



**JIM**  
2013

# Journées d'informatique musicale

13, 14 et 15 mai 2013

# Actes Proceedings

<http://www.mshparisnord.fr/JIM2013/>

Édités par  
Anne Sèdes  
et Alain Bonardi





# Journées d'informatique musicale

13, 14 et 15 mai 2013

# Actes Proceedings

<http://www.mshparisnord.fr/JIM2013/>

Édités par  
Anne Sèdes  
et Alain Bonardi



Maison des Sciences de l'Homme  
Paris Nord

# Comités JIM 2013

## Journées d'Informatique Musicale

### COMITÉ D'ORGANISATION

- Anne Sèdes, CICM, université Paris 8, Maison des Sciences de l'Homme Paris Nord
- Alain Bonardi, CICM, université Paris 8
- Guilherme Carvalho, CIRRA 21, université Montpellier Paul Valéry, chercheur associé au CICM
- Antonio de Sousa Dias, CITAR, université Catholique de Porto, chercheur associé au CICM
- Horacio Vaggione, professeur émérite, CICM, université Paris 8, Maison des Sciences de l'Homme Paris Nord
- Antonio Lai et José Manuel Lopez Lopez, département de musique de l'université de Paris 8, pour le concours ERASMUS

### COMITÉ DE PILOTAGE

- Daniel Arfib
- Gérard Assayag
- Marc Chemillier
- Myriam Desainte-Catherine
- Dominique Fober
- Mikhaïl Malt
- Yann Orlarey
- François Pachet
- Laurent Pottier
- Julien Rabin
- Jean-Michel Raczinski
- Anne Sèdes

### COMITÉ DE LECTURE

- Daniel Arfib
- Nicolas d'Alessandro
- Gérard Assayag
- Nadine Baptiste
- Karim Barkati
- Olivier Baudouin
- Alain Bonardi
- Jean Bresson
- Guilherme Carvalho
- Marc Chemillier
- Thierry Coduys
- Benoît Courribet
- Myriam Desainte-Catherine
- Frédéric Dufeu
- Dominique Fober
- Hugues Genevois
- Vincent Goudard
- Mikhaïl Malt
- Yann Orlarey
- François Pachet
- Laurent Pottier
- Julien Rabin
- Jean-Michel Raczinski
- Anne Sèdes
- Antonio de Sousa-Dias
- Horacio Vaggione

# Sommaire JIM 2013

Comités des JIM 2013 .....	4
Présentation des JIM 2013 .....	5
Sommaire des JIM 2013 .....	6

## SESSION 1 : MODÉLISER ET REPRÉSENTER

**Antoine Vincent, Alain Bonardi et Bruno Bachimont**

*ÉTUDE DES PROCESSUS COMPOSITIONNELS : UN LANGAGE POUR REPRÉSENTER LES PROCESSUS DE PRODUCTION SONORE .....* 9

**Patrick Saint-Dizier**

*QUELQUES ÉLÉMENTS DE LINGUISTIQUE ET DE TRAITEMENT AUTOMATIQUE DES LANGUES POUR UNE MODÉLISATION DU DISCOURS MUSICAL .....* 19

**Raphael Foulon et Jean Bresson**

*UN MODÈLE DE CONTRÔLE POUR LA SYNTHÈSE PAR FONCTIONS D'ONDES FORMANTIQUES AVEC OM-CHANT .....* 31

## SESSION 2 : SYSTÈMES ET RÉALISATIONS

**Matthieu Jacquot**

*GESTION DES FLUX DE MESSAGES ENTRE APPLICATIONS GRÂCE À UN PLUGIN DE SERVEUR OSC DANS LE PROCESSUS COMPOSITIONNEL ÉLECTROACOUSTIQUE ET MIXTE.....* 39

**Florent Berthaut, David Janin et Myriam Desainte-Catherine**

*LIBTUILE : UN MOTEUR D'EXÉCUTION MULTI-ÉCHELLE DE PROCESSUS MUSICAUX HIÉRARCHISÉS .....* 45

**Dorian Cazau, Olivier Adam et Marc Chemillier**

*SYSTÈME DE CAPTATION OPTIQUE POUR LA TRANSCRIPTION AUTOMATIQUE DE LA MUSIQUE DE CITHARE MALGACHE MARDVANY.....* 51

## SESSION 3 : APPRENDRE EN MANIPULANT

**Philippe Galleron et Alain Bonardi**

*CRÉATION MUSICALE ASSISTÉE PAR ORDINATEUR EN MILIEU SCOLAIRE  
PENSER DES CADRES ET DES OUTILS POUR UNE APPROCHE PRATIQUE PAR L'ÉCOUTE ET LA MANIPULATION .....* 59

**Fabrice Guédy, Frederic Bevilacqua, Dominique Fober et Loïc Robert**

*ETUDIER ET JOUER UNE ŒUVRE MUSICALE EN MANIPULANT DES OBJETS .....* 69

**Jules Françoise, Ianis Lallemand, Thierry Artières, Frederic Bevilacqua, Norbert Schnell et Diemo Schwarz**

*PERSPECTIVES POUR L'APPRENTISSAGE INTERACTIF DU COUPLAGE GESTE-SON.....* 77

## SESSION 4 : PARTITIONS INTERACTIVES

### **Sabina Covarrubias Acosta**

*POUR UNE ECRITURE MULTIMEDIA INTERACTIVE DANS LA COMPOSITION MUSICALE*..... 85

### **Dominique Fober, Stéphane Letz et Yann Orlarey**

*PROGRAMMATION ÉVÉNEMENTIELLE DE PARTITIONS MUSICALES INTERACTIVES*..... 91

## SESSION 5 : APPROCHES MUSICOLOGIQUES

### **Pierre Couprie**

*CARTES ET TABLEAUX INTERACTIFS : NOUVEAUX ENJEUX POUR L'ANALYSE DES MUSIQUES ÉLECTROACOUSTIQUES*. 97

### **Felipe Merker Castellani et Anne Sèdes**

*EXTENSION DU MUSICAL VERS L'INTERMÉDIALITÉ : ANALYSES DES APPROCHES DE THIERRY DE MEY ET DE GEORGES APERGHIS*..... 103

### **Laura Zattra**

*LES ORIGINES DU NOM DE RIM (RÉALISATEUR EN INFORMATIQUE MUSICALE)* ..... 113

### **Tom Mays et Pedro Bittencourt**

*MODULATEURS EN ANNEAU ET SAXOPHONE : LE DISPOSITIF D'ÉCRITURE MIXTE ET L'INTERPRÉTATION PARTICIPATIVE DANS L'ŒUVRE LE PATCH BIEN TEMPÉRÉ II* ..... 121

## SESSION 6 : FAUST

### **Romain Michon**

*FAUST2ANDROID : UNE ARCHITECTURE FAUST POUR ANDROID* ..... 131

### **Stéphane Letz, Dominique Fober et Yann Orlarey**

*COMMENT EMBARQUER LE COMPILATEUR FAUST DANS VOS APPLICATIONS ?*..... 137

## SESSION 7 : MISES EN ESPACE DU SON

### **Thierry Dilger**

*GASPR : GRAPHICAL AUDIO SPATIALIZATION PROGRAM WITH REAL TIME INTERACTIONS*..... 141

### **Tim Otto Roth**

*SONAPTICON - SPACE AS AN ACOUSTIC NEURONAL NETWORK*..... 149

### **Otso Lahdeoja**

*SONS SOLIDIENS ET ESPACES AURAUX COMPOSITES*..... 153

## ATELIERS DE PROGRAMMATION

### **Myriam Desainte-Catherine, Nicolas Hincker, Pascal Baltazar et Théo de La Hogue**

*ATELIER AUTOUR DU SÉQUENCEUR I-SCORE*..... 159

### **Sinan Bokesoy**

*DEVELOPMENT OF A MORPHING TOOL FOR THE « COSMOSF » SYNTHESIZER* ..... 161

## INSTALLATIONS ET POSTERS

### **Guillaume Loizillon**

*ZOOPHONIE : NOTES SUR UNE INSTALLATION SONORE ET SES DEVELOPPEMENTS* ..... 171

### **Marc Billon et Anne Sèdes**

*MUSIQUES VISUELLES, COMPOSITION (MUSICALE) AUDIOVISUELLE*..... 177

### **Alexis Crawshaw, Robert Lamp et Nicolas Fdida**

*AVANT-PREMIÈRE DE « LA GALERIE D'ONDES » : UNE INSTALLATION SONORE ET VISUELLE POUR EXPLORER L'INTERSENSORIALITÉ DES INFRASONS*..... 183

