SysMus09

November 18-20 | Ghent | Belgium

Proceedings of the Second International Conference of Students of Systematic Musicology

Edited by Luiz Naveda

IPEM – Institute for Psychoacoustics and Electronic Music
Ghent University

Blandijnberg 2, Ghent, B-9000, Belgium | Tel: +32 (0)9 264 4126
Fax: +32 (0)9 264 4143 | http://www.ipem.ugent.be/sysmus09

ISBN:

Luiz Naveda (Ed.)

SysMus09

Proceedings of the Second Conference of Students of Systematic Musicology

IPEM – Institute for Psychoacoustics and Electronic Music

Ghent University

November 18-20, November, 2009, Ghent, Belgium.

Blandijnberg 2, Ghent, B-9000, Belgium | Tel: +32 (0)9 264 4126 | Fax: +32 (0)9 264 4143 | http://www.ipem.ugent.be/sysmus09

Preface

SysMus09 is the Second International Conference of Students of Systematic Musicology – a conference for PhD students, entirely organized by PhD students in the field of systematic musicology. SysMus09 is hosted by IPEM - Institute for Psychoacoustics and Electronic Music, Department of Musicology (Ghent University, Belgium) at the Bijloke Music Center from 18 to 20 November, 2009. The initiative, which was launched by the Department of Musicology at University of Graz (Austria) in 2008, aims at connecting students and researchers from the broad spectrum of music research fields to envisage the forthcoming generation of research in systematic musicology.

The focus of this conference is on one of the major challenges for researchers in systematic musicology, namely, how to deal with the inter-disciplinary nature inherent to our own field. In order to promote this question, the committees have been engaged in promoting a conference that is marked by academic independence, a large interdisciplinary span and an international scope. The review committee, composed of advanced PhD students and post docs, was responsible for 93 anonymous reviews and for the proposition and selection of keynotes. These efforts lead to this volume of proceedings, which includes 18 papers (66% of acceptance) and 27 authors representing 17 countries. The diversity of areas is represented by topics as diverse as music cognition, music performance, cross-cultural studies, acoustics, psychoacoustics, and music and movement.

On behalf of SysMus09, we wish to thank our keynote speakers, professors Ian Cross (University of Cambridge) and Marcelo Wanderley (McGill University), as well as professor Marcel Brass (Ghent University) for their contribution to our conference programme. We also wish to thank our programme chairs, senior advisors and Manuela Marin for the support and care about the concept of the conference.

Luiz Naveda

Ghent, November 2009

Organization

Conference Chair

Luiz Naveda, *Ghent University*

Organizing Committee

Aagje Lachaert, *Ghent University*
Alexander Deweppe, *Ghent University*
Bart Moens, *Ghent University*
Chia-Fen Wu, *Ghent University*
Chris Müller, *Ghent University*
Denis Amelynck, *Ghent University*
Edith Van Dyck, *Ghent University*
Frank Desmet, *Ghent University*
Ivan Schepers, *Ghent University*
Katrien Debouck, *Ghent University*
Katty Kochman, *Ghent University*
Leen De Bruyn, *Ghent University*
Luc Nijs, *Ghent University*
Micheline Lesaffre, *Ghent University*
Michiel Demey, *Ghent University*
Nuno Diniz, *Ghent University*
Pieter Coussement, *Ghent University*
Pieter-Jan Maes, *Ghent University*

Senior advisors

Marc Leman, *Ghent University*
Mark Reybrouck, *K.U. Leuven*
Dirk Moelants, *Ghent University*
Leon van Noorden, *Ghent University*

Review Committee

Alexandre Bouënard, *McGill University*
Jonna Vuoskoski, *University of Jyväskylä*
Kaarina Marjanen, *University of Jyväskylä*
Kaisu Krohn, *The University of Helsinki*
Kristian Nymoen, *University of Oslo*
Kurt Jacobson, *Queen Mary University of London*
Malte Münster, *Universität Hamburg*
Manuela Marin, *University of London*
Marc Thompson, *University of Jyväskylä*
Martin Morgenstern, *Universität Halle-Wittenberg*
Michael Hove, *Cornell University*
Nicolau Werneck, *USP - Brazil*
Olivia Ladinig, *Universiteit van Amsterdam*
Rafael Ferrer Flores, *University of Jyväskylä*
Simon Van Damme, *K.U.Leuven*
Tal-Chen Rabinowitch, *University of Cambridge*
Vinoo Alluri, *University of Jyväskylä*

Table of Contents

Conference papers

Sonification of a Quantum Mechanical System inspired by a Quantum Computation Model

Anna Saranti
Institute of Electronic Music and Acoustics
University of Music and Dramatic Arts Graz
Inffeldg. 10/3, 8010 Graz, Austria
anna.saranti@student.kug.ac.at

David Pirrò
Institute of Electronic Music and Acoustics
University of Music and Dramatic Arts Graz
Inffeldg. 10/3, 8010 Graz, Austria
pirro@iem.at

Gerhard Eckel
Institute of Electronic Music and Acoustics
University of Music and Dramatic Arts Graz
Inffeldg. 10/3, 8010 Graz, Austria
eckel@iem.at

ABSTRACT

In this paper, we describe the sonification of a quantum mechanical system and the processes that occur as a result of its quantum mechanical nature and interactions with other systems. The quantum harmonic oscillator is not only regarded as a system with sonifiable characteristics but also as a storage medium for quantum information. By representing sound information quantum mechanically and storing it in the system, every process that unfolds on this level is inherited and reflected by the sound.

Keywords

Sonification, shapelet

1. INTRODUCTION

The quantum harmonic oscillator is one of the most fundamental quantum mechanical systems. It describes as in classical mechanics the motion of an object subjected to a parabolic potential [1, pp. 54–63]. As every other quantum mechanical system it is described by its Hamiltonian, which for this system is solvable with known eigenstates and eigenvalues. Any state of the system can be expressed as a superposition of its eigenstates.

The quantum harmonic oscillator provides a physical realization of a quantum computer model [2, pp. 283–287] where quantum information is stored in the state of the quantum harmonic oscillator and then processed through its intrinsic time evolution or through coupling with the environment. The sonification choices that were adopted in this work could also be associated with these information processing operations.

At a first step sound information is stored quantum mechanically in the system's state. Letting the system evolve in time or interact with other systems affects the state and thereby the stored information. The deformation of the stored sound reflects the characteristics and properties of the system and the processes that occur. In the cases where the eigenvalues and eigenstates are affected, their sonification could also add more insight to the phenomena.

The motivation for this approach is to gain a first insight to quantum computational storage operations through sound. Quantum mechanical memory has in general different properties from the classical [2, pp. 13–17], which can be highlighted through sonification. The impact of an external disturbance to the stored quantum information is a fairly complex procedure with interdependencies that can be perceived coherently through sound.

The part of the stored quantum information which is classically accessible through quantum measurement and the impact of the measurement operations in the classically retrieved part can be also acoustically represented with the use of this approach. The best known model of a quantum mechanical memory unit is the qubit [2, pp. 13–17] which is abstract and unbounded from the properties of the physical system that realizes it. The harmonic oscillator quantum computer model bases on the features of the underlying system and therefore the representation of the quantum information is directly interconnected with the system properties.

2. QUANTUM HARMONIC OSCILLATOR

2.1 Description of the system

Every quantum mechanical system's total energy is described by its Hamiltonian $\hat{H}$. Leaving the time evolution of the system aside and concentrating on the description of the system for a specific time point, the time-independent Schrödinger equation is [1, pp. 19–32]:

$$\hat{H}\psi(x) = E\psi(x) \quad (1)$$

where $\hat{H}$ is the Hamiltonian of the system, $\psi(x)$ the wavefunction that represents the state of the system and E the eigenvalues of $\hat{H}$. The value of $|\psi(x)|^2$ expresses the probability density of finding the oscillating object at the position x [3, pp. 54–57]. The eigenvalues that satisfy the equation (1) are quantized and represent the eigenenergies of the quantum harmonic oscillator:

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega, \quad n = 0,1,2,\ldots \quad (2)$$

The eigenstates that satisfy the equation (1) are mathematically expressed with the help of the Hermite Polynomials $H_n(x)$:

$$H_n(x)=(-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}, \quad n=0,1,2,\ldots \quad (3)$$

The eigenstates $\psi_n(x)$ which satisfy the equation (1) are weighted Hermite polynomials and represent the eigenfunctions of the quantum harmonic oscillator:

$$\psi_n(x)=\left(\frac{a}{\sqrt{\pi}\, n!2^n}\right)^{\frac{1}{2}} H_n(\alpha x)\, e^{\frac{-\alpha^2}{2}x^2}, \quad \alpha=\sqrt{\frac{m\omega}{\hbar}} \quad (4)$$

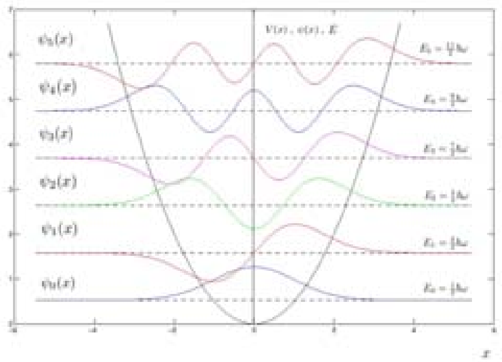


Figure 1: The first six (from 0 to 5) Eigenenergies and Eigenfucntions of the Quantum Harmonic Oscillator. The Eigenergies are depicted with equally spaced vertical lines. The corresponding Eigenfunctions are shifted on the y-axis each one with offset the particular Eigenenergy. On the same plot is also the parabolic potential V.

The eigenfunctions of the quantum harmonic oscillator constitute a complete and orthonormal basis. Therefore any state of the system which is represented by the wavefunction can be written as a linear combination of its eigenstates.

$$\psi(x,t)=\sum_{n=1}^{\infty} c_n \psi_n(x) \quad (5)$$

The sum of all probabilities should sum up to one. The c_n coefficients are complex numbers that are called probability amplitudes [4] and fulfill the normalization condition:

$$\sum_{n=1}^{\infty} |c_n|^2=1 \quad (6)$$

2.2 Shapelet Basis Expansion Method

The description of the audio signals that is realized in this work bases on their decomposition onto the eigenfunctions of the quantum harmonic oscillator thus falling to the category of a non-parametric signal expansion method [5, pp. 9–21]. The signal can be expanded as a linear combination of the basis functions $\psi_n(x)$. The coefficients for a signal y can be obtained from the following equation:

$$y=\sum_n c_n \psi_n \rightarrow c_n = B^{-1} y \quad (7)$$

Where B is the matrix that contains the eigenfunctions $\psi_n(x)$ of the quantum harmonic oscillator. The $\psi_n(x)$ functions are called Shapelets [6], [7], because they form a perturbed version of a Gaussian function as shown by the equation (4). Shapelets have a main difference to wavelets, namely the various shapes of the basis functions. The wavelet transform basis functions are the same up to a scaling factor. On the other side the Shapelet basis functions are of different size and form.

The first step of the sonification procedure is to store an audio signal into the quantum harmonic oscillator using the overlap - add method [8, pp. 237–238]. The signal is multiplied by a sliding window of length N. The successive windowed signal frames are expanded as a linear combination of the eigenfunctions of the same quantum harmonic oscillator.

The number of the eigenfunctions of the quantum harmonic oscillator is theoretically infinite but in this work only finite number of eigenfunctions are implemented depending on the needs for a good analysis and resynthesis. Throughout this work for a windowed part of a signal with number of samples N the number of the coefficients c_n used was also N. An extended testing of this basis for audio signal processing applications can be made in future works.

2.3 Harmonic Oscillator Quantum Computation Model

The computational building block of a quantum mechanical memory is the qubit whereas the memory of a quantum computer consists of several qubits [2, pp. 13–17]. As with the classical bit, the qubit is realized on a physical system. The primarily difference is that this physical system is in a level where quantum mechanical phenomena are apparent and determine the properties of the storage. A detailed description of the qubit, its properties and its differences with the classical bit are beyond the scope of this paper and are not needed for the understanding of this work.

What is essential for the comprehension of this approach is that the state of the quantum harmonic oscillator is in correspondence with the state of a quantum mechanical memory created from qubits. In the quantum harmonic oscillator model one possible physical implementation of the qubits is made in such a way that the state of the whole memory can be expanded as a linear combination of the eigenfunctions $\psi_n(x)$ [2, pp. 283–287]

The analogy is for every 2^N eigenstates that are used for the expansion of the signal represent the quantum information storage capability of N qubits because they create an equivalent complex Hilbert space.

It is assumed that the system can be prepared in a desired state through an initialization procedure. Special attention needs to be drawn to the fact that the coefficients that are computed for the expansion of the stored audio signal not only need to fulfill the equation (7) but also the normalization condition (6).

3. CLOSED SYSTEM TIME EVOLUTION

A quantum system that evolves without coupling to the environment is called closed or isolated [2, pp. 81–84]. The time-dependent Schrödinger equation (8) describes the evolution of the closed system in time [1, pp. 19–32].

$$i\hbar \frac{\partial \psi(x,t)}{\partial t} = \hat{H} \psi(x,t) \quad (8)$$

where $\hbar$ is the Planck's constant. The time evolution is a procedure that changes the state of the system but leaves the eigenenergies and eigenfunctions unaffected. If the wavefunction of the system $\psi(x,0)$ at time $t_0 = 0$ is described by the equation [2, pp. 13–17]:

$$\psi(x,0)=\sum_{n=1}^{\infty} c_n^{(0)}\psi_n(x) \quad (9)$$

where $c_n^{(0)}$ are the coefficients of the input sound according to the $\psi_n(x)$ basis at time $t_0 = 0$, then after time t each coefficient will be multiplied by a different complex exponential term:

$$c_n^{(t)} = c_n^{(0)} e^{-iE_n t/\hbar} \quad (10)$$

where E_n is the n-th eigenenergy. The state of the system will change accordingly:

$$\psi(x,t)=\sum_{n=1}^{\infty} c_n^{(t)}\psi_n(x)=\sum_{n=1}^{\infty} c_n^{(0)} e^{-iE_n t/\hbar}\psi_n(x) \quad (11)$$

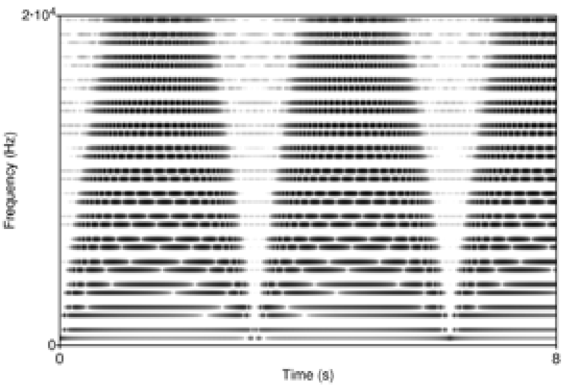


Figure 2: Spectrum of a sinusoidal of frequency 440 Hz expanded over the Shapelet basis, then resynthesized and time-evolved. According to the time-evolution described by equation (11), the spectrum presents modulating frequency components that rise through the periodic change of the expansion coefficients.

Every time that a windowed sound segment is stored in the quantum oscillator, the coefficients $c_n^{(0)}$ for this window are computed with respect to the basis of the eigenfunctions $\psi_n(x)$. Each coefficient is then multiplied with its corresponding exponential term. The real and imaginary part of the time evolved coefficients are separately used for the resynthesis of the sound and produce two individual tracks that are merged in a stereo file. The phenomenon produces a modulation in the spectral domain which repeats itself after a specific period of time (fig. 2).

In this implementation the time variable t doesn't flow continuously at each sample but for every window increases by the value of the hop size that is used in the overlap-add. With the use of a scaling factor on the time parameter the phenomenon can be heard at various speeds.

The time evolution implements a unitary transformation and therefore is the main procedure that can be used for the realization of quantum gates in this computational model [2, pp. 283–287]. With the additional use of perturbation of the eigenfunctions as described in the next chapter, information processing is achieved.

4. OPEN SYSTEM

4.1 Overview of Perturbation Theory

When a quantum mechanical system has strong interactions with the environment it is called open [2, pp. 353–354]. Solving such systems i.e. finding their eigenenergies and eigenfunctions, is a complex and difficult procedure. Therefore an approximation method needs to be used. The perturbation theory is one of them and can be applied when a system with a solvable Hamiltonian $\hat{H}_0$ is subjected to a relatively weak disturbance $\delta\hat{H}$ in regard to the value of $\hat{H}_0$ [1, pp. 133]. Thus, the Hamiltonian of the overall system can be written as an addition of the exact solvable $\hat{H}_0$ and the disturbance $\delta\hat{H}$:

$$\hat{H}=\hat{H}_0+\delta\hat{H}$$

The fact that this disturbance is small enough assures that there are only going to be slight changes $\delta\psi$ and δE on the wavefunction and the energy of the system. The eigenenergies and eigenfunctions can be expressed with the help of power series:

$$E_n^{(k)}=\frac{1}{k!}\frac{d^k E_k}{d\lambda^k},\quad k=0,1,2,\ldots \quad (13)$$

$$\psi_n^{(k)}=\frac{1}{k!}\frac{d^k \psi_k}{d\lambda^k},\quad k=0,1,2,\ldots \quad (14)$$

The perturbation $\delta\hat{H}$ corresponds to a Hamiltonian that is mathematically represented by a Hermitian matrix. In the case of the quantum harmonic oscillator with Hamiltonian $\hat{H}_0$ we can think of a disturbance $\delta\hat{H}$ that is a result of adding or removing some energy from the system. Throughout this work, the use of approximation approaches other than the perturbation theory are not addressed, but this could be a topic that can be further explored.

There are two types of perturbation approaches: the time-independent and the time-dependent. The time-independent procedure describes the system's behavior when the disturbance is constant, whereas the time-dependent deals with systems that are subjected to a time-varying disturbance.

4.2 Time-independent or Rayleigh-Schrödinger Method

4.2.1 Description of the process

The undisturbed or principal system will have an exact solution according to the time-independent Schrödinger equation [1, pp. 134–140]:

$$\hat{H}_0\psi_n^{(0)}(x)=E_n^{(0)}\psi_n^{(0)}(x) \quad (15)$$

The zero at the superscript of $E_n^{(0)}$ denotes that the eigenenergies are from the undisturbed system whereas the n at the subscript of the eigenenergies shows the correspondence of the n-th eigenenergy to the n-th eigenfunction. After the disturbance is applied the Hamiltonian will change by means of the equation (12) where the term $\delta\hat{H}$ is replaced by λV and thus the Schrödinger equation for this system will be:

$$\hat{H}\psi_n(x)=\left(\hat{H}_0+\lambda V\right)\psi_n(x)=E_n\psi_n(x) \quad (16)$$

Where $\hat{H}$ is the Hamiltonian, $\psi_n(x)$ the eigenfunctions and E the eigenenergies of the of the disturbed system. The λ term is a factor that controls the disturbance intensity and can take values with range from 0 to 1 which represent no perturbation to full perturbation accordingly. Just because the disturbance is weak, the eigenenergies E_n of the disturbed system will not deviate very much from the eigenenergies $E^{(0)}$ of the undisturbed. The same property holds for the eigenfunctions. The power series expansion will be in accordance with the equations (13) and (14).

$$E_n=E_n^{(0)}+\lambda E_n^{(1)}+\lambda^2 E_n^{(2)}+\ldots \quad (17)$$
$$\psi_n=\psi_n^{(0)}+\lambda\psi_n^{(1)}+\lambda^2\psi_n^{(2)}+\ldots \quad (18)$$

The superscripts 0, 1, 2... denote the zero-th, first and second term of the power series. The zero superscript is the unperturbed one. The n at the subscript of the eigenenergies shows the correspondence of the n-th eigenenergy to the n-th eigenfunction.

The derivation of the solution occurs by inserting the equations (17) and (18) into (16). The expression of the first order term of the eigenfunction's correction as a linear combination with respect to the orthonormal basis that is formed from the eigenfunctions $\psi_n^{(0)}(x)$ of the unperturbed system, leads to the first and second correction to the energy of the system:

$$E_n^{(1)}=\psi_n^{(0)}(x)^{\dagger}V\psi_n^{(0)}(x),\quad E_n^{(2)}=\frac{\sum_m |\psi_m^{(0)}(x)^{\dagger}V\psi_n^{(0)}(x)|^2}{E_n^{(0)}-E_m^{(0)}} \quad (19)$$

The first term of the eigenfunction correction is expressed by the following equation:

$$\psi_n^{(1)}(x)=\frac{\sum_m \psi_m^{(0)}(x)^{\dagger}V\psi_n^{(0)}(x)}{E_n^{(0)}-E_m^{(0)}}\psi_m^{(0)}(x) \quad (20)$$

The sum is made over all n eigenfunctions except the m. Higher terms can be obtained iteratively but are not used for the implementation to reduce the computational complexity of the implementation.

4.2.2 Audification Choices

For the audification of this perturbation kind various disturbance types that correspond to different Hermitian matrices V were used. One example of a used perturbation corresponds to a constant electrical field with a potential that has linear dependency from the displacement x which is added to the parabolic potential. The λ factor can also be used to control how intense the development of the disturbance phenomena will be.

By applying the same disturbance type V many times consecutively, a slow deformation of the shape of each of the eigenfunctions can be examined at first. The eigenenergie's values are also slightly deviating from their initial value, each one differently but consistently as a whole. Each one of the perturbation types produces a characteristic change which is clearly recognizable.

A phenomenon that occurs in every tested disturbance type, is a deformation of the eigenfrequencies and eigenvalues after the application of sufficient many consecutive time-independent perturbations. Suddenly the system starts to decompose and after a while it collapses. The eigenenergies E_n value range grow and the implemented simulation eventually stops. The eigenfunctions $\psi_n(x)$ are also greatly deformed at the same time because as the eigenenergie's and eigenfunction's changes are closely linked as expressed also from the equations (19) and (20).

The alteration of the eigenfunctions can be made independently hearable by a direct mapping of the eigenfunctions in the time axis, where each channel holds one eigenfunction.

Figure 3 shows the deformation of the eigenfrequencies in subsequent perturbations. One crucial point in which audification and sonification are profitable over a visual representation is the fact that the eye cannot grasp the interconnections and the consistency of the changes of all eigenfunctions as an integrated entity.

As mentioned before, the eigenfunction's transformations can be also made recognizable by analyzing the windowed part of the audio signal as a linear combination of the eigenbasis. In every step of the overlap-add procedure a time-independent perturbation is applied which alters the eigenfunctions in a way that they may not constitute a complete orthogonal basis anymore. Despite this fact, the coefficients are computed as if the underlying basis was orthonormal. By this means the deformation of the sound is an indication for the decomposition of the eigenfunctions and their orthonormality.

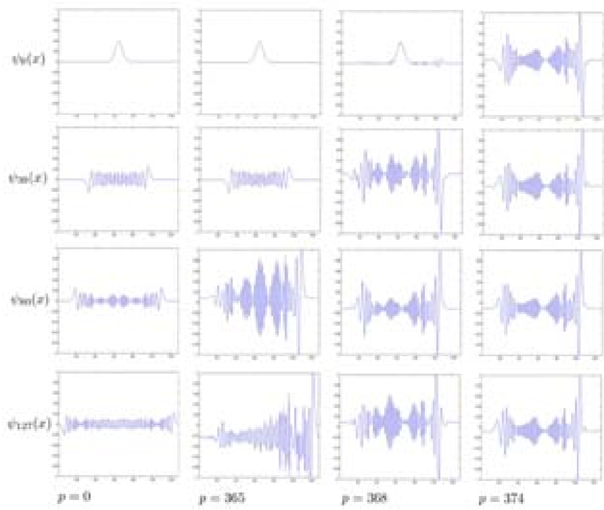


Figure 3: The deformation of the Eigenfunctions ψ_0, ψ_{39}, ψ_{80} and ψ_{127} after perturbing successively several times. In the leftmost part of the figure the Eigenfunctions are unperturbed and at the rightmost are

one step before collapsing. The value of p denotes the number of time-independent perturbations that are already applied to the system. The analogy between the amplitude of the perturbed Eigenfunctions is not in direct proportion with their computed values after the dissolving starts, due to normalizing conditions.

Perturbations of small intensity have no recognizable audible effects in the stored sound. The effects are starting to take place only a little before the collapsing occurs. Because of the rich content of the eigenfunction's alterations, a sonification procedure that would be more reflective of the phenomenon could be addressed.

4.2.3 Sonification Choices

Just because the eigenenergies of the unperturbed system are equally spaced as presented in equation (2), the idea of a projection of their values on the frequency plane have arisen. With an appropriate scaling factor the eigenenergies can be seen as the frequencies of sinusoidals that before any perturbation create a harmonic sound.

Each time the perturbation is applied the values of the frequencies of the sinusoidals are slightly changed. To make the sound effect more recognizable, the amplitude of all the sinusoidal components of the spectrum was set to the same value and then was filtered with the spectral envelope of a vowel of small duration with the help of cepstrum analysis [8, pp. 319–321].

As it can also be seen in figure 4 the first times the perturbation is applied the spectrum of the sound has a characteristic development. After a critical number of perturbations the decomposition of the system begins and an inharmonic sound is produced.

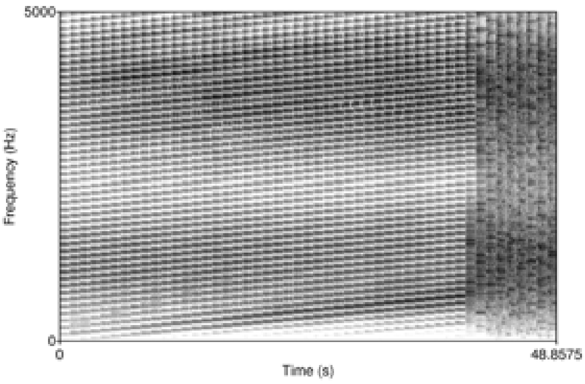


Figure 4: Spectrum of the sonified Eigenenergies after perturbing successively several times. At each frame that is easily seen in the figure by the vertical characteristic, the new Eigenenergies are computed. For the first perturbations the spectrum maintains a recognizable structure and after sufficiently many the dissolution of the spectral structure is apparent.

4.3 Time-dependent or Dirac Method

4.3.1 Description of the process

In this case the perturbation is denoted with the $V(t)$ operator that is assumed to be small in regard to the Hamiltonian H_0 of the undisturbed system and the time duration of the disturbance reaction to the system is considered to be small enough. The eigenenergies and eigenfunctions of the system will also change with time. The Hamiltonian will be the addition of the unperturbed solvable and the time-dependent term [1, pp. 149–153]:

$$\hat{H}(t)=\hat{H}_0+V(t) \quad (21)$$

The time-dependent Schrödinger equation is in this case:

$$i\hbar\frac{\partial\psi(x,t)}{\partial t}=\hat{H}(t)\psi(x,t)=\left(\hat{H}_0+V(t)\right)\psi(x,t) \quad (22)$$

and cannot be solved by separating the spacial and temporal parts with the use of variable separation. That is the reason that in this case the solution cannot be implemented with the approach of the time-independent case. In analogy with the methodology of the time-independent case, the wavefunction of the system will be expanded as a linear combination of the basis of the unperturbed system's eigenfunctions whereas the solution involves the detection of the expansion coefficients.

$$\psi(x,t)=\sum_{m=1}^{\infty}c_m(t)\psi_m^{(0)}(x) \quad (23)$$

The coefficients $c_m(t)$ are represented by a mathematical expression that includes both the time-evolution term that is caused from the unperturbed Hamiltonian H_0 combined with the time-dependent transformation $a_m(t)$ that is generated from $V(t)$:

$$c_m(t)=a_m(t)e^{\frac{-iE_m^{(0)}t}{\hbar}} \quad (24)$$

The $a_m(t)$ terms are expanded with the help of power series. The equation (23) is solved to:

$$\psi(x,t)=\sum_{m=1}^{\infty}a_m(t)e^{\frac{-iE_m^{(0)}t}{\hbar}}\psi_m^{(0)}(x) \quad (25)$$

where the $a_m^{(1)}(t)$ is the first correction term of the $a_m(t)$ expansion:

$$a_m^{(1)}(t)=\frac{-i}{\hbar}\int_0^t V_{nm}(t')e^{\frac{-i(E_n-E_m)t'}{\hbar}}dt' \quad (26)$$

and $V_{nm}(t)$ expresses the term:

$$V_{nm}(t)=\int_x \psi_m^{(0)}(x)^{\dagger}V(t)\psi_n^{(0)}(x)\ dx \quad (27)$$

The further higher terms are computed iteratively but are not used in the implementation of this work due to their computational complexity.

The term $a_m^{(1)}(t)$ in equation (26) represents the fact that the system possesses a kind of memory. The integration is always computed from the time point where the perturbation started. Even if the disturbance stops its action the effects of the interaction are "remembered" and maintained in the system.

This phenomenon is inherited to the stored sound.

4.3.2 Audification Choices

The time-dependent perturbation only affects the state of the system. Therefore an insight to the influence of the disturbance can be only made through the audification of the stored sound. More specific, for every windowed segment coefficients the first order correction term is computed and added as shown in (27). The resynthesized sound with respect to the basis of the unperturbed eigenfunctions contains the changes that are implied by the disturbance.

So far the type of perturbations $V(t)$ that were used could be decomposed as a product of a constant Hermitian matrix V and a function of time $f(t)$. The V term contains the spatial dependency and is in analogy with the systems that were used in the time-independent perturbation and the $f(t)$ which expresses the time dependency and contains combinations of linear, step and sinusoidal functions.

In the signals treated with a time-dependent perturbation there is always an existing component that evolves in time according to the unperturbed Hamiltonian as seen in (24) and a component that evolves under the influence of the perturbation. These two evolutions interfere with each other and create recognizable interference patterns in the spectral domain. Specially in the case where $f(t)$ is almost constant for a specific duration, a periodic component which acoustically is clearly separated from the evolution modulation appears as shown in figure 6.

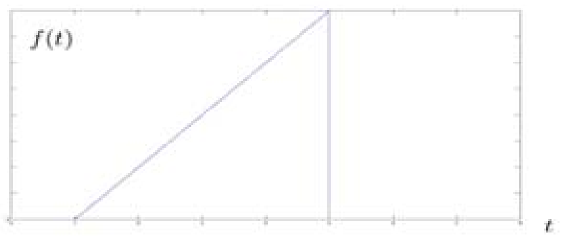


Figure 5: The time-dependency $f(t)$ of the perturbation that was used for the creation sound in figure 6

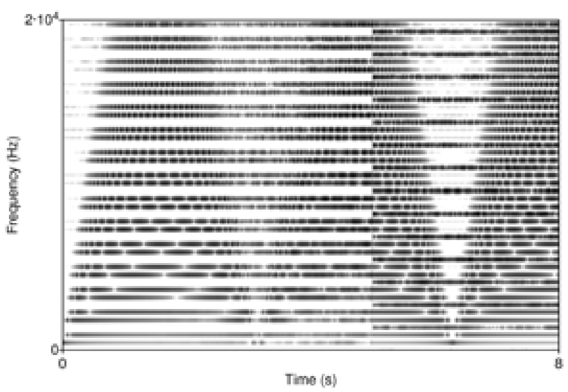


Figure 6: Spectrum of a sinusoidal with frequency 440 Hz when opposed to a time-dependent perturbation with V a Hadamard matrix and time-dependency $f(t)$ as in the figure 5. The development of the sound has similarities with the time-evolution spectrum of the figure 2 in the beginning but the perturbation effects gradually grow. At the time point $t=5\,sec$ the perturbation stops but its effects remain as a constant periodical component.

By using perturbations with different types of V parts and same time dependency $f(t)$ it became apparent that the developed sounds reflect more the time dependency than the V component.

5. IMPLEMENTATION DETAILS

Two different software packages that implement the functionality mentioned above have been programmed, one in Matlab and one in C for a PD external. The GNU Multiple Precision Arithmetic Library (GMP) [9] provided an efficient solution to the computation of the factorial which is the main obstacle for the implementation of the eigenfunctions $\psi_n(x)$ for large n. For the efficient computation of the externals, specially in the analysis-resynthesis stage, the CBLAS [10] and CLAPACK [11] were included.

6. REFERENCES

[1] Ehlotzky, F. *Quantenmechanik und ihre Anwendungen.* Springer-Verlag, Berlin Heidelberg, Germany, 2005.

[2] Chuang, I.L., and Nielsen, M.A. *Quantum Computation and Quantum Information.* Cambridge University Press, New York, USA, 2000.

[3] Schwabl, F. *Quantenmechanik QM I Eine Einführung.* Springer-Verlag, Berlin Heidelberg, Germany, Revised Edition 2005.

[4] Deutsch, D., Ekert, A., and Lupacchini, R. Machines, Logic and Quantum Physics. *Bull. Symbolic Logic*, 6, 3, (Sep. 2000), 265–283.

[5] Goodwin, M.M. *Adaptive Signal Models – Theory, Algorithms and Audio Applications.* Kluwer Academic Publishers, Boston Dordrecht London, USA, 1998.

[6] Refregier, A. Shapelets: I. A Method for Image Analysis. *Mon. Not. Roy. Astron. Soc.*, 338, 1, (Jan. 2003), 35-47.

[7] Coffey, M.W. Properties and possibilities of quantum shapelets. *J. Phys. A: Math. Gen.*, 39, 4, (Jan. 2006), 877–887.

[8] Zölzer, U. *Digital Audio Effects.* John Wiley & Sons, Chichester West Sussex, England, Reprinted 2005.

[9] GNU Multiple Precision Arithmetic Library. http://gmplib.org/, last accessed 17 October 2009

[10] CLAPACK (f2c'ed version of LAPACK). http://www.netlib.org/clapack/, last accessed 17 October 2009

[11] CBLAS Library. http://www.gnu.org/software/gsl/manual/html_node/GSL-CBLAS-Library.html, last accessed 17 October 2009

The well-tempered reverberator: an instrument for composing rooms intuitively

Berit Janssen
University of Hamburg
Neue Rabenstraße 13
20354 Hamburg
0031-6-10651789
berit.janssen@gmail.com

ABSTRACT

The modelling of room acoustics can be performed with a number of different approaches, all of which entail some advantages as well as disadvantages, which mostly leads to a trade-off between acoustic realism and computing time. An algorithm that allows for dynamic change between spaces must therefore combine different techniques to optimize the quality of representation as well as performance. Such an optimization must be powered by insights into the human perception of reverberation. Psychoacoustic research exists, but is mostly limited to inside spaces, or even only concert halls, and is therefore not representative of all possible acoustic spaces. An experimental design in which inside and outside spaces were investigated is presented. Paired impulse responses from these spaces were rated by listeners on a difference scale. A multidimensional scaling analysis of the data points to the early energy decay as the most decisive feature in distinguishing spaces from each other. Similar experiments should follow, which can help to develop a comprehensive theory of reverberating spaces. Not only the implementation of findings from such experiments, but also the development of an intuitive user interface is required to perfect the reverberation instrument.

Keywords

Psychoacoustics; Room Acoustics

1. INTRODUCTION

Reverberation is a vital element of human communication: be it speech, or be it music. The acoustic conditions of performance spaces – indoors and outdoors – have influenced musical composition for centuries. Ever since digital technology allows for the recreation of virtually any room, it is conceivable that the acoustic space itself can be employed as a compositional parameter. Most reverberation tools today either employ statistical, fast methods, even though they cannot model some acoustic phenomena, or they "sample" a real or virtual space in the form of impulse responses. To reverberate a signal with an impulse response, it must be convolved, which has the disadvantage of being less fast and versatile than the statistical methods.

A reverberation instrument must allow for dynamic manipulation of reverberance, but also enable users to locate their sounds in any acoustic space; hence, a trade-off between both methods must be achieved. To achieve a good compromise, it is indispensable to determine which aspects of the temporal and spectral fine-structure of a reverberance can be perceived, and which cannot be perceived by the human auditory system. Only in this way can data be effectively reduced, without a loss in perceptual quality. Most psychoacoustic research of the past decades, however, has been devoted to improving concert hall acoustics. The question as to what makes acoustics "good" has dominated, rather than an unbiased desire to understand the way humans perceive and categorize all kinds of different acoustic spaces.

Different ways to close this gap of knowledge need to be followed, one of which is the investigation of impulse responses through multidimensional scaling. The advantage of this method is its use of difference judgements, which have more inter-subjective reliability than judgements requiring a rating according to verbal descriptions.

The dimensions in which acoustic spaces are categorized can then be linked to objective measurements of room acoustics, and finally be implemented in a reverberation software. In order for this software to become a usable reverberation instrument, however, the objective parameters need to be linked to a meaningful user interface.

An overview of different approaches to artificial reverberation is given first, then the psychoacoustic findings on reverberation are considered. An experiment by the author is presented, which however is only the first step towards the free musical control of reverberation.

2. ARTIFICIAL REVERBERATORS

In principle, reverberation can be seen as a series of repetitions of the original signal, which have a different amplitude. In order to reproduce this phenomenon, two approaches are possible: the impulse response approach, which strives to reproduce the temporal and spectral structure of a real space's reverberance; and the stochastic approach, which assumes that the density and intensity of reflections in real spaces can be convincingly reconstructed using probabilistic methods.

The methods based on impulse responses include either the measurement of a real room or space, or the calculation of the impulse response for a virtual space. The impulse response can then be convolved with a target signal to reverberate it, generally leading to very convincing results. However, both the obtaining of an impulse response and the ensuing convolution have their problems and limitations, as will be discussed below.

The stochastic approach mainly relies on delay lines, which can replicate room reflections as repetitions of a signal at a fixed delay time and gain. A single delay line will not sound convincing, so different combinations of delay lines are in use. These techniques can provide very fast computation speeds, but do not always produce convincing results. It is of course conceivable that the complexity of a delay line network increases to the extent that its impulse response corresponds to the impulse response of a real room, hence the two approaches are ultimately connected.

2.1 Problems of the impulse response approach

A space's impulse response is the series of reflections with which it reacts to a sound emission. This can be considered both in the time and in the frequency domain, as either the arrival time and intensity of reflections, or as a space's characteristic boost of some, and attenuation of other frequencies. Theoretically, the measurement of impulse responses requires an impulse of infinitesimal length and infinite energy. Such a signal does not exist, but can be approximated by a crack from a spark or a starter gun. Alternatively, the response of the room can be measured in the frequency domain, by using a so-called sine sweep: an oscillator stepping through all frequencies. But this approach, too, has its limitations, as the reproducing loud speakers will not respond to all frequencies in the same way. Distortions of the impulse response must therefore be accepted if they are measured in a real space, but the deviations from the "ideal" response are not too severe.

Some of these distortions can be avoided in a digital model of a space. However, the computation of the impulse response suffers from other problems, as the spreading, scattering and attenuation of sound needs to be modelled. Moreover, effects such as diffraction of longer wave forms around edges and obstacles complicate the modelling. Most algorithms generating impulse responses from a virtual space make use of geometric methods, such as the image source method. In this approach, the wavefronts are constructed as rays, and walls and obstacles are treated as acoustic mirrors. With increasing order of reflections, however, the number of possible paths increases exponentially, leading to very long computation times. Other virtual methods to measure impulse responses, such as ray tracing, finite element or finite boundary methods, suffer likewise from potentially high computational expenses.

After the impulse response of a space has been determined by any of the above-mentioned methods, another problem ensues: the convolution of a target signal with the impulse response requires a large number of multiplication and addition operations, which cannot be performed in real time. The computation can be accelerated by transforming both the target signal and the impulse response to the frequency domain; a multiplication in the frequency domain is equivalent to a convolution in the time domain. Hence, fast fourier transforms permit the reverberation of a signal in time windows, which can be played back while the next segment is reverberated. However, the interpolation between such windows remains a problem. Moreover, the options to manipulate an impulse response are limited: no dynamic transition from one reverberance to another is possible.

2.2 Problems of the stochastic approach

Room reflections can also be implemented as a series of delays. Assuming an exponential decay of energy, a feedback delay line with a gain smaller than unity can reproduce a room's energetic behaviour in response to a signal. However, more complexity is needed to achieve a convincing reverberator. The most well-known filter network consisting of different feedforward and feedback delay lines is the Schroeder reverberator sometimes called Freeverb, which combines a number of parallel comb filters, which generate spectral density, with a series of allpass filters, which generate temporal density. [1] However, the choice of parameters – delay time and gain – remains largely "half an art and half a science." [2]

The control of filter networks is not limited to the combination of simple delay lines. New approaches include tapped delay lines, i.e. delay lines with multiple outputs, or multiple input, multiple output (MIMO) networks, which connect the inputs and outputs via matrix transfer functions. [3] Waveguides, two-dimensional delay lines often used for modelling travelling waves along strings or pipes, can also be employed to model paths of travelling waves in a room. Yet the advantage of the stochastic approaches – their fast computation time and realtime applicability – decreases with every increase of complexity.

A good, realistic reverberance is no doubt achievable at real time speed with filter networks. Especially the tail end of an exponential decay can be modelled sufficiently well, whereas many algorithms calculate the first reflections by means of an image source algorithm. However, what about non-exponential decays, colouration or flutter echoes? Rooms and spaces often show such behaviour (see **Fig.1**), but reverberators operating on delay line principles do not. It may be true that we do not need such reverberances if we want to master a vocal or instrumental recording. [4] Yet even if spaces are acoustically "bad", in that they distort the signal to an unclear, unintelligible result, they are part of the range of human experience, and can very well have more than interesting applications for non-instrumental music such as the electro-acoustic genre.

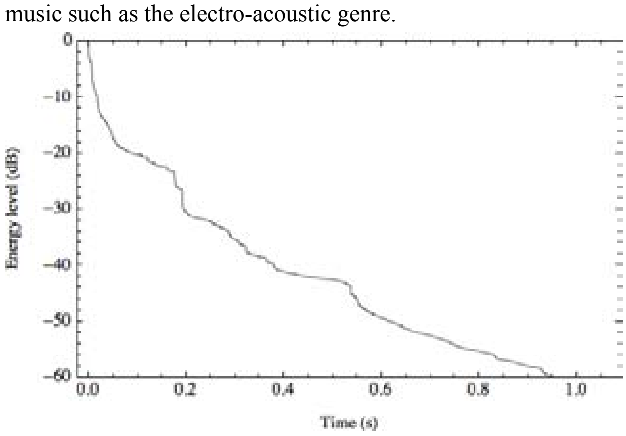


Figure 1. Example for a non-exponential decay as found in an open courtyard: statistical methods are difficult to apply.

2.3 Hybrid methods

A reverberation instrument needs to provide a control of all possible spaces: the unique acoustic features of concert halls, bathrooms, forests, courtyards, streets and any other imaginable reverberance. On the other hand, the control of these spaces should not be only a matter of opening an impulse response from a library: realtime manipulation and dynamic transition between spaces is needed. From the problems pointed out above, it is evident that it is not sufficient to limit oneself to one of the outlined approaches: only through a clever combination of different implementation methods can a convincing, versatile and dynamic reverberator be achieved. The choice of methods must then be guided by the way humans perceive reverberation: which aspects of reverberation are those that enable us to distinguish and categorize spaces, which aspects of reverberation can be implemented at reduced resolution, since they are mostly inaudible? Psychoacoustic findings are needed to answer these questions.

3. PSYCHOACOUSTIC RESEARCH ON REVERBERATION

Investigations on the perception of reverberation have been conducted since the 1960s. Most of the early studies focussed on concert halls, which were either rated during a concert, or in listening experiments with recordings. The subjective responses were then related to objective differences of the reverberances.

Later research also employed artificial reflections, such as impulse responses generated by image source algorithms, or early reflections of a synthesized sound field. In this way, the stimuli can be controlled in a much more sophisticated way, and the listeners' responses can be related to objective measures more precisely.

3.1 Early studies on concert halls

One of the first to investigate the perception of reverberation, Leo Beranek established a number of categories which he found to be important in the quality judgements for concert halls. He rated a number of different concert halls based on these semantic descriptions during orchestral performances, and subjected the ratings to a factor analysis. This showed that the quality judgements were most prominently linked to the length of the time interval between direct sound and the first reflections. [5] It has been criticized, however, that his results are merely based on his own judgements, which makes their global validity doubtful. Yet his semantic categories are still an important achievement, showing the width of sensations attached to reverberation.

Other researchers took an approach comparable to Beranek's, such as Hawkes and Douglas [6], or Wilkens and Lehmann. [7] The advantage of such a method is the weighting provided by the factor analysis: this highlights the most important semantic categories and facilitates the interpretation of the results. On the other hand, these studies suffer from the fact that acoustic experiences are very hard to verbalize: even though subjects may perceive reverberation in similar ways, they may diverge considerably in their word choices to describe differences, yielding very non-uniform data.

In 1974, Gottlob and Siebrasse [8] chose a different approach, asking for preference judgements between two impulse responses from different concert halls. This eliminates the problem of semantic ambiguity, and a factor analysis can be performed on the data. The resulting four factors were found to be correlated with the objective measures reverberation time *T*, definition *D*, interaural cross-correlation *IACC*, and the ratio of early lateral to total early energy.

The findings of these and other studies on concert halls are very valuable to gain an insight into the many parameters that govern the perception of reverberation. However, it is doubtful whether these parameters are also the most salient features of spaces other than concert halls, such as small rooms with prominent resonances, or outside spaces. Hence, studies focussing on more generic spaces constitute another step forward towards a comprehensive understanding of reverberation.

3.2 Studies on artificial reverberation

In a relatively unknown experiment, Berkley [9] investigated speech stimuli derived from an image source algorithm. He used a rectangular virtual room, of which the room size was kept constant, while the reverberation time (i.e. absorption of the walls) and source-receiver distance was varied. He collected difference judgements on paired impulse responses in a listening test, which were subjected to a multidimensional scaling algorithm. This yielded a two-dimensional map, of which one dimension was related to the reverberation time, the other to the timbral development of the reverberance.

In the 1980s, Ando [10] was able to investigate room reflections in even more detail: using wave field synthesis, he had control over the delay and energy of single reflections. He played a piece of orchestral music with different settings for the strength and delay of early reflections to the participants of his experiments. The responses yielded a series of orthogonal factors of subjective preferences, among which the reverberation time, the sound pressure level *SPL* and the interaural cross-correlation were found to be most prominent.

3.3 Applicability of results to the proposed reverberation instrument

As these studies exemplify, most of the psychoacoustic research on reverberation is limited to inside spaces, if not even focussed on concert halls. The results of these studies cannot be employed to understand the cognitive representation of outside spaces, or other spaces showing a significantly different reverberant behaviour. Since human experience is not limited to rectangular rooms or concert halls, it is important to understand how all kinds of spaces are handled by the human auditory system. Only then a comprehensive theory on human categorization of reverberation can be achieved.

Moreover, many studies employ an assessment of the reverberation quality, i.e. they assume that there is "good" and "bad" reverberation. In terms of a reverberation instrument, however, such a distinction is not applicable, since any space can be of musical interest, and should be within the range of the instrument. Hence, the question should be more directed at the overall differences of spaces: how do humans distinguish and cognitively navigate through their acoustic environment?

4. COGNITIVE REPRESENTATION OF REVERBERATING SPACES

An experiment was performed in order to overcome some of the limitations to earlier studies, and to arrive at a more complete picture of human perception and cognitive representation of reverberation.

Impulse responses from inside and outside spaces were presented to listeners in pairs, which were asked to give difference judgements. A multidimensional scaling analysis was performed. These methodological choices are shortly discussed.

The result is a two-dimensional map representing the perceived differences with satisfactory accuracy, of which one dimension could be clearly correlated to the early decay time, whereas the second dimension was more difficult to interpret, as will be discussed below.

4.1 Experimental design

A number of impulse responses was recorded in inside and outside spaces, using a starter gun as an impulse. Two small-diaphragm condenser microphones (*Studio Projects C4*) with omnidirectional capsules were used for recording, and the data was transferred to a notebook running the sequencer software *Audacity* via a *RME Fireface 200.* The resulting impulse responses were normalized, the direct sound was removed, and nine reverberances were selected: five stemming from outside, four from inside spaces.

The impulse responses were then combined to a total of 36 pairs, allowing for an inter-comparison of all stimuli. These pairs were then presented in mono (the secondary microphone having pre-eminently measurement purposes) via speakers to a total of 49 participants, all of which were Musicology students and lecturers at the University of Hamburg. They were asked to rate the pairs on a sale ranging from "very similar" to "very dissimilar".

The responses were then averaged and subjected to the multidimensional scaling routine ALSCAL, as implemented in SPSS.

4.2 Methodology

Difference judgements are a good instrument to access the cognitive parameters by which complex phenomena are categorized. They do not necessitate long training sessions of the participants, or the limitation to "expert" listeners for the performance of listening experiments. Reverberation is a phenomenon that every non-hearing-impaired human being is familiar with, even if not everyone would be able to express the perceptional changes caused by a change of reverberation in accurate room acoustic terms.

Difference judgements have been successfully employed by Grey in his 1977 study on timbre, the perception and categorization of which is also complex, and hard to analyze semantically. [11] He presented his participants with stimuli of re-synthesized instrument timbres at a fixed pitch, in which the initial transient was missing so as to focus on the steady-state timbre.

While difference judgements could doubtlessly also be performed on the basis of more meaningful stimuli, such as music or speech, the present experiment also limits itself to the comparison of rather abstract sounds, namely impulse responses. This is meant to assure that other cognitive processes, such as melody or speech recognition, do not interfere with the difference rating. Moreover, certain sound sources may seem more natural in some reverberances than in others, and lead to further distortions.

Difference judgements are collected in order to perform a multidimensional scaling analysis. This process fits the perceived difference to an n-dimensional map ordering the stimuli according to distances. With increasing dimensions, the fitting of data becomes more feasible, but the interpretation of the map is more difficult.

There are different multidimensional scaling routines which iteratively minimize the difference between the perceived and the mapped distances. The success of the mapping can be expressed in different indexes, as for instance Kruskal's stress factor, with 0 as the optimal fit. [12]

4.3 Results

98.2% of the variation in the participants' responses could be explained by a two-dimensional map, with Kruskal's stress value at s = 0.06, as can be seen below:

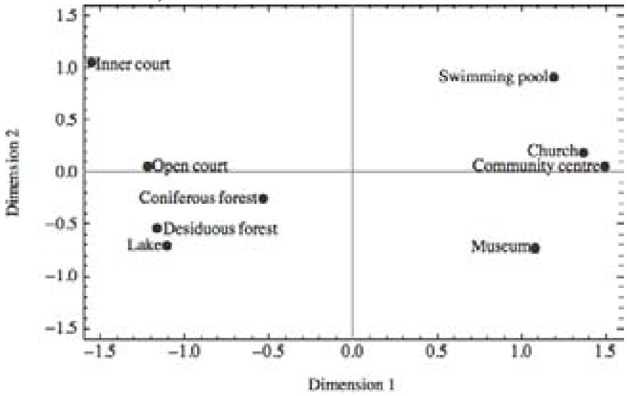


Figure 2. MDS results of a listening experiment on impulse responses

4.4 Discussion

Dimension 1 of the shown representation was found to be highly correlated (R=0.94) with the early decay time of the impulse responses. This seems not to be surprising at first sight, since various other studies have established the importance of the reverberation time, and early decay time and reverberation time are closely related. The early decay time considers the first drop by -10 dB of the energy, whereas the reverberation time generally focusses on the drop from -15 to -35 dB. For exponential decays, these measures will allow to estimate the point in time at which the reverberance drops below the hearing threshold at -60 dB. However, for non-exponential decays, as may be found in many rooms or outside spaces, the measured reverberation time or early decay time is very short, whereas the actual reverberance can be of a duration comparable to inside spaces. Everyone who shouted or clapped their hands in a forest can picture this long, but very subdued reverberance.

From the current data it is evident that it may be not the total length of the reverberation at all which is important for the categorization of reverberation, and that it is mainly the early decay, which is steady for closed rooms, but mostly rapid for outside spaces, which is one of the main cues according to which spaces are distinguished. The importance of this result for the synthesis of artificial reverberation still needs to be verified.

Dimension 2 was more difficult to analyze: it correlates weakly (R=0.52) with the integrated autocorrelogram of the impulse response, a measure of periodicities. Periodicities can lead to a filtering of the spectrum due to interferences. This yields the assumption that Dimension 2 is related to a timbral component, detecting spectral or temporal variations. However, no convincing measure for such effects has been established for reverberation so far, therefore audible timbral effects cannot necessarily be measured.

4.5 Outlook

Due to the limited number of stimuli used in this experiment, the full picture of reverberation needs to be investigated in a range of similar experiments. Moreover, since technical restrictions did not allow for binaural recordings, the influence of sound incidence cannot be assessed from the present data. Future research will remedy these shortcomings.

5. FIRST STEPS INTO AN ACOUSTIC LANDSCAPE

The way to an easily controllable reverberation instrument has just begun: the hypotheses extracted from multidimensional scaling experiments need to be tested in a more systematic way. Third-party software may be employed for this, but also the first versions of the reverberation instrument itself.

Moreover, the effect of reverberation on meaningful sonic objects needs to be investigated: does the shaping through a reverberance have a different impact on a speech signal than on traffic noise, for instance? Do experiences with typical spaces in which we expect to find specific sound sources play a role in our evaluation of reverberance?

The relationship between sounds and acoustic spaces needs to be understood in order to grasp the musical material opened up by a sophisticated control of reverberation. Only then a meaningful interface can be designed, which will allow composers, musicians and other users to control reverberation intuitively as a musical instrument.

REFERENCES

[1] Moorer, J.A. "About this Reverberation Busines." In: *Foundations of Computer Music.* Eds. Curtis Roads and John Strawn. Cambridge, MA: MIT Press, 1985. 605-639.

[2] Zölzer, Udo. "Reverberation." In: *DAFX – Digital Audio Effects.* ed. Udo Zölzer. Chichester: Wiley, 2002. 178.

[3] Smith, J.O. Physical Audio Signal Processing, December 2008 Edition. *http://ccrma.stanford.edu/~jos/pasp/.* Online book, accessed 5-10-2009.

[4] Blesser, Barry. "An Interdisciplinary Synthesis of Reverberation Viewpoints." Journal of the Audio Engineering Society, **49.10** (2001), 867-903.

[5] Beranek, Leo L. "Concert hall acoustics – 1992." Journal of the Acoustical Society of America **92.1** (1992), 1-40, reporting on his work as originally published in 1966.

[6] Hawkes, R.J., Douglas, H. "Subjective Acoustic Experience in Concert Auditoria." Acustica **24.5** (1971), 235-250.

[7] Lehmann, P. and H. Wilkens. "Zusammenhang subjektiver Beurteilungen von Konzertsälen mit raumakustischen Kriterien." Acustica **45** (1980), 256-268.

[8] Schroeder, M.R., D. Gottlob and K.F. Siebrasse. "Comparative study of European concert halls: correlation of subjective preference with geometric and acoustic parameters." Journal of the Acoustical Society of America **56** (1974), 1195-1201.

[9] Berkley, David A. "Normal Listeners in Typical Rooms." In: Studebaker, Gerald, and Irving Hochberg (eds.): *Acoustical Factors Affecting Hearing Aid Performance.* Baltimore: University Park Press, 1980. 3-24.

[10] Ando, Yoichi. *Architectural acoustics: blending sound sources, sound fields, and listeners.* New York: Springer, 1998.

[11] Grey, John M. "Multidimensional perceptual scaling of musical timbres." Journal of the Acoustical Society of America **61.5** (1977), 1270-1277.

[12] Janssen, Jürgen. *Statistische Datenanalyse mit SPSS für Windows. Eine andwendungsorientierte Einführung in das Basissystem und das Modul Exakte Tests.* Berlin: Springer, 2007.

Peruvian orchestral music 1945–2005. Identities in diversity

Clara Petrozzi, PhD
University of Helsinki
Linnustajantie 2 A 10
02940 Espoo, Finland
+358 414818175
clarapetrozzi@gmail.com

ABSTRACT

Peruvian orchestral music since 1945 consists of works in various styles and techniques, but also shows an interest of many composers for expressing local, national or regional identities through the use of elements from popular and traditional musics. The official culture policies have kept the diverse musical traditions separated, ignoring Peruvian society's multicultural character and also the possibilities that orchestral music offers. Studying and performing this repertoire and analyzing it from a cross-cultural point of view should give new light on the country's art music history and its original features.

Keywords

Orchestral music since 1945, Peruvian music, cross-cultural composition, multiculturalism, identity.

1. INTRODUCTION

This paper studies the case of orchestral music in Peru as an example of the cross-cultural musical practice existing in a multicultural society originated by colonization. It observes how the state's cultural policies have taken European models in the founding of music institutions. At the same time, art music composers have turned to traditional and popular musics as a means for constructing an own identity, be it local, national or regional. An analysis methodology that takes into account the interaction of Western art music with other local music cultures should give tools to write an art music history of former colonies that respects their particular conditions, create a repertoire for local orchestras and get local musicians and composers acquainted with their own art music, which is mainly unknown. It would also help to recognize the aesthetic, technical and musical contributions to art music of the compositions written in these countries [1, 3, 8, 12].

Orchestral music and the orchestra institution in general –not only in former colonies– are particularly sensitive to changes in cultural policy, since they depend strongly on the economical support given by the state or private sponsorship. Therefore, the study of the evolution of this institution and the music produced for it might reveal important things about a country's cultural policy and the role that art music and orchestral music have in the life of its people.

2. HISTORICAL BACKGROUND

Peruvian orchestral music tradition starts in the Colony, having its origins in the instrumental groups that performed during theatre plays and in church following the Spanish model. Only in 1814 took place in Lima the first concert with no connection to the stage. During the 19th century, theatre was still the main context for musical life in the city. In these years symphony orchestras already existed and the first classical and romantic style works for orchestra were written. But because of the agitated political situation in the years right before the Emancipation and Independence (1821) many of the compositions were lost, and the most important composers left the country. Opera and zarzuela were predilect genres during the rest of the century, which foreign companies visited regularly. Other important musical institutions were the constantly founded philarmonic societies, which also supported some orchestras, and salon and domestic music with the piano as their main instrument. In Cusco, former capital of the Incas, there was a new sensibility that received the influence of the indigenous music in the middle and upper classes of the society; the so-called pre-indigenist composers combined this music with stylistic features of European or coastal origins. The proximity to the Andean culture caused a renewed interest in local traditional and popular music, and the beginning of musicological research on it at the turn of the century[1]. [5.]

In Peru, the beginning of the 20th century is called the "Aristocratic Republic", because civil conservative governements tried to modernize the country at the workers' expenses, causing unending strikes and protests. Foreign capital penetrated the country and formed enclaves in mining and sugar agriculture zones. The coast and the Andean regions developed in different ways: the coast became more capitalistic, when the Andes mantained arcaic ways of production and suffered exploitation. Governments tried to improve workers' situation, but they couldn't stop the protests and strikes, and the repression caused an increase of the influence of left-wing movements. The first anarchist circles appeared in 1905. [11.]

The cultural policy at the beginning of the 20th century defined prehispanic heritage's protection and created some major cultural institutions.[2] Important events included the celebration of the centenary of the Independence in 1921, the archeological discoveries of prehispanic art and the foundation of the *Amauta*

[1] Musicological research started with José Castro's (1872–1945) *Sistema pentafónico en la música indígena y precolonial* (1897).

[2] Among these were the Peruvian Historical Institute (1903), the Peruvian authors' collection (1906), the Alcedo National Academy of Music (1908) and the National Academy of Arts (1919).

periodical in 1926. The anarchist movement founded cultural circles for urban literate workers, mainly of European descendance. A general change in the perception of the Andean world gave birth to *indigenismo*, intellectual movement that had a great influence on Peruvian cultural life for some decades. Indigenismo thinkers were themselves non-indigenous Spanish-speakers, and paternalism is one of the features of this trend, that wants to protect the native, assimilate him to Western culture and "raise" his cultural level to their own. [6.]

Academic music was practiced by professionals –many of them of European origin– and active *amateurs,* some of which could even teach and sustain musical groups of different kinds.[3] Most orchestras functioned in the theatres or for special concerts. The members of the Lima's Philharmonic Society (1907) founded an orchestra that had much influence in the local musical life. To direct it and the Music Academy the government brought back a Peruvian musician who was living and working in Germany, Federico Gerdes (1873–1953), who re-introduced much of the Classical and Romantic symphonic and chamber music repertoire to the public in Lima. This policy reflects the intention of still doing things "as in Europe", entrusting the important posts to people residing abroad. [6.]

A similar process of institutionalization was carried out in other major cities. In Arequipa, several active composers wrote songs, short piano pieces and other salon music. The musical life included concerts, lyric *soirées* and other spectacles, in institutions such as the Society for Musical Progress, the Musical League, musical clubs and the Philharmonic Society. The Cusco Philharmonic Association gave concerts of European Classical and Romantic music, and from 1915 to 1920 acted the Company of Inca Dramas, which performed theatre and dances in quechua language. Similarly, the Peruvian Mission of Inca Art (1923–1928) toured with great success in the cultural region that includes the southern Andes and the Titicaca area in Peru with Bolivia and northern Argentina. The composers in these cities used local popular music in their works, sometimes difficult to establish as art music. Orchestral works of the period include opera, zarzuela and symphonic music.[4]

After the First World War, many Latin American composers moved to Paris to study with Nadia Boulanger or Paul Dukas, and took part in festivals and concerts in Barcelona, Venice, London and Amsterdam in the 1920s.[5] Many operas and other music theatre compositions were written in these years, but they have stayed in oblivion or are lost.[6] The public for these compositions were middle and upper class city dwellers who could afford paying theatre tickets.

Between the World Wars, the cultural policy included measures such as the legal recognition of the indigenous communities, the first law on cultural patrimony and the Archaeological Species Register. The iniciatives in the musical area came from private people, who sometimes got later some support from the government.[7]

The new century saw also the arriving of new media, influencing greatly the activities of the orchestras in the cities and meaning the gradual decay of music theatre genres. The negative effects of radio and cinema on the orchestras caused the protests of the musicians. The first decades of the century also saw the influence of new popular genres coming from other American countries: Argentinian tango, Mexican music (mainly through the cinema), Cuban son, and American charleston, foxtrot and jazz. The first commercial radio station was opened in 1925 in Lima, and Radio Nacional was founded in 1937. In the first period of the radio the stations had their own orchestras or invited local orchestras to participate in live performances and programmes. The programmation included both classical and popular genres from Europe, other Latin American countries and Peru.

The international crisis of 1930 had effects also in Peru, causing rebellions and dictatures. Manuel Prado was elected president for the period during WWII, when the country was allied to the Allied Forces. Before and during the war many European musicians came to Peru, being the most important the Belgian Andrés Sas (1900–1966) who came in 1924, and the German Rodolfo Holzmann (1910–1992), who arrived in 1938.[8] This

[3] Vicente Stea (Italia, 1884–Lima 1943), for instance, founded the Society for String Quartet, directed many orchestras and composed orchestral music including a Symphony in G (1905) and *Sinfonía Autóctona* (1931).

[4] José María Valle Riestra (1859–1925) was trained in Europe and could face projects of bigger scale as the opera *Ollanta* (versions 1900 and 1920), with influences of Italian opera, Wagner and also included local musical elements. During the war between Chile and Peru (1879–1883) his home was destroyed and many of his works got lost. He wrote also an *Elegy* (1908) in memory of the war victims. He composed some zarzuelas, *genre* which had its golden age in Peru in the second decade of the century. Also the famous piece *El cóndor pasa* is part of a zarzuela (1913), composed by Daniel Alomía Robles (1871–1942), important because of its ant-imperialist libretto, and its author because of his etnomusicological and etnographical work on Andean music and its diffusion in Peru and North America.

[5] Theodoro Valcárcel (1900–1942) studied in Milan since 1914, returning to Peru in 1920. He was named Chief of the Folklore Section at the National Academy, worked in the Ministry of Education's Art Section and later at the Peruvian Art Institute's Music Office. Another composer that received the government's support was Alfonso de Silva (Callao 1903–1937). He got a stipendium to study in Europe, but lived there a bohemian and short life, leaving *Lieder* and three orchestral works, of which his suite *Instantes* (1923) is considered by Pinilla (1985, 155) one of the best of the period. Another composer who studied abroad with a grant of the Government in 1920 and had sufficient technical knowledge as to write two operas (*Cajamarca* and *Francisco Pizarro*) and other orchestral works such as a popular *Marinera y tondero* was Ernesto López Mindreau (Chiclayo, 1892–1972).

[6] At least we know about operas written by Mariano Béjar Pacheco, Policarpo Caballero, Carlos Valderrama, Enrique Fava Ninci, Napoleone Maffezzoli, José Benigno Ugarte, Luis Pacheco de Céspedes and Renzo Bracesco. Zarzuela composers were even more.

[7] For example, important institutions were the Instituto Bach (1929), the Academy Sas-Rosay and the Orchestra d'archi directed by the Italian Virgilio Laghi.

[8] Luis Pacheco de Céspedes (1895–1984) had been studying and working in France and came back in the 40s. His wife, the American dancer Kaye MacKinnon, founded the Ballet Peruano and Pacheco wrote several ballets with national subjects from the pre-Inca and Inca cultures and using themes or rhythms of Andean, Amazonian and coastal musics. He is

migration was one of the reasons that the Philharmonic Society's members argued to insist on the necessity of founding a national symphony orchestra, after a succesful concert series organized by Lima's local government. The orchestra was founded in 1938 and impulsed the creation of new orchestral works. The musicians of the new Orquesta Sinfónica Nacional had also to teach in the Alcedo Academy. The OSN started its activities with great success. An open air auditorium was built in 1939. In the first eight years the orchestra offered 600 concerts and premiered about 300 works. In this way, orchestral music changed its status from an upper class hobby to a cultural service directed to the whole population, under the jurisdiction of the Ministry of Education until 1962. [6, 9.]

During World War II the relative economic welfare thanks to the rise of metal prices benefitted the country. The government started the legislation of the musical sector since 1939, establishing a five members Council to decide on musical education policy. In 1941 a law on Public Education created an Artistic Education and Cultural Extension Directory. In 1942 the government established the National Art Prizes and published programs for the public music instruction in primary and secondary schools. In 1943, a new Music Culture Board made up of seven members took the lead of the National Academy and of the Symphony Orchestra. This Board had to decide on all the principal matters such as nominating professors and deciding the budgets for the Academy and the Orchestra, and creating choirs and chamber music groups. This was the first attempt on behalf of the government to develop a coherent policy for the musical sector. In these years official status was given to the small orchestras existing in Arequipa (South) and Trujillo (North) as well as the music schools in the same cities. [9.]

3. ORCHESTRAL MUSIC AFTER 1945

After WWII, the international politics again influenced art music events in the region. During the 1950s, USA developed an active Panamerican cultural policy. The Organization of the American States (OAS) organized festivals, tours and concerts since the 1940s, gave grants, and published music, catalogues of colonial music and a bulletin since 1957. Other regional institutions founded were the Inter-American Institute of Ethnomusicology and Folklore, and the Inter-American Music Council, 1956. Electronic music laboratories were founded in several countries. Béhague [2] states that the 50s marked the beginning of a new era for Latin American music, whereas the political and economical dependence continued, but Latin America faced a more cosmopolitan world.

In Peru, these years witnessed big migrational waves from the countryside to the cities and especially to Lima. This gave Andean popular music new possibilities of spreading: the first recordings were made in the forties, the first radio programmes dedicated to Andean music started in the 50s, and a lot of song books were published. In Lima, art music composers could get accquainted with the Andean tradition more easily. The official cultural policy, however, did not pay attention to these changes, in spite of the interest showed by composers in combining these traditions and studying vernacular musics. In 1957 Radio Nacional was commissioned the recording of all the concerts given by the OSN. TV Channel 7 of the Ministry of Education started the next year due to an agreement with UNESCO, followed by several commercial stations.

The 40s and 50s saw the coming on the scene of an interesting group of composers,[9] which received the name of "Generation of the 50s" eventhough they did not share a common aesthetic. Nearly all of them have written orchestral compositions. At postgraduate studies abroad they learned modernist techniques and styles from dodecaphony to aleatoric music, graphical notation and happening. One of the most important events in the sixties was the foundation of the Centro de Altos Estudios Musicales of the Instituto Di Tella in Buenos Aires, led by Alberto Ginastera and functioning from 1963 to 1971.[10] [4.] Latin American composers could contact there not only the most important composers of that time, but also other colleagues of the region.[11]

The searching of an identity different to the European took in some cases a Latin American dimension.[12] Some Peruvian composers sought their own musical roots in the traditional music of their regions,[13] in popular art or in the features of the geography and the accomplishments of pre-Inca cultures,[14] and combined them with modernist compositional techniques. Many experimental trends were tested by Latin American composers, but they did not get enough working opportunities, and many of them migrated. Positive events were the nine Festivals for Interamerican Music organized by the OAS in Washington between 1958 and 1978, where some hundred works were premiered, and the three Festivals of Latin America and Spain in Madrid (1964–1970). [2].

4. MUSIC AND REVOLUTION

The sixties were a time of social rebellions in South America. Following the Cuban revolution, many leftist movements got support from young artists, also in Peru. Art music composers joined this kind of projects, and this decade can be seen as a

excepcional for his almost exclusive use of the orchestral medium.

9 José Malsio (1925–2007), Enrique Iturriaga (b.1918), and Rosa Alarco (1913–1980). Iturriaga was the first composer graduated of the National Conservatory (name of the Academia Alcedo since 1946). His work for voice and orchestra *Canción y muerte de Rolando* (1947), written when he was still a student, won a National Prize and was also significant for putting new Peruvian poetry into music. After them Enrique Pinilla (1927–1989) continued studies in Madrid and Berlin, Armando Guevara Ochoa (b. 1927) in the US and Celso Garrido-Lecca (b. 1926) in Chile. To this group belonged also Francisco Pulgar Vidal (b. 1929), César Bolaños (b. 1931), Edgar Valcárcel (b. 1932) and Olga Pozzi Escot (b. 1931).

10 In the institute teached José Vicente Asuar, Earle Brown, Luigi Dallapiccola, Mario Davidovsky, Luis De Pablo, Roman Haubenstock-Ramati, Witold Lutosławski, Bruno Maderna, Gian Francesco Malipiero, Olivier Messiaen, Luigi Nono, Eric Salzman, Domingo Santa Cruz, Roger Sessions, Vladimir Ussachevsky, John Vincent.

11 Peruvian César Bolaños directed the electronic laboratory of the CLAEM since its foundation in 1963. Other Peruvians working in the CLAEM were Edgar Valcárcel and Alejandro Núñez Allauca (b. 1943).

12 This is the case in some works of Celso Garrido-Lecca.

13 As Edgar Valcárcel and the music of the Puno region.

14 For instance in the symphony *Chullpas* by Francisco Pulgar Vidal or in *Elegía a Machu Picchu* by Celso Garrido-Lecca.

time of relatively high politization of the arts. In all South America there was a debate on modernity versus the searching for own roots and identity in art, strengthened by these social and political movements. The European modernism was felt strange by local artists, because it took them far from their own social reality, which was demanding a position statement from the artist. The romantic use of vernacular melodies did not satisfy the modern aestethical ideals of the composers any more. That is why, after a short period of experimentalism in avant-garde trends and techniques, Latin American composers turned their view back to the popular musics of their region, in a moment that pluralism was gaining presence internationally. According to Juan Orrego-Salas [14] the Latin American vanguardist feels alienated of his own reality, and seeks as a solution the synthesis between the local cultural roots and the international tradition.

In Peru, the cultural policy was delegated to a new institution, the Casa de la Cultura. Composer Enrique Pinilla led its office for music and cinema in 1963–1969, organizing archives and concerts, and conducting national works as a guest conductor at the OSN, but he could not fulfill all his plans in this post. The main problems were the lack of budget and the personalism persisting in the decisions of the cultural authorities and institutions. In 1963 an official status was given to the National Folklore School, which existed since 1949. The orchestras were put under the Casa de la Cultura's jurisdiction until 1971. Reorganizations of the OSN took place in 1957 and in 1965, followed by its dissolution and its new founding. Its principal conductor changed also many times.[15] In 1965 a Culture Law was given, where new subjects were discussed, as author's rights or tax deduction. [9.]

A military government took the power in 1968 and executed a strong ideological anti-imperialist and nationalist policy, banning for instance rock music as alienating. The government tried to promote cultural elements that had been neglected: indigenous languages, handicrafts and popular musics. The Andean culture and the Inca resistance were important for its propaganda. In 1971, the Organic Law for the Education Sector created the National Culture Institute (INC), which tried to promote the arts and the creation but had severe budget limitations. Some composers supported this policy but got disappointed with its limited results and the bureaucracy implied. Andean popular music was seen also by composers as a symbol for social solidarity and was used in some works. Nationalist works were also written: for instance, the government called for a symphony contest where the national anthem had to be used as source material. Art music was not prosecuted, but the economic crisis hit its institutions hardly. Still, in the first years of the 70s there were new music concerts organized and some of the most important works of the period were written in these years.[16] [4].

The INC influenced directly the orchestral activities in the country. In 1972 ambitious objectives for the OSN were defined by this instance [9], but the orchestra did not receive sufficient resources to accomplish them. The policy of the INC preferred national conductors, and since then, the OSN has been led by Peruvian as main conductors. The lack of economical resources was evident in the infrastructure of the orchestra and in the musicians' salaries. The National Conservatory was integrated into the INC, and its name changed to National Music School. Some new activities took place in it, such as workshops of Popular Song and Musicology. During the rest of the decade, in coincidence with to this turn towards popular musics and the international pluralism, avantgarde works were gradually abandoned and forgotten. At the same time, the art music institutions fell into a severe crisis that continued for several years.

5. CRISIS AND PLURALISM

The return to democracy in 1980 was followed by the disapointment caused by the economic crisis and the violent conflict between maoist Sendero Luminoso guerrilla and the army forces. These difficulties caused the migration of many musicians and damaged seriously the musical institutions. The return to democracy meant, again, organizational changes: a new status was given to the state orchestras in 1984 as part of the Cultural Activities Department of the INC. As one of their main purposes stood the diffusion of Peruvian works, creation of chamber music groups, recordings and commission of new works. [9.] Without sufficient resources, these functions could not be fulfilled. Instead, the orchestra suffered the lack of instruments and furniture, transportation problems, lack of funds for comissions, lack of archive and statistics, and a decrease in the musicians' salary. [13.]

During Alan García Pérez's first period (1985–1990) inner violence continued. The crisis got worse by natural catastrophes and borderline problems. In spite of some state iniciatives in the cultural area, the music institutions did not receive relief to its problems. The private sector offered some services in musical education, which was not covered by the conservatories in all levels, organization of music festivals, master classes and even comissioning of works. Some private orchestras had an intermittent activity. An opera association started in 1980 with the presentation of opera seasons with Peruvian and foreign artists and orchestras, and a private classic music radio was started in 1983. Cultural centers of many countries had a big role supporting local music activities. Because of the violence, centralism became more pronounced and province orchestras worked in isolation. The artistic production of the 80s reflects these problems. The number of young composers studying and working in the country decreased, and few composers could write for orchestra. [15.]

The aesthetic of the orchestral works written in this period reveals postmodern features such as the use of techniques of quotation and collage; some composers turned back to tonal languages or combined tonal and non-tonal elements. Historical styles and genres –such as colonial Baroque– were also evoqued by other composers.[17]

Alberto Fujimori's government (1990–2000) defeated Sendero Luminoso, but commited crimes of corruption and against Human Rights. The revelation of these illegalities led to Fujimori's escape to Japan and to the provisional government of Valentín Paniagua, who called for elections, won by Alejandro

[15] After the first director, the Austrian Theo Buchwald, came the Rumanian Jean Constantinescu, the German Hans Günther Mommer, the Peruvian Armando Sánchez Málaga and the Mexican Luis Herrera de la Fuente.

[16] In orchestral music, for instance, Edgar Valcárcel's *Checán II*, Francisco Pulgar Vidal's *Elegía Apu Inka* and Enrique Pinilla's piano concerto were all written in 1970. The group Nueva Música presented chamber music programs with works of Peruvian and other Latin American composers.

[17] Colonial themes have been used by Francisco Pulgar Vidal, Pedro Seiji Asato and others.

Toledo (2001–2006). A Commission of Truth was established and revealed the terrible results of the years of violence.

6. GLOBALIZATION

Peru initiated its recovering from this conflict in a world that had changed after the falling of the wall in Berlin (1989), of the Soviet Union (1991) and of the birth of the "global village" as a consequence of the progress in communications and transport, and of the digitalization process. The movements of capitals caused financial crisis in Mexico, Brasil and Argentina. In Peru, the liberalization of the economy increased contacts with other countries; the changes in world politics brought many musicians to Peru from Cuba, the former Soviet Union and East Europe countries.

Musically, the international trend to mix types of music such as classical and rock, and the *world music* used some strategies which were common to Latin American music cultures since colonial times: borrowing, mixing and hybridization. One of the problems of orchestral music in the country is that there has been no possibility to officially study orchestral conducting. Still, orchestras exist also in the around 15 high level music schools and conservatories throughout the country, one private conservatory in Lima and some universities which are now giving the possibility to study music. In 2008 the National Conservatory received the status of university level.

The organization of the state orchestras was again changed in 1993, 1994 and 2001 [9]. In fact, the lack of resources continued and Peruvian musicians had to seek working opportunities abroad. In 1991 a private association, Renacimiento, founded a Philharmonic Orchestra in the capital city, which lasted only a few years because the private sponsors did not continue supporting its work. The child- and youth orchestras' movement is steadily growing; in some places, with a strong social function as they work with children of very poor economical situation.

The liberalization of the economy brought also many new instances to the market of basic music education: private schools and music schools, some of them including orchestral work. New organizations in the musical sector include the Círculo de Composición del Perú (CIRCOMPER), a composers' collective with members residing in the country and abroad [7]. One interesting feature of the music of the new millenium is the rise in the number of composers; another is the renewed writing of operas after many decades. New media and new tools in the writing of music have been very important facilitating the work of Peruvian composers, their access to information and to resources, and to working possibilities such as competitions, calls for scores or the distribution of their music or the possibility to show it to others through Internet, websites or blogs.

The official text *Lineamientos y programas de política cultural del Perú 2003–2006* (Guidelines and programs of Peru's cultural policy 2003–2006) shows a new emphasis in the multicultural, multiethnical and multilingual character of the country and the need to respect it as a basic element of the cultural policy. It expresses the need for a Ministry of Culture that would include some areas now not included in the INC, such as the National Library, the Science and Technology Council, the editorial sector and the cinema. This Ministry would ellaborate cultural policy plans for the different regions considering their specific needs. [10.]

7. CONCLUSIONS

It is not evident that a new document on cultural policy or the eventual creation of a Ministry of Culture will guarantee a better destiny for Peruvian art music than it has had until the present day. The official recognition of a multicultural state implies a status of membership in society also for art music, still seen by some as "elitistic", "foreign" or "alienating" in a kind of reverse cultural racism. This also should create pressure on art music institutions to recognize the multicultural and multilingual/multiethnical reality and its richness and let it be reflected in its practices and study programmes, even more, since this happens in the work of many Peruvian composers.

These works should be studied and performed, since they express in a unique way the many identities that coexist in the Peruvian society: the local, the national, the Latin American, and the composer's individual language as a part of the so-called universal or international art music world.

The functioning of the music institutions in Peru should be also studied, since it shows that a great amount of personalism and inefficiency are damaging the work of the musicians. A look on the Peruvian culture policy of the last decades shows a constant changing of the bureaucratic organizations in charge of the musical life, and a rift between these organizations and the musicians working in the field, whose interests are not taken into account.

History shows that orchestral music is influenced by the cultural policy in many ways. The quality of the orchestral institution is reflected in the number and character of the works written for the medium. The possibility of having the work played is very important for the composers. This is shown also in the number of concertant works, commisioned or dedicated to specific soloists who can arrange the performance of the work. In turn, the amount of support given to culture, to the arts and to particular forms and expressions of arts depends on political and ideological decisions. The economic crisis of the country is reflected in the resources given to the cultural sector. [14, 15.]

Supporting this art form in Peru –or in other former colonies– has sense, because Western art music is a tradition brought to Peru in its day by Europeans, but it has already several centuries of own history in the region. It has already become one of the components of the country's multicultural mosaic. As the works of the Peruvian composers show, their orchestral music has original features that are born from the fertile dialogue between local popular and traditional musics and the "big tradition", as well as contributes to it [14]. As an art form, it has the right to receive the same support from society as any other art expression. Orchestral music, both its practice and its consume, can be important for the new urban population that is no longer connected with the musical traditions of their towns of origin, as an alternative to the offer of commercial mass media.

8. REFERENCES

[1] Adler, C. Cross-cultural hybridity in music composition: Southeast Asia in three works from America. On line: http://members.cox.net/christopheradler/writings.html, 1998, last visited 1.5.2008.

[2] Béhague, G. *Music in Latin America: An Introduction.* Prentice Hall, Englewood Cliffs, NJ, 2000, 285–286.

[3] Blake, M. The present-day composer refuses to budge: case studies in new South African orchestral music. *South African Journa of Musicology* 25, 2005, 127-143.

[4] Bolaños, C., Estenssoro, J.C., Iturriaga, E., Pinilla, E., Quezada Macchiavello, J. and Romero, R. (eds.) *La música en el Perú.* Patronato Popular y Porvenir pro Música Clásica, Lima, 1985, 165-200.

[5] Op. cit., 121-123.

[6] Op. cit., 127-143.

[7] Circomper. Círculo de composición del Perú. On line: http://www.geocities.com/circomper2/, last visited 21.4.2008.

[8] Chapman, J. The creative practice of composition in a cross-cultural musical setting. *Speculation and Innovation: applying practice led research in the Creative Industries.* Queensland: Queensland University of Technology. On line: www.speculation2005.qut.edu.au/papers/ChapmanJ.pdf, 2006.

[9] Coloma Porcari, C. *La Orquesta Sinfónica Nacional (1938-2002).* Instituto Nacional de Cultura, Perú, Centro Nacional de Información Cultural, Lima, 2005. On line: http://inc.perucultural.org.pe/serv5.shtml, last visited 24.3.2005.

[10] Consejo Nacional de Cultura. *Lineamientos y programas de política cultural del Perú 2003-2006.* Instituto Nacional de Cultura, Lima, 2002.

[11] Cotler, J. *Clases, estado y nación en el Perú.* Lima: Instituto de Estudios Peruanos, 1978, 132–254.

[12] McGraw, A.C. Balinese Musik Kontemporer: An Other Avant-Garde. facultystaff.richmond.edu/~amcgraw/writing/other_avant_garde.pdf, last visitde 6.6.2008.

[13] Meza, L.A. La Sinfónica en tres movimientos (segunda parte). *La casa de cartón de OXY,* segunda época, 9:58-65, 1996.

[14] Orrego-Salas, J. Técnica y estética. *América Latina en su música.* Isabel Aretz, ed. México D.F.-París: siglo veintiuno editores-Unesco, 1993 [1997], 174-198.

[15] Petrozzi, C. *La música orquestal peruana de 1945 a 2005. Identidades en la diversidad.* Ph.D. Thesis, University of Helsinki, Helsinki, Finland, 2009. On line: http://urn.fi/URN:ISBN:978-952-10-5690-1, 2009.

Expan: a tool for musical expressiveness analysis.

Campolina, Thiago
CEGeME - EMUFMG
Av. Antônio Carlos, 6627
Belo Horizonte – MG – Brasil
thicampolina@yahoo.com.br

Loureiro, Maurício
CEGeME - EMUFMG
Av. Antônio Carlos, 6627
Belo Horizonte – MG – Brasil
mauricio@musica.ufmg.br

Mota, Davi
CEGeME - EMUFMG
Av. Antônio Carlos, 6627
Belo Horizonte – MG – Brasil
davialvesmota@yahoo.com.br

ABSTRACT
Variations of note duration, articulation, intensity, pitch and timbre are commonly used by performers as resources to communicate their intentions of expressiveness and intelligibility. A systematic analysis of descriptor parameters that encapsulate such variations may indicate existing connection between the interpreter's manipulation of these variations and the way we perceive his or her expressiveness intentions. In this paper, we propose an application designed to facilitate the analysis of a set of descriptor parameters of musical expressiveness, aimed to both research and pedagogical settings.

Keywords
Musical expressiveness; musical performance; music content processing and extraction; musical signals segmentation.

1. INTRODUCTION
Studies on musical expressiveness have demonstrated that musicians use small variations of duration, articulation, intensity, pitch and timbre to communicate to the listener aspects of the music they interpret [2, 3]. By comparing performances of different musicians, as well as different interpretations by the same musician, these deviations can be perceived with a surprising clarity, even by non specialized listeners. Systematic analysis of data retrieved from audio signals supports that our ability to perceive variations of acoustical parameters from different performances are connected to some extent to how musical sentences are we organized. Accurate and robust expressiveness-related data acquisition together with a friendly interface designed for users with no specialized computational skills, will not only enhance the potentials of this tool for scientific research, but also mediate its use in educational environments such as instrument pedagogy, musical analysis and music education.

2. FEATURES
The application was initially conceived to compare short monophonic musical excerpts with identical number of notes. Therefore, the score presented to subjects, loaded as a MIDI file, should not contain trills or any kind of unmeasured ornaments. Multiple tracks can be loaded or recorded, processed and have their data displayed in separate panels for individual analysis or side by side for track comparisons. The panel displays time evolution individual or multiple descriptor parameters, which can be selected by the user from a set that includes note parameter values such as note duration, intra-onset-interval (IOI) duration, articulation index, legato index, as well as intra-note feature vectors, such as pitch, intensity, spectral centroid, spectral irregularity and others. As this is a prototype, we focus only two descriptor parameters related to note transitions, *articulation index* and *legato index*, as described below. A description of the data acquisition and processing techniques used will be discussed in the sequence.

3. CONCEPTS AND METHODS
3.1 Note Segmentation
Segmentation is not a trivial problem, even on monophonic musical signals, especially if we consider the subjectivity of note discrimination. In order to estimate inter and intra-note descriptor parameters, the signal was segmented into musical notes according to the score, then each note was further segmented into attack, sustain and release. Note onsets, note offsets, end of attack and beginning of release were detected on the RMS envelope averaged for 1024 samples (hamming window with a hop size of 256).

Onsets and offsets were detected by searching the minimum RMS between two consecutive values crossed by an adaptive threshold, as suggested by De Poli [1], calculated as the average energy in a neighborhood of 1 s of each RMS frame. The estimate of the fundamental frequency changes, with a pitch threshold below a half tone, helped the segmentation in cases where the detection of onsets and offsets was not possible by means of energy level only, such as legato notes.

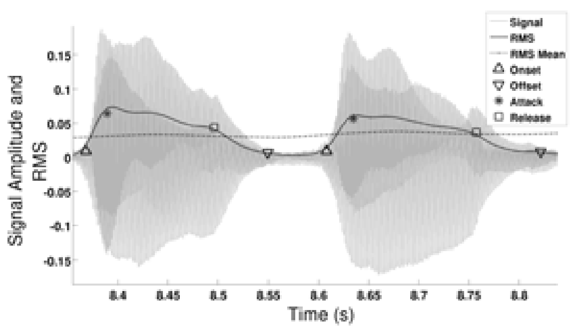


Figure 1. Detection of note onsets, end of attack, beginning of release and offsets, on the RMS envelope

End of attack and beginning of release were detected for the inner segmentation of the note. It doesn't exist in the literature a measurement method that can describe the attack unequivocally [4]. End of attack is commonly defined as the first amplitude maximum after the note onset, which was detected by searching for maximum local variations of the first derivative of the RMS signal. However, while this definition is adequate to describe the attack in most situations, it fails in cases where maximum amplitude is reached much later in the sustained segment of the note.

Due to the presence of transients during note transitions, the end of attack can be assumed as the point where the correlation among harmonics amplitude is reestablished after note onsets. In cases of soft attacks or notes tightly bound to the previous one, this assumption appears to be an adequate alternative to represent the instant where the sustain part of note begins. Therefore, the end of attack was determined by the spectral flux, a measure of the degree of variation of the spectrum over time, as in equation 1 [5]:

$$SF = \frac{1}{T}\sum_{p=1}^{T}\left|r_{p,p-1}\right| \quad (1)$$

where $r_{p,p-1}$ is the correlation coefficient among spectral amplitudes calculated at instants t_p e t_{p-1} of the signal of duration T, windowed as in the calculation of the RMS envelope.

The spectral flux confirmed most of the end of attack estimated by energy variation with an average difference of 6.4 ms between the 2 methods and was able to detect these points where the energy method was not.

The beginning of release was detected as the first amplitude maximum before the note offset.

3.2 Descriptors for Note Transitions

The intentional manipulation of note transitions is an important issue in the communication of expressiveness intentions. In order to analyze the quality of the transition between consecutive notes, two descriptors are employed: *articulation index* and *legato index*.

3.2.1 Articulation Index

Note durations and attack quality are well controlled by the musician, which are closely related to the performer's intentions of expressiveness and intelligibility in a particular room (the performer's ability to control note durations depends on ambience reverberation conditions). The *articulation index* (equation 2), defined as the ratio between note duration DR (time interval between note onset and offset) and the time interval between both note onsets (known as intra-onset-interval - IOI), was used to describe the intentional manipulation of note duration by the performer (figure 2):

$$AI(i) = 1 - \frac{DR(i)}{IOI(i)} \quad (2)$$

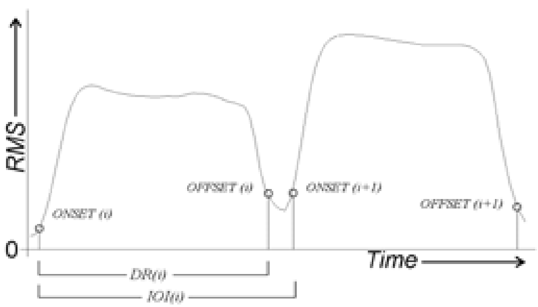


Figure 2. Parameters used to calculate the Legato Index

This index is appropriate to describe transitions between detached notes, usually produced in the clarinet with abrupt interruptions of the air flow by slightly beating the tongue on the reed. With this action, the player controls the quality of the attacks as well as the duration of each note. On the other hand, this index is not adequate for describing the quality of *legato* transitions, as it assumes values close to 0 for most *legato* notes, and does not considers the energy envelope behavior in the transition region.

3.2.2 Legato Index

Legato notes on the clarinet are produced by means of a single blow without interrupting the air flow during the transition. To investigate the quality of transitions between notes played with the intention to be *legato*, we use a descriptor suggested by Maestre [4], which appears to be related to the musician's abilities, to the ambience reverberation conditions and to the acoustic characteristics of the instrument. The legato quality is measured by a comparison to an ideal *legato* without any decrease of energy, represented by the straight line traced from the beginning of release of a note to the end of attack of the subsequent note, as showed in figure 3.

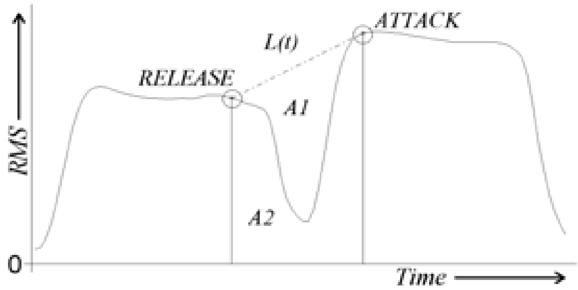


Figure 3. Areas used to calculate the Legato index.

The *legato index* is calculated by means of the area A_1 between this line and the energy curve and the total area $A_1 + A_2$ below this straight line, as in equation 3:

$$LI = 1 - \left[\frac{A_1}{A_1 + A_2}\right] \quad (3)$$

4. RESULTS AND DISCUSSION

4.1 Segmentation

Table 1 shows the error calculated between the automatic and manual note segmentation of a set of recordings of different musical excerpts played by the same musician, comprising 2986 notes. Errors were calculated with different tolerances sizes: 0, 1 and 2 RMS frames (sample rate of 44100 Hz).

Table 1. Error of onsets and offsets automatic detection.

Tolerance: number of frames (ms)	Error: number of notes (%)		Success: number of notes (%)	
	Onsets	Offsets	Onsets	Offsets
0	215 (7.2%)	187 (6.3%)	2784 (92.2%)	2808 (94.1%)
1 (23.2 ms)	154 (5.2%)	153 (5.1%)	2845 (95.3%)	2842 (95.2%)
2 (46.4 ms)	102 (3.4%)	125 (4.2%)	2897 (97.0%)	2874 (96.3%)

The signal can be played back simultaneously with a MIDI synthesized version of the excerpt with the estimated note onsets and offsets as an efficient tool for note segmentation and pitch detection verification. A different set of 54 samples consisted of six different musical excerpts executed by three clarinetists, each with three different induced intentions was also tested. Table 2 shows the accuracy of this segmentation.

Table 2. Automatic note segmentation accuracy.

Total Number of Notes	1927
Extra Notes	44 (2.28%)
Missing Notes	49 (2.54%)
Notes out of sync	3 (0.16%)
Wrong Pitch	10 (0.52%)

These data were also used to compare the two methods of attack detection, by spectral flux and RMS variation. Figure 4 shows a histogram of the time differences between the two methods in number of frames of 5.3 ms (sample rate of 48000) of 267 notes. Positive values indicate that the detection of the end of attack by the spectral flux was later than by RMS variation (mean = 6.4 ms). It was verified that time differences grater than approximatelly 10 frames, most of than negative, correspond to long notes (500 ms or longer), where RMS variation was not able to detect the end of attack. Spectral flux

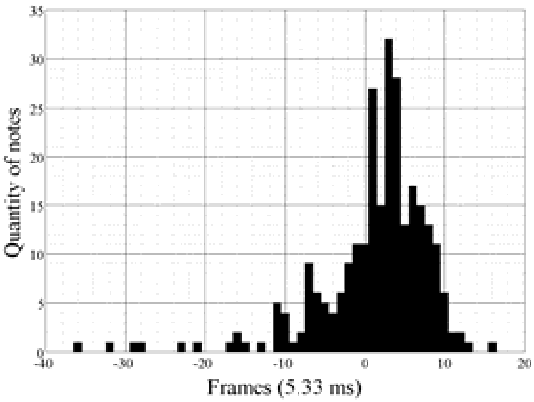


Figure 4. Histogram of the difference between the ends of attack determined by the spectral flux and RMS energy variation, in frames of 5.3 ms.

4.2 Interface of Expan

The interface of Expan is intended to provide a simple way for comparing descriptor values for similar notes of different samples, in order to enable objective analyses of large amounts of data. Figure 5 shows the first version of the graphical interface. Three different performances of the score displayed on the top panel are displayed in the panels below. The first one plots the signal wave form and the other two the *legato index* and the *articulation index*, respectively.

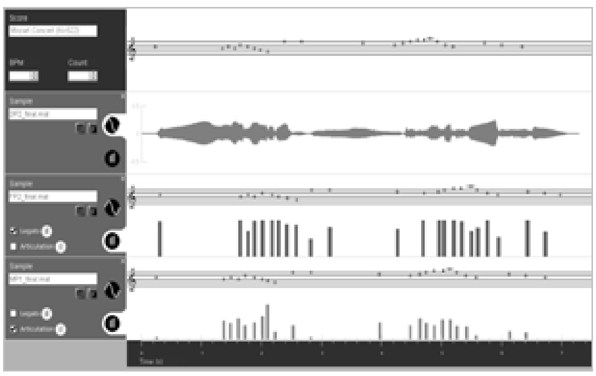


Figure 5. Graphical user interface

4.3 Examples

The following examples shown in figures 7 and 8 correspond to an excerpt of the first movement of Mozart's Clarinet Concerto in A Major, Kv 622 (figure 6).

Figure 6. Excerpt from the first movement of Mozart's Clarinet Concert in A Major (Kv 622)

4.3.1 Example: Articulation Index

Figure 7 shows an example of this descriptor calculated for two samples played by 2 different performers. The zoomed detail displays the transition a note repetition (F5). Repeated notes in clarinet are usually articulated with an interruption on the air flow. A high value in this descriptor is related to strongly detached notes. We can also observe is the similarity of values of the index in this transition for both performers, which might indicate similar articulation intentions.

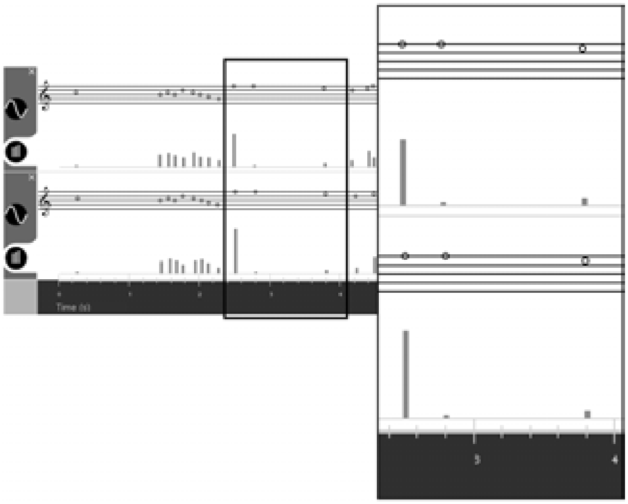


Figure 7. Example of the Articulation Index

4.3.2 Example: Legato Index

The values of this index are related to energy continuity along note transitions and may reflect acoustic characteristics of the clarinet, especially in relation to fingering configuration. Figure 8 shows two groups of fast note transitions in the same register of the instrument (4/4 measure at 120 BPM). In Group 1 we notice lower index values for the transition from C5 to B4 (2nd and 5th transitions) that might be related to the fact that B4 activates the longest vibration mode of the instrument, which is reached by padding the largest and furthest below positioned hole of the instrument, an action more difficult to be accomplished than those presented by other transitions in this passage.

In Group 2 we can observe lower index values in the transition from A5 to F5 (6th transition), which requires the closing of 2 holes simultaneously using fingers of distinct hands, the most complex finger movement of this passage. These results show that this index may present potentials to reveal inherent characteristic of specific note transitions, which may related to aspects such as fingering complexity.

The verification of these assumptions still needs further investigation, which is out of the scope of this paper.

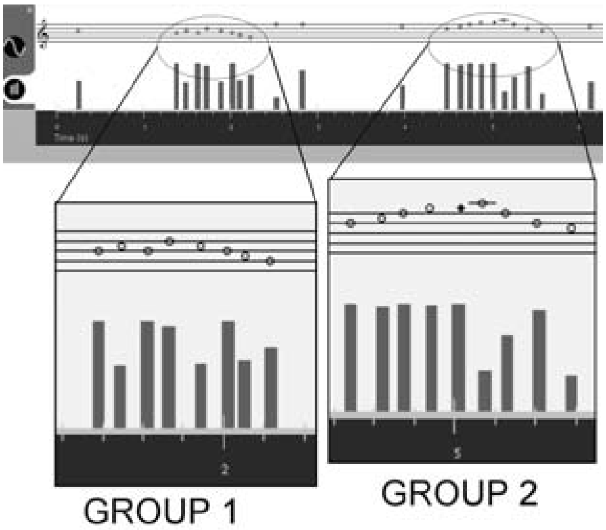


Figure 8. Example of the Legato Index

5. CONCLUSION AND FURTHER WORKS

Expan is intended to automatize the extraction of descriptors parameters and facilitate musical expressiveness analysis methodologies in dynamics situations such as musical instrument lessons. It was developed in MATLAB. In this first version, the processing time for a sample 10 seconds long takes about half minute in a mid-end personal computer. The processing time will be improved with the migration to other languages such as C++, which could lend to a real time implementation.

6. ACKNOWLEDGEMENTS

We would like thank the clarinets who participated on the experiments and the interaction designer Daniel L. Silva for helping with the interface design. This project was made possible with the sponsorship of the National Council for Scientific and Technological Development - CNPq.

7. REFERENCES

[1] De Poli, G. and Mion, L., "From audio to content," In Unpublished *book*, G. De Poli, Ed. Padova: Dipartimento di Ingegneria Dell'Informazione - Università degli Studi di Padova, 2006.

[2] Gabrielsson, A., "Music Performance Research at the Millenium," *Psychology of Music*, vol. 31, pp. 221-272, 2003.

[3] Juslin, P. N., "Cue utilization in communication of emotion in music performance: relating performance to perception," *Journal of Experimental Psychology: Human perception and performance*, vol. 26, pp. 1797-1813, 2000.

[4] Maestre, E. and Gómez, E., "Automatic Characterization of Dynamics and Articulation of Expressive Monophonic Recordings," in *Audio Engineering Society 118th Convention Paper*, pp. 8, Barcelona, 2005.

[5] McAdams, S., "Perspectives on the Contribution of Timbre to Musical Structure," *Computer Music Journal*, vol. 23, pp. 85-102, MIT Press, 1999.

The acoustic consequences of movement restraint on six professional western contemporary popular singers

Gemma Turner
Australian Centre for Applied Research in Music Performance
Sydney Conservatorium of Music
University of Sydney, Australia
gtur8984@uni.sydney.edu.au

Dianna T. Kenny
Australian Centre for Applied Research in Music Performance
University of Sydney, Australia
61 2 0425 358 275
d.kenny@usyd.edu.au

ABSTRACT
In western contemporary popular (WCP) styles of singing, high sound levels must be achieved, usually in the presence of a high level of body movement in performance. However, in recording studios, under stage direction or in the process of participation in empirical voice research, singers are often required to stand still. There is an assumption that singers will be able to achieve the same sound output under these conditions. This study assessed whether the restraint of body movements of six professional WCP singers would alter their acoustic output, in this case, their sound pressure level (SPL). 3-D video footage of the body movements were collected simultaneously with sound recorded through a head-mounted microphone while all sang the same R & B song in two conditions, with and without movement. The recordings were analysed for SPL and the results of the two conditions were compared. There was a significant reduction in the SPL at peak points in the 'no movement' condition but also throughout the phrase. This indicates that WCP singers need to be cautious about excessive restraint of their body movements if they wish to achieve optimal singing outcomes.

Keywords
Contemporary popular singing, body movement, SPL

.

1. INTRODUCTION
It is commonly assumed that WCP singers should be able to reproduce their usual stage standard in terms of vocal performance while standing still. This is exemplified by the concept taught by many classical pedagogues of the "noble posture", a stance that is maintained throughout the entire breath cycle [13] even though it has been observed empirically that posture differs during inspiration and expiration [9] and in particular just prior to a sudden increase in sound output, a phenomenon referred to as prephonatory posturing [19, 27]. In some prior singing studies, static postures have been studied, largely with regard to the issue of body alignment, particularly of the head and neck [31]. Although adjustments from one posture to the next would of necessity require body movement,

this phase of behaviour has been rarely noted or discussed, possibly because the classical singers observed in most studies do not move as much as WCP singers but also because of technological limitations. New technologies for observing body movement may change this in the future but still present technical problems such as altering subject behaviour and presenting synchronisation problems with sound recording [7].

However, significant anterior-posterior (AP) movements by a highly disciplined professional classical soloist were noted in a methodological study on microphone use for voice research [1] indicating that even classical singers may be moving more than previously thought. To our knowledge, prior to our research, the effects of movement restraint on sound output have not been tested.

Sound pressure levels (SPL - measured in decibels) is a common measure of voice function [2]. High SPLs are required within WCP music [3]. WCP singers achieve high levels within these styles using higher levels of muscle effort at the larynx [16] and higher internal pressures [22]. This combination creates a higher vocal load for WCP singers than for classical singers at equivalent sound levels [23], which can cause voice damage if not approached correctly [10]. This makes knowledge gained from the study of the physiological production of SPL of direct practical use to singers wanting to avoid vocal injury.

This study challenges the assumption that a static physical posture can be used in all singing situations on the basis of knowledge of the interactions of upper body behaviour with respiratory [6, 27], laryngeal [15] and vocal tract function [11 & Shimada, 1999, 18]. It is well-established that subglottal pressure (Ps) is the main factor in SPL regulation [25] and that this is controlled by interactions between the respiratory and laryngeal systems [21]. There is a linear relationship between subglottal pressures generated beneath the vocal folds in the respiratory system and SPL variation, which means that higher Ps is required with increases in SPL [26]. SPL can be voluntarily altered on three levels - in the vocal tract, at the larynx and in the respiratory system [28], elements which, unfortunately, are impossible to separate fully without multiple, simultaneous and invasive procedures which would impede movement and create discomfort in the singer [4]. An example of this would be direct measurement of Ps by tracheal puncture [5] . Other indirect methods of calculating Ps have been of questionable accuracy [20] and have only been conducted when subjects were standing still. In view of the preliminary nature of this study it was considered that observation of SPL levels in different conditions would give us an indirect indication of the effect of movement on Ps production if all other factors at the point of comparison (pitch, vowel, vocal register) remained the same [24] . Information gained could then be applied to more highly controlled studies in the future. We conducted a preliminary study in which we examined the singers' movement behaviours and the relationships between these and SPL variation [29]. We analysed the antero-posterior (AP) torso movement direction frequency and characteristics at the point of maximum sustained SPL and found that the most common movement of the torso was in a posterior direction. It was also found that this was more likely to occur at high SPL points than low SPL points to a statistically significant degree. Results showed that there was an association between the AP torso movement of WCP singers and dynamic variation indicating a possible function of the movement in the voice production of SPL variation. One function of movement may be that it facilitates abdominal muscle contraction to raise Ps [29].

In light of this knowledge, we then hypothesized that if the posterior torso movement on the high SPL note was being used to create the high SPL, restraining the singers from moving would reduce SPL at those points. We also decided we would need to measure longer samples for SPL to assess whether movement restraint would have an effect on dynamics at other points in a sung phrase with points of both high and low vocal loading.

2. METHOD

Participants were six singers of contemporary popular music, 1 male and 5 female, ranging in age from 21 to 46 years with between 5 and 28 years of professional experience of both stage and recording studio performance.

To avoid using cues that might trigger atypical SPL output strategies [12] or body postures [31], verbal instructions were minimized by providing a vocal model of an R & B song on a recording to all participants to learn. Singers were permitted to alter the key of the song to suit their habitual range and style of delivery. However, they were requested to maintain a similar dynamic range to the original recorded version provided to them. Singers were encouraged to sing and behave as if in performance. No other verbal cues were given. A head-mounted microphone was used to keep the singers' distance from the microphone constant.

After singing in the first condition, the singers were then requested to sing again, behaving as if they had been given the instruction to remain still in a spotlight by a director. This imagery was chosen to provide a realistic musical rationale for a stationary performance.

3. RESULTS

One line of the song excerpt contained notes of both high pitch and highest SPL, therefore representing the most demanding section of the song in terms of maximal voice production. This was referred to as the Peak Phrase (PP). SPL was calculated on this line. A note of the highest SPL within the PP was referred to as Peak Notes (PN) in the subsequent analysis.

For all singers, the maximum SPL attained on a PN occurred in the M condition. Individually, singer mean differences on the PN ranged from at 0.96 dB which is at just noticeable difference (JND) [32] to 4.6 dB making the output reductions in the NM condition acoustically significant. SPL values and difference scores (N-NM) were consistently lower in the NM condition for each singer. The mean difference was 2.37dB (SD=1.3); t=4.46 (df=5), p (two-tailed) =.007. The paired t-test assessing the averaged differences for all six singers between the M and NM conditions, averaged over the three takes in each condition showed that the mean difference of 2.87dB (SD=1.3) was statistically significant [t=4.42 (df=5), p (one-tailed) =.039].

SPL calculations of the entire Peak Phrase in both conditions can be seen in Figure 1.

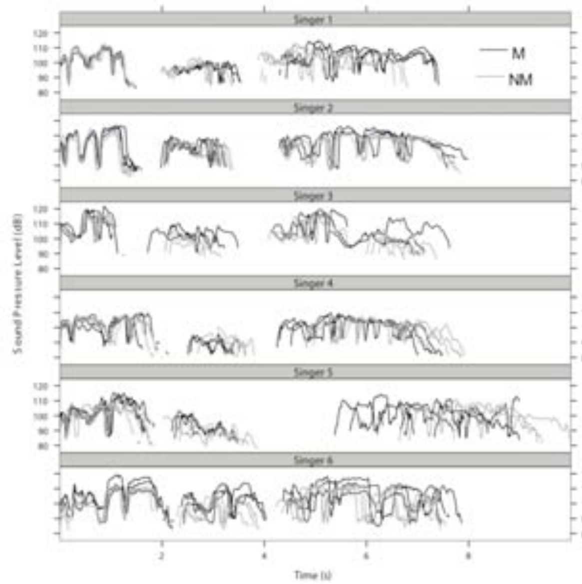


Figure 1. SPL of Peak Phrase in Movement (M) and No Movement (NM) conditions

SPL output was reduced at many points in the phrase in the NM condition and not just at points of high SPL.

4. DISCUSSION

The results indicated that vocal SPL is reduced when movement is reduced. This occurred even though the singers were not restricted in the way they could sing the song, and were permitted to lower the key in accordance with their habitual pitch range. Even though the singers sometimes reduced the two-octave pitch range of the song, thereby minimizing the vocal load and making the song easier to sing, the relationship between movement and SPL remained significant. The singers varied widely in age and length of experience and yet all experienced an audible reduction in SPL indicating that this effect was robust.

These results indicate that there may be either a direct physical function in the movement of these singers that assists with the production of the high SPL required for these styles of singing or that the manner of singing is entrained with movement in such a way that it cannot be easily separated without a reduction in SPL. The fact that SPL was reduced at lower pitches and at quieter points in the dynamics where Ps demands are lower indicates that the effects of movement restraint are not solely related to interference with Ps generation.

There are several possible explanations for this effect, which we have discussed in more detail elsewhere [30]. One possibility is that the NM condition interfered with the autophonic response, whereby the SPL of one's own voice is regulated on a multisensory level by factors such as the sensation of muscle effort [8, 14, 17]. The restraint of active body behaviour may have interfered with that response by altering the sensory feedback required for the mechanism to work. This would also explain the alterations at all SPL levels.

5. CONCLUSION

The results of this study indicate that restraint of body movement leads to reduced SPL levels in WCP singing**.** These results have significant practical implications for those working with singers during the conduct of voice research, in professional recording studios and on stage in that the placement of physical constraints on singers may have detrimental effects on SPL control that may be insurmountable even in very experienced singers. Those working with WCP singers need to be cautious in restraining their body movements if they wish optimal singing results to be achieved with regard to sound level.

6. REFERENCES

[1] Cabrera, D., Davis, P.J., Barnes, J., Jacobs, M. and Bell, D. Recording the operatic voice for acoustic analysis. *Acoustics Australia*, *30* (3). 103-108.

[2] Corthals, P. Sound pressure level of running speech: percentile level statistics and equivalent continuous sound level. *Folia phoniatrica et logopaedica*, *56* (3). 170-181.

[3] Dibben, N. The role of peripheral feedback in emotional experience with music. *Music Perception*, *22* (1). 79-115.

[4] Finnegan, E.M., Luschei, E.S. and Hoffman, H.T. Estimation of alveolar pressure from direct measures of tracheal pressure during speech. *NCVS Status and Progress Report* (June). 1-10.

[5] Finnegan, E.M., Luschei, E.S. and Hoffman, H.T. Modulations in respiratory and laryngeal activity associated with changes in vocal intensity during speech. *Journal of Speech, Language, & Hearing Research*, *43* (4). 934-950.

[6] Gandevia, S.C., Butler, J.E., Hodges, P.W. and Taylor, J.L. Balancing acts: Respiratory sensations, motor control and human posture. *Clinical and Experimental Pharmacology and Physiology*, *29* (1-2). 118-121.

[7] Giovanni, A., Assaiante, C., Galmiche, A., Vaugoyeau, M., Ouaknine, M. and Le Huche, F. Forçage vocal et posture: études expérimentales chez le sujet sain [Vocal forcing and posture: experimental studies on normal subject] [French]. *Revue de Laryngologie Otologie Rhinologie*, *127* (5). 285-291.

[8] Ho, A.K. Volume perception in Parkinsonian speech. *Movement Disorders*, *15* (6). 1125-1131.

[9] Hodges, P.W., Gurfinkel, V.S., Brumagne, S., Smith, T.C. and Cordo, P.C. Coexistence of stability and mobility in postural control: evidence from postural compensation for respiration. *Exp Brain Res*, *144* (3). 293-302.

[10] Hollien, H., Report on vocal registers. in *Stockholm Music Acoustics Conference*, (Stockholm, 1983), Royal Swedish Academy of Music

[11] Honda, K., Hirai, H., Masaki, S. and Shimada, Y. Role of vertical larynx movement and cervical lordosis in F0 control. *Language and Speech*, *42* (4). 401-411.

[12] Huber, J.E., Chandrasekaran, B. and Wolstencroft, J.J. Changes to respiratory mechanisms during speech as a result of different cues to increase loudness. *Journal of Applied Physiology*, *98* (6). 2177-2184.

[13] Hudson, B. The effects of the Alexander technique on the respiratory system of the sing/actor Part II: implications for training respiration in singer/actors based on concepts of Alexander technique. *Journal of Singing*, *59* (2). 105-110.

[14] Ingham, R.J., Warner, A., Byrd, A. and Cotton, J. Speech effort measurement and stuttering: investigating the chorus reading effect. *Journal of Speech, Language, and Hearing Research*, *49* (3). 660-670.

[15] Kooijman, P.G., de Jong, F.I., Oudes, M.J., Huinck, W., van Acht, H. and Graamans, K. Muscular tension and body posture in relation to voice handicap and voice quality in teachers with persistent voice complaints. *Folia phoniatrica et logopaedica*, *57* (3). 134-147.

[16] Koufman, J.A., Radomski, T.A., Joharji, G.M. and Russell, G.B. Laryngeal biomechanics of the singing voice. *Otolarygology - Head and Neck Surgery*, *115*. 527-537.

[17] Lane, H.L., Catania, A.C. and Stevens, S.S. Voice level: autophonic scale, perceived loudness and effects of sidetone. *Journal of the Acoustical Society of America*, *31* (2). 160-167.

[18] MacCurtain, F. and Welch, G., Vocal tract gestures in soprano and bass: a xeroradiographic-electro-laryngographic study. in *SMAC (Stockholm Music Acoustics Conference) 83*, (1983), Royal Swedish Academy of Music, 219-238.

[19] McFarland, D.H. and Smith, A. Effects of vocal task and respiratory phase on prephonatory chest wall movements. *Journal of Speech and Hearing Research*, *35* (5). 971-982.

[20] Plant, R.L. and Hillel, A.D. Direct measurement of subglottic pressure and laryngeal resistance in normal subjects and in spasmodic dysphonia. *Journal of Voice*, *12* (3). 300-314.

[21] Sawashima, M., Niimi, S., Horiguchi, S. and Yamaguchi, H. Expiratory lung pressure, airflow rate, and vocal intensity: data on normal subjects. in Fujimura, O. ed. *Vocal Physiology: voice production, mechanisms and functions*, Raven Press Ltd, New York, 1988, 416-422.

[22] Schutte, H.K. and Miller, D.G. Belting and pop, nonclassical approaches to the female middle voice: some preliminary considerations. *Journal of Voice*, *7* (2). 142-150.

[23] Sundberg, J., Gramming, P. and Lovetri, J. Comparisons of pharynx, source, formant, and pressure characteristics in operatic and musical theatre singing. *Journal of Voice*, *7* (4). 301-310.

[24] Sundberg, J., Leanderson, R., von Euler, C. and Knutsson, E. Influence of body posture and lung volume on subglottal pressure control during singing. *Journal of Voice*, *5* (4). 283-291.

[25] Sundberg, J., Titze, I.R. and Scherer, R. Phonatory control in male singing: a study of the effects of subglottal pressure, fundamental frequency, and mode of phonation on the voice source. *Journal of Voice*, *7* (1). 15-29.

[26] Tanaka, S. and Gould, W.J. Relationships between vocal intensity and noninvasively obtained aerodynamic parameters in normal subjects. *Journal of the Acoustical Society of America*, *73* (4). 1316-1321.

[27] Thorpe, W.C., Cala, S.J., Chapman, J. and Davis, P.J. Patterns of breath support in projection of the singing voice. *Journal of Voice*, *15* (1). 86-104.

[28] Titze, I.R., Shirley, G.I., Long, R., Stathopoulos, E., Ramig, L.O., Carroll, L.M. and Riley, W.D. Messa di voce: an investigation of the symmetry of crescendo and decrescendo in a singing exercise. *NCVS Status and Progress Report* (June). 145-152.

[29] Turner, G. and Kenny, D.T. A preliminary investigation into the association between body movement patterns and dynamic variation in Western contemporary popular singing. *Musicae Scientiae*, *In press*.

[30] Turner, G. and Kenny, D.T. Restraint of body movement potentially reduces peak SPL in western contemporary popular singing. *Musicae Scientiae*, *In press*.

[31] Wilson Arboleda, B.M. and Frederick, A.L. Considerations for maintenance of postural alignment for voice production. *Journal of Voice*, *In press*.

[32] Zwicker, E. and Fastl, H. Just-Noticeable Sound Changes. in Zwicker, E. and Fastl, H. eds. *Psycho-acoustics: facts and models*, Springer, 1999, 176.

The Occurrence of Music Performance Anxiety in Early Childhood

Hélène Boucher
McGill University
2840, Lyall
Montréal (Québec, Canada)
011-514-705-9958
helene.boucher@mail.mcgill.ca

Not available in the web proceedings

Not available in the web proceedings

Postmodernism and Musicology – A Methodical Misunderestimation

Jan Giffhorn
Univ. for Music and Performing Arts
Vienna
Gabelsbergergasse 2/18,
1020 Vienna, Austria
0043 681 1084 6372
giffhorn@gmx.net

ABSTRACT

This contribution deals with the consequences of the postmodern turn on the musicological disciplines. Due to the dominant position of historical musicology, it was not until the early nineties that a postmodern branch, the so-called 'New Musicology', could arise. An examination of a dispute between C.Rosen, dogmatic proponent of the historical section, and L.Kramer, proponent of the New Musicology, reveals that their conflict is largely based on methodical matters; the unhistorical and rather subjective approach of New Musicology to phenomena such as musical meaning, musical experience and musical performance contrasts with the modernist claim for objectivity. But even the traditional approach turns out to be contradictory. The merits of postmodern strategies and concepts in musicology lie in widening the methodical horizon and in entailing interdisciplinarity, so that the boundaries between systematical and historical musicology and ethnomusicology become permeable.

Keywords

Postmodernism, Musical Meaning, Musical Experience, Historical Musicology, Systematic Musicology, New Musicology, Rosen, Kramer, McClary, Eggebrecht

1. INTRODUCTION

In 1986, German musicologist Hans-Heinrich Eggebrecht proclaimed:

"Everything scientific which is directed toward music is (or should be) a science that is concerned with history; that is to say, a historical musicology (music-science), and everything that cannot be understood in these terms is merely an aid to this end." cited in [3, p.69], translation: Erik Leidal.

The essence of this statement is an undeniable dominance of historical musicology; the historical discipline is to play the lead whereas all others, like systematic musicology or ethnomusicology, are in the best case assistants. Yet the naivete of Eggebrecht's statement is impressive. All the more, considering that the 'postmodern turn' had already massively affected the humanities on a huge scale. 1986: That is twenty years after Foucault's *The Order of Things,* eight years after of Lyotard's *The Postmodern Condition*, three years after Sloterdjik's *Critic of Cynical Reason,* and only three years before Fukuyama's *The End of History?* It was not until the early nineties that postmodernist thought began to influence musicology. This contribution aims to research why musicology had been ignorant towards new influences.

2. POSTMODERNISM

2.1. In Default Of A Serious Definition

This is only one of many 'pomophobe' jokes:

"How many postmodernists does it take to screw in a light bulb? Even the framing of this question makes a grid of patriarchal assumptions that reveals a slavish devotion to phallocentric ideas – such as, technical accomplishment has inherent value, knowledge can be attained and quantities of labor can be determined empirically, all of which makes a discourse which further marginalizes the already disenfranchised."

Another fascinating 'object of interest' is the so-called "Postmodernism Generator", an online-application (www.elsewhere.org/pomo) written by Andrew C. Bulhak from Monash University. It is able to generate texts in a postmodern style immediately. Having read one of the instantly produced essays, one is told that "the essay you have just seen is completely meaningless and was randomly generated by the Postmodernism Generator…This installation of the Generator has delivered 3.875.992 essays since 25/Feb/2000 18:43:09 PST, when it became operational." – Given the fact that googling the term "Postmodernism" lead only to 2.640.000 results, the amount of nearly 4.000.000 arbitrarily created and hollow essays is surprising, if not shocking.

2.2. Postmodern Features

It is still controversial among scholars as well as among artists whether the term 'postmodernism' refers to an epoch, a style, a condition, a cultural-political preference, an intellectual movement or simply an attitude [2, p.20-28] [17, passim]. Moreover, it is still at issue whether postmodernism is a further stage of modernism or whether it is the overcoming of modernism. Another option is to construe postmodernism as countermovement to modernism, as anti-modernism [16, p.102ff]. It is impossible to name all features of postmodernism, but some of them have already been subtly portrayed in the joke at the beginning of the text: "Patriarchal assumption" juxtaposed with "slavish devotion to phallocentric ideas" alludes to the postmodern criticism of power and the key role of sex, gender and sexuality; the statement that "technical accomplishments" do not have "inherent value" criticizes one of the central ideals of modernism. Moreover, the possibility of knowledge is challenged, it cannot "be attained." Discarding the possibility of objective knowledge and truth in favor of a subjectivity, postmodernism brings into focus the conditions and circumstances of the attempt of gaining knowledge; if knowledge is subjectively 'contaminated', the circumstances of this very process become the topic of research and part of a provisional truth. The modern ideal of objectivity is replaced by rather cultural constructions.

2.3. The 'Postmodern Turn' In Musicology

The entering of postmodern thoughts, methods and strategies into musicology was an affront to all that musicology had come to stand for. It meant to discard the 'Eggebrecht-doctrine' in favor of a concept in which terms like 'history' or 'progress' are no longer of significance.

Instead of assuming a formal and structural core of music which was said to bear a musical meaning, postmodern musicology focuses on the role of the listener and performer. The truth is no longer embedded exclusively in the piece but in the social environment. The composer's aims are not privileged to any further towards those of the interpreter or even of the listener: Relevant meaning is created in a performance. Thus, according to postmodern musicology, musical expression and musical meaning cannot only lie in formal aspects of a musical work and therefore be independent form structural considerations. That is why the extramusical elements gain importance in musical analysis. Due to the importance of the claim that subjectivity replaces the autonomy of objectivity, the distinction between high and low culture is no longer pertinent. New methods are necessary which do not respect the boundaries of traditional musicology. The resulting 'New Musicology' widens the methodical horizon and refers to interdisciplinary approaches. The boundaries between historical and systematical musicology and ethnomusicology become permeable, for example in connecting sociological and structural dimensions of musical works with each other. [1], [14], [15]

The following 'case' reveals some of the problems that result from the conflict between old and new musicological approaches.

3. THE CASE ROSEN AGAINST KRAMER

In June 1994, Charles Rosen, author of rather historical books like *The Classical Style* (1971) or *The Romantic Generation* (1998), published an essayistic review in The New York Review of Books. In his review of several book Rosen took the opportunity to deal with the status quo of historical musicology and the rise of New Musicology. Rosen is taking offense at one particular proponent of the New Musicology, namely Lawrence Kramer who later replied aggressively to Rosen's "attacks", as well as Rosen did in turn to Kramer's replique.

One might argue that this source is of little relevance for the present discussion because it has been published fifteen years ago; it might be outdated, especially since quite a lot of publications dealing with musicology and postmodernism was yet to come at that point, among them Kramer's *Classial Music and Postmodern Knowledge* (1995) or Susan McClary's *Conventional Wisdom – The Content of Musical Form* (2001). But the debate between traditional and New Musicology, as it becomes apparent in these articles, is still unresolved in academic discourse. For more see [1], [3], [6], [11].

3.1. Old Vs. New Musicology

At the outset of his essay, Rosen's perspective is characterized by an ambivalent attitude towards necessary renewals and reforms in musicology. On the one hand, he recognizes a "real intellectual crisis" which is mainly due to a separation of musicological scholarship and the 'real life'. According to Rosen, New Musicology grew out of a desire to close this gap. Thereby, the new approach seems to be an appropriate means "to drag musicology out of its isolation." [all Rosen-quotes refer to [12]

But on the other hand, Rosen leaves no doubt that he is extremely suspicious of New Musicology, particularly of its methods and topics of research. Criticizing Kramer Rosen uncovers a disproportion between Kramer's musical grasp on the one side and his outstanding scholarly versatility on the other. Although he emphasizes Kramer's "passion for new ideas and the facility with which he can juxtapose music and contemporary developments in other fields", he leaves no doubt about Kramer's "weak grasp of the experience of music". After disenchanting Kramer's analysis of the opening of Haydn's *Creation*, Rosen sums up his view towards Kramer's inability to offer a close reading in the statement that "his grasp of cultural history, and his evident love of music and delight in its manifestation are not matched by a sensitivity to the ways in which music can be perceived rather than analysed on paper." This affects his earnest attempt "to unite formal analysis and cultural history."

With regards to the tasks of New Musicology, Kramer agrees with Rosen to great extent, claiming that New Musicology "seeks to retrieve music from its traditional self-containment and to uncover its many worldly meanings" [all Kramer-quotes refer to 10]. But he points out that the "nostalgic" Rosen belongs to the 'old school' and does not understand the new concepts. According to Kramer, Rosen "speaks for what was once the dominant culture of classical music, a culture that too eagerly monumentalizes what it values." In order to clarify his point he focuses on two central concepts: (musical) meaning and (musical) experience: "Rosen grounds musical understanding in a normative Musical Experience in which what is heard and what is true are synonymous and simultaneous."

Another major point for discussion between them is the function of contextualisation. Among the new methods of New Musicology, which Rosen criticizes, is Kramer's contextualization as part of a "defective" theory. Nothing can substitute for the merit of a close reading. And contextualizing in Rosen's mind is nothing but a distraction from that goal; thus, contextualisation is not an appropriate means of analysis. Rosen refers to an argument about close reading between Kramer and Tomlinson, which had been published in *current musicology* [9, p.18-40]. In the article Kramer considers close reading as an integral part of an analytical process, saying that its function is "to trace out the interrelations of musical pleasure, musical form and ideology." Tomlinson is of the opinion that it is of no use at all, which in Rosen's mind exposes Tomlinson as one who fears modernism per se: "He [Tomlinson] cannot accept the term objective, because today that is a word equivalent to 'modernist', 'postitivist', 'aestheticist' and 'phallocentric'." Rosen holds against them that a close reading is not only necessary but that "the intense concentration on the text of a work of music is a joy, sensuous as well as intellectual." So if Kramer is right, "close reading" or formal musical analysis is of little or perhaps even of no importance at all.

3.2. Contradictions

Thus, the basic conflict comes down to what constitutes valuable intellectual inquiry, in particular about discovering a musical meaning.

This is no small matter; throughout their debate, one can see how both 'candidates' intend to reveal their opponents dogmatism: Rosen accuses Kramer of abusing the concept of contextualisation in order to disguise his lacking practical ability to analyse structurally. That is, according to Rosen, Kramer is only simulating knowledge with the help of a method that was invented solely for this purpose. Kramer on the other hand insinuates that Rosen's approach to musical experience is a little more than naive. According to Kramer, Rosen wants to believe in a normative way of hearing music and endow the musicologist with the authority necessary to begin the pursuit of objective knowledge.

Indeed, I find the last point to be apt. There is a subtle yet distinct arrogance in Rosen's words. His tone is polite, but often patronizing; and sometimes even contradictory: He radically criticizes his opponent as a kind of amateur, but he states that the public's interest in Kramer's writings is "a hopeful sign". He mentions that "no interpretation is likely to be definitive" [13], deals with the "inherent instability of musical meaning" and complains about "composers, performers, and listeners [who] are being shut out by the latest trends." Yet Rosen's liberal statement that he does not want "to prescribe the way music can be listened to" seems more tactically than honest. On the one hand, Rosen's claim for objectivity and normativity cannot be denied; on the other, his statements sound similar to one of the most famous 'slogans' of postmodernism, "Cross the border – close the gap!"[7, p.57] His argument is neither fish nor fowl: He ascribes to tenants of a historical paradigm that has lasted for decades, if not centuries, but he feels the crisis of his discipline and fears the subversive aspects inherent to New Musicology because they could put an end to the supremacy of the modernist world view.

4. NEW WORLD ORDER?

New Musicology has never intended to be 'everybody's darling'; perhaps, it has always been its intention to provoke debate with traditional musicology. In this regard, postmodern musicology is much more anti-modernist than postmodern philosophy.

The fruits of the labour of New Musicology have been partly successful in that once firmly defined boundaries are now more porous and interdisciplinarity is common place.

And even Eggebrecht was somehow aware of the dubious nature of his position. In the chapter *Über die Gliederung der Musikgeschichte* in one of his last books *Musik im Abendland*, he states:

"History is something that exists only because we think it and tell it … of course no one believes me when I say that, and I don't believe it completely myself – well, only a bit. From this bit, however, I assume for the time being – more in jest – that it's the truth." [4, p.169], translation: Erik Leidal.

It appears that there is according to Eggebrecht enough room in this discourse to allow place for a Subject.

REFERENCES

[1] Babich, B. E., Postmodern Musicology, In: V. E. Taylor, C. Winquist (Eds.), *Encyclopedia of Postmodernism*, Routledge, New York, 2001.

[2] Behrens, R. , *Postmoderne,* Europäische Verlagsanstalt, Hamburg, 2004.

[3] Butt J., The Postmodern Mindset, Musicology and the Future of Bach Scholarship, In: *Understanding Bach*, Vol.1, 2006, www.bachnetwork.co.uk/ub1/butt.pdf (Sept.30, 2009).

[4] Eggebrecht H.H., *Musik Im Abendland,* Durchgesehene Taschenbuchausgabe, Piper, München, 1996.

[5] Elschek, O., Systematic Musicology – anachronism or challenge? A perspective on disciplinary and scientific developments, In: Albrecht Scheider (Ed.), *Systematic and cognitive Musicology. Concepts, Methods, Findings.* Reihe Hamburger Jahrbuch für Musikwissenschaft, Band 24, Peter Lang Verlagsgruppe, Frankfurt a.M. et al., 2008, 63-87.

[6] Emmerson, S., *Where next? New music, new musicology,* Music, Technology and Innovation Research Centre, De Montfort University, Leicester, UK, 2007, www.ems-network.org/IMG/pdf_EmmersonEMS07.pdf pdf (Sept.30, 2009).

[7] Fiedler, L., Überquert die Grenze, schließt den Graben!, In: Wolfgang Welsch, *Wege aus der Moderne – Schlüsseltexte der Postmoderne-Diskussion,* Akademie Verlag, Berlin, 1994, p.57-74.

[8] Foxon, T. *Some thoughts on 'Critical Musicology'*, 2005, http://www.musicalresources.co.uk/SomethoughtsonCriticalMusi.php (Sept.30, 2009).

[9] Tomlinson G., Musical Pasts and Postmodern Musicologies: A Response to Lawrence Kramer, In: Goering, E.J. (Ed.), *Current Musicology,* No. 53 'Approaches to the disciplines', 1992, 18-41.

[10] Kramer, L., In Response to 'Music à la Mode', In: *The New York Review of Books,* Vol.41, No.15, Sept. 22, 1994, http://www.nybooks.com/articles/2143 (Sept.30, 2009).

[11] Lochhead, J./Auner, J. (Eds.), *Postmodern Music/Postmodern Thought,* Routledge, New York, London, 2002.

[12] Rosen C., Music à la Mode, In: *The New York Review of Books,* Vol.41, No.12, June 23, 1994, http://www.nybooks.com/articles/15080 (Sept.30, 2009).

[13] Rosen C., Reply to Lawrence Kramer 'In response to 'Music à la Mode', Vol, 41, No.15, Sep.22, 1994, http://www.nybooks.com/articles/2143 (Sept.30, 2009)

[14] Smith C., Aspects of Modernism and Postmodernism in Music, In: *Ars Nova – Unisa Musicologica,* Vol. 33&34, 2001/2002, 36-40

[15] Walker J., The Deconstruction of Musicology – Poison or Cure?, In: *Music Theory Online*, Vol. 2, No.4, 1996, http://mto.societymusictheory.org/issues/mto.96.2.4/mto.96.2.4.walker.html.

[16] Welsch W., *Ästhetisches Denken,* Reclams, Stuttgart, 1990/^{2}2003.

Towards Human-Machine Sonic Interactions: Mapping Human Dance Gestures onto Robots

João Lobato Oliveira
INESC-Porto/LIACC-FEUP
Rua Dr. Roberto Frias
4200-465 Porto, Portugal
jmldso@gmail.com

Luiz Naveda
IPEM-Dept. of Musicology
Rozier, 44
9000 Ghent, Belgium
Luiz.Naveda@Ugent.be

Fabien Gouyon
INESC-Porto
Rua Dr. Roberto Frias, 378
4200-465 Porto, Portugal
fgouyon@inescporto.pt

Luis Paulo Reis
LIACC-FEUP
Rua Dr. Roberto Frias, s/n
4200-465 Porto, Portugal
lpreis@fe.up.pt

ABSTRACT
The presented work is founded on the loop between sound, rhythmic perception, movement, and interaction through dance expression. It consists on the study of human dance movements, by capturing performed individual dancing sequences, and mapping them to a humanoid robot, considering the robot's corporeal constraints and its sensorimotor topology. This shall promote more natural human-robot social interactions, founded on the intimate dependence of the coordination between sounds and movements of two interactors. At a higher level this study intend to investigate the effects of rhythmic synchrony in dance-oriented human-robot social interactions as an alternative to design robots that can perceive and behave according to natural rhythms in more open-ended interactions.

Keywords
Human Motion Capture, Dance, Music, Rhythm, Robotics, Artificial Intelligence, Human-Robot Interaction.

1. INTRODUCTION
Robotics researchers worldwide have devoted many years of work in the study of interactive skills for increasing the robots' sociability and enhancing their communication. More and more AI researchers are then looking for more natural and intimate forms of interaction which may depend on the subtle coordination imposed by sound and movement within two interactors [1], increasing the interest in the notion of embodiment applied to robotics and virtual agents.

It has been proposed that the relation between intelligence and the morphology of the body may rely in a dynamic interaction between body, brain and environment [2]. In this context, dance is seen as a collection of expressive gestural behaviours that emerge from a body's morphology, shaped by the corporeal responses to musical stimuli and cultural environment.

Psychological and sociological research even suggests that the embodiment of rhythm, through dance movements, is an elementary and pervasive organizing principle for social human interaction, representing a communication channel beyond spoken language [3].

Given such, we intend to develop social intelligent robotic agents that can map intermodal rhythms, perceived from environmental observations, to dance motions; by imitating others and applying an individual sense of creativity while following the musical rhythm. This paper exposes our research plans on studying the effects of rhythmic synchrony in dance-oriented human-robot social interactions, resulting from designing humanoid robots that can perceive and behave according to natural rhythms in more open-ended interactions.

Accordingly, this paper briefly describes our methodology for capturing human dance motion primitives and for mapping this motion capture data to humanoid robots, for the embodiment of dance in synchrony to musical rhythm. At a lower extent we may address the consequent problem of keeping stable sensorimotor signatures over different modalities [4]. To sustain the relational involvement our robotic system must ultimately alternates between dance movements strongly coupled to cross-modal rhythms, and more detached creative motions.

The paper structure is as follows. The next section presents a study on relevant techniques and implementations regarding dancing interactive robots. Section 3 briefly describes our research plan endeavouring the robot embodiment of real human dance movements, by mapping dance sequences previously extracted from human performances. Finally section 4 concludes this paper by presenting the main conclusions and describing our further research.

2. RELATED RESEARCH
Academically, dancing robots and human-robot musical interaction are moreover common terms. In an increasing number of research labs around the world, researchers follow a quest to find the perfect solution to achieve a rhythmic perceptive dancing robot that could interact with humans. Besides, recent generations of robots ever more resemble humans in shape and articulatory capacities [4] motivating researchers to design interactive dancing robots that can mimic the complexity and style of human choreographic dancing. This section presents some notable late solutions on (interactive) robot dancing, and robot audition methodologies; including virtual agents that also rhythmically interact with humans, over different modalities.

Nakazawa, Nakaoka *et al.*, [5], [6], from Tokyo University, presented an approach that lets a biped robot, HRP-2 imitate the spatial trajectories of complex motions recurring to a motion capture system. To do so, they developed a learning-from-observation (LFO) training method that enables a robot to acquire situated knowledge from observing human demonstrations, by relying on predesigned task models which represent only the actions that are essential to mimicry. This method was applied in the performance of Japanese traditional folk dance imitating a female human dancer. For such, and because the leg and upper body have different purposes as well as different motor constraints, the authors applied different strategies to design task models for leg and upper-body motions. The leg motions stably supported the robot body, while the upper-body motions expressed dancing patterns; concatenating and adjust these two motion types in the final stage. Finally to generate executable motion, considering balance issues, they used a dynamic filter to compensate the zero-moment point (ZMP) and the yaw-axis moment, and conducted skill refinement to resolve other kinematic problems, such as self-collision.

Despite the flexibility of motion generation, this method didn't account for the timing adjustment of dancing movements needed to interact with auditory environments, i.e., in response to musical rhythm. To overcome this limitation, later, Shiratori, Nakakoa *et. al* [7] concerned on modelling the proper modifications on upper body motion of dance performances to follow the given speed of the played music. Their model was based on the insights that high frequency components are gradually attenuated depending on the music speed, and that important stop motions are preserved even when high frequency components are attenuated. For that purpose, and to satisfy the joint limitations of the robot, the authors analyzed motion data at varying musical speeds by using a hierarchical motion decomposition technique – hierarchical B-spline. Following the referred previous work, all motion data was captured at 120 fps by an optical motion capture system, produced by Vicon, and each joint angle was calculated using quaternion algebra, and then converted into a 3D logarithmic space. Their final tests were also performed in HRP-2 showing successful upper-body robot dance performance at different speeds, while satisfying the criteria for balance maintenance.

Subsequently, Shiratori *et. al*, extended their previous work by presenting a new approach for synthesizing virtual (character) dance performances matched to input music, and based on emotional aspects [8], [9]. The dance was performed by combining the candidate motion segments whose rhythmic and intensity features match those retrieved from music. The motion features were also retrieved from a human dance motion capture database (based on the speed of a performer's hands, feet, and centre of mass), which also supplied the spline function for interpolating the best matched motion segment, consequently endeavouring human-like dance sequences.

Similarly, Penasse and Nakamura [10] presented a new approach to generate musical adaptive dance motions, by interpolating dance key frames synchronized to rhythmic onsets. Each dance motion primitive is so attached, in real-time, to the downbeats of the music, being then interpolated to express complete dance sequences. Their interpolation approach recurs to a cubic spline function that constrains motion patterns to the bi-dimensional space of their virtual character.

Other animated applications have been designed also focusing on the non-verbal bodily interaction between human subjects and virtual characters. Reidsma *et. al* developed a virtual dance partner, Virtual Dancer [11], as an AI embodied agent for keeping an autonomous dance-oriented interaction with humans. Their Virtual Dancer environment consisted on an agent that dances together with a human subject to the musical rhythm, by aligning motion keyposes with the beats of the music while adjusting itself to the user's movements. The moves of the virtual dancer are meaningfully chosen from a database previously obtained from a motion capture system. The multi-modal beat-onsets are detected in real-time by observing the movements of the human dancer and their feet activity, registered in a dance pad; engaged with the analysed musical beats. Their computer vision implementation extracts global characteristics about the movements of the human dancer, such how much he moves around or how much he waves with the arms; by analyzing the user's silhouette, his centre of mass and his radial activity. By alternating patterns of following the user or taking the lead with new moves, the system attempts to achieve a mutual dancing interaction where both human and virtual dancer influence each other.

In a more educational perspective, Tanaka *et al.* from Sony, used QRIO for developing a dancing robot that interacts with children [12], [13]. Their implementation used a posture mirroring dance method which, based on an Entrainment Ensemble Model (EEM), relies on cyclic repetitions of sympathetic and dynamic behaviours. To keep a synchronous interaction with children the authors used a "Rough but Robust Imitation" visual system through which QRIO mimics the detected human movements, in an entrainment process.

In 2007, Aucoutuier *et al.* [14] used a robot designed by ZMP, called MIURO, in which they built basic dynamics through a special type of chaos (chaotic itinerancy - CI) to let the dancing behaviour emerge in a seemingly autonomous manner, exhibiting a variety of periodic motion styles alternating from detached independent movements to others strongly attached to musical rhythm. The robot motor commands were generated in real time by converting the output from a neural network that processes a pulse sequence corresponding to the beats of the music. Each neuron was (biologically) inspired in the FitzHugh-Nagumo model to generate chaotic itinerant behaviour among low-dimensional local attractors through higher dimensional chaos. The resulting dancing revealed a strong comprise between synchronization and autonomy which ultimately enhance the long-term interest.

Advocating that in real natural environments, in order to convey meaningful musical understanding onto intelligent robots one must research more robust algorithms related to robotics audition [15], Takeda *et. al* studied alternative real-time beat trackers, robust enough to contour environmental and self noises (like motor gears), and its own musical expression performed through stepping, singing and scatting; while adapting to tempo changes [16]. For such requirements, their beat tracking algorithm applied a spectro-temporal pattern matching, with a short length window for tempo adaption, and a semi-blind ICA to improve the signal-to-noise ratio of the input audio signals (captured by embedded microphones).

Their beat tracker was then applied in distinct musical interactive experiments with humans through a set of validatory experiments, using Honda's ASIMO biped humanoid robot in the pursuit of a musical comprehensive and expressive robot. In the first test the robot stamped its feet in time with the musical beats [17]. In the second experiment a human subject clapped his hands (for one minute) and the robot scatted according to this clapping [18]. In a third final experience [19] the robot stepped in time with the subject's clapping while scatting, therefore separating beats from its own musical expression and self-noises. Their evaluation suggested that the implemented beat tracker was able to adapt to tempo changes and contour the noisy audio signals, as proposed. The experiments also shown that the accuracy of beat-tracking is of most importance to guide a human-robot musical interaction, turning it on easy and enjoyable relationships.

Focused on a more active perspective of robot action according to human intention, Takeda *et al.* [20] proposed a dance partner robot, referred to as MS DanceR, consisting on a platform for realizing effective human-robot coordination with physical interaction. MS DanceR consists on an omni-directional mobile robotic interface which moves along dance-step trajectories during a ballroom dance, including a force/torque sensor (Body Force Sensor) to realize compliant physical interaction between the robot and a human, by measuring the user leading force-moment. For estimating the next intended dance step according to the human lead at the transition, a Step Estimator module used a set of hidden Markov models (HMMs) to stochastically model the time series of the leading force-moment. According to step transition rules, based on motion constraints, the Motion Generator module then generated cooperative dancing movements in a close-loop physical interaction with the human subject.

Considering that rhythmicity is a holistic property of social interaction, Michalowski *et al.* [1], [21], [22] investigated the role of rhythm and synchronism in human-robot interactions, and their application in pedagogical and therapeutical scenarios. To do so they developed perceptive techniques and generated social rhythmic behaviours in non-verbal interactions through dance between Keepon, a yellow puppet-like robot [1], [21], or Roillo, a robotic virtual platform, [22], and children. For perceiving rhythm over different modalities, their implementation used musical signal-processing techniques to detect the musical tempo, clapping, or drum-beats; and computer-vision methods, and accelerometers (and pressure sensors) to enable the perception of repetitive movements by people's heads, arms, or bodies.

Ellenberg *et al.* [23] used Robonova as their humanoid robot dancing platform. In their implementation they applied a real-time beat tracker, based on Klapuri's algorithm [24], for synchronizing the robot's dance motion in phase with the music. Their application coordinated dance by choosing a random series of gestures, from a motion library, and linearly interpolating them, point-to-point, for generating a dancing sequence, while assuring the robot's balance. oreograph gesture sequences, extending the system to a basic choreographic tool.

As a first step beyond our robot dancing concept we already built a humanoid dancing robot, based on Lego Mindstorms NXT kits, capable of embodying synchronized dance movements in response to a low-level musical rhythmic perception model, based on real-time note onsets detection [25]. Through this workplan, we intend to give a step further and incite the performance of realistic human-like dance robotic movements. This behaviour shall promote more natural human-robot social interactions founded on the intimate dependence of the coordination between sounds and movements of two interactors.

3. WORK DESCRIPTION

This research mainly consists on the study of human dance movements, by capturing individual dancing sequences in real human dance performances. To this extent we intend to use motion capture state-of-the-art technologies for retrieving individual body segments' motion trajectories, and going through the full process of mining the gathered data to extract relevant information towards the embodiment of human-shaped robot dance movements. These dance movements may then be mapped to a humanoid robot, and subsequently aligned to the beats in the music with recurrence to a real-time beat tracking algorithm, implemented for such purpose; ultimately promoting synchronous human-robot sonic interactions, in an entrainment process.

3.1 MoCap Data Gathering and Extraction of Relevant Information

Optical Motion Capture is an efficient method for accurately tracking and recording 3D motions from any live (dance) performance.

By using mocap technologies, like the OptiTrack [26] system, from NaturalPoint, we should gather relevant human dance motion data that would be used as dance primitives.

In order to analyze and fully process each captured motion primitive we would also recur to proper software, that besides retrieving all motion capture data would allow for its proper analysis. Ultimately some built-in routines may be used for studying the rhythmic structure beyond the retrieved motion patterns and for exploring different degrees of the body's engagement with music.

Within this work module we may either retrieve dance keyposes, through a complete corporeal description; or full dancing sequences, describing closed-loop trajectories.

3.2 Robot Movements Mapping

At this point, we shall generate robot dance motion primitives by mapping the previously acquired motion data into a biped humanoid robot. This motion data is basically expressed in the robot as joint angle trajectories, consisting on transposing each captured marker, from dance motion primitives, to the correspondent robot's body junction (see Figure 1 for different robot joint types). The generation of complete dancing poses or sequences would then derive from the interconnection of all joint angles calculated by inverse kinematics of all captured marker positions. All this process must account for the robot body's constraints at the actuator's capacity, the mechanical structure and its limited degrees of freedom, still far from the more the ones expressed by the human body. In order to increase stability and overcome biped balancing issues some dynamics techniques may also be applied to reduce the amount of moments generated by movements of the robot upper body (see *sec. 3.4* for a detailed description over the embodiment process).

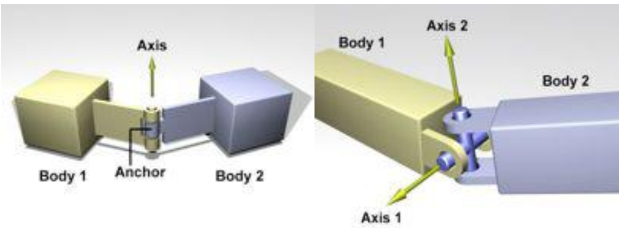


Figure 1. Robot actuator joints: a)Hinger joint; b)Universal joint.

All this process shall be firstly simulated in a proper robotic simulator modelling the dynamics beyond the robot body's motion in real environments. For this effect we should use SimSpark [27], a generic open-source robot simulator used in academic and education, developed as part of the RoboCup initiative, for simulating multi-agent systems in three-dimensional environments. For physical realistic dynamics it uses the Open Dynamics Engine (ODE) [28], which offers fast rigid body simulations, collisions detection, and supports the use of articulated body structures; furthermore integrating joint motors for stabilize the simulation.

This approach will provide a range of human dance expressions for "humanized" dance movements. Every dance primitive shall be stored in a style-organized dataset, for being previously meaningfully selected during the dance embodiment stage.

3.3 Dance Embodiment

In order to infer some musical and rhythmic meaning to the performed dance, the succession of dance poses must be aligned with the beats of any music, as exposed in Figure 2.

This alignment shall be achieved through applying our beat tracking algorithm (see *3.3.1*) that can predict beat patterns along the music. To generate the transitions from one key pose to another an interpolation function can be used, which may consider the robot's corporeal constraints and the respective sensorimotor topology (in *3.3.2*).

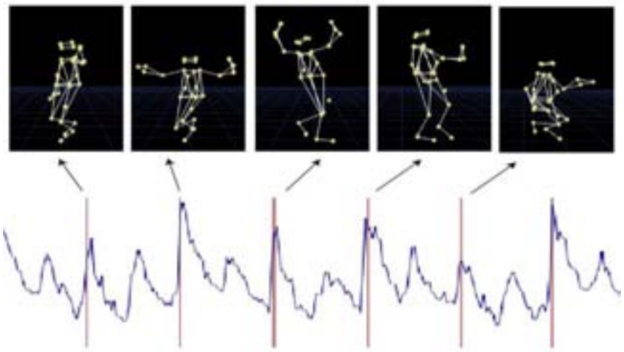


Figure 2. Synchronizing dance key poses to beats in the music.

3.3.1 IBT – Real-Time Beat Tracker

IBT -standing for INESC Beat Tracker- is an algorithm implemented in the MARSYAS platform [29] that predicts beats from music audio signals. In the line of the BeatRoot system [30], and differently from many existing algorithms, IBT does not track beats in a repeated induction loop, but rather implements a process of reconciliation between concurrent predictions and the actual data. Yet, distinctively from BeatRoot, the algorithm works in a causal fashion, processing sequentially incoming values of an (almost-) continuous data representation featured by the signal's spectral flux. After an initial pulse induction step, which computes the auto-correlation (ACF) of the first few seconds of data, a series of agents are generated, representing different hypotheses regarding beat positions and tempo. These hypotheses are then used as predictions propagated onto incoming data. Around each of its beat predictions, each agent is evaluated by a central agency (the "referee") with respect to the goodness-of-fit between predictions and local maxima in the observed data. In order to cater for possible variations in tempo and/or timing, agents' tempo and beat hypotheses are adjusted around each prediction. New agents are created on-the-fly, and others are killed, when large discrepancies are observed between existing agents' predictions and local maxima in the data. The referee keeps a running evaluation of all agents and makes decisions regarding which is the best agent at each data frame, whose beat prediction is validated.

3.3.2 Sensorimotor Synchronization

Ultimately we must threat the rhythmic coordination between perception and action while embodying human-like movements retrieved from motion capture data, and aligned to the perceived beat onsets. To approach this synchronism and embodiment we must deal with several psychological and mechanical issues such as intention, motor rate limits, negative mean asynchrony, variability, perturbation, among others [31]; as well as considering the reciprocal and dynamical coupling among brain (control), body, and environment, centred on the notions of self-organization and embodiment [32].

This task may be decomposed into two subsequent stages: dance key-pose synthesis aligned with the musical rhythmic metrical structure; and the transition between dance primitives along the performance. Concerning these issues an event aligner technique must be applied for mapping the dance primitives with the musical beats, while accounting for the robot's degrees-of-freedom (DoFs) and its motors' rate limits; and an interpolation function allied with (inverse) kinematics must treat the transition between motion keyposes.

Following the referred literature, some combined techniques have already been proposed to contemplate both issues. In order to overcome the robot joint limitations while keeping the motion rhythm, Shiratori *et. al* [7] used a temporal scaling technique for attenuating high frequency components of captured human motion components, as the musical tempo become faster; while preserving the extracted keyposes. This method consisted on the use of a hierarchical B-spline interpolation for control frequency resolution, by only setting control points at desired temporal intervals, denoted by musical tempo. This process also preserved the continuity of motion acceleration by iteratively applying constraints for optimizing the inter-joint kinematics and the robot balance.

Other techniques, more focused on legged motion and balance issues, used the captured motion data to specify a trajectory for the zero moment point (ZMP) in the control system rather than using it explicitly as desired joint angles [5], [6]. This method modified the original motion in a way that the upper body is horizontally translated so that a pseudo ZMP followed the desired one. Since translating the upper body is an approximation of translating the whole body, applying this method was iterated until translation converged. For interpolating the joints within each posture the authors additionally applied inverse kinematics techniques.

Alternatively, for keeping the sensorimotor synchronization, Yoshii *et. al* used a feedback step control method for adjusting the robot's step intervals tuned with the beat times occurrence, while accounting for musical tempo deviations [17].

Concerned with keeping the dance kinematic continuity Ellenberg *et. al* considered the full dance as a state machine transiting from a sequence of motion states [23]. This transitions consisted on the point-to-point interpolation between successive poses with recur to intermediate states to maintain stability, such as switching supporting legs, or walking forward. In order to keep synchrony with the musical beats, their beat tracker was optimized to predict beat times in the needed anticipatory fashion to generate each key-pose, from a pose database, on time.

4. CONCLUSIONS AND FUTURE WORK

This paper describes our research focused on multidisciplinary aspects of interactive sound and music production/perception, with a special focus on a robotics-based embodiment methodology. This work specifically consists on a model for capturing individual dance gestures, in real human dance performances, and the posterior mapping of gathered motion capture data in humanoid robots, for performing realistic human-like dance movements in synchrony to musical rhythm.

The further use of this approach shall "humanize" robot movements, which at a higher level would promote social human-robot sonic interactions, through dance, considering the interaction as an entrainment process based on rhythmic synchrony, imitation and patterns. For such, we shall research and ultimately design an interactive and flexible multi-agent framework through which a group of (heterogeneous) humanoid robots may collaborate by embodying human-like dance movements, in synchrony to a cross-modal perception of rhythm, depicted from auditory and visual stimuli.

5. REFERENCES

[1] Michalowski, M. P., Marek P., H., Kozima and H., Sabanovic 2007. *A Dancing Robot for Rhythmic Social Interaction*. In proceedings of the 2nd Annual Conference on Human-Robot Interaction (HRI), Washington, DC, March 2007, pp. 89–96.
[2] R. Pfeifer and F. Iida, 2006. *Morphological Computation for Adaptive Behavior and Cognition*. International Congress Series, 1291: pp.22-29.
[3] Kendon, A. 1970. *Movement Coordination in Social Interaction: Some Examples Described*. Acta Psychologica, 32:100–125.
[4] Aucouturier, J.-J. *Cheek to Chip: Dancing Robots and AI's Future*. IEEE Intelligent Systems, vol. 23, no. 2, pp. 74-84, Mar./Apr. 2008.
[5] Nakaoka, S., Nakazawa, A., Kanehiro, F., Kaneko, K., Morisawa, M., Ikeuchi, K. *Learning from Observation Paradigm: Leg Task Models for Enabling a Biped Humanoid Robot to Imitate Human Dance*. Int'l J. Robotics Research, vol. 26, no. 8, pp. 829–844, 2007.
[6] Nakaoka S. *Generating Whole Body Motions for a Biped Humanoid Robot*. Master Thesis of Information Science and Technology in Computer Science, at Graduate School of Information Science and Technology, University of Tokyo, 2003.
[7] Shiratori, T., Kudoh, S., Nakaoka, S., Ikeuchi, K. 2007. *Temporal Scaling of Upper Body Motion for Sound Feedback System of a Dancing Humanoid Robot*. In proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), October 2007.
[8] Shiratori, T., Nakazawa, A., Ikeuchi, K. 2006. *Dancing-to-Music Character Animation*. In proceedings of Eurographics, September 2006.
[9] Shiratori, T., Ikeuchi, K. 2008. *Synthesis of Dance Performance Based on Analyses of Human Motion and Music*. IPSJ Transactions on Computer Vision and Image Media, vol. 1, no.1 (also IPSJ Online Transactions, vol.1, no. 0), pp. 34-47, June 2008.
[10] Penasse, V. Nakamura, Y. 2003. *Dance Motion Synthesis with Music Synchronization*. In Robotics Society of Japan 2003, v. 21.
[11] Reidsma, D., van Welbergen, H., Poppe, R., Bos, P., Nijholt, A. 2006. *Towards Bi-directional Dancing Interaction*. In proceedings of the International Conference on Entertainment Computing (ICEC 2006), Cambridge, UK, LNCS 4161, Springer, pp. 1–12.
[12] Tanaka, F., Fortenberry, B., Aisaka, K., Movellan, J. 2005. *Plans for Developing Real-time Dance Interaction Between QRIO and Toddlers in a Classroom Environment*. In proc. of 4th IEEE International Conference on Development and Learning (ICDL), Osaka, Japan, pp. 142-147.
[13] Tanaka, F., Suzuki, H. 2004. *Dance Interaction with QRIO: A Case Study for Non-boring Interaction by using an Entertainment Ensemble Model*. In proceedings of the 2004 IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN), Kurashiki, Japan, pp. 419-424.
[14] Aucouturier, J.-J. and Ogai, Y. 2007. *Making a Robot Dance to Music Using Chaotic Itinerancy in a Network of FitzHugh-Nagumo Neurons*. In proceedings of the 14th International Conference on Neural Information Processing (ICONIP), Kitakyushu, Japan.
[15] Takeda, R., Nakadai, K., Komatani, K., Ogata, T., Okuno, H. G. 2007. *Robot Audition with Adaptive Filter Based on Independent Component Analysis*. In proceedings of the 25th Annual Conference of the Robotics Society of Japan (in Japanese), pp. 1N16.
[16] Murata, K., Nakadai, K., Takeda, R., Okuno, H. G., Torii, T., Hasegawa, Y., Tsujino, H. 2008. *A Beat-Tracking Robot for Human-Robot Interaction and Its Evaluation*. In proceedings of IEEE-RAS Int'l Conf. on Humanoid Robots (Humanoids), December 2008.
[17] Yoshii, K., Nakadai, K., Torii, T., Hasegawa, Y., Tsujino, H., Komatani, K., Ogata, T., Okuno, H. G. 2007. *A Biped Robot that Keeps Steps in Time with Musical Beats While Listening to Music with Its Own Ears*. IROS-2007, IEEE/RSJ, pp. 1743–1750.
[18] Mizumoto, T., Takeda, R., Yoshii, K., Komatani, K., Ogata, T., Okuno, H. G. 2008. *A Robot Listens to Music and Counts Its Beats Aloud by Separating Music from Counting Voice*. In proceedings of IROS'2008, pp. 1538-1543.
[19] Murata, K., Nakadai, K., Yoshii, K., Takeda, R., Torii, T., Okuno, H. G., Hasegawa, Y., Tsujino, H. 2008. *A Robot Singer with Music Recognition Based on Real-Time Beat Tracking*. In proceedings of 9th Int'l Conf. on Musical Information Retrieval (ISMIR 2008), pp. 199-204.
[20] Takeda, T. Hirata, Y., Kosuge, K. 2007. *Dance Step Estimation Method Based on HMM for Dance Partner Robot*. IEEE Trans. Industrial Electronics, vol. 54, no. 2, pp. 699–706.
[21] Michalowski, M. P., Marek, P., Kozima, H. 2007. *Methodological Issues in Facilitating Rhythmic Play with Robots*. In proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Jeju, Korea.
[22] Michalowski, M. P., Sabanovic, S., Michel, P. 2006. *Roillo: Creating a Social Robot for Playrooms*. In proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), pp. 587–592, University of Hertfordshire, UK, September 2006.
[23] Ellenberg, R., Grunberg, D., Oh, P., Kim, Y. 2008. *Exploring Creativity through Humanoids and Dance*. In proceedings of the 5th International Conference on Ubiquitous Robotics and Ambient Intelligence (URAI), November 2008.
[24] Klapuri, A. P., Eronen, A. J., Astola, J. T. 2006. Analysis of the Meter of Acoustic Musical Signals. Audio, Speech and Language Processing, IEEE Transactions on [see also Speech and Audio Processing, IEEE Transactions on], vol. 14, no. 1, pp. 342–355.
[25] J. Oliveira, F. Gouyon, L. P. Reis, 2008. *Towards an Interactive Framework for Robot Dancing Applications*. In Proceedings of ARTECH, 4th International Conference on Digital Arts, pp. 52-59, 7, 8 November, Porto, Portugal.
[26] *NaturalPoint's OptiTrack*: http://www.naturalpoint.com/optitrack.
[27] Obst, O., Rollmann, M. 2005. *SPARK - A Generic Simulator for Physical Multiagent Simulations*. Computer Systems Science and Engineering 20(5). SimSpark Available online at: http://simspark.sourceforge.net.

[28] Smith, R. *ODE - Open Dynamics Engine*.
[29] Tzanetakis, G., Cook, P. 2000 *MARSYAS: A Framework for Audio Analysis*. Organized Sound 4(3). MARSYAS available at: http://marsyas.sness.net.
[30] Dixon, S. 2001. *Automatic Extraction of Tempo and Beat from Expressive Performances*. Journal of New Music Research 30, pp. 39–58.
[31] Repp, B. H. 2005. *Sensorimotor Synchronization: A Review of the Tapping Literature*. In Psychonomic Bulletin & Review, 12, pp. 969–992.
[32] Pfeifer, R., Lungarella, M., & Iida, F. 2007. *Self-Organization, Embodiment, and Biologically Inspired Robotics*. Science, 318(5853), 1088.

Sound Amplification Technology and the Acousmatic Experience

Jos Mulder

University of Technology Sydney, Faculty of Arts and Social Sciences

Room 16.22 P.O. Box 123 Broadway

NSW 2007 Australia

+61 295142019

Johannes.Mulder@uts.edu.au

ABSTRACT

This paper aims to link the use of sound amplification technology to the Acousmatic theory as conceived by Pierre Schaeffer. The argument pursued here is part of my multi-disciplinary PhD thesis that researches the influence of the use amplification technology on music from Semiotic, Sociology of Music and Aesthetic perspectives.

KEYWORDS

Music Technology, Amplification, Acousmatics, Aesthetics.

1. INTRODUCTION

There are many publications that deal with the mediatisation of music. The research discourse is mainly about mediatisation through reproduction, broadcasting, and the possibility of creating and recreating sounds (for instance as described by Pierre Schaeffer). After reviewing this literature I have identified a gap in this discourse: technology applied to make musical instruments and voices louder, for instance to make them available to a larger audience at a concert, seems to be taken for granted in many publications. The issue of a potential influence of amplification technology on musical practices is not addressed. However, nowadays it is very common to amplify music at concerts, making the question of its impact on that music a very interesting one. In my PhD research project I take a multi-disciplinary approach to study the influence of amplification technology on music. Apart from describing the well-researched relation of sound amplification to Acoustics I am relating this thesis to theories from Semiotics, Sociology of Music and Aesthetics. The aim is to develop a theoretical framework and meta-language that will benefit both the praxis of music and the sciences that study music and its culture.

2. ACOUSMATIC THEORY

Pierre Schaeffer's theorem of *Acousmatics*, or *Reduced Listening*, very accurately describes the influence that the developing technologies of capturing and reproducing sounds had and have had on music (and other aural arts). Influenced by Husserl's phenomenology Schaeffer proposes that listening reduces a signal from its source. This reduction allows us to perceive sound as a phenomenon in itself. The possibility of detaching a sound from its source by technological means has made this phenomenological experience apparent. In his magnum opus *Traité des objects musicaux* [10] Schaeffer tried to develop a language for describing sounds *in themselves* [3] based upon this reduced listening.

2.1.1 See no evil hear no evil

Schaeffer borrows the term Acousmatic from Pythagoras' students or *akousmatikoi* who were listening to their teacher from behind a screen, to prevent visual distraction from his words[1]. According to Schaeffer, a sound becomes detached from its source when we record it. For instance when walking down a street we come across a barking dog, the source of that sound is evident. But when we record that sound with a microphone and some sound recording device, we are able to separate that sound from its source. We can now play back that dog's bark whenever and wherever we want. Possibly the person who first recorded those barks has a recollection of which dog where was the source of that sound, but most people confronted with that played back bark have no idea. The bark is now, in Schaeffer's terms a *concrete* sound. We can use it as a ring-tone, a sample in a hip-hop track or as an element in a composition with other concrete sounds, which we he have come to know as "Musique Concrète". The recording media in Schaeffer's days (first on discs, later magnetic tape) allowed for a materialization of the recorded sound into a solid concrete object. Materialized in a much more immediate way than recorded sound nowadays, using digital technology, available for direct manipulation. The term concrete also suggest the idea of working directly or concretely with sound, rather than in traditional composing where a system of abstract notation represents the sounds to be made concrete [7]. This may not sound like a big deal, but lets not forget, Schaeffer drew up this theory in the 1940's, when the possibilities of generating sounds and using them for musical purposes where limited[2]. His discovery of the Acousmatic nature of sound came after working with "closed grooves" on a vinyl record. After a modification of the groove on a disk it repeats the same section

[1] According to Chion [3] the writer Jérôme Peignot called the word Acousmatic to his attention. Historical sources (again from Chion) include a poem *Acousmate* by Guillaume Appollinaire and the book *Stromateis* by Clement of Alexandria ca. 250 BC.

[2] Limited to early electronic instruments like the Hammond Organ (1935), the Trautonium (1928) and the Theremin (1920) see for instance Chabade [2] and Bongers [1].

over and over again. Later when magnetic tape became available this was easier to realize by using tape loops.

2.1.2 Wagner and cats

Before the technical detachment of sound became possible Wagner tried to get rid of the visual aspects of the source of the sound: nothing should distract the spectator of the action on stage. In the Bayreuther *Festspielhaus*, built after Wagner's specifications (and still home to many performances of Wagner opera's), the orchestra pit is largely hidden or "sunk" under the stage and fenced of by a screen[3]. A century later a theatre orchestra was banned from pit or theatre all together: in the musical Cats the orchestra (or band) plays in a different room (sometimes even in a different building) and the sound is relayed electronically, a video connection allowing the conductor to see the stage and vice versa for the performers on stage. For performances of Stockhausen electronic music such as *Gesang der Jünglinge,* the composer suggests to dim all the lights with the exception of a little bright moon projected on the proscenium. If you are uncomfortable listening with your eyes closed you can look at the "moon", ensuring very little visual distraction.

2.1.3 Listening without seeing: music

Using Schaeffer's Acousmatic theory in a context of amplified music seems a strange idea, as in most cases at a concert the visual connection from the audience with the performers is not hindered by anything. On the contrary, in concert halls the lay-out and lighting are approached in such a way that we may always see what is happening on stage. In the sense of Schaeffer's descriptive account using his theory for amplified sound is not an option. However there is a more prescriptive approach to the Acousmatic experience: in his *The Aesthetics of Music* Roger Scruton [11] makes the Acousmatic experience crucial for perceiving music as music. For Scruton, both thought and awareness of the cause or the source are excluded when we listen to music. In other words: music becomes detached from its source when we listen musically. This also suggests that for listening to music as music it makes no difference whether we listen to a symphony in a concert hall, or to that same symphony on the radio or played back from a record. In Scruton's account we hear music as Acousmatic, and everything we hear non-Acousmatic as "just" sound. Scruton embarks on his argumentation by referring to Schaeffer:

> "In listening […] we spontaneously detach the sound from the circumstances of its production and attend to it as it is in itself: this, the 'acousmatic' experience of sound, is fortified by recording and broadcasting, which completes the severance of sound from is cause that has already begun in the concert hall."

For the scope of this paper I am interested in that beginning of "the severance", where one or more microphones pick up the sound (of instruments). In the case of amplification this sound is relayed instantaneously through loudspeakers in the same room.

2.1.4 Listening without seeing: film

Michel Chion, in the context of film extends the idea of Acousmatics also to voice and sound effects [3]. In this paper I will restrict myself to music and explore the potential of the term "Acousmatic" for understanding the way in which amplification detaches a musical sound from its source.

[3] As a consequence the sound of the orchestra is dampened, avoiding overpowering of the singers by the orchestra.

3. Acousmatics and amplification

As we have seen above, Schaeffer's theory does not cover what happens if we capture sound (with a microphone) and instead of storing or broadcasting it on the radio, directly relay it through a loudspeaker in the same room. In other words reproducing the sound instantaneously in the same room and so amplifying that sound. This seems to be a special case of broadcasting, where both the sound source and its immediate reproduction are audible in the same space at roughly the same time. To make this more apparent, consider the following ways of listening to a violinist A playing composer B's first Suite in recital hall C, with an imaginary curtain to fulfil the non-visual requirement:

1. We could sit in the hall and listen to the violinist's performance.
2. We could sit at home and listen through a loudspeaker to a direct radio broadcast of that performance captured with a microphone close to the instrument[4].
3. We could sit at home and listen, through the same loudspeaker, to a recording made with the same close microphone of the same performer, of a same performance of that work in the same hall.
4. We could sit in the hall and listen to that same recording, through the same loudspeaker, set up in the hall for that purpose.
5. We could sit in that hall and listen to the same violinist playing the same work picked up by the same microphone, and relayed to that loudspeaker in the hall so that this amplified sound is perceived equally loud as the acoustic source, the violin.

There are many more possibilities to add to these examples, such as listening to a recording or a radio broadcast of it on our mobile phone whilst sitting in the hall, but for the scope of this paper these examples suffice. Using the combination concert hall and "classical" music is by way of example; I don't intend to single out any sort of music. The effects described here apply to any sound with a mechanical source, that is picked up by a microphone. An instrument, a human voice or even a loudspeaker as the sound of an electric guitar is often "picked-up" by a microphone pointed at the guitar's amplifier's speaker.

3.1.1 Acousmatic implications

There are a number of factors in the five different listening situations that are relevant to a discussion of the Acousmatic theory: our listening location (hall C or at home), the use of a (close) microphone for recording, broadcast and amplification; and the loudspeaker used for listening to these mediations. Apart from the musical parameters that differ every time violinist A plays this particular piece by B in hall C there is another obvious difference between a recording played back in the hall, and the performance in person: the absence of the player and her or his instrument and, as a consequence, of the acoustic source.

[4] The microphone and loudspeaker and paraphernalia are of a theoretic ideal type that doesn't introduce any colouring or phase distortion. For argument sake I will assume that said machinery is analogue, as the use of digital devices introduce a latency that will unnecessarily confuse this discussion.

3.1.2 Recording introduces an infinite temporal displacement

The examples where a recording of the performance is being used differ from the other examples in that the sound from the instrument is picked up and stored before played back. This introduces a time factor indefinitely larger than the amount of time we usually associate with the physics of sound in a hall, or a direct broadcast via the radio or a direct relay through a loudspeaker.

3.1.3 Amplification introduces a finite temporal displacement

Only in the first example is A's violin the source of the sound, transmitted through hall C where we sit and listen to it (for Scruton's aesthetic theory this would be an Acousmatic Experience too). In situations 2, 3 & 4 we are listening Acousmatic in the sense that Schaeffer described: the sound of the violin is detached from the body that produces it and of the room it was initially transmitted in. When a sound is amplified, as in the last example, not only do we hear the acoustic source (the violin), at almost the same time we hear the sound that is picked up by the microphone and relayed through the speaker. There is a time factor of another magnitude here: the electronic signal travels much faster (electrons travel at nearly the speed of light) through the cables to the loudspeaker as the sound waves emitted from the instruments travel trough the air (at the speed of sound)[5]. Depending on the relative distances of the listener to the instrument and the loudspeaker, the reproduced sound from the latter can be perceived even before the sound of the original source. The law of the first wave front describes the physical and psycho-acoustic workings of this phenomenon. This law, also known as the precedence effect or Haas effect, will be treated in section 4.

3.1.4 Electronic source causes spatial displacement

Amplifying an acoustic source, such as a violin, creates a new electronic source, the loudspeaker. We can hear the electronic source at the same time, after or before we perceive the sound from the acoustic instrument. This has as a consequence that it alters the spatial parameters of how we hear the acoustic source in the concert hall. The electronic source creates its own reflections, in a different relation to the room's acoustic than those caused by the acoustic source. Although coming from distinctly different sources, we may or may not be able to discern between the reverb related to the acoustic or the electronic source, depending on the relative distance between us, and the two sources. This enhances the spatial displacement of the violin caused by the finite displacement in time.

3.1.5 Level of Amplification

Apart from the relative distance between the listener and the two sources there is another important parameter in this discussion: the level of amplification. In the last example I proposed a level of amplification that makes the electronic source perceived just as loud as the acoustic. The situation changes dramatically if we amplify the violin at a barely audible level, or, if we amplify the violin to such extremes that the acoustic source becomes inaudible. The level of amplification and its perceived impact, in relation to the acoustic source form a continuum, which I would like to describe as an "Acousmatic Continuum". The way a theory of amplification subscribes to Schaeffer's Acousmatic theory changes accordingly. As long as we hear the violin as an acoustic source we do not perceive the violins sound as "detached" (in Schaeffer's vocabulary). But when the acoustic source is overshadowed by the sound from the loudspeaker, making it harder to hear whether we are listening to violinist A or a recording of his performance, we could consider it as an Acousmatic experience.

4. AMPLIFIED SOUND AND TIME COMPENSATION

The psychoacoustic effect that comprises of several different auditory phenomena is known as the Haas effect (after Helmut Haas [6]). Differences in phase, and level between perceived sounds are, within certain limits, decisive as where we locate a sounds source. Practically this means that, again within certain limits, if we perceive the direct "acoustic" sound of violin A before we perceive the amplified sound of violin A coming from a loudspeaker in the concert hall, we hear violin A as the location of the sound, not the loudspeaker. The critical limits are circa 30 milliseconds in time (as a reference, we start perceiving sounds as an echo when the difference is larger than about 50 milliseconds) and circa 10 dB (SPL) in level. The problems related to the temporal displacement of sound caused by the difference in speed (of sound in air and electrons in a conductor) where acknowledged as early as 1925, in a patent filed by inventor B.S. McCutchen (US patent nr. 1642040) who proposed a delay mechanism (that didn't work however). The first practical electronic delay systems became available with the tape-echo in the late 1960s, which has occasionally been used to compensate timing differences in amplification systems. With the advent of digital (sound) technology since the 1980s it has become much easier and nowadays it is common to use digital delay devices to "time align" acoustic source(s) and loudspeakers. However advanced this may sound, the time alignment is optimized for only one place in an auditorium, as in every seat the relative distance to loudspeaker and acoustic source is slightly different. It is important to consider that this time alignment allows for compensation of the difference in the speed of sound in air (from the acoustic source) and the speed of sound in the electronic domain. It does not account for time differences between aural and visual perception.

4.1.1 Time compensation and video

At performances where the action on stage is both amplified and enhanced by video screens or projections another temporal displacement becomes apparent. The light of the images travels to us at the speed at light, whereas the sound (from the instruments or the loudspeakers) travels to us at the speed of sound. There is experimental evidence that we are unable to detect difference in arrival time between sound and light inferior to 40–50 ms [8]. This means that audiovisual synchronicity is not problematic when the distance is no larger then roughly 20 meters [9]. However larger auditoriums, or in a sports stadium the visuals may not be synchronous with the sound. To compensate for this the video needs to be delayed, relative to the audio. Again this can only be optimized for one area in the auditorium. In this we also find the reason why the difference in aural and visual perception of the same event on stage (the violinist exciting the strings of his instrument with a bow) has no direct relevance to a discussion of the Acousmatic experience.

5. AN ACOUSMATIC CONTINUUM?

In Scruton's words, the *severance* of sounds begins in the concert hall. This suggests a process rather than a singular

[5] The speed of light is roughly one million times higher as the speed of sound in air.

event. As I have written above, within the realm of amplification, the temporal displacement is depending on the distance from the listener to the acoustic source and the loudspeaker. And the spatial displacement is depending of the temporal displacement and the level of amplification. If we fix the position of the listener in the middle of the hall, with ideal situation of an equal distance to acoustic source and loudspeaker and regard the difference in time between the sources as a constant, the level of amplification is the remaining controllable factor in the acousmatic process. This will allow us to plot out the spatial displacement as a function of the level of amplification. With low-level amplification the acoustic source can still be perceived, the amplification "helping" a little bit in making the acoustic source appear louder. With a high level of amplification the sound of the acoustic source will no longer be perceivable as such as the it is "overpowered" by the loudspeakers.

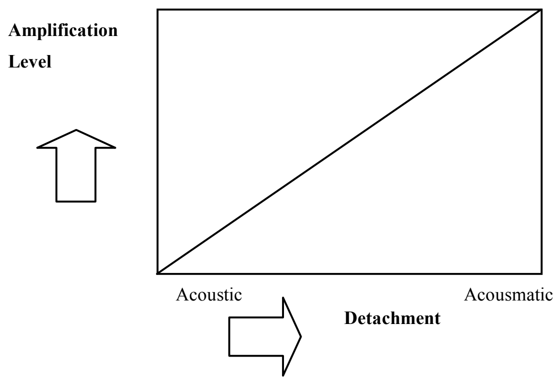


Figure 1 Amplification level vs. detachment

There is a risk of oversimplifying the theory by contrasting the Acousmatic with the Acoustic, as this diagram suggests. Michel Chion, who studied and worked with Schaeffer, wrote a guide to accompany Schaeffer's *Treaté des objects musicaux*, which was translated very recently [4]. In the first section that deals with the Acousmatic he points out:

> "We must take care not to misinterpret the acousmatic situation, for example by making a distinction between the "objective" – what is behind the curtain – and the "subjective" – "*the listener's reaction to these stimuli*" in an over–scientific simplification of the phenomenon. On the contrary "*the acousmatic involves a reversal of the normal course of events (…) it is no longer a question of knowing how a subjective listening interprets or distorts 'reality' or of studying reactions to stimuli; the listening itself becomes the origin of the phenomenon to be studied. (…) The question: 'What am I hearing?… What precisely are you hearing?" is turned back on to the subject, in the sense that he is being asked to describe, not the external references of the sound he perceives, but his perception itself.*" (Italics from translation)

Once more this directs us to the problem of applying the Acousmatic theory to amplified sound, as the visual connection remains intact. Being able to see the (acoustic) source of the sound will make it problematic to listen to the sound as a phenomenon on its own and hence not strictly acousmatic. The above diagram suggests that there is a continuous relationship between the level of amplification and whether we perceive the detached sound as Acousmatic. Such a relation, an Acousmatic Continuum would be a useful paradigm for a theoretical discussion of amplified sound. However, with the visual connection intact, the auditory experience can only be Acousmatic in a Schaefferian sense when the amplification is so loud that we can't hear the acoustic source anymore. But for now I will use the term Acousmatic Continuum to indicate the relation between the detachment of sound and the level of amplification.

5.1.1 Acousmatic Continuum and music

In the complexity of the multi-modal experience that for instance a concert of music with amplification is, the notion of a Continuum may provide us with a framework that helps us understand the relation between amplified sound and its acoustic source. In a practical way I find the following diagram helpful, depicting different sort of music as a function of the amount of amplification that is used.

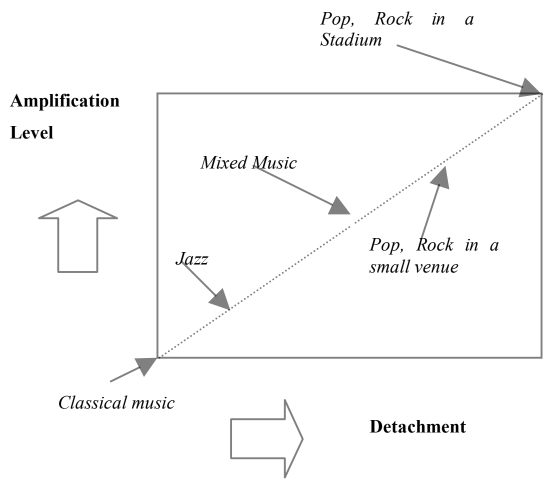


Figure 2 Amplification level and amplified music

Under the lemma Classical music we can find what is usually un-amplified music, for instance a performance of a string quartet in a concert hall. Jazz usually has some amplification, for instance the double bass and guitar are usually amplified and most jazz-vocalists use a microphone.

5.1.2 Mixed music

Somewhere in the middle of the diagram we could place what is often called "Mixed Music". This is usually defined as a work for orchestra and pre-recorded "Tape" (and hence Acousmatic) material. Often with those kinds of performances the orchestra or acoustic instruments are amplified a little bit to allow for better blending of the acoustic and Acousmatic or sounds. Examples are Stockhausens *Kontakte*, or *Hymnen* (Region III, mit Orchester), Reich's *Different Trains* or Nono's *La Fabbrica Illuminata*. Simon Emmerson [5] suggests that this kind of combined acoustic/acousmatic music could possibly disappear as:

> "There is a feeling amongst many performers that the medium itself is inappropriate for the flexibility demanded by different acoustics, venues, audiences and other musical variables (p108)"[6].

[6] Apart from these inflexibilities, Emmerson points out, there is the problem of synchronisation, (unless synchronised timing between the tape and performer a not essential, or even undesired). For instance, in Stockhausens Hymnen with Orchestra (3rd region) the conductor has to know the

Inappropriate or not, the practise of “Mixed Music” shows us the complexity of the relation between room acoustics, acoustic (generated on an instrument or with a human voice) and electronic (re-generated by a loudspeaker) sounds.

5.1.3 *Amplified music*

The next step up the acousmatic scale is Pop or Rock music in a small venue, or a bar. Most instruments are perceived coming through the loudspeakers but often louder instruments like horns or drums are still acoustically louder than the amplification. At a stadium concert, where everything we hear of the performance on stage, comes from loudspeakers we can consider what we heard as an Acousmatic experience–under “live” conditions. In a way we are listening to a recording (mixed in real time) that is played back instantaneously. And sometimes it is hard to judge whether what we see and hear is really performed live or is it all “playback” with musicians miming to pre-recorded material.

5.1.4 *Different functions of amplification*

Emmerson is one of the first authors to systematically look into the amplification of sound in a musical context. In his book [5] *Living Electronic Music* he identifies six functions of amplification that he considers vital to discussing “live music requiring electronic technology for its presentation”, in other words amplified music. These are (in my words paraphrasing the author):

Table 1 Six Functions of Amplification

1	Balance	Correcting Acoustic imbalances
2	Blend	Integrating timbres
3	Projection/ Spatialisation	Zooming in perceptually/intentional (ac ousmatic) dislocation.
4	Perspective	Suggesting different acoustics (dislocation of space and time)
5	Coloration	Intentional distortion
6	Resonance	Use of feedback

For a discussion of the acousmatic we are mainly concerned with the first two functions or reasons of amplification, and the spatial dislocation (whether intentional or not) that goes with it.

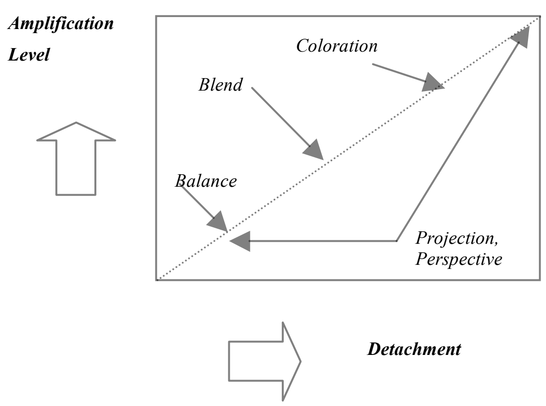


Figure 3 Amplification functions

In the above merger of the Continuum and Emmerson’s functions we can see that some of these functions are fully dependant of the level of amplification and some to a lesser extent. Projection (and Spatialisation) can (theoretically) be realized at any level and Perspective as soon as the added synthesized acoustic can be heard. Full spatialisation, to give a sound a wholly different location in a room (this goes beyond panning between a “left and right” loudspeaker, can be realized when we no longer hear the acoustic source and the electronic sound has become detached. Blend and Balance are more particularly defined in each situation (provided that agreement on an acceptable balance is possible). Coloration (here different from applying filters to the amplified sound) by intentionally distorting a sound electronically will naturally need a certain amount of amplification to come to the desired effect.

6. FUTURE RESEARCH

In the near future I hope to realize some experiments that will allow for empirical verification of the above. The experiments would be conducted with both speech and music while parameters could be the amplification level and the distance between the acoustic (instrument) and the electronic (loudspeaker) source. Participants will be registering when the localization shifts from the acoustic to the electronic source. Other parameters could include time compensation (to engage the Haas effect) and a “Pythagorian” screen, hiding the source from the audience to single out visual clues. Expected outcomes should include a critical distance between an amplified instrument and the loudspeaker that provides the amplification. In relation to the Continuum described above this experiment can stipulate two different phases of that Continuum, first dislocating the source and after that making the source inaudible due to a very high level of amplification. Along this line the experiments could be extended to position the six different functions of amplification as mentioned by Emmerson.

7. CONCLUSION

Amplification of acoustic sources such as musical instruments including the human voice can be realised by immediate and local electronic reproduction of that source. The (finite) temporal and spatial displacement caused by the instantaneous reproduction of the sound in the same room suggests a detachment of the sound and its source. This detachment works much in the way of the Acousmatic experience as described by Pierre Schaeffer. However being able to see the acoustic source makes this a problematic exercise for the Schaefferian requirement of “hearing without seeing”. The sound becomes detached from its acoustic source, depending on the level of amplification that is applied. Louder amplification means a sound becomes more detached from its acoustic source and its initial mode of transmission. Because of this interrelationship I suggest we should think of relation between the detachment of sound and the level of amplification as a continuum. Calling this continuum Acousmatic would be going too far as the visual connection with the source is not lost. Another term, equally evocative as Acousmatic, will have to be found to describe this unique and important relation. The Continuum can be used as a tool to identify and position different sorts of (amplified) music and sound amplification functions.

“Elektronische Musik” by heart to be able to distil cue’s, timing and tempo information, while conducting an orchestra.

8. REFERENCES

[1] Bongers, Bert. Electronic Musical Instruments: Experiences of a New Luthier. *Leonardo Music Journal*, vol. 17 (2007), pp 9-16

[2] Chabade, Joel. *Electric Sound: The Past and Promise of Electronic Music*. Prentice-Hall, New Jersey 1997.

[3] Chion, Michel. *The Voice in Cinema.* Columbia University Press, New York, 1999.

[4] Chion, Michel. *Guide To Sound Objects. Pierre Schaeffer and Musical Research.* Translated by John Dack and Christine North, downloaded via www.ears.dmu.ac.uk (visited Oct. 1, 2009), 1983/2009, pp12.

[5] Emmerson, Simon. *Living Electronic Music*. Ashgate Publishing Limited, Aldershot, 2007.

[6] Haas, Helmut. The Influence of a Single Echo on the Audibility of Speech. *J. Audio Eng. Soc.*, Vol. 20 (Mar. 1972), pp145-159.

[7] Hamilton, Andy. *Aesthetics & Music*. Continuum, London, 2007.

[8] I.J. Hirsh and C.E. Sherrick, Jr, Perceived order in different sense modalities. *J. Exp. Psychol*. 62 (1961), pp. 423–432.

[9] Lewald, J and Guski,R. Auditory-visual temporal integration as a function of distance: no compensation for sound-transmission time in human perception. *Neuroscience Letters*,Vol 357, Issue 2, 4 (March 2004), 119-122.

[10] Schaeffer, Pierre. *Treaté des Objets Musicaux*. Éditions du Seuil, Paris 1966.

[11] Scruton, Roger. *The Aesthetics Of Music*. Clarendon press, Oxford, 1997.

Learning in a Musical Child-Computer Interaction – An approach in Embodied Cognitive Science of Music

Julia Wewers
Department of Systematic Musicology, University of Cologne
Albertus-Magnus-Platz
50923 Köln
0049-221/5109581
julia-wewers@gmx.de

ABSTRACT

In this paper I will argue for the relevance of measuring of musical child-computer interactions for analysis of schema-based musical learning and developmental stages of music cognition within the cognitive music research paradigm.

Keywords

Embodied cognitive science of music, human-computer interaction, new media art, development, schema theory.

1. INTRODUCTION

Due to a fast technological development of interfaces and computer programs in the last years, new musical instruments and interactive sound installations were created giving innovative possibilities of sound production. This claims also for new perspectives in interdisciplinary music research that could include these techniques. In this paper I will develop an approach combining theories of embodied cognitive science with findings of musical development of children, using methods from current research in Human-Computer and Human-Robot Interaction.

2. THEORETICAL BACKGROUND

2.1 Developmental psychology and schema theory

In developmental psychology the interaction of a subject with its environment is the basis of constructive and schema-theoretic models of learning. According to Jean Piaget [12] knowledge about the world is acquired via accommodation and assimilation, therefore through adaptation within the near environment, building up cognitive schemata. First based on reflexes and spontaneous movements, a child constitutes a schema-network that becomes more and more differentiated over the years. Also logical-mathematical functions were acquired through the corporal interaction with objects and through cognitive actions that are defined as "self-regulation". This means that the subject is a self-organizing system that constitutes its knowledge of the world in an active way [9], but under environmental effects.

One example of schema-based learning is presented by Michael Arbib who already gave attention to the acquisition of language using models of schemata within Artificial Intelligence research [1]. He also focuses on the steady action-perception cycle within a subject-environment interaction that underlies cognitive language acquisition as a constructive process. The development of musical knowledge and behavior that are not completely innate, as studies demonstrate[1], could also be described by the construction of different types of schemata, like motor schemata [2][13][2] and image schemata [5].

Some theories of developmental psychology expect learning due to phases and steps of life. Piaget's research claims for a theory of a child's cognitive stages that are rooted in the equilibration between self-organization and the perceived world (through assimilation and accommodation) and that is defined by different steps of complexity of their logical-mathematical thinking (Fig.1).

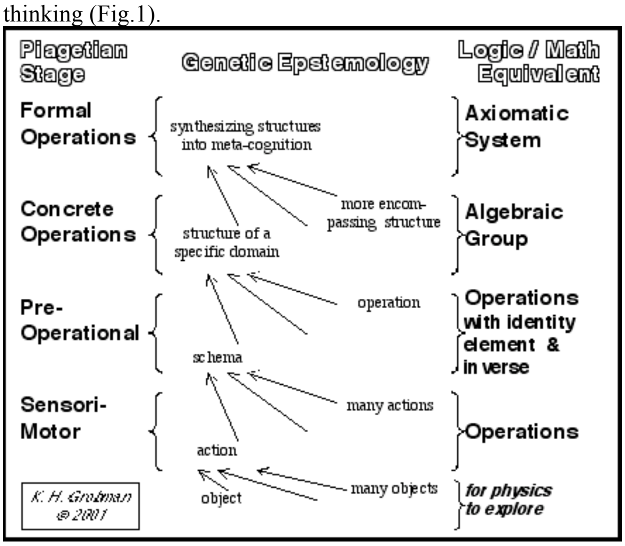


Figure 1. Logical-mathematical knowledge in the Piagetian stages and schema acquisition. [17]

In this context Laurel J. Trainor studied musical learning in early "critical periods" of childhood. She showed for example that "children younger than about 7 years do not yet represent

[1] e.g. [12]

[2] the term "motor schema" is used in different ways by J. Piaget and M. Arbib.

melodies in terms of their harmonic implications"[3]. This classification of (musical) capabilities into steps of life therefore seems to be a possible approach in research of musical development.

2.2 Embodied Cognitive Science of Music

Within the paradigm of embodied cognitive science of music cognitive music processing is understood as embedded in and influenced by a situation [7]. The situatedness of the body (e.g. the implicit knowledge of "Who am I?", "Where am I?", "How can I act on the actual surrounding circumstances?") has impact on the perception and production of music.

One feature of the theory of embodiment is a continuous action-perception cycle that couples body-movements with cognitive features (like "image-schemata") [5]. Action and perception influence each other at every time transcending a bodily border that in dualistic theories lies between the mind and the external world. Therefore corporal gestures, like playing an instrument or dancing, are significant outcomes of expressive music production that could give information about the underlying processes.

Concerning schema-based leaning, the interaction of the subject with its environment plays a central role in Piaget's theories for adaptation through assimilation and accommodation. The construction and organization of schemata and knowledge would not exist without bodily experiences of environmental objects. Especially the logical-mathematical thinking bases necessarily upon the cognitive abstraction of time and space relations of different events and things (e.g. object permanence). In the research on the construction of music specific schemata consequently a musical interaction has to be the subject of study, because cognitive development is always coupled with interactive behavior.

For experiments concerning embodied cognitive science of music a method has to be chosen that enables a fast action-perception sequence and gives the possibility to measure body movements at an adequate point within this cycle.

A short summary of research trends of cognitive psychology within HCI and HRI should demonstrate that instruments with tactile and auditory feedback as well as installations mapping body movement onto sound are very qualified to fulfill this need.

2.3 Human-Computer and Human-Robot Interaction

The current research fields of Human-Computer Interaction (HCI) and Human-Robot Interaction (HRI) concentrate on the usability of these computer-based systems by humans, on interface design as well as on physical and cognitive ergonomics. In the context of Music and New Media Art musical HCI/HRI (MHCI/MHRI) intends to observe these interactions and to determine their psychological causes within a musical environment. That offers new possibilities to establish and verify theories about musical learning, since alternative instruments and sound installations are supposed to be used by everybody immediately and intuitively.

HCI and HRI focus on the one hand on the usability of computer based systems concerning the speed and effectiveness of learning to solve tasks. On the other hand a subject of research is the impact of the system's design on the cognitive processes underlying the interaction. In MHCI and MHRI were described above all the possibility of creative and expressive usage of interactive systems with sound output and the underlying mental processes that are fundamental for artistic behavior [7].

Sound installations are as well based on the interaction cycle of body movements and especially auditory feedback. Under the terms of focusing on the music-specific motion-sound cycle the paper claims for the usage of sound installations as part of MHCI research for different reasons. It offers for example the possibility to measure the musical behavior of even very young children because they are still incapable to handle traditional or alternative musical instruments. Furthermore, regarding the embodiment paradigm a sound installation allows the implication of the whole body in the sense of being embedded within a situation.

In this context the role of sound installations raises the question whether a visible object like a robot or an alternative instrument has to be involved in the musical interaction that should be the subject of research. Concerning MHRI, research like for example on bodily rhythmic synchronization via observations of musical child-robot interactions has already been done [10]. In their studies Michalowski et al. use robots especially designed for children to achieve a playful and attractive musical interaction (e.g. the yellow dancing robot "Keepon" (Fig.2)). The authors focus on the development of human social interaction that is rooted in competencies like e.g. detecting and tracking human faces, recognizing the prosodic features of their mother's voice, etc. Michalowski et al. emphasize the role of rhythm in the aspect of joint attention that is a basis for human-human interaction.

In contrast to these more sociological studies the present paper drives attention on the specific musical mental processes that underlie the rhythmical movements and claims therefore for another focus of cognitive music research.

Figure 2. "Keepon", an interactive robot, used for therapy and research on rhythmic child-robot interaction. [10]

3. MAIN CONTRIBUTION

3.1 Two levels of the research subject

It is necessary to distinguish between two different levels of cognitive musical development. The first one focuses on the action-perception coupling in a music-generating situation and affords therefore an analysis of a short time span. Musical

[3] [14], 273.

bodily behavior is here supposed to be based on schemata and eventually on innate musical universals like rhythmic capabilities. How will the usage of an unknown musical "instrument" and the connected creative behavior be learned? What is the role of the body or gestures in this process? When will they explore that their movement has impact on their environment? In this context theories of embodied music cognition are formulated to explain musical learning.

The second level will concentrate on the ontogenetic research of musical development in childhood. Can musical behavior in a sound installation give information about mental musical capabilities in different ages? Is it possible to support a theory of cognitive stages comparing expressive motion with steps of logical-mathematical thinking?

I propose the measuring of children's behavior within a sound installation for the study of musical learning on the first mentioned level. Additionally the results of these experiments can be used for a comparative study between children of different ages for research on musical learning of the second level (developmental stages). One aim of the study will be to compare the data of intuitive expressive motion with findings of children's cognitive development according to stages and to discover differences between the musical competence of children of certain ages. Will children of diverse ages learn differently fast to manipulate sound with their body? And to what creative and expressive degree will they behave at which age?

3.2 Focus on the rhythmical aspect of musical interaction

Furthermore, the idea of rhythm as a basis for the construction of the first schemata can be found in theories of Jean Piaget: "[…] that the organism is never passive, but is spontaneously and globally active in a rhythmic way."[4] Here the question arises, whether mental organized structures that are also the basis for rhythmic body movements are innate and to what extent they could contribute to the (especially musical) development. So possibly underlying rhythmic abilities could be interesting for both, music research and developmental psychology. Future research will therefore focus on the organization of mental structures especially underlying rhythmical elements of behavior.

4. METHODOLOGY

The elaboration of an adequate experiment design seems to be the most challenging task in the upcoming process of my research, for it has to conform to stringent conditions, like e.g. the immediateness of sound mapped onto motion.

4.1 Proposed experiment setting and measurement

As empirical methods I propose the measurement of body movements of children in a sound installation, e.g. with camera-based motion tracking. A sound installation involves the need of a fast action-perception-coupling because of immediate auditory feedback to the own movements and allows sound production using gestures of the whole body to conform to principles of the embodiment paradigm. In addition digital music systems of new technologies are helpful because spontaneously created "musical artifacts" can be recorded and would then be available for further interpretation.

[4] [13] 17; translation by the author.

An experiment design using the technique of an installation would be more appropriate for the proposed study than interactions with traditional or alternative music instruments for it allows even toddlers to produce sound immediately via body movements. One should also note that experiments in HCI/HRI using new technologies qualify especially for children, due to their tendency to step up spontaneously and naturally to new systems.

The three-dimensional movement within the experiment-room should be recorded via cameras as well as the body acceleration. A software program, like e.g. the open software platform *EyesWeb* [3] (Fig. 3) maps changes of direction and speed of the recorded objects on multi-parametric sound that will be given back immediately via loudspeaker. This ensures a fast auditory feedback to the body movements. I consider it important, that the mapped sound does not consist of melodies or preprogrammed rhythmic patterns that change in pitch, loudness, etc., depending on the movements. The mapping should rather allow a wide range of simple sounds varying in their parameters to offer as much creativity in music production as possible. Additionally a transparent mapping of dimension and acceleration of movement to the sound parameters is necessary to make enable also reproductions of created patterns. Afterwards the recorded sound as an auditory artifact can be analyzed, as well as the resulting body-imaging given by the software program.

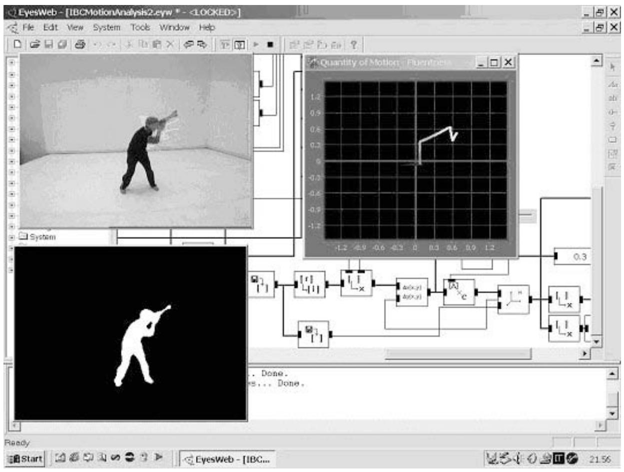


Figure 3. Motion analysis and mapping of gestures with the software program *EyesWeb.* [18]

Another important condition of the experiment will be that children act alone and without a task within the installation to focus on the most natural and spontaneous behavior in this musical subject-environment interaction. They should be able to concentrate on their gestures and the musical effects and to feel free in creative and expressive music production.

4.2 Observation

For the analysis of body movements in a child-computer interaction or within a sound installation also the method of observation could be considered. In contrast to other qualitative methods like interviews or questionnaires a "structured observation" enables the survey of qualitative data in a structured form very shortly after the motion. Regarding the inability of young children to describe their sensations verbally, it also gives the possibility to measure features underlying a performance, e.g. the degree of expression.

But it is necessary to emphasize, that structured observation can only be accepted within empirical scientific research, if it fulfills some conditions [11]: On the one hand it is inevitable to have an adequate system of categories, that assigns observed body movements clearly to classes. On the other hand observers have to be trained and calibrated so that they can be regarded as statistic reliable measuring instruments.

Given the fact hat structured observation as a qualitative method is not yet developed sufficiently within musical human-computer interaction [16], the method of the research proposed in this paper should be limited to survey and evaluation of digital data.

5. CONCLUSIONS

The aim of this paper was to give a new perspective within embodied cognitive science of music. Relating on different theories of developmental psychology and human-computer interaction a connection is proposed between mental music processing and musical motion.

Analyzing the behavior of children within a sound installation could possibly bring findings about the cognitive organization of musical capabilities, especially of rhythmic features. Theories of schema-based musical learning and acquisition of musical abilities according to developmental stages have yet to be proved in the future using alternative empirical methods like technologies of New Media Art, e.g. camera-based motion tracking.

6. IMPLICATIONS

Cognitive Musicology is here regarded as an interdisciplinary field of research that aims to explore the evolutionary and ontogenetic sources of musical behavior as well as the mental processes underlying an expressive and creative act of music production. With my paper I want to show a new perspective in the plurality of studies based on the embodied cognition paradigm.

Breaking these new grounds I hope to get findings that could be relevant in Human-Computer Interaction, Cognitive Science of Music and Developmental Psychology. Further research could focus on creativity and expressivity of children using sound installations and could therefore contribute to Aesthetics.

7. ACKNOWLEDGMENT

I would like to thank the Department of Musicology of the University of Cologne, and especially Son-Hwa Chang and Henrik Niemann for their motivating and critical help.

8. REFERENCES

[1] Arbib, M.A., Conklin, E.J., Hill, J.C. *From schema theory to language.* Oxford University Press, Oxford, 1987

[2] Arbib, M.A., Hesse, M.B. *The construction of reality.* Cambridge University Press, Cambridge, 1986.

[3] Camurri, A. et al. Multimodal analysis of expressive gesture in music and dance performances. In A. Camurri, G. Volpe (Eds.) *Gesture-based Communication in Human-Computer Interaction*, LNAI 2915, Springer Verlag, 2004, 20-39.

[4] Ghallager, S. *How the Body Shapes the Mind.* Oxford Univerity Press, New York, 2005.

[5] Johnson, M. *The body in the mind. The bodily basis of meaning, imagination, and reason.* University of Chicago Press, 1987.

[6] Kiefer, C., Collins, N., Fitzpatrick, G. HCI Methodology For Evaluating Musical Controllers: A Case Study. In *Proceedings of the New Interfaces for Musical Expression (NIME08)*, Genova, 2008.

[7] Kim, J.H. et al. Embodied Musical Human-Computer and Human-Robot Interaction. In: *Proceedings of 25th Tonmeistertagung*, Leipzig, 2009.

[8] Leman, M. *Embodied Music Cognition and Mediation Technology.* The MIT Press, Cambridge, Massachusetts, 2008.

[9] Maturana, H., Varela, F. *Autopoiesis and Cognition: the Realization of the Living.* D. Reidel, Boston, 1980.

[10] Michalowski, M. P., Sabanovic, S., and Kozima, H. A dancing robot for rhythmic social interaction. In *Proceedings of the ACM/IEEE international Conference on Human-Robot interaction (Arlington, Virginia, USA, March 10 - 12, 2007). HRI '07*. ACM, New York, NY, 2007, 89-96.

[11] Robson, C. *Real World Research - a resource for social scientists and practitioner-researchers. Second Edition.* Blackwell Publishing, Oxford, 2002, 309-345.

[12] Piaget, J. *Meine Theorie der geistigen Entwicklung.* Beltz Verlag, Weinheim/Basel, 1970 (2003).

[13] Piaget, J., Inhelder, B. *Die Psychologie des Kindes* Deutscher Taschenbuchverlag, München, 1972 (1993).

[14] Trainor, L.J. *Are there Critical Periods for Musical Development?* Wiley Periodicals, Inc. Dev Psychobiol 46: 262 – 278, 2005.

[15] Schögler, B., Trevarthen, C. To sing and dance together. From infants to jazz. In Bråten S. (Ed.) *On Being Moved: From Mirror Neurons to Empathy*. John Benjamins Publishing Company, Amsterdam/Philadelphia, 2007.

[16] Sobtzick et al. Strukturierte Beobachtung als wissenschaftliche Methode zur Untersuchung musikalischer Mensch-Computer-Interaktion. In Kain, S., Struve, D., Wandke, H. (Eds.) *Workshop-Proceedings der Tagung Mensch & Computer 2009.* Logos, Berlin, 2009, 176-182.

[17] www.devpsy.org/teaching/theory/piaget_ge.html

[18] www.infomus.org/EyesWeb

The influence of the meaning of lyrics on the expressed emotion of music valence

Kazuma MORI
Graduate School of Human Science,
Osaka University
Yamada oka 1-2 Suita
81-06-6877-5111
moka@hus.osaka-u.ac.jp

Keywords
music, lyric, expressed emotion, circumplex model, valence

1. BACKGROUND

The main motivation of music listening is emotional response elicited music [11]. There are two main classification of music; instrumental and song. Across culture, the majority of everyday listening includes words [3]. Otherwise phrased, the majority of everyday listening is song. Then, there are strongly linked that the emotional response elicited music and lyrics in song[1]. There were many psychological music and emotion studies from pioneer study of Hevner [5], however, few have focused on music and emotion with lyrics [7]. In previous related studies, the researchers compared emotional response to the instrumental music and to the instrumental music accompanied vocals in order to measure the influence of the lyrics (e.g. [1], [18]). But vocals have acoustic features of voice and meaning of lyrics, and possibly both of these influence the expressed emotion of instrumental music. According to Morton & Trehub [13], when singing is different in the same lyrics, that to say, the acoustic features of voice are different in the same lyrics, listener's emotional response is also different. Batcho [2] reported the lyrics presented by text gave emotional response to the participant. From these studies, both acoustic features of voice and meaning of lyrics may influence emotional response. Then, these components should separate to study music and emotion with lyrics. The present study used materials that music with lyrics in a foreign language that the participants are not understood and a written Japanese translation of the lyrics[2]. To use these materials, the present experiment examines the influence of the lyrics under conditions where the acoustic feature of voice is separated from the meaning of the lyrics. Accordingly, the present study reveals the function of lyrics in emotional response elicited music.

Since the time of ancient Greece, emotional response of music was two types: expressed emotion of music and felt emotion of music. It is important to make a distinction between the two types, because these have each different characteristic, and measuring felt emotion is more difficult than measuring expressed emotion [8]. The present exploratory study focuses on the expressed emotion of music.

Gabrielsson & Juslin [4] summed up previous research of the expressed emotion of music using multivariate analysis (e.g. factor analysis), and pointed out these researches repeatedly yielded two meaning space: arousal and valence. These meaning spaces may correspondent arousal and valence of circumplex model [15]. This model has arousal and valence axis, and put emotion in four quadrants: Activation-Positive (AP), Deactivation-Positive (DP), Deactivation-Negative (DN) and Activation-Negative (AN) (see, Figure 1). Mori, Nakamura, Yasuda, Shoda [12] examined the influence of the lyrics on the expressed emotion of the music. That study measured festive, calm and sad expressed emotion. The materials were art works that the music expressed mainly festive emotion and the lyrics expressed mainly calm and sad emotion. Thinking based circumplex model, festive emotion was included activation-positive quadrant, calm emotion was included deactivation-positive quadrant and sad emotion was included deactivation-negative quadrant. Mori et al [12] results showed the influence of the lyrics on the expressed emotion of the music was weak in the deactivation-positive quadrant. The Other hand, the influence of the lyrics on the expressed emotion of the music was strong in the deactivation-negative quadrant. These results suggested that the influence of the lyrics on the expressed emotion of the music was difficult when the lyrics and the music arousal are different. But the influence of the lyrics on the expressed emotion of the music was easy when the lyrics and the music valence are different. However, because not to measure emotion in activation-negative quadrant and to use the materials that were only one pattern combination of the expressed emotion of the music and the lyrics, it was unsatisfied to reach such a conclusion. To resolve these problems, there need the experiment that measure the expressed emotion include all quadrant of circumplex model and use the materials that were some different combination of the music and the

[1] In this paper, we define song has meaningful lyrics.

[2] In this paper, the music is defined the acoustic stimuli wherein the lyrics are in a foreign language that can't be understood by the participants and accompanied by instrumental music. The lyrics is defined written in the Japanese language formed the visual stimuli.

lyrics about valence. Such experiment can examine more comprehensively whether the influence of the lyrics on the expressed emotion of the music was easy when the lyrics and the music valence are different. Then, the present experiment measure the expressed emotion based on all quadrants of circumplex model, and focus on the music and the lyrics valence was congruent or incongruent.

2. AIMS

The present study investigates the influence of the meaning of lyrics on the expressed emotion of the music, particularly on the aspect of valence.

3. METHOD

3.1 Participants

52 participants from the Osaka university community were recruited for the experiment (22 men and 30 women). None of the participants majored in music. The mean age was 22.6 (*SD*=1.96, *range*=20-28). All the participants were volunteered.

3.2 Experimental Condition

To investigate the influence of the lyrics on the expressed emotion of the music, there were two conditions that only the music was presented (M) and the music + the lyrics were presented (ML). These two conditions were main condition. Moreover, as sub condition, there was the condition that only the lyrics were presented (L).

3.3 Experimental Design

Two factorial designs: main condition (M / ML) × emotional quadrant (Activation-Positive [AP] / Deactivation-Positive [DP] / Deactivation-Negative [DN] / Activation-Negative [AN]). Both these factor are within participants.

3.4 Materials

The materials selected pop/rock art works commercially available. Because college students easily show emotional response to pop/rock [9], the materials were selected from this genre. The art works were the music expressed only positive emotion. The Positive music used two types, activation and deactivation, on the basis of circumplex model. These lyrics were valence congruent with the music (positive valence) or valence incongruent with the music (negative valence). Then, the art works were the music and the lyrics valence congruent or incongruent. In following, as Table1, each works type call C-AP (Congruent Activation-Positive), C-DP (Congruent Deactivation-Positive), I-AP (Incongruent Activation-Positive) and I-DP (Incongruent Deactivation-Positive).

The materials select, to begin with, was done by experimenter's subjective judge. Experimenter was screening of 500 works of art works, 12 works which may have C-AP, C-DP, I-AP and I-DP type combination of the expressed emotion of the music and the lyrics were selected. Next, the 12 works were pilot tested with 9 graduates by the following method of the present study experiment. The pilot test was done each M and L. The materials ranged in length from 82.5 to 118.5s (*M* = 102.2s). All the music was faded out last 5s. The music selected criteria was that the AP quadrant or the DP quadrant of the rating scales were higher mean rating than any other quadrants. The lyrics were selected in these music. The lyrics selected criteria was that positive quadrants (the AP and the DP quadrant) mean ratings was higher negative quadrants (the DN and the AN quadrant) mean ratings (congruent works), or negative quadrants mean ratings was higher positive quadrants mean ratings (incongruent works). The art works fulfilled these criteria was six out of 12. The C-AP and the C-DP were each one, the I-AP and the I-DP were each two. Then, in the present experiment, six materials were used. (Appendix)

Table 1. Materials types and quantity matrix

Type	Combination	Music	Lyrics	Quantity
C-AP	Congruent	Activation-Positive	Positive	1
C-DP		Deactivation-Positive	Positive	1
I-AP	Incongruent	Activation-Positive	Negative	2
I-DP		Deactivation-Positive	Negative	2

3.5 Apparatus

All the experiment conducted the sound-proof room in the Osaka University. The presentation of the materials was controlled Microsoft Visual C++ 2008 programming in personal computer (FMVNFB70WC). The music played from the speaker (SONY, TA-DVD-V88ES) through the amplifier (SONY, SS-AL5), as all the music about 62dBA ($L_A$eq). This sound level may propriety level. The distance of the speaker and the participants was about two meters. The lyrics were presented in the center of the 19 inch monitor's screen (I/O DATA, LCD-AD193GW). The text of the lyrics changed constantly, in sync with the vocals of the music. The text format was SHIFT JIS and size was 50 point. In the pilot test, the participants reported that changing speed of the lyrics was not too quick. The distance of the monitor and the participants was about one meter.

3.6 Questionnaires

The rating scales were used the following three scales for each quadrant of emotional circumplex model [15]. "Activation-Positive" (AP) quadrant: exciting, festive, and sunny; "Deactivation-Positive" (DP) quadrant: calm, warm, and relaxing,; "Deactivation-Negative" (DN) quadrant: melancholic, sorrowful, and lonely; "Activation-Negative" (AN) quadrant: stressful, unsettling, and disconcerting. The ratings were based on a seven point scales, where "0" represented "not at all" and "6" represented "very strongly". Not to rate the same quadrant continuously, the 12 scales were lined up randomize, and presented all tests the order. If the participants had episode memory for the music, memory bias may make emotional response differ other participants [10]. The participants were asked whether know or not the materials in M and L. If the participants understand language meaning in M, ML was not different M for the participants. The participants were asked how understand the meaning of lyrics in M by using percentage. Furthermore, gender and age or other questions were asked.

3.7 Procedure

All the participants were subjected to the three experimental conditions. The participants were tested in one or groups of two or three. First, the participants were subjected to M. In the same day, they were subjected to L. We put some minutes break between M and L. More than 2 weeks ago, the participants subjected to ML. In every condition, first of the experiment, the practice material was played and the participants were shown how to complete the rating scales of questionnaire to familiarize with the response format. The six experimental materials in each condition were then presented, and participants rated the

Table 2. Component loadings of the rating scales in the principle component analysis

Rating scales	Component1	Component2
Exciting	-.47	.68
Festive	-.64	.61
Sunny	-.76	.34
Warm	-.63	-.22
Relaxing	-.43	-.59
Calm	-.28	-.72
Sorrowful	.74	-.33
Lonely	.82	-.19
Mealancholic	.84	.01
Disconcerting	.73	.30
Unsettling	.68	.38
Stressful	.50	.54
Eigenvalue	5.1	2.5
% of Variance	42.1	21.1

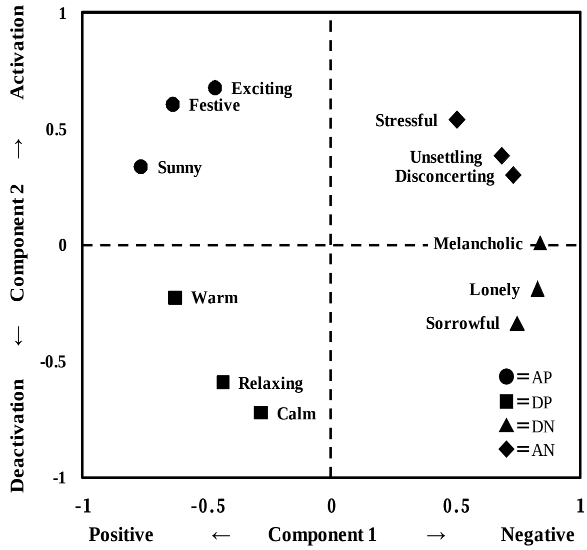


Figure 1. There was the plot of component loadings. The first component interpreted as valence axis, the second component interpreted as arousal axis.

12 scales and answered other questions each immediately after the presentation. After the participants had completed questionnaires, the next material was presented. The six materials of each condition were presented with counterbalance across the experiment. In M and ML, participants were instructed to rate for music.

4. RESULTS

The ratings of two participants who knew the material were excluded from the following analysis. From the report of the participants, it was confirmed that none of them could comprehend the meaning of lyrics in M (M = 2%, SD = 0.12). All of the other ratings were analyzed as follow.

4.1 Data Reduction

The unrotated principal component analysis yielded two components. The components yielded until two, because this analysis purpose to set the rating scales in two dimensions. The cumulative accounted variance was 63.2%. (Table2). The first component interpreted as valence axis because the rating scales for the AP and the DP quadrant load negatively, the DN and the -AN quadrant load positively on this components. The second component interpreted as arousal axis because the rating scales for the DP and the DN (except 'melancholic' loading +0.01) quadrant load negatively, the AP and the AN quadrant load positively on this components (Figure 1). These results showed four quadrants could be identified each other for the participants. The coefficient of reliability was conducted according to three rating scales for each quadrant. Cronbach's alpha coefficients showed high values, the AP quadrant: α = 0.88, the DP quadrant: α = 0.87, the DN quadrant: α = 0.88, the AN quadrant: α = 0.86. These analyses showed the rating scales in each quadrant exhibited sufficient internal consistency. Then, the mean ratings were calculated for each quadrant in each condition for each type of the art work (Figure 2).

4.2 Congruent Work

A 2(main condition: M / ML) × 4(emotion quadrant: Activation-Positive / Deactivation-Positive / Deactivation-Negative / Activation-Negative) repeated-measures ANOVA was conducted each type in order to study the influence of the lyrics on the expressed emotion of the music. In the C-AP work, the result showed no interaction, F (2.19, 109.13) = 0.44, *n.s.* The result also showed no main effect of condition, F (1, 49) = 1.00, *n.s.* Then, the influence of the lyrics on the expressed emotion of music was not found. In the C-DP work, the 2 × 4 ANOVA showed the interaction, F (1.88, 92.02) = 9.81, $p < .001$. There was a simple main effect for the condition of the AP quadrant and the DN quadrant, F (1, 49) = 5.14, 12.12, respectively, $p < .03$. Then, the expressed emotion of the music was changed by the influence of the lyrics in the AP quadrant and the DN quadrant. Moreover, in the DN quadrant, the ML ratings (M = 1.6, SD = 1.1) were marginally lower than the L ratings (M = 2.0, SD = 1.3), t (49) = 1.79, $p < .08$. This is a difficult to think the influence of lyrics in the DN quadrant. This may because the music deactivation valence became clearly positive by the influence of the lyrics expressed positive emotion. The positive lyrics might decrease the deactivation-negative emotion of the music.

4.3 Incongruent Work

In the I-AP work, the 2 × 4 ANOVA showed the interaction, F (2.16, 105.73) = 109.49, $p < .001$. There were a simple main effect for the conditions of all quadrant, F (1, 49) = 103.15, 45.97, 88.00, 24.08, respectively, $p < .01$(F value order was the AP quadrant, the DP quadrant, the DN quadrant and the AN quadrant). In the I-DP work, the 2 × 4 ANOVA also showed the interaction, F (2.35, 115.09) = 59.38, $p < .001$. There were a simple main effect for the conditions of each quadrant, F (1, 49) = 7.33, 47.88, 47.19, 25.86, respectively, $p < .01$(F value order was the AP quadrant, the DP quadrant, the DN quadrant and the AN quadrant). Then, in incongruent works, the influence of the lyrics on the expressed emotion of music was clearly found. Thinking the influence of the lyrics in more detail, there may be the differences of it in each quadrant. To investigate the differences, the Euclidean distance was calculated between the each participant's rating for ML and M, and between those for ML and L. The distances were calculated for each quadrant and each type of the art works (Table 2). Only for the DP quadrant

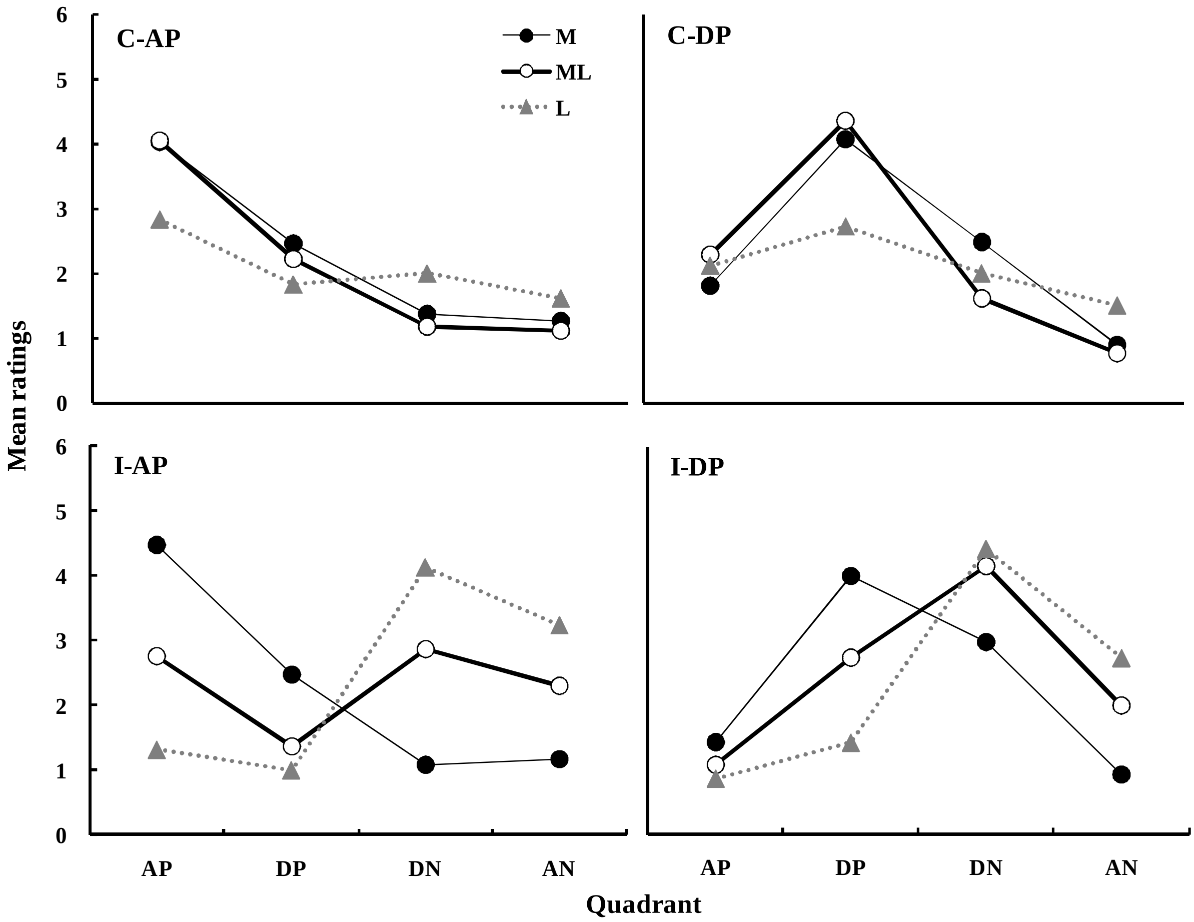


Figure 2. Mean ratings for each condition of each work types.

of the I-AP work, the distance between ML and L was smaller than that between ML and M, t (49) = 2.57, p < .02. Only for the DN quadrant of the I-DP work, the distance between ML and L was smaller than that between ML and M, t (49) = 2.81, p < .01. Then, in only these quadrants, the expressed emotion of the music + the lyrics neared the lyrics, that to say, the influence of the lyrics was strong. In any other quadrants of both types, the expressed emotion of the music + the lyrics located center of the music and the lyrics, that to say, the influence of the lyrics was mid. These results showed, in incongruent works whether the music activation or deactivation, the music deactivation valence was clearly changed positive into negative (see, Figure 2).

Table 3. Mean Euclidian distance (SD) of between M and ML or L and ML.

Type	Interval of Euclidean distance	Quadrant			
		AP	DP	DN	AN
I-AP	ML−M	1.8 (1.1)	1.4 (0.9)	1.9 (1.2)	1.7 (1.0)
	ML−L	1.6 (1.1)	0.9 (0.7)	1.4 (1.0)	1.5 (0.9)
I-DP	ML−M	0.9 (0.6)	1.5 (1.0)	1.3 (1.0)	1.5 (1.0)
	ML−L	0.7 (0.5)	1.5 (1.1)	0.8 (0.5)	1.5 (1.0)

5. DISCUSSION

The present study examined the influence of the meaning of lyrics on the expressed emotion of the music under conditions where the acoustic feature of voice is separated from the meaning of lyrics.

First, the results showed the influence of the lyrics on the expressed emotion of the music was strong when the lyrics and the music valence are different. This result support author's previous study [12]. Valence incongruent lyrics may clearly change the expressed emotion of music. Furthermore, the influence of the lyrics was different between the AP works (the C-AP and the I-AP) and the DP works (the C-DP and the I-DP). In the AP works (the music mainly expressed the activation-positive quadrant of emotion), when the music and the lyrics valence were congruent, the expressed emotion of the music was not influenced by the lyrics. When the music and the lyrics valence were incongruent, the expressed emotion of the music was influenced by the lyrics, but the deactivation valence was more influenced than the activation valence. On the other hand, in the DP works (the music mainly expressed the deactivation-positive), whether the music and the lyrics valence were congruent or incongruent, there was the influence of the lyrics. When the lyrics and the music valence were congruent, the expressed emotion of the music in the deactivation-positive quadrant was maintained, and the expressed emotion of the music in the deactivation-negative quadrant became lower than the lyrics. When the lyrics and the music valence were incongruent, the expressed emotion of the music in the deactivation-positive quadrant became weakly, and only the expressed emotion of the music in the deactivation-negative quadrant clearly strong. These results suggest the influence of lyrics may especially strong the expressed emotion of music deactivation valence. Iwai & Adachi [6] suggested sad (deactivation-negative emotion) lyrics easily influenced the expressed emotion of music. However, this study only used happy (activation-positive emotion) and sad (deactivation-negative emotion) lyrics. The present study used the lyrics expressed across all four quadrants of emotion in circumplex model, and found lyrics may easily influence the music deactivation valence. As is well known, calm (deactivation-positive emotion) expressed by instrumental music was confused with sorrow (deactivation-negative emotion) [4]. The present study results suggest the meaning of lyrics may have a function that clearly distinct these two expressed emotion. Juslin [7] said that "the combination of words and music offers very precise means of communication of emotions, in which the two 'channels' —verbal and non-verbal— complement each other." The meaning of lyrics may have such a function especially in deactivation music.

However, the problem of the present study was that materials were few, and works type quantity was one-sided. The future study should use many more materials and equally type quantity. Furthermore, this experiment used only positive music. Although fear (activation-negative emotion) is rarely conveyed by art music [19], to support this experiment suggest, there need the experiment that use negative music. And the limitation of the present study is, first, the lyrics presented from visual. The difference of emotional response between only the music and the music + the lyrics might reflect the difference between auditory and visual material reception. Noulhiane, Mella, Samson, Ragot & Pouthas [14] pointed out the way emotional response process between audition and vision may different. Schmidit & Trdinor [16] also pointed out the brain area of emotional response process between audition and vision may different. To eliminate the influence of material reception from auditory and vision, the future study need the way present only audition to examine the influence of the meaning of lyrics. Second, although many styles of music express different emotions as time unfolds [17], the present study did not reflected time series information in emotional response. Same as the information of music change momentarily, the information of lyrics change momentarily. With measure emotional response along time series, the more detail could get the influence of the meaning of lyrics on the expressed emotion of music.

6. CONCLUSIONS

The present study suggests the part of the influence of lyrics on the expressed emotion of music. In previous related studies of music and emotion with lyrics, it is not necessarily the case that these results reflected the influence of the meaning of lyrics because these studies were not separate the acoustic features of voice and the meaning of lyrics. On the contrary, this study successful suggests the influence of the meaning of lyrics on the expressed emotion of music. Because music with lyrics the genre that makes up the majority of our everyday listening, this study may give trivial but significant finding in music and emotion study field.

7. IMPLICATIONS

When we listen to music, deactivation music may strongly communicate meaning of lyrics.

8. REFERENCES

[1] Ali, O. S., & Peynircioğlu, F. Z. Songs and emotions: are lyrics and melodies equal partners? *Psychology of Music*, *34*, 4 (Oct. 2006), 511-534.

[2] Batcho, K. I. Nostalgia and the emotional tone and content of song lyrics. *American Journal of Psychology*, *120*, 3 (Fall. 2007), 361-381.

[3] Clarke, H. L. The basis of musical communication. *The Journal of Aesthetics and Art Criticism*, *10*, 1952, 242-246.

[4] Gabrielsson, A., & Juslin, P. N. Emotional expression in music. In *Handbook of Affective Sciences*, R. J. Davidson., K. R. Scherer., & H. H. Goldsmith., Eds. Oxford University Press, Oxford, NY, 2003, 503-534

[5] Hevner, K. The affective character of the major and minor modes in music. *American Journal of Psychology*, *47*, 1935, 103-118.

[6] Iwai, K., & Adachi, M. Japanese college students' emotional responses to J-POP-like songs on physiological and psychological measures. In *Proceedings of the 10th International Conference of Music Perception & Cognition* (Hokkaido, Sapporo, August 25-29, 2008). ICMPC Press, ICMPC10, 2008, 490-495.

[7] Juslin, P. N. From mimesis to catharsis : expression, perception, and induction of emotion in music. In *Musical Communication*, D. Miell., R. MacDonald., & D. J. Hargreaves., Eds. Oxford University Press, Oxford, England, 2005, 85-115.

[8] Juslin, P. N., & Laukka, P. Expression, perception, and induction of musical emotions: a review and a questionnaire study of everyday listening. *Journal of New Music Research*, *33*, 3 (2004), 217-238.

[9] Juslin, P. N., Lijeström, S., Västfjäll, D., Barrades, G., & Silva, A. Emotional reaction to music in everyday life: music, listener, and situation. *Emotion*, *8*, 5 (Oct. 2008), 668-683.

[10] Juslin, P. N., & Västfjäll, D. Emotional response to music: the need to consider underlying mechanisms. *Behavioral and Brain Sciences*, *31*, 5 (Dec. 2008), 559-621.

[11] Krumhansl, C. L. Music: a link between cognition and emotion. *Current Directions in Psychological Science*, *11*, 2002, 45-50.

[12] Mori, K., Nakamura, T., Yasuda, S., & Shoda, H. The influence of lyrics on impressions about music, In *Proceedings of the 6th Conference of Japanese Society for Cognitive Psychology* (Chiba, Chiba, May 31-June 1). JSCP Press, Japanese, 2008, 100.

[13] Morton, J. B., & Trehub, S. E. Children's judgments of emotion in song. *Psychology of Music*, *35*, 4 (Oct. 2007), 629-639.

[14] Noulhiane, M., Mella, N., Samson, S., Ragot, R., & Pouthas, V. How emotional auditory stimuli modulate time perception. *Emotion*, *7*, 4 (Nov. 2007), 697-704.

[15] Russel, J. A. A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 6 (Dec. 1980), 1161-1178.

[16] Schmidt, L. A., & Trainor, L. J. Frontal brain electrical activity distinguishes valence and intensity of musical emotions. *Cognition & Emotion*, *15*, 4 (Jul. 2001), 487-500.

[17] Schubert, E. continuous measurement of self-report emotional response to music. In *Music and Emotion: Theory and Research*, P. N. Juslin., & J. A. Sloboda., Eds. Oxford University Press, Oxford, England, 2001, 393-414.

[18] Stratton, V. N., & Zalanowski, A. H. Affective impact of music vs. lyrics. *Empirical studies of the arts*, *12*, 2 (Oct. 1994), 173-184.

[19] Terwogt, M. M., & Van Grinsven, F. Musical expression of moodstates. *Psychology of Musi*c, *19*, 2 (Oct. 1991), 99-109.

9. Appendix

C-AP work

1. Tout envoyer en l'air (KYO, 2008), Model number in Japan: BVCP-21585, Language: French, Time: 112.5s

C-DP work

1. Em Meus Bracos (Various Artists, 2008), Model number in Japan: VUCC-10009, Language: Portugal, Time: 107.5

I-AP works

1. OCH JAG GRÄT MIG TILL SÖMNS EFTER ALLA DAR (SÄKERT!, 2008), Model number in Japan: XQBZ-1603, Language: Swedish, Time: 102.0
2. Regresa A Mi (Thalia, 2004) , Model number in Japan: VJCP-68628 , Language: Spanish, Time: 96.5

I-DP works

1. Eclipse (Lisa Ono, 2005), Model number in Japan: TOCT-90030, Language: Portugal, Time: 114.5
2. RETOUR A GINZA (Nathalie, 2000), Model number in Japan: AICT-1173, Language: French, Time: 116.5

Rock to the Music: Theory Exploration and Electroencephalography (EEG) Investigation

Li-ching Wang
Graduate Institute of Musicology
National Taiwan University
No.1, Sec. 4, Roosevelt Road,
Taipei, 10617 Taiwan
+886-2-33664699
r96144005@ntu.edu.tw

Chen-gia Tsai
Graduate Institute of Musicology
National Taiwan University
No. 1, Sec. 4, Roosevelt Road,
Taipei, 10617 Taiwan
+886-2-33664699
tsaichengia@ntu.edu.tw

ABSTRACT
The relationship between musical rhythms and body movements is evident, but the underlying mechanisms are not fully understood. It has been suggested that the vestibular system plays an important role in the perception of rhythm, exemplified by simply rotating the head. This kind of movement enables listeners to follow rhythm more easily, especially syncopated rhythm. Body movements may facilitate beat-chunking, which decreases the cognitive needs of perceiving complex rhythm. The present study, examines the phenomenon of "swaying with music", and at the same time evaluates electroencephalography (EEG) and movements. The result demonstrates that the movements of different parts of the body change the brain wave on the frontal midline (Cz). This result suggests the key role of the supplementary motor area in integrating auditory information and vestibular sensation. We argue that the perception of rhythm requires hierarchical management of subdivided beats through chunking, and vestibular stimuli are the most important components contributing to chunking for syncopated rhythm.

Keywords
Musical rhythm, movement, chunking, vestibular stimuli, EEG.

1. BACKGROUND
There is a close relationship between music and body movement. Even those who do not receive musical training involuntarily tap their feet or nod their heads with music. Moving with music is a phenomena observed in many cultures [3]. Biomusicologist Ian Cross has pointed out that "generalizable definition of music would refer to music's two roots in sound and movement."[7].

Rhythm is the most important musical aspect that affects body movements. We usually move on beat when tapping or dancing with music. A beat is the basic unit of measuring music time, suggesting the received rhythm in the time section of the same length. Regular and alternate up and down beats can be grouped and received in the form of meters. The arrangement of beats can be received through meters, but perceptually the beat, the most prominent rhythmic level, is the level that people choose to move with. It is hypothesized that metric structure and the sense of beat stimulate the body movement with rhythm and the auditory-motor interactions [6]. The feedback of auditory input and movement enables music players to count beats more accurately because they both fill the structure into the blank spaces between beats [26]. Human beings have the ability of synchronizing movements with stimuli, especially with auditory stimuli [22]. Among the motor-related stimuli and sensory feedback, acceleration of a part of the body is the main factor which enables rhythmic synchronization [14].

Human beings start to feel the rhythm early in their infant stage. Infants have the abilities to move with rhythm. Besides, being moved rhythmically by adults affects the rhythm perception of infants [18]. For adults, voluntary body movements also affect rhythm perception. Simply tapping the foot in different ways results in different rhythm perception [19]. Therefore, the interaction between body movement and auditory input not only develops very early but also plays an important role in rhythm perception.

Besides auditory feedback, music players also perceive other feelings when they are performing; for example, tactile sense, proprioception, and vestibular sense which also play roles in the music performance [1]. The tactile feedback on the fingertips resulting from musical instrument playing makes performers sense beats more accurately [9]. The proprioception resulting from the changes of the relative positions between body portions and musical instruments or between different body portions [2], and the vestibular feedback resulting from head movements both enable performers to control the rhythm more easily [18][19]. Listeners can also consolidate the rhythmic structures in their minds through different feedbacks when moving with the music.

It is very often observed that singers lead listeners to clap hands or wave arms in concerts, but different singers move in different ways. Some may tap down-beat while some tap up-beat. It has been pointed out that more inaccuracies result when participants are asked to tap up-beat (syncopation, anti-phase) than down-beat (synchronization, in-phase). Besides, tapping up-beat not only takes more reaction time [11], but also needs more attention [4]. If the frequency of the music beat increases to more than 2Hz, participants will switch the up-beat movement into down-beat one for it is difficult to maintain syncopation [10].

Electroencephalography (EEG) and Megnoencephalography (MEG) studies have shown that the power of the beta band decrease when participants are doing up-beat tasks, meaning syncopation tasks demand more cognition needs [5][16]. Not only the movement patterns such as up-beat and down-beat result in different EEG recordings, the body portions that move with music also affect the consequence [17].

Trainor and colleagues suggested that the vestibular system plays an important role in the perception of rhythm [24]. Vestibular stimuli can be produced through simply rotating the head [25]. Thus the movement of the head is of the most importance when swaying with music, for it can stimulate the vestibular system, enabling listeners to follow the rhythm more easily, especially syncopated rhythm [20][12].

Recent studies show that tapping rhythmically activates the mesial frontal lobe and the premotor cortex, while the EEG-EMG (Electroencephalography-electromyography) coherence happens in 15-30 Hz [21]. This kind of rhythmic tapping mode enables the spinocortical signal feedback, and the SMA is a part of the network generating rhythmic mode. These studies also pointed out that the spinocortical feedback modes of hand and foot movements are similar. A study of EEG and fMRI also revealed that the auditory-evoked potentials can be most conspicuously detected in the SMA [15] which is responsible for the supplementary motor area (SMA) and the vestibular stimuli project region.

2. AIMS

Although previous studies have suggested that body movements affect the perception of musical rhythm, the mechanisms causing music performers or listeners to sway with music are still unclear. The aim of the present study is to reveal the relationships between the types of body movements and the patterns of physiological signals.

3. MAIN CONTRIBUTION

Based on previous neuroscientific investigations, the present study compares the EEG signals, acceleration of body movement, and skin conductance level (SCL) during music listening and when moving with music. This study is expected to advance our knowledge of rhythm processing and movement coordination.

3.1 Methods

3.1.1 Participants

Fifteen students (3 males and 12 females, all between 19-27 years old) in National Taiwan University participated in this experiment. All participants maintain the habit of listening to music.

3.1.2 Stimuli

Three sections of music with different degrees of syncopation were chosen, of which two had higher degrees of syncopation, and the other (steady and timbre-invariant drum strokes) had a lower degree. The degree of syncopation is defined by the model proposed by Longuet-Higgins and Lee (See Figure 1.) [13][8]. According to Fitch and Rosenfeld, the higher the degree of syncopation, the more complex the rhythm [8]. Therefore, we used three pieces of music with different degrees of syncopation as the stimuli. The transcriptions of these stimuli are shown in Figure 2. Their tempo was chosen to be around 2.2 Hz, as it is difficult for people to maintain the up-beat in tempos faster than 2 Hz [10].

3.1.3 Apparatus

Music was played using an earphone connected with a computer in a controlled room in the Graduate Institute of Musicology. EEG signals were collected at the SMA projected region Cz on each participant's scalp. Motion measurements were carried out by a one-dimensional accelerometer. Epidermal activity (EDA) was also measured. The signals of EEG, accelerometer and EDA were all recorded by Biopac Student Lab PRO.

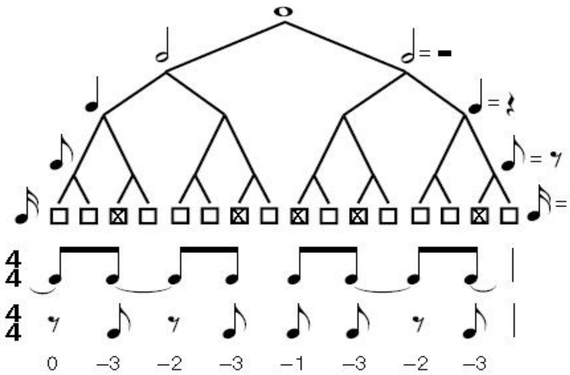


Figure 1. A Longuet-Higgins and Lee (1984) rhythm tree with an example rhythm and corresponding event weights. Every note is given a weight according to their hierarchical status, with stronger beats having more weight. Syncopations occur when a rest or tied note is preceded by a sounded note of a lower weight. The difference in weights is the syncopation value for each such pair, and the total syncopation value for the rhythm is the degree of syncopation (Fitch and Rosenfeld, 2007).

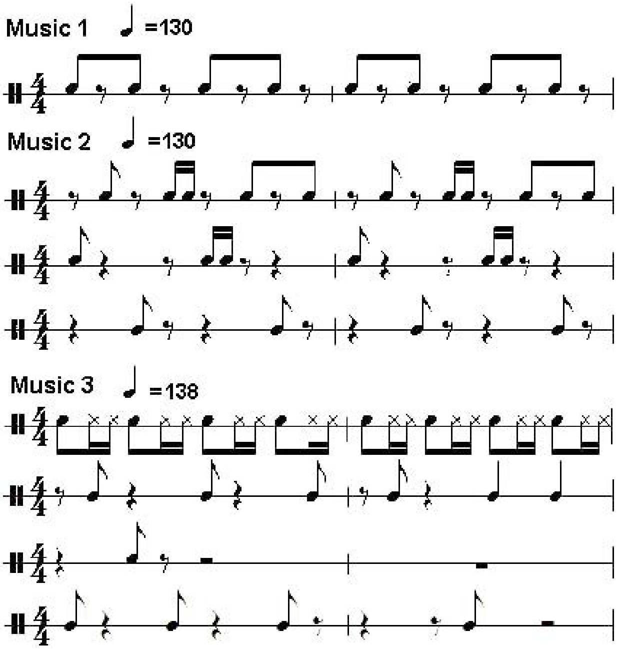


Figure 2. The transcriptions of the three stimuli. They have several percussive timbres and were processed through Cooledit software. Music 1 is the simplest stimulus without syncopation while the others contain more complex rhythms with syncopation. Music 3 is more difficult in both rhythm and tempo.

3.1.4 Procedure

The experiment focused on two main factors in three musical stimuli with different degrees of syncopation: the portions of the body that actively move, and how the movements correspond to the musical beats. The participants were asked to listen to these stimuli and move in various ways.

The experiment included two stages. The first stage contained three sections: resting without music for 45 seconds; being regularly shaken without music for 15 seconds; and listening to three musical stimuli, each for 25 seconds. The reason of including the condition of "passively pushed" was to generate vestibular stimuli. There were no active muscle contractions during section one.

In the second stage, the participants listened to three musical stimuli randomly and were instructed to accomplish the tasks below: nodding on down-beats, nodding on up-beats, tapping the right index finger on down-beats, tapping the right index finger on up-beats, tapping the right big toe on down-beats, and tapping the right big toe on up-beats. The participants were asked to sit on a chair in a relaxed position and with eyes closed. EEG, accelerometer and EDA recordings lasted for 30 seconds, from 5 seconds before the music to 25 seconds after the music played. Table 1 shows a list of all the tasks in the experiments.

All participants were only allowed to move one part of his/her body. For example, when a subject was asked to nod his/her head, he/she should not move his/her fingers or toes. In order to minimize EEG artifacts, every movement type was limited to a small range.

After these two stages of data recording, the participants were asked to talk about their feelings and their music interests/habits.

Table 1. Lists of all the tasks in the experiment. Abbreviations. M0: stands for music number zero which has the lowest degree of syncopation. M1 and M2 stands for music number one and two which have higher degrees of syncopation. H: head, F: finger, T: toe, D: down-beat, U: up-beat.

Stage 1:					
Rest without music for 45 seconds					
Being shaken without music rhythmically for 15 seconds					
M0 for 26 seconds		M1 for 25 seconds		M2 for 25 seconds	
Stage2: each for 30 seconds					
HM0U	HM0D	HM1U	HM1D	HM2U	HM2D
FM0U	FM0D	FM1U	FM1D	FM2U	FM2D
TM0U	TM0D	TM1U	TM1D	TM2U	TM2D

3.1.5 Signal Processing

The EEG signals were analyzed in the frequency domain and the time domain. In the time domain, we focused on the phase relationship between the periodic body movements and the low-passed EEG signals. Therefore, we compared the beats, the EEG signals, and the accelerometer signals. In the frequency domain, we used FFT (Fast Fourier Transform) to extract the spectral features of EEG signals, such as the harmonic peaks in the theta, alpha and beta bands.

Average SCL for every task, including nodding on down-beats, nodding on up-beats, tapping the right index finger on down-beats, tapping the right index finger on up-beats, tapping the right big toe on down-beats, and tapping the right big toe on up-beats, was calculated in terms of EDA integration divided by total activity time using Matlab software. Their ratios were compiled into statistics using SPSS software.

3.2 Results

In this study, the EEG signals when participants were listening to the music without any passive or active movements reveal no specific harmonic peaks which were observed in Mayhew and his colleagues' experiment [15].

The EEG signals when the participants were passively pushed showed that the periodical movements without active muscle contraction induced quasi-sinusoidal waves, which also appear in the EEG signals when the participants actively nod their heads along with the beats. Figure 3 shows the EEG recording of a participant being pushed, and the EEG recording and the accelerometer recording when she was actively nodding her head (EEG signal low-pass filtered at the cutoff frequency of 20 Hz). The time domain analysis of the EEG recordings also shows that when the movement acceleration is crossing zero, the EEG voltage reaches its maximum.

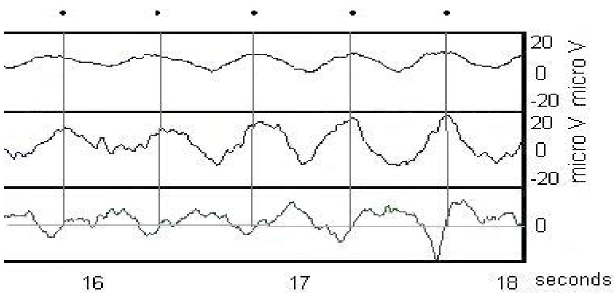


Figure 3. EEG recording of one participant being pushed on the beat (first row), EEG recording of the same participant nodding on the beat (second row), and a accelerometer recording simultaneous with that of the second row (third row). The dots on the upper frame indicate the times when the acceleration of the head is crossing zero. As the vertical lines show, when the movement acceleration is crossing zero, the EEG voltage reaches its maximum.

The time domain analysis of the EEG recordings when participants were nodding their heads, tapping their fingers and toes reveals different brain wave patterns. We observed that the EEG signal for nodding heads is especially periodic (Figure 4).

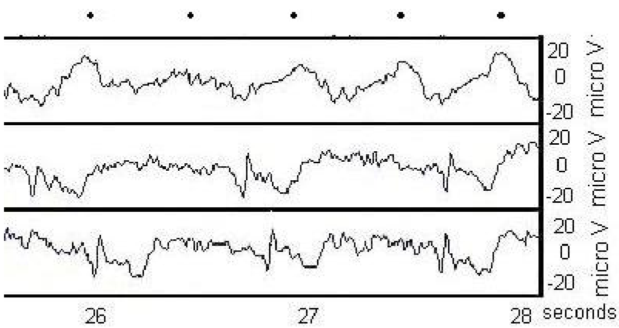


Figure 4. EEG recordings of one participant nodding her head (first row), tapping her right index finger (second row), and tapping her right big toe (third row) with music, all on the down-beat. The dots on the upper frame indicate the times when the acceleration of the head is crossing zero.

The time domain analysis of the accelerometer signals on the head, finger, and toe shows different moving patterns for each part of the body (Figure 5).

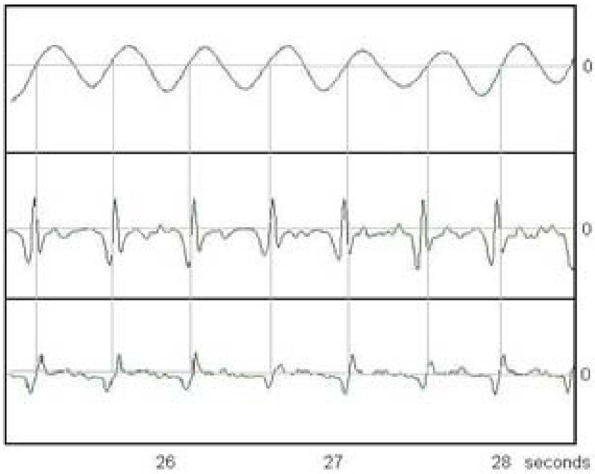


Figure 5. The time domain analysis of the accelerometer signals on the head, right index finger, and right big toe indicates different moving patterns for different parts of the body. The vertical lines show the times when the head movement acceleration is crossing zero.

The frequency domain analysis of the EEG recordings reveals that periodic head or right big toe movements induce harmonic peaks in the theta band. These peaks reflect the performance of the tasks. The EEG spectra reveal the difference between subjects in every task (Figure 6). In some participants, the 1 alpha wave decreases when tapping on the up-beat, which is consistent with previous research [5][16]. The fundamental frequency is apparent in the task of nodding. It is most unclear in the tasks of tapping the right index finger. The harmonic peaks in the theta band of the EEG spectrums during active nodding are similar to those during passively-pushed conditions.

The EDA analysis shows that in most cases, skin conductance level dramatically increased right after participants started to move with music. Three-way ANOVA reveals a significant main effect of body portions which actively moved (sig. = 0.003 < 0.05), and *post hoc* Tukey HSD test shows the effect between head and toe movements is significant (sig = 0.002 < 0.05). The EDA changes of a representative participant are shown in Figure 7.

3.3 Discussion

This study explores the relationships between the types of body movements and the patterns of physiological signals. EEG and SCL recordings may reflect different strategies for movement coordination to accomplish different tasks and the cognitive resources they needed. Swaying the head resulted in more clear

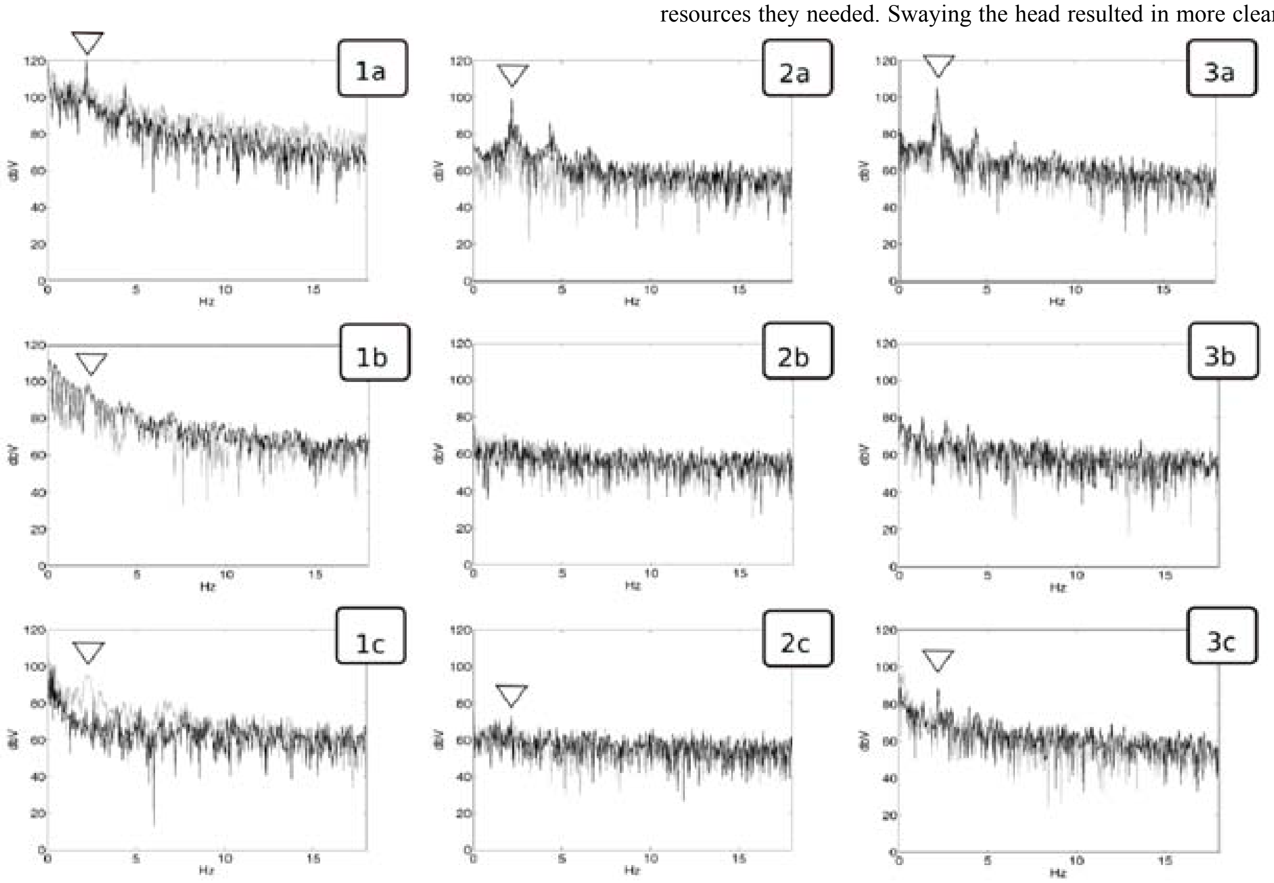


Figure 6. EEG spectrums of three participants (1, 2, and 3) taken when nodding their heads (a), tapping their right index fingers (b), and tapping their right big toes (c) with music on the down-beat (light) and on the up-beat (dark). The musics the participants were listening to were of the same tempo with 2.17Hz beat frequency. As observed, there is a remarkable fundamental peak in 2.17 Hz in the EEG recorded during nodding. The results of tapping right index fingers and right big toes are different.

harmonic peaks in the EEG spectra above the SMA, whether the motion was on the up-beat or down-beat. The fact that the SMA is also the vestibular stimuli represented area may provide an explanation of why people feel it is more natural to sway their heads with music, compared to other body movements. When movement acceleration is zero, the EEG voltage reaches its maximum, and since zero is passed through in a very short time, the brain must perceive the duration very clearly and precisely. The message of "head acceleration passing the zero point" is important for the brain.

The harmonic peaks of EEG spectra in the theta band reflect the periodicity of head/toe movements. It is observed that the performance of nodding or tapping the toe on the up-beat is better than that of nodding or tapping on the up-beat. The power of theta/alpha/beta bands for tasks may reveal different levels of attention and cognitive loading. The feeling and the feedback to the tasks differed between participants, and the differences were also shown in their EEG recordings.

The presence of harmonic peaks in the EEG spectrums when the participants' heads were moving or moved may be due to vestibular stimulation. The signals from the vestibular system when it is stimulated by head movements which were directly transmitted to SMA therefore influenced the EEG signals in the Cz channel. Compared with the stimuli from the vestibular system, the stimuli from finger and toe movements resulting in tactile sense are apparently not strong enough to affect the EEG recording. This study is also the first EEG study to ask participants to actively nod their heads with music. To ascertain that the EEG signal is induced from the vestibular stimuli rather than the EMG of neck muscle contraction, we examined the EEG recorded during passively-pushed conditions without any muscular activity. We found the EEG signals of both passive and active head motions were similar. This finding provided more evidence for the idea that the SMA integrates auditory input [15] and vestibular stimuli [25] for movement coordination.

SCL reveals the participant's arousal level. Its increase after starting the movement may be due to the difficulty of the tasks or the feeling of excitement. According to interviews and observations during the experiment, it is more difficult for the participants to tap on the up-beat, but the difficulty of nodding, tapping with fingers or toes varies across participants. Most of the participants indicated that index finger movement is the most flexible one followed by big toe movement, but they felt more natural to move their heads with the music. Some participants who have the habit of dancing also mentioned that they strongly wanted to move their legs and bodies. SCL analysis shows that the EDA (epidermal activity) of the participants when moving their heads is the lowest compared to tapping other parts of the body, whether the movement is done on the up-beat or the down-beat. It may suggest that head movement with music is more spontaneous, and therefore less cognitive resources are needed

4. CONCLUSION

It is suggested that the vestibular system can be facilitated through head motion which increases the feeling of rhythm [24]. In addition, limb movements also enhance the feedback of musical rhythm [21]. The present study shows that the EEG signals during movement with music depends on which part of the body (head nodding, toe tapping) they move to music. The head/toe movements tend to reinforce the feeling of steady beats [21], and the frontal midline region receives sensory signals of these movements. This result supports the hypothesis that syncopated rhythms invoke more body movements from the listeners; the perception of rhythm requires hierarchical organization of subdivided beats through chunking, and vestibular stimuli are the most important components contributing to chunking for syncopated rhythm. The reason that syncopated rhythms cause body movements more easily may be that they stimulate chunking [23]. Another major finding is that the EEG voltage reaches its maximum on the beat of the movement, and at the same time the acceleration of head movement passes zero.

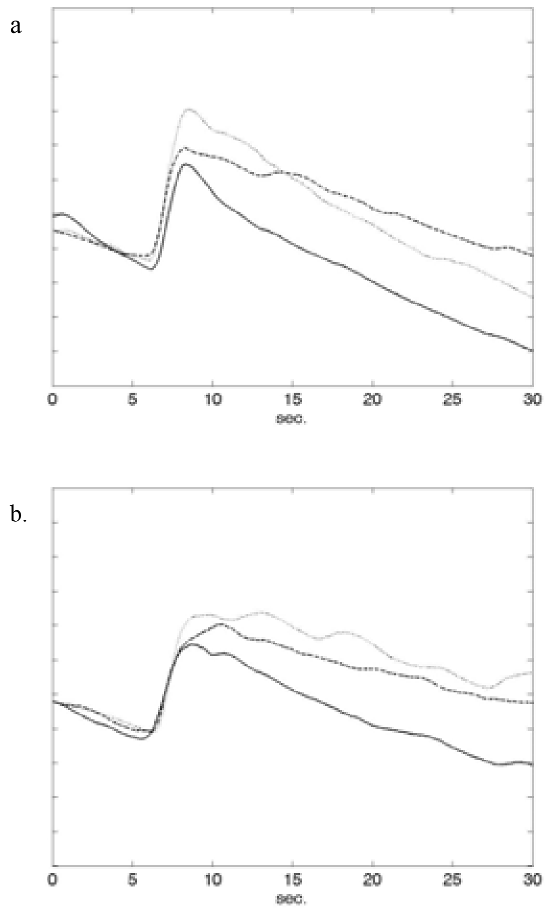


Figure 7. The EDA recording when nodding head (gray dotted lines), tapping right index finger (black dashed lines), and tapping right big toe (black solid lines) with music in the down-beat (a) and up-beat (b) tasks.

5. IMPLICATION

EEG signals have been seldom recorded during head motion. The present study shows that similar EEG signals on the frontal midline can be induced by periodic head motion and toe motion. Although the meaning of these frontal midline activities remains to be clarified, we suggest that EEG recordings may be a useful tool for exploring the neural and cognitive processes involved in moving the body with music. Whereas the present study provides preliminary information of the physiological responses while moving with music, future investigations are needed for clarifying the integrative role of the SMA in movement coordination.

6. ACKNOWLEDGMENTS

We would like to thank to three anonymous reviewers for their suggestions in the early draft, and Feng-pu Yang for his assistance in the experimental study.

7. REFERENCES

[1] Aschersleben, G. Temporal control of movements in sensorimotor synchronization. *Brain and Cognition,* 48, (2002), 66-79.

[2] Balasubamaniam, R. Trajectory formation in timed repetitive movements. In M. L. Latash and F. Lestiennen (Eds.), *Motor control and learning* (pp. 47-64). New York: Springer, 2006.

[3] Brown, S. Biomusicology, and three biological paradoxes about music. *Bulletin of Psychology and the Arts*, 4, (2003), 15–17.

[4] Carson, R. G., Chua, R., Byblow, W. D., Poon, P., and Smethurst, J. Changes in posture alter the attentional demands of voluntary movement. *Proceedings of the Royal Society, Series B- Biological Sciences,* 266, (1999), 853-857.

[5] Chen, Y., Ding, M., and Kelso, J. A. S. Alpha (10 Hz), Beta (20 Hz) and Gamma (40 Hz) networks in the human brain and their functions in a visuomotor coordination task revealed by MEG. *Society for Neuroscience,* 25, (1999), 1893.

[6] Chen, J. L., Zatorre, R. J., and Penhune, V. B. Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythm. *NeuroImage, 32,* 4 (Oct. 2006), 1771-1781.

[7] Cross, I. Music, cognition, culture and evolution. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp. 42–56). New York, NY, USA: Oxford University Press. 2003.

[8] Fitch, W. T., and Rosenfeld, A. J. Perception and production of syncopated rhythms. *Music Perception,* 25(1), (2007), 43-58.

[9] Goebl, W., and Palmer, C. Tactile feedback and timing accuracy in piano performance. *Experimental Brain Research*, 186, (2008), 471-479.

[10] Jantzen, K. J., Fuchs, A., Mayville, J. M., Deeke, L., and Kelso, J. A. S. Neuromagnetic activity in alpha and beta bands reflect learning-induced increases in coordinative stability. *Clinical Neurophysiology,* 112, (2001), 1685-1697.

[11] Kelso, J. A. S., DelColle, J. D., Schöner, G. Action-perception as a pattern formation process. In : Jeannerod M, (Ed.). *Attention and performance XIII.* (pp. 139-169). Hillsdale, NJ: Erlbaum. 1990.

[12] Levitin, D. J., and Tirovolas, A. K., Current advances in the cognitive neuroscience of music. *Annals of the New York Academy of Sciences, 1156* (Mar. 2009), 211-231.

[13] Longuet-Higgins, H. C., and Lee, C. S. The rhythmic interpretation of monophonic music. *Music Perception, 1*(1984) 424-441.

[14] Luck, G., and Sloboda, J. A. Exploring the spatio-temporal properties of the beat in simple conducting gestures using a synchronization task. *Music Perception*, 25, (2008), 225-239.

[15] Mayhew, S. D., Dirckx, S. G., Niazy, R. K., Iannetti, G. D., and Wise, R. G.. EEG signature of auditory activity correlate with simultaneously recorded fMRI responses in humans. *Neuroimage*, doi:10.1016/j.neuroimage.2009.06.080.

[16] Mayville, J. M., Fuchs, A., Ding, M., Cheyne, D., Deeke, L., and Kelso, J. A. S. Event-related changes in neuromagnetic activity associated with syncopation and synchronization timing tasks. *Human Brain Mapping,* 14, (2001), 65-80.

[17] Pfurtscheller, G., Brunner, C., Schlögl, A., and Lopes da Silva, F.H. Mu rhythm (de)synchronization and EEG single-trial classification of different motor imagery tasks. *Neuroimage* 31, (2006), 153-159.

[18] Phillips-Silver, J., Trainor, L.J. Feeling the beat: movement influences infants' rhythm perception. *Science,* 308, (2005), 1430.

[19] Phillip-Silver, J., and Trainor, L.J. Hearing what the body feels: auditory encoding of rhythmic movement. *Cognition*, 105, (2007), 533-546.

[20] Phillips-Silver, J., and Trainor, L. J. Vestibular influence on auditory metrical interpretation. *Brain and Cognition, 67* (Jun. 2008), 94-102.

[21] Raethjen, J., Govindan, R. B., Binder, S., Zeuner, K. E., Deuschl, G., and Stolze, H. Cortical representation of rhythmic foot movements. *Brain Research, 1236* (Oct. 2008), 79-84.

[22] Repp, B. H. Sensorimotor synchronization: a review of the tapping literature. *Psychonomic Billetin and Review*, 12, (2005), 969-992.

[23] Sakai, K., Hikosaka, O., and Nakamura, K. Emergence of rhythm during motor learning. *Trends in cognitive Sciences, 8,* 12 (Dec. 2004), 547-553.

[24] Trainor, L. J., Gao, X., Lei, J., Lehtovaara, K., and Harris, L. R. The primal role of the vestibular system in determining musical rhythm. *Cortex, 45,* 1 (Jan. 2009), 35-43.

[25] Wang, L. J., Wei, J. H., Zhang, D., Dong, W. J., Guo, K. P., and Hu, M. O. The effects of vestibular stimuli on brain cognitive processing: an ERP study. *Neuroreport, 15*, 18 (Dec. 2004), 2757-2760.

[26] Wohlschläger, A., and Koch, R. Synchronization error: An error in time perception. In P. Desain and L. Windsor (Eds.), *Rhythm perception and production* (pp. 115-127). Lisse, The Netherlands: Swets and Zeitlinger.2000.

Chaos and complexity in musical development?

Manuela M. Marin
Department of Musicology & Language Development and Cognitive Science Unit
University of Graz, Austria
Heinrichstrasse 36, A-8010 Graz
0043 316 380 2475
marin@edu.uni-graz.at

ABSTRACT

Nonlinear dynamical systems (NDS) theory has recently been successfully applied to developmental psychology. Although NDS theory has provided fruitful insights into issues related to language development, an NDS approach to musical development has not been pursued so far. Musical enculturation to pitch and rhythmic structures follows a clear developmental trajectory in infancy and early childhood. Issues related to developmental change are largely ignored in these discussions. Here, I will suggest that change in musical development can be better understood by applying principles such as attractors, bifurcations, chaos, self-organization and emergence. The child's brain is regarded as a complex system consisting of many components (motor, sensory, cognition, language and music components, etc.) which are dynamically related to the environment (caretakers, cultural artifacts, etc.). NDS theory may account for developmental change by regarding stages of the developing system as attractor states. Moreover, comparisons to language development reveal that changes in the development of music and language may occur at similar stages during development, indicating that the respective dynamical systems may interact and overlap at various developmental stages.

Keywords

Musical development, nonlinear dynamic systems theory, chaos, complexity, language development

1. INTRODUCTION

Nonlinear dynamic systems (NDS) theory has entered the realm of the physical sciences already 50 years ago, which may be considered as the beginning of a paradigmatic shift in science in general. More recently, the NDS approach has also been successfully – although comparatively scarcely – used to describe and explain phenomena in the fields of neuroscience, psychophysics, sensation, perception, cognition, and developmental, social, organizational and clinical psychology [1]. Surprisingly, musical development has not been discussed within this framework so far, although NDS theory has contributed to a better understanding of first, second and multilingual language development [2-5].

Moreover, NDS theory has been used in the study of motor development [6] and emotional development [7, 8]. This paper suggests to ask new types of questions specifically referring to developmental change during the enculturation to a musical (and linguistic) system by presenting theoretical and methodological approaches that are based on the principles of attractors, bifurcations, chaos, self-organization and emergence.

2. NONLINEAR DYNAMICS AND COMPLEXITY IN DEVELOPMENT

According to Paul van Geert [5], there are currently three approaches to development taken within NDS theory. First, it can be regarded as a theory of embodied and embedded action, largely developed by Esther Thelen and Linda Smith [6]: "The dynamic system at issue is the continuous coupling between the organism and its environment, showing a time evolution that takes the form of intelligent action" [5, p.184]. Second, the qualitative properties of dynamic systems are emphasized, for example, by Marc Lewis [7, 9, 10]. The developmental system is characterized by non-linear behavior, self-organization, emergence and attractor states. Third, van Geert himself is a representative of the position that applies NDS theory to describe and investigate changes in time, focusing on time evolution, rules and functions [5]. Alternatively, one can differentiate between mathematical and methaphorical dynamic systems approaches. In the former, the primary aim is the identification of variables (control parameters) which regulate changes in a system by applying differential equations and/or computer simulations, whereas the latter approach is qualitative and focusses on detailed descriptions of the developmental changes in a system [11]. It can be argued that aspects of these various approaches within the framework of NDS theory can also be applied to the study of musical development.

One of the most challenging tasks in developmental psychology is the description and explanation of change. Researchers are often satisfied with the mere description of developmental stages and specific developments [12], thus the field currently suffers from a lack of theoretical sophistication. This is rather surprising given that Piaget reflected on the basic mechanisms of development many years ago. Specifically, current reviews on musical development do not discuss developmental change within a theoretical framework [13, 14]. An NDS approach could unite the great number of empirical findings of the last decades in developmental psychology in general [15], and in musical development in particular.

Weisstein [16, p.501] defines dynamical systems theory as "a means of describing how one state develops into another state over the course of time," which in mathematical terms can be expressed as

$$y_{t+1} = f(y_t), \tag{1}$$

implying that the next state (at time t +1) is a function (*f*) of the preceeding state (at time t). These recursive relationships stand in strong contrast to the traditional static relationships defined by

$$y_i = f(x_i), \qquad (2)$$

expressing the relationship between a dependent (y) and an independed (x) variable, which for a variable x_i creates a corresponding variable y_i. The important conceptual difference between these two approaches is that "a dynamic model recursively generates a time series (a state and the next state and the next...), [whereas] a static model generates a sample or population of individuals that are in principle independent of one another. [...] Statements about populations do not necessarily apply to the individuals in the population" [15, p. 244-245], which can be regarded as a common homology error in the behavioural sciences. Importantly, van Geert formulates the basic developmental change function as follows: "All changes of the system occur through information that is moderated through the system" [15, p. 250].

A human being or a society are made up of innumerable interacting variables, which "co-determine each other's time evolution" [5, p. 181]. This forms the basis of a system's complexity [17]. Such complex systems have the ability to self-organize, thus dynamic systems can also be defined as "systems that change over time and that can autonomously generate complexity and form" [18, p. xii]. Self-organization can be defined as "irreversible processes in non-linear dynamical systems, leading to more complex structures of the overall system by the cooperative activity of subsystems" [19, p. 118]. Self-organization is flexible and adaptive and a specific form of emergence [5]. Dynamical systems are nested, so every system is part of another system, and they settle in attractor states, which are temporary and can be simple, complex or chaotic [17, 18]. Attractor states can be compared to phases of physical matter, which depend on a single parameter or a confluence of parameters. From a developmental perspective, attractor states can be related to the notion of stage (e.g., level, phase), which reflects the idea of internal coherence and thus certain stability: "Developmental stages form attractor states in that they are represented by habitual, coherent patterns of performance, skill, or action that self-organize spontaneously in the person's habitual context, niches, or living spaces." [15, p.262]. Specifically, developmental stages are defined by a system's major control parameters (cumulative amount of experience) and determined by a system's control parameters (brain maturation). Alternatively, attractor states can be discussed within the context of a critical transition model, in which critical states are those in which transitions may occur due to external influences that lead to major changes in states by the dissipation of stress that has built up in a system [15].

3. METHODOLOGICAL PRINCIPLES

It can be claimed that modern mainstream psychology is mainly interested in linear cause-effect relationships, which does not require dynamic methodology [20]. Which methodologies can be applied to study development from a dynamic systems' perspective? A fundamental methodological assumption is that "developmental research *must* be longitudinal, that is, in order to study change processes" [11, p.442] and variability. Furthermore, the developing child is conceptualized as an open system, which makes the input-output model of the infant-environment relationship redundant: the infant-environment relationship is a complex and non-linear process [11]. This description of a mutual relationship, an ever-ongoing exchange between the child and its environment, has already been described by Piaget by pointing out that adaptation is a two-way process: "adaptation of the child to the environment (accommodation) and adaptation of the environment to the child (assimilation)" [11, p.432]. However, Piaget regarded adaptation as internal to the organism. In contrast to open systems, closed systems can exist without being in an exchange relationship with their environment because they are conceived of as being designed or pre-programmed and context independent. Equally important, developmental analysis largely depends on a detailed description of the day-to-day experiences over time (relational-historical approach), enabling a coverage of the infant's development from a historical perspective which can also be related to real-time events [11]. In other words, every event in a dynamic system has to be seen as a product of the immediate environment and the previous history of the system; subsequent events thus cannot be regarded as being equivalent but are time-dependent. Therefore, a sequence of observations of events in a dynamic system violates the assumption of the independence of the observations, which makes general statistical approaches such as ANOVAs or parametric correlations unacceptable for data analysis [21]. Another important methodological issue concerns the significance of single variables. In general, psychological research measures variables that are assumed to be significant in themselves. One can usually use these variables and their (linear) combination to regress on a dependent variable. Conversely, a dynamic systems approach entails that local interactions of variables are significant and not single variables [21]. Byrne [22] even propagates the notion of the "death of the variable", while others prefer the notion of collective variables, that is variables which refer to the interaction between multiple elements of a system [23]. Larsen-Freeman and Cameron also point out that the nature of explanation differs considerable from conventional approaches: "Complexity theory works at the system level, and explanation is in terms of the system's behavior, not at the level of individual agents or elements (unless, of course, they are the system that is under investigation)" [23, p.201]. In this light, causality has to be re-defined and one type of causality applicable within the context of dynamic systems theory is called co-adaptation, which "describes a kind of mutual causality, in which change in one system leads to change in another system connected to it, and this mutual influencing continues over time" [23, p. 202]. Due to this interconnectedness of sub-systems and their continually changing behaviour over time, variability of a system's behaviour is not regarded as noise or measurement error but as an intrinsic quality of a system [24]. The systematic study of changes in variability (and stability) within a system is thus one of the most challening tasks in the investigation of developmental processes. For this purpose it is essential to conceptually reduce the complexity of a state space to a simple state space in order to be able to observe the qualitative changes [5].

Larsen-Freeman and Cameron [23] proposed several modified research methodologies for studies on language development based on the issues addressed above. These methodologies could also be easily applied to research on musical development and include

- Ethnography
- Formative and design experiments/action research
- Longitudinal, case studies, time-series approaches
- Microdevelopment
- Computer modelling

- Brain imaging, and
- Combinations of methodologies

How can non-independent data be analysed? A common approach in the context of dynamic systems theory is to describe and model the temporal trajectory underlying the dynamic process rather than to compare the system's states in given moments. The process is thus considered as a whole and insights are, for example, gained from meta-trend analyses [21]. Growth mixture models have recently become popular in order to compare the temporal evolution of single cases. In spite of the fact that growth mixture models are able to capture the temporal dynamics of a systems, they are still based on the conventional assumption of linearity, which makes it unsuitable to describe phenomena of dissipative dynamics and emergence of order [25]. Instead, sets of differential equations and the investigation of the dimensionality of the phase space are commonly used to analyse processes in dynamical systems. For instance, logistic growth models, a model of growth under limited resources, have been successfully applied to second language development [5].

4. UNDERSTANDING MUSICAL DEVELOPMENT

Musical development is generally discussed from two perspectives, namely enculturation and formal training [13]. Here, I am focusing on the former and suggesting that the acquisition of a musical system's spectral (pitch) and temporal (rhythmic) structures can be better understood by applying the principles ad methodologies of NDS theory. Like in language development, enculturation proceeds from universal to system-specific processes. Current research has primarily investigated when and how infants acquire or lose certain abilities on their way to achieve implicit knowledge about their musical system by applying behavioural and neuroscienctific methods without considering questions related to developmental change [13, 14, 26, 27]. The perception of musical pitch structure is learned in several stages, probably based on a universal sensitivity to consonance which enables the learning of scales and later, in the case of Western music, the learning of harmonic relations. Perception of harmony can be regarded as the highest level of enculturation to pitch structure and may be already present at the age of 3-4 years [28]. Rhythm perception precedes meter perception in infancy. However, it is still unclear whether temporal and spectral properties of music develop in parallel [13].

By drawing on current findings from other developmental research, I argue that musical development can be meaningfully discussed in the context of dynamic systems theory. For instance, longitudinal studies and computer simulations could help to develop musical growth functions and explain phase transitions in a musical complex system with developmental dynamics. In principle, research could focus on any aspect of the development of music perception and production and could follow any methodological approach in accordance with NDS theory. In other words, there are many ways by which NDS theory could be applied to the investigation of musical development from infancy up to early childhood. I suggest that it is up to the researcher to choose a specific research question and suitable methodologies, depending on the age of the children. As a first step, repeated-measure designs with a focus on the analysis of changes in variability could provide insights into the changing dynamics of musical development. In such an endeavour, special emphasis should be placed on stages of development that have already been identified as crucial in conventional studies (e.g., loss of sensitivity to specific features of foreign musical systems).

Comparisons between music and language development [29, 30] can be specifically addressed in the discussion of the emergence of musical and linguistic syntactic abilities [31]. In adults, neuroscientific and behavioural findings have shown that there is an overlap in the processing of musical and linguistic syntax [32-34]; therefore, it can be assumed that the development of shared neural processing resources may lead to similar stages and transitions during early development.

5. CONCLUSION

In order to explain developmental change, NDS theory has been successfully applied in the general field of developmental psychology, and especially in language development. By building on these theoretical and methodological approaches, I have suggested to discuss and explore musical development under this theoretical framework. Moreover, music and language share many features and neural resources, which leads to the conclusion that changes in enculturation to a musical and linguistic system may best be understood under a common framework of NDS theory by longitudinal studies that screen for linguistic and musical changes in perception, production and neural development.

6. ACKNOWLEDGEMENTS

I wish to thank Annemarie Peltzer-Karpf and Bruno Gingras for their helpful comments.

7. REFERENCES

[1] Guastello, S. J., Koopmans, M. and Pincus, D. *Chaos and complexity in psychology*. Cambridge University Press, Cambridge, 2009.

[2] Hohenberger, A. and Peltzer-Karpf, A. Language learning from the perspective of non-linear dynamical systems. *Linguistics*, 47, 2, 481-511.

[3] Larsen-Freeman, D. On the complementarity of Chaos/Complexity Theory and Dynamic Systems Theory in understanding the second language acquisition process. *Bilingualism: Language and Cognition*, 10, 1, 35-37.

[4] Mohanan, K. P. *Emergence of complexity in phonological development*. York Press, Inc., York, 1992.

[5] van Geert, P. The Dynamic Systems Approach in the Study of L1 and L2 Acquisition: An Introduction. *The Modern Language Journal*, 92, 2, 179-199.

[6] Smith, L. B. and Thelen, E. Development as a dynamic system. *Trends in Cognitive Sciences*, 7, 8, 343-348.

[7] Lewis, M. D. Bridging emotion theory and neurobiology through dynamic systems modeling. *Behavioral and Brain Sciences*, 28, 2, 169-194.

[8] Camras, L. A. and Witherington, D. C. Dynamical systems approaches to emotional development. *Developmental Review*, 25, 3-4, 328-350.

[9] Lewis, M. D. The promise of dynamic systems approaches for an integrated account of human development. *Child Development*, 71, 1, 36-43.

[10] Lewis, M. D. and Granic, I. *Emotion, development and self-organization: Dynamic systems approaches to emotional development*. Cambridge University Press, Cambridge, 2000.

[11] Vedeler, D. and Garvey, A. P. *Dynamic methodology in infancy research*. Springer, Dordrecht et al., 2009.

[12] Van Geert, P. A dynamic systems model of basic developmental mechanisms: Piaget, Vygotsky, and beyond. *Psychological Review*, 105, 634-677.

[13] Hannon, E. E. and Trainor, L. J. Music acquisition: effects of enculturation and formal training on development. *Trends in Cognitive Sciences*, 11, 11, 466-472.

[14] Trehub, S. E. and Hannon, E. E. Infant music perception: Domain-general or domain-specific mechanisms? *Cognition*, 100, 1, 73-99.

[15] Van Geert, P. *Nonlinear complex dynamical systems in developmental psychology*. Cambridge University Press, Cambridge, 2009.

[16] Weisstein, E. W. *CRC concise encyclopedia of mathematics*. CRC Press, Boca Raton, 1999.

[17] De Bot, K., Lowie, W. and Verspoor, M. A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and Cognition*, 10, 1, 7-21.

[18] Smith, L. B. and Thelen, E. *A dynamic systems approach to development: Applications*. MIT Press, Cambridge MA, 1993.

[19] Ebeling, W. *Chaos - Ordnung - Information: Selbstorganisation in Natur und Technik*. Harry Deutsch, Frankfurt Main, 1991.

[20] Toomela, A. *How methodology became a toolbox - and how it escapes from that box*. Springer, Dorndrecht et al., 2009.

[21] Salvatore, S., Lauro-Grotto, R., Gennaro, A. and Gelo, O. *Grasping the dynamic nature of intersubjectivity*. Springer, Dordrecht et al., 2009.

[22] Byrne, D. *Interpreting quantitative data*. Sage, London, 2002.

[23] Larsen-Freeman, D. and Cameron, L. Research Methodology on Language Development from a Complex Systems Perspective. *The Modern Language Journal*, 92, 2, 200-213.

[24] van Geert, P. and van Dijk, M. Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior and Development*, 25, 4, 340-374.

[25] Laurenceau, J. P., Hayes, A. M. and Feldman, G. C. Statistical and methodological issues in the study of change in psychotherapy. *Clinical Psychology Review*, 27, 715-723.

[26] Trainor, L. J. and Trehub, S. E. A Comparison of Infants' and Adults' Sensitivity to Western Musical Structure. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 2, 394-402.

[27] Hannon, E. E. and Trehub, S. E. Tuning in to musical rhythms: Infants learn more readily than adults. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 35, 12639-12643.

[28] Marin, M. M. Effects of early musical training on musical and linguistic syntactic abilities. *Annals of the New York Academy of Sciences*, 1169, 187-190.

[29] Jackendoff, R. Parallels and nonparallels between music and language. *Music Perception*, 26, 3, 195-204.

[30] McMullen, E. and Saffran, J. R. Music and Language: A Developmental Comparison. *Music Perception*, 21, 3, 289-311.

[31] Patel, A. D. *Music, language, syntax and the brain*. Oxford University Press, Oxford, 2008.

[32] Koelsch, S. Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, 15, 2, 207-212.

[33] Slevc, L. R., Rosenberg, J. C. and Patel, A. D. Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin and Review*, 16, 2, 374-381.

[34] Fedorenko, E., Patel, A., Casasanto, D., Winawer, J. and Gibson, E. Structural integration in language and music: Evidence for a shared system. *Memory & Cognition*, 37, 1, 1-9.

Music's Golden Section – An Empirical Approach

Michelle Phillips
Centre for Music and Science, University of Cambridge,
West Road, Cambridge, UK
Telephone number: 0044 (0)7811 971869
mep41@cam.ac.uk

ABSTRACT

The golden section has attracted increasing attention from music analysts since the 1970s, owing largely to Ernö Lendvai and Roy Howat and their work on the music of Bartók and Debussy respectively. However, of the 120+ such studies to date, not one attempts to support any generic claim that the work's aesthetic attractiveness is enhanced by this arithmetic proportion *empirically*, i.e. no experimental investigation as to whether we can *hear* the golden section, whether consciously or non-consciously, has ever been undertaken.

Such experimental work must necessarily incorporate recent research into perception of large-scale musical form and musical time, and also findings and methods from relevant empirical investigations in psychology - a field which has seen extensive scientific occupation with this proportion since Gustav Fechner's experiments with golden section rectangles in 1876.

This paper will outline the relevant background and the surrounding debate, as well as the considerations and methods appropriate for empirical work on musical time and the aural salience of the golden section.

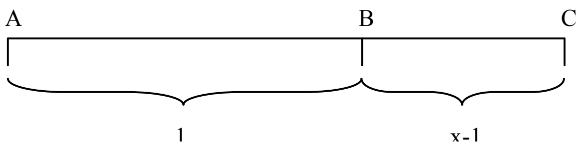
Keywords

Golden section, musical time, large-scale form, music and proportion, music and mathematics.

1. INTRODUCTION

Finding structures in notation tells us very little about how minds experience the music. – Huron, ESCOM 2009

The golden section is obtained by dividing a whole into two parts such that the ratio of the smaller to the larger section is equal to that of the larger to the whole (BC:AB = AB:AC):



Solving the resulting quadratic equation ($x^2 - x - 1 = 0$) gives a ratio of 1:1.618 to three decimal places, a ratio termed the 'golden section'.

Also known as the golden mean, ratio, number and divine proportion, and arithmetically linked to the Fibonacci series, the golden section possesses intriguing mathematical properties. Yet its popularity extends beyond mathematics: the golden section is represented in nature (for example, in seed formation in pine cones and sunflowers); the field of experimental psychology continues to critically examine and build on Gustav Fechner's [21] work with 'golden rectangles' and to address the question of whether the golden section is indeed an artefact of visual aesthetic preference; social psychologists claim human judgments of positive and negative attributes in others conform to this ratio; and dentists are taught the aesthetic appeal of teeth proportioned according to the golden section during training. Popular culture also devotes attention to the concept: the social networking website *Facebook* includes the groups 'oh beautiful 1.618, please don't ever terminate' and 'I Love The Golden Ratio and Fibonacci Numbers'; and the online micro-blogging service *Twitter* includes members using the identifiers 'GoldenRatio' '1x62 Golden Ratio', 'Fib_numbers' and 'Sezione_Aurea', amongst many others bearing relation to the golden section.

Over the last 30 years the golden section has received increasing attention in the field of music analysis, encompassing music from Gregorian chant to Madonna. This popularity is largely owing to Ernö Lendvai's [34] work on the music of Béla Bartók, first published in English in the 1970s, and Roy Howat's [29] subsequent monograph, *Debussy in Proportion.*

There are more than 120 separate examples of music analysis which incorporate the golden section, all with one surprising uniting factor – not one of these studies tackles the golden section in music *empirically*. That is, although these studies claim to identify the proportion within the musical notation, and often make generic claims regarding the aesthetic value lent to the work by the manifestation of this ratio, not one demonstrates that this ratio may be *heard*. Indeed, few consider this potential perceptual obstacle to such conclusions. It is hence not known whether the golden section is aurally perceptible, nor has any study yet attempted to address this.

Such an experimental approach requires consideration of multiple complex issues. Fundamentally, can empirical approaches ever sufficiently recreate and address aesthetic behaviour? The aesthetic experience is one which may not necessarily be reduced to a set of principles or induced second hand [16], and thus empirical approaches to this question are difficult at best, and perhaps impossible '...since a subject has agreed to [do] the experiment he feels obliged to do something; this doing need not represent aesthetic behaviour' [38, p. 510].

Although much empirical work has been performed on the perception of local musical characteristics (metre, phrase), perception of large-scale form (entire sections, or pieces) is little studied, and consequently little understood. Empirical work undertaken to date in this area has suggested that musical form may not in fact be perceived globally at all, but rather only on the local level [43], and that altering a work's tonal structure may go unnoticed by trained and untrained musicians alike [13]. Other musical features may take priority in perception over global structure: for example, a study by Deliège *et al*. [17] finds weak sensitivity to large-scale form compared to cues perceived on the music's surface. However, studies have also indicated some perception of large-scale form by musicians and non-musicians [33, 5], finding that listeners demonstrated highly veridical judgements of location of a musical extract within the whole, and that there was high inter-subject agreement about the features marking segment boundaries [12].

Study of perception of large-scale form also necessitates familiarity with one relatively new and as yet under-explored area of research; the experience of musical time beyond the perceptual present. Compared to the work as experienced during music analysis, the listener does not perceive the musical bars or sections in objective and absolute relation to one another. Rather, as extensive studies of psychological time have demonstrated, experience of duration depends on multiple complex factors, such as context of the event, level of arousal, body temperature, preference and method of timing estimation [8, 25, 27]; in other words, a watched kettle does take longer to boil, time does fly when you're having fun, and the years do pass more quickly as you age. Recent empirical work in music psychology has begun to explore theories of psychological time, and has demonstrated that musical features (such as harmonic modulation, tempo variation, or tonality) may alter the listener's perception of elapsed duration of a section of music [3, 21, 22, 31, 39]. Hence the question of whether the golden section can be perceived aurally, in addressing the listener's experience of length of musical sections in relation to one another, necessitates an understanding of the notion of musical time.

The present paper will discuss the concept of musical time, i.e. how and to what extent our experience of elapsed duration during musical engagement may be shaped by musical factors, and outline studies in this area to date. An overview will then be provided of empirical research to come and methodology to be adopted. Initial results are expected at the end of October 2009.

2. BACKGROUND

2.1 The golden section

Music analysis may apply the golden section to music either in relation to musical *pitch* (the 'vertical' golden section), or *duration* (the 'horizontal' golden section). The former involves analysis of pitch intervals in accordance with the Fibonacci sequence, and the latter, the subject of the current research, raises questions of perception of large-scale form and musical time. This is the more prevalent form of golden section analysis, whereby the analyst most commonly identifies a significant point in the work and claims to have discovered, by counting bars, beats, or seconds, that this point divides the work in the ratio of the golden section.

Music analysis may adduce that the ratio has been used intentionally as a compositional tool, in order to structure the work, or that the golden section is manifested regardless of compositional intention (i.e. the ratio occurs intuitively to the composer), and the work in question may owe part or all of its effect, aesthetic or otherwise, to this quality.

The ratio is indisputably intentionally employed by many composers as a way to structure rhythms, motifs, durations and significant points in a composition, or as an initial model to order compositional ideas, which is then often deviated from as the music is developed. No evidence exists that composers were aware of the ratio pre-20th century (indeed there is no evidence Bartók was aware of it), but there are multiple examples of composers stating their intentional use after this and into the 21st century. Post-1900, and largely post-Lendvai and Howat's landmark studies, this technique was widely employed (composers include Hindemith, Xenakis, Schoenberg, Ligeti, Stockhausen and Nono), and continues to be used extensively today (for example by composers such as Murail, Gubaidulina, Tutschku, Knussen, Maxwell Davies, Holliger, Wuorinen, Sáry, Weir and Turnage). Nevertheless, claims by composers that they have successfully manipulated the listener's aural experience as the music unfolds in time remain unsubstantiated, though they may be informative about the composer's own musical conceptions.

Regardless of the question of compositional intention, music analysis has adopted the golden section as one of its tools. Of the 127 examples currently in existence, over a third relate the concept to pre-20th century works (deduced through systematic analysis of these studies). That is, the golden section is thought to have been employed in composition practise *non-consciously*. Moreover, existing golden section writing in relation to music often claims that such proportioning contributes to the music's effect on the listener: e.g., Evans [20] claims that 'the subtle qualities of dynamic balance attributed to φ [phi] contribute to the work's [Mozart's *Symphony in G Minor*] beauty and popularity' (p. 307). Such studies often fail to question the assumption that arithmetically-deduced relations within the musical score are perceived by the listener. Indeed, Howat [29] himself does not avoid this pitfall, claiming in his study that Debussy's proportional systems 'show ways in which the forms are used to project the music's dramatic and expressive qualities with maximum precision' (p. 1).

Such theorising regarding the ratio's apparently universal value has become irrationally commonplace, given that the question of whether the golden section is audible has not been sufficiently addressed, or experimentally explored. To borrow Cook's [14] terminology, the music analyst's search for the ratio appears frequently to be 'a symptom of analytical desperation' (p. 166). Such desperation bring with it a risk of confirmation bias: 'there are so many ways of taking your sections that I doubt whether any musical composition can avoid golden ones somewhere' [44, p. 19]; it is notable that only one extant study claims to have looked for golden section relations, but found none [4].

The assumption that the golden section may be aurally perceived could be thought to stem from other areas of research; outside of the field of music, the golden section has received extensive treatment in terms of *visual* salience, and there is some evidence (although much disputed) for its visual perceptibility [42]. Empirical investigations have also claimed to have evidenced haptic perception of the golden section [28].

Aural perception, however, may not be governed by the same features as those involved in visual and haptic processing. In particular, the aural experience unfolds in time, rather than being available in its entirety in one moment (as is the case for visual art). Hence an understanding of experience of time and duration is integral to the notion of aural perception.

2.2 Psychological time

'Time' may be thought of as the seconds, minutes, days and years by which our routines and lives are regulated. This is clock time, measured according to the radioactive decay of the caesium atom in the atomic clock. However, this objective form of measuring and controlling events in time is not representative of the form in which humans experience such events. Rather, psychological time is pliable and subjective, and the speed at which we experience events varies according to multiple factors, including context, level of arousal, preference, pharmacological influence, body temperature and engagement in the task at hand.

The study of psychological time is not a new area of interest; indeed, William James recognised the importance of this distinction between clock time and psychological time in *The Principles of Psychology*, questioning the nature of humans' 'special sense for pure time' [18, p. 1944]. James also recognised that any investigation of psychological time will rely on one of two timing paradigms; the prospective paradigm (or 'experienced duration'), in which the subject is aware that an estimation of elapsed duration will be required; and the retrospective paradigm (or 'remembered duration'), in which the subject is unaware that such judgement will be required. The former relies largely on levels of attention, as the more attention is paid to the passing of time rather than the task at hand, the longer will be the estimate of elapsed duration. The latter relies heavily on memory resources, as the subject must recall the sequence of events experienced and gauge time elapsed.

There are two main models which attempt to account for retrospective time estimation. The first, Ornstein's [40] storage-size hypothesis, suggests that judgement of elapsed duration depends on the capacity in memory to store the events in the time frame. The second, and currently more widely accepted, is Block's [7] contextual change hypothesis, which proposes that remembered duration relies on number of changes in the timed interval, and the more changes recalled, the longer the estimate.

It is the prospective condition which has received greater empirical attention, as this allows a larger amount of experimental data to be collected (the retrospective task may only be undertaken once per subject, as they are then aware that the task involves timing judgements). Broadly, the study of psychological time in the prospective condition takes one of two approaches. The first and earliest has its origins in Gibbon *et al*.'s [24] internal clock (or pacemaker) model of temporal processing. This model is based on the theory of a pacemaker that emits pulses during a timed event and an accumulator which stores the pulses in working memory and compares these to pulse counts stored in reference memory once the event is over. The internal clock model has undergone significant development over recent decades. In particular, level of arousal (a state of increased physiological activation associated with emotion) has been accounted for (arousal is thought to speed up the rate of pulses emitted, and therefore increase estimates of elapsed duration), and a gate has been added which allows pulses through to the accumulator when attention is paid to time rather than the salient event (which results in more pulses being stored, and thus a longer estimation of elapsed duration). Hence this model is also referred to as one of 'time-sharing' or 'resource allocation', or as the 'stopwatch paradigm'. These developments culminated in the 'attentional-gate' design which forms the basis of current models [47]. One very recent addition to this is offered by Buhusi and Meck [10], who propose that a 'context' component be added to the attentional-gate model, incorporating attention and memory resources. Such a component accounts for the finding that, not only has the level of attention diverted from counting pulses been shown to shorten the experience of elapsed duration, but the decay of these pulses in working memory also appears to contribute to this effect. This model offers a promising platform for further internal clock designs, particularly with its focus on memory resources, a component previously – and surprisingly – given little consideration.

No neural basis for this internal clock model has yet been found. Rather, a second approach to the area of research has emerged over the last decade or so, and this is a model of timing based on neural networks. Such a model assumes that, rather than pulses emitted by an internal pacemaker, neural oscillations may form the basis of timing mechanisms. Additionally, convincing evidence has recently emerged that different time scales (i.e. hundredths of milliseconds vs. seconds) may engage different neural processes. Thus, timing processes are thought to be 'duration-dependent' [46]. Although there is some agreement that mechanisms do indeed differ at the sub- compared to the supra-second level, consensus on further, finer distinction has not been reached. Koch [32] refers to Mauk and Buonomano's four categories of duration [as cited on 32]: microseconds, milliseconds, seconds, and circadian rhythms (p. 1908); and Pöppel [41] argues for the 'perceptual present' as being cross-culturally true at 2-3 seconds. However, it appears such finite distinction may not be appropriate or accurate, as Lewis and Miall's [37] failed empirical attempt to locate the 'break points' between one process and the next demonstrates.

From this duration-dependent line of study, new and adapted models have emerged to accommodate and explore differing time frames. For example, Buonomano *et al.*'s [11] state-dependent network (SDN), relating to durations at the sub-second level, proposes that temporal events are encoded in the context of the previous neural event (i.e. temporal processing is mapped within the existing state of the neural network) and finds that interval discrimination is impaired in the processing of short intervals (100ms compared to 300ms). Another model which posits neural oscillations as the basis of temporal processing is Eagleman and Pariyadath's [19] attempt to account for duration encoding relating to sub-second time scales, in which the more energy required to neurologically encode a stimulus, the longer the perceived duration. This is not the only study to propose such a relation. Bueti and Walsh's [9] 'a theory of magnitude' (ATOM) suggests that the experience of magnitude of time is processed in the parietal cortex, along with all other magnitudes, such as size, space, number and speed. Support given for such a theory includes behavioural evidence that magnitude judgments share resources, for example, increased intensity in a value often increases experienced duration [48].

Studies using brain imaging, timing tasks, and subjects with brain injury have shown that multiple brain structures are involved in temporal processing; thus there appears to be no one temporal mechanism for the processing of time, but rather timing seems to involve multiple brain areas and neural networks, perhaps depending on the duration timed and the context of the timing task [32]. Areas thought to be most salient include the parietal cortex, the prefrontal cortex, the supplementary motor area, the cerebellum and the basal ganglia.

The study of the brain areas involved in temporal processing has also been influenced by consideration of the effect of *emotion* on timing. Although the attentional-gate model of timing acknowledges a potential influence of levels of arousal on the experience of time, that influence remains under-specified within the model; the extent of the impact of emotion on duration experience has only recently begun to receive serious attention. Craig's [15] model begins with the premise that the anterior insular cortex (AIC) is involved with subjective awareness and processing emotionally salient moments in the brain, and this series of moments allow a view of the sentient self, in each immediate 'now'. Therefore (and here supported by fMRI studies), the AIC is a crucial locus in the human brain for the perception of time. Craig proposes that it is these moments which are then used to track the passing of time, rather than the pulses of the internal clock – hence experience of time is dependent on emotional state. Craig also touches on music in this regard: 'One especially fascinating aspect is that this structure can provide an emergent basis for the uniquely human faculty of music. That is, if music is viewed as the rhythmic temporal progression of emotionally laden moments, which this model directly instantiates, then this model provides a ready basis for the neural representation of music as a coherent emotional construct' (p. 1937). An emotionally-based theory of temporal processing is also explored and proposed by Droit-Volet and Gil [18], who demonstrated that subjects recalling whether faces shown to them were in a 'long' or 'short' duration condition overestimated the duration of the faces demonstrating anger and fear, i.e. those emotions that were considered most arousing. This finding is discussed as supporting the notion that, rather than the internal clock becoming *distorted* during arousing experiences, the clock is fine-tuned to adapt to contextual events in the environment.

To summarise these recent theories: there is some agreement that an internal clock-like mechanism could exist for the timing of seconds, but smaller durations are currently discussed largely in terms of neural bases of timing. However, it is clear that there are multiple and conflicting theories and models of temporal processing, and currently no single one seems to be unambiguously supported by the available evidence. Furthermore, the distinctions previously proposed (atomic vs. psychological time) may prove to constitute over-simplifications; for example, van Wassenhove's [45] discussion of the multisensory processing of time (i.e. including vision) concludes with the suggestion that psychological time may not necessarily be the same as neural time, and thus a new category of timing may be required.

It is also apparent in light of recent research regarding psychological time, particularly empirical work undertaken during 2009, that experience of elapsed duration is a complex and multi-faceted process. The timing of an event is governed by a broad range of factors, both extrinsic and intrinsic to the human body, which are currently little understood and have only recently begun to receive due attention in empirical investigation. Whether an internal clock, contextual change, or neural network-based model underlies temporal processing appears to vary from event to event, depending on multiple factors including arousal and level of attention.

2.3 Musical time

One popular and cross-cultural activity, for which emotional response and attentional demand are highly relevant, is that of music listening. Attention to the nature of the experience of musical time over large timescales is relatively new and as yet under-explored; however, it is vital to any consideration of a potentially aurally detectable proportional relation. The question of aural perception of the golden section relies on an understanding of the concept of elapsed duration during music listening, rather than on score analysis, as has previously been the case; 'the psychology of performing and listening is not found in strict interpretations of musical scores, but rather in understanding how activities change over time' [30, p. 218].

It has been asserted that the golden section is 'audible', with claims that either such ratios have been uncovered through music listening rather than score analysis [1], or that the analytically defined ratios are audible [2]. Notably, Lendvai himself claims his interest in golden section proportions in Bartók's music began during a concert performance which compelled him to count the beats during listening, and claim in a later publication: 'The sine qua non of golden section is sensation (perception with the aid of the senses) [...] It fulfils its task only if it can be perceived' [35, p. 256-7]. Despite such claims, little account has been taken in golden section analysis of the notion of elapsed duration and experience of musical time.

So what is 'musical time' and could it be represented by existing models of psychological time as described above? Mari Riess Jones [30] refers to experienced time during everyday events as 'garden variety time' (as opposed to 'high church time', which represents time *abstracted* from these everyday events). Musical events, as one form of everyday event, present new challenges to current models of psychological time, as 'time in music is difficult to tease apart from the structure of the event itself [...] In short, temporal production and estimation have special meanings in musical contexts' (p. 213-4). Jones discusses the inadequacies in internal clock models for representing music time, and calls for new models to properly represent experience of music events: 'Somewhere between convention established in the two disciplines of psychology and music is a middle ground that may permit new conceptions of psychological time as it figures in music' (p. 214).

Jones' [30] model incorporates 'the performer as an artistic *interpreter* of music events' (p. 229), who manipulates time to 'achieve successful communication of musical ideas by intentionally violating certain ratio-based time norms [as] time proportionalities have an intriguing function in musical communication' (p. 222). This model provides a valuable starting point for the development of a representation of psychological time which reflects music listening.

Empirical studies of music's effect on perception of elapsed duration are few, but increasingly common, and are beginning to explore and gauge the relevance of existing models of timing to musical events. This area has wide-reaching social implications; if music is deemed to influence experience of time, its use or application in time-sensitive settings could be strategically controlled and manipulated. For example, waiting time in an airport or hospital, duration in a queue on the telephone, or time experienced in a retail outlet, could all be influenced by the acoustic environment.

Forde Thompson *et al.* [23] tackled the notion of musical time empirically in their short article in *Canadian Acoustics* 'Musical influences on the perception of time', and found subjects' judgements of the duration of each musical excerpt retrospectively to be highly veridical. However, other studies have demonstrated that elapsed duration and time in music may be contingent on complex musical qualities.

Kellaris and Kent [31] found that estimates of elapsed duration varied significantly from clock time during music listening, and, moreover, musical modality influenced subjects' experience of time. The study employed a bespoke composition in a major, minor, and atonal condition, and asked subjects for their estimations of elapsed duration, finding that: 'Music pitched in a major key produced the longest duration estimates and the greatest disparity between actual (i.e., clock) time and perceived time. Music pitched in a minor key produced a significantly shorter average duration estimate. Atonal music produced the shortest, and most accurate, average duration estimates' (p. 373). Also notable is that the piece which elicited the greatest affective response, the major condition, did not result in the shortest estimate, i.e. time did not "fly" during this arousing piece. These results are in line with those discussed above (see 'Timing and Emotion'), such as the finding in Droit-Volet and Gil [18] that arousing experiences lead to overestimations of elapsed duration.

The effects of music on the judgement of elapsed duration in a casino setting formed the basis of a study by Noseworthy and Finlay [39]. Their findings demonstrate that estimates of elapsed duration are found to be higher and more accurate when the auditory environment includes music (along with the ambient casino sound) than without, particularly so when the music was of a slow tempo and was played relatively loudly. The latter finding is discussed by the authors in light of the greater attention demanded by music at a higher volume.

Familiarity of musical stimuli may also have an effect on experience of time. Bailey and Areni [3] tested estimates of elapsed duration with familiar and unfamiliar music, with or without a memory task, and found those waiting idly estimated shorter durations during listening to familiar (rather than unfamiliar) music, and those performing a memory task thought a longer duration had elapsed in the familiar music condition.

The most recent study of musical time, by Firmino *et al*. [22], required participants to listen to three bespoke chord sequences of 30 seconds each, which modulated from the key of C major to one each of G major, F major, or G flat major and back to the home key, and estimate the elapsed duration of each. The investigators found that, significantly, participants judged the extract which modulated to the more distant key (G flat major) to have the shortest elapsed duration, and concluded that 'This finding demonstrates that tonal modulation is a relevant parameter that affects retrospective time estimation in music' (p. 207).

In summary, these studies demonstrate that features of music and musical structure *do* appear to influence the experience of elapsed time. Moreover, the extent and nature of this influence varies depending on multiple factors, including the volume, tempo, modality, familiarity and harmonic variation of the music.

3. HYPOTHESIS

Although studies conducted to date have focussed on experience of elapsed duration of an entire musical work, extract, or stimulus, there is clearly scope for exploring experience of time *within* sections of a musical whole. For example, a work which begins in a major key and modulates to the relative minor at its midpoint would perhaps not be experienced as having two sections of equal length, but rather the judgement of relative duration of one section to the other would be influenced by the tempo of each.

The relevance of musical time and studies of large-scale musical form to the golden section debate in music is clear. Given the complexity and variety of factors which appear to influence subjective time during music listening, the claim that golden section proportioning identified through music analysis may be aurally perceived and hence contribute to the aesthetic effect of the work seems at best highly dubious and at worst fundamentally inaccurate.

It is predicted that golden section divisions, which may be salient in the musical score, are not audible by the listener. Rather, experience of duration of musical works and sections is governed by a variety of complex factors, which have only recently begun to be revealed by empirical investigation.

4. METHOD

Empirical work will begin with a task involving the prospective estimation of duration of six musical extracts, each in one of either a major, minor, or atonal condition, and in a familiar or unfamiliar condition. Data will be collected from the general public in attendance at the Centre for Music and Science 'open day' on 24th October 2009 as part of the Cambridge Festival of Ideas.

5. INITIAL FINDINGS

There are no initial findings at date of publication. It is hoped that these will be available for presentation by November 2009.

6. ACKNOWLEDGMENTS

I would like to extend my sincere thanks to Dr Ian Cross and colleagues in the Centre for Music and Science for their invaluable advice and assistance.

7. REFERENCES

[1] Anderson, M. First Performances - London, Royal Albert Hall: Proms 2006, Benjamin, Rihm, Brahms-Glanert, Elgar-Payne. *Tempo, 61, 239* (2007), 53-55.

[2] Atlas, A. W. Gematria, Marriage Numbers, and Golden Sections in Dufay's "Resvellies vous". *Acta Musicologica, 59, 2* (1987), 111-126.

[3] Bailey, N., & Areni, C. S. When a few minutes sound like a lifetime: Does atmospheric music expand or contract perceived time? *Journal of Retailing, 82*, *3* (2006), 189-202.

[4] Barazzoni, B. Kurtág's Music: The Spectacle of Nature, the Breath of History, from op. 7 to 27. *Studia Musicologica Academiae Scientiarum Hungaricae, 43*, *3/4* (2002), 253-267.

[5] Bigand, E. *More about the musical expertise of musically untrained listeners.* Paper presented at the Conference on Neurosciences and Music: Mutual Internactions and Implications on Developmental Functions, Venice, Italy (2002, Oct 25-27).

[6] Bigand, E., & Poulin-Charronnat, B. Are we "experienced listeners"? A review of the musical capacities that do not depend on formal musical training. *Cognition, 100, 1* (2006), 100-130.

[7] Block, R. A. Models of Psychological Time. In R. A. Block (Ed.), *Cognitive Models of Psychological Time* (pp. 1-35). Hillsdale, New Jersey: Lawrence Erlbaum Associates (1990).

[8] Block, R. A., & Zakay, D. Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review, 4, 2* (1997), 184-197.

[9] Bueti, D., & Walsh, V. The parietal cortex and the representation of time, space, number and other magnitudes. *Phil. Trans. R. Soc. B*, *364, 1525* (2009), 1831-1840.

[10] Buhusi, C. V., & Meck, W. H. Relative time sharing: new findings and an extension of the resource allocation model of temporal processing. *Phil. Trans. R. Soc. B, 364, 1525* (2009), 1875-1885.

[11] Buonomano, D. V., Bramen, J., & Khodadadifar, M. Influence of the interstimulus interval on temporal processing and learning: testing the state-dependent network model. *Phil. Trans. R. Soc. B, 364, 1525* (2009), 1865-1873.

[12] Clarke, E. F., and Krumhansl, C. L. Perceiving Musical Time. *Music Perception, 7, 3* (1990), 213-52.

[13] Cook, N. The Perception of Large-Scale Tonal Closure. *Music Perception, 5* (1987), 197-206.

[14] Cook, N. *Analysing Musical Multimedia*. Clarendon Press, Oxford, 1998.

[15] Craig, A. D. Emotional moments across time: a possible neural basis for time perception in the anterior insula. *Phil. Trans. R. Soc. B, 364, 1525* (2009), 1933-1942.

[16] Cross, I., & Tolbert, E. Music and Meaning. In S. Hallam, I. Cross & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology*. New York: Oxford University Press, 2009.

[17] Deliège, I., Melen, M., Stammers, D., & Cross, I. Musical schemata in real-time listening to a piece of music. *Music Perception, 14, 2* (1996), 117-159.

[18] Droit-Volet, S., & Gil, S. The time-emotion paradox. *Phil. Trans. R. Soc. B, 364, 1525* (2009), 1943-1953.

[19] Eagleman, D. M., & Pariyadath, V. Is subjective duration a signature of coding efficiency? *Phil. Trans. R. Soc. B, 364, 1525* (2009), 1841-1851.

[20] Evans, B. Number as Form and Content: A Composer's Path of Inquiry. *Leonardo, 25* (1992), 303-311.

[21] Fechner, G. T. Various Attempts to Establish a Basic Form of Beauty: Experimental Aesthetics, Golden Section, and Square. *Empirical Studies of the Arts, 15, 2* (1997, translated from 1876), 115-130.

[22] Firmino, É. A., Bueno, J. L. O. et al. Travelling Through Pitch Space Speeds up Musical Time. *Music Perception, 26, 3* (2009), 205-209.

[23] Forde Thompson, W., Russo, F. A. et al. Musical influences on the perception of time. *Canadian Acoustics 21, 3* (1993), 83-84.

[24] Gibbon, J., Church, R. M., & Meck, W. H. Scalar Timing in Memory. *Annals of the New York Academy of Sciences, 423* (1984), 52-77.

[25] Gil, S., Rousset, S., & Droit-Volet, S. How Liked and Disliked Foods Affect Time Perception. *Emotion, 9, 4* (2009), 457-463.

[26] Green, C. D. All That Glitters - A Review of Psychological-Research on the Aesthetics of the Golden Section. *Perception 24, 8* (1995), 937-968.

[27] Gruber, R. P., & Block, R. A. Effect of caffeine on prospective and retrospective duration judgements. *Human Psychopharmacology-Clinical and Experimental, 18, 5* (2003), 351-359.

[28] Hintz, J. M., & Nelson, T. M. Haptic aesthetic value of the golden section. *British Journal of Psychology, 62, 2* (1971), 217-223.

[29] Howat, R. *Debussy in Proportion*. Cambridge University Press, Cambridge, 1983.

[30] Jones, M. R. Musical Events and Models of Musical Time. In R. A. Block (Ed.), *Cognitive Models of Psychological Time* (pp. 207-238). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc, 1990.

[31] Kellaris, J. J., & Kent, R. J. The Influence of Music on Consumers' Temporal Perceptions: Does Time Fly When You're Having Fun? *Journal of Consumer Psychology, 1, 4* (1992), 365-376.

[32] Koch, G., Oliveri, M., & Caltagirone, C. Neural networks engaged in milliseconds and seconds time processing: evidence from transcranial magnetic stimulation and patients with cortical or subcortical dysfunction. *Phil. Trans. R. Soc. B, 364, 1525* (2009), 1907-1918.

[33] Lalitte, P., & Bigand, E. Music in the moment? Revisiting the effect of large scale structures. *Perceptual and Motor Skills, 103, 3* (2006), 811-828.

[34] Lendvai, E. *Béla Bartók: An Analysis of His Music*. Kahn and Averill, London, 1971.

[35] Lendvai, E. Remarks on Roy Howat's 'Principles of Proportional Analysis'. *Music Analysis, 3, 3* (1984), 255-264.

[36] Levinson, J. Concatenationism, architectonicism, and the appreciation of music. *Revue Internationale De Philosophie, 60, 238* (Dec 2006), 505-14.

[37] Lewis, P. A., & Miall, R. C. The precision of temporal judgement: milliseconds, many minutes, and beyond. *Phil. Trans. R. Soc. B, 364, 1525* (2009), 1897-1905.

[38] McManus, I. C. The Aesthetics of Simple Figures. *British Journal of Psychology, 71, 4* (1980), 505-524.

[39] Noseworthy, T. J., & Finlay, K. A Comparison of Ambient Casino Sound and Music: Effects on Dissociation and on Perceptions of Elapsed Time While Playing Slot Machines. *Journal of Gambling Studies, 25* (2009), 331-342.

[40] Ornstein, R. E. *On the Experience of Time*. Harmondsworth: Penguin, 1969.

[41] Pöppel, E. Pre-semantically defined temporal windows for cognitive processing. *Phil. Trans. R. Soc. B, 364, 1525* (2009), 1887-1896.

[42] Russell, P. A. Testing the aesthetic significance of the golden-section rectangle. *Perception, 29, 12* (2000), 1413-1422.

[43] Tillmann, B. and Bigand, E. The Relative Importance of Local and Global Structures in Music Perception. *The Journal of Aesthetics and Art Criticism, 62, 2* (2004), 211-22.

[44] Tovey, D. F. *Essays in Musical Analysis, Volume I-II, Symphonies*. London: Oxford University Press, 1935.

[45] van Wassenhove, V. Minding time in an amodal representational space. *Phil. Trans. R. Soc. B, 364, 1525* (2009), 1815-1830.

[46] Wittmann, M., & van Wassenhove, V. The experience of time: neural mechanisms and the interplay of emotion, cognition and embodiment. *Phil. Trans. R. Soc. B, 364, 1525* (2009), 1809-1813.

[47] Zakay, D., & Block, R. A. Temporal cognition. *Current Directions in Psychological Science, 6, 1* (1997), 12-16.

[48] Woodrow, H. A Quantitative Study of Rhythm - The Effect of Variations in Intensity, Rate and Duration. In R. S. Woodworth (Ed.), *Archives of Psychology* (Vol. 2, pp. 335-401). New York: The Science Press, 1909.

Pop, Sampling and Postmodernism: The Case of Radiohead's *Idioteque*

Nina Jukić
Department of Musicology, Music Academy, University of Zagreb
Kašinska 41, 10360 Sesvete
Zagreb, Croatia
+385/98/9202-042
nina_jukic@yahoo.com

ABSTRACT
This research focuses on the song *Idioteque* by Radiohead, which features a prominent sample from Paul Lansky's computer composition *mild und leise*, which itself interpolates the opening motif from Richard Wagner's *Tristan und Isolde*. Assuming that sampling is a postmodern technique and referring to the principles of semiotics, this paper aims at showing whether or how the concept of postmodernism is applicable to Radiohead's work, through the analysis of the mentioned music pieces. Also, it tries to explain why Radiohead can be considered an important phenomenon in contemporary popular culture.

Keywords
Paul Lansky, Pop, Postmodernism, Radiohead, Sampling

1. INTRODUCTION
Radiohead are an English rock band formed in 1985, with their debut album *Pablo Honey* released in 1993. Its members are Thom Yorke (lead vocals, rhythm guitar, piano, beats), Jonny Greenwood (lead guitar, keyboard, other instruments), Ed O'Brien (guitar, backing vocals), Colin Greenwood (bass guitar, synthesizers) and Phil Selway (drums, percussion). With their first hit single *Creep* Radiohead began to receive attention in the media and gain success on the charts, thanks to their style then similar to the so-called grunge music that was popular in the early nineties and emerged from Seattle, USA. Their third studio album, *OK Computer* (1997) brought them worldwide fame and great critical acclaim – it introduced many new elements into Radiohead's music and art, but it was not so experimental (like the next two albums) that it would not be proclaimed a classic. After a difficult period of creative block, they came up with material for the following two albums, *Kid A* (2000) and *Amnesiac* (2001), which brought a more radical change in sound, style and attitude towards music industry. Despite of Radiohead's quite experimental approach to music and music consumption during the *Kid A*/*Amnesiac* era, these albums still gained great commercial success and critical appraisal. Since then, the band has continued to develop their innovative approach, which also includes the most recent album, *In Rainbows* which was released through the band's own website as a digital download for which customers could make whatever payment that they wanted, including nothing. They push boundaries in the visual arts as well by collaborating with the artist Stanley Donwood. Radiohead have always been politically, socially and ecologically sensible, criticizing the alienation, dehumanization and the technological development of the modern world, which are the most common topics of their songs. And despite of all the success and popularity, they have still remained quite an "underground", alternative image with no scandals, no excesses, no extravagant live shows, and not drawing attention to their private lives.

Radiohead are at the same time a very untypical "rock band" and still one of the most famous in the world. That is why they make an interesting research topic in contemporary popular music studies. How is it possible that a music group can be considered "popular" and "artistic" at the same time? Are those two categories not supposed to be separate? This research hopes to find the answer in the concept of postmodernism, by applying it to *Idioteque*, a song by Radiohead which is an excellent example of the band's complexity. *Idioteque* is interesting not just because it questions the usual opposition in popular music between "rock" and "electronica", but also because it features a prominent sample from Paul Lansky's computer composition *mild und leise*, which itself interpolates the opening motif from Richard Wagner's *Tristan und Isolde*. This paper will try to explain why sampling is considered to be a postmodern technique, how and why did Lansky cite Wagner and how and why did Radiohead sample Lansky, why postmodernism can be a concept relevant for contemporary culture and why Radiohead, if postmodern, are a relevant phenomenon in this culture.

2. POSTMODERNISM
The term "postmodernism" was coined in the writings of Fiedler (1969), Jencks (1975), Lyotard (1979) and Welsch (1997). It was first used in literary criticism, and then transmitted to the field of architectural theory. Since the early 1980s the term has been used to describe the general social, economic, politic and especially artistic conditions and their relation to modernism. According to Welsch, the main feature of postmodernism (not as a new epoch following Modernism, but more as a "project of modernity", in words of Habermas) is pluralism, from which all other characteristics can be derived. [1] In the optimistic understanding of postmodernism (Welsch), its main characteristics are not arbitrariness ("anything goes") and indifference, but "the simultaneity of the unsimultaneous", the closeness of the most different things. In the arts, postmodernism is realized through the constant effort to exceed the limits and taboos of modernism. This does not mean that artists give up on innovation – it means that innovation is not

being achieved anymore by increasing the complexity or differentiation of the material, but by combining of the already familiar, by citing what was once considered to be an "already used material". [1] In this way, the repetition-taboo of modernism is broken and the old is creatively embedded in a new context. The postmodern aesthetics is based on the stylistic device of double coding and a shift in the understanding of art and commercialism. This brings us to the notion of postmodernism in popular music industry, where it is possible that art and commercialism go hand in hand. The economical aspect of art is no longer denied, but instead provocatively accentuated. [1] Through citation, pastiche, self-reference, irony, intertextuality and the exposure of the working process, the illusions of organic unity, "high" art and the originality of the author become deconstructed.

2.1 Pop and Postmodernism

In his "Postmodernism and Popular Culture" (1992), Strinati has suggested that there are five main characteristics of postmodernism: the breakdown of the distinction between culture and society, an emphasis on style at the expense of substance and content, the breakdown of the distinction between high culture (art) and popular culture, confusions over time and space and the decline of the "meta-narratives". [7] In this account, postmodernism entails a new fluidity between culture and society. Society cannot be defined without taking account of cultural factors. There is also a new fluidity between forms of popular culture and high culture, which has led to difficulties in defining precisely what these different forms of culture actually refer to. For example, in the late 1960s, the artist Andy Warhol collaborated with the rock band the Velvet Underground, which itself included a classically trained musician John Cale and the more pop-orientated writing of Lou Reed. The resulting product is difficult to classify in high or popular culture terms. It has been suggested that postmodernism has now become a "cultural dominant". [7] If this is the case, it would affect popular music. Thus, it has been suggested by some analysts that the breaking of the boundaries between previously separated genres of popular music evidences the development of postmodernism. The example would be Radiohead's interest in both rock music, electronica and classical music, especially avant-garde music of the 20th century (Jonny Greenwood, classically educated multi-instrumentalist, is a big fan of Messiaen, Penderecki, Ligeti and others), and even jazz. Likewise, the break with the narrative of the conventional song could be argued to be postmodern. In many Radiohead's songs since *Kid A* we can find fragmentized lyrics, repeated without syntactical necessity, instrumentalized and digitally manipulated voice that is barely decipherable, etc. This was possible because they decided to step into the realm of electronic and experimental music, and mixing it innovatively with their rock roots. A related case has been made for the interconnection of music and image in contemporary music video which some have seen as a postmodern form. Radiohead have proved to be very innovative and creative in this field as well – from the series of "blips" or very short "antivideos" set to portions of tracks from *Kid A* and *Amnesiac* to the recent *House of Cards* video made with no cameras or lights, but instead with 3D plotting technologies which collected information about the shapes and relative distances of objects, so the video was created entirely with visualizations of that data.

The benefits of a concept like postmodernism are that it illustrates the linkages and parallels between different areas of culture which are influencing each other, and facilitates the identification of overall trends in society and culture, showing in broad outline how society is changing. [7]

2.2 Sampling and Postmodernism

The act of taking a portion (sample) of one sound recording and reusing it as an instrument or a different sound recording of a song has its roots in the art techniques of Dadaism and pop art, such as collage and readymade, while many of the manipulations offered by samplers directly relate to the techniques developed in *musique concrète*. The idea of creating interesting sounds by using electronic machines to generate signals, which can be converted into acoustic energy through the use of an amplifier and loudspeaker, is confined to the twentieth century. Indeed, affordable mass-produced synthesizers only began to appear in the late 1960s, as a direct result of the development of transistor technology and the principle of voltage control. The transition from analogue synthesis to digital synthesis heralded significant changes in both sound and practice. The use of sampling in popular music introduced many new ways of dealing with prerecorded material (editing, playing back, looping, etc.) [9]. These sonic manipulations owe something to the way vinyl records are used rhythmically by DJ's ("scratching"), as first developed by hip hop musicians during late 1970s. Specifically, the ways that samples are used in pop often tend to emphasize the rhythmic characteristics of the music. Samples also enable musicians to introduce sounds that are acoustically impure (often with limited bandwidth and dynamic range) and use them rather like sonic *objets trouvés* (readymades): their imperfections provide their aesthetic significance. It is for these reasons that sampling has become so important and significant in recent years: sampling, in spite of its background in digital technology, offers ways of countering the sonic impact of purely machine-driven and machine-generated music, and reintroducing sounds (albeit digital recordings) of human-made sound. Sampler becomes a musical instrument. In the context of postmodernism, where eclecticism, mixing of codes, parody and pastiche are considered to be its main features and stylistic devices [7], sampling becomes a powerful tool as well, because it most obviously follows the postmodern principle of intertextuality and serves as a proof to the breakdown of the distinction between high culture (art) and popular culture [1].

3. PAUL LANSKY'S *MILD UND LEISE*

3.1 Background

mild und leise, an 18-minute composition by Paul Lansky (born 1944), is one of the first pieces of music ever made with a computer. Lansky composed it in 1973 on the IBM 360/91 mainframe computer, using a second-generation computer synthesis language, Music360, written by Barry Vercoe specifically to take advantage of the capabilities of the machine. Lansky explained: "This IBM mainframe was, as far as I know, the only computer on the Princeton University campus at the time. It had about one megabyte of memory, and cost hundreds of thousands of dollars (in addition to requiring a staff to run it around the clock). At that point we were actually using punch cards to communicate with the machine, and writing the output to a 1600 BPI digital tape which we then had to carry over to a lab in the basement of the engineering quadrangle in order to listen to it." [6] This means that the sounds were written in digital form on nine-track magnetic tapes, about an inch wide, with the diameter of an LP, and holding only about 16 megabytes of data, which then had to be taken to a separate facility for digital-to-analog conversion. The process of composing *mild und leise* was very slow and difficult. It was not unusual for the machine to take an hour to compute a minute of sound, and this was working at a sampling rate of 14000 (as opposed to the standard rate today of 44100). [5] "[*mild und leise*] uses FM synthesis, which had just been worked out at Stanford, and later became the staple of Yamaha's DX7 series of synthesizers, and also a special purpose filter design program written (in Fortran IV) by Ken Steiglitz." [6]

mild und leise was eventually published on a record titled "Electronic Music Winners", on a Columbia/Odyssey LP in 1976, after it won a contest held by the *International Society for Contemporary Music* (ISCM).

Musically speaking, *mild und leise* actually began as an exploration of the famous "Tristan Chord" from Richard Wagner's opera *Tristan und Isolde*: "The piece is based on the 'tristan chord' and its inversions, hence the title. I worked out a multi-dimensional cyclic array based on this chord as the harmonic basis of the piece (…)". [6] The title of the piece comes from the beginning of Isolde's famous aria at the end of the opera.

3.2 The Tristan Chord in *mild und leise*

The "Tristan Chord" is a chord made up of the notes F, B, D# and G#. More generally, it can be any chord consisting of these same intervals. It is so named as it is the very first chord heard in *Tristan und Isolde*. (see Figure 1)

Figure 1. The beginning of *Tristan und Isolde*

The score of *Tristan und Isolde* has often been cited as a landmark in the development of Western music. Wagner uses throughout *Tristan* a remarkable range of orchestral color, harmony and polyphony and does so with a freedom rarely found in his earlier operas. The "Tristan Chord" is of great significance in the move away from traditional tonal harmony as it resolves to another dissonant chord. With this chord, Wagner actually provoked the *sound* or structure of musical harmony to become more predominant than its *function*, a notion which was soon after to be explored by Debussy and others. The chord and the figure surrounding it are well known to have been parodied and quoted by a number of later musicians. Berg derived the figure in his *Lyric Suite* from his twelve-tone compositional material; in *Golliwog's Cakewalk*, Debussy mockingly quotes the gloomy "Tristan Chord" in the middle of a lighthearted piece, while in *Pelléas at Mélisande* he includes it in a setting of the phrase "je suis triste". Wagner's use of musical color also influenced the development of film music. For instance, the opening of *Tristan und Isolde* was added to Luis Buñuel and Salvador Dalí's Surrealist film *Un chien andalou*.

What makes the "Tristan Chord" so interesting to musicians and theoreticians even today is its ambiguity - it can be interpreted and analyzed in several different ways, some more complicated than the others.

Lansky explains that he thought of the chord as "a linking of the musical intervals of a minor third and a major third, by a minor third (E#-G#, B-D#, linked by G#-B). (see Figure 2) (…) There are more sophisticated and appropriate ways think of how this harmony functions, but I was more concerned with interval networks at the time." [5]

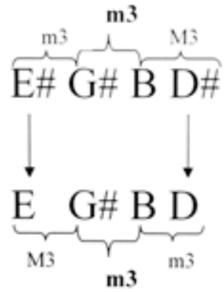


Figure 2. The two chords from the beginning of *Tristan und Isolde* as understood by Paul Lansky in his *mild und leise*

If we understand the "Tristan Chord" in this way, as a linking of the musical intervals of a minor third and a major third, by a minor third, it is clear that the chord following it can also be analyzed similarly: this second chord then is actually a linking of the intervals of a major third and a minor third, by a minor third. In this simplified schema, it can be seen that the two chords from the beginning of *Tristan und Isolde* are symmetrically related to each other (m3 + m3 + M3 and M3 + m3 + m3), with the minor third being the centre of symmetry in both chords.

In the process of exploring this network of relations, Lansky then combined two major thirds at the interval of a minor third (M3 + m3 + M3) (see Figure 3), to come up with the chord sequence that Radiohead sampled in their song *Idioteque*.

Figure 3. The chord from *mild und leise* which was used to build the sequence sampled in *Idioteque*

The section that Radiohead used consists of four different registral spacings of this four-note chord, but in *Idioteque* the original E, G#, B, D# is transposed a semitone lower (Eb, G, Bb, D) to accommodate Thom Yorke's (the singer's) tessitura (see Figure 4). The top two voices move up from the middle register, while the lower two voices move down. The similar sense of direction is present at the beginning of *Tristan und Isolde* (see Figure 1).

Figure 4. The sequence from *mild und leise* as sampled in *Idioteque*

But, in *Idioteque* this chord sequence plays quite a different role than it does in *mild und leise*, which will be explained in the following chapters.

4. RADIOHEAD'S *IDIOTEQUE*

4.1 Background

Idioteque is the eighth track on Radiohead's fourth studio album *Kid A* that was released in 2000. Following their third album *OK Computer* (1997), *Kid A* featured a minimalist style with less overt guitar parts and more abstract, electronic sounds, as well as diverse instrumentation including the ondes Martenot, programmed electronic beats, strings, and jazz horns and was thus considered too experimental and too "difficult" by some mainstream critics as well as fans, longing for a return to Radiohead's previous traditional-rock-band style. Not only was the music of *Kid A* something completely different from their previous works (incorporating much more synthetic sound and rhythm), unexpected among their fans, and thus a risky economical "move" (and risky economical moves are not very usual or welcome in most of today's popular music), but they also did not release any singles or videos to promote the album. Instead, they released a series of "blips" set to portions of tracks that were played on music channels and released freely on the Internet, being a useful deconstruction of the standard MTV-style video. Mark B. N. Hansen writes of *OK Computer* as of a referential-allegorical mode, and the shift to *Kid A* describes as the shift to expressive-symbolic mode. [3] This shift has to do with Radiohead's experimentation with techno music techniques like sampling and digital sound processing, but also with their willingness to deterritorialize its rock sound (by instrumentalization of the voice) in a process that yields an expansion in the sonic terrain of its music, and on the other hand, to bring this expanded terrain back to bear on the performance-oriented model of rock itself (as opposed to techno) – by, for example, employing atypical instruments (for a "guitar-based rock band") like the ondes Martenot and other analogue synthesizers capable of "simulating" the effect of digital processing in a live environment.

The voice of Thom Yorke in *Kid A* and *Amnesiac* (2001) is being manipulated by different digital techniques, like vocoder and Autotuner. It is not surprising that he turned to the rhythms of techno music and the techniques of digital composition in search of a solution to his "disgust with his own voice" in that period. [3] In *Kid A*, his voice is being mostly fragmented. This process of instrumentalization of the human voice was employed already with Stockhausen's *Gesang der Jünglinge* (1995-56).

Idioteque is the most "techno" song on *Kid A* in its sound, as well as the first so dramatically "techno" Radiohead song ever. It eschews the luxurious harmonic language and rhythmic and formal complexities of the songs from *OK Computer*. Instead of testosterone-driven guitars of earlier CD's, now there is a simpler and more abstract harmonic and formal language and a greater reliance on machine-made sounds. *Idioteque* consists of a strong machine driven beat that does not change throughout the song (it just stops for a few seconds at one moment, to accentuate its return even stronger), Thom Yorke's vocals, some electronic sounds and noises and samples from two pieces: Paul Lansky's *mild und leise* and Arthur Kreiger's *Short Piece*. Lansky wrote: "The piece puzzled me at first. I had never heard anything like it. Its profile was strange: sections were repeated many times, the tunes were relatively simple by Radiohead standards, and the textures were extremely unusual, with little deference to the slicker side of electronica. It took a number of hearings to begin to understand it, and eventually to genuinely like it." [5]

4.2 Analysis of *Idioteque*

While the looping sample from *mild und leise* made of four chords provides the entire and the only harmonic background to the song, samples from Kreiger's *Short Piece* are more difficult to recognize, being scattered all throughout the song and thus play a more "decorative" role in *Idioteque*. Although, there is a distinctive and prominent 11 seconds long sample from *Short Piece* at the very beginning of *Idioteque* (in *Short Piece* it can be found at 00:59-01:10), preparing the *mild und leise* sample to appear for the first time right after it.

Thom Yorke's vocals in the song, interestingly enough, have not been digitally processed like in many other songs on *Kid A*. This is surprising at first hearing, because one would expect to hear typical "techno" vocals in such a typical "techno" song. But, neither *Idioteque* is a techno song, nor Radiohead are a typical band. Yorke sings with his natural voice, sliding into his famous falsetto during the chorus, but the voice is nevertheless instrumentalized. It is the melody and the lyrics that give the impression that it might be a machine singing, a sort of a robot. Also, his own voice singing "come first, women and children come first…" has been sampled and looped in the background occasionally to provide another rhythmical layer to the song – which is another proof of using the voice as an instrument. Yorke uses a simple melody and repetition and fragmentation of the lyrics in order to create the effect of a non-human, machine-like voice:

Who's in a bunker?
Who's in a bunker?
Women and children **first**
And the children first
And the children

I laugh until my head comes off
I swallow until I **burst**
Until I burst
Until I

(Etc.)

The bolded lines signify the repetitions which occur at the end of verses, with lyrics being fragmentized by gradually omitting words and being sung on the same note (Eb, or sometimes F-Eb) with which the line that is being repeated and fragmentized ended ("first", "burst"…). This creates a sort of an echo-effect – with his natural singing voice Yorke actually imitates electronic sound manipulation processes like looping. An interesting paradox occurs when these "fake looping" lines are actually – looped (the already mention line: "come first, women and children come first…").

The choice of melody is also important for this effect and it depends largely on the harmony of the sample from Lansky's *mild und leise*. There are no real pitch-producing engines in *Idioteque* aside from Yorke's voice (with Ed O'Brien's during the chorus, and sometimes during the verse) and the computer chords. The choice of notes for the tune reveals a very sensitive and imaginative response to the color of the chords. While the chords all project a similar distinct color, since they all consist of different spacings of the same four notes, the choices of pitch in the tune reflect a subtle parsing of the chords into two overlapping triads, an Eb major triad and a G minor triad (see Figure 5).

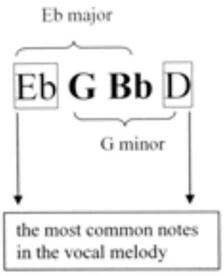


Figure 5. The role of the sampled chord sequence from *mild und leise* in the harmony and melody of *Idioteque*

The last of the four chords in the sequence has Eb and Bb in the bass, which more forcefully projects an Eb major flavor than other spacings. When this final chord arrives the notes of the tune move to Eb, while at other times D is more central. Eb and D are the two pitches that are not in the intersection of the Eb major and G minor triads. Thus the choice of notes for the tune, mainly using the pitches unique to the Eb major or G minor triads, pushes the song into a world where these triads oscillate and alternate. It is thus difficult to assert that the song is "in" Eb major or G minor. It is simply defined by the color of these two chords. [5] But, it is not important what key the song is in. The harmonies forge their own syntactical unity. It can be said that the chords are thought of as textures rather than aspects of progression and the main focus of the moment is on the sonic qualities of the ensemble. The basic (and only) harmony is merely an overlapping Eb major and G minor triad, and there is no functional sense of progression or key. The function of the static harmony is that it acts like a placeholder, a single harmonic landscape, a background that puts other musical elements into focus, like the accumulation of novel electronic effects and sounds. The tune of the verse section is relatively flat, oscillating gently between D and Eb (as mentioned earlier), but in the refrain things change when voice leaps to a high Bb and the descending tune fills in the span of a perfect 5th, using a slower rhythmic pulse, matching that of the computer chords. The melodic and rhythmic relations and contrasts between verse and refrain are thus intensified and made more effective. [5] In *Idioteque* (which is 5 minutes and 8 seconds long), Radiohead used only a few seconds from *mild und leise* (seconds 44 to 53) to create the looping sample. But, this chord sequence in *Idioteque* plays quite a different role than in *mild und leise* – due to looping, it receives a certain rhytmical quality in the way that it is not the case in *mild und leise*, it becomes the harmonic basis for the vocal melody and, being a smooth contrast to the relentless strong repeating machine-like sound of the electronic drum beat, it contributes significantly to the song's texture.

Idioteque is a perfect example of the Radiohead's postmodern breaking of the boundaries between previously separated concepts. [7] "(…) Radiohead should be celebrated for its ability to mix together categories normally opposed – analog and digital, rock and techno, breath-based and machine beat – in ways that expose ever deeper sonic affinities between noise and music, and deploy them to expand rock itself." [3] In *Idioteque*, Yorke's fragile human voice is superimposed on the strong machine rhythm, and although the music seems quite danceable, the lyrics are definitely not the kind one would find in typical dance music of today. They are paralleled in the visual artwork for the album *Kid A* by Stanley Donwood, the artist responsible for all Radiohead artwork, and Yorke, under the pen name "Tchock". Donwood's paintings depict a wasteland covered by sheets of ice and snow, with fires in distant forests and genetically modified bears and other mysterious shapes taking control of human civilization. Lines such as "Ice age coming", "We're not scaremongering, this is really happening" and "Mobiles skwrking, mobiles chirping, take the money and run" create an apocalyptic imagery which also functions as a critique of the modern society, climate change and alienation in the age of digital technology and virtual communication.

4.3 The Songwriting Process

Thom Yorke said in an interview: "*Idioteque* wasn't my idea at all; it was Jonny's. Jonny handed me this DAT [Digital Audio Tape] that he'd gone into our studio for the afternoon... and, um, the DAT was like 50 minutes long, and I sat there and listened to this 50 minutes. (…) Then there was this section of about 40 seconds long in the middle of it that was absolute genius, and I just cut that up and that was it." [10] Apparently Greenwood spent one afternoon throwing together a collage of sounds against his synthesizer's simulated drum track. These included *mild und leise*, music by Penderecki, Art Krieger's *Short Piece* (from the same LP as *mild und leise*, and which Greenwood found in a second hand record shop), street noises, radio noises and other sounds. He then gave the mix to Yorke, who extracted the sequence from *mild und leise,* composing *Idioteque* around it. The notion that Radiohead would take an excursion into the domain of experimental music in composing *Idioteque* is fascinating and significant. Yorke had also said that he wrote the lyrics for *Kid A*/*Amnesiac* by writing down some unconnected words and phrases, cutting them and mixing them up in a hat, and then deciding on the order by pulling them out of the hat. This sort of compositional process, where some element of the composition is left to chance, is already known from aleatoric music and composers such as John Cage. The influences of post-war serialism, Cage, Fluxus, electronic music, rock music and so on, have expanded the domain of the musically acceptable to include events and activities that were previously regarded as extraneous to the texture of music. It's only necessary to contemplate the now commonplace feel of bleeps and bloops at various points during *Idioteque*, for example, or of the electronic media noise at seams of so many Radiohead songs (or to see Jonny Greenwood waving a portable radio around at concerts, tuning random sounds into the ongoing fabric) to see an instance in which noises in the sonic surface of a composition are no longer regarded as novel. Also, the use of technology in all its forms – reproductive, as in recording, and generative, as in the use of guitars, synthesizers and computers – has shaped the way composers work. Radiohead's working method in *Idioteque* reveals an adventurous attitude toward the compositional process, technology and the musically acceptable. It is very interesting to see the band visit the domain of experimental and avant-garde music. Its motivation for doing so is not common in the world of popular music and is undoubtedly directly related to the artistic crisis that followed the success of *OK Computer*. [5]

5. SAMPLING/CITING AND SEMIOTICS

We have explained the way that Lansky cited Wagner and the way that Radiohead sampled Lansky. But, what does all that mean? Why is that important? Because in *Idioteque* a double postmodern phenomenon of "the old in a new context" is achieved – first, the Wagner's "Tristan Chord" becomes reinterpreted in a completely new medium of a computer composition, Lansky's *mild und leise* and second, a section from this piece gets a completely new function in the overall context of Radiohead's *Idioteque.* Thanks to sampling, in the (rock?) band's song there is the (art?) music of both Lansky and

Wagner – this fact is again the proof of Radiohead's postmodernity, as defined earlier. Much pop music sampling relies on identification of specific samples on the part of the listener. In the case of *mild und leise* and *Idioteque*, the question of source identification is an interesting one, because it shows how these two pieces of music are different in their relation to their source of inspiration. Semiotics can help to explain this. According to Saussure, a sign is composed of the signifier and the signified, with the signifier being the material representation of a sign and the signified being the concept indicated by the signifier. In this light, a distinction can be made between "citation" and "sample". While sample is the repetition of the signifier, citation is trying to repeat the signified as well, in other words, the meaning. [2] In *mild und leise*, its title functions as a sign – it tells us that the piece refers to Wagner's *Tristan* (signified) and thus we can try to connect this information to the music itself (signifier). In *Idioteque*, though, the sample from *mild und leise* is taken out of its original context; it becomes a mere signifier and is repeated as such – without the obvious signified, unless we find out that it is a sample of Lansky. When we hear *Idioteque* or read its title, there is nothing to tell us that we are listening to a sample of Lansky who cited Wagner. We do not hear Wagner in *Idioteque*, but it *is* there. Although different in the sense explained above, the compositional processes of both *mild und leise* and *Idioteque* are still quite similar, because both Lansky and Radiohead reached for inspiration or were intrigued by somebody else's music – it does not really matter if it was a chord from an opera or a vintage early electronic music LP.

6. ART OR POP - OR BOTH?

What is popular music? If you had asked Theodor Adorno, he would have said that popular musical compositions follow familiar patterns and introduce little originality. They belong to the culture industry, thus they are not authentic, unlike "serious" music, which does have artistic value. Some fifty years later, Nick Hornby agreed on this, but his question was primarily "what popular music *should* be like?". He gave his answer to this question in his notorious review of Radiohead's album *Kid A* ("Beyond the Pale", *New Yorker*, October 30, 2000), where he complained that the album is too experimental, too arty and too non-commercial. What Hornby wants Radiohead to be is a stereotypical rock band who should give you "your money's worth". But what Radiohead did with *Kid A* and continued to do since then is quite the opposite. They resist fitting into stereotypes and mainstreams of pop/rock music in many different areas – from experimentation and innovation in music and visual art to challenging the music industry's requests for a sellable music product; from usage of many different technologies besides the traditional guitar-band line-up and referring to influences in contemporary art music to questioning the notion of the rock star as a sex symbol. Radiohead challenge the distinction between the so called art and pop music in a truly postmodern manner.

So, is Radiohead a pop band or not? They belong to neither the world of pop nor high culture in the traditional sense, but at the same time, they somehow belong to *both* of the worlds. Radiohead are not trying to deny that they are inevitably a part of the global economic system – rather, as Yorke said in an interview: "If you're interested in actually being heard, you have to work within the system." [8] After all, Radiohead is *just* a rock band (as the band often reminds us – and this attitude, that artist is not a genius, is yet another proof of postmodernism). But, the content of Radiohead's art should be lauded for suggesting the space necessary to even discuss the possibility of resistance. [8] Because, their goal is to effectively criticize and transform values, rather than simply make money as entertainers. [4]

7. CONCLUSION

Postmodern theory is a very useful tool for the comprehension of current popular music and culture, especially if it is applied to the work of a band such as Radiohead, because it helps to explain the linkages and parallels between different areas of culture which are influencing each other. Radiohead are far from being a stereotypical rock group and they often visit the domain of experimental and avant-garde music, which is not common in the world of pop music. In a truly postmodern manner, Radiohead thus challenge the distinction between the so called art and pop music. All this makes them a relevant and interesting research topic.

8. REFERENCES

[1] Appen, R. von and Doehring, A. Künstlichkeit als Kunst. Fragmente zu einer postmodernen Theorie der Pop- und Rockmusik, in: Phleps, T. (ed.), *Beiträge zur Popularmusikforschung 27/28: Populäre Musik im kulturwissenschaftlichen Diskurs II*, ASPM, Karben, 2001, 13-33.

[2] Bonz, J. Sampling: eine postmoderne Kulturtechnik, in: Jacke, Ch., Kimminich, E. and Schmidt, S. J. (ed.), *Kulturschutt. Über das Recycling von Theorien und Kulturen*, transcript Verlag, Bielefeld, 2006, 323-353.

[3] Hansen, M. B. N. Deforming Rock: Radiohead's Plunge into the Sonic Continuum, in: Tate, J. (ed.), *The Music and Art of Radiohead*, Ashgate, Hants, 2005, 118-138.

[4] Kaye, B. "Kid A" as a Musing on the Postmodern Condition, in: Forbes, B. W. and Reisch, G. A. (eds.), *Radiohead and Philosophy: Fitter Happier More Deductive,* Open Court, Chicago, 2009, 241-250.

[5] Lansky, P. My Radiohead Adventure, in: Tate, J. (ed.), *The Music and Art of Radiohead*, Ashgate, Hants, 2005, 168-176.

[6] Lansky, P. my radiohead adventure, (Accessed 11 October 2009), http://www.music.princeton.edu/paul/radiohead.ml.html.

[7] Longhurst, B. *Popular Music and Society*, Polity Press, Cambridge, 1995.

[8] Schneiderman, D. "We got Heads on Sticks / You got Ventriloquists": Radiohead and the Improbability of Resistance, in: Tate, J. (ed.), *The Music and Art of Radiohead*, Ashgate, Hans, 2005, 15-37.

[9] Warner, T. *Pop Music – Technology and Creativity: Trevor Horn and the Digital Revolution*, Ashgate, Hants, 2003.

[10] Thom Yorke Talks About Life in the Public Eye, a radio interview originally broadcast on 12 July 2006, (Accessed 14 October 2009), http://www.npr.org/templates/story/story.php?storyId=15226006.

Marching to the same beat: entrainment and persuasion

Sarah Knight
University of Cambridge
Centre for Music and Science
Faculty of Music
11 West Road
Cambridge
CB3 9DP
UK
slk34@cam.ac.uk

ABSTRACT
This paper presents a theoretical discussion of the role of entrainment in oratory, and specifically its ability to aid communication and persuasiveness through both its attentional and also its interpersonal and social functions. Explicit links are thereby forged between musical performance and performative speech.

Keywords
Entrainment; musical performance; oratorical performance; persuasion.

1. INTRODUCTION
Entrainment has been shown to play an attention-directing role in music perception, and, thanks to its ability to create prolonged and highly accurate synchrony between interactants, various researchers also attribute it a key role in social and interpersonal bonding. It is not, however, solely a musical phenomenon – it has been shown to function in the perception of certain speech types, too.

Various researchers describe the important part played by speech rhythm in persuasive oratory, and some hint at a role for entrainment, but this role is overwhelmingly characterised as purely attentional – a means of communicating speech content more effectively. It is proposed here, however, that the affiliative powers of entrainment are also at work in such situations, increasing the speaker's persuasiveness independently of the speech's semantic content.

Links between musical performance and performative speech have long been theorised; for example, Darwin [14] wrote in 1874 that "We must suppose that the rhythms...of oratory are derived from...musical powers". The theoretic framework presented here takes steps toward formalising such links with respect to structure, cognitive processes and functionality.

2. NONVERBAL PHATIC COMMUNION
Nonverbal cues, such as gesture and prosody, have been shown to play a key role in human communicative interactions. This role is multifaceted. In speech-based interactions, nonverbal cues may serve to clarify linguistic meaning. This clarification may take the form of, amongst other things, stress-based word differentiation (e.g. PERmit (n.) vs. perMIT (v.)), the disambiguation of discourse function (e.g. distinguishing a question from a statement), or the use of content-related gestures [17; 18; 31]. However, both prosody and gesture may additionally function in a non-linguistic role, carrying information to the listener regarding the speaker's attitude and physical state [11; 17]. There is also a function of nonverbal cues which is unrelated to the communication of information – linguistic or otherwise – and is instead concerned entirely with promoting interpersonal affiliation and social bonding. As such, this third function can be viewed as something like a nonverbal equivalent of Malinowski's "phatic communion" [27], or even of Austin's "speech acts" [2] – a direct, real-time means of actively forging and maintaining social ties.

Two primary components of this nonverbal phatic communion are behavioural mimicry (shared actions across time) and synchrony (shared actions in shared time). As Bavelas et al [5] point out, behavioural mimicry and/or synchrony are a display of "sameness" – both personal and attitudinal, of social and interpersonal affiliation and of topical agreement. However, as both Bavelas et al [5] and La France [23] stress, this display is not necessarily an expression of an existing state; rather, it is *constitutive* of that state. That is, such cues actively create a sense of rapport and agreement, and a lack of them can, in Bavelas' words, "signal discord *or even cause it*" ([5], p. 334; my italics). Moreover, by creating a sense of mutual goals and action plans, these behaviours contribute to the creation of *shared intentionality* between interactants – a state of collaborative interactivity defined by shared goals and action plans, and suggested to be a motivating factor in human cognition in general [39] and human communication in particular [10]. It is in these senses, then, that such cues are active and performative: through mimicry and synchrony interlocutors "perform" their sameness, and this performance in turn actively creates the shared intentionality, the social and interpersonal bonds, and the moves towards agreement of which such cues appear to be a demonstration. Furthermore, this creation is an immediate, "bottom-up" process which takes place without the mediating step of information decoding.

Illustrations are numerous. For example, it is well-documented that interlocutors mirror each other's movements, facial expressions and postures near-synchronously – a process which increases empathy, rapport and sense of agreement [3]. This process was identified as early as 1964, when Scheflen noted that "congruence in posture indicates similarity in views" ([35], p. 328). Scheflen's work was taken up by LaFrance, who confirmed empirically that posture sharing and rapport were positively related and, crucially, suggested that posture sharing has causal priority – that is, that shared bodily alignments establish affiliation as well as merely reflecting it [23]. In recent years this causal priority has been verified [24], and the topic is still being explored; for example, Shockley, Richardson and Dale [38] studied coordination of body sway and gaze between interlocutors, and suggested that such phenomena promote joint alignment of cognition as well as of action. Aside from posture sharing, there is evidence that speakers in a dyad adjust to one another's gestural rhythm, creating an "interspeaker rhythm" hypothesised to aid "personal rapport" [29]. In the speech signal itself one finds the process of intonational "accommodation", by which a speaker matches his prosodic features to those of his interlocutor and thus enhances affiliation [31]. Synchrony, although rarer in most speech-based interactions, has nevertheless been shown to occur. Gill's work on "Body Moves", for example, describes instances of physical synchronisation between collocutors at moments of topical agreement; at these points the speaker/listener divide is temporarily broken down through the creation of joint action, suggesting a shared psychological state [16].

Given that mimicry takes the form of shared actions, but synchrony puts shared actions into shared time, it seems plausible that synchrony will establish a stronger sense of shared intentionality and more powerful interpersonal bonds than mimicry; and indeed, recent research suggests that synchrony does promote greater affiliation [20]. However, synchrony is difficult to sustain accurately over multiple actions. Nevertheless, there is a powerful means by which humans can achieve prolonged and relatively precise synchrony: the process of entrainment.

3. WHAT IS ENTRAINMENT?

Entrainment refers to the synchronisation of a biological, internal, regular oscillator with an external, periodically-recurring event. In two interdependent processes, it both allows expectancies to be generated about the timing of future events and also directs attention, matching attentional peaks to the external events [4; 21; 25]. Humans may make their internal entrainment manifest through entrained physical actions [22]; for example, many people tap their feet to the regular "beat" when listening to a piece of music. However, internal entrainment needn't necessarily have physical correlates; it has been shown to be present in babies too young to actively synchronise their movements to a beat [43] and one's own experience shows that it is perfectly possible to "keep up with" the beat in a piece of music without tapping one's foot.

4. WHY DO WE ENTRAIN?

As mentioned above, entrainment both allows the generation of expectancies about the timing of future events and also directs attention, matching attentional peaks to events. However, through promoting entrained physical action, it is also a powerful means of evoking and maintaining synchronous action between interactants exposed to the same external periodic stimulus. It thus seems likely that entrainment serves a second function, related to social and interpersonal bonding. And indeed, of the various theories that exist as to the evolutionary role of entrainment, almost all put forward some kind of affiliation-based explanation. For example, Cross and Woodruff [10] propose that music developed as a means of negotiating social uncertainty; that is, music's semantic indeterminacy allows participants to hold their individual interpretations of a situation without forcing confrontation between differing positions, while the temporally-structured group interaction afforded by entrainment nevertheless simultaneously creates affiliation. Bispham [6] stresses that rhythmic behaviours accompany and underlie all social interactions, and suggests that, through creating a shared temporal framework for interaction, entrainment gives rise to collective emotionality, shared experience and group cohesiveness. Kirschner and Tomasello [22] also emphasise the social role of entrainment, putting its efficacy down to its ability to create joint attention – the condition underlying shared intentionality [37]. Hove [19] suggests that the interpersonal synchrony made possible by entrainment could "lead to interpersonal empathy and understanding" (p. 30), while McNeill [30] describes "muscular bonding" – an "enhancement of social cohesion" (p. 4) and "euphoric fellow-feeling" (p. 2) caused by group movement in unison with a regular pulse. Meanwhile, infant/caregiver entrainment has been suggested by researchers to foster intersubjectivity and to give rise to shared emotions and experience [28; 40; 41; 9]. Indeed, entrainment forms a central part of Malloch's notion of "communicative musicality" – the process by which the combination of "musical" factors in human communication creates an emotional "companionship" [28]. In empirical studies, Hove and Risen [20] showed that participants displayed greater affiliation with experimenters when the two had engaged in synchronised tapping than when the tapping had been asynchronous or the experimenter had merely observed. Wiltermuth and Heath [42] present evidence that groups who engaged in synchronised physical activity (i.e. entrained to one another's actions) felt a stronger sense of group cohesion and trust than those who had not. And Macrae et al [26] showed that, when the physical gestures of a conversing dyad were synchronised to a regular pulse, participants' memory for utterances and facial appearance was facilitated, suggesting that entrainment had enhanced social interaction.

Most of these theories deal with stimuli that could be considered "musical" – perhaps unsurprisingly, since an external, periodically-recurring event (the "beat") forms a major component of musical structure. However, many modern musical experiences are isolated and/or passive in nature – the strongly unidirectional concert hall context with its highly constrained audience, for example, or solitary iPod listening – and thus seem far removed from the powerfully social interactions central to the hypotheses discussed above. Nevertheless, entrainment by its very nature still takes place in these contexts; and, indeed, this is one of the ways in which music is hypothesised to remain meaningful when removed from the social domain. As Cross and Woodruff [10] describe, "even in contexts that involve apparently passive listening..." such cues "afford the experience of joint action, joint attention and joint intention" between the listener and the virtual "persona" of the music (p 7). Thus the affiliative effects of entrainment persist even when overt synchrony and/or mirroring are not possible.

In conclusion, then, it seems that entrainment is a particularly potent means of creating and maintaining social ties across a range of communicative contexts. In particular, entrainment serves to maintain a sense of shared intentionality and "interpersonal" bonding even in the apparently passive context of many musical encounters in the modern world.

5. ENTRAINMENT TO SPEECH

Although it plays a key role in music perception, entrainment is not limited to stimuli we might consider "musical"; it has also been shown to take place when speech signals contain high levels of periodicity [13]. Furthermore, the models proposed for speech- and music-based entrainments share basic elements [15; 33], reflecting the qualitative similarity of the two processes. However, entrainment can only take place when perceptual regularities are present, and such regularities appear to be very rare in normal speech [11; 32]; in other words, levels of periodicity in everyday conversations are generally not high enough to afford entrainment. As a result, interactants rely on mimicry and brief moments of less rigidly structured synchrony as their main nonverbal routes to affiliation.

Not all speech types, however, are "normal" everyday conversation, and this paper focuses on one such specialised speech type: public persuasive oratory in the modern Western sense, most frequently seen in political speeches and debates. Both anecdote and descriptive accounts suggest that this type of speech is characterised by a higher occurrence of sections of perceptual periodicity than everyday speech. And indeed, the authors who describe such periodicity do so in a way that strongly suggests entrainment processes. For example, Quené and Port [33], in a paper on speech-based entrainment, suggest that the "regular timing of stressed syllables...helps to get a spoken message across to the listener" (p. 11). Cummins [12] describes how "effective public speakers carefully time their speech to maximally engage an audience", and that this engagement "often manifests as physical entrainment" among listeners (p. 2). Atkinson [1] explains how the temporal regularity of rhetorical devices allows the audience to "come in on time" (p. 59) – a turn of phrase strongly suggesting the expectation-generating function of entrainment. And several researchers emphasise the crucial role played by rhythmic regularity in determining how successfully preachers engage their congregation [11; 34]. These studies thus suggest that oratory is one speech type which provides – albeit for only brief sections – the perceptual regularities necessary for entrainment to take place; moreover, they imply that this improves communicative efficacy by directing listeners' attention at these points. But, as described above, attention-direction is not the sole function of entrainment – it also has a strong interpersonal and social function. Such research therefore begs the question: could oratory also benefit from this second function of entrainment? And could there be positive effects on persuasiveness, as well as simply on message communication?

6. PERSUASION, ORATORY AND ENTRAINMENT

The term "persuasion" is generally used to refer to those processes governing the formation and change of attitudes, where "attitude" denotes a psychological tendency to evaluate an attitude object in a particular way [8]. Persuasion is generally treated in the literature as a "top-down" process involving learned socio-cultural heuristics [e.g. 7; 36]. However, persuasion processes also seem likely to be influenced by the more immediate and "bottom-up" phatic domain. To explain: nonverbal behaviours have been shown – as detailed above – to be highly successful at creating at least a *sense* of agreement, shared interests and joint intentions between communicators; and for a speaker to induce in a listener a sense of shared accord and mutual aims seems highly likely to influence that speaker's persuasiveness. Moreover, the notion of creating shared intentionality, with its underlying shared goals and shared action plans, is strongly suggestive of the move towards shared attitudes which lies at the heart of persuasion.

Of course, the performance context of most oratory – in which a large audience remains mostly quiet and still – renders interactive mimicry and synchrony unlikely, if not impossible, thus making internal entrainment processes the primary nonverbal affiliative tool. However, this is not an implausible scenario: after all, the relatively passive oratorical performance context is shared with many modern musical performances – for example, in the concert hall – and, as noted above, music in these contexts relies on the persistence of entrainment's affiliative effects to remain efficacious in the phatic domain. It is in this sense, then, that entrainment is hypothesised to aid speaker persuasiveness: any listener entrainment that arises through perceptible regularities in an orator's speech or gesture would help to foster within the listener a sense of social affiliation, shared intentions and accord between speaker and listener. This in turn seems likely to encourage a convergence of speaker/listener attitudes, thus positively influencing a speaker's persuasiveness.

7. EMPIRICAL WORK

A basic preliminary test of these hypotheses, intended to lay the groundwork for more detailed empirical studies, is currently underway. This test relies on perceptual judgement rather than objective measurement, and is designed to swiftly deliver an indication of the value of performing more comprehensive tests using similar stimuli. To summarise: two experienced public speakers are videoed delivering a number of set speeches spontaneously, and these performances are taken as a baseline. The speakers are then trained to deliver the same set of speeches with increased levels of periodicity at appropriate moments. The speeches are matched as far as possible for length, content and location of periodicity. The videos are then shown to two groups of subjects, one of which rates the performances for rhythmic characteristics and the other of which rates the performances for persuasiveness. Correlations can therefore be sought between perceptions of rhythmicity and perceptions of persuasiveness. Although admittedly a blunt tool, this initial investigation should provide at least some indication as to the validity of, and possible directions for, more detailed and time-consuming testing involving, among other things, objective measurements of the speech signal.

8. CONCLUSION

The theory presented here combines previous hypotheses about the attentional role of entrainment in oratory with a new hypothesis about entrainment's interpersonal, persuasive function. In doing so, it takes initial steps towards formalising a long-hypothesised link between musical performance and performative oratory, suggesting that this link is, at least in part, not merely a superficial perceptual-structural similarity, but rather that these two communicative types developed to serve similar functions in terms of both listener-internal attentional control and also interpersonal affiliation and social dynamics. The theory is in the process of being tested empirically.

9. ACKNOWLEDGEMENTS

Many thanks to Dr Ian Cross and my colleagues from the Centre for Music and Science, University of Cambridge, for all their help and advice.

10. REFERENCES

[1] Atkinson, M. (1984) *Our Masters' Voices* (London: Methuen)

[2] Austin, J. L. (1971) *How to do things with words* (London: Oxford University Press)

[3] Bargh, J. A. and Chartrand, T. L. (1999) "The unbearable automaticity of being", *American Psychologist 54*, pp. 462–479.

[4] Barnes, R. and Jones, M.R. (2000) "Expectancy, Attention, and Time", *Cognitive Psychology 41*, pp. 254-311.

[5] Bavelas, J.B., Black, A., Lemery, C.R. and Mullett, J. (1987) "Motor mimicry as primitive empathy" in N.

Eisenberg and J. Strayer (Eds.), *Empathy and its development* (Cambridge: Cambridge University Press)

[6] Bispham, J. (2006) "Rhythm in Music: What is it? Who has it? And why?" *Music Perception 24 (2)*, pp. 125-134.

[7] Chaiken, S. (1980), "Heuristic versus systematic information processing and the use of source versus message cues in persuasion", *Journal of Personality and Social Psychology 39*, pp. 752-766.

[8] Chaiken, S., Wood, W. and Eagly, A. H. (1996) "Principles of persuasion" in E. T. Higgins and A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (New York: Guilford Press), pp. 702-742.

[9] Cowley, S. J., Moodley, S. and Fiori-Cowley, A. (2004) "Grounding signs of culture: primary intersubjectivity in social semiosis", *Mind, Culture Activity 11*, pp. 109–132.

[10] Cross, I. and Woodruff, G. E. (2008) "Music as a communicative medium" in R. Botha and C. Knight (Eds.) *The Prehistory of Language* (Oxford: Oxford University Press)

[11] Crystal, D. (1975) The English tone of voice: essays in intonation, prosody and paralanguage (London: Edward Arnold)

[12] Cummins, F. (2009) "Rhythm as entrainment: The case of synchronous speech", *Journal of Phonetics 37*, pp. 16-28.

[13] Cummins, F. and Port, R., (1998) "Rhythmic constraints in stress timing in English", *Journal of Phonetics 26*, pp. 145-171.

[14] Darwin, C. (1874) *The Descent of Man* (accessed via http://www.gutenberg.org/; last accessed 19/05/09)

[15] Gasser, M., Eck, D. and Port, R. (1999) "Meter as mechanism: A neural network model that learns metrical patterns." *Connection Science, 11 (2)*, pp. 187-216.

[16] Gill, S. P. (2004) "Body moves and tacit knowing" in B. Gorayska and J. L. Mey (Eds.), *Cognition and technology* (Amsterdam: John Benjamin)

[17] Gussenhoven, C. (2002) *The phonology of tone and intonation* (Cambridge: Cambridge University Press)

[18] Holle, H. and Gunter, T. C. (2007) "The role of iconic gestures in speech disambiguation: ERP evidence", Journal of Cognitive Neuroscience 19 (7), pp. 1175-1192.

[19] Hove, M. J. (2008) "Shared circuits, shared time, and interpersonal synchrony", *Behavioral and Brain Sciences 31 (1)*, pp. 29-30.

[20] Hove, M. J. and Risen, J.L. (in press) "It's All in the Timing: Interpersonal Synchrony Increases Affiliation".

[21] Jones, M. R. and Boltz, M. (1989) "Dynamic Attending and Responses to Time", *Psychological Review 96 (3)* pp. 459.

[22] Kirschner, S. and Tomasello, M. (2009) "Joint drumming: Social context facilitates synchronization in preschool children", *Journal of Experimental Child Psychology 102*, pp. 299–314.

[23] LaFrance, M. (1979) "Nonverbal Synchrony and Rapport: Analysis by the Cross-Lag Panel Technique", *Social Psychology Quarterly 42 (1)*, pp. 66-70.

[24] Lakin, J.L. and Chartrand, T.L. (2003) "Using nonconscious behavioural mimicry to create affiliation and rapport", *Psychological Science 14 (4)*, pp. 334-339.

[25] London, J. (2004) *Hearing in Time* (Oxford: Oxford University Press)

[26] Macrae, C. N., Duffy, O. K., Miles, L. K. and Lawrence, J. (2008) "A case of hand waving: Action synchrony and person perception", *Cognition 109*, pp. 152–156.

[27] Malinowski, B. (1994) "The Problem of Meaning in Primitive Languages" in J. Maybin (Ed.), *Language and literacy in social practice* (Clevedon: Multilingual Matters in association with the Open University)

[28] Malloch, S. N. (1999) "Mothers and infants and communicative musicality", *Musicae Scientiae Special Issue (1999-2000)*, pp. 29-57.

[29] McClave, E. (1994) "Gestural Beats: The Rhythm Hypothesis", *Journal of Psycholinguistic Research 23 (1)*, pp. 45-66.

[30] McNeill, W. H. (1995) *Keeping Together in Time: Dance and Drill in Human History* (Cambridge, Mass.: Harvard University Press)

[31] Nolan, F. (2006) "Intonation" in B. Aarts and A. McMahon (Eds.), *Handbook of English Linguistics* (Oxford: Blackwell).

[32] Port, R. F., Cummins, F. and Gasser, M. (1995) "A dynamic approach to rhythm in language: Toward a temporal phonology". Technical Report 150, Indiana University Cognitive Science Program.

[33] Quené, H and Port, R. (2005) "Effects of Timing Regularity and Metrical Expectancy on Spoken-Word Perception", *Phonetica 62*, pp. 1-13.

[34] Rosenberg, B. A. (1970) "The Formulaic Quality of Spontaneous Sermons", *The Journal of American Folklore 83 (327)*, pp. 3-20.

[35] Scheflen, A. E. (1964) "The significance of posture in communication systems", *Psychiatry 27*, pp. 316-31.

[36] Schultheiss, O. C. and Brunstein, J. C. (2002) "Inhibited power motivation and persuasive communication: A lens model analysis", *Journal of Personality 70*, pp. 553-582.

[37] Sebanz, N., Bekkering, H. and Knoblich, G. (2006) "Joint action: Bodies and minds moving together", *Trends in Cognitive Sciences 10*, pp. 70–76.

[38] Shockley, K., Richardson, D. and Dale, R. (2009) "Conversation and Coordinative Structures", *Topics in Cognitive Science 1(2)*, pp. 305-319.

[39] Tomasello, M., Carpenter, M., Call, J., Behne, T. and Moll, H. (2005) "Understanding and sharing intentions: The origins of cultural cognition", *Behavioral and Brain Sciences 28*, pp. 675-691.

[40] Trevarthen, C. (1999) "Musicality and the intrinsic motive pulse: evidence from human psychobiology and infant communication", *Musicae Scientiae Special Issue (1999-2000)*, pp. 155-211.

[41] Trehub, S. (2003) "Musical predispositions in infancy: an update" in R. Zatorre and I. E. Peretz (Eds.), *The cognitive neuroscience of music* (Oxford: Oxford University Press)

[42] Wiltermuth, S. S. and Heath, C. (2009) "Synchrony and Cooperation", Psychological Science 20 (1), pp. 1-5.

[43] Winkler, I., Háden, G. P., Ladinig, O., Sziller, I. and Honing, H. (2009) "Newborn infants detect the beat in music", *Proceedings of the National Academy of Sciences of the United States of America 106 (7)*, pp. 2468–

Experimental music in England in the late 1960s and early 1970s: the aesthetics of political and social conscience

Settimio, Fiorenzo Palermo
Middlesex University
14 Coppetts Road
London N10 1JY
0447734843304
sfpalermo@yahoo.co.uk

ABSTRACT
In this paper I attempt at understanding the aesthetic decisions of a small group of experimental composers active in England at the end of the 1960s and beginning of 1970s by taking into consideration the political, social and economical circumstances which informed them

Keywords
Experimental English music; free improvisation, music and politics; art and social conscience

1. INTRODUCTION
Towards the end of the 1960s and the beginning of 1970s a handful of composers based in England experimented with unconventional techniques in music. Cornelius Cardew (1936-1981), Hugh Davies (1943-2005), and Keith Rowe (1940-) among others seemed to abandon established paths for composition and performance. The Scratch Orchestra, founded by Cardew, Howard Skempton (1947-) and Michael Parsons (1938-), adopted graphic scores and the rendition of popular classic pieces by musicians and non musicians as most prominent methods; Hugh Davies built his own instruments on which to improvise on, either solo or as a member of the ensemble Gentle Fire or the Music Improvisation Company, with Derek Bailey (1930-2005) and Evan Parker (1944-) among others; Keith Rowe applied extended techniques to playing the guitar, conducting extemporizations with the group AMM, which involved Edwin Prévost (1942-), Lou Gare (1939-) and Cardew, as well as featuring the participation of Parker.

2. METHODOLOGY
Despite the apparent rupture with the traditions of the concert hall, such practices can be arguably rooted in musical history. For example the touring exhibition Eye Music [1], with Hugh Davies as curator, traced the origin of experimental and avant-garde graphic scores to elements already present in the past practice of notating classical music. Such comparisons are helpful to engender acceptance of an anomalous method in writing music, but seem unnecessary when discussing the motives behind their development. The diversity of the modi operandi of this group of composers is testament to their personal search for alternative viable routes in music. Each one's idiosyncrasies reveal their uniqueness. Nonetheless their collaborative activity suggests a common ground in their practice that perhaps transcends the specific methodology employed, involving factors at the origin of such production. The similar trajectory that some of these musical projects followed appears to be more than a coincidence, for example a marked alteration in their operation in 1972, points to a deep connection with the events taking place outside the musical context. An ulterior and substantial reason to concentrate on extra musical events to define the compositions and performance by this heterogeneous collection of composers is their public interest for matters relating to the development of socio-political theories and their practical application. Therefore my approach to studying their work will be a systematic analysis of their aesthetic choices with reference to the political, social and economical circumstances of that period in history, proposing a direct connection relating the two. I will focus on the work by Cardew, Davies in primis, including other composers operating outside the conventions of the concert hall, such as Daphne Oram (1925-2003) and Henri Chopin (1922-2008). I will concentrate on their activity approximately between 1968 and 1972, which seemed crucial years, both artistically and historically.

3. MUSIC SERVE IMPERIALISM
3.1 Refrain
During those years there seemed to be a general feeling that mobilized many to take action against government policies they disagreed with, in particular it was notable the protest to end the Vietnam War. Such activism to contrast authorities was reflected in Cornelius Cardew's criticism of the work by masters of the American experimental movement and of the European avant-garde, like John Cage and Karlheinz Stockhausen. In 1972 Cardew was asked to introduce the BBC broadcasting of Stockhausen's *Refrain*, perhaps also in virtue of Cardew's past as assistant to the German composer in the early 1960s. Cardew took this chance to launch an unprecedented attack towards a work he claimed to serve obsolete and socially unacceptable beliefs. Cardew declared to consider Stockhausen's pieces as part of the same exploitative and oppressive hierarchy, which he individualized as inherent to imperialism. He followed this inflammatory words with an exhortation to fight back, in equal measure, the war in Vietnam, Stockhausen's music, but also the same ideals that may have been conceived in each one's mind[2]. It seemed obvious that political theory, and in particular Marxism was all encompassing: investing music, shaking the fundamental structure of its composition and performance with the same

vehemence that students at the time were questioning the establishment in the States and Europe.

3.2 Treatise

This political conscience, perhaps a product of the reaction after World War II for authoritarian forms of government such as Fascism and Nazism, seemed to translate in Cardew's embracing of graphic scores. In fact his work Treatise appears to thwart the composer's authority in imparting specific instructions to the musicians, conferring notation a more inspirational role than of dictation. In fact the stave, notes and clefs are stretched, stylized and displaced to assume a prominent visual quality rather than of musical accuracy. Cardew seemed to seek an active participation from the performers, rather than impart them with prescriptive directions of how to conduct a rendition of the piece. He used this score to stimulate the musicians, whom would have responded by translating their own personal experience, background and individuality to what they believed the lines and dots on the page represented. With Treatise Cardew appeared to have given the means to performers to claim their contribution to a composition. Another option to abolish such hierarchical distribution of roles in music was to compose and perform own material to also ensure artistic continuity in the execution of a piece. Gentle Fire distinguished themselves as one of the few ensembles to play work composed by individual members or collectively, for example *Quintet* (1969) by Davies or *Group Composition V* (1972) [3].

3.3 From Scratch

Fighting authority, or at least protesting for a more egalitarian society was a characteristic feature of the sentiment of the time. Minority groups made their voice heard through rebelling to the status quo. From riots in the States which involved the black community, to Stonewall, to the student protests in Paris in 1968, the recognition of equal human rights echoed. Cornelius Cardew identified certain elitism even among those practicing music. He condemned the rampant classism inherent to the classical orchestra. The Scratch Orchestra involved his experimental music class at Morley College and some of his student at the Royal Academy of Music, where he was Professor, but also admitted untrained musicians, amateurs and those who were unable to read music. Accepting proposals for pieces to perform from any of the members, The Scratch Orchestra advocated thus a greater openness to the practice of music. Such aperture was extended to the music. In the Draft Constitution, published on The Musical Times on June 1969 he writes that each member of the group needed to have a notebook (or 'scratchbook') where to annotate musical parts, which they were able to perform for an indefinite amount of time and in the manner that suited them most, including verbally, graphically and musically among others[4]. The disregard of virtuosity, in this context, was aimed at a more direct expression in music, but also as a critique of the alleged technocracy in music.

4. THE COMPOSER'S ROLE

4.1 An Analogy

The Scratch Orchestra seemed to aim at acting as a microcosm of society realizing the tenets of Marxism. The role of music was to offer the opportunity to test possible solutions of organizing society. Such ideal was shared by Daphne Oram who claimed that the arts it was imperative to conduct an investigation by representation of the current and foreseeable social order. Such reflection in the arts would offer a reenactment, which in turn would serve as an opportunity to monitor development and try out possible solutions to actual problems, without involving the whole of society in the consequences [5]. Such declaration seemed to suggest a different understanding of the role of the composer. The image of an artist detached by society to create an art that was pure appeared surpassed. Oram built a system, the Oramics, capable of reproducing line drawings into music, through a computerized process, decades before the digital revolution. This also seemed an opportunity to involve those who had no musical training in composition The experimental composers in England were involved socially, giving their support to the community. For example Davies organized many workshops for children to stimulate their approach to making music, which he deemed free from preconceived notions [Davies, private communication]. Davies himself explored creative ways to compose music by inventing instruments. His first musical artifice, the shozyg, was devised in 1968. Davies's vast collection of purpose built instruments included various objects unusual in the context of musical performance, either used as ready-made instruments or as parts to integrate in his designs. These included kitchenware items, for example egg slicers, and discarded goods, such as bed springs. Springs in particular constituted the resounding body most used in Davies's artifices. According to him they represented an object with incredible acoustic potential, but neglected in the orchestral instrumentarium [Davies, private communication].

4.2 Recycling

Davies attitude towards instrument building reflected the growing attention towards ecology. In fact in those years the effect of the wastage resulting from mass production was highlighted as an issue because of its proven harm to the ambient. Recycling was introduced to contain the irreparable damage. The process of finding object, described by Davies as 'serendipity' [6], became for Davies a method of composition. In the same way as a classical composer picked notes from the available range, Davies made his choice among an array of junk produced by society. Found objects and scores thus became proof that what is ordinarily discarded as useless in the consumerist's attitude can still be used for as high a purpose as producing art. Unlike Duchamp though, Davies's critique was directed towards society at large, rather than a specific group, such as the art world and the creative process, or the music business.

5. IMPROVISATION

5.1 Freedom

Mass consumption was allegedly encouraged by the economical boom which invested the world after World War II. The rise in capital also created a lucrative market for the production of commercial music. Pop music appeared to be an example of manufactured, disposable commodity in the same way as the goods produced in series by other industries. This trend in music had its opponents among the English experimentalists, who sought to practice a music which avoided the formulaic characteristics of mainstream songs. Improvisation seemed to offer an opportunity to avoid clichés given its unprogrammatic nature. In the same manner as it had given jazz players the freedom from exhausted canons, it created a frame within which experimentalists could explore new sonorities with respects of their political and social beliefs.

5.2 Equality

The adoption of improvising techniques allowed the participants to act as composers and performers at the same time. The collective basis upon which extemporizations would take place ensured the equal status of each contributor. It was in this context that many composers collaborated towards a common goal, despite their different background. For example the Music Improvisation Company involved the guitarist Bailey, the saxophonist Parker and the Davies on live elctronics. Live electronics was an umbrella-term that comprised the amplification of objects and the use of unusual instrumentation, at which Davies had become a specialist. Prévost accounts the advantages of adopting improvisation over a more conventional method of composition by stressing the communicative interrelationship between musicians, engendering an unprecedented creative input by its participants. He defines this quality 'dialogical', stressing the openness to individual expression within a collective environment, and suggests this form of composing music as a viable educational tool in the teaching of music to younger people [7]. The openness to resources in this musical practice earned itself the appellative of 'free', since a diversity of techniques and methods were accepted as well as sounds. With regards to this current, which developed towards the late 1960s, Ben Watson comments: "Free improvisation represents a practical attempt at a universal musical language so far unsurpassed. Like Finnegans Wake or the writings of Karl Marx, the legibility of Bailey's oeuvre is dependent less on its finely wrought and conscientious inner structure (…), than on the potential of society to recognize the substantial freedoms capitalism's conquest of nature implies- but which its social relations deny" [8].

6. AMPLIFICATION

6.1 Integration

Such seemed the spirit of the time in advocating for integration of different ethnicities in a growing multi-cultural society. The emphasis was on unheard-of sounds, symbolizing the new, exotic and in certain respects the previously silent. With this regard amplification constituted a vehicle to give prominence to acoustic events of reduced intensity. The advances in technology made this option readily available. Amplification defined the sound of rock'n'roll, which exploited this process to magnify its power and urgency. For the experimental composers in England it came to symbolize a social and political act.

6.2 Silence

The native French poet Henri Chopin, who lived in England since 1968, identified a political valence to silence. In fact he claimed that, thanks to this medium, the hardly audible sound, such as a leaf falling from a tree, can be heard, and can be made to matter. He concluded that silence was a misconstruction by authoritarian governments to serve their interests. He envisaged silence as synonymous with repression of minorities and the obliteration of their voice [Chopin, personal communication]. Such disregard convinced him to leave France on a voluntary exile in protest of the current political administration. His understanding of silence was therefore deeply rooted in social and political dimension. In his work, such as Sol Air (1961-64) [9] amplification of words, of the human voice, is directed to expand utterances into sweeping sounds, walking a fine balance between poetry and music. John Cage also described his experience of the state of silence, during a visit to an anechoic chamber. He also agreed that acoustic events were inescapable although low in intensity, but he seemed to relate the phenomenon solely to music and how it could actually escape historicization, by deducting his theory of non-intentional sounds [10].

7. CONCLUSION

It seems apparent that the sort of commitment of the composers mentioned was broader than seeking emotional expression in music and more involving than supporting a cause, going as far as investing their aesthetic decisions. This seemed to transform music from an art detached from the world, an art for art's sake, to a vehicle for expressing social and political concern, from an individual and collective point of view. Education, political theories and social organization were reviewed, practiced, experimented through their music and methodology. Such endeavors call for recognition of their work as fruit of the times, deeply embedded with the events that characterized the period. By 1972 the choices these experimental composers had made seemed to catch up with them. Amidst a change of mood from the optimism and confidence of the economic boom, which the first oil crisis brought about, the downturn affected the resources independent groups and composers needed to continue performing. Lacking an established network to support them, ensembles like Gentle Fire disbanded. The political ideas which had ignited the formation of The Scratch Orchestra and AMM became object of discord. The increasingly militant Cardew and Rowe, who adhered to Maoism, experienced tension with their associates. As a result the Orchestra collapsed and AMM went through a phase of rupture. Skempton commented how such political ideas from initially liberating had become apparently restrictive of their work [Skempton, personal communication]. Nonetheless many of these composers continued their activity concentrating on their individual development. For example Davies devised many installation pieces where the audience would play his purposely-placed resounding objects. One of his last commissions was in a cave in Slovenia in 2003. On returning from such an experience he commented on how his focus in designing the work was directed towards the approachability to the visitors of the site [Davies, private communication]. Despite all the time past and the many changes, it seemed that music, in his understanding, was still serving the people.

8. ACKNOWLEDGMENTS

Thanks to Dr. John Dack for his support, to Howard Skempton and Michael Parsons for their time. In memory of Hugh Davies and Henri Chopin.

9. REFERENCES

[1] Davies, H., Lawson, J., Regan, M. *Eye Music - the Graphic Art of New Musical Notation: (Exhibition Touring) Sheffield, Mappin Art Gallery, 7 June to 13 July 1986 ... (and elsewhere)*. Arts Council, London, UK, 1986.

[2] Cardew, C. Stockhausen Serves Imperialism. In *Stockhausen Serves Imperialism.* Latimer New Dimensions Limited, London, UK, 1974, 47.

[3] Hugh Davies Collection, National Sound Archive, London, C1193/54, C1193/27.

[4] Cardew, C. A Scratch Orchestra: Draft Constitution. In *25 Years from Scratch* (ICA, London, 20th November 1994). LMC, London, UK, 20-21.

[5] Oram, D. *An Individual Note.* Galliard, London, 1972, 57.

[6] Roberts, D. Hugh Davies: Instrument Maker. In *Contact 17* (Summer '77). Contact Magazine, York, UK.

[7] Prévost, E. Improvisation: Music for an Occasion. In *No Sound Is Innocent.* Copula, Essex, UK, 1995, 167-178.

[8] Watson, B. *Derek Bailey and the Story of Free Improvisation*. Verso, London, UK, 2004, 146.

[9] Hugh Davies Collection, National Sound Archive, London, C1193/93.

[10] Cage, John *Silence.* Boyars, London, UK, 1978.

Wave Field Synthesis by an Octupole Speaker System

Tim Ziemer
University of Hamburg
Institute of Musicology
Neue Rabenstraße 13
20354 Hamburg
Tel.: 0049 - (0)40 – 428 38 4863
Tim.Ziemer@studium.uni-hamburg.de

ABSTRACT
This paper deals with the analysis and reconstruction of wave fields of musical instruments.

Keywords
Acoustics, wave field synthesis, sound field synthesis, WFS, octupole, octahedron, sound radiation, spatial sound radiation of musical instruments, circular microphone array, dsp, digital signal processing, holography, Mathematica.

1. INTRODUCTION
Any musical instrument has its own frequency-dependent sound radiation. Amplitude and phase relations are different in each point in space. Stereophonic and Dolby surround technologies can be used to reconstruct the sound radiation of an instrument for one specific spot in the room, the so-called "sweet spot"[1]. This paper describes our attempt to reconstruct the wave field of instruments for the whole room.

To reproduce the character of the sound radiation for the whole adjacencies it is necessary to measure the wave field of the instrument. This is done by a circular microphone array around the instrument. The wave field consists of amplitudes which can be various in each direction, caused by wave shadows and/or interferences between several oscillating domains of the instrument/speaker/tone generator. The frequency dependent wave field is to be reproduced by loudspeakers.

An octahedron-shaped speaker system, spotted where the instrument was placed, shall lead to the same wave field as generated by the instrument, so a listener cannot differ between the real instrument and the speaker system, not even by moving around in the room. The wave field reconstruction will be done by manipulating one recording separately for all speakers.

The origin of this research was a group work for a seminar. I thank my fellow students for the preparatory work and the support in assembling the equipment. All graphics were designed with Wolfram Mathematica 6.
Correspondences and/or suggestions concerning this article shall be addressed to the Institute of Musicology of the University of Hamburg to forward it to Tim Ziemer.

[1] Cf. R. Rabenstein, S. Spors, *Sound Field Reproduction*, 1095.

2. RESEARCH
2.1 Research Objective
This paper describes the current research to determine the wave field of musical instruments and to reconstruct it by an octahedron-shaped speaker system. It is a research for a seminar, partly falling back on diverse seminar groups' procedures of workings and research with other aims.

The radiation of the speaker system shall be similar to the wave field of the instrument so that one gets the impression of the real instrument being played in one location of the room. Hence the natural spatial sound of the instrument is not limited to one single spot, it is possible to listen to realistic sounding music with a group of people, or even move within the room without losing the realistic spatial character of the instrument.

2.2 Breadboard Construction
2.2.1 The Microphone Array
A circular microphone array is installed in a free-field room. It is specially built for measuring wave fields of musical instruments. It is hung up in plane, parallel to the floor grid in a height of 1m. It has a 1m radius and contains 128 equidistant microphones, one every 0.05m (more accurate: $\sin\frac{\pi}{64} \approx 0.049m$).

2.2.2 The Speaker System
An octahedron-shaped speaker array was specially built for the wave field reconstruction. This platonic solid consists of eight isosceles triangular faces. Each face is parallel to the opposing. The side length is 0.3m, so the height of one pyramid is 0.21m ($\frac{\sqrt{2}}{2} \times 0.3m \approx 0.212m$). All face centers have a 0.12m distance to the center of the solid ($\frac{\sqrt{6}}{6} \times 0.3m \approx 0.122m$). The face centers are also the centers of the loudspeaker membranes with a 0.08m radius. Oscillating all in phase, they create an almost monopole shaped wave field. The opposing speakers oscillating inversely phased, create an octupole-shaped wave field. It is possible to create all its intermediate steps and combinations.
The lower 0.03m of the octahedron is sawed off for the cables to go through and for the speaker array to be mounted on a right circular cylinder with a 0.03m diameter. This pedestal with its quadratic 0.3m² wide base is 0.91m high, so the center of the speaker array is in height of the microphone array.

2.2.3 The Measuring

An instrumentalist is placed with his instrument in the middle of the microphone array. Just after the recording of the microphone array starts, the instrumentalist plays one note. The recording durability accounts to exactly two seconds. The sample rate is 48kHz, the sample depth is 16Bit. So we get 128 recordings à 96 000 Samples.

A two-dimensional measurement of the wave field is realizable. The array records a two second sample of the sound of the instrument but only a one-second cutoff of the signal - the quasi-stationary state of the oscillation - is considered in this research, since the transient state has an innate sound radiation.

2.2.4 The Coordinates

A spherical coordinate system is applied, with the origin in the middle of the circular array. The microphones *Y* and the speakers *X* aren't considered as areas, but as their midpoint spots.

To simplify matters, only eight microphones are to be considered for the wave field reconstruction, equivalent to the speakers, but still all 128 are used for measuring.

The microphone array is in plane, so the polar angle θ is 0 and the radius r is equal for each microphone coordinate (1m). The azimuth angle Φ varies $\frac{\pi}{4}$ from each microphone to the adjacent microphones. The Speaker-coordinates have also equal radii (0.12m), but are arranged in two parallel planes with $\theta_{1-4} = \frac{\pi}{4}$, $\theta_{5-8} = -\frac{\pi}{4}$, and $\Phi_{1,5} = 0$, $\Phi_{2,6} = \frac{\pi}{2}$, $\Phi_{3,7} = \pi$, $\Phi_{4,8} = \frac{4}{3}\pi$.

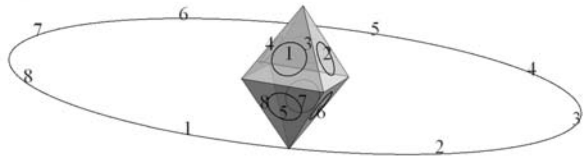


Figure 1. Schematic graphic of the dimensions of the speaker system and the microphone array. The octahedron contains the speakers $X_{i=1,\dots,8}$, the circular array consists of 128 equidistant microphones $Y_{a=1,\dots,128}$. The signal of eight equidistant microphones (each sixteenth) $Y_{j=1,\dots,8}$ ($\triangleq Y_{a=1,17,\dots,113}$) will be used for the reconstruction.

A distance matrix *M* contains the distances between the speakers and the microphones. After transforming the polar coordinates to Cartesian coordinates ($x = r \times \cos\Phi \times \cos\theta$; $y = r \times \sin\Phi \times \cos\theta$; $z = r \times \sin\theta$) the distance can be derived easily by accumulating the norm $M_{ij} = |X_i - Y_j|$. The matrix contains 64 values.

$$M = \begin{bmatrix} |X_1 - Y_1| & \cdots & |X_8 - Y_1| \\ \cdots & \cdots & \cdots \\ |X_1 - Y_8| & \cdots & |X_8 - Y_8| \end{bmatrix}$$

2.3 Disquisition

With the 128 recorded signals, it is possible to analyze the radiation characteristic of the instrument. Since the radiation is dependent on frequency (e.g. low tones are often monopole-shaped, while high tones tend to be more sophisticated), the spectrum is divided into nine separately-examined octave bands $f_{1,\dots,9}$ around the center frequency 62.5Hz (respectively 63Hz) and its octaves (125Hz, 250Hz, 500Hz, 1kHz, 2kHz, 4kHz, 8kHz, and 16kHz). The dominant frequency of each band is determined and stands representatively for the whole frequency region. The speaker system, replacing the instrument, shall reproduce the sound radiation.

2.3.1 Calibration

The microphones are adjusted manually, so they might have a slightly different trim. This deviation is to be corrected by calibrating the array. Therefore a tone generator with a monopole-shaped radiation is placed in the center of the array, to play a reference-tone. Obviously the measured amplitudes would have to be equal, the wave field is supposed to be a circle. Each microphone receives the signal with equal volume. The amplitude deviations are the results of misadjustments, and/or linear distortions or a malfunctioning microphone (which must not be considered for the measuring). Multiplying the measured amplitudes of the frequency (shown in figure 2.a) by their reciprocal values $C_{a=1,\dots,128}$ (see figure 2.b) leads to a circular wave field, so a calibration vector $C_{j=1,\dots,8}$ containing the reciprocal values of the amplitudes at the eight microphones $Y_{j=1,\dots,8}$ is the factor to correct the misadjustment.

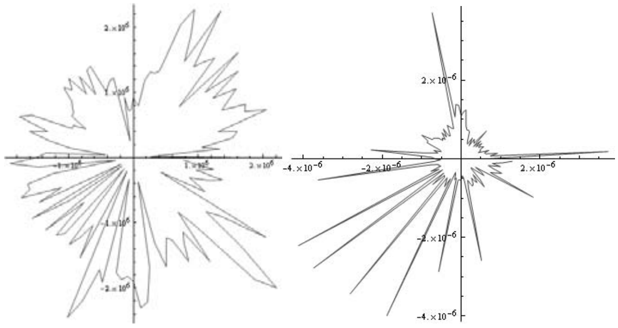


Figure 2.a Amplitude of a 500Hz tone for each microphone. Microphone one is on the positive side of the abscissa, the following anticlockwise to the 128th microphone right below the abscissa.

Figure 2.b shows the reciprocal values of 2.a. Multiplying them leads to a perfect circle (see figure 3.).

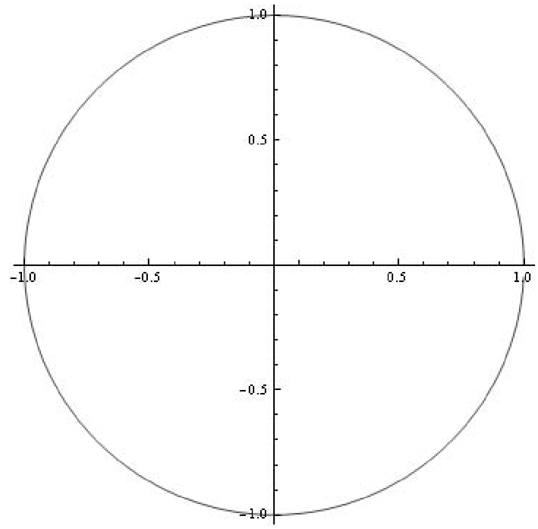


Figure 3. Product of figure 2a and 2b.

2.3.2 Dissolving Away the Speaker Wave Field

Hence the speakers have individual frequency-dependent sound radiation and a wave shadow behind the octahedron-shaped construction and the pedestal, the wave field of the loudspeaker system has to be determined and accounted for the reconstruction of the wave field of the instrument. To identify the radiation, a sinusoidal sound is played on each speaker separately for each frequency band. The product of the amplitudes A' emanating from the Speakers $X_{1,...,8}$, measured at the microphone spots $Y_{1,...,8}$ multiplied by the C vector divided by the measured maximal amplitude affords the speaker radiation matrix R for one single octave band.

$$R_{fij} = \frac{A'_{X_i Y_j} \times C_j}{\max\limits_{i=1,\dots,8;j=1,\dots,8}\left(A'_{X_i Y_j}\right)}$$

The R matrix implicates the C vector and contains 64 values between 0 and 1 which depict the percentage of the volume that achieves the microphones. The rows are the results for one microphone, the columns are the results for one speaker.

$$R_f = \frac{1}{\max\limits_{i=1,\dots,8;j=1,\dots,8}\left(A'_{X_i Y_j}\right)} \begin{bmatrix} A'_{1_{X_1 Y_1}} \times C_1 & \cdots & A'_{1_{X_8 Y_1}} \times C_1 \\ \cdots & \cdots & \cdots \\ A'_{1_{X_1 Y_8}} \times C_8 & \cdots & A'_{1_{X_8 Y_8}} \times C_8 \end{bmatrix}$$

The R-Matrix also implicates the natural amplitude decay which is caused by distance and follows the inverse distance law $\frac{1}{r}$.

2.3.3 Fourier Analysis

Now one frequency of the spectrum is analyzed for the complex amplitude ($B_{1,...,8}$) in each recorded signal. A FFT of a one-second cutout of the recording allows an identification of frequencies with the accuracy of 1Hz and isolates the quasi-stationary from the transient state.

For one microphone, there are 48000 Fourier values with a real part and an imaginary part. The last 24000 values are the conjugate-complexes axis mirrors of the first 24000, only value 1 and 24001 are independent, 24001 is also the axis of reflection. The 2nd value equates to the 48000th, the 24000th equates to the 24002 th. The real parts are simply axis mirrors, the imaginary parts are negative axis mirrors ($\times (-1)$). (For example the 151th value is 0.5 + i150, and the 47851th value is 0.5 - i150. Together these spectral components build a 75Hz frequency.)

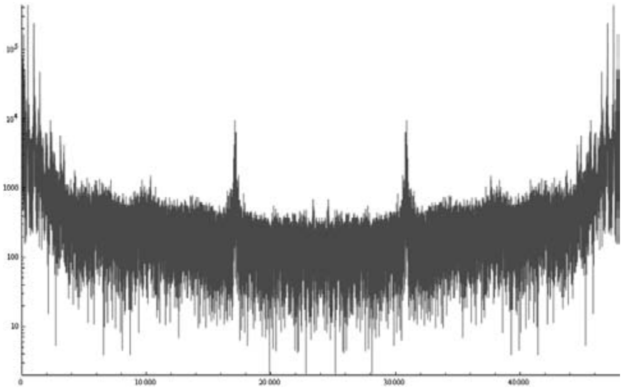


Figure 4. A logarithmical plot showing the absolute values of the Fourier spectrum of a one-second shakuhachi-sample cutoff of one microphone.

2.3.4 The Wave Field

Plotting the product of one specific spectral value from the Fourier transformation of each of the 128 microphones multiplied by the C-value for each microphone in a polar plot, one gets the two-dimensional wave field of the instrument (see Figure 6).

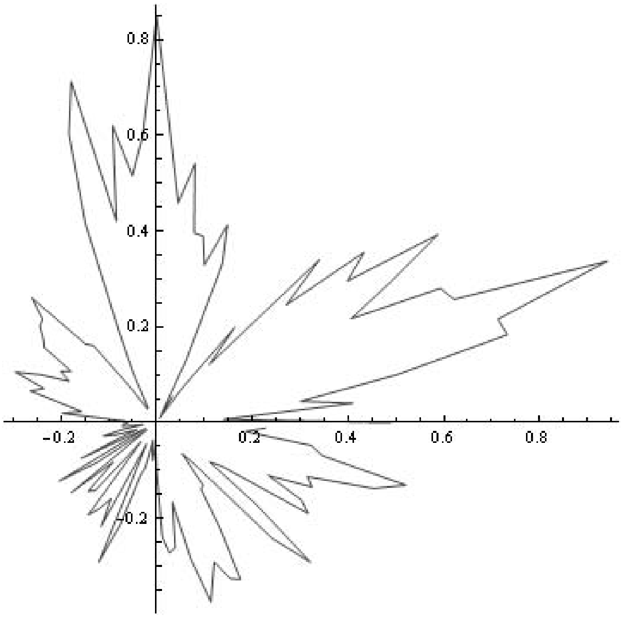


Figure 5. The radiation of a 14421Hz frequency of a shakuhachi without the microphone calibration.

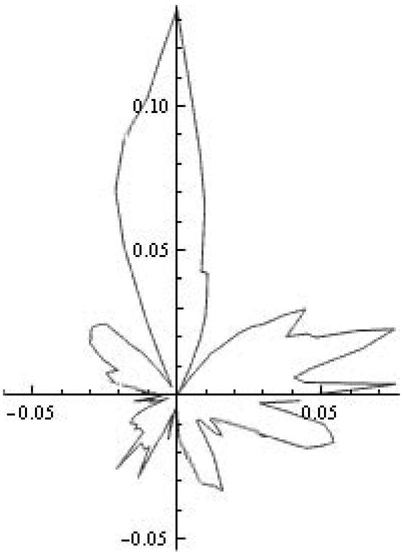


Figure 6. Radiation of the same signal multiplied by the C-Matrix. The malfunctioning microphones have been omitted. One can see that the wave field is abraded.

2.3.5 The Research Core

The determined complex amplitudes ($B_{1,...,8}$) are to be the product of the yet unknown complex amplitude of the speaker system ($A_{1,...,8}$), multiplied by the phase shift PS from each speaker to each microphone (which is dependent on the wave number k_f), due to the travel time of the signal, bearing in mind the amendatory R matrix.

The wave number

$$k_f = \frac{2\pi f}{c}$$

is the quotient of angular frequency and sonic velocity. Diverse distances lead to several travel times from each speaker to the different microphones. So the phase relations are to be a periodic function of wave number and distance matrix:

$$PS_{fij} = e^{ik_f M_{ij}}$$

The K-matrix

$$K_{fij} = R_{fij} \times PS_{fij} \times AD_{ij}$$

combines these values and implies all changes a sound would experience when played by the speakers in the middle of the array. It contains 64 complex values, since the phase shift is an e-function. Multiplying the changes K by the amplitudes of the speakers A leads to the measured amplitudes at the microphones B'.

These amplitudes have to be multiplied by the C-vector to get the real, calibrated values.

$$B_{fj} = B'_{fj} \times C_j$$

$$\overbrace{\begin{bmatrix} K_{X_1Y_1} & \cdots & K_{X_8Y_1} \\ \cdots & \cdots & \cdots \\ K_{X_1Y_8} & \cdots & K_{X_8Y_8} \end{bmatrix}}^{K_f} \cdot \overbrace{\begin{bmatrix} A_1 \\ \cdots \\ A_8 \end{bmatrix}}^{A_f} = \overbrace{\begin{bmatrix} B_1 \\ \cdots \\ B_8 \end{bmatrix}}^{B_f}$$

Nine of these equations are to be formed per instrument (one per frequency band, because the phase shifts and the radiations of the speakers are frequency dependent). The rows of this equation are linearly dependent on each other. Dissolving the equation leads to values for $A_{1,\ldots,8}$.

It is disproportionally more difficult to be solved when more microphones than speakers are used, because the linear system of equations is over-determined. To use more microphones, impreciseness deliverables have to be accepted.

The equation leads to nine complex amplitudes per speaker, respectively to eight complex amplitudes per spectral component. Setting the ignored spectral components 0 and initiating the nine determined amplitudes and their mirrors per speaker (the mirrors with a negated imaginary part), an inverse Fourier transform should create a time signal with the right amplitudes and phases at the eight microphone spots when played by the speakers.

But since the nine spectral components representatively stand for a whole octave, all spectral components of the frequency band have to be multiplied by the complex amplitude, and the axis mirrors by the complex amplitude with a negated imaginary part. For all determined speaker amplitudes per frequency band are multiplied by the same spectral components (the spectrum of the octave band of one (mono-)signal), their relations are not affected. Merely the spectrum is enriched by the further spectral components of the octave bands (so the emerging sound consists of the whole spectral components and not only of nine frequencies).

The modified octave bands and their mirrors have to be stringed together (leaving the first component and the first mirror component unchanged) and retransformed to a time signal by an inverse Fourier transformation. This procedure could be done to reconstruct the wave field of our entire measured and calculated instruments by manipulating any dry recording of the same type of instrument. Only differentiations in the wave field due to playing techniques are not considered yet, as well as the distinctively radiating transient state.

2.3.6 Examination of the Functionality

The sine-in-one-direction-test is one way to find out if the theory practically works. One amplitude B is set to a big value (0.451677 + i109.543) for one frequency, while the others are virtually 0. The achievement should lead to a tone which can be heard clearly in one direction, while the rest of the plane around the speakers remains almost silent.

As a second examination of the achievement, one of the measured violin samples is to be manipulated and played by the speaker system, which is placed in the center of the microphone array. This sound is to be measured and the signals of the eight chosen microphones are to be compared to the recordings of the real instrument. There should be an accurate correlation.

For a representation of a whole instrument it is reasonable to repeat the anterior steps for enough notes to cover the whole frequency range of the instrument.

2.4 Findings

The sine-in-one-direction-test with a great value at microphone one leads to a clear, loud sound in front of the first speaker, whereas the amplitudes in the other directions around the octupole speaker system were clearly lower in plane of the microphone array.

A comparison between the recording of the real and the reconstructed violin shows diverse conformities. The direction where the back of the instrumentalist was has a lack of higher frequencies in both recordings. The frontal area is loud and has a wide spectrum. The volume ratios of the reconstruction don't always correspond to the original recordings.

More measurements and comparisons will be accomplished. Further research shall lead to a more precisely reconstructions and solve weaknesses and problems. The use of more instruments and microphones shall confirm the hitherto findings and show the exactness of the wave field reconstruction. The functionality is not definitely accurate, and potential sources of errors are to be found and nonconformities to be annihilated.

3. PROSPECTS

If the reconstruction of the wave field works in each aspect for the eight measured points, the next step would be an accretion of applied microphones of the array. For the current setup, a reconstruction using all 128 microphones is reasonable if the radiation in between the eight microphones differs too much from the original wave field (thereby one has to accept a fault tolerance, because of the over-determined linear system of equations).

Adding a second, vertical hemicycle-array for a three-dimensional wave field reconstruction, would be the final aim.

Since the transient phenomenon has a differing radiation from the quasi-stationary state, it has to be concerned with separately, and the results merged with the quasi-stationary conditions. That would cover all wave field characteristic of single instruments.

A database integrated in the speaker system, including the phase- and amplitude-shifts of the most common instruments, could enable the speakers to reconstruct a realistic wave field from a mono signal, so no eight-channel soundcard will be needed.

Reproducing the wave field of diverse instruments simultaneously would lead to the impression that these instruments stand at the same spot. Towards this weakness the measuring of whole ensembles or bands would help, and even widen the utility area.

A real time conversion would make the speaker system suitable for everyday use. The speaker could be used to produce music with prospects to utterly new spatial conditions, so it has potential for specially programmed software to produce stereo music ("spatial music") in a whole new way.

This subject has potential for decades of research and development.

4. REFERENCES

[1] Baalman, M. A. J. *On Wave Field Synthesis and electro-acoustic music, with a particular focus on the reproduction of arbitrarily shaped sound sources*, on http://deposit.d-nb.de/cgi-bin/dokserv?idn=989646890&dok_var=d1&dok_ext=pdf&filename=989646890.pdf, 5- 09- 2009, Berlin 2007

[2] Berkhout, A.J., de Vries, D., and Vogel, P. Acoustic control by wave field synthesis, in: *Journal of the Acoustical Society of America. Vol 95, no.* 5 (May 1993), 2764-2777.

[3] Berkhout, A.J., de Vries, D., and Vogel, P. Array technology for acoustic wave field analysis in enclosures, in: *Journal of the Acoustical Society of America. Vol 102, no. 1* (November 1997), 2757-2770.

[4] ChuanXing, B., XinZhao, Ch., and Jian, Ch. Patch nearfield acoustic holography based on the equivalent source method, in: *Science in China Series E. Technological Science*, Hefei 2008, 11-110

[5] Hayek, S. I. Nearfield Acoustical Holography, in: *Handbook of Signal Processing in Acoustics*, 1129-1139

[6] Oikawa, Y.o. Spatial Information of Sound Fields, in: *Springer Handbook of Speech Processing*, New York 2008, 1403-1421

[7] Rabenstein, R., Spors, S. Sound Field Reproduction, in: *Springer Handbook of Speech Processing*, New York 2008, 1095-1114

[8] Rabenstein, R., Spors, S., and Steffen, P. Wave Field Synthesis Techniques for Spatial Sound Reproduction, in: *Topics in Acoustic Echo and Noise Control*, Berlin et al. 2006, 517-545

[9] Rimas, A., Freed, A., Kassakian, P., and Wesel, D. A compact 120 Independent Element Spherical Loudspeaker Array with Programmable Radiation Patterns, on http://archive.cnmat.berkeley.edu/Research/AES2006/speakerarray/icosaspeakers.htm, 5- 09- 2009, Berkeley 2006

Author Index

SysMus09 is a conference of students of Systematic Musicology with focus on the ongoing research developed by PhD and advanced Master students from the SysMus discipline and other disciplines engaged in music research.

It is a privileged forum to disseminate new research initiatives and create international networks of research. In this second edition, we discussed the potential and problems of the interdisciplinary tasks that challenges all spheres of SysMus and most part of the research in music. The conference includes the publication and presentation of peer-reviewed papers, keynotes with top researchers in the area, workshops and social activities hosted in Ghent, Belgium, one of the most enchanting Flemish cities.