

Klingt Sinus blau und Sägezahn rot?

Eine Untersuchung zu Crossmodal Correspondences bei der Wahrnehmung von synthetischen Wellenformen

Gabriel Feller, Christoph Reuter
Musikwissenschaftliches Institut der Universität Wien, Österreich

Hintergrund

Crossmodal Correspondences beschreiben allgemein verbreitete Verknüpfungen zweier Sinnesmodalitäten. Die Tonhöhe lässt sich z.B. mit der Größe eines Objekts, der räumlichen Höhenanordnung, der Helligkeit, mit Bewegungsrichtungen oder mit Gewicht in Verbindung bringen [1]. Bereits 1971 testeten Bernstein et al. mithilfe von Sinuswellen anhand von vier Personen mögliche Assoziationen von bestimmten Farben (rot und blau) zu Tonhöhen (100 Hz und 1000 Hz), jedoch ohne Ergebnis [2]. Ward et al. konnten 2006 bei Hörversuchen mit Sinustönen und anderen Klängen einen Zusammenhang zwischen Tonhöhe und Helligkeit sowie zwischen Klangfarbe und damit assoziierter Farbe erkennen [3]. Nach Paris und Spence (2012) werden Sinusklänge eher mit runderen visuellen Stimuli und Rechteckschwingungen eher mit kantigeren visuellen Stimuli [4].

Ziele und Fragestellungen

- Welche farblichen Crossmodal Correspondences treten bei synthetischen Wellenformen auf?
- Können bisherige Untersuchungen zur crossmodalen Wahrnehmung von verschiedenen synthetischen Wellenformen bestätigt werden?
- Gibt es signifikante Unterschiede in der Beschreibung der Wellenformen?

Methode

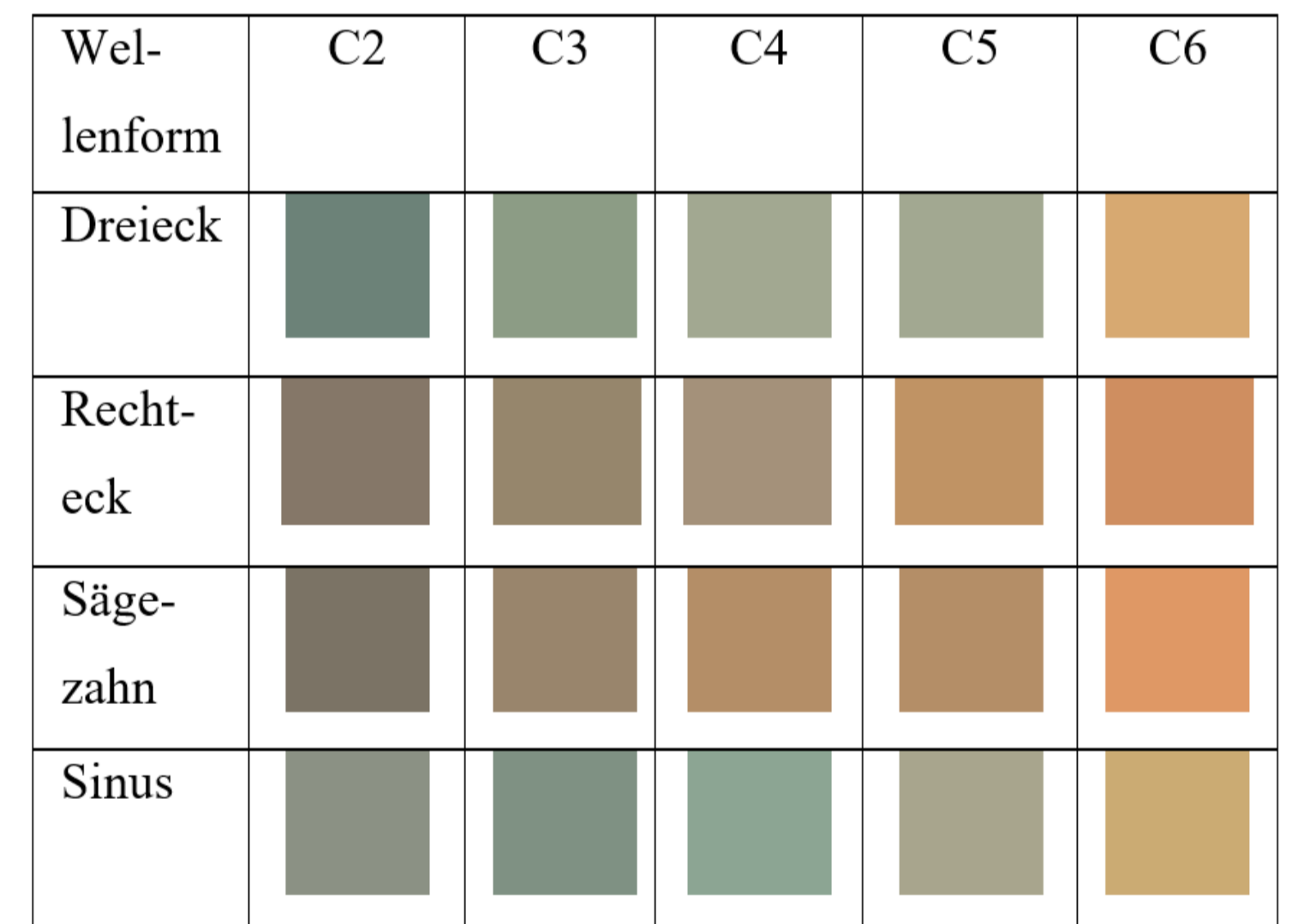
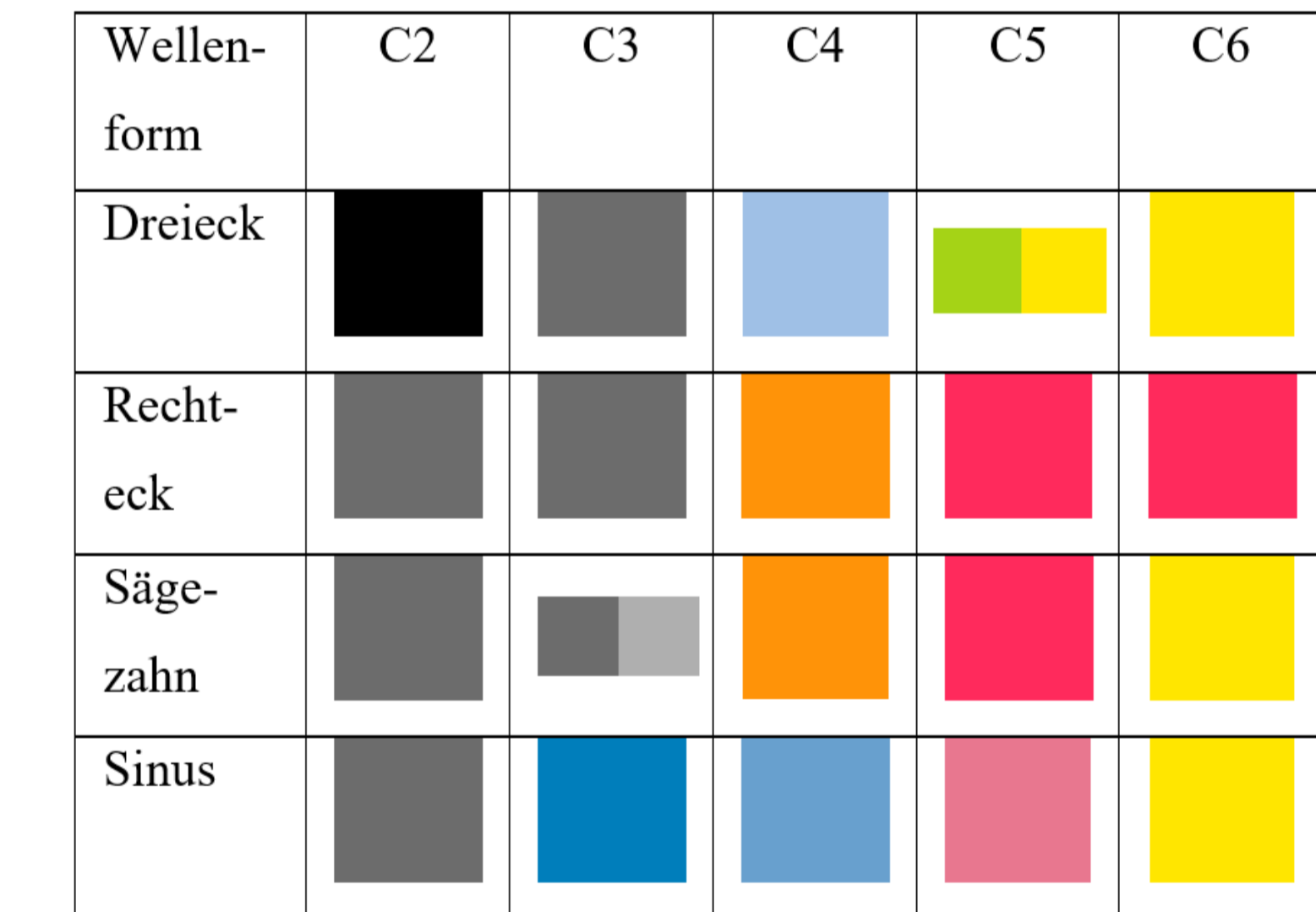
44 Versuchspersonen (14 ♂, 30 ♀, Alter: 18-80, \bar{x} 37,8 Jahre, SD=16,5) ordneten in einem Onlineexperiment jeweils 5 lautheitsangeglichene Sinus-, Dreiecks-, Rechteck- und Sägezahnschwingungen (Dauer: 4 s, auf C2, C3, C4, C5, C6) drei subjektiv am besten passenden Farben aus dem Berkeley Color Project zu [6][7]. Alle ausgewählten Farben wurden nach Stimmenanzahl gewichtet im RGB-Farbmodell gemittelt und in den HSV-Farbraum übertragen.