## ATELIERS

### **Julien Colafrancesco, Pierre Guillot, Elliott Paris, Anne Sèdes et Alain Bonardi**

*LA BIBLIOTHEQUE HOA, BILAN ET PERSPECTIVES*..... 187

### **Alain Bonardi, Bruno Bossis, Pierre Couprie, Frédéric Dufeu, Mikhail Malt et Laurent Pottier**

*ATELIER : OUTILS POUR L'ANALYSE DE LA MUSIQUE ÉLECTROACOUSTIQUE*..... 199

### **Jean-Baptiste Thiebaut, Sean Soraghan, Lauren J Ianni et Rolet Lamb**

*MULTI-DIMENSIONAL CONTROLLERS, EXPRESSIVITY, SOFTWARE: AN INTEGRATION PROBLEM*..... 201

## CONCERT

### **Nicole Canham, Carlos Lopez Charles**

*AN EVOLVING COLLABORATION: PERFORMER AND COMPOSER APPROACHES TO CREATING VISUAL MUSIC*..... 209

# ÉTUDE DES PROCESSUS COMPOSITIONNELS : UN LANGAGE POUR REPRÉSENTER LES PROCESSUS DE PRODUCTION SONORE

*Antoine Vincent*  
Univ. de Technologie de Compiègne  
Heudiasyc UMR 7253 CNRS  
& IRCAM  
antoine.vincent@hds.utc.fr

*Alain Bonardi*  
Université Paris 8  
CICM EA 1572  
& IRCAM  
alain.bonardi@ircam.fr

*Bruno Bachimont*  
Univ. de Technologie de Compiègne  
Heudiasyc UMR 7253 CNRS  
bruno.bachimont@utc.fr

## RÉSUMÉ

Les pratiques de composition ont évolué avec les nouvelles technologies, ce qui n'est pas sans poser des problèmes de représentation des œuvres sonores : le manque d'écriture abstraite d'un objet sonore peut empêcher sa rejouabilité dans le respect des intentionnalités du compositeur. La méthode de préservation actuelle qui reste la référence est la mise à jour régulière et perpétuelle des œuvres afin d'en assurer la rejouabilité. Notre idée est d'accompagner ces mises à jour en proposant un accès intelligent aux informations nécessaires. Pour cela, nous avons créé un langage de représentation des processus de production sonore, riches en informations intéressantes afin de capter certaines intentions du compositeur. Ce langage, réalisé dans le cadre du projet Gamelan<sup>1</sup>, avec l'aide d'experts du domaine et créé à partir de plusieurs cas réels, est entouré d'outils de suivi de production et de navigation au sein d'un flux de données, en vue de proposer à la fois l'acquisition et le tri des données récupérées durant la production sonore, ainsi que l'extraction et la navigation au sein des informations représentées dans la modélisation.

## 1. INTRODUCTION

La production sonore est en perpétuelle évolution : de tout temps, des nouvelles pratiques ont émergé, n'ayant pour seules limites que celles des compositeurs. Durant le dernier siècle, les technologies ont donné de nouvelles opportunités de création d'œuvres musicales : l'alliance séculaire entre le son produit et sa représentation est désormais rompue, car il est dorénavant possible de travailler directement les matériaux sonores sans utiliser les notes.

Cette évolution amène les spécialistes du domaine de la production musicale à chercher comment retrouver une représentation qui permettrait de décrire comment reproduire un son créé directement. Cette question se pose notamment au moment où les outils sont frappés d'obsolescence et difficile à remobiliser : il faut savoir comment prendre en compte l'évolution des techniques qui

permettent de produire des sons. L'obsolescence technologique rend les objets musicaux virtuels très fragiles dans le temps si nous ne savons pas comment reproduire deux fois le même son, c'est-à-dire s'il n'est pas possible de prescrire l'usage des outils techniques pour le produire.

Dans la première section de notre article, nous présenterons le problème actuel autour du manque de représentation abstraite d'un objet sonore et son impact sur la préservation des œuvres musicales contemporaines, puis notre méthodologie de création d'un langage permettant de représenter la production d'une œuvre. Dans la deuxième partie nous exposerons les étapes de réalisation de notre travail et notamment les différents aspects de base du langage. Enfin nous terminerons sur un exemple d'utilisation du langage pour représenter certaines informations provenant d'une production sonore afin d'illustrer la portée actuelle du travail réalisé et justifier l'intérêt de notre approche.

## 2. MOTIVATION : LE BESOIN DE REPRÉSENTATION

### 2.1. La production sonore aujourd'hui

#### 2.1.1. Contexte

La production sonore a largement évolué ces dernières décennies. L'introduction des technologies, puis leurs évolutions, ont entraîné une transformation des pratiques de composition. Elles ont notamment ouvert le champ des possibles dans la manipulation des sons. Mais leur représentation n'a pas réussi à suivre ce bouleversement : la *partition*, l'un des substrats les plus importants permettant de représenter les notes est en partie caduque : il n'est pas possible de représenter un *son* comme une *note*.

Les piliers importants de la représentation des œuvres musicales sont :

- la *partition*, représentation abstraite nécessaire pour rejouer l'œuvre ;
- les *instruments*, qui peuvent maintenant être électroniques ou numériques, et donc largement soumis à l'obsolescence technologique ;
- la *lutherie*, pratique permettant d'acquérir et de conserver le savoir-faire autour de la fabrication les

1. Projet ANR Gamelan : un environnement pour la Gestion et l'Archivage de la Musique et de l'Audio Numériques – <http://www.gamelan-projet.fr/>

instruments, mise à mal depuis l'apparition des technologies bien souvent propriétaires ;

- le *conservatoire*, qui enseigne les savoir-faire, notamment l'écriture musicale et l'utilisation des instruments, qui tente de s'adapter aux nouvelles pratiques de composition (exemple du Master Réalisateur en informatique musicale à l'Université Jean Monnet de Saint-Etienne).

Parmi ceux-ci, certains ne sont plus en phase avec les nouvelles pratiques de production sonore. En effet, il faut mettre en perspective la création actuelle : de plus en plus d'acteurs peuvent être impliqués dans une production, dans des lieux géographiques très éloignés ; les home studios permettent aussi la création par des compositeurs moins experts, parfois amateurs, qui peuvent aussi travailler avec des technologies complètement fermées et donc impossible à reproduire et à représenter. Or, ces évolutions des pratiques entraînent une perte des substrats stables dans le temps. La partition ne permet pas de représenter les sons ; il n'est plus possible d'avoir une organologie des instruments, la classification utilisée pour les instruments dit classiques est caduque avec les nouveaux instruments qui se multiplient de manière exponentielle ; et la lutherie ne permet pas d'apprendre à reproduire les instruments technologiques (propriétaires ou non) car ce savoir-faire n'est pas accessible. Les conservatoires évoluent plus rapidement : ils s'adaptent plus facilement aux changements des pratiques, et abordent notamment les difficultés de préservation avec les technologies.

Quand nous parlons de *préservation*, nous parlons d'offrir la possibilité de rejouabilité : ainsi, préserver une œuvre, c'est anticiper une future *re-création*. Dans le cas de la musique dite classique, tout était assez simple, et nous le voyons aujourd'hui. La partition, l'abstraction de l'œuvre rédigée par le compositeur lui-même était garante de l'authenticité, et permettait sans difficulté de relire l'œuvre (à condition de conserver les compétences de lecture et de jeu). Les instruments classiques, même s'ils ont subi des modifications mineures au cours du temps, sont toujours présents et permettent de rejouer les œuvres. Or, comment penser la préservation quand nous perdons l'écriture, ainsi que le savoir-faire sur la création et l'utilisation des instruments ?

### 2.1.2. Constat

Notre idée n'est pas de chercher une alternative aux pratiques de préservations actuelles [4]. Il est commun actuellement d'utiliser le concept de préservation active, c'est-à-dire que la préservation passe par les mises à jour régulières des œuvres : porter les œuvres à chaque nouvelle représentation permet d'anticiper les éventuelles évolutions technologiques afin de maintenir les matériaux numériques opérationnels. À chaque évolution des technologies, les mises à jour régulières permettent d'anticiper les transformations qui pourraient être brutales : le meilleur moment pour faire évoluer l'œuvre est le moment où cohabitent les deux technologies car il est possible de

comparer les deux résultats sonores afin de coller au maximum à l'identité de l'œuvre sonore.

