits anzunehmenden Gruppierungen werden im Dendrogramm der Clusteranalyse erkennbar. Es ist deutlich zu erkennen, dass fünf Cluster gefunden wurden. Da allerdings sechs Instrumente getestet wurden, wurden sechs Cluster erzwungen, wodurch das Querflötencluster sozusagen künstlich zweigeteilt wurde. Die Celloklänge sind ebenfalls exklusiv in einem Cluster vereint. Fagott, Horn und Posaune sind relativ gleichmäßig auf zwei Cluster verteilt, wobei sich der höchste Fagottklang (F5) ins ansonsten ebenfalls „reine“ Klarinettencluster „verirrt“ hat. Die beiden Fagott/Horn/Posaune-

Cluster scheinen sich nach Oktavlage, also Tonhöhe aufzuteilen. In einem Cluster liegen alle Klänge ab B3 (Fagott) aufwärts, im anderen alle Klänge ab A3 (Horn) abwärts. Für Horn und Fagott liegt dies nahe am unteren Ende des ersten Formantbereichs von 300–500 Hz, beim Fagott zudem genau auf einer in der Literatur ziemlich übereinstimmend genannten Registergrenze (zusammengefasst bei [9]).

Die MDS-Konfiguration in vier Dimensionen (Stress-I (Kruskal) = .747) zeigt die im Dendrogramm sichtbaren Cluster sehr klar abgegrenzt auf der von Dimension 1 und Dimension 2 aufgespannten Ebene. Auf der von Dimension 3 und 4 aufgespannten Ebene scheinen die Klänge auf den ersten Blick chaotisch verteilt zu sein. Auffällig ist, dass sich die Querflötenklänge v.a. im I. Quadranten, die Klarinettenklänge dagegen v.a. im IV. Quadranten finden. Besonders das Horn, aber auch Fagott und Posaune sind auf Dimension 4 sehr homogen, während sie auf Dimension 3 sehr heterogen und jeweils in sich tendenziell in absteigender Tonhöhe sortiert sind. Das Cello hingegen ist auf Dimension 4 homogen und auf Dimension 3 gespreizt, wobei es ebenfalls tendenziell in absteigender Tonhöhe sortiert ist.

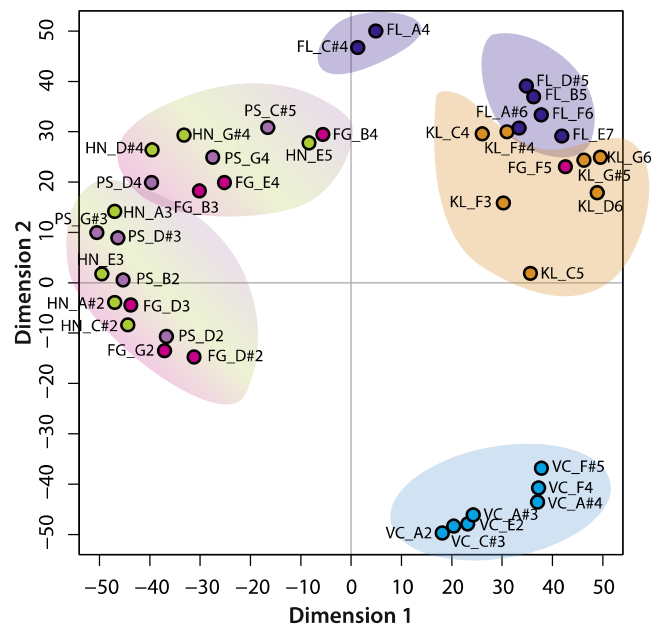


Abbildung 3: MDS-Ebene Dimension 1 und 2. Die mittels Clusteranalyse gefundene Gruppierung scheint sich ausschließlich auf diese beiden Dimensionen zu beziehen.