Darüber hinaus gaben die Versuchspersonen mit Hilfe von Schieberegler auf 10 Skalen zwischen 0 und 1 an, inwieweit der jeweils gehörte Klang den Adjektiven aus der Studie von Grill et al. [5] zur sprachlichen Beschreibung von Klangtexturen entspricht.

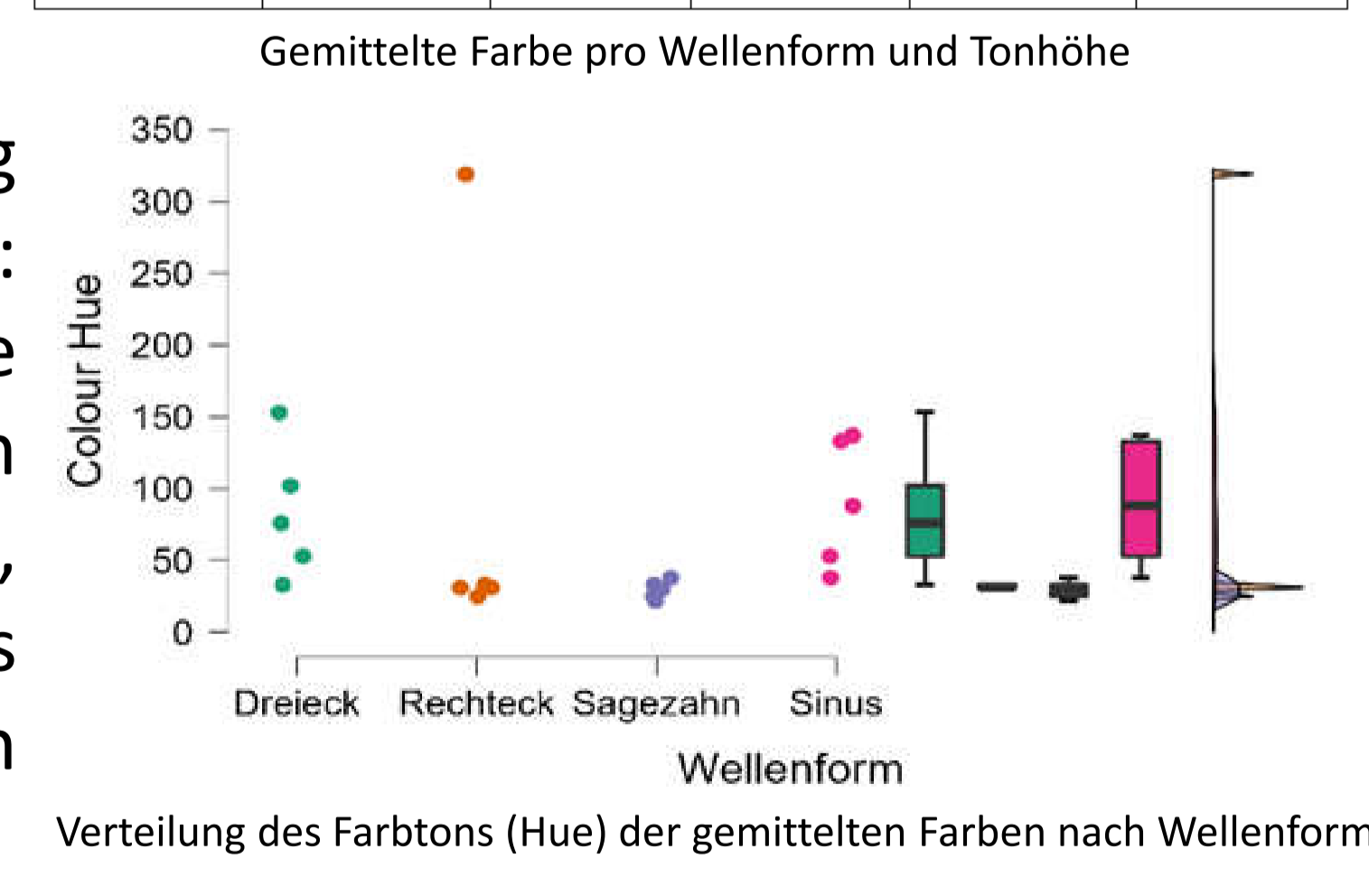
Via Signalanalyse [8][9][10][11][12] wurden pro Klangbeispiel jeweils 61 berechenbare Audiomerkmale extrahiert. Über Korrelations- und Regressionsanalysen wurden Zusammenhänge zwischen den Farbdimensionen, den Einordnungen in die Adjektivliste und den errechneten Audiomerkmale ermittelt.

Ergebnisse

Tiefe Töne werden häufig mit Schwarz und Grau assoziiert, hohe Töne eher mit warmen Farben (Gelb und Rot). Bei den Sinus- und Dreiecksschwingungen zeigen sich Tendenzen, kalte Farben (Blau, Grün) im mittleren Tonbereich zu wählen. Bei Sägezahn und Rechteck sind es dort eher Orange- und Rottöne.



Dies zeigt sich auch nach der Mittelung der gewählten Farben pro Klangbeispiel: Während sich teiltonreiche Klänge (Rechteck und Sägezahn) konzentriert im roten Bereich aufhalten (Hue bei 0), verteilen sich teiltonärmere Klänge (Sinus und Dreieck) weiter zu den grün-blauerer Farbtönen hin (Hue bei 90 und darüber).



Ein t-Test zeigt, dass der Blau- und Grün-Anteil (RGB-System) bei Rechteck- und Sägezahnklängen signifikant geringer ist als bei Sinus- und Dreiecksklängen (Blauanteil: $t=-5,804$, $p<0.001$; Grünanteil: $t=-3.627$, $p=0.002$).

Literatur

[1] Louden, S. J.: Cross-Sensory Perception in Music and Visual Media. A Neurogenetic Approach to Multimedia Counterpoint. Buffalo (2018). <> [2] Bernstein, I. H., Eason, T. R.: Hue-Tone Sensory Interaction. A Negative Result. In: Perceptual and Motor Skills 33/3 (1971), 1327-1330. <> [3] Ward, J., Huckstep, B., Tsakanikos, E.: Sound-Colour Synaesthesia. To What Extent Does It Use Cross-Modal Mechanisms Common to Us All? In: Cortex 42 (2006), 264-280. <> [4] Parise, C. V., Spence, C.: Audiovisual crossmodal correspondences and sound symbolism. A study using implicit association test. In: Experimental Brain Research 220/3-4 (2012), 319-333. <> [5] Grill, T., Flexer, A., Cunningham, S.: Identification of perceptual qualities in textural sounds using the repertory grid method. In: Proceedings of the 6th Audio Mostly Conference (2011), 67-74. <> [6] Ambros, S. et al.: Sind crossmodal correspondences kulturell erlernt? Zuordnung von Farben und Tönen in Madagaskar. In: Fortschritte der Akustik - DAGA 2021. Deutsche Jahrestagung für Akustik (2021), 1076-1079. <> [7] Palmer, S. E. et al.: Music-color associations are mediated by emotion. In: PNAS 110/22 (2013), 8836-8841. <> [8] Lartillot, O. et al. (2008). A Matlab Toolbox for Music Information Retrieval. In: C. Preisach et al. (Hrsg.), Data Analysis, Machine Learning and Applications (S. 261-268). Berlin, Heidelberg: Springer. <> [9] Lartillot, O. (2019). Miningsuite: A comprehensive Matlab framework for signal, audio and music analysis. Proceedings of the 16th SMC (S. 489-489), Málaga. <> [10] Bogdanov, D et al. (2013). Essentia: An Audio Analysis Library for Music Information Retrieval. Proceedings of the 14th ISMIR Conference (S. 493-498), Curitiba. <> [11] Pearce, A., et al. (2017). Timbral attributes for sound effect library searching. AES Proceedings, Erlangen. <> [12] McFee, B., et al. (2015). Librosa: Audio and music signal analysis in python. Proceedings of the 14th python in science conference (S. 18-25), Austin.

Ergebnisse

Aus der Korrelationsanalyse der klanglichen Merkmale mit den Werten der damit assoziierten Farben wird deutlich:

- Je höher die Tonhöhe, desto höher der Rot-Anteil sowie auch der Grün-Anteil und die farbliche Helligkeit.
- Je heller die Klangfarbe, desto höher der Rot-Anteil, die Sättigung und die farbliche Helligkeit und desto geringer der Grün- und Blau-Anteil sowie der Farbton (Hue). Das gilt umgekehrt für Merkmale der klanglichen Wärme.
- Je geräuschhafter und rauher die Klangfarbe, desto stärker der Rot-Anteil, die Sättigung und die Helligkeit und umso geringer der Grün- und Blau-Anteil.

RGB/HSV-Farbwerte korrelieren mit	Rot-Anteil (R)	Grün-Anteil (G)	Blau-Anteil (B)	Farbton (H)	Sättigung (S)	Helligkeit (V)
Klangliche Merkmale der Tonhöhe						
Frequenz	0,902	0,533				0,523
Klangliche Merkmale der Helligkeit						
Zero Crossing Rate	0,933	0,491		-0,454		0,584
Spectral Strongpeak	0,796	0,638				
Spectral Centroid	0,472	-0,460	-0,820		0,597	0,455
Spectral Rolloff		-0,571	-0,818		0,597	
Spectral Flatness		-0,510	-0,798		0,588	
Spectral Spread		-0,501	-0,829		0,603	
Spectral Entropy		-0,562	-0,814		0,577	
Brightness		-0,497	-0,819		0,591	
Spectral Bandwidth		-0,531	-0,819		0,590	
Spectral Decrease	0,479					
Timbral Sharpness	0,589		-0,907		0,586	0,512
Timbral Brightness	0,648		-0,803		0,580	0,549
Spectral Energy (4000-20000 Hz)	0,576		-0,682		0,494	0,524
Spectral Energy (800-4000 Hz)	0,737					0,472
Spectral Contrast (800-1600 Hz)		0,643				
Spectral Contrast (800-1600 Hz)	0,678	0,671				
Spectral Contrast (1600-3200 Hz)	0,575					
Spectral Contrast (3200-6400 Hz)	0,481					
Spectral Kurtosis			0,719		-0,512	
Spectral Skewness		0,488	0,787		-0,563	
Klangliche Merkmale der Wärme						
Spectral Contrast (200-400 Hz)	-0,462					
Spectral Contrast (0-200 Hz)	-0,770	-0,555		0,489		
Timbral Warmth	-0,904		0,525	0,519		-0,632
Timbral Booming	-0,912		0,574	0,478		-0,612
Timbral Depth	-0,924	-0,543		0,461		-0,536
Spectral Energy (20-150 Hz)	-0,492			0,468		
Spectral Energy (150-800 Hz)						
Klangliche Merkmale der Geräuschhaftigkeit						
Harmonic Percussive Ratio		0,655	0,733		-0,490	
Percussive Energy		-0,773	-0,532			
Spectral Complexity		-0,790	-0,512			
Timbral Hardness	0,549		-0,845		0,596	0,479
Timbral Roughness		-0,709	-0,837		0,550	
Spectral Flux (>1000 Hz)	0,510		-0,805		0,527	0,480
Spectral Flux (<100 Hz)	-0,578	-0,778				
Spectral Flux		-0,562	-0,790		0,510	
Roughness (Vassilakis)	0,472		-0,619			
Roughness (Sethares)		-0,645	-0,699			

Positive und negative Korrelationen zwischen klanglichen Merkmalen und Farbwerten der assoziierten Farben (es wurden nur Korrelation mit $r<-0,4$ bzw. $r>0,4$ sowie $p<0,05$ einbezogen)

Bei den Zuordnungen der Klangbeispiele zu den Adjektivlisten von Grill et al [5] zeigten sich ebenfalls signifikante Unterschiede zwischen teiltonreichen (Sägezahn und Rechteck) und teiltonarmen Schwingungen (Sinus und Dreieck) bei 7 von 10 Adjektivpaaren:

Sinus/Dreieck ↔ Sägezahn/Rechteck	t (bei p<0,05)
glatt, weich ↔ rau	11,61
tonhaltig ↔ geräuschhaft	10,174
homogen, gleichförmig ↔ heterogen, differenziert	8,185
natürlich, analog ↔ künstlich, digital	6,771
regelmäßig, geordnet ↔ unregelmäßig, chaotisch	6,675
dicht, flächig ↔ spärlich, punktuell	3,234
kantig, zerrissen ↔ rund, kompakt	-12,422

So werden teiltonarme (Sinus/Dreiecks-) Klänge überzufällig häufig als rund, kompakt, glatt, weich, tonhaltig, homogen, gleichförmig, natürlich, analog, regelmäßig, geordnet, dicht und flächig beschrieben, während teiltonreiche Klänge signifikant oft als kantig, zerrissen, rau, geräuschhaft, heterogen, differenziert, künstlich, digital, unregelmäßig, chaotisch, spärlich und punktuell beschrieben werden.

Interessanterweise lassen sich über die ja ebenfalls recht deutlichen Farbassoziationen zu den Klangbeispielen auch signifikante Zusammenhänge zwischen den Farbwerten der zugeordneten Farben und den Adjektivpaaren zur Klangfarbenbeschreibung herstellen:

RGB/HSV-Farbwerte korrelieren mit	Rot-Anteil (R)	Grün-Anteil (G)	Blau-Anteil (B)	Farbton (H)	Sättigung (S)	Helligkeit (V)
glatt, weich ↔ rau		-0,721***	-0,888***		0,565**	
tonhaltig ↔ geräuschhaft		-0,698***	-0,889***		0,553*	
homogen, gleichförmig ↔ heterogen, differenziert		-0,799***	-0,831***			
natürlich, analog ↔ künstlich, digital	0,563**		-0,897***		0,631**	0,567**
regelmäßig, geordnet ↔ unregelmäßig, chaotisch		-0,894***	-0,717***			
dicht, flächig ↔ spärlich, punktuell	0,796***		-0,728***	-0,649**		0,559*
kantig, zerrissen ↔ rund, kompakt		0,623**	0,892***		-0,586**	
hoch, hell ↔ tief, dunkel	-0,959***	-0,53*				-0,597**
statisch, starr ↔ dynamisch, bewegt	-0,693***	-0,739***				
nahe, klar ↔ fern, verschwommen	-0,801***					-0,566**

Zusammenhänge zwischen Adjektivpaaren und den zu den damit verbundenen Klangfarben assoziierten Farbwerten (es wurden nur Korrelation mit $r<-0,4$ bzw. $r>0,4$ einbezogen; * = $p<0,05$; ** = $p<0,01$; *** = $p<0,001$)

Entsprechend lässt sich für die Adjektiv-Zuordnung sagen:

- Je glatter, weicher, tonhaltiger, homogener, gleichförmiger, regelmäßiger, geordneter, runder und kompakter die Klangfarbenbeschreibung, desto mehr Grün- und Blau-Anteile besitzt die mit diesem Klang verbundene Farbassoziation.
- Je höher, heller, statischer und starrer die Klangfarbenbeschreibung, desto mehr Rot- und Grün-Anteile besitzt die mit diesem Klang verbundene Farbassoziation.
- Je näher, klarer, künstlicher, digitaler, spärlicher und punktueller die Klangfarbenbeschreibung, desto mehr Rot-Anteile besitzt die mit diesem Klang verbundene Farbassoziation (dies gilt ebenso für den Farbton (Hue), der sich immer weiter in den Rotbereich bei 0 verschiebt, je spärlicher und punktueller die Klangfarbe beschrieben wird).
- Je rauher, geräuschhafter, künstlicher, digitaler, kantiger und zerrissener die Klangfarbenbeschreibung, desto gesättigter ist die mit diesem Klang verbundene Farbassoziation.
- Je künstlicher, digitaler, spärlicher, punktueller, höher, heller, näher und klarer die Klangfarbenbeschreibung, desto heller ist die mit diesem Klang verbundene Farbassoziation.

Zusammenfassung

- In mittleren Tonhöhen werden teiltonreiche Klänge eher mit Orange-/Rottönen assoziiert und teiltonarme eher mit Grün-/Blautönen.
- Je heller die Klangfarbe, desto höher der Rot-Anteil, die Sättigung und die farbliche Helligkeit und desto geringer der Grün- und Blau-Anteil sowie der Farbton (Hue)
- Je geräuschhafter und rauher die Klangfarbe, desto stärker der Rot-Anteil, die Sättigung und die Helligkeit und umso geringer der Grün- und Blau-Anteil.
- Teiltonreiche und -arme Klänge können mit Hilfe von Adjektivlisten eindeutig beschrieben werden, die damit assoziierten Farben lassen sich ebenfalls auf diese Beschreibungen übertragen.