Nous souhaitons ainsi accompagner au mieux les mises à jour. Les constantes re-créations restent la référence des mises à jour, nous voulons simplement aider ces travaux afin qu'ils se déroulent dans les meilleures conditions possibles, car nous voyons qu'aujourd'hui il est parfois difficile d'effectuer les migrations des œuvres. En règle générale, il est toujours possible de rejouer une œuvre, mais le coût peut être important : il est par exemple tout à fait possible qu'il faille faire des choix entre plusieurs rendus sonores, et que par manque d'information le choix effectué ne soit pas celui qui correspondait le plus aux souhaits originaux du compositeur.

## 2.2. La création d'un langage de représentation d'une création sonore

### 2.2.1. Problématique

Notre objectif est de travailler sur les informations nécessaires afin d'accompagner au mieux les migrations ou autres techniques de mises à jour des œuvres car il est nécessaire de disposer des bonnes informations, c'est-à-dire utiles pour des prises de décision. Les informations abstraites et complètes se trouvaient sous la forme de la partition. Or, comment pouvons-nous représenter les informations sans la partition ? Ou encore, quelle alternative pouvons-nous trouver à la partition ? Pour créer cette alternative, que devons-nous tenter représenter ? La réponse première serait de chercher à poser une abstraction de l'œuvre à plat, comme le fait la partition pour la musique plus classique : mais aucune tentative jusqu'alors tentée n'a donné de résultat viable. Certains travaux permettent de représenter une part de l'œuvre (ASTREE<sup>2</sup>) mais pas la totalité [3]. Nous préférons nous concentrer sur la documentation des processus de production.

Le processus de production d'une œuvre sonore est riche en informations : c'est pendant la composition que les informations sont les plus accessibles, les plus objectives et les plus complètes possibles. C'est durant ce processus que nous aurons accès au compositeur, et à toutes les traces extractibles des différentes phases de la création de l'œuvre. Il est possible d'extraire du processus de production des données, automatiquement ou non, qui permettent d'étudier les choix opérés. Il est ainsi nécessaire de travailler sur l'organisation de toutes les traces que nous pouvons récupérer du processus de production, qui peuvent rapidement s'accumuler pour former un ensemble de données très volumineux. Il faut se poser la question des informations que nous souhaitons (et celles que nous ne souhaitons pas) conserver, et travailler sur leur accessibilité afin de les offrir à l'étude pour faciliter la compréhension de la création d'une œuvre.

2. Projet ASTREE : analyse et synthèse de traitements temps réel

### 2.2.2. Objectif

Notre idée concernant la classification et l'accessibilité des informations est la création d'un *langage* qui permettra d'exploiter toutes les traces. Ce langage de représentation du processus de production permettra ainsi de s'intéresser à la traçabilité de l'information. Toute étape du processus pourra être transcrite, réécrite via le langage, pour ainsi indexer l'information de telle manière qu'elle sera réexploitable. Le but restant d'offrir la possibilité d'accéder à l'information afin d'accompagner au mieux les migrations des œuvres ou leur compréhension.

Un langage, qu'il soit informatique ou non, est défini avec trois briques de base (que nous trouvons par exemple dans les principes des compilateurs [1]) :

- le *lexique* : c'est le *vocabulaire*, c'est-à-dire la création des mots, du dictionnaire ;
- la *syntaxe* que nous pouvons assimiler à la *grammaire* : c'est l'art de composer des phrases qui sont correctes au niveau de leur structure, même si elles sont des *non-sens* ;
- la *sémantique* concerne le travail sur le sens : une phrase correcte syntaxiquement doit véhiculer une idée, et être compréhensible, voire même *bien formée*.

Pour élaborer ce langage, nous allons ainsi devoir travailler sur la définition d'un vocabulaire, base essentielle pour être capable d'exprimer des informations. Il faut aussi élaborer des règles syntaxiques pour pouvoir relier les morceaux d'informations et les organiser, les préciser ; en bref établir des phrases, une grammaire. Ce langage sera utilisé dans un environnement développé dans le cadre du projet Gamelan qui a pour objectif d'élaborer un méta-environnement, permettant d'extraire les informations durant la production sonore, en suivant le travail des compositeurs et offrant la possibilité de naviguer au sein de la production.

## 2.3. Approche itérative et concrète

### 2.3.1. Méthodologie

Nous souhaitons élaborer un cadre formel pour décrire les œuvres, c'est-à-dire les modéliser. Nous devons ainsi poser les bases nécessaires à la création d'un langage, en commençant par l'élaboration d'un vocabulaire commun aux différents types de production étudiés, puis définir la syntaxe qui permet de construire des phrases porteuses d'informations utiles pour notre objectif de préservation. Mais nous travaillerons aussi à élaborer une manière de prescrire le bon usage du langage ainsi développé, toujours en vue de capter les informations porteuses de sens et supprimer le bruit, comme le ferait un guide de bonnes pratiques.

Pour cela, nous nous attacherons à travailler directement avec les acteurs de la production sonore, suivant une approche itérative et concrète : à chaque œuvre que nous aurons la chance d'étudier, nous remettrons notre travail à plat pour intégrer les nouvelles informations [9]. Il semble

très important aussi d'inclure les experts dans ce travail de représentation, car c'est par un travail collaboratif que nous pourrions isoler les informations qui serviront à la compréhension de la production.

### 2.3.2. Classification des œuvres étudiées

Afin de définir un corpus d'étude, nous allons déjà nous attacher à déterminer un moyen de distinguer les œuvres sonores. Dans le cas de la musique classique, une classification de la production sonore était présente ; or avec les œuvres numériques, nous ne pouvons plus nous fier à une telle organologie. Afin de différencier les différentes productions des œuvres sonores, nous reprenons une idée développée sur la modélisation des dispositifs de musique numérique [7], que nous couplons avec l'idée des processus de production sonore, pour obtenir une classification selon deux axes.

Le premier axe reprend assez intuitivement l'idée du processus : nous pouvons distinguer les différentes étapes d'un processus créatif ou celles des processus potentiels de re-création des œuvres. Nous pouvons dès lors distinguer les différentes étapes des processus de manière assez classique car issues des standards du monde audiovisuel (en adoptant notamment les notions de *pré-production*, de *production* et de *post-production*, historiquement définies comme étant les phases de préparation, de génération des matériaux puis d'activité de montage du produit final).

Le second axe concerne les dispositifs numériques utilisés, en précisant les technologies employées pour la création de l'œuvre, puis les transformations effectuées entre cette création originale et les re-créations suivantes. Nous reprenons trois familles de dispositifs :

- la *mémoire* regroupe ce qui concerne l'utilisation de contenus, comme le travail sous ProTools. Nous pouvons y placer des échantillonnages ou des captations sonores, et même d'autres données ;
- l'*algorithme* représente une catégorie de traitement du signal, comme Max/MSP. La synthèse FM et les transformations temps réel trouvent leur place ici ;
- l'*interface homme-machine* contient les éléments manipulés, pour le déclenchement ou la représentation en situation de concert. Les instruments classiques ou les claviers MIDI utilisés pour réaliser les performances sont des exemples.

Notre classification, succincte, ne tente pas de créer un standard ou même de justifier les choix effectués, mais elle nous permet de chercher des œuvres qui sont différentes, que nous pourrions choisir d'analyser car nous semblent intéressantes dans le but de couvrir notre terrain d'étude.

### 3. LANGAGE : LE COUPLE DIMPO & PATRONS DE CRÉATION

#### 3.1. Digital Music Production Ontology (DiMPO)

##### 3.1.1. Spécifier un vocabulaire commun

La première étape, dans la réalisation de notre travail, consiste à développer le vocabulaire de base. Ce lexique doit être clairement posé, car il doit être commun :

- à des mondes différents en apparence, c'est-à-dire le monde informatique (représentation des connaissances) et musical (pratiques de production) ;
- à des processus de production qui peuvent sembler différents d'une œuvre à une autre (un objet sonore est un prototype [5]).

Notre premier travail consiste à élaborer un vocabulaire qui ne posera aucune ambiguïté entre tous les protagonistes. Nous pouvons citer le simple exemple de *création*, terme portant un sens très précis dans le monde de la production sonore, alors qu'il est très commun pour l'informatique. D'autres termes plus précis comme *patch* ou *montage* sont très différents en fonction du contexte : pour éviter tout problème dans le choix des termes, nous allons travailler la sémantique des termes choisis.

Pour ce travail sémantique, à partir des termes sélectionnés après avoir étudié les flux de production et en avoir abstrait les connaissances qui semblent importantes de conserver, nous allons utiliser les principes différentiels [2] : l'objectif est de classer dans une *taxinomie* tous les concepts présents, ce qui est une première étape nécessaire dans le but de poser une terminologie de la production sonore, mais en ajoutant pour chaque concept quatre principes qui permettent de justifier sa position dans l'arbre créé, tout en le définissant au mieux :

1. la similitude entre le concept fils et le père (*similarity with parent*, SWP) ;
2. la similitude entre le concept et ses frères (*similarity with siblings*, SWS) ;
3. la différence entre le concept et ses frères (*difference with siblings*, DWS) ;
4. la différence entre le concept fils et père (*difference with parent*, DWP).

Nous voyons sur la figure 1 un exemple de hiérarchisation de nos concepts avec l'application des principes différentiels. Ce travail nous permet de justifier l'arborescence que nous développons dans notre taxinomie [8].

##### 3.1.2. Élaborer les règles syntaxiques

À partir de la taxinomie ainsi créée, nous pouvons travailler à créer les syntaxes qui permettront de représenter les connaissances. L'idée ici est le développement d'une *ontologie*, dans laquelle nous ajoutons des propriétés, des relations et des restrictions, qui permettent d'exprimer les informations et de poser la grammaire qui complète le

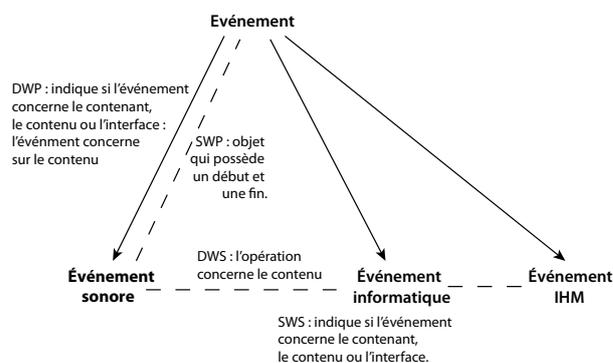


Figure 1. Application des principes différentiels sur l'Événement Sonore.

lexique élaboré précédemment. Par l'étude des pratiques, nous affinons nos choix de concepts en cherchant à spécifier pour chacun d'entre eux ce que nous souhaitons ajouter : nous pouvons préciser des informations (avec une propriété) ou le lier à un autre concept (via une relation). Les restrictions permettent de limiter l'utilisation des relations afin de préciser l'expressivité de l'ontologie au niveau des concepts les plus importants, en éliminant les non-sens les plus évidents. Avec les relations que nous avons ajoutées, nous pouvons par exemple exprimer une relation d'*appartenance* ou de *destination* (figure 2).

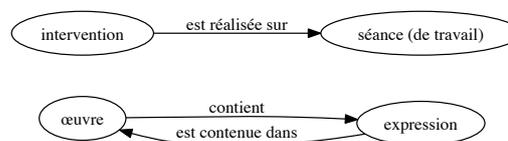


Figure 2. Exemple de relations entre les concepts.

Lors de l'élaboration de l'ontologie DiMPO (*Digital Music Production Ontology*), nous travaillons en nous basant sur de l'existant. Nous avons par exemple intégré les concepts de vCard qui permettent de gérer un carnet d'adresses : ainsi, une part d'interopérabilité est prévue à la base. Nous avons aussi une relation forte entre DiMPO et des concepts de FRBR (spécifications fonctionnelles des notices bibliographiques) : nous reprenons les quatre concepts de base qui permettent de définir une classification des livres (figure 3).

Prenons un exemple pour illustrer le modèle FRBR et la similitude avec les œuvres sonores :

- l'*œuvre* est pour les livres le concept le plus abstrait : *Wizard's First Rule* de Terry Goodking. Nous retrouvons la même chose dans le monde sonore : *Nuages Gris* de Franz Liszt ;

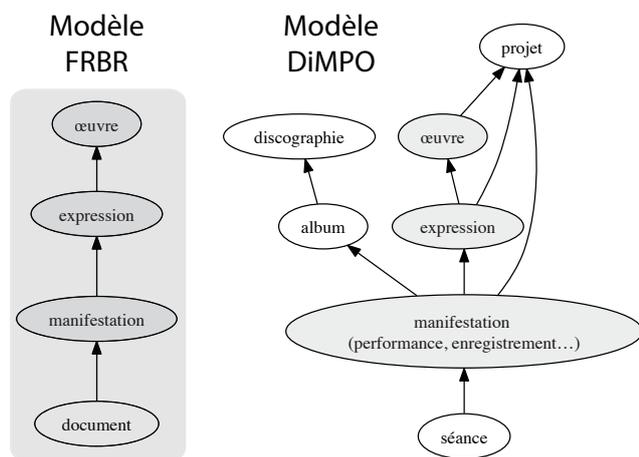


Figure 3. Similitude entre le modèle FRBR et DiMPO.

- l'*expression* est un niveau plus précis présentant l'une des expressions du livre : le texte traduit en français par Jean-Claude Mallé, *La Première Leçon du Sorcier*. Pour le versant musical, nous aurions l'œuvre *Nuages Gris* interprétée par Emmanuelle Swiercz et enregistré en décembre 2010 à l'Espace de Projection de l'IRCAM ;
- la *manifestation* est le concept exprimant la manifestation de la version : le livre sorti en janvier 2003 par l'éditeur Bragelonne avec la couverture de Keith Parkinson. Avec notre exemple musical, nous pourrions parler de la commercialisation de *Nuages Gris* sur le CD *Liszt Voyageur* de 2012 chez le label Intrada ;
- le *document* représente l'exemplaire physique : ici l'exemplaire du livre ou le CD possédé. Dans le cas musical, ce concept n'est pas toujours présent : une œuvre peut se manifester par une performance en situation de concert sans enregistrement. Et l'objet musical n'est pas facilement bornable, c'est d'ailleurs tout l'enjeu de nos travaux ; ce qui explique aussi tous les concepts nécessaires en vue de détailler la production d'une œuvre.

### 3.2. Patrons de créations

#### 3.2.1. Prescrire le bon usage du modèle

Avec DiMPO, nous sommes capables d'exprimer des connaissances en utilisant tous les concepts extraits de notre étude des processus de production. Mais il n'est pas aisé d'utiliser correctement ce cadre conceptuel, notamment lorsque nous n'avons pas participé à son développement. Nous n'avons pas trop restreint son utilisation pour coller avec l'infinité des processus qui existent et laisser la possibilité d'en représenter une majorité. Ainsi nous pouvons nous demander quelle méthode serait nécessaire et intéressante afin de préconiser un bon usage du cadre formel, sachant que nous avons toujours besoin de travailler sur un nombre important de flux de production distincts.

Notre idée est l'utilisation du patron : largement utilisé

dans le monde de la couture, il permet de concevoir un vêtement à partir du modèle. En informatique, la notion de patron de conception (*design pattern*) est très répandue, notamment en développement logiciel : il est issu de l'expérience des concepteurs et décrit, généralement sous forme de diagrammes, les grandes lignes d'une solution, que nous pouvons adapter en fonction des besoins. Ils permettent ainsi de capitaliser les connaissances et préconiser des bonnes pratiques.

Nous reprenons cette idée, déjà portée sous la forme de *patrons d'indexation* qui permettraient de préconiser la bonne indexation de document audiovisuel [6], et créons des *patrons de création* qui eux préconiseront la bonne utilisation de DiMPO en fonction du processus de production et de l'étape au sein de celui-ci. Ces patrons, issus des grandes étapes et de la similitude entre les processus de production, permettent de représenter la généricité extraite des différentes œuvres, et ainsi de construire une modélisation de l'œuvre suivie qui possèdera les informations minimales à connaître.

#### 3.2.2. Graphe-patron

Prenons un exemple de graphe-patron, qui représente une étape extraite du processus de production, que nous pourrions nommer le *montage* d'une œuvre. Cette étape, nous pouvons l'imaginer comme celle qui concerne le mixage : nous avons déjà le matériel de base nécessaire, et nous l'agençons, le corrigeons, le modifions pour générer une première version de l'œuvre. Cette définition peut difficilement représenter tous les types de musiques, mais tente de s'en approcher.

Pour dessiner les graphes, nous commençons par lister les concepts nécessaires pour représenter les informations de l'étape du processus souhaitée en les prenant de DiMPO. Dans notre cas, nous allons avoir besoin de ces termes :

- *objet informationnel*, c'est l'élément de base car c'est celui qui représente l'objet modifié durant cette étape du processus ;
- un deuxième *objet informationnel* qui inclut généralement le premier (la plus petite unité sémantique est en principe un sous-élément d'une unité globale) ;
- un dernier *objet informationnel*, qui est en fait le premier dans un état antérieur : c'est ce qui permet de gérer tous les états de l'objet suivi ;
- des propriétés permettant surtout de *dater* et de gérer des *positions* (surtout temporelles) et/ou des *paramètres* de configuration ;
- l'*objet informationnel* est présent dans un *objet virtuel* ;
- cet *objet virtuel* est lui-même présent au sein d'un *objet processus* ;
- ce *processus* faisant partie d'une *manifestation* ;
- la *manifestation* étant une *expression* d'une *œuvre*.

Avec ces concepts, et via les relations de DiMPO, nous pouvons dresser le graphe-patron représentatif de l'étape (figure 4). Ainsi nous avons le patron qui permet de modéliser l'œuvre et mettre la représentation à jour à chaque

nouvelle action dans cette étape du processus, avec les concepts nécessaires en vue de retrouver une bonne instantiation. Cette dernière portera les informations minimales pour une future exploitation.

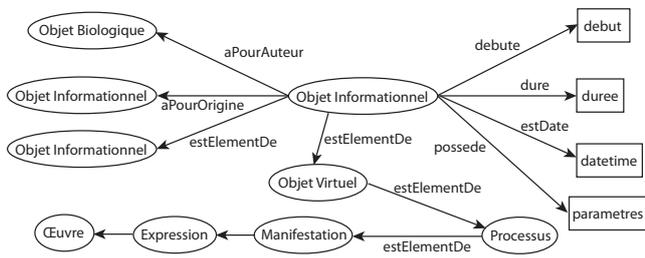


Figure 4. Patron assemblage.

### 3.3. Déclinaison de patrons

#### 3.3.1. Représenter des informations factuelles

Avec le patron précédent, nous avons une couverture transversale des processus créatifs : il fonctionne sur la plupart des cas étudiés. Il est ainsi garant de la généralité des points communs : il réduit les différents cas étudiés et en extrait les similitudes pour assurer que nous sommes capables de capter au minimum les informations essentielles des flux de production. Ce patron utilise les concepts de niveau moyen du modèle (les concepts *noyaux*), qui sont assez abstraits pour s'adapter à différents types d'œuvres.

Certaines limites apparaissent : la capture de la généralité est contraire à la recherche d'informations sémantiquement parlantes, car nous captions des données qui ne sont pas assez porteuses de sens : il faudrait pour cela utiliser les concepts du niveau bas de l'ontologie développée (concepts *domaine*). En utilisant ces concepts, les informations seraient factuelles, bien plus porteuses de sens et donc utiles. L'idée développée va être de spécialiser le patron en fonction du type de l'œuvre ou du processus : en déclinant le patron, il gèrera directement les informations les plus intéressantes : cette idée permet de capturer des données beaucoup plus précises.

Les cas de figures sont assez différents dans ce besoin de déclinaison :

- dans les cas les plus classiques, nous dérivons un patron pour en créer un autre qui servira pour un ou plusieurs types de production sonore ;
- nous aurons parfois uniquement le patron générique, car il répond déjà à des questions assez précises et est applicable sur l'ensemble des types de production ;
- ou nous n'aurons tout simplement pas de patron générique, pour une pratique qui est exclusive à un type bien particulier : le patron sera ainsi créé (dans sa version déclinée) uniquement pour un cas de figure.

Nous remarquons qu'une fois encore nous avons la possibilité de capturer à la fois la généralité et la spéci-

ficité des flux de production, ce qui permet ainsi d'avoir un système à la fois flexible et précis dans son utilisation.

#### 3.3.2. Déclinaison du patron de création

En reprenant le graphe patron de l'assemblage, nous allons examiner une déclinaison pour un enregistrement en vue d'une publication sur support. Après une première étape d'enregistrement en studio, nous avons les différents artefacts qui seront utilisés en vue d'effectuer le montage. Dans ce cas de figure, nous allons chercher à descendre dans les concepts les plus précis possibles de DiMPO à partir du patron générique :

- l'*objet informationnel* central suivi est la *région*, la plus petite entité sélectionnée, la plus précise lors du montage dans un séquenceur ;
- la *région* est un élément d'un autre *objet informationnel* qui est la *piste* ;
- la *région* peut provenir d'une autre *région*, ce qui est particulièrement utile pour avoir un historique des transformations ;
- les propriétés sur les paramètres ne sont pas nécessaires de base, en revanche la *position* et la *durée* sont très importantes ;
- la *région* et la *piste* sont tous deux des éléments du *fichier de session* (un *objet virtuel*) et la *région* étant un élément de la *piste*, la relation de transitivité nous permet de simplifier et de marquer uniquement le lien entre la *piste* et le *fichier de session* ;
- le *fichier de session* est un élément d'un *processus*, plus particulièrement d'une *séance*, une unité temporelle de travail sur l'œuvre ;
- les items *manifestation*, *expression* et *œuvre* ne changent pas.

Nous atteignons un graphe-patron en figure 5 qui permet de travailler avec des informations sémantiquement parlantes et précises, donc beaucoup plus utiles pour étudier et comprendre l'étape de montage sur une œuvre.

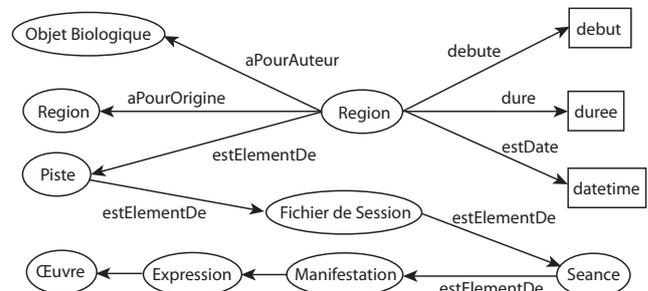


Figure 5. Patron assemblage décliné pour le montage sous un séquenceur.

## 4. EXPÉRIMENTATION : LISZT VOYAGEUR

### 4.1. L'exemple développé

#### 4.1.1. L'œuvre en question

Nous utiliserons l'enregistrement de l'œuvre *Nuages Gris* de Franz Liszt, interprétée par Emmanuelle Swiercz et enregistrée à l'Espace de Projection de l'IRCAM en 2010 pour l'album *Liszt Voyageur*. Cet exemple ne pose pas à proprement parler de problème de préservation mais est un cas très intéressant pour le développement du modèle et des patrons car nous avons eu l'opportunité d'étudier la globalité de son processus de production. Nous avons représenté son champ de couverture sur la figure 6 : nous avons un processus, que nous avons suivi dans sa totalité, qui couvre le domaine de la *mémoire*, car nous avons une mise sur support avec une activité d'enregistrement en studio. Nous avons ajouté sur la figure les documentations générées durant la production (en ne poussant pas la précision pour ne pas cataloguer exhaustivement toutes les informations) et un bilan de ce qu'il reste après la production dans le dossier d'archive constitué initialement.

Nuages Gris E. Swiercz		Création (2010)			Documentation restante
		Pré-prod.	Production	Post-prod.	
Mémoire	Œuvre				
	Documentation générée	- Documentation sur les intentionnalités (emails, enregistrements de répétitions préparatoires, enregistrements de référence choisis) - Montage du projet - Documents organisationnels (plannings, intervenants)	- Partitions annotées - Acquisitions initiales - Fichiers projet - Informations de montage (représentation et évaluation de ce qui a été enregistré, choix de montage) - Photographies & films	- Livret de l'album - Master - Liste du matériel utilisé	- Livret de l'album - Acquisitions initiales - Liste du matériel utilisé - Présence des protagonistes - Notes d'intention
Algo.	Œuvre				
	Documentation générée				
IHM	Œuvre				
	Documentation générée				

Figure 6. Couverture de *Nuages Gris*

Nous pouvons nous demander quelles sont les informations que nous souhaiterions retrouver plus tard, à partir de ce qui a été généré durant le processus et en fonction de ce qui existe déjà dans le dossier d'archive. Si la préservation n'est pas un objectif intéressant, l'étude du processus peut nous permettre de retrouver les souhaits de l'interprétation : quelle était l'idée défendue, l'intention souhaitée dans cette version de l'œuvre ? Pourquoi avoir choisi certains passages plutôt que d'autres ? Quel était l'objectif artistique ciblé ?

Sans la présence des principaux protagonistes, une telle expertise permet à des musicologues de proposer une vision des intentionnalités, ainsi qu'une portée pédagogique sur les choix techniques et sonores dans le montage d'une œuvre musicale : la capitalisation des connaissances à des portées différentes, mais systématiques.

#### 4.1.2. Le fichier de logs

Dans le cadre du projet Gamelan, un logiciel est capable de suivre une production et de renvoyer toutes les

actions effectuées durant le montage d'une œuvre sonore sous la forme d'un fichier de logs avec les informations récupérées automatiquement. Conserver uniquement ce fichier n'est pas très utile : en l'état, il est inexploitable car il est impossible de naviguer dans le document de façon efficace. En effet, il y a beaucoup trop de données, qu'il faut analyser et classer afin d'être compréhensibles. L'utilisation de nos patrons pour ajouter les informations dans une modélisation d'œuvre nous permettra de supprimer le bruit généré par l'excès d'information. Dans l'extrait proposé en figure 7, quelques informations noyées dans la masse des données semblent utiles : nous allons maintenant utiliser le patron sur le fichier pour extraire et représenter les informations qui nous intéressent.

### 4.2. Utilisation du patron de création

#### 4.2.1. Les informations représentées

En utilisant le patron en figure 5, à chaque action nous allons mettre à jour la modélisation de l'œuvre. Sur la figure 8, nous avons les éléments liés à l'import d'un son. Avec l'utilisation du patron, les informations récupérées automatiquement sont ajoutées dans la modélisation. Certaines informations peuvent être ajoutées manuellement (il est difficile de connaître l'identité exacte de l'utilisateur automatiquement). Nous savons ici où est inséré le son et il est raccordé à l'ensemble des concepts qui permettent de savoir précisément dans quelle œuvre et à quel moment de la production nous sommes.

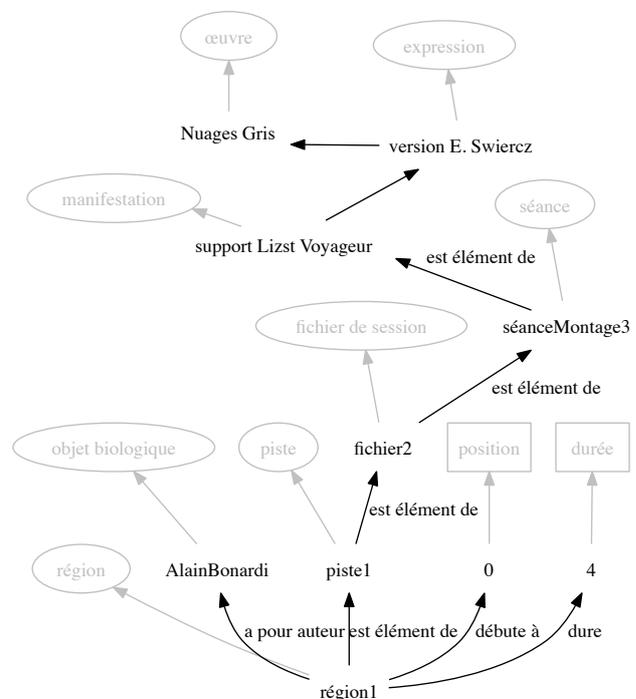


Figure 8. Import d'un son sur une piste du logiciel.

```

@ CurrentSelection Selected Tracks : "piste1", "2_bounceNuagesAmbR", "3_bounceNuagesCpL", "4_bounceNuagesCpR", "5_bounceNuagesApL", "6_bounceNuagesApR"; Begin =
"39.102404"; End = "40.124082"
@ Operation Cut to the clipboard
@ StopPlaying
@ ImportAudio/Users/bonardi/Desktop/GAMELAN_LISZT/1_bounceNuagesAmbl.aup /Users/bonardi/Desktop/GAMELAN_LISZT/6_bounceNuagesApR.wav
@ CurrentSelection /Users/bonardi/Desktop/GAMELAN_LISZT/1_bounceNuagesAmbl.aup Selected Tracks : "piste1"; Begin = "0.000000"; End = "4.000000"
@ Operation /Users/bonardi/Desktop/GAMELAN_LISZT/1_bounceNuagesAmbl.aup Imported /Users/bonardi/Desktop/GAMELAN_LISZT/6_bounceNuagesApR.wav@ CurrentSelection
"777.919861"
@ CurrentSelection /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Selected Tracks : "Audio Track"; Begin = "761.060917"; End = "777.919861"
@ Operation /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Created new audio track
@ CurrentSelection /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Selected Tracks : "Audio Track"; Begin = "761.060917"; End = "777.919861"
@ Operation /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Created new audio track
@ CurrentSelection /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Selected Tracks : "Audio Track"; Begin = "761.060917"; End = "777.919861"
@ Operation /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Created new audio track
@ CurrentSelection /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Selected Tracks : "Audio Track"; Begin = "761.060917"; End = "777.919861"
@ Operation /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Created new audio track
@ CurrentSelection /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Selected Tracks : "Audio Track"; Begin = "761.060917"; End = "777.919861"
@ Operation /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Created new audio track
@ CurrentSelection /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Selected Tracks : "Audio Track"; Begin = "761.060917"; End = "777.919861"
@ Operation /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Created new audio track
@ CurrentSelection Selected Tracks : "piste1", "2_bounceNuagesAmbR", "3_bounceNuagesCpL", "4_bounceNuagesCpR", "5_bounceNuagesApL", "6_bounceNuagesApR"; Begin =
"39.102404"; End = "40.124082"
@ Operation Cut to the clipboard
@ StopPlaying
@ CurrentSelection /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Selected Tracks : "Audio Track", "Audio Track", "Audio Track", "Audio Track",
"Audio Track", "Audio Track"; Begin = "22.804733"; End = "61.907136"
@ Operation /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Pasted from the clipboard
@ CurrentSelection /Users/bonardi/Desktop/GAMELAN_LISZT/repertoirePourSession/montageNuagesGris_02.aup Selected Tracks : "Audio Track", "Audio Track", "Audio Track", "Audio Track"

```

Figure 7. Extraits des logs provenant du montage de *Nuages Gris*.

Sur la figure 9, nous retrouvons l'import couplé à l'utilisation de la commande *couper* qui génère ainsi deux nouvelles régions. Nous avons ici les informations concernant les deux actions : la lisibilité est déjà difficile. Il est simple d'imaginer la complexité des relations que nous aurons après une séance de montage même simple. Mais la différence concrète avec le fichier de logs est l'indexation des informations : avec le langage, nous pouvons écrire des requêtes pour extraire les informations souhaitées en sachant qu'elles sont présentes.

#### 4.2.2. Analyse des informations

En utilisant la figure 9, nous pouvons imaginer les informations extractibles. Nous avons déjà spécifié que cet exemple n'avait pas un grand intérêt dans un souci de préservation (tous les substrats nécessaires pour le rejouer existent); mais il est très intéressant pour tester DiMPO et les patrons car c'est un cas relativement simple et que nous avons bien étudié. Ainsi, nous allons plutôt parler de l'intérêt pédagogique : un musicologue qui veut comprendre les intentionnalités que l'interprète a souhaité exprimer ici.

Nous pouvons, avec les deux seules premières étapes, savoir quels sont les sons choisis (importés) et donc retrouver les fichiers d'origine. Ce qui permet à partir d'une bibliothèque des sons de les comparer avec ceux qui n'ont pas été sélectionnés pour comprendre les choix. Ensuite, dans le montage, la découpe d'une région nous permet aussi de suivre plus finement les passages qui sont été sélectionnés. De la même manière, nous pourrions retrouver les effets appliqués (par exemple, les *cross-fader*) ou les traitements audio pour normaliser les sons.

### 4.3. Discussion

#### 4.3.1. Portée des informations représentées

Les informations sont exploitables grâce à DiMPO qui offre un format de représentation de tous les types d'information et pose les bases d'un langage qui peut être utilisé avec les patrons. Notre liste de patrons est totalement extensible et déclinable : ils sont dissociés de DiMPO. Le modèle laisse assez de liberté pour exprimer des informations qui n'ont pas été pensées à la base, le peu de restrictions n'empêche pas d'ajouter plus de données que prévues initialement. Dans notre exemple, la portée de la préservation n'est pas des plus intéressantes, mais dans d'autres cas de figure, si un fichier ne peut plus être ouvert à cause de l'obsolescence technologique, nous aurons les métadonnées, les interlocuteurs privilégiés, les pièces documentaires centralisées et accessibles, ainsi que des informations sur la production et l'interprétation de l'œuvre.

Les informations que nous pouvons récupérer depuis la modélisation peuvent être très complètes. Il est alors nécessaire d'avoir un moyen de médiation technologique entre cette modélisation et l'utilisateur. C'est l'un des objectifs du projet Gamelan que d'offrir un outil permettant de naviguer au sein d'une production, notamment en élaborant des requêtes qui permettent de poser des questions à partir des patrons existants.

#### 4.3.2. Limites connues

La première limite évidente du travail concerne la couverture des processus : nous sommes conscients de ne pas couvrir la totalité des pratiques de production sonore, mais nous avons tenté de travailler le plus possible avec le monde réel. L'élaboration et l'évaluation de DiMPO ont été faites à partir d'exemples concrets, en tentant de prendre un petit panel d'œuvres assez différentes et potentiellement représentatives des principales pratiques actuelles. Nous avons ainsi travaillé avec les besoins des acteurs du monde de la production sonore, ce qui nous autorise à penser que nous avons au minimum les informations

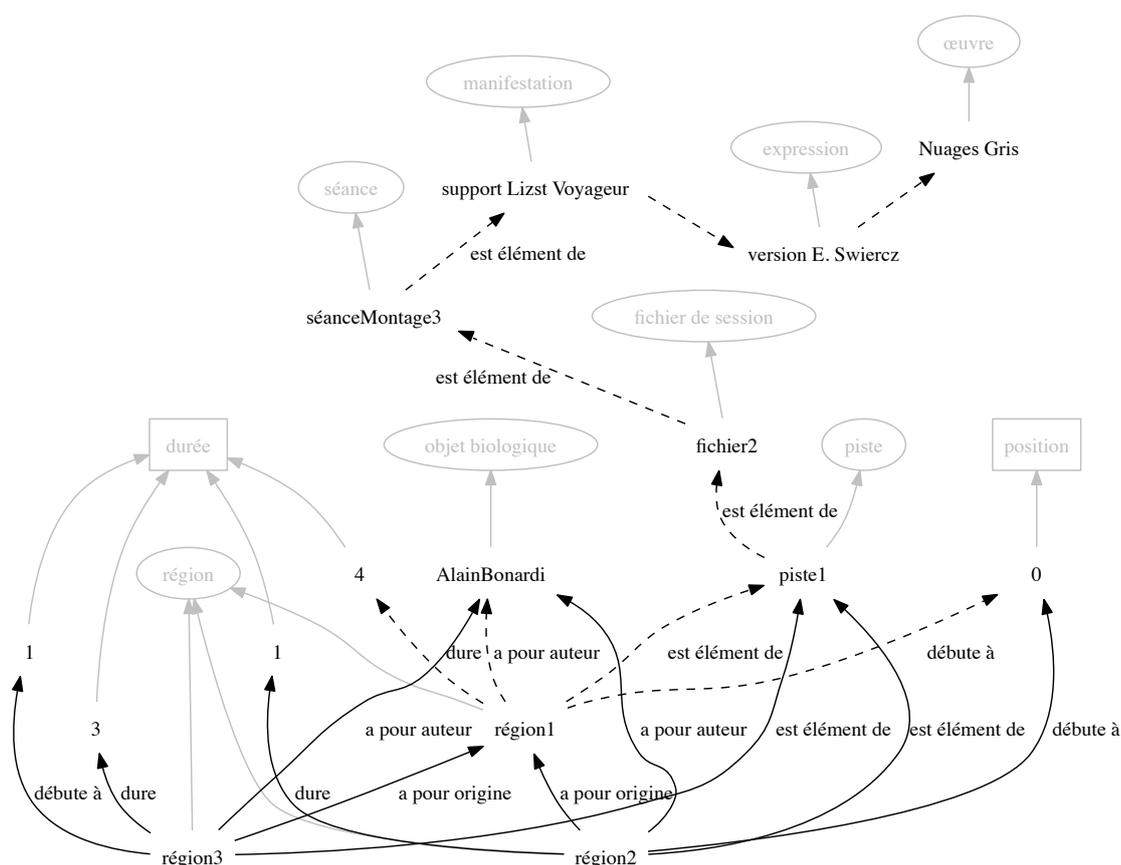


Figure 9. Graphe du modèle instancié avec deux actions (en pointillés l'import et en plein le couper).

essentiels ; le travail de normalisation sémantique permettant aussi de proposer une terminologie compréhensible par tous.

Il est difficile de mesurer directement la puissance de notre solution : ce travail a été réalisé dans le cadre d'un projet borné dans le temps, et notre approche incrémentale, gourmande en temps, mais efficace en qualité, est longue et fastidieuse. La réalisation dans le cadre du projet nous permet d'élaborer les outils permettant l'archivage intelligent des informations et surtout un accès après la phase de remplissage de la modélisation. L'accès aux informations n'est pas forcément le plus aisé, mais l'indexation correcte des données permet d'être sûr, modulo l'adaptation de l'outil, de pouvoir créer des requêtes pour les extraire.

## 5. CONCLUSION

Nous avons présenté les problématiques apportées par les nouvelles technologies, notamment autour de la gestion des informations : elles sont de plus en plus nombreuses, et donc difficiles à archiver correctement. Elles sont d'autant plus importantes qu'elles ont un intérêt grandissant face au recul des formes d'écriture abstraite associées à la partition. Actuellement, pour rejouer une pièce, il est souvent nécessaire de la mettre à jour, pour l'adapter aux technologies du moment : les informations sont alors

la clé pour effectuer une migration dans de bonnes conditions.

Le langage de représentation développé avec l'aide des experts et en étudiant différentes productions en cours, et présenté dans cette article vise à aider l'archivage des données qui seront nécessaires. Il a pour vocation à représenter les informations pertinentes, et aider à les séparer du bruit. Une fois modélisées, les informations sont exploitables, notamment avec les outils développés dans le cadre du projet Gamelan. Les patrons préconisent un bon usage du cadre conceptuel posé par DiMPO, ce qui assure la présence et la bonne classification des données et permet de s'assurer de la qualité des informations enregistrées.

Le travail sémantique effectué sur les concepts du modèle permet de gérer un vocabulaire qui semble commun à différentes pratiques de composition et compréhensible pour toute personne qui étudiera DiMPO. Le langage est peu restrictif, ce qui le laisse ouvert à des usages non prévus lors du développement ou l'ajout d'autres concepts qui sembleraient nécessaires. De même, il est toujours possible d'ajouter ou de décliner des patrons pour prévoir de nouveaux usages. Ainsi nous avons un langage évolutif, qui nous espérons, pourra être à la base de nouveaux systèmes d'archivages qui auront pour vocation à penser la préservation dès le début des productions sonores.

## 6. REFERENCES

- [1] Aho, A. V., Lam, M. S., Sethi, R., Ullman, J. D., *Compilers : principes, techniques, & tools (2nd Edition)*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2006.
- [2] Bachimont, B., *Ingénierie des connaissances et des contenus : le numérique entre ontologies et documents*. Science informatique et SHS, Hermès, Paris, 2007.
- [3] Bonardi, A., "Approches pratiques de la préservation/virtualisation des œuvres interactives mixtes : En Echo de Manoury", *Actes des 17<sup>es</sup> Journées d'Informatique Musicale (JIM 2011)*, Université Jean-Monnet, Saint-Etienne, 2011.
- [4] Bonardi, Barthélemy, J., *The Preservation, Emulation, Migration, and Virtualization of Live Electronics for Performing Arts : an Overview of Musical and Technical Issues*. ACM Journal on Computing and Cultural Heritage, volume 1, issue 1, article 6, New York, 2008.
- [5] During, E., "Prototypes", Entretien avec Franck Madlener, *L'Etincelle*, Paris : Ircam, 2010.
- [6] Isaac, A. Bachimont, B., Laublet, P., "Indexation de documents audiovisuels : ontologies, patrons de conception et d'utilisation", *Acte des 16<sup>es</sup> Journées Francophones d'Ingénierie des Connaissances (IC 2005)*, Nice, 2005.
- [7] Lemouton, S., Ciavarella, R. Bonardi, A., "Peut-on envisager une organologie des traitements sonores temps réel, instruments virtuels de l'informatique musicale ?", *Actes de la 5<sup>e</sup> Conférence de Musicologie Interdisciplinaire (CIM 09)*, Paris, 2009.
- [8] Vincent, A. Bachimont, B., Bonardi, A., "Modéliser les processus de création de la musique avec dispositif numérique : représenter pour rejouer et préserver les œuvres contemporaines", *Actes des 23<sup>es</sup> Journées Francophones d'Ingénierie des Connaissances (IC 2012)*, Centre des Cordeliers, Paris, 2012.
- [9] Vincent, A., Bonardi, A., Rousseaux, F., "Du sauvetage à la préservation des œuvres musicales créées avec dispositif électronique", *Actes des 18<sup>es</sup> Journées d'Informatique Musicale (JIM 2012)*, Université de Mons, Belgique, 2012.

# QUELQUES ÉLÉMENTS DE LINGUISTIQUE ET DE TRAITEMENT AUTOMATIQUE DES LANGUES POUR UNE MODÉLISATION DU DISCOURS MUSICAL

*Patrick Saint-Dizier*  
IRIT-CNRS Toulouse,  
stdizier@irit.fr

## RÉSUMÉ

Dans ce bref article prospectif, nous proposons quelques éléments propres à la linguistique formelle et au traitement automatique des langues qui permettent de modéliser finement plusieurs éléments du discours musical. Nous introduisons les structures de traits typés pour représenter des thèmes. Nous introduisons alors quelques opérations majeures que l'on peut effectuer sur celles-ci dans le cadre de la musique. Nous présentons et adaptons ensuite le cadre de la pragma-dialectique afin de représenter les articulations discursives d'une oeuvre à plusieurs niveaux de détail. Nous concluons sur l'emploi de la plateforme <TextCoop> que nous avons développée pour l'analyse de ces structures musicales.

## 1. INTRODUCTION

Les relations entre les formalismes et les outils d'analyse de la langue naturelle et de la musique ont fait l'objet de plusieurs investigations dans le passé selon le point de vue de la perception (Katz et al 2009) (Jackendoff 2009), ou selon un point de vue syntaxique, en particulier génératif (Jackendoff et al 2005), (Lerdahl et al 1996)). Les différentes analyses montrent, à juste titre, que les formalismes et outils pertinents de part et d'autre sont relativement éloignés et que les analogies (il s'agit bien de deux langages) sont relativement éloignées ou abstraites. C'est particulièrement le cas, par exemple, de la syntaxe et de la morphologie en langue qui trouvent peu de contreparties directes en musique. C'est moins vrai en ce qui concerne l'analyse du discours : l'analyse du discours en langue et celui pratiqué par les spécialistes de l'analyse musicale ont en effet des similarités marquées.

Dans cet article, nous proposons d'abord une modélisation de thèmes (mélodies) par le biais de structures de traits typés largement utilisées dans les formalismes du traitement des langues. Nous montrons que des opérations assez simples permettent de rendre compte des variations apportées à un thème ainsi que des opérations classiques du contrepoint (formes miroir, rétrogrades, par augmentation, etc.). Les exemples étudiés dans cet article sont élémentaires, cependant ce formalisme permet aussi de représenter des structures et des processus plus complexes de la musique contemporaine. En utilisant ces structures

typées, nous montrons à partir de deux oeuvres comment on peut développer une analyse du discours musical sur des niveaux de granularité variable (Peeters et al 2009). Le discours reflète la structure globale d'une oeuvre à travers les relations entre les thèmes et leurs développements. Pour ce faire, nous nous appuyons sur les développements récents de la théorie pragma-dialectique, une théorie linguistique du discours basée sur les interactions entre entités (locuteurs ou bien thèmes ou voix dans le discours musical). Enfin, nous montrons brièvement quelques éléments de modélisation et d'implémentation dans le cadre de notre plateforme d'analyse du discours, <TextCoop>, basée sur la logique et la notion de programme logique.

## 2. LES STRUCTURES DE TRAITS TYPÉS POUR REPRÉSENTER DES THÈMES

Les structures de traits typés sont largement répandues dans les formalismes de description de la structure grammaticale d'une langue. Elles sont à l'origine de plusieurs formalismes basés sur l'unification. Nous montrons ici comment un thème (pour simplifier ici une mélodie, mais toute structure peut être traitée de la même manière) se représente par une séquence de structures typées. Nous montrons ensuite comment des opérateurs se définissent sur une telle structure pour modéliser ou réaliser des opérations classiques : variation, passage au mode mineur, ou bien des opérations du contrepoint (formes miroir, etc.). La définition d'une structure correspondant à une mélodie élémentaire ou à un énoncé élémentaire a fait l'objet de plusieurs études, par exemple (Bimbot et al 2012), (Schauer 2006).

### 2.1. Quelques définitions

Une structure de traits typés est une structure récursive de la forme attribut-valeur où la valeur est un type, qui peut lui-même être une suite de paires attribut-valeur. Dans la langue naturelle, des attributs sont, par exemple, la catégorie syntaxique, la morphologie, qui se décompose en sous-attributs tels que genre, nombre, et voix, temps et mode pour les verbes. La catégorie syntaxique est associée à un ensemble de valeurs qui définissent son type : {nom, verbe, adjectif, préposition, ...}. Une telle structure se représente en général de la façon suivante pour un nom tel que *chaise* :

```
[cat= n,
morphologie = [genre = féminin,
               nombre = singulier],
type-sémantique = objet-physique ]
```

Les attributs peuvent être définis dans un ordre quelconque en respectant les hiérarchies. L'absence d'un attribut signifie dans notre cas qu'il n'est pas pertinent pour ce que l'on représente. Lorsque plusieurs objets sont décrits par une telle structure (par exemple les mots d'une phrase), il est possible d'exprimer des contraintes ou des relations entre ces objets ou bien de gérer ces structures par le biais d'équations, par exemple dans des règles de réécriture. On peut ainsi dire que le genre et le nombre d'un déterminant (ou article) sont identiques à celui du nom qu'il introduit. La même contrainte s'exerce entre le sujet et le verbe d'une proposition.

Les systèmes à base de traits ont fait l'objet d'une théorisation poussée dans un cadre logique (Carpenter 1992), ils sont aussi implémentés dans des langages logiques permettant la description déclarative et à un haut niveau d'abstraction de structures linguistiques. Ces structures peuvent alors être traitées par une grande variété de stratégies d'analyse et d'algorithmes.

## 2.2. Représenter une mélodie par le biais d'une structure de traits

Nous restons ici dans un cadre volontairement très simple au niveau des exemples pour mieux introduire les concepts. Ce cadre peut être étendu à de nombreuses autres caractéristiques (par exemple le timbre ou l'accentuation) selon le même principe. Une mélodie, qui forme un thème, est une séquence de notes ordonnée temporellement, dans le cadre d'un mètre et d'une tonalité donnés.

Un certain nombre de caractéristiques d'une note peuvent être représentées indépendamment du reste du thème. Parmi ces caractéristiques, une structure d'attributs du type suivant (sans vouloir être exhaustif) peut être introduite :

```
[nom , hauteur, durée, ajout-durée,
altération]
```

Le nom est simplement, dans notre exemple, à prendre dans la gamme de do à si, la hauteur est un nombre de 1 à 8 en général qui situe l'octave, la durée est plus complexe : au niveau du type, nous avons une structure ordonnée en progression arithmétique qui va de la ronde (ou la carrée) à la quadruple croche, les triolets ou autres regroupements sont indiqués de la même manière. L'attribut ajout-durée indique les extensions : point, double-point par exemple. Altération comprend le dièse, le bémol, double dièse et bémol et pas d'altération. La première note de la mélodie au soprano donnée en annexe 3 (op. 9 de Brahms) est le fa $\sharp$ , il se représente comme suit, dans un format simple et lisible :

```
note: [nom = fa, hauteur = 5, durée = noire,
sous-durée = aucune, altération = # ]
```

Cette notation peut paraître simpