Proceedings of the 12th International Conference on Music Perception and Cognition and the 8th Triennial Conference of the European Society for the Cognitive Sciences of Music, July 23-28, 2012, Thessaloniki, Greece
Cambouropoulos E., Tsougras C., Mavromatis P., Pastiadis K. (Editors)

Beyond Helmholtz: 150 Years of Timbral Paradigms

Kai Siedenburg,*[1] Christoph Reuter,#[2]

* *Austrian Research Institute for Artificial Intelligence, Austria*
# *Musicological Institute of the University of Vienna, Austria*
[1]`kai.siedenburg@gmx.de`, [2]`christoph.reuter@univie.ac.at`

ABSTRACT

This article locates von Helmholtz's groundbreaking research on timbre and a few of its historical implications in terms of musical and mathematical coordinates. Through pinpointing on selected timbre-related examples it describes how music aesthetic ideals, mathematical theories and acoustics research systematically interdepend. After repositioning Helmholtz's work with respect to Fourier's theorem, two musical perspectives are considered, Schoenberg's vision of *Klangfarbenmelodie* and Xenakis's quest for sonic granularity. It is moreover suggested to regard a American National Standards definition as a late echo of Helmholtz's reign. The evolution of the multi-dimensional-scaling-based timbre space model is briefly outlined before observing a plurality of mathematic approaches which seems to mark current research activities in acoustics.

I. INTRODUCTION

In October 1862, Hermann von Helmholtz finished the first edition of his "*Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*" (*"On the Sensations of Tone as a Physiological Basis for the Theory of Music"),* the date as indicated in the foreword; the book was officially published one year later. The work, certainly the most influential contribution to music psychology and acoustics of the 19th century, initiated the modern science of the perception and physics of Klangfarbe - a concept that was preceded by the notion of timbre[1] in the anglo-american and french literature.

Then what does it refer to, this tone color, this peculiar quality of tone? At an early stage of his investigation, Helmholtz states

"The quality of tone [Klangfarbe] depends on the form of vibration". (Helmholtz, 1875, p. 32)

Although this statement might seem of minor novelty, it summarizes much of what timbre research generally deals with, namely analyses of forms of acoustic vibrations and its perceptual consequences. Such analyses are conducted automatically and without effort by the hearing apparatus each and every moment. However, acousticians require mathematical tools and theories, with which the manifold differences of these forms of vibrations can be appropriately approached. In the case of Helmholtz, the application of Fourier's theorem, originally conceived for the study of heat transfer, turned out to be a promising choice and an enormously successful turn for the scientific study of timbre.

In 1822 Fourier had published his Théorie analytique de la chaleur, which meant a breakthrough for 19th century mathematical physics. One of its contributions was the conjecture that any periodic function can be represented as an infinite series of trigonometric functions. The Fourier series expansion of a periodic signal f then can be written as

$$f(t)=a_0+\sum_{n=1}^{\infty} a_n\cos(nt)+b_n\sin(nt)$$

or in more modern but equivalent terminology using complex exponentials

$$f(t)=\sum_{n=1}^{\infty} c_n e^{2\pi i n t}$$

Fourier's theorem had great impact on 19th and 20th century science - it is not an overestimation to consider it as a methodological paradigm in the sense of Kuhn (1954). Georg Simon Ohm was the first to apply it to acoustics, (cf. Muzzulini, 2006). As it will be discussed in detail later on, Helmholtz extended its influence with respect to the description of sound to the physical and physiological domain, that is, he showed that it serves as a powerful mathematical model for acoustic vibrations and hearing. However, his On the sensations... was eminent enough to become a paradigm in acoustics itself.

Today, we face the situation that timbre is a controversial concept and commonly considered as the least understood basic aspect of music-theory and psychology (see e.g. Krumhansl, 1989; McAdams 2009). It thus seems worth to recapitulate parts of the notion's trajectory of 150 years. Without aiming for a comprehensive history of post-Helmholtzian timbre research, this article rather focusses on the discussion of a few examples which appear to be characteristic for a more general evolution.

Two hypotheses will form the golden thread of these elaborations. Firstly, it will be argued that timbre research since Helmholtz is realized in an interdepending framework of acoustic, musical and mathematical thought. Secondly, the musical approach towards timbre has gone through major changes, i.e. the portions of the "form of vibration" which are musically relevant are not the same as in Helmholtz's time any more. Therefore, after revisiting parts of Helmholtz's initial contribution, two sections featuring music-theoretical speculations of the composers Arnold Schoenberg and Iannis Xenakis shall contribute to the musical point of view. By briefly relating Helmholtz's work to more recent developments as Grey's timbre space model, its refinements as well as to current developments, it is hoped to outline a radiant picture of 150 years of timbral paradigms.

---

[1] Of course, there are differences in the connotation of *timbre, Klangfarbe* and *the quality of tone,* the latter being the 1875 suggestion of the translator Alexander Ellis, cf. Muzzulini (2006). However, these are of minor impact for the purpose of our discussion and we will hence assert a synonymic relation of these terms.

II. “ON THE SENSATIONS OF TONE”

Herrmann von Helmholtz leaves a grand scientific heritage. His *Die Lehre von den Tonempfindungen als physiologische Grundlage der Musik* [*On the sensations of tone as a physiological basis for the study of music*][2] (Helmholtz, 1863) turned out be one of the most influential works on acoustics and music perception ever written. Besides its research accomplishments in terms of revealing new insights about acoustics and human physiology and perception, the book featured three major methodological novelties which would turn out to become significant paradigm shifts of the scientific study of music of his day, (cf. Reuter, 2011). Firstly, Helmholtz clearly took the stance of the empirical sciences to investigate basic musical and aesthetic questions. He thus confronted theories on musical aesthetics with inductive and reductionist reasoning. Furthermore, the book drew a comprehensive picture of physiological acoustics and music perception, thereby integrating results which used to be widely dispersed over various different articles. Thirdly, and perhaps most importantly, Helmholtz delivered a unique multidisciplinary work of research, which he himself dubbed as an *“unusual attempt to pass from natural philosophy into the theory of arts”*. (p. 8)

Fluently working with the most advanced concepts of Mathematics, Physics, Physiology and Musicology of his time, he gave a fine lesson in what has nowadays become common sense: certain questions require the consideration of different research perspectives. As Karl Popper put it 100 years later,

“We are not students of some subject matter, but students of problems. And problems may cut right across the borders of any subject matter or discipline.” (Popper, 1963, p.88) [3]

In terms of the basic structure of “On the sensations...”, the book tends to move from fundamental to complex acoustic phenomena. Its first part, *“On the composition of vibrations”*, introduces the concepts of overtones and tone-color and explicates the physiological basis of hearing and music perception. The second part entitled “On the interruption of harmony” moves on to explain phenomena as combinational tones, beats and to lay a foundation of the notion of consonance and dissonance. Finally, and most remote to Helmholtz's initial point of departure (namely physiology), he goes on to elaborate on the implications of his findings for *“The relationship of musical tones”*. This is intended to serve as nothing less than a scientific foundation of music theory and aesthetics.

The first part, comprising 6 chapters in which he detailedly lays out his theory of musical sound, thereby interweaving a mathematical, physical and physiological perspective on sound, will be of particular interest for our endeavor.

A. Helmholtz's *Klangfarbe*

At an early stage of the first chapter, Helmholtz recalls the text-book definition of the *“three points of difference in musical tones”* which are to be distinguished according to their *force or loudness*, *pitch or relative height* and *quality* [Klangfarbe]. Giving a preliminary (but as will be argued later, consequential) hint of what he means by the third phenomenon, he explicates

“By the quality of a tone [Klangfarbe] we mean that peculiarity which distinguishes the musical tone of a violin from that of a flute or that of a clarinet, or that of the human voice, when all these instruments produce the same note at the same pitch. We have now to explain what peculiarities of the motion of sound correspond to these three principal differences between musical tones.” (p. 16-17)

While the physical correlates of pitch and loudness were to be easily located in the parameters of the duration of the periodicity and its amplitude, this still left infinitely many parameters for the a sound's timbre to vary. However, the application of Fourier's theorem could significantly reduce this complexity:

“[...] these propositions of Fourier may be thus expressed for the purpose of their application to the theory of sound: Any vibrational motion of the air in the aural passages, corresponding to a musical tone, may be always, but for each case only in one single way, exhibited as the sum of a number of simple vibrational motions, corresponding to the partial tones of this musical tone.”(p. 52-53)

What was regarded as relevant for a sound's timbre was the way its partial tones, i.e. its amplitudes and phases, are configured. The infinite complexity of vibrational movement of musical sound was hence reduced by Fourier's theorem to a finite number of parameters.

The evidence he gathered through his experiments in sound synthesis and physiology led him to the conclusion that Fourier's theorem was not just a mathematical formalism, but beard physical and physiological reality. For instance, he demonstrated that by using resonators tuned to certain frequencies it was possible to filter out and amplify certain partial tones from a compound sound. That means, the partial tones were really there, not just mathematical fiction. From the physiological perspective, he observed that

“there must be different parts of the ear which are set in vibration by tones of different pitch and which receive the sensation of these tones.” (p. 215),

thus providing the physiological analogue of the idea of the ear as a frequency analyzer. In his quest for the reduction of complexity, he even went one step further:

“The quality of the musical portion of a compound tone [Klangfarbe] depends solely on the number and relative strength of its partial simple tones, and in no respect on their difference of phase.” (p. 184)

---

[2] In case confusion is impossible, we will skip the reference to Helmholtz book in the following, and cite the english translation (Helmholtz 1875) by Alexander Ellis while only referencing its respective page number.

[3] Despite the latter being obvious today, most academic institutions are still strictly organized in terms of disciplines. The multidisciplinary scientific study of music hence lives a scarce and scattered life while going under its many names of systematic/systemic/empirical/cognitive musicology, music psychology/technology/computing, (psycho)acoustics, auditory neuroscience, to name a few. Strictly speaking, Helmholtz thus cannot be considered as a founding father of an internationally established discipline (as his approach does not allow for such), but certainly of an enormously influential attitude towards interdisciplinary music research, (cf. Niemoeller, 2003).

This statement, perhaps the quintessential Helmholtzian approach towards musical timbre, locates the essence of timbral quality in the number and strength of a tone's partials. At the same time, phase information is discarded as being irrelevant for musical timbre, hence promoting the validity of Ohm's Law in acoustics.

This basic representation of timbre was then used to thoroughly characterize the timbre of musical instruments. For instance, while simple sinusoids were considered to have a smooth character without any roughness, tones without even overtones as those of the clarinet or the stopped organ pipes were described as hollow. The strong presence of partial tones higher than the sixth or seventh were moreover classified as leading to a cutting and rough appearance (p. 172-173).

B. “The musical portion”

Was this intended as a general theory of hearing? Interestingly, Helmholtz's object of study, “the musical portion of a tone”, was narrowly specified:

“[...] a musical tone strikes the ear as a perfectly undisturbed, uniform sound which remains unaltered as long as it exists, and it presents no alternation of various kinds of constituents.” (p.12)

In physical terms this meant

“The sensation of a musical tone is due to a rapid periodic motion of the sonorous body; the sensation of a noise to non-periodic motions.” (p. 13) And later on, he adds *“When we speak in what follows of a musical quality of tone, we shall disregard these peculiarities of beginning and ending, and confine our attention to the peculiarities of the musical tone which continues uniformly.”* (p. 109)

Helmholtz thus distinguished stationary musical sound from the non-stationary rest (noise) by means of definition. In fact, from todays perspective his musical tones seem like non-existent idealizations of the so-called *steady state* part of sounds. Why this severe constriction of generality?

An answer might be found in the triangular inter-dependence of acoustical, mathematical and musical thought: On the one hand, the focus on strictly periodic sounds was a necessary step for employing classical Fourier theory which could only analyse strictly periodic signals.

On the other, this definition of *musical* sound seemed acceptable, given the aesthetic ideals of the time and the circumstance that most of the sonic vocabulary could be considered as being (quasi-) harmonic, i.e. harmonic with a degree of idealization or with some negligible error. Helmholtz clearly expresses his sonic preferences when writing about instruments producing inharmonic overtones as kettledrums, bells, and cymbals.

“Such tones are applicable for marches and other boisterous music, principally intended to mark time. But for really artistic music, such instruments as these have always been rejected, as they ought to be, for the inharmonic secondary tones, although they rapidly die out, always disturb the harmony most unpleasantly [...]” (p. 119)

In other words, the aesthetic ideals of his time allowed for a very restricted definition of timbre without losing much of musical value.

These constraints had the peculiar consequence that parts of the acoustic energy that any instrument emits would have to be considered as musical per se, while others would need to be discarded as irrelevant noise, even when they contribute to timbral discrimination. Helmholtz recognizes this by admitting,

“[...] the tones of most instruments are usually accompanied by characteristic irregular noises, as the scratching and rubbing of the violin bow, the rush of wind in flutes and organ pipes, the grating of reeds. These noises, with which we are already familiar as characterizing the instruments, materially facility our power of distinguishing them in a composite mass of sounds.” (p. 101)

Still, these components were discarded as a phenomenon of secondary relevance, as they were incommensurable with the Fourier paradigm that Helmholtz was about to extend to the domain of acoustics and music perception.

III. MUSICAL INTERACTIONS

A. Klangfarbenmelodien

“Tone-color melodies (Klangfarbenmelodien)! How acute the senses that would be able to perceive them! How high the development of spirit that could find pleasure in such subtle things! In such a domain, who dares ask for theory!” (Schoenberg, 1978, p. 422)

These emphatic words constitute the closing passage of Schoenberg's influential *Harmonielehre*, the composer's major treatise on harmony. Of more than 400 pages of the english version, he dedicates the very last to the topic of *Klangfarbe* and its potential role in composition, noticing that *“[...] our attention to tone colors is becoming more and more active.”* (Schoenberg, 1978, p. 422).

Schoenberg's understanding of tone colors was deeply rooted in the Helmholtzian paradigm of timbre. Disregarding temporal morphologies (from today's point of view naturally inscribed in any instrumental timbre), he considered the essence of these colors to lay in the superposition principles of partials, (cf. Muzzulini, 2006). By dismissing the dominating role of the fundamental frequency he (somewhat diffusely) arrives at the conclusion that pitch is only one sub-dimension of tone color. In his own words, *“Pitch is nothing else but tone color measured in one direction.”* (Schoenberg, 1978, p.421)

Alfred Cramer recalls in a remarkable synthesis of Schoenberg's early atonal work that the turn of the 20th century notion of *Klang* was often used as synonym of *chord* (Cramer, 2002). The term thus naturally beared ambiguity, which became manifest in a conception of Klangfarbe as Schoenberg's. Cramer continues to argue that *Klangfarbenmelodie* should not exclusively be regarded as referring to successions of single instrumental timbres, but in a more general sense, as a *“progression of chords of varied formation not necessarily grounded in the harmonic series.”* (Cramer, 2002, p.2) This reading of Klangfarbenmelodie as a harmonic principle - although to a certain extend reducing its innovative spirit - fits well into the historical context, namely by conceiving chords as colors.

Curiously, Schoenberg's famous *5 Pieces for Orchestra* op. 16 (1909) already seems to bear explorations of such colorful harmonic progressions in its no. 3, firstly called also *Akkordfärbungen* (chord colorings), cf. (Jena 2011, p. 330), later (after 1925) *Farben - Sommermorgen an einem See* (colors - summer morning by a lake). In the beginning section, especially in the first four measures, the change of the pitch content in the sustained chords is reduced up to a minimum while its orchestration fluently shifts back and forth. Although scholars disagree whether this should be considered as a true realization of a Klangfarbenmelodie, see e.g. (Dahlhaus, 1989), the fact that the composer neither mentions this piece, nor discusses those of other composers[4] suggests that his enthusiasm of the introductory quote must have mainly been directed towards an underlying musical *logic* of patterns of tone colors, which he believes to be only developed in the future:

“Now, if it is possible to create patterns out of tone colors that are differentiated according to pitch, patterns we call 'melodies', progressions, whose coherence (Zusammenhang) evokes an effect analogous to thought processes, then it must also be possible to make such progressions out of the tone colors of the other dimension, out of that which we call simply 'tone color', progressions whose relations with one another work with a kind of logic entirely equivalent to that logic which satisfies us in the melody of pitches.” (Schoenberg, 1978, p. 421)

Where laid the roots of this famous vision? Was it yet another foresight of an ingenious innovator? Carl Dahlhaus presumed that Schoenberg had already read Helmholtz's “On the sensations...”, too similar were the two masters' hypotheses about the nature of consonance and tonality. Dahlhaus argued that Schoenberg shared two important Helmholtzian theorems, which contradicted the prevailing music-theoretical views of their times: Firstly, there naturally is a gradual but not a categorical difference between consonance and dissonance. Secondly, tonality is not a natural law of music, but a principle of musical form, (cf. Dahlhaus, 1989). He further speculated that Schoenberg's vision of a logic of Klangfarbenmelodien might have been derived from Helmholtz's elaborations on the feeling of melodic similarities which the latter had considered to be based upon shared partials of successive tones (Helmholtz, 1863, p. 556). On the other hand, Helmholtz had expressed his belief in the human capacity of developing a timbral sensitivities in listening (Helmholtz, 1863, p. 107). In this light, Schoenberg's vision of Klangfabenmelodie appears as a direct projection of Helmholtzian thought into a musical future.

Later on, the acousticians Meyer-Eppler et al. (1959) showed that it is possible to generate purposive melodies by only using sequences of timbres (or better formants). Hereby he moved the Schoenbergian aesthetic paradigm towards a physically tangible and mathematically describable perspective.

B. Sonic Granularity

Roughly half a century later after the writing of *Farben*, a new generation of composers was asking new musical questions. The composer Iannis Xenakis was one of them, and undoubtedly became one of the most influential figures in instrumental and electroacoustic music of the 20th century. He was trained as engineer and architect before fully dedicating his life to music. This is commonly considered to be of great impact for his work as a composer since he kept on “transposing” advanced scientific models to the musical domain.

It was Xenakis' second major instrumental composition *Pithoprakta* (1955-56) which featured a novel, previously unimaginable sonority. Its bars 52-59, even called the “historical measures” (Antonopoulos, 2011), consist of masses of thousands of pizicatti-glissandi, played by the string section in a duration of less than 20 seconds. Over this period, 46 string parts produce an average of 25 notes each.

The initial pitch, duration and glissando speed of each of these grains were determined by using the Maxwell-Boltzmann-distribution, modeling the speed of gas-particles in the context of statistical physics, cf. (Antonopoulos, 2011). Xenakis' thus composed a cloud of non-stationary sonic particles, realized through plucked glissandi of very short duration, depicting a snapshot of large numbers of discrete and independent air particles.

The piece exhibits a variety of other granular textures and non-stationary glissandi - stationary sounds are rare exceptions, in fact. Then what are the timbral qualities of the employed sounds regarded from a Helmholtzian perspective? How to define the spectral steady state of a tone which lasts not more than 200 ms without exhibiting constant pitch? Helmholtz's theory of Klangfarbe simply did not account for such acoustic events. Expressed in other words, *Pithoprakta* made the insufficiencies of Helmholtz's definition of timbre as well as classical Fourier-theory as a mathematical/acoustical foundation of music very clear.

Xenakis’ quest for non-stationarity was not limited to the compositional domain. After he had finished electroacoustic pieces employing tiny slices of recorded sounds (*Concrete PH*, 1958) and electronically generated sonic grains (*Analogique A/B*, 1959) (where both pieces employ techniques which later would be termed *granular synthesis*) he published a theoretical foundation of his granular vision in 1960 stating that

“All sounds represent an integration of corpuscles, of elementary acoustic particles, of sound quanta. Each of these elementary particles possesses a double nature: the frequency and the intensity (the life-time of each corpuscle being minimum and invariable).” (Xenakis, 1960)

This statement perhaps best summarizes Xenakis' granular ontology of sound which refutes the 100 year-lasting reign of the Fourier-based paradigm of timbre. As an alternative he seems to suggest a quantum-based point of view, therewith

---

[4] Also compare Schoenberg's remarks on Webern's use of Klangfarbenmelodie (Schoenberg, 1984). He fully neglects to use the term for Webern's realizations and insists on an emphatic usage of the word, requiring its own musical form and type of organization.

transposing the wave-particle debate from physics to the musical arena. Giving a thorough account of the statement's historical trajectory, Solomos characterizes the latter aspect as the first constituents of a *granular paradigm* (Solomos, 2006), which originates in Xenakis' work but is elaborated by later 20th/21st century composers as Agostino Di Scipio or Curtis Roads. Here, Solomos uses the notion of a paradigm in a wide sense, encompassing both intuition and rationality, being a comprehensive aesthetic and scientific “vision of the world”.

Specifying this worldview, Xenakis notes (Xenakis, 1992, p. 244)

“Now, the more the music moves toward complex sonorities close to 'noise' the more numerous and complicated the transients become and the more their synthesis from trigonometric functions becomes a mountain of difficulties [...]. It is as though we wanted to express a sinuous mountain silhouette by using portions of cycles.”

These words were technically concretized in his writings on a *Markovian Stochastic Music Theory*, cf. (Xenakis, 1960), (Xenakis, 1992), (of which he considered *Analogique A/B* as applications). Here, Xenakis claimed that any sound could be represented as a collection of sinusoids of a short duration which were to be ordered along their dimensions of pitch height and amplitude on a *screen*. The temporal morphology of a sound could then be described in a *book*, i.e. a temporally ordered collection of screens. Being aware of the shortcomings of these initial conceptions[5], he confined the scope of the theory by writing (Xenakis, 1960)

“This description of the micro-structure of acoustic signals is used as the starting-point of the musical realization and must be understood rather as to be an intuitive representation than of scientific consistence. But it can be considered as to be a first approach towards the ideas introduced into the theory of information by Gabor.”

Explicitly referring to the physicist Denis Gabor, Xenakis' conception paralleled research the former had published in the late 1940ths. Gabor had claimed in his seminal paper *Theory of Communication* (Gabor, 1946) that any signal f could be expanded as

$$f(t)=\sum_{(n,m)\in\mathbb{Z}^2} c_{n,m}\, e^{2\pi imt}\, e^{-\pi(t-n)^2}$$

for appropriately chosen coefficients c. That is, any sound could now be interpreted as a superposition of translated and sinusoidally modulated Gaussian functions - arbitrary sound events could be formally described, analysed and synthesized, both in terms of their temporal *and* spectral behavior. Gabor's conjecture was finally proven in the 1970s, cf. (Grochenig, 2001, p. 142), and can be considered as the initial work of modern time-frequency (alternatively called Gabor-) analysis. The latter has become an evolving academic area, which for instance provides the mathematic foundations of acoustic signal processing.

For some reasons, Xenakis claimed later on that he had developed these ideas all by himself, while being departed from the work of Einstein, cp. (Solomos, 2006), and was hence denying any influence of Gabor's. However, regarding to his 1960 statement from above it seems rather plausible that Gabor's ideas found their way to Xenakis via the acousticians Moles or Meyer-Eppler, compare the discussions of (Roads 2004, DiScipio 2006, Solomos 2006).

Meyer-Eppler had been propagating the idea of considering both, time and frequency evolution of a sound already at least in 1955, (see Meyer-Eppler, 1955), and it is speculated that Xenakis had attended a lecture of his in the same year, cf. (Solomos 2006). Moreover, Xenakis explicitly references Meyer-Eppler in (Xenakis, 1960). He does so in *Formalized Music*, too, but the complete reference at the end of the book is missing, in fact. However the details of this historical trace exactly took place, this example again highlights the interdependence of musical, acoustical and mathematical thought, especially in a time which with Kuhn (1954) might be considered as a musico-scientific paradigm-shift.

IV. TIMBRE SPACES

A. Defining Timbre

Ironically, the portion of Helmholtz's theory of timbre that perhaps received the greatest international resonance goes back to a very preliminary statement of his whole work. In 1960, the American National Standards Institute summarized

“Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar.” (ANSI, 1960)

Notice how this definition paralleled the initial Helmholtzian statement (see above), given around 100 years before. The definition merely abstracts from Helmholtz's example of a violin, clarinet or flute to general instruments. In this light, it appears as a compactly phrased reference to *On the sensations*, or put in another way, as an artifact of the 100 year-lasting Helmholtzian paradigm of timbre.

One of the harshest refutations of this definition came from Albert Bregman, the famous researcher of human audition, with his work on *Auditory Scene Analysis* (Bregman, 1990) commonly considered to be the founder of a new subdiscipline of hearing research. In this very work he writes,

“The problem with timbre is that it is the name for an ill-defined wastebasket category.” After quoting the ANSI definition from above he continues, *“This is, of course, no definition at all. For example, it implies that there are some sounds for which we cannot decide whether they possess the quality of timbre or not.”* He concludes by stating *“I think the definition of the American National Standards Association [sic] should be this: 'We do not know how to define timbre, but it is not loudness and it is not pitch.' ”* (Bregman, 1990, p. 92-93)

---

[5] Technically speaking, his writings implied the use of rectangular window functions in the Short-Time Fourier-Transform, which are known to have terrible properties of time-frequency localization. Also, using a seperate amplitude axis, allowing for grains with the same frequency having different amplitudes, does not seem to be very useful as this makes the representation non-unique and grains of the same frequency would add up in hearing, anyway.

These authoritative words ridiculed the definition-by-negation and the musical conservatism of the ANSI statement, as it could not account for the timbre of unpitched instruments and was unable to name any concrete physical quantities relating to timbre perception. Moreover, the refutation was corresponding to a musical environment that had changed significantly – now it seemed completely counter-intuitive to disqualify the important class of unpitched instruments as not having a timbral quality, contrary to Helmholtz's narrow notion of the "musical portion" of a tone.

Notwithstanding, what perhaps is most remarkable about this attempt of defining timbre is not that an institution like the ANSI had possibly made an erroneous or incomplete conjecture, but that there still are as many research papers on hearing and acoustics using the ANSI definition as an accepted (or possibly implicit) starting point of their investigations as there are papers employing its refutation as a point of departure. Does there simply not exist a better general definition of comparable compactness? The aspects of the infinite dimensions in which signals can differ, i.e. the aspects of timbre which are of musical relevance were simply not as clear any more as in Helmholtz's time. The trajectory that the ANSI definition took thus displays the lack of certainty that the acoustics community faces when dealing with timbre. This uncertainty appears to be intimately bound to the tension between the persistence and the constraints of Helmholtz's heritage of Klangfarbe.

B. The Timbre Space Model

Bregman's timbral skepticism becomes even more remarkable when it is confronted with the substantial amount of research which had already been conducted on the physical correlates of timbre perception in the (English) research literature. A particularly successful model of timbre perception has been the *timbre space* as promoted in John Grey's Stanford dissertation (Grey, 1975). The model made use of the newly developed (and in the meanwhile well-known) statistical method of *multidimensional scaling* (MDS) (Shepard, 1962) which takes a pairwise dissimilarity matrix of stimuli as input and generates a spatial representation of these stimuli which represents the original distance relations. The interpretation of the resulting spatial dimensions then remains a task for the individual researcher.

Grey used this method to construct a spatial configuration of timbre stimuli from scratch, so to say, that is from plain pairwise dissimilarity judgements of subjects which did not presuppose any kind of semantic categories for the judgement. His stimuli set consisted of 16 instrument timbres, all at the same pitch (E-flat 4) and loudness, re-synthesized via additive synthesis using simple line-segment approximations of the partials' time-variant amplitudes. Thus, these sounds were in fact artificial stimuli, based on a filter-bank analysis (alternatively: STFT/Gabor transform) of real instruments. Grey obtained a three-dimensional MDS solution. He interpreted the dimensions, i.e. the physical correlates of timbral discrimination, as being related to temporal features, the spectral energy distribution, as well as spectro-temporal features as the synchronicity of a sound's partials during its onset (Grey, 1975).

The MDS-based timbre-space model, founded on the assumption of a continuous cognitive representation of timbre with a few underlying dimensions, has been refined significantly by various researchers, both in terms of the employed stimuli and the flexibility of the underlying statistical method. Perhaps the most refined account of this type is the study of McAdams et al. (1995) which derives an MDS solution with three dimensions in addition to a degree of specificity for each stimuli, i.e. a measure of the acoustic uniqueness of a stimuli that is not accounted for by the other continuous perceptual dimensions. In this way instrument specific features as e.g. the return of the hopper of the harpsichord can be integrated into the continuous spatial model (McAdams, 1995). Here and in most recent MDS studies on timbre, the resulting dimensions were interpreted in terms of the three physical dimensions of spectral energy distribution, temporal evolution and spectro-temporal behavior. Highly correlating audio-descriptors were identified as the *spectral centroid*, indicating the point of gravity of the spectrum, the (*log*) *attack-time,* referring to the rapidity of a sound's onset, and the *spectral flux*, being a measure of the successive spectral differences of a sound over time.

These results outline a perceptual model of timbre which is symmetric in time and frequency. This matches the developments in the mentioned mathematical discipline of time-frequency analysis, the formal foundation of any modern signal analysis and e.g. the STFT. Here signals are not considered exclusively in terms of time *or* frequency but as *inhabitants* of a time-frequency plane, where both dimensions interact with "equal rights", see e.g. (Gröchenig, 2001).

The timbre space model has been repeated by different studies using behavioral measures, as e.g. in the learning of timbral relations, cf. (Krumhansl, 1989; McAdams, 1992; McAdams, 1999; Tillmann and McAdams, 2004). Furthermore, the three dimensions of the timbre space seem to be separately represented in auditory memory, as neurobiological research, using the mismatch negativity component of auditory event-related potentials, suggests (Caclin et al., 2006).

Despite this noticeable progress in the elaboration of the timbre space model, it is far from being commonly accepted as a sufficient description of the phenomenon of timbre (Reuter, 1998; Reuter 2002). Major points of criticism concern the constrained generality of the experimental setting (comparing mostly artificial stimuli, only on the same pitch, only in one dynamic level, in an untypical instrument register, only one articulation etc.), and in particular the circumstance that the spectral dimension represented via the spectral centroid can hardly account for formants. These are found to be crucial for the timbre of brass instruments, for instance, while being only sufficiently represented by taking into account the relation of 2-3 spectral maxima, i.e. not only one central point of gravity.

More generally comparing Anglo-Saxon and German timbre research, the main dilemma lies in the following circumstances: Even though Helmholtz was aware of the starting transients and typical accompanying noises of musical instruments he did not believe in their relevance concerning timbre perception as noted above (Helmholtz, 1863, p. 102, 114-118). It was one of his successors Carl Stumpf who observed the timbre of vowels and musical instruments not only in a narrow (Helmholtzian) sense (i.e. only the combination of the partials and their amplitudes) but also by taking account of starting transients, typical noises etc. (Stumpf, 1890, p. 390). Furthermore Stumpf suggested that the generation and perception of timbral properties also depends on the chosen pitch and dynamics. For this reason his investigations always employed not only one pitch/dynamic level but many. His students (e.g. Wolfgang Köhler (1909), Gerhard Albersheim (1939), Carl Erich Schumann (1929) and also their pupils like Hans Peter Reinecke (1953), Jobst Fricke (1965) etc.) were very aware of these circumstances. While Helmholtz' "*Lehre von den Tonempfindungen"* had been translated into English by Alexander J. Ellis, a translation of Stumpf's standard works "*Tonpsychologie"* (1883 and 1890) and *"Die Sprachlaute"* (1826) is still missing, however, as well as the translation of one of the books of other German-speaking successors of Helmholtz and Stumpf.

Consequently, the English research literature which has been most influential on this field since World War II is lacking this particular aspect of knowledge of musical timbre. Its salience can be immediately observed by examining an instrument in all its pitch and dynamic levels (instead of only one or a few pitches): timbre perception of musical instruments is determined to a certain extend by steady pitch-independent formant areas (Stumpf, 1926, S. 382 ("Nebenformanten"), Schumann, 1929; Fricke, 1975), whose origin and characteristics at pitch and dynamic changes has been resolved today to a great extent (Fricke, 1975, Voigt, 1975, Auhagen, 1984 Reuter, 1995). These timbral effects appear to be in accordance with the principles of the Auditory Scene Analysis (Bregman, 1990) as well as with recommendations in the orchestration treatises of the last centuries (Reuter 2000; Reuter 2002). The latter remark moreover highlights an example of the influence of language (barriers) on the evolution of timbre research.

V. CONTEMPORARY PERSPECTIVES

A. Beyond Fourier: A Thousand Ways to Transform a Signal

The timbre space model has not been established as *the* exclusive model of timbre perception of the late twentieth century. Nor stands its technological basis, time-variant signal analysis and synthesis at the end of its evolution.

Today, we rather encounter a multitude of different competing methods for signal transformation which are based on increasingly flexible mathematical frameworks. Instead of trying to prove in Helmholtzian manner that certain mathematical models are uniquely suited for representing sound, research on mathematical signal processing nowadays often faces a trade-off between psychoacoustic realism and mathematical flexibility. Naturally, the closer a model is fitted to account for the properties of hearing and perception the more complicated becomes its mathematical treatment, normally. The modern playground of signal transforms includes notions like *wavelet-bases, non-stationary Gabor-transforms, cascaded filterbanks, sparse representations, dictionary learning,* etc. (see Mallat, 2009 for an introduction).

A time-frequency representation which features exactly such a trade off is the so called *Constant Q Transform* (CQT), originally introduced by Brown (1991). It features a geometric instead of a linear frequency resolution and therewith particularly well matches human perception. Moreover, it can be configured to perfectly fit to the 12 half-tones of the western musical scale, i.e. each point in the representation corresponds to a half tone. For these reasons, it has been of particular interest for the audio processing community recently (which is at the same time struggling with some of its mathematical limitations, namely the missing matrix structure of its time-frequency coefficients). Progress in the formal understanding of time-frequency transforms has even improved the practical implementation of the CQT, as a newly developed approach allows for processing without reconstruction error as well as accelerated computation (Velasco, 2011).

B. Echos from Schoenberg

Naturally, it did not take long that musicians exploited and developed related these novel mathematical ideas for their creative purposes. The piece *A Letter from Schoenberg* (2006) by the Austrian composer Peter Ablinger, produced roughly a century after Schoenberg's *Farben* (1909), is an example of an ingenious musical application of a constant-Q-type filterbank with geometric frequency resolution. Written as part of his cycle *Quadraturen* and developed in collaboration with the Intitute of Electroacoustic Music and Acoustics (IEM) in Graz, the piece takes an original recording of a letter dictated by Arnold Schoenberg and re-composes it for an automated piano (i.e. a real piano played by a computer-driven mechanism). The voice is analyzed, the analysis is slightly processed and the generated representation, in which each coefficient or "pixel" (as Ablinger calls it himself) refers to one half-tone, re-orchestrated for a very precise computer-driven piano (Ablinger, 2012). The words which are instrumentally synthesized are additionally provided as text via video screening. A perceptually intriguing piece of music is the result: Without the text/video, the speech seems very difficult to understand. With the help of the text, however, one can easily construct the correspondence between sounds and words. Moreover, one clearly has the impression that Schoenberg is recognizable, the spectro-temporal morphology of his voice, even his peculiar German accent speaking in the English language. Despite the strong vocal resemblance, the listener hears that the sounds are produced by a piano. As Douglas Barret (2009) wonders

> "*The listener is caught up in the task of trying to place the source of this ghostly apparition: if not from the piano, from just where does Schoenberg speak? And by what mechanism are we able to hear Schoenberg's voice?"*

Parts of the spectro-temporal structure of *A Letter from Schoenberg* clearly refer to Schoenberg's voice, while other cues are without doubt identifiable as stemming from a piano. This timbral ambivalence gives rise to a peculiar perceptual tension, which in fact musically questions a seemingly well defined relation between continuous sonic qualities and the categorical referentiality of musical timbre.

Needless to say, Ablinger is not the only musician playing with such new tensions which have become possible through novel tools in signal analysis and transformation. In an attempt at "letting an orchestra speak", the composer Jonathan Harvey collaborated with researchers from the *Institute Recherche et Coordination Acoustique / Musique (IRCAM)* in Paris. A special software tool was developed suggesting combinations of instrumental sounds which as a whole closely resemble a given initial sound recording (Carpentier 2010). This tool which is still being further developed and was used as a technological basis of Harvey's 2008 piece *Speakings* (Nouno, 2009), again a piece exploring a timbral field between speech and instrumental music.

C. Closing Words

This article located Helmholtz's groundbreaking research on timbre and a few of its historical implications in terms of musico-mathematical coordinates. Through pinpointing on selected timbre-related musical and scientific examples it was described how aesthetic ideals, mathematical theories and acoustics research systematically interdepend.

After repositioning Helmholtz's work with respect to Fourier's theorem, two musical interactions were considered, Schoenberg's vision of *Klangfarbenmelodie* and Xenakis's quest for sonic granularity. It was moreover suggested to regard the 1960 ANSI definition as a late echo of Helmholtz's reign. The evolution of the MDS-based timbre space model was briefly outlined before observing a plurality of mathematic approaches which seems to mark current research activities in acoustics.

Pierre Schaeffer remarks that

"Musical ideas are prisoners, more than one might believe, of musical devices." (Schaeffer, 1977)

It seems reasonable to rephrase this statement and conclude by stating that *research in acoustics is contingent, more than one might believe, upon mathematical concepts and aesthetic ideals.*

ACKNOWLEDGMENT

This work was supported by the Austrian Science Fund (FWF) : T384-N13 and the WWTF project Audio-Miner (MA09-024).

REFERENCES

Ablinger, P. (2012, April 15). A letter from Schoenberg. Retrieved April 23, 2011, from http://ablinger.mur.at/txt_qu3schoenberg.html.

American Standarts Association (1960), *American Standarts Association Terminology* (including mechanical shock and vibration), New York.

Antonopoulos, A. (2011). Pithoprakta: The historical measures 52-59 new evidence in glissando speed formalization theory & theoretical applications. In *Proceedings of the Xenakis International Symposium*, Southbank Centre London.

Auhagen, W. (1987) Dreiecksimpulsfolgen als Modell der Anregungsfunktion von Blasinstrumenten. In *Fortschritte der Akustik, 13. DAGA'87,* (pp. 709–712), Aachen.

Bregman, A.S. (1990). Auditory Scene Analysis: The Perceptual Organization of Sound. MIT Press.

Brown, J.C. (1991) Calculation of a constant q spectral transform. *Journal of the Acoustical Society of America,* 89(1), 425-423.

Barrett, G. (2009). Between noise and language: The sound installations and music of peter ablinger. *Mosaic: a Journal for the Interdisciplinary Study of Literature,* 42(4).

Cramer, A. (2002). Schoenberg's Klangfarbenmelodie: A principle of early atonal harmony. *Music Theory Spectrum,* 24(1), 1-34.

Dahlhaus, C. (1989). *Schoenberg and the new music.* Cambridge University Press.

Caclin, A., Brattico, E., Tervaniemi, M., Naatanen, R., Morlet D., Giard, M.-H., & McAdams, S. (2006). Seperate neural processing of timbre dimensions in auditory sensory memory. *Journal of Cognitive Neuroscience*, 18(12), 1959–1972.

Carpentier, G., Tardieu, D., Harvey, J., Assayag, G. & Saint-James, E. (2010). Predicting timbre features of instrument sound combinations: Application to automatic orchestration. *Journal of New Music Research,* 39(1), 47–61.

DiScipio, A. (2006). Formalization and intuition in analogique a et b. *Definitive Proceedings of the "International Symposium Iannis Xenakis"* (Athens, May 2005).

Gabor, D. (1946). Theory of communication. *Journal of the Institution of Electrical Engineers Part III*, 93(26), 429–457.

Grey, J.M (1975) *An Exploration of Musical Timbre*. PhD thesis, Stanford University.

Groechenig, K. (2001). *Foundations of Time-Frequency Analysis.* Birkhauser.

von Helmholtz, H. (1963). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik.* Friedr. Vieweg & Son.

von Helmholtz, H. (1875). *On the sensations of tone as a physiological basis for the study of music.* Longmans.

Jena, S. (2011), Klangfarbenmelodie - Anmerkungen zu einer Utopie und ihrer Geschichte. In A. Schmidhofer, & S. Jena (Eds.): *Klangfarbe - Vergleichend-systematische und musikhistorische Perspektiven*. Peter Lang, pp. 329-341.

Kuhn, T. (1996, 1954). *The structure of scientific revolutions.* University of Chicago Press.

Krumhansl, C. (1989). Why is musical timbre so hard to understand? In: S. Nielsen & S. Olsson (Eds.), *Perception of electroacoustic sound and music.* Elsevier.

Mallat, S. (2009) *A Wavelet Tour of Signal Processing - The Sparse Way*. Academic Press.

McAdams, S. & Giordano, B. (2009). The perception of musical timbre. In: *Oxford Handbook of Music.* Oxford University Press.

McAdams, S. (1999), Perspectives on the Contribution of Timbre to Musical Structure, *Computer Music Journal* 23/3, 85–102.

McAdams, S., Windsberg, S., Donnadieu, S., de Soete, G. & Krimphoff, J. (1995). Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes, *Psychological Research* 58, 177-192.

McAdams, S. & Cunible, J.-C. (1992). Perception of timbral analogies. *Philosophical Transactions of the Royal Society: Biological Sciences*, 336, 383-389.

Meyer-Eppler, W., Sendhoff, H. & Rupprath, R. (1959) Residualton und Formantton. *Gravesaner Blätter* 14, 70–83.

Meyer-Eppler, W. (1955). Sichtbar gemachte Musik. *Gravesaner Blaetter* 1, 27–31.

Muzzulini, D. (2006). *Genealogie der Klangfarbe.* Peter Lang.

Nouno, G., Cont, A., Carpentier, G., & Harvey, J. (2009). Making an orchestra speak. In *Proceedings of the Sound and Music Computing (SMC'09) Conference,* (pp. 277–282), Porto, Portugal.

Popper, K. R. (1963). *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge and Kegan Paul.

Reuter, C. (2011). Akustische Fundierung der Musiktheorie: Kommentar zu Helmholtz Lehre von den Tonempfindungen. In W. Auhagen, V. Busch & J. Hemming (Eds.), *Kompendium Systematische Musikwissenschaft*, (pp. 51–59), Laaber.

Reuter, C. (2002). Klangfarbe und Instrumentation. Peter Lang.

Reuter, C. (2000). Verschmelzung und partielle Verdeckung. *In Fortschritte der Akustik, DAGA 2000* (pp. 176–177), Oldenburg.

Reuter, C. (1998). Wie und warum in der Geschichte der Klangfarbenforschung meistenteils am Klang vorbeigeforscht wurde. In Niemoeller (Ed.): *Perspektiven und Methoden einer Systemischen Musikwissenschaft. Bericht ueber das Kolloquium im Musikwissenschaftlichen Institut der Universitaet zu Köln* (pp. 293-301), Peter Lang.

Reuter, C. (1995) *Der Einschwingvorgang nichtperkussiver Musikinstrumente*. Peter Lang.

Roads, C. (2004). *Microsound*. The MIT Press.

Schoenberg, A. (1978, 1911) *Theory of Harmony (Harmonielehre)*. University of California Press, Wien.

Schumann, K.E. (1929) *Physik der Klangfarben*, Berlin.

Shepard, R. (1962) The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika*, 27(2), 125–140.

Solomos, M. (2006). *The granular connection* (Xenakis, Vaggione, Di Scipio...). Unpublished manuscript.

Stumpf, C. (1890) *Tonspsychologie*, Band 2, Hirzel.

Stumpf, C. (1926) *Die Sprachlaute*, Springer.

Tillmann, B. & McAdams, S. (2004). Implicit learning of musical timbre sequences: Statistical regularities confronted with acoustical (dis)similarities. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 30(5), 1131–1142.

Velasco, G, Holighaus, N., Dörfler, M. & Grill, T. Constructing an invertible constant-q transform with nonstationary gabor frames. *Proceedings of DAFX11 Paris.*

Voigt, W. (1975) *Untersuchungen zur Formantbildung in Klängen von Fagott und Dulzianen*. Bosse.

Xenakis, I. (1992) *Formalized Music: Thought and mathematics in composition*. Pendragon Press.

Xenakis, Iannis (1960). Elements of stochastic music. *Gravesaner Blaetter*, 18, 84–105.