

Die Farbe macht die Musik

Akustische und perzeptuelle Grundlagen der Klangfarbenwahrnehmung

Kai Siedenburg

Die Klangfarbe verleiht einzelnen Klangereignissen charakteristische Qualitäten und liefert Hinweise auf die Klangquelle und -erzeugung. So erlauben klangfarbliche Merkmale etwa Instrumente im Mix eines Orchesterklangs heraus zu hören oder die Stimme einer bekannten Person zu erkennen. Helmholtz's klassische Konzeption der Tonempfindungen legte den Grundstein der Klangfarbenforschung und zeigte, dass die relative Stärke der Partialtöne für die Klangfarbe wesentlich ist. Moderne Arbeiten haben den Begriff um viele weitere Facetten ergänzt. So wissen wir heute, dass auch temporale und spektrale Merkmale in die Klangfarbenwahrnehmung hineinspielen. Dieser Artikel bietet einen Überblick zu klassischen Ergebnissen und neueren Entwicklungen der Klangfarbenforschung. Nach einer Diskussion der Definition von Klangfarbe werden insbesondere semantische Attribute der Klangfarbe, die Methode der Klangfarben-Räume und die akustische Merkmalsextraktion vorgestellt. Eingehend wird der zentrale Aspekt der klangfarblichen Helligkeitswahrnehmung und ihre Wechselwirkung mit Tonhöhe und Lautheit diskutiert. Letztlich wird eine auditorische Illusion auf Basis von klangfarblicher Helligkeit präsentiert.

Einleitung

Was wäre die Musik ohne den Kontrast von Klängen verschiedener Instrumente? Man könnte wohl behaupten, Symphonieorchester, Big Bands und die meisten Formen von Musikensembles verdanken ihre Existenz der menschlichen Faszination für Klangfarben. Gleiches gilt für Instrumentenbauer, Schöpfer von Synthesizern sowie verschiedensten anderen Mitteln der elektronischen und computerbasierten Klangerzeugung. Die Klangfarbe als Aspekt der auditorischen Wahrnehmung ist also von zentraler Bedeutung in unserer Musikkultur. Dies ist keineswegs ein neuzeitliches Phänomen. Die berühmten Ausgrabungen von frühen Musikinstrumenten des Tübinger Archäologen Nicholas Conard aus Höhlen auf der schwäbischen Alb [1], bei denen aus Vogelknochen und Mammutelfenbein hergestellte Flöten gefunden wurden, zeigen, dass der Mensch schon vor mehr als 35.000 Jahren eine Faszination für Klang und klangliche Innovation hegte. Man stelle sich vor: draußen herrscht Eiszeit, drinnen wird den Klängen der neuesten Flöten gelauscht. Die Klangfarbe steht insgesamt für eine Sammlung

Acoustical and perceptual foundations of timbre

Timbre lends individual sound events a characteristic quality and provides information on the sound source and sound excitation. For example, timbral cues allow instruments to be heard in the tangle of an orchestral sound or to identify the voice of a familiar person. Helmholtz's classic conception of tone perception laid the foundation for timbre research and showed that the relative strength of the partial tones is essential for timbre. Modern works have added many other facets to the term, so that today we know that temporal and spectral features also play a role in timbre perception. This article provides an overview of classical results and recent developments in research on timbre. After a brief discussion of the definition of timbre, the article reviews semantic attributes of timbre, the method of timbre spaces, as well as acoustic feature extraction for describing timbral properties. The central aspect of timbral brightness perception and its interactions with pitch and loudness are discussed in detail. Finally, an auditory illusion based on timbral brightness is presented.

von Attributen der auditorischen Wahrnehmung, die einzelnen Klangereignissen neben der Tonhöhe und Lautheit eine charakteristische Qualität verleihen und Hinweise auf die Klangerzeugung und -quelle liefern. Klangfarbe erlaubt es uns, den Klang einer Flöte von dem einer Stimme zu unterscheiden, Klänge als warm, rau oder schrill zu empfinden oder die Stimme einer Freundin in einem lauten Restaurant wieder zu erkennen. Die Klangfarbe vieler akustischer Musikinstrumente variiert mit dem (Tonhöhen-)Register und der Spielstärke (d.h. der musikalischen Dynamik), was bedeutet, dass die meisten akustischen Klangerzeuger eine Vielzahl verschiedener Klangfarben hervorbringen können. Die wichtigsten der Klangfarbe zugrunde liegenden akustischen Merkmale sind anhand der Form der spektrotemporale Hüllkurve (grober Energieverlauf über Zeit und Frequenz) eines akustischen Signals ermittelbar. Dies beinhaltet Merkmale wie den spektralen Schwerpunkt (spectral centroid), der als eine Art Erwartungswert des linearen Amplitudenspektrums berechnet werden kann und die Helligkeit eines Klanges beeinflusst, die Dauer des Einschwing-

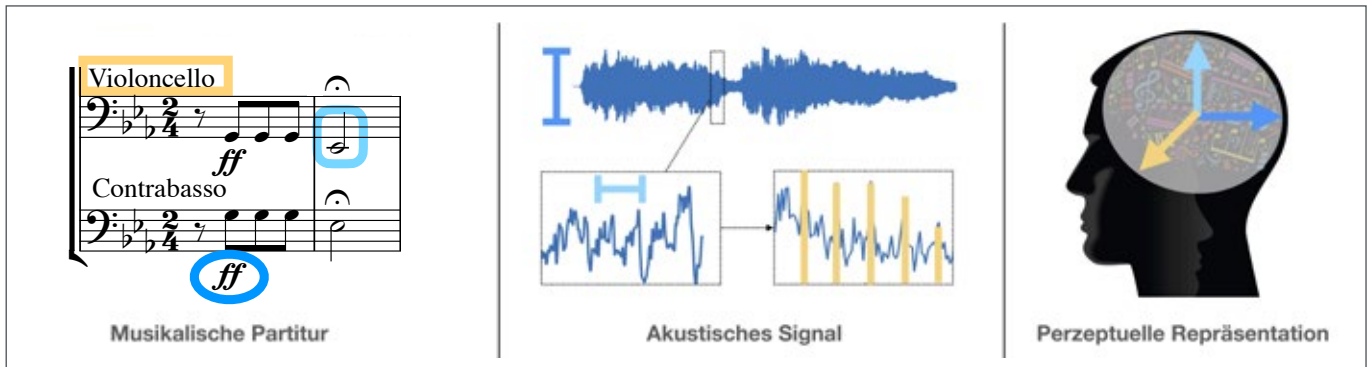


Abb. 1: Die musikalischen Dimensionen Instrumentation, Tonhöhe und Dynamik beeinflussen maßgeblich die akustischen Eigenschaften des Spektrums, der Periode und der Amplitude mit den entsprechenden Wahrnehmungsdimensionen der Klangfarbe, Tonhöhe und Lautstärke.

vorgangs (attack time), die die Impulsivität eines Klanges abbildet oder feinere zeitliche oder zeitlich-spektrale Modulationen, die die Rauigkeit eines Klanges bestimmen. Im weiteren Verlauf des Artikels werden verschiedene für die Klangfarbenwahrnehmung relevante akustische Merkmale vorgestellt und diskutiert.

Zum Begriff der Klangfarbe

Hörforscher haben sich lange schwer damit getan, den Begriff der Klangfarbe allgemeingültig zu definieren (in diesem Artikel verwende ich den Begriff Klangfarbe synonym mit dem englischen timbre). Schon Hermann von Helmholtz sieht im Vergleich zu anderen auditorischen Wahrnehmungsdimensionen insbesondere bei der Klangfarbe Erklärungsbedarf: *“Was wir unter Stärke des Tons und unter Tonhöhe verstehen, brauche ich nicht zu erklären. Unter Klangfarbe verstehen wir diejenige Eigenthümlichkeit, wodurch sich der Klang einer Violine von dem einer Flöte, oder Clarinette, oder menschlichen Stimme unterscheidet, wenn alle dieselbe Note in derselben Tonhöhe hervorbringen.”* (S. 20) Diese trotz allem recht eingängige Definition mag als Vorgänger der berühmt berüchtigten Definition des American National Standards Institute gesehen werden: *“Timbre. That attribute of auditory sensation which enables a listener to judge that two non-identical sounds, similarly presented and having the same loudness and pitch, are dissimilar [sic]. NOTE-Timbre depends primarily upon the frequency spectrum, although it also depends upon the sound pressure and the temporal characteristics of the sound”* [2, S. 35]. In seinem umfassenden Buch zur auditorischen Szenenanalyse kritisierte Albert Bregman diese Definition heftig [3]: *“This is, of course, no definition at all. [...] The problem with timbre is that it is the name for an ill-defined wastebasket category. [...] I think the definition ... should be this: ‘We do not know how to define timbre, but it is not loudness and it is not pitch.’ [...] What we need is a better vocabulary concerning timbre.”* (S. 92–93) Kommentare wie diese ließen viele Forscher im

Zweifel, ob der Begriff überhaupt zu etwas gut ist.

Mit einem konstruktiven Mindset lässt sich jedoch feststellen, dass der Begriff der Klangfarbe keineswegs Abfalleimer, sondern vielmehr ein Sammelbegriff ist, in dem verschiedene Funktionen der auditorischen Wahrnehmung zusammengefasst sind. Sicherlich bedarf es weiterer begrifflicher Unterscheidungen, um die einzelnen Aspekte des Sammelbegriffs genauer zu spezifizieren [4]. Die Klangfarbe kann Schallereignisse qualifizieren und identifizieren. Das bedeutet, die Klangfarbe liefert Cues, um auditorischen Ereignissen eine bestimmte Qualität zu verleihen (die raue Stimme, der fette Bass). Zudem liefert die Klangfarbe Hinweise zur Identifikation oder Spezifikation der Identität einer Klangquelle und der Art und Weise der Klangerzeugung (die Geige, die gezupft wurde). Hierbei ist festzuhalten, dass (ähnlich wie bei der Tonhöhenwahrnehmung) Klangfarbe nur für verschmolzene (d. h. als zusammen gehörig wahrgenommene) auditorische Ereignisse ermittelt wird. Keineswegs sollte man also etwa von der Klangfarbe der Beatles sprechen; der sogenannte Sound einer bestimmten Band oder eines musikalischen Genres entsteht gerade durch das Zusammenspiel und die Konfiguration der Klangfarben perzeptuell separater (aber nicht notwendigerweise zeitlich separater) Klangereignisse.

Ein weiterer wichtiger Punkt ist, dass sich im Falle vieler Musikinstrumente die Klangfarbe als Funktion von (Tonhöhen-)Register und Spielstärke systematisch verändert. Beispielsweise hat eine Klarinette in ihren verschiedenen Registern sehr verschiedene klangfarbliche Ausprägungen. Es macht also wenig Sinn, von der „einen“ Klangfarbe der Klarinette zu sprechen, denn es sind in Wirklichkeit viele. Ähnliches gilt für Klangerzeuger ohne klar ausgeprägte Tonhöhen, wie z. B. Trommeln oder Gongs. Selbst aber für synthetisch erzeugte Klänge (somit strikt kontrolliert) wurden bereits zahlreiche perzeptuelle Wechselwirkungen zwischen den Attributen Tonhöhe, Lautstärke und Klangfarbe festgestellt. Diese

akustischen und perzeptuellen Wechselwirkungen werden im weiteren Verlauf des Artikels unter die Lupe genommen. Als Vorgriff sei gesagt, diese Wahrnehmungsdimensionen sind natürlich nicht voneinander unabhängig, sondern bilden ein zusammenhängendes Inventar.

Zur Definitionsfrage ließe sich schließlich noch die Perspektive von Stephen McAdams heranziehen [5], der schon im Titel der bisher größten Konferenz zur Klangfarbe im Jahr 2018, *Timbre is a many-splendoured thing*, die Vielseitigkeit des Phänomens betont. Die Frage, was Klangfarbe nun im Kern sei, so McAdams, könnte aber auch in eine Sackgasse führen. Produktiver sei die Frage, was man mit Klangfarbe in Musik und Forschung anstellen kann. Dies könnte man unterstreichen, hat doch der Begriff der Klangfarbe in den letzten Jahren für ein dynamisches Forschungsgeschehen Pate gestanden, das viele akademische Disziplinen miteinander verbunden hat, wie z. B. die Akustik, Signalverarbeitung, Musikinformatik, Musikpsychologie und auditorische Neurowissenschaften.

Themen des Artikels

In diesem Übersichtsartikel werden zuerst einige klassische Grundpfeiler der Klangfarbenforschung dargestellt: Die Helmholtzsche Konzeption der Klangfarbe, grundlegende Arbeiten zu semantischen Attributen für Klangfarben sowie sogenannte Klangfarben-Räume. Daran anknüpfend werden neuere akustische Modelle der Klangfarbenwahrnehmung diskutiert. Akustische Merkmale oder Deskriptoren wie z. B. der spektrale Schwerpunkt werden in diesem Abschnitt im Detail erklärt. Letztlich wird auf die Frage eingegangen, inwiefern ein wichtiger Aspekt der Klangfarbe, nämlich die auditorische Helligkeit, als unabhängig und andersartig von der Tonhöhenwahrnehmung betrachtet werden kann. Dieser Überblick erhebt keineswegs einen Anspruch auf Vollständigkeit; interessierte Leser mögen hierfür weitere Quellen konsultieren [6–8].

Grundpfeiler der Klangfarbenforschung

Klangfarbe im Helmholtzschen Sinne

1863 veröffentlichte Hermann von Helmholtz die erste Ausgabe seines Buches *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik* [9]. Die Arbeit wird heute als der einflussreichste Beitrag zur Hörforschung des 19. Jahrhunderts betrachtet. Helmholtzs wichtigstes konzeptionelles Werkzeug war Fouriers Theorem, das bekanntlich postuliert, dass jede periodische Funktion als unendliche Reihe trigonometrischer Funktionen dargestellt werden kann. Ohm und Helmholtz wandten das Theorem auf die Beschrei-

bung von Schall an und demonstrierten damit seine Nützlichkeit für akustische Probleme [10]. Hier implizierte das Fouriersche Theorem, dass die unendliche Komplexität von Schwingungsbewegungen, die Klängen innewohnen, auf eine endliche Anzahl von Parametern reduziert wurde, nämlich die Amplituden und Phasen von trigonometrischer Funktionen, die den Partialtönen entsprechen.

Durch Versuche in Klangsynthese und Physiologie kam Helmholtz zu dem Schluss, dass das Fouriersche Theorem die physikalische und physiologische Realität exakt beschreibt. Er verwendete hierfür z. B. gestimmte Resonatoren, um Partialtöne aus einem Gesamtklang herauszufiltern und zu verstärken. Damit zeigte er, dass Partialtöne nicht nur mathematische Fiktion, sondern physikalische Realität waren, die manipuliert und erlebt werden konnten. In Bezug auf die Physiologie stellte Helmholtz fest, dass es verschiedene Teile des Ohrs geben muss, die durch Töne unterschiedlicher Frequenz in Schwingung versetzt werden. Damit begründete er die einflussreiche Idee des Ohrs als Mechanismus zur Frequenzanalyse [11]. Bezüglich der Klangfarbe postulierte Helmholtz, dass die Qualität des musikalischen Teils eines zusammengesetzten Tons ausschließlich von der Anzahl und der relativen Stärke seiner Teiltöne abhängt und in keiner Weise von deren Phasendifferenz (oder anderen zeitlichen Merkmalen). Nach Helmholtz ist die Klangfarbe also durch die Form des Spektrums eines Tons festgelegt. Diese Perspektive war für seine Zeit bahnbrechend, da dadurch dem Klangunterschied zwischen verschiedenen Instrumenten eine konkrete akustische Basis gegeben wurde.

Hier ist festzuhalten, wie eng Helmholtz sein Studienobjekt definiert hatte, in dem er sich auf den von ihm sogenannten „musikalischen Teil eines Tons“ beschränkte. Diesen definierte er als das, was heute wohl als steady-state eines Tons bezeichnet werden würde, d. h. als eine stationäre Phase des Tons, in der wenige Änderungen auftreten. Durch diese unrealistische Annahme völlig stationärer Klänge war sein Begriff der Klangfarbe in der Tat eine starke Vereinfachung dessen, was heute als Klangfarbe verstanden wird. Offensichtlich wurden zeitliche Aspekte der Klangfarbe wie z. B. der zeitliche Verlauf oder Einschwing-Transienten bei diesem Ansatz nicht berücksichtigt. Doch war sich Helmholtz dieser Tatsache durchaus bewusst: *“Wenn wir im Folgenden von musikalischer Klangfarbe reden, sehen wir zunächst von diesen Eigenthümlichkeiten des Anfangs und Endes ab, und berücksichtigen nur die Eigenthümlichkeiten des gleichmäßig andauernden Klanges.”* (S. 116) Das bedeutet, dass Klangfarbe im Helmholtzschen Sinne nur ein kleiner Teil dessen ist, was heute unter dem Phänomen verstanden wird.

Semantische Attribute von Klangfarben

Kein anderer Aspekt eines Tons regt solch reichhaltige semantische Assoziationen an wie die Klangfarbe. Schon Helmholtz war an den hier zugrundeliegenden akustischen Merkmalen interessiert: *“Es stellen sich dabei gewisse allgemeine Regeln heraus für diejenigen Anordnungen der Obertöne, welche den in der Sprache als weich, scharf, schmetternd, leer, voll oder reich, dumpf, hell, usw. unterschiedenen Arten der Klangfarbe entsprechen.”* (S. 118) Diese metaphorischen Beschreibungen mögen nicht entscheidend für die Wahrnehmung der Klangfarbe sein – man kann Klangfarben wahrnehmen, erkennen oder sich vorstellen ohne sie mit Worten beschreiben zu müssen. Jedoch spielen diese Begriffe sicher eine wichtige Rolle in der Phänomenologie und Ästhetik der Klangfarben. Außerdem könnte die Art und Weise, wie über Klangfarbe gesprochen wird, weitere Informationen über die Wahrnehmungsdimensionen der Klangfarbe offenlegen.

Systematische Forschung zur Klangfarbe und ihren semantischen Assoziationen entstand in den 1970er Jahren. In empirischen Studien erwies sich die semantische Differentialmethode (SD) [12] als überaus nützliches Werkzeug. Semantische Differentiale sind verbal verankerte Skalen, die typischerweise entweder aus zwei entgegengesetzten beschreibenden Adjektiven wie hell – dunkel oder aus einem Adjektiv und seiner Negation wie hell – nicht hell aufgebaut sind. Eine Reihe von Klängen kann nun nacheinander anhand mehrerer solcher Skalen beurteilt werden, die dann durch Faktoranalyse auf eine kleine Anzahl von Faktoren und Faktorladungen reduziert werden. Dafür werden oft die Teststimuli in Tonhöhe, Lautstärke und Dauer angeglichen.

In einer einflussreichen SD-Studie verwendete von Bismarck [13] im Jahr 1974 synthetische Spektren, die Vokal- und Instrumentenklänge imitierten. Diese wurden von Probanden auf verbalen Skalen beurteilt, die zur Beschreibung solcher Klangfarben als geeignet erschienen. Die Faktoranalyse ergab einen vierdimensionalen Raum. Dessen erste Dimension wurde durch das Differential stumpf – scharf gebildet, das fast die Hälfte der Gesamtvarianz in den Daten erklärte und gut mit dem spektralen Schwerpunkt der Testtöne korrelierte. Der unter anderem von Zwicker weiterentwickelte Begriff der Schärfe [14] geht auf diese Untersuchungen zu semantischen Attributen der Klangfarbe zurück.

Mit dem Schärfebegriff ergab sich jedoch für die englisch sprechende Forschungsgemeinschaft ein Übersetzungsproblem. In einem Experiment mit englisch sprechenden Probanden stellten Kendall und Cartrette [15] fest, dass Bewertungen auf dem Differential dull – sharp unzuverlässig waren. Dies wurde

darauf zurückgeführt, dass das Wort sharp im Englischen in Bezug auf Tonhöhe und Intonation, aber nicht für Klangfarbe benutzt wird.

Jüngere SD-Studien mit englisch sprechenden Probanden haben gezeigt, dass der Begriff der Helligkeit ein weiteres zentrales semantisches Attribut für Klangfarben darstellt [16]. Die Helligkeit zeigt eine Verwandtschaft mit Schärfe, und Helligkeitsurteile korrelieren mit dem spektralen Schwerpunkt von Tönen. Neben der Helligkeit haben sich außerdem die semantische Dimension der Reichhaltigkeit (fullness) und der Rauigkeit (roughness) als zentral erwiesen [17].

Die meisten SD-Studien sind in der Regel sehr sparsam in Punkto der Anzahl an Dimensionen, die sie hervorbringen. Es gibt jedoch auch Studien, die explizit die Vielfalt der sprachlichen Beschreibungsweisen betonen. Eine jüngere Studie dieser Art schlug ein Modell mit 20 potentiell redundanten semantischen Dimensionen vor [18]: Polternd / tief, leise / singend, wässrig / flüssig, direkt / laut, nasal, schrill / laut, perkussiv, rein / klar, metallisch, kratzig / körnig, klingelnd, funkelnd / brillant, luftig / atmend, resonant / lebendig, hohl, holzig, gedämpft / verschleiert, anhaltend / gleichmäßig, offen, konzentriert / kompakt. Ob diese Liste an Begriffen nun zwingend notwendig für die Beschreibung der orchestralen Farbenvielfalt ist, sei dahin gestellt. Bemerkenswert ist jedoch, wie viele Attribute darin auftauchen, die in erster Linie erstmal nichts mit Klang zu tun haben. Viele Begriffe sind vom Sehsinn, Tastsinn oder etwa aus Begrifflichkeiten für Materialeigenschaften abgeleitet und werden dann scheinbar mühelos mit klanglichen Eigenschaften assoziiert. Was die perzeptuelle und neuronale Grundlage dieser erstaunlichen Assoziationsfähigkeit ist, bleibt eine offene Forschungsfrage.

Klangfarben-Räume

Obwohl Studien mit semantischen Differentialen dank der Methode der Faktoranalyse ein Stück weit von einzelnen Begriffspaaren abstrahieren können und somit generelle semantische Faktoren aufzudecken versuchen, sind die elementaren Eingangsdaten immer begrifflich geprägt. Es existieren aber auch alternative Methoden, die ganz ohne semantische Beschreibungen und Urteile auskommen. Hierfür begannen Studien in den 1970er Jahren die Technik der sogenannten multidimensionalen Skalierung [19,20] (MDS) zu benutzen. Gefüttert mit einer Matrix von Unterschiedlichkeitsdaten erzeugt MDS eine räumliche Konfiguration von Punkten, welche die ursprünglichen Unterschiedlichkeitsdaten abbildet. Das bedeutet, sehr unterschiedliche Datenpaare werden vom MDS-Algorithmus räumlich weit entfernt angeordnet, sehr ähnliche Paare entsprechend

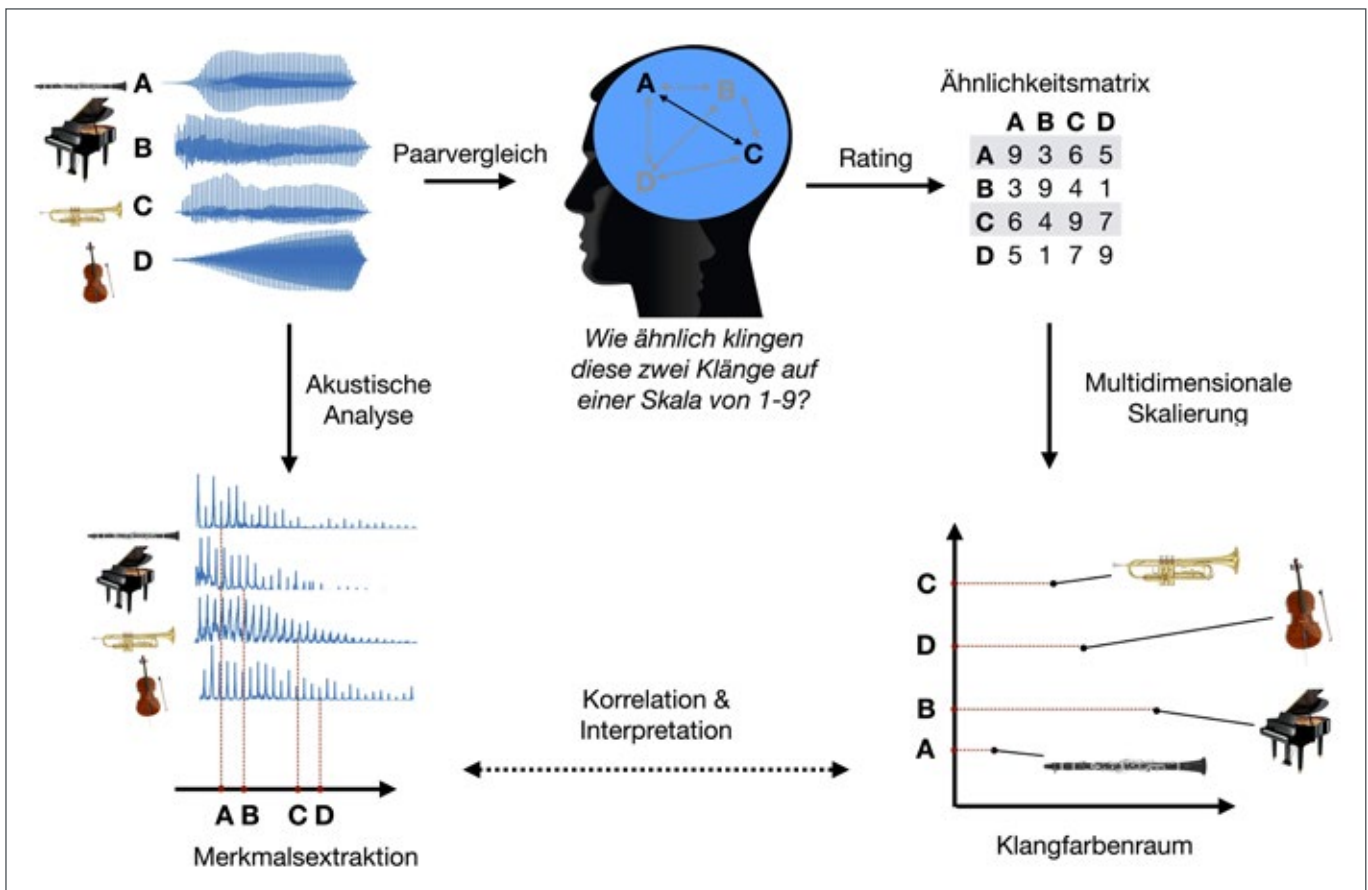


Abb. 2: Schematik zum methodischen Ansatz der Klangfarben-Räume.

nah aneinander. Wenn nun Probanden eine Reihe von Tönen bzgl. ihrer Verschiedenheit bewerten, so ist es mit MDS möglich, aus diesen Daten eine räumliche Darstellung abzuleiten. Diese Darstellungen werden oft Klangfarben-Raum oder timbre space genannt, siehe Abbildung 2. Um die Beeinflussung der Beurteilungen durch andere auditorische Attribute auszuschließen, werden die zu beurteilenden Töne normalerweise in Tonhöhe, Lautstärke und Dauer angeglichen. Mit MDS ist es also möglich, auf agnostische Weise, d. h. ohne Vorannahme von relevanten akustischen Merkmalen oder semantischen Attributen, wesentliche akustische und perzeptuelle Merkmale der Klangfarbenwahrnehmung zu explorieren. Die MDS-Methode war ein wertvolles Werkzeug für die moderne Klangfarbenforschung von Musikinstrumentenklängen. Doch wurde MDS ebenso für die Charakterisierung von Sprachqualität [21], von Industrieproduktklängen und des Sounddesigns [22] sowie der Musikwahrnehmung von Trägern von Cochlear-Implantaten [23] genutzt. Wichtige Beiträge entstanden in den 1970er Jahren. Nach ersten Anwendungen von Reinier Plomp [24] und David Wessel [25], verwendete John Grey [26] synthetisch emulierte Klänge von Orchesterinstrumenten, die mittels additiver Synthese auf der Basis von zeitvarianten Amplituden- und Frequenzverläufen von Par-

tialtönen erzeugt wurden. Greys Klangfarben-Raum hatte eine dreidimensionale Struktur. Seine Interpretation der zugrundeliegenden akustischen Dimensionen bezog sich auf die spektrale Energieverteilung für die erste Dimension, die Einschwing synchronität von Partialtönen für die zweite Dimension und dem spektralen Schwerpunkt während des Einschwingvorgangs als dritte Dimension.

Vielleicht die einflussreichste klassische MDS-Studie wurde von Stephen McAdams und Koautoren verfasst [27]. In dieser Studie mit 18 synthetischen Instrumentenklängen wurden viele Möglichkeiten des sogenannten CLASCAL-Algorithmus der MDS [28] ausgenutzt, welche die Modellierung verschiedener Probandenklassen möglich macht (diese gewichten die verschiedenen MDS-Dimensionen unterschiedlich). Ausserdem konnte jedem einzelnen Stimulus ein sogenannter Spezifitäts-Wert zugewiesen werden, der die idiosynkratische Verschiedenheit von Stimuli quantifiziert, d. h. eine Verschiedenheit, die nicht gut anhand den von allen Stimuli geteilten kontinuierlichen Dimensionen der MDS dargestellt werden kann. Drei Audio-Deskriptoren wurden als beste Kandidaten für die psychophysische Interpretation der MDS-Dimensionen angesehen: die logarithmische Einschwingzeit, der spektrale Schwerpunkt sowie der spektrale Flux (mittlere Korrelationen zwi-

schen benachbarten Kurzzeitamplitudenspektren). Somit suggerierte diese Studie, dass temporale, spektrale und spektrotemporale Merkmale für die Wahrnehmung von Instrumentenklängen wesentlich sind. Mittlerweile haben eine Reihe von MDS-Studien bestätigt, dass der spektrale Schwerpunkt und die Einschwingzeit für die Ähnlichkeitswahrnehmung die zwei wichtigsten akustischen Korrelate darstellen [29]. Die Einschwingzeit scheint besonders wichtig für Stimulus-Sets zu sein, die sowohl kontinuierlich und impulsiv angeregte Töne enthalten, da diese zwei Klassen von Tönen von der Einschwingzeit robust separiert werden. In diesem Sinne ergänzten diese Studien den Helmholtzschen Ansatz, indem sie zeigten, dass die zeitlichen Merkmale wie die Einschwingzeit ein salientes klangfarbliches Merkmal ist. Gleichzeitig wirft die geringe Dimensionalität von vielen der so generierten Klangfarbenräume – normalerweise beobachten Studien etwa zwei bis drei Dimensionen – Zweifel an der Vollständigkeit dieser Räume auf. Auf der einen Seite ist es leicht vorstellbar, dass klangfarbliche Nuancen von diesen wenigen Dimensionen nicht erfasst werden. Auf der anderen Seite könnten diese resultierenden niedrigdimensionalen Räume möglicherweise auch Einschränkungen in der Fähigkeit der Hörer widerspiegeln, Ähnlichkeits-Bewertungen für viele mehrere Wahrnehmungsfaktoren gleichzeitig vorzunehmen [30].

Eine subtile, aber wichtige Grundvoraussetzung von MDS ist, dass die wahrgenommene Ähnlichkeit zwischen zwei Tönen unabhängig von der Präsentationsreihenfolge sein sollte: Ob erst Klang A, dann Klang B oder erst Klang B, dann Klang A gehört wird, sollte keinen Unterschied machen, sonst wäre eine räumliche und somit von vorn herein symmetrische Darstellung unzulässig. Hier ist es wichtig zu wissen, dass asymmetrische Ähnlichkeitsurteile in der Psychologie nicht ungewöhnlich sind. In seinen klassischen Arbeiten legte Amos Tversky [31] überzeugend dar, dass Ähnlichkeitsurteile oft gerichtet sind, so dass etablierte, saliente Stimuli als Referenzen agieren: Wir sagen *Nord-Korea ist wie das kommunistische China* (jedoch nicht andersherum) oder *Ich spiele Fußball wie Lionel Messi* (jedoch wahrscheinlich nicht andersherum). In der Klangfarbenforschung wurde die Symmetrie der Ähnlichkeitsbewertungen erst jüngst überprüft [32]. In der Studie wurden zwei Kategorien von Stimuli benutzt, nämlich akustische und synthetische Instrumententöne. Im Ergebnis konnte festgestellt werden, dass Urteile in der Regel symmetrisch sind. Allerdings wurde beobachtet, dass Ähnlichkeitsurteile dedizierte Asymmetrien aufwiesen, wenn die Töne eines zu bewertenden Paares aus verschiedenen Kategorien (akustisch vs. synthetisch) stammten. Dies deutet

darauf hin, dass die Voraussetzung von Symmetrie und damit die Anwendbarkeit des MDS-Ansatzes insbesondere für Studien mit heterogenen Stimuli nicht uneingeschränkt gültig ist.

Akustische Merkmale

Audio-Deskriptoren

Ein wichtiges Mittel, um Hörversuche mit den zugrundeliegenden akustischen Eigenschaften der Stimuli zu verknüpfen, ist die Extraktion von akustischen Merkmalen anhand von Audio-Deskriptoren. In Abbildung 3 wird die Berechnung von einigen für die Beschreibung von Klangfarben zentralen Merkmalen angedeutet. In der oberen Reihe wird ein einfacher harmonischer, amplitudenmodulierter Tonkomplex dargestellt (500 Hz Grundfrequenz), in der unteren Reihe ein aufgenommener Klavierton (C4, 262 Hz Grundfrequenz). Die elementarste Darstellung des Signals ist die Wellenform (links). Die graue gestrichelte Linie deutet hier das Ende des Einschwingvorgangs der Töne an. Die logarithmierte Dauer des Einschwingvorgangs (log attack time) wurde in zahlreichen Klangfarben-Räumen für die Beschreibung der Impulsivität von Tönen benutzt [33].

Eine weitere elementare Darstellung des Signals ist das mittels der Fourier-Transformation gewonnene Amplituden-Spektrum. In Abbildung 3 (Mitte links) korrespondiert die graue Linie mit dem spektralen Schwerpunkt (spectral centroid). Dieser liegt für den synthetischen Tonkomplex zwischen dem vierten und fünften Partialton und beim Klavierton genau auf dem zweiten Partialton. Der spektrale Schwerpunkt entspricht numerisch dem ersten Moment der spektralen Leistungsdichte des Signals und korreliert (jedoch nicht als einziger Deskriptor) mit der wahrgenommenen Helligkeit von Tönen [34]. Ein weiterer auf der Spektralverteilung basierender Deskriptor ist das Stärkeverhältnis von geraden und ungeraden Partialtönen (odd-even-ratio) [35], welches z. B. für die hohl-klingenden Töne des tiefen Registers der Klarinette relevant ist.

In Abbildung 3 (Mitte rechts) sind die Spektrogramme der Signale dargestellt. Im Spektrogramm ist der Verlauf der akustischen Energie über die Zeit klar sichtbar, was z. B. die statische Natur des synthetischen Signals und das schnelle Abklingverhalten der hohen Partialtöne des Klaviers illustriert. Der sogenannte spectral flux Deskriptor [27] quantifiziert solche Veränderungen über die Zeit, nämlich durch Korrelation benachbarter Kurzzeitamplitudenspektren (d. h. Spalten des Spektrogramms). In der Literatur werden die erwähnten Deskriptoren und viele andere Deskriptoren meist auf Basis des Spektrogramms berechnet, womit diese Deskriptoren noch eine zeitliche Dimension erhalten.

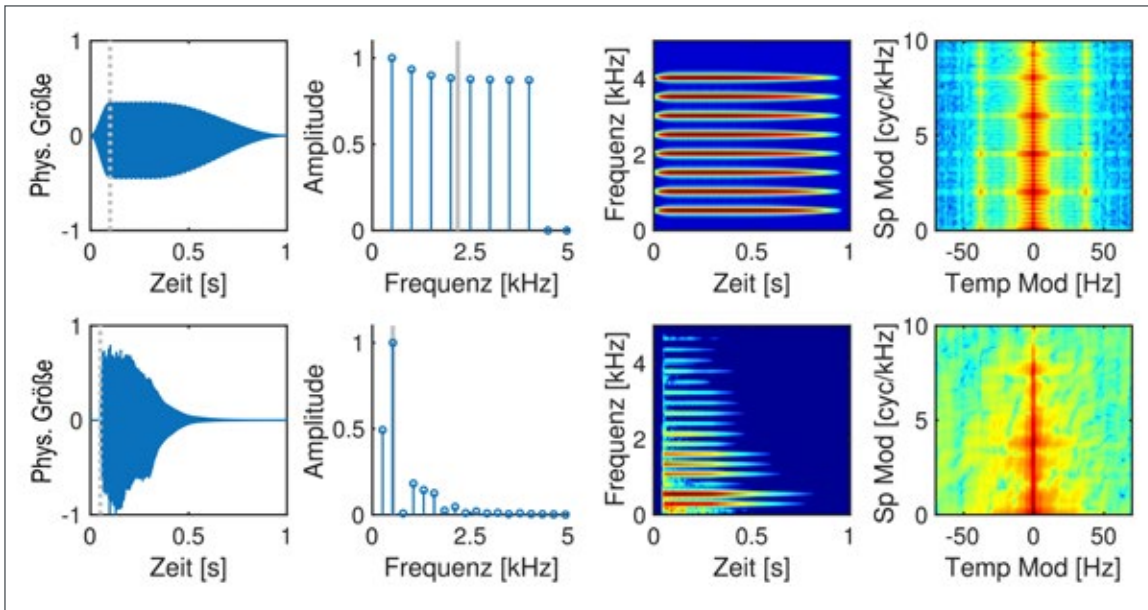


Abb. 3: Repräsentationen eines akustischen Signals. Oben: synthetisches Signal (F_0 : 500 Hz), unten: Klavierton (F_0 : 262 Hz). Von links nach rechts: Wellenform, Amplitudenspektrum, Spektrogramm, Modulationsspektrum.

Eine große Anzahl weiterer numerischer Deskriptoren wurden in der Literatur und insbesondere im Bereich des Music Information Retrieval benutzt, um klangfarbliche Aspekte von Instrumententönen zu beschreiben. Dazu gehören z. B. die spektrale Breite (spectral spread), die Rauschhaltigkeit, die spektrale Entropie, die Inharmonizität, der spektrale Crestfaktor oder die Stärke von Frequenzmodulationen [33]. Es zeigt sich jedoch, dass viele dieser Deskriptoren stark miteinander korrelieren und nur begrenzt unabhängige Informationen codieren. Beispielsweise beschreiben der spektrale Schwerpunkt und die spektrale Breite als der erste und zweite Moment der Spektralverteilung prinzipiell unterschiedliche akustische Eigenschaften. Doch da die Spektren von Instrumententönen in Richtung der tiefen Frequenzen beschränkt sind, sind diese Deskriptoren in der Praxis oft hoch korreliert – die spektrale Breite ist hoch, wenn auch der Schwerpunkt hoch ist. In einer Analyse einer Reihe von Instrumententönen fanden Geoffrey Peeters und Koautoren zehn statistisch relativ unabhängige Cluster von Deskriptoren [36]. Insofern unterliegen akustische Deskriptoren einer Art abnehmenden Grenznutzen.

Modulationsspektrum

Die Nutzung klassischer Audio-Deskriptoren setzt voraus, dass klangfarbliche Qualitäten anhand einzelner akustischer Maße bestimmt werden können. Die Helligkeit ist hierfür ein gutes Beispiel, die anhand des spektralen Schwerpunkts recht gut vorhergesagt werden kann. Ein konzeptuell anders ausgerichteter Ansatz könnte aber auch lauten, dass die Klangfarbe holistisch als Gestalt wahrgenommen wird, d. h. durch

ein komplexes spektrotemporales Muster. Diese Idee hat in jüngerer Zeit an Popularität gewonnen und wird oft mit Hilfe des Modulationsspektrums (siehe Abbildung 3, rechts) operationalisiert. Letzteres bildet die gemeinsame temporale und spektrale Variation im Spektrogramm ab und kann u. a. durch eine zweidimensionale Fourier-Transformation vom logarithmierten Amplituden-Spektrogramms berechnet werden [37]. Auf der x-Achse sind zeitliche Modulationen (in Hz) aufgetragen, auf der y-Achse finden sich spektrale Modulationen (Perioden/kHz). Für den synthetischen Ton zeigt das Modulationsspektrum Energiemaxima auf der y-Achse an, die bei Vielfachen von 2 Perioden/kHz liegen. Diese entsprechen dem Kehrwert der Grundfrequenz des Signals (0,5 kHz) und sind Folge der harmonischen Partialtonstruktur des Signals. Das Modulationsspektrum zeigt außerdem klare vertikale Komponenten bei ± 40 Hz an, was der dem Signal aufgeprägten Amplitudenmodulation entspricht und dessen Rauigkeit widerspiegelt. Diese Modulationen sind jedoch weder in der Wellenform noch dem Spektrogramm eindeutig erkennbar. Das Modulationsspektrum könnte konzeptionell als eine Art Erweiterung der vielfältig genutzten Mel-frequency cepstral coefficients gesehen werden. Während Letztere die Form des Spektrums codieren (d. h. spektrale Modulationen), codiert das zweidimensionale Modulationsspektrum die Form des Spektrogramms über die Zeit und die Frequenz ab (d. h. spektrotemporale Modulationen).

Luc Arnal und Koautoren [38] zeigten, dass mit Hilfe des Modulationsspektrums die menschlichem Schreien inhärente Rauigkeit robust charakterisiert werden konnte und dass Aktivität in der Amygda-

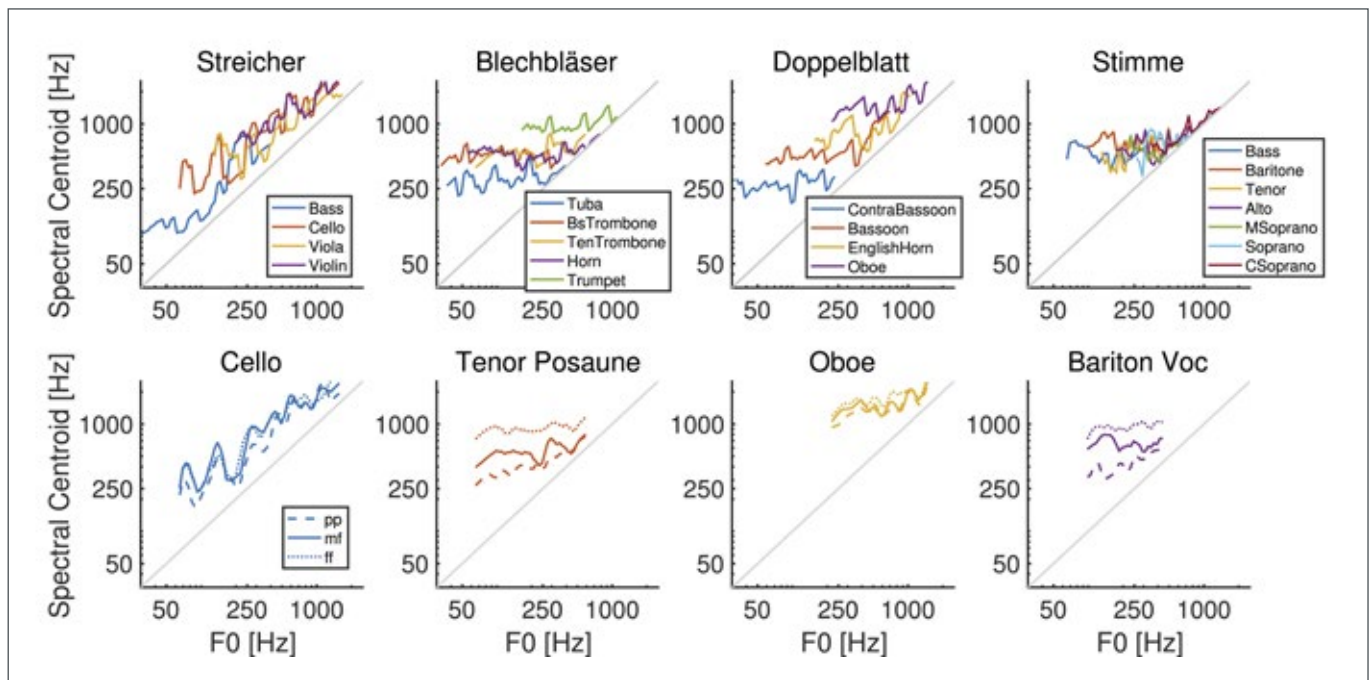


Abb. 4: Messung des spektralen Schwerpunkts bei verschiedenen Instrumenten und Instrumentenfamilien gemittelt über Spielstärken pianissimo (pp), mezzo forte (mf) und fortissimo (ff) (oben) sowie für einzelne Instrumente und Spielstärken (unten).

la, dem Emotionszentrum des Gehirns, mit dem Auftreten von Energie im Modulationsspektrum im Bereich um 70 Hz korrelierte. In Bezug auf Klangfarbenräume zeigten Taffeta Elliot und Koautoren [39], dass die Position von Tönen mit Hilfe von Modulationsspektren und eines Regressionsmodells ebenso gut vorausgesagt werden konnte wie mit klassischen Audio-Deskriptoren. Für die Identifikation von Musikinstrumenten gibt es mittlerweile ebenfalls auf der Modulationsanalyse beruhende Ansätze, die relativ präzise die Identifikationsleistung humaner Probanden nachbilden können [40].

Zusammenfassend könnte man also sagen, die derzeitige Herausforderung in der Nutzung von der akustischen Merkmalsanalyse scheint nicht zu sein, möglichst viele Informationen aus dem Signal zu extrahieren. Die Herausforderung scheint eher darin zu liegen, Ansätze zu finden, die für die jeweils betrachtete Aufgabe gerade relevanten Informationen aus Audio-Deskriptoren oder Modulationsrepräsentationen herausfiltern können, so dass diese für die Nutzer interpretierbar bleiben.

Zusammenhang von Klangfarbe, Tonhöhe und Lautstärke

Eine Limitation vieler Arbeiten zur Klangfarbenwahrnehmung ist, dass diese oft bei nur einer Tonhöhe und einer Spielstärke durchgeführt wurden. Dies mag ein erster Startpunkt sein, erlaubt die Angleichung der Tonhöhe und Lautheit es doch auszuschließen, dass Messergebnisse durch Faktoren konfundiert werden, die erst einmal nicht spezifisch für

die Klangfarbe sind. Dies würde der negativen Definition der Klangfarbe als Überbleibsel des Tons bei angeglicherer Tonhöhe und Lautstärke entsprechen. Dieser Ansatz bedarf jedoch der Ergänzung, da viele akustische Eigenschaften von Instrumententönen stark abhängig von diesen Produktionsfaktoren sind. Im Folgenden werden zuerst akustische und danach perzeptuelle Facetten dieses Phänomens beleuchtet.

Akustischer Zusammenhang von Spektrum, Grundfrequenz und Spielstärke in Instrumententönen

In Abb. 4 wird der Zusammenhang von Tonhöhe und spektralem Schwerpunkt sowie Spielstärke illustriert. Die Abbildung zeigt Messungen für die Instrumentengruppen der Streicher, Blechbläser, Doppelblatt-Bläser und der (klassischen) Gesangsstimme. Die entsprechenden Aufnahmen sind der Vienna Symphonic Library (www.vsl.co.at) entnommen. In der oberen Reihe werden Messungen gemittelt über drei Stufen der Spielstärke, unten werden Messungen für einzelne Instrumente spezifisch für jede Spielstärke gezeigt.

Offenbar gibt es für die verschiedenen Instrumentengruppen einen andersartigen Zusammenhang von Grundfrequenz F_0 und spektralem Schwerpunkt. Bei den Streichern steigt der Schwerpunkt trotz einiger Varianz in den Messungen relativ linear an. Höhere Töne werden tendenziell auf leichteren Saiten gespielt, was zu einer Erhöhung des spektralen Schwerpunktes führt. Deutlich zu sehen ist hier, dass im tieferen Register der Kontrabass (oben links, blaue Linie) einen deutlich tieferen spektralen

Schwerpunkt verglichen zum Cello (rote Linie) besitzt, was die vergleichsweise dunkle Klangfarbe des Kontrabasses begründet.

Bei den Blechbläsern und den Doppelblatt-Instrumenten sieht man zumindest bis auf die höchsten Register keine so starke Verschiebung des spektralen Schwerpunktes mit der Tonhöhe. Dies entspricht dem Umstand, dass diese Instrumente recht gut als Quelle-Filter-Modell beschrieben werden können. Die Tonhöhe beeinflusst allein das Quell-Signal, während der spektrale Schwerpunkt als grober Lageparameter des Spektrums zum Großteil durch den Filter, d. h. die physikalischen Eigenschaften des Instruments, determiniert wird. Für diese beiden Instrumentengruppen ist es auch deutlich, dass kleinere Instrumente in der Regel höhere spektrale Schwerpunkte haben, was dazu passt, dass z. B. die Trompete eine deutlich hellere Klangfarbe verglichen mit der Posaune oder gar der Tuba hat. Die Gültigkeit des Quelle-Filter-Modells voraussetzend hat Roy Patterson [41] entsprechend vorgeschlagen, dass die Instrumentenfamilie die Form der spektralen Einhüllenden bestimmt und sich diese Form mit kleiner werdenden Instrumenten der gleichen Familie in Richtung höherer Frequenzen verschiebt. Für die Gesangsstimme trifft dies in den hier betrachteten Messwerten aber nicht zu; man kann hier unabhängig vom Stimmregister sehr ähnliche Schwerpunkte beobachten.

Die Analyse zeigt ebenfalls, dass Erhöhungen der Spielstärke zu Änderungen im Spektrum von Tönen führt. Mit größerer Spielstärke erzeugte Klänge haben nicht nur prinzipiell eine größere Energie bei allen Klangkomponenten, das Spektrum breitet sich auch zu höheren Frequenzen aus, da mehr Schwingungsmodi des physikalischen Systems angeregt werden. Das impliziert insbesondere den höheren spektralen Schwerpunkt, wie in Abbildung 4 (unten) abgebildet. Doch erscheint die Stärke dieser Änderungen jedoch von Instrument zu Instrument unterschiedlich. Während die Tenor-Posaune und der Bariton extreme bis zu 500 Hz große Änderungen aufweisen, sind die Unterschiede bei dem Cello oder der Oboe geringer.

Die Messungen des spektralen Schwerpunkts können natürlich nur einen ersten groben Anhaltspunkt liefern. Für viele weitere Illustrationen der akustischen Eigenschaften der Töne von Orchesterinstrumenten sei auf das Buch von Jürgen Meyer [42] verwiesen. Was diese Darstellungen jedoch belegen, ist, wie viele Eigenheiten im Verlauf allein schon des spektralen Schwerpunkts über Tonhöhe und Spielstärke auftreten. Möglicherweise sind es gerade diese akustischen Eigenheiten oder deren Kovarianz, die es Hörern überhaupt ermöglicht, Musikinstrumente

anhand nur kurzer Töne oder Phrasen im Gewirr des Orchester- oder Ensembleklangs zu erkennen.

Perzeptuelle Wechselwirkung von Helligkeit, Tonhöhe und Lautstärke

Man könnte nun fragen, ob der eben dargestellte Zusammenhang zwischen akustischen Parametern in der Wahrnehmung eine Analogie findet. Inwiefern sind die elementaren auditorischen Wahrnehmungsdimensionen miteinander verwoben? Diese Frage hat mittlerweile zu spannenden Forschungsarbeiten in der Psychoakustik, Musikwahrnehmung und kognitiven Neurowissenschaften geführt.

Eine Studie, die den Einfluss der Helligkeit auf die musikalische Intervallwahrnehmung sehr deutlich macht, kommt von Russo und Thompson [43]. Probanden mit und ohne musikalischer Ausbildung beurteilten hier musikalische Intervalle als ab- oder aufsteigend. Bei kongruenter Variation der Helligkeit, d. h. gleichzeitiger Aufwärts- oder Abwärtsbewegung von Grundfrequenz und spektralem Schwerpunkt wurden die Intervalle von beiden Probandengruppen ausnahmslos als größer beurteilt im Vergleich zu inkongruenten Bewegungen. Das bedeutet, wenn sich beide Parameter in verschiedene Richtungen bewegten, fand eine Reduktion der wahrgenommenen Größe der Bewegung statt. Dies liefert bereits wichtige Hinweise, dass die Attribute Tonhöhe und klangfarbliche Helligkeit nicht getrennt verarbeitet werden, sondern perzeptuell interagieren.

Emily Allen und Andrew Oxenham [44] untersuchten, inwiefern Helligkeits- und Tonhöhenwahrnehmung sich gegenseitig beeinflussen können, wenn beide Dimensionen durch die Variation des spektralen Schwerpunkts oder der Grundfrequenz strikt kontrolliert werden. Insbesondere wurde in dieser Untersuchung durch vorangehende Messungen von Diskriminationsschwellen dafür gesorgt, dass die experimentell genutzten Einheiten von spektralem Schwerpunkt und Grundfrequenz für alle Probanden dieselbe Wahrnehmungsstärke besaßen. Im Hauptexperiment mussten Probanden synthetische Stimuli anhand von Differenzen des spektralen Schwerpunkts oder der Grundfrequenz unterscheiden, während Variation in der jeweils anderen Dimension von den Probanden ignoriert werden sollte. Die Ergebnisse zeigten erneut, dass die Unterscheidungsschwellen umso geringer waren, je kleiner die Variation in der zu ignorierenden Dimension war. Musiker zeigten insgesamt eine feinere Unterscheidungsfähigkeit, doch der Unterschied zu Probanden ohne musikalisches Training hing nicht von dem getesteten Parameter ab (spektraler Schwerpunkt vs. Grundfrequenz). In einer Folgestudie [45] wurden die neuronalen Korrelate dieser Unterscheidungsfähigkeit mit Hilfe der

funktionellen Magnet-Resonanz-Tomographie untersucht. Im Ergebnis konnte keine stabile anatomische Unterscheidung der Regionen des auditorischen Kortex festgestellt werden, die sich für die sensorische Kodierung von Grundfrequenz und spektralem Schwerpunkt verantwortlich zeigten.

Während die Wahrnehmung von Helligkeit und Tonhöhe eng verwandt erscheinen, was auch auf die teilweise tonotopische Repräsentation beider Attribute im auditorischen Kortex zurückzuführen sein mag, stellt sich doch auch für Lautheits-basierte Aspekte der Wahrnehmung die Frage nach Wechselwirkungen. Eine Studie ist hier der Frage nachgegangen, inwiefern die Klangfarbe die Beurteilung der Spielstärke (d. h. musikalische Dynamik) beeinflusst. Für die Wahrnehmung der Stärke eines Tons, d. h. ob er als zart und leise oder laut und kraftvoll empfunden wird, ist natürlich nicht nur der Wiedergabepegel verantwortlich, sonst würden ja Dynamikunterschiede in komprimierter Musik verschwinden, was nicht der Fall ist. Wie oben erläutert verändert sich die Spektralverteilung über Stufen der Spielstärke hinweg und der spektrale Schwerpunkt verschiebt sich nach oben. Doch zu welchem Maß diese klangfarblichen Cues und der Wiedergabepegel die wahrgenommene Dynamik beeinflussen, war lange unbekannt. Fabiani und Friberg [46] untersuchten diesen Zusammenhang. In einem Experiment testeten sie verschiedene Tonhöhen, Schallpegel, Instrumentierungen (Klarinette, Flöte, Klavier, Trompete und Violine) und den Dynamikstufen entsprechenden klangfarblichen Cues (d. h. mit Tönen, die mit den Spielstärken pp, mf und ff aufgenommen waren). Probanden wurden gebeten, die wahrgenommene Dynamik jedes Stimulus auf einer Skala von pp bis ff anzugeben. Ergebnisse zeigten, dass interessanterweise der Wiedergabepegel wie auch die klangfarblichen Cues eine ähnlich große Rolle für die Bewertung von Dynamik spielten. Die einzige Ausnahme war hier die Flöte, wo ein geringerer Einfluss der Klangfarbe festgestellt wurde.

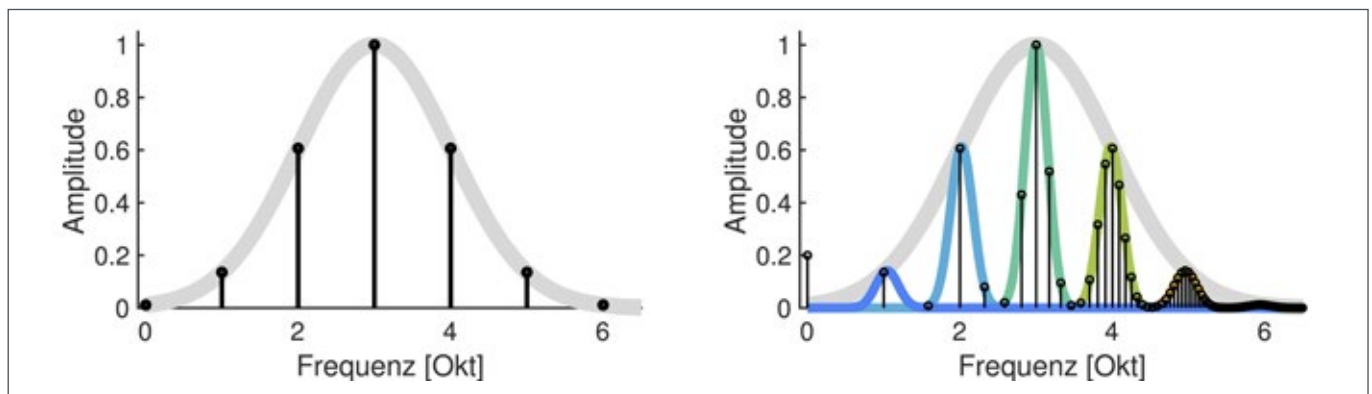
Die Tonhöhe war hauptsächlich für die Klarinette, Flöte und das Klavier relevant, wo höhere Tonhöhen höhere dynamische Bewertungen erhielten. Urteile zur Dynamik von Instrumententönen basieren daher sowohl auf der Lautstärke und Klangfarbe, als auch in geringem Maße auf der Tonhöhe.

Insgesamt zeigen sich also direkte Wechselwirkungen von sowohl klangfarblicher Helligkeit und Tonhöhe, als auch Klangfarbe und Facetten der Lautstärkewahrnehmung. Die weitere Erforschung dieses Themenkomplexes stellt eine wichtige Herausforderung für zukünftige Arbeiten der Psychoakustik dar.

Eine auf Helligkeitswahrnehmung basierende Shepard-Illusion

Die enge Verwandtschaft der Tonhöhe und Helligkeitswahrnehmung wird im Folgenden anhand der berühmten Shepard-Illusion illustriert. Seit ihrer Einführung vor mehr als 50 Jahren [47] spielen Shepard-Töne eine wichtige Rolle bei der Untersuchung der Wahrnehmung von Tonhöhen. Shepard-Töne sind Tonkomplexe, die durch die Überlagerung von Partialtönen im Oktavabstand konstruiert werden. Die Amplituden dieser Partialtöne sind durch eine Gaußsche Funktion bestimmt. Abbildung 5 zeigt eine Schematik des resultierenden Spektrums. Diese Konstruktionsweise hat zur Folge, dass ein um eine Oktave verschobener Shepard-Ton mit dem ursprünglichen (nicht verschobenen) Shepard-Ton identisch ist. Dies ermöglicht die Paarung der Partialtöne im Oktavabstand und der Gauß-förmigen spektralen Hüllkurve. Somit sind alle Intervalle zwischen Shepard-Tönen immer kleiner als eine Oktave. Wenn nun Shepard-Töne, z. B. in Halbtönen, schrittweise nach oben verschoben werden, merkt der Zuhörer meist nicht, dass das Karussell nach einer Oktave wieder am Ursprung angekommen ist, denn jeder einzelne Schritt führte ja aufwärts. Letztlich kann somit der Eindruck einer schier endlos aufsteigenden Bewegung entstehen. Dieses Phänomen ist

Abb. 5: Spektrum des klassischen Shepard-Tons (links) und des vom Autoren für die Helligkeitswahrnehmung modifizierten Stimulus (rechts). Auf der x-Achse ist die Frequenz auf logarithmischer Skala mit der Einheit Oktaven aufgetragen, auf der y-Achse die Amplitude.



die sogenannte Shepard-Illusion (online sind eine Vielzahl von Demonstrationen verfügbar). Selbstverständlich funktioniert die Shepard-Illusion auch mit absteigenden Tonfolgen oder sogar kontinuierlichen Glissandi.

Ein weiterer faszinierender Aspekt der Shepard-Töne ist die Mehrdeutigkeit in der Richtungswahrnehmung von sequentiell abgespielten Paaren von Shepard-Tönen im Tritonus- bzw. Halb-Oktav-Abstand. Bei diesen Intervallen ist es im Durchschnitt in etwa gleich wahrscheinlich, dass Hörer Verschiebungen nach oben oder unten berichten [48]. Neuere Studien haben hier insbesondere ein Augenmerk auf die Rolle des vorhergehenden auditorischen Kontextes gelegt. Claire Chambers, Daniel Pressnitzer und Kollegen haben gezeigt, dass ein solcher Kontext die Wahrnehmung von Halb-Oktav-Intervallen stark beeinflussen kann [49]. Wenn der Kontext aus Shepard-Tönen bestand, die eher höher lagen als der erste Ton des Halb-Oktav-Intervalls (z. B. +2, +4, +3 Halbtöne), so wurde das Halb-Oktav-Intervall von der Vielzahl von Probanden nach oben gehend wahrgenommen. Wenn nun tiefer liegende Shepard-Töne als Kontext präsentiert wurden (z. B. -3, -4, -2 Halbtöne), so wurde das Halb-Oktav-Intervall überwiegend als nach unten gehend gehört.

Interessierte Leserinnen und Leser mögen sich fragen, was dies mit Klangfarben zu tun hat. Die Antwort lautet, dass das gleiche Wahrnehmungsmuster für die Wahrnehmung von Helligkeit demonstriert werden kann [50]. Eine solche Analogie liefert wichtige Hinweise für das Verständnis des Zusammenhangs von Tonhöhen und Helligkeitswahrnehmung. In Abbildung 5 (rechts) findet sich das Spektrum des entsprechenden Stimulus (Klangbeispiele können auf der Website des Autors gefunden werden: uol.de/en/medipysics-acoustics/divisions/music-perception-and-processing/research). Wie sich erkennen lässt, sind die Partialtöne dieses adaptierten Shepard-Stimulus nun nicht mehr im Oktavabstand, sondern Teil einer natürlichen Obertonreihe (d. h. natürliche Vielfache der Grundfrequenz). Die Amplituden dieser Partialtöne werden nun durch zweierlei bestimmt: Einerseits der bereits bekannten globalen spektralen Einhüllenden, andererseits von einer lokalen Einhüllenden, die Oktav-zyklisch ist. Die Verschiebung dieser lokalen spektralen Einhüllenden bewirkt nun eine Veränderung der Helligkeit, jedoch nicht der Grundfrequenz. Da somit die Wahrnehmung der Tonhöhe konstant gehalten wird, ergibt sich insgesamt eine Oktav-zyklische Verschiebung der Helligkeit eines Tons bei gleichbleibender Tonhöhe. Die Shepard-Illusion funktioniert also nicht nur für die Tonhöhenwahrnehmung, sondern auch im Bereich der Helligkeitswahrnehmung. In Hörexperimenten hat sich

gezeigt, dass Halb-Oktav-Verschiebungen der lokalen spektralen Einhüllenden nun ebenso geteilte Meinungen hervorruft wie für klassische Shepard-Töne. Und sogar die Auswirkungen eines vorhergehenden Kontextes sind identisch. Ob der zuletzt präsentierte Ton als mehr oder weniger hell im Vergleich zum Vorgänger gehört wird, entscheidet maßgeblich der vorhergehende Kontext [50].

Diese überraschende Analogie unterstreicht die Ähnlichkeiten zwischen Tonhöhen und Helligkeitswahrnehmung. Interessanterweise haben neuere Arbeiten zur Tonhöhenwahrnehmung in Frage gestellt, ob die Tonhöhenwahrnehmung durch einen einheitlichen Mechanismus zur Ermittlung der Periodizität eines Tons verstanden werden kann [51]. Vielmehr scheint es, dass insbesondere die perzeptuelle Verarbeitung von kleineren Tonhöhen-Intervallen besser durch einen „spectral tracking“ Mechanismus erklärt werden kann, der konzeptionell nicht mehr stark verschieden von denen der Helligkeitswahrnehmung zugrundeliegenden Verarbeitungsschritten erscheint.

Es ist zu bemerken, dass eine gewisse Skepsis bzgl. der klassischen Unterscheidung von Tonhöhe und Klangfarbe nichts ganz Neues ist. Schon der Komponist Arnold Schoenberg räsonierte am Ende seiner Harmonielehre im Jahr 1911 [52] in eine ähnlich gehende Richtung: *Ich kann den Unterschied zwischen Klangfarbe und Klanghöhe, wie er gewöhnlich ausgedrückt wird, nicht so unbedingt zugeben. Ich finde, der Ton macht sich bemerkbar durch die Klangfarbe, deren eine Dimension die Klanghöhe ist. Die Klangfarbe ist also das große Gebiet, ein Bezirk davon ist die Klanghöhe.* Anschließend spekulierte er, dass doch auch Melodien bestehend aus dem Kontrast verschiedener Klangfarben in kompositorischer Reichweite wären. Obwohl Klangfarbenmelodien nicht Teil des Standard-Repertoires der Musik geworden sind, so ist doch die kompositorische Relevanz der Klangfarbe im 20. und 21. Jahrhundert insgesamt enorm gewachsen und ist dabei, auch in der Musikwissenschaft und Musiktheorie eine zunehmende Rolle zu spielen [53–55].

Zusammenfassung und Ausblick

Dieser Artikel hat einen Überblick über klassische Ergebnisse und neuere Entwicklungen der Klangfarbenforschung gegeben. Die Klangfarbe umfasst eine Sammlung von Attributen der auditorischen Wahrnehmung, die einzelnen Klangereignissen charakteristische Qualitäten verleiht und Hinweise auf die Klangquelle und -erzeugung liefert. Helmholtz klassische Konzeption der Tonempfindungen legte ein Grundstein für die Akustik und Musikwahrnehmung. Diese Konzeption besagt, dass die relative Stärke der Partialtöne für die Klangfarbe wesentlich

ist. In der Behandlung von neueren Entwicklungen wurde nach einem Überblick über akustische Merkmalsextraktion der Fokus auf die klangfarbliche Helligkeit gelegt, die ein zentraler Aspekt der Klangfarbe darstellt. Die akustische Grundlage der Helligkeitswahrnehmung wurden erläutert und ihre Wechselwirkungen mit der Tonhöhe und Aspekten der Lautheit illustriert. Letztlich wurden Ähnlichkeiten zwischen der Tonhöhen- und Helligkeitswahrnehmung anhand einer neuartigen, auf die Helligkeit übertragene Shepard-Illusion diskutiert. Die Klangfarbenforschung erscheint derzeit recht produktiv und dynamisch. Wie oben angedeutet gilt es noch viele elementare Aspekte der Klangfarbenwahrnehmung besser zu verstehen. Hier verdienen insbesondere die Wechselwirkungen mit anderen auditorischen Attributen Beachtung sowie eine gründlichere akustische Modellierung von Instrumententönen als Funktion der Tonhöhe und Spielstärke. Doch es gilt, auch bisher noch unbeantwortete Fragen zu stellen, z. B. welche Rolle die Klangfarbe in der Musikwahrnehmung von Schwerhörigen spielt und ob Klangfarbenforschung dazu beitragen kann, Hörgeräte für Musik zu optimieren. Diese Fragen versprechen spannende und vielfältige Herausforderungen für die nächsten Jahre.

Literatur

- [1] Conard, N. J.; Malina, M.; Münzel, S. C.: New flutes document the earliest musical tradition in southwestern Germany. *Nature* 460 (7256), pp. 737–740, 2009.
- [2] ANSI: Psychoacoustic terminology: Timbre. 1960 / 1994.
- [3] Bregman, A. S.: *Auditory Scene Analysis: The Perceptual Organization of Sound*. MIT Press, Cambridge, MA, 1990.
- [4] Siedenburg, K.; McAdams, S.: Four distinctions for the auditory “wastebasket” of timbre. *Frontiers in Psychology* 8 (1747), 2017.
<http://doi.org/10.3389/fpsyg.2017.01747>.
- [5] Thoret, E.; Goodchild, M.; McAdams, S., eds.: *Timbre 2018: Timbre is a many splendored thing*. McGill University, 2018.
- [6] Siedenburg, K.; Saitis, C.; McAdams, S.; Popper, A. N.; Fay, R. R.: *Timbre: Acoustics, Perception, and Cognition*. Springer Handbook of Auditory Research. Springer Nature, Heidelberg, Germany, 2019.
- [7] Schneider, A.: Perception of timbre and sound color. In: *Springer Handbook of Systematic Musicology*, pp. 687–725. Springer, 2018.
- [8] Handel, S.: Timbre perception and auditory object identification. In: *Hearing*, edited by B. C. Moore. *Handbook of Perception and Cognition*, pp. 425–461. Academic Press, San Diego, CA, 1995.
- [9] von Helmholtz, H.: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. Friedr. Vieweg & Sohn, Braunschweig, Germany, 1863.
- [10] Muzzolini, D.: *Genealogie der Klangfarbe*. Peter Lang, 2006.
- [11] Lyon, R. F., *Human and Machine Hearing: Extracting Meaning from Sound*. Cambridge University Press, Cambridge, MA, 2017.
- [12] Osgood, C. E.: The nature and measurement of meaning. *Psychological Bulletin* 49(3), p. 197, 1952.
- [13] von Bismarck, G.: Timbre of steady sounds: A factorial investigation of its verbal attributes. *Acta Acustica united with Acustica* 30(3), pp. 146–159, 1974.
- [14] Fastl, H.; Zwicker, E.: *Psychoacoustics: Facts and models*. 3rd ed., Springer, 2006.
- [15] Kendall, R. A.; Carterette, E. C.: Verbal attributes of simultaneous wind instrument timbres: I. von Bismarck’s adjectives. *Music Perception* 10(4), pp. 445–467, 1993.
- [16] Saitis, C.: The semantics of timbre. In: *Timbre: Acoustics, Perception, and Cognition*, edited by K. Siedenburg, C. Saitis, S. McAdams, A. N. Popper, and R. R. Fay, pp. 119–149. Springer, 2019.
- [17] Zacharakis, A.; Pasiadis, K.; Reiss, J. D.: An interlanguage study of musical timbre semantic dimensions and their acoustic correlates. *Music Perception: An Interdisciplinary Journal* 31(4), pp. 339–358 (2014).
- [18] Reymore, L.; Huron, D.: Using auditory imagery tasks to map the cognitive linguistic dimensions of musical instrument timbre qualia. *Psychomusicology: Music, Mind, and Brain* 30(3), pp. 124–144, 2020.
- [19] Shepard, R.: The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika* 27(2), pp. 125–140, 1962.
- [20] Kruskal, J. B.: Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 29(1), pp. 1–27, 1964.
- [21] Kreiman, J.; Gerratt, B. R.; Precoda, K.; Berke, G. S.: Individual differences in voice quality perception. *Journal of Speech, Language, and Hearing Research* 35(3), pp. 512–520, 1992.
- [22] Susini, P.; Lemaitre, G.; McAdams, S.: Psychological measurement for sound description and evaluation. In: *Measurement with persons: Theory, methods, and implementation areas*, edited by B. Berglund, G. B. Rossi, J. T. Townsend, and L. R. Pendrill, pp. 227–253. Psychology Press, New York, NY, 2012.
- [23] Marozeau, J.; Lamping, W.: Timbre perception with cochlear implants. In: *Timbre: Acoustics, Perception, and Cognition*, pp. 273–293. Springer, 2019.
- [24] Plomp, R.: Timbre as a multidimensional attribute of com-

- plex tones. In: Frequency Analysis and Periodicity Detection in Hearing, edited by R. Plomp and G. F. Smoorenburg, pp. 397–414. Suithoff, Leiden, The Netherlands, 1970.
- [25] Wessel, D. L.: Psychoacoustics and music: A report from Michigan State University. PACE: Bulletin of the Computer Arts Society 30, pp. 1–2, 1973.
- [26] Grey, J. M.: Multidimensional perceptual scaling of musical timbres. The Journal of the Acoustical Society of America 61(5), pp. 1.270–1.277, 1977.
- [27] 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(3), pp. 177–192, 1995.
- [28] Winsberg, S.; De Soete, G.: A latent class approach to fitting the weighted Euclidean model, CLASCAL. Psychometrika 58(2), pp. 315–330, 1993.
- [29] McAdams, S.: The perceptual representation of timbre. In: Timbre: Acoustics, Perception, and Cognition, edited by K. Siedenburg, C. Saitis, S. McAdams, A. N. Popper, and R. R. Fay, pp. 23–57, Springer, 2019.
- [30] Siedenburg, K.; Fujinaga, I.; McAdams, S.: A comparison of approaches to timbre descriptors in music information retrieval and music psychology. Journal of New Music Research 45(1), pp. 27–41, 2016.
- [31] Tversky, A.: Features of similarity. Psychological Review 84(4), pp. 327–352, 1977.
- [32] 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. <http://doi.org/10.3389/fpsyg.2015.01977>.
- [33] Caetano, M.; Saitis, C.; Siedenburg, K.: Audio content descriptors of timbre. In: Timbre: Acoustics, Perception, and Cognition, edited by K. Siedenburg, C. Saitis, S. McAdams, A. N. Popper, and R. R. Fay, pp. 297–333. Springer, 2019.
- [34] Schubert, E.; Wolfe, J.: Does timbral brightness scale with frequency and spectral centroid? Acta Acustica united with Acustica 92(5), pp. 820–825, 2006.
- [35] Caclin, A.; McAdams, S.; Smith, B. K.; Winsberg, S.: Acoustic correlates of timbre space dimensions: A confirmatory study using synthetic tones. The Journal of the Acoustical Society of America 118(1), pp. 471–482, 2005.
- [36] Peeters, G.; Giordano, B. L.; Susini, P.; Misdariis, N.; McAdams, S.: The Timbre Toolbox: Extracting audio descriptors from musical signals. The Journal of the Acoustical Society of America 130(5), pp. 2.902–2.916, 2011.
- [37] Chi, T.; Ru, P.; Shamma, S.: Multiresolution spectrotemporal analysis of complex sounds. The Journal of the Acoustical Society of America 118(2), pp. 887–906, 2005.
- [38] Arnal, L. H.; Flinker, A.; Kleinschmidt, A.; Giraud, A.-L.; Poeppel, D.: Human screams occupy a privileged niche in the communication soundscape. Current Biology 25(15), pp. 2.051–2.056, 2015.
- [39] Elliott, T.; Hamilton, L.; Theunissen, F.: Acoustic structure of the five perceptual dimensions of timbre in orchestral instrument tones. The Journal of the Acoustical Society of America 133(1), pp. 389–404, 2013.
- [40] Siedenburg, K.; Schädler, M. R.; Hülsmeier, D.: Modeling the onset advantage in musical instrument recognition. The Journal of the Acoustical Society of America 146(6), pp. EL523–EL529, 2019.
- [41] Patterson, R. D.; Gaudrain, E.; Walters, T. C.: The perception of family and register in musical tones. In: Music perception, edited by M. Riess Jones, R. R. Fay, and A. Popper, pp. 13–50. Springer, New York, NY, 2010.
- [42] Meyer, J., Akustik und musikalische Aufführungspraxis: Leitfaden für Akustiker; Tonmeister; Musiker; Instrumentenbauer und Architekten. Bochinsky, Bergkirchen, Germany, 1995.
- [43] Russo, F.; Thompson, W.: An interval size illusion: The influence of timbre on the perceived size of melodic intervals. Attention, Perception, & Psychophysics 67(4), pp. 559–568, 2005.
- [44] Allen, E.J.; Oxenham, A. J.: Symmetric interactions and interference between pitch and timbre. Journal of the Acoustical Society of America 135(3), pp. 1.371–1.379, 2014.
- [45] Allen, E.J.; Burton, P. C.; Olman, C. A.; Oxenham, A. J.: Representations of pitch and timbre variation in human auditory cortex. Journal of Neuroscience 37(5), pp. 1.284–1.293, 2016.
- [46] Fabiani, M.; Friberg, A.: Influence of pitch, loudness, and timbre on the perception of instrument dynamics. The Journal of the Acoustical Society of America 130(4), pp. EL193–EL199, 2011.
- [47] Shepard, R. N.: Circularity in judgments of relative pitch. The Journal of the Acoustical Society of America 36(12), pp. 2.346–2.353, 1964.
- [48] Deutsch, D.: The processing of pitch combinations. In: The Psychology of Music, edited by D. Deutsch, 3 ed., pp. 249–325. Academic Press, San Diego, CA, 2013.
- [49] Chambers, C.; Akram, S.; Adam, V.; Pelofi, C.; Sahani, M.; Shamma, S.; Pressnitzer, D.: Prior context in audition informs binding and shapes simple features. Nature Communications 8, p. 1.5027, 2017.
- [50] Siedenburg, K.: Timbral Shepard-illusion reveals perceptual ambiguity and context sensitivity of brightness perception. Journal of the Acoustical Society of America 143(2), pp. EL93–EL98, 2018.
- [51] McPherson, M. J.; McDermott, J. H.: Diversity in pitch perception revealed by task dependence. Nature Human Behaviour, 2(1), pp. 52–66, 2017.
- [52] Schoenberg, A.: Theory of Harmony [Harmonielehre]. University of California Press (R. E. Carter, Trans. from original German edition, 1911), Berkeley, CA, 1911/1978.
- [53] Barriere, J.-B.: Le timbre: Métaphore pour la composition, pp. 293–299, 1991.
- [54] Rehding, A.: Timbre/techné. In: The Oxford Handbook of Timbre, edited by A. Rehding and E. Dolan, Oxford University Press, 2018.
- [55] Fink, R.; Latour, M.; Wallmark, Z.: The relentless pursuit of tone: Timbre in popular music. Oxford University Press, 2018. ■



Dr. Kai Siedenburg
*Carl von Ossietzky Universität,
 Department für
 medizinische
 Physik und Akustik,
 Oldenburg*