Dimension 1 korreliert hoch mit Deskriptoren der spektralen Energieverteilung wie Spectral Skewness ($r = -.892^{***}$), Brightness ($r = .764^{***}$) und Spectral Centroid ($r = .714^{***}$). Ebenso zeigt sich ein moderater Zusammenhang mit der Tonhöhe ($r = .590^{***}$). Dimension 2 korreliert besonders mit Deskriptoren geräuschhafter Spektralanteile wie der Harmonic/Percussive Ratio ($r = .612^{***}$), Roughness (n. Vassilakis, $r = -.561^{***}$) und der Zero Crossing Rate ($r = .508^{***}$). Dimension 3 korreliert ebenfalls mit Deskriptoren, die geräuschhafte Spektralanteile beschreiben: Inharmonicity ($r = .492^{***}$), Spectral Entropy ($r = .440^{**}$), Roughness (n. Sethares, $r = .483^{**}$) sowie Roughness (n. Vassilakis, $r = .434^{**}$). Dimension 4 korreliert v.a. mit Faktoren der Lautheit bzw. Lautstärke wie Loudness in Sone ($r = -.493^{***}$) sowie dem RMS ($r = -.308^{*}$) und überdies ebenfalls mit der Tonhöhe ($r = .341^{*}$).

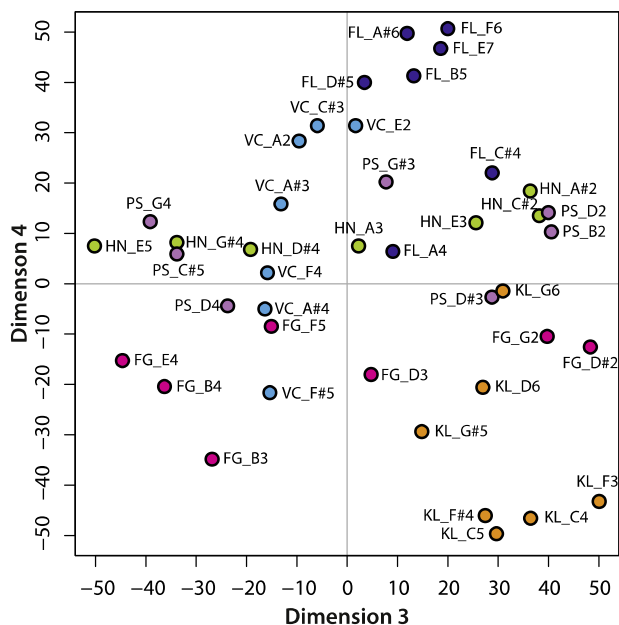


Abbildung 4: MDS-Ebene Dimension 3 und 4. Es ist praktisch keinerlei instrumentenabhängige Verteilung der Klänge erkennbar.

Instrumentenbezeichnungen

Von den wenigen NM haben sich nur drei zur Benennung ihrer angelegten Instrumente entschlossen. In einem dieser Fälle wurden die drei angelegten Instrumente als „Schnarrend“, „Plastik“ und „Klar“ bezeichnet. In beiden anderen Fällen wurden, evtl. mit Ausnahme einer „Mundharmonika“, übliche und zu erwartende Instrumentenbezeichnungen gewählt. Insgesamt (M und NM) kann man die Bezeichnungen in folgende vier Kategorien einteilen: (1.) Instrumentennamen, darunter erwartbare wie „Querflöte“ oder „Cello“, übergenaue wie „Altflöte“ oder „Kontrafagott“ sowie weniger passende wie „Akkordeon“ oder „Klavier“. (2.) Instrumententypen wie „Holzbläser“ oder „Rohrblatt“ sowie ausgefallenerer wie „Pfeife“ oder „Pfeiffelchen“. (3.) Analogiebildungen/Beschreibungen wie „gedämpft“, „hauchig/floetenhaft“, „penetrant“, „Schnarrend“ oder „entenhaft“. (4.) Schwer zu interpretierende Bezeichnungen wie „bleck“ (Vermutlich ein Tippfehler in „blech“), „hb“ (möglicherweise abgekürzt für „Holzbläser“) oder aber „Fluff“, „Plastik“ und „undefinierbar“. Wie das folgende Diagramm zeigt, sind dies aber eher Randerscheinungen. Die häufigsten Bezeichnungen sind akkurat:

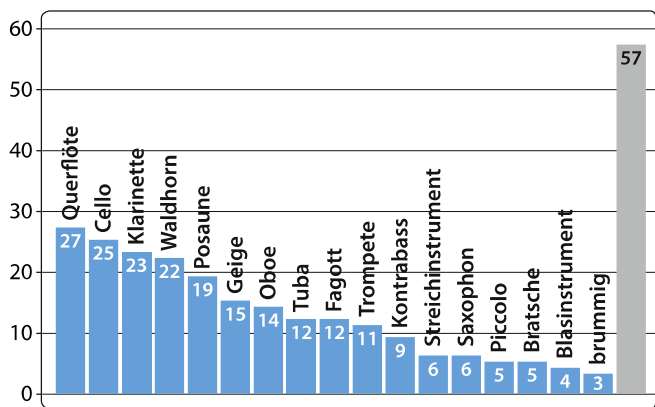


Abbildung 5: Die häufigsten Instrumentenbezeichnungen

Die Bezeichnungen, unter denen die Klänge der tatsächlich getesteten Instrumente meistens einsortiert wurden, sind insgesamt ziemlich genau, in der häufigsten Bezeichnung sogar, abgesehen vom Horn, durchgehend korrekt. Auffällig ist auch hier, dass Cello, Klarinette und Querflöte mit größerer Sicherheit korrekt benannt wurden als Fagott und Blechbläser. Hier war stets das Horn die zweithäufigste Wahl.

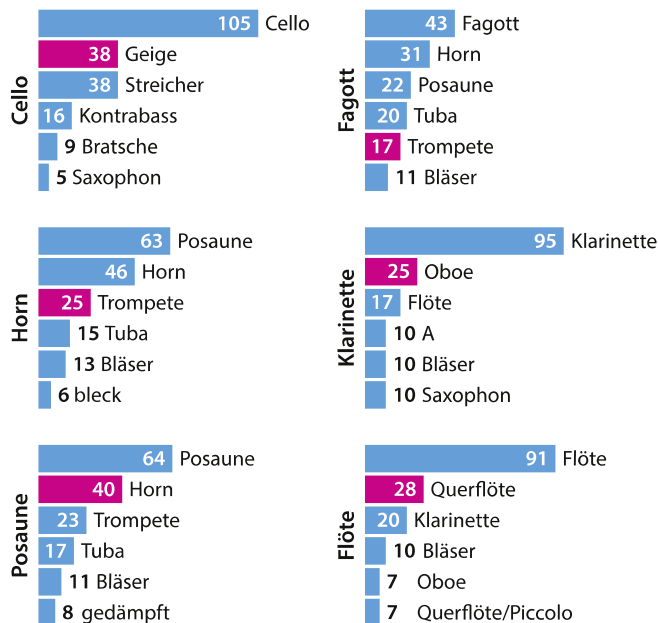


Abbildung 6: Die jeweils sechs häufigsten Bezeichnungen für die Klänge der getesteten Instrumente. Magenta: jeweils der Wert, unterhalb dem die Summe aller Bezeichnungen kleiner ist als inkl. und darüber.

Überträgt man diese Verteilungen auf die errechneten Cluster, zeigt sich folgende Verteilung: Die Trennung von Streichern und Bläsern ist eindeutig. Das Cello-Cluster wurde mit Abstand am häufigsten als „Cello“ (95 Mal) und am zweithäufigsten als „Geige“ (16) bezeichnet (38,7 % aller möglichen Bezeichnungen). Das Klarinettencluster mit dem Fagottklang wurde ebenfalls überwiegend richtig als Klarinette (79) bzw. seltener als Flöte (13) bezeichnet (28,1 %). Fasst man die beiden rechnerischen Querflötencluster zu einem zusammen, überwiegen mit Flöte (55) und Querflöte (23) insgesamt ebenfalls die korrekten Bezeichnungen. Die beiden Cluster, auf die sich Horn, Posaune sowie Fagott aufteilen, sind uneindeutiger benannt. Im höheren (B3 aufwärts) sind Horn (54) und Trompete (30) die häufigsten Bezeichnungen (22,8 %) und im tieferen Fall (A3 abwärts) Posaune (80) und Tuba (42) mit zusammen 27,1 %.

Fazit

Zusammenfassend lässt sich festhalten, dass die aus bekannten, auf psychoakustischen Ähnlichkeiten beruhenden Hörexperimenten gewohnten Zuordnungsschwierigkeiten nicht erkennbar sind. Die urteilsbasierten Cluster bilden die verglichenen Musikinstrumente insgesamt überraschend gut ab. Da die Versuchsteilnehmer(innen) in dieser Studie explizit zu Ähnlichkeitsurteilen vor dem Hintergrund der Instrumentenzugehörigkeit aufgerufen waren, liegt der Schluss nahe, dass vorhandene, erlernte mentale Konzepte von Musikinstrumenten, sofern aktiviert, die Bewertung von Instrumentalklängen beein-

flussen bzw. dass die Identität von Musikinstrumenten nicht allein auf psychoakustischen Klangeigenschaften basiert.

Dies könnte ein weiterer Hinweis darauf sein, dass der Begriff „Klangfarbe“ in seiner aktuellen Verwendung ungeeignet ist, da Musikinstrumente und folglich auch Instrumentalfarben mehr als die Summe ihrer Einzelklänge zu sein scheinen. Ein Vorschlag, den semantischen Komplex hinter dem Begriff Klangfarbe zu entzerren und den wahrnehmungspsychologischen Gegebenheiten auch terminologisch Rechnung zu tragen, wurde bereits 2018 an gleicher Stelle vorgestellt [3]. Diesem zufolge wäre der Begriff „Klangfarbe“ auf die Ebene einzelner Klänge beschränkt, während für einzelne Teiltöne bzw. Musikinstrumente die Begriffe „Tonfarbe“ bzw. „Instrumentalfarbe“ zur Verfügung stünden (für eine detaillierte Aufstellung vgl. [3]).

Der zu erwartende Einfluss musikalischer Kompetenz auf die Genauigkeit der Instrumentenerfassung konnte, aufgrund der niedrigen Nichtmusiker(innen)zahl im Rahmen der aktuellen Studie nicht erschöpfend untersucht werden.

Ausblick

Klarere Aussagen zum Einfluss musikalischer Kompetenz mit einer ausreichend großen Teilnehmer(innen)zahl von beiden Enden des musikalischen Kompetenzspektrums ist einer der Anknüpfungspunkte für weitere Überlegungen in diese Richtung. Ein weiterer Ansatz könnte es sein, mit gezielten wahrnehmungspsychologischen und/oder Kognitionsaufgaben zu ermitteln, ob das „mentale Konzept“ von Musikinstrumenten auf Ebene der Mustererkennung bzw. Gestaltbildung entsteht oder ob höhere kognitive Prozesse im Spiel sind. Letzteres wäre möglicherweise mit entsprechend angepassten kontrastiven Aufmerksamkeitsstests wie den bekannten nach Stroop benannten Tests möglich [10].

Literatur

- [1] Marozeau, J.; Cheveigné, A. de; McAdams, S.; & Winsberg, S.: The dependency of timbre on fundamental frequency. *JASA* 114(5) (2003), 2946–2957.
- [2] Handel, S. & Erickson, M. L.: Sound Source Identification: The possible role of timbre transformations. *Music Perception* 21/4 (2004), 587–610.
- [3] Siddiq, S.; Reuter, C.; Czedik-Eysenberg, I.; & Knauf, D.: Die physikalischen Korrelate von Instrumentalklangfarben. *Fortschritte der Akustik – DAGA 2018, München* (2018), 1695–1698.
- [4] Deutsch, W.; Rösing, H.; Födermayr, F.: Klangfarbe. *MGG2, Sachteil, Bd. 5 (Kas–Mein)*. Kassel, (1996) Sp. 138–170.
- [5] Siedenburg, K.; Jones-Mollerup, K.; & McAdams, S.: Acoustic and Categorical Dissimilarity of Musical Timbre: Evidence from Asymmetries Between Acoustic and Chimeric Sounds. *Frontiers in Psychology* 6/1977 (2016).
- [6] Siddiq, S.; Reuter, C.; Czedik-Eysenberg, I.; & Knauf, D.: Towards the physical correlates of musical timbre(s). *Proceedings of the ICMPC15/ESCOM10, Graz* (2018), 411–415.
- [7] Stumpf, C. Die Sprachlaute. Experimentell-phonetische Untersuchungen nebst einem Anhang über Instrumentalklänge. Berlin, Springer, 1926.
- [8] Siddiq, S.; Reuter, C.; Czedik-Eysenberg, I.; & Knauf, D.: Timbre Space reloaded: Tonhöhe und Dynamik als Teil der Klangfarbenempfindung. *Fortschritte der Akustik – DAGA 2017, Kiel* (2017), 194–197.
- [9] Reuter, C. Klangfarbe und Instrumentation. *Geschichte – Ursachen – Wirkung*. Frankfurt/M, Lang, 2002.
- [10] Stroop, J. R.: Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 18/6 (1935), 643–662.