Sind crossmodal correspondences kulturell erlernt? Zuordnungen von Farben und Tönen in Madagaskar

Sarah Ambros, Jörg Jewanski, Christoph Reuter, August Schmidhofer

Institut für Musikwissenschaft, Universität Wien, A-1090 Wien, E-Mail: Jewanski@gmx.de

Einleitung

Assoziationen zwischen den Sinnen mit konstanter Ausprägung werden als *crossmodal correspondences* bezeichnet [1]. Sie sind in der Wahrnehmung in den unterschiedlichsten Ausprägungen verankert. Innerhalb von auditiv-visuellen Zuordnungen sind es u.a. folgende konstante Assoziationen (tabellarische Überblicke: [1], S. 979; [2], S. 258):

- Tonhöhe und Helligkeit (hohe Töne werden mit hellen Farben assoziiert, tiefe Töne mit dunklen Farben) [3][4]
- Lautstärke und Helligkeit (laute Töne werden mit hellen Farben assoziiert, leise Töne mit dunklen) [3][4]
- Töne und Form (hohe Töne werden mit eckigen Formen assoziiert, tiefe Töne mit runden Formen) [3][4]
- Tonhöhe und Formgröße (hohe Töne werden mit kleinen Formen assoziiert, tiefe Töne mit großen Formen) [5].

Crossmodal correspondences zwischen instrumentalen Klangfarben und Farben/Formen wurden bisher kaum untersucht [6][7][8][9]. Historisch betrachtet, geht die auch bei Dichtern und Bildenden Künstlern häufig erwähnte Zuordnung zwischen dem Klang einer Trompete und der Farbe Rot (scarlet) auf Quellen des späten 17. Jahrhunderts zurück. So berichtet der englische Philosoph John Locke in seinem *Essay Concerning Humane Understanding*: "A studious blind Man ... bragg'd one day, That he now understood what Scarlet signified. ... It was like the Sound of a Trumpet" [10], S. 199; [vgl. 11].

Nicht endgültig geklärt ist die Frage, ob es sich bei crossmodal correspondences um eine (kulturell) erlernte oder angeborene Fähigkeit handelt [12]. Während konstante Assoziationen zwischen Lautstärke und Helligkeit bereits im frühesten Kindesalter beobachtet werden können, was auf angeboren hindeutet [13], sind Verbindungen zwischen Tonhöhe und Formgröße erst im späteren Alter nachweisbar, was auf erlernt hinweist [5]. Crossmodal correspondences zwischen Tonhöhe und Helligkeit wurden auch bei Schimpansen beobachtet, so dass zumindest diese Form im sensorischen System von Primaten verankert ist [14].

Gründe für die Existenz von crossmodal correspondences sind von der jeweiligen Ausprägung abhängig. Diejenigen Formen, die möglicherweise angeboren sind, resultieren vermutlich aus den neuronalen Strukturen für Reizkodierung: Eine Erhöhung der Stimulusintensität läuft parallel mit einer generellen erhöhten neuronalen Aktivität, die unabhängig von der Modalität des Reizes ist. Jene Formen, die bislang als erlernt erkannt wurden, resultieren vermutlich aus Erfahrungen mit der Umwelt [1].

Hinsichtlich der Frage der Entstehung ist weiterhin offen, ob möglicherweise erlernte crossmodale Verbindungen nur innerhalb einer Kultur erscheinen oder universell die gleiche Ausprägung aufweisen. Die vorliegende Studie stellt kulturelle Aspekte in den Vordergrund und untersucht mögliche Zusammenhänge und Unterschiede bei der Zuordnung von hohen und tiefen Tönen und unterschiedlichen Klangfarben in Verbindung mit Farbassoziationen.

Ziele

Das übergeordnete Ziel dieser empirischen Studie ist die Erforschung potenzieller Auswirkungen kultureller Einflüsse auf crossmodal correspondences hinsichtlich der Zuordnung von Klangfarbe und Farbassoziation. Des Weiteren ermöglicht diese Replikationsstudie [vgl. die Ausgangsstudie mit Europäer:innen: 8; 9] einen Vergleich von europäischen und nicht-europäischen Versuchsteilnehmer:innen.

Folgende Fragen wurden gestellt:
1) Kann eine konstante Verbindung von Klangfarben und Farbe bei nicht-europäischer Bevölkerung beobachtet werden?
2) Inwiefern unterscheiden/ähneln sich die Zuordnungen von Nicht-Europäer:innen und Europäer:innen?
3) Welche Unterschiede und Gemeinsamkeiten können bei der Farbauswahl von Kindern und Erwachsenen beobachtet werden?
4) Gibt es Parallelen zu bisher durchgeführten Studien?

Methode

Teilnehmende und Studienablauf

63 Teilnehmende aus Toliara (einer Stadt im Südwesten Madagaskars) nehmen an der Studie teil: 33 männliche und 30 weibliche Personen im Alter von zehn bis 60 Jahren (Mittelwert: 27,3).

Die Proband:innen sind in sechs Gruppen mit je zehn Personen (fünf Männer, fünf Frauen) sowie zu einer Gruppe bestehend aus drei Männern aufgeteilt. Die Gruppen werden nacheinander befragt, die Teilnehmer erhalten eine Aufwandsentschädigung. Zu den Klangfarbe-Farbe-Zuordnungen werden zusätzlich grundlegende demographische Daten wie Alter und Geschlecht erfragt.

Akustische Stimuli

Die 60 Hörbeispiele beinhalten zehn Instrumente (Flöte, Oboe, Klarinette, Fagott, Horn, Trompete, Posaune, Tuba, Violine und Cello) aus der *Vienna Symphonic Library*. Alle Stimuli werden in gleicher Laustärke sowie jedes Instrument in drei verschiedenen Tonhöhen (E2 bis E6) abgespielt. Die Klänge werden in randomisierter Reihenfolge gespielt und wiederholen sich, um die Übereinstimmung der einzelnen Farbzuordnungen zu überprüfen. Die Instrumentennamen als auch die Tonhöhen werden nicht mitgeteilt, um symbolische Assoziationen zu vermeiden. Die Artikulation der Instrumente ist legato. Die Ausgabe der Klänge erfolgt über Kopfhörer.

Visuelle Stimuli

Wie bereits in anderen Studien zu audio-visuellen Assoziationen [15][16] sowie der Ausgangsstudie [8][9] werden als visuelle Stimuli die 38 Farben des *Berkeley Color Projects* gewählt (Abbildung 1), die farblich leicht angepasst und anders angeordnet sind, um die nicht vorhandene Erfahrung der Teilnehmer:innen mit Fragebögen und Online-Erhebungen zu berücksichtigen: die Fragebögen werden händisch mit Kugelschreiber ausgefüllt.

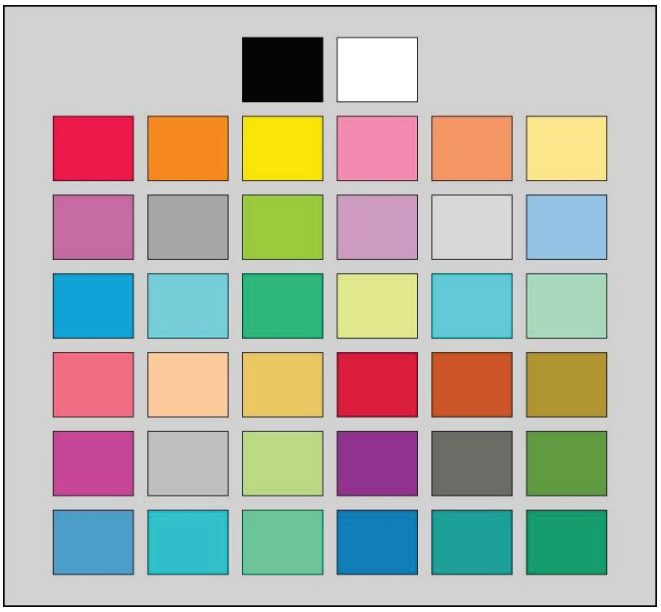

Abbildung 1: *Berkeley Color Project*-Farbpalette

Die Farben sind am Fragenbogen auf weißem Hintergrund dargestellt. Der Fragebogen umfasst zehn Seiten mit sechs Spalten für akustische Stimuli pro Seite (60 Klänge). Pro Klang sollen von den Teilnehmer:innen jeweils drei Farben zugeordnet werden (Abbildung 2).

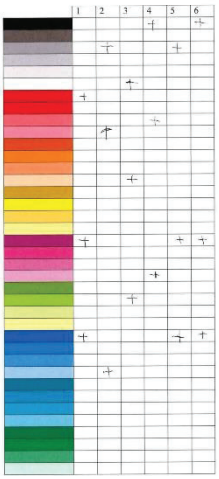

Abbildung 2: Erste Seite des Fragebogens. 38 Farben wurden in Zeilen dargestellt sowie 6 akustische Stimuli pro Seite in Spalten.

Evaluation

Um klare Farbzuordnungen einzelner Stimuli aufzuzeigen, werden Häufigkeiten für jede Instrumentengruppe, sowie zusätzlich für jede Tonhöhe berechnet. Die Häufigkeiten werden zudem separat für 51 erwachsene Versuchspersonen, sowie für 12 teilnehmende Kinder ausgegeben, um beide Ergebnisse miteinander zu vergleichen. Fleiss' Kappa (κ) wird ermittelt, um die Interrater Reliabilität der Teilnehmenden zu analysieren. Da jeder Stimulus doppelt aufgeführt ist, wird Cohens Kappa für alle 63 Versuchspersonen errechnet, um die individuelle Übereinstimmung der Farbantworten zu überprüfen. Alle Farbzuordnungen werden gemittelt und auf Korrelationen bezüglich verschiedener Klangeigenschaften analysiert.

Studienergebnisse

Die Interrater Reliabilität fällt mit $\kappa = 0.012$ ($p < 0{,}001$) gering aus, möglicherweise kann dies auf die zu breite Farbauswahl von 38 Farben zurückgeführt werden. Auch die individuelle Übereinstimmung bei Klangfarbe-Farbe-Zuordnungen der Proband:innen ist durchgehend niedrig.

Häufigkeiten

Hier zeigt sich, dass die Farbe Schwarz mehrfach assoziiert wird – die Klänge der Flöte, der Posaune, des Horns sowie des Fagotts werden in allen Tonlagen am häufigsten mit dieser Farbe beschrieben. Ein möglicher Zusammenhang der Farbwahl mit der Position auf dem Fragebogen (sie stand ganz oben) müsste in weiteren Studien genauer untersucht werden.

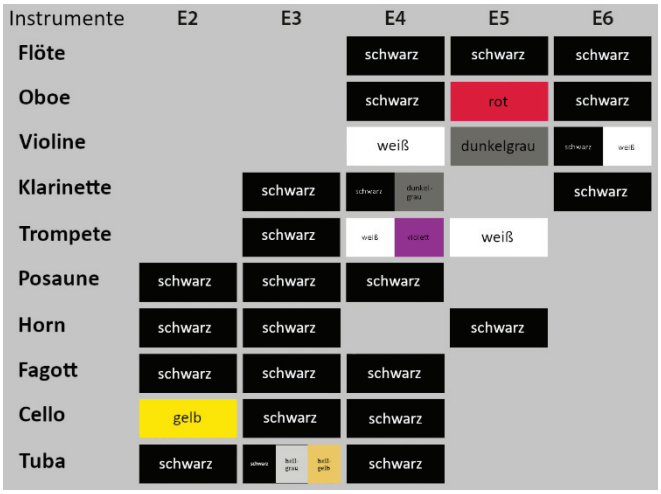

Abbildung 3: Farbhäufigkeiten pro Instrument und Tonhöhe. Häufig assoziiert wurde die Farbe Schwarz sowie in tiefen Lagen der Streichinstrumente zudem helle Farben.

Bei Betrachtung einiger Instrumenten-Farbzuordnungen können weitere Auswahltendenzen einzelner Farbgruppen aufgezeigt werden: Holzblasinstrumente, wie beispielsweise die Oboe werden häufiger mit starken Rottönen (Orange, Rot) assoziiert. Die Trompete wird vermehrt mit den Farben Schwarz, Violett und in hoher Tonlage E5 etwa mit Weiß in Verbindung gebracht. Auch bei den Streichinstrumenten lassen sich eindeutige Zuordnungen festhalten: sowohl dem tiefsten Klang der Violine als auch des Cellos ordnen die Teilnehmenden mehrheitlich helle Farben (Weiß, Gelb) zu (Abbildung 3).

Vergleich zwischen Kindern und Erwachsenen

Die Annahme, dass sich die Farbzuordnungen der Kinder von jenen der Erwachsenen unterscheiden könnten, wird bei den meisten Instrumenten bestätigt. Insbesondere die Zuordnungen der Oboe, Violine, Klarinette, Posaune und dem Horn unterscheiden sich stark voneinander: während bei den erwachsenen Proband:innen die Farbe Schwarz den

Holzblasinstrumenten sowie zudem den meisten Blechblasinstrumenten am häufigsten zugeordnet wird, assoziieren die Kinder nahezu einstimmig die Farbe Rosa mit den Klängen der Oboe und Violine. Ebenfalls hellere Farbtöne werden der Klarinette (Gelb) sowie dem Horn (helles Orange) zugeordnet (Abbildung 4).

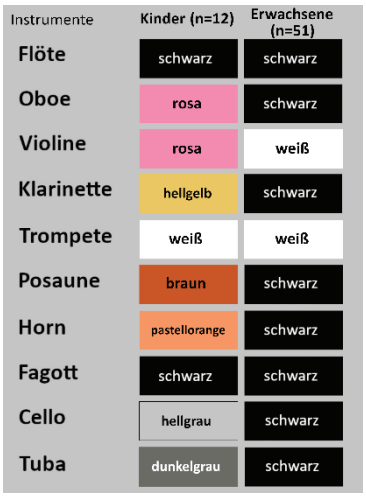

Abbildung 4: Instrumenten-Farbassoziationen bei Kindern und Erwachsenen.

Korrelationen

In der Ausgangsstudie konnte ein positiver Zusammenhang zwischen steigender Tonhöhe und der Helligkeit der zugeordneten Farbe festgestellt werden [8][9]. In der vorliegenden Studie können bei Korrelationsberechnungen von RGB-Farbmittelwerten moderat bis große Korrelationen einzelner Eigenschaften gezeigt werden. Die Berechnung des Korrelationskoeffizienten r nach Pearson ergibt eine stark positive Korrelation zwischen Tonhöhe ($r=0{,}591$; $p< 0{,}001$), Tonhelligkeit ($r=0{,}535$; $p> 0{,}003$) und Spectral Centroid ($r=0{,}561$; $p< 0{,}002$) in Bezug auf die Farbhelligkeit. Auch in Bezug auf die Farbsättigung der gemittelten Farbzuordnungen kann eine positive Korrelation von Tonhöhe, Helligkeit, Spectral Centroid und dem Frequenzband zwischen 2 und 4 Kilohertz beobachtet werden ($p< 0{,}023$). Hinsichtlich des Farbtons auf der anderen Seite ergibt sich eine moderat negative Korrelation der drei erstgenannten Klangeigenschaften (Tabelle 1).

Tabelle 1: Korrelationen von Klang- und Farbeigenschaften auf Basis von berechneten RGB-Farbmittelwerten.

	Farbton	p-Wert	Farbsättigung	p-Wert	Farbhelligkeit	p-Wert
Tonhöhe	-0,433	0,019	0,420	0,023	0,591	0,001
Helligkeit	-0,499	0,006	0,563	0,001	0,535	0,003
Spectral Centroid	-0,578	0,001	0,584	0,001	0,561	0,002
Härte	0,332	0,078	-0,352	0,061	-0,433	0,019
khz2to4Energie	-0,472	0,010	0,576	0,001	0,518	0,004

Vergleich mit europäischen Ergebnissen

Bei Europäer:innen im Vergleich zu Nichteuropäer:innen unterscheiden sich die Studienergebnisse hinsichtlich mehrerer Faktoren: Während bei Europäer:innen eine steigende Tonhöhe mit helleren Farbzuordnungen korrelierte (Abbildung 5), kann ein solches Ergebnis bei den Madagass:innen auf Ebene einzelner Klänge lediglich bei der Farbwahl der Trompete gezeigt werden (siehe Abbildung 3).

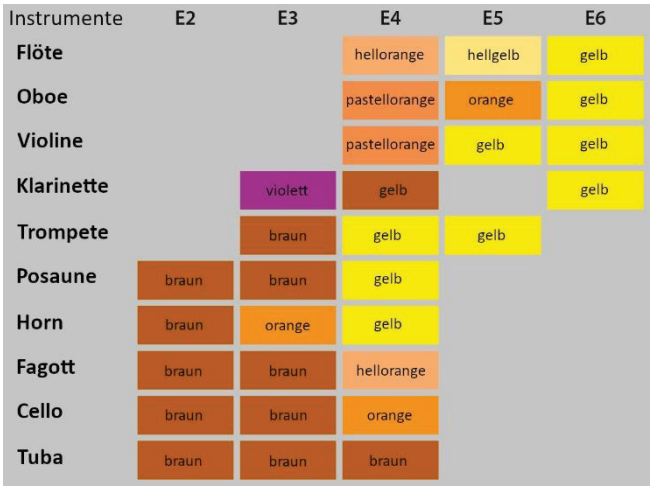

Abbildung 5: Ermittelte Farbhäufigkeiten pro Instrument und Tonhöhe von Europäer:innen [9].

Sprachanalogien und kulturelle Aspekte

Eine mögliche Erklärung der unterschiedlichen Ergebnisse von Europäer:innen und Nichteuropäer:innen könnte in der Sprache bzw. der jeweiligen Kultur zu finden sein. Während im deutschen Sprachgebrauch Beziehungen wie hoch-hell und tief-dunkel gängig sind, bezeichnet das Volk der Shona in Simbabwe beispielsweise hohe Töne als dick, tiefe als dünn. In Zentralafrika andererseits werden hohe Klänge als schwach und tiefe als stark beschrieben [17]. Möglicherweise handelt es sich bei dieser Form von crossmodal correspondences um eine westlich ankulturierte Fähigkeit, während in nicht-westlichen Kulturen die Farbe anderen auditiven Attributen (z.B. der Klangfarbe) zugeordnet wird. Die Verbindung von hohen Tönen mit hellen Farben könnte dementsprechend kulturell geprägt und sprachabhängig sein, müsste jedoch durch weitere Forschungen in und außerhalb Europas genauer untersucht werden.

Fazit und Ausblick

Es konnten keine konstanten Verbindungen von einzelnen Klängen und Farben gezeigt werden. Auch bereits bekannte crossmodal correspondences (z.B. steigende Tonhöhe und Farbhelligkeit) ließen sich nicht eindeutig beobachten. Jedoch unterschieden sich die Farbassoziationen einzelner Instrumente bei madagassischen Kindern und Erwachsenen stark voneinander.

Aus der westlichen Literatur bekannte Instrumenten-Farbassoziationen wie beispielsweise die Farben Rot und Gelb, die häufig dem Klang einer Trompete zugeordnet werden, konnten in dieser Studie mit Madagass:innen nicht aufgezeigt werden.

Die Ergebnisse sprechen möglicherweise dafür, dass es sich bei dieser Art von crossmodal correspondences um keine angeborene, universell gültige Fähigkeit handelt, sondern verweisen auf erlernte Faktoren, welche durch Erfahrungen mit der Umwelt bzw. der Kultur und Sprache geprägt werden. Es bedarf jedoch weiterer Untersuchungen an nichteuropäischer Bevölkerung. Das Verbinden von quantitativen mit qualitativen Forschungsmethoden kann ebenso tiefergehende Erkenntnisse liefern wie der Einfluss der Darstellung von einzelnen Farben sowie der von

kulturellen Farbsymboliken und ein Vergleich von europäischen und nicht-europäischen Musikinstrumenten.

Die Autor:innen bedanken sich bei Isabella Czedik-Eysenberg für ihre Unterstützung im Bereich der Audiofeature-Analyse und bei Saleh Siddiq für seine Hilfe bei der Darstellung und Berechnung der Farbwerte.

Literatur

[1] Spence, C.: Crossmodal correspondences: A tutorial review. Attention Perception & Psychophysics 73 (2011), 971–995

[2] Louden, S. J.: Cross-sensory perception in music and visual media: A neurocognitive approach to multimedia counterpoint. PhD Diss. Faculty of the Graduate School of the University at Buffalo, State University of New York, 2018

[3] Marks, L. E.: On Associations of Light and Sound: The Mediation of Brightness, Pitch, and Loudness. The American Journal of Psychology 87/1 (1975), 173–188

[4] Marks, L. E.: On cross-modal similarity: Auditory–visual interactions in speeded discrimination. Journal of Experimental Psychology: Human Perception and Performance 13/3 (1987), 384–394.

[5] Marks, L. E.; Hammeal, R. J. & Bornstein, M. H.: Perceiving similarity and comprehending metaphor. Monogram Society Research Children Development 52/1 (1987) (special issue)

[6] Ward, J.; Huckstep, B. & Tsakanikos, E.: Sound-colour synaesthesia: To what extant does it use cross-modal mechanisms common to us all? Cortex 42 (2006), 264–280

[7] Adeli, M.; Rouat, J. & Molotchnikoff, S.: Audiovisual correspondence between musical timbre and visual shapes. Frontiers in Human Neuroscience 8/352 (2014), 1–11

[8] Reuter, C.; Jewanski, J.; Saitis, C.; Czedik-Eysenberg, I.; Siddiq, S.; Kruchten, S. & Oehler, M.: Colors and Timbres – Consistency and Tendencies of Color-Timbre Mappings in non-synesthetic Individuals. Poster presented at the Conference "Musik im audiovisuellen Kontext" – 34. Jahrestagung 2018 der Deutschen Gesellschaft für Musikpsychologie (DGM), Universität Gießen, 7.-9. September 2018

[9] Siddiq, S.; Czedik-Eysenberg, I.; Jewanski, J.; Saitis, C.; Kruchten, S.; Sakhabiev, R.; Oehler, M. & Reuter, C.: Colored Timbres. Do Crossmodal Correspondences between Musical Instrument Sounds and Visual Colors Rather Depend on Pitch instead of Timbre? [submitted]

[10] Locke, J.: An Essay Concerning Humane Understanding. Basset, London, 1690

[11] Jewanski, J. & Reuter, C.: About Blue Flutes and Red Trumpets. A Tutorial Review on the Historical Development of Timbre-Color Mappings [in preparation]

[12] Spence, C. & Deroy, O.: Crossmodal correspondences: Innate oder learned? i-Perception 3 (2012), 316–318

[13] Lewkowicz, D. J.; Turkewitz, G.: Cross-modal equivalence in early infancy: Auditory–visual intensity matching. Developmental Psychology 16 (1980), 597–607

[14] Ludwig, V. U.; Adachi, I. & Matsuzawa, T.: Auditory mappings between high luminance and high pitch are shared by chimpanzees (Pan troglodytes) and humans. PNAS (2012). doi: 10.1073/pnas.1112605108

[15] Palmer, S.; Schloss, K.; Xu, Z. & Prado-Léon, L.: Music-color associations are mediated by emotion. PNAS 110/22 (2013), 8836–8841

[16] Whiteford, K. L.; Schloss, K. B.; Helwig, N. E. & Palmer, S. E.: Color, music, and emotion: Bach to the blues. i-Perception 9/6 (2018), 1–27

[17] Eitan, Z. & Timmers, R.: Beethoven's last sonata and those who follow crocodiles: Cross-domain mappings of auditory pitch in a musical context. Cognition 114 (2010), 405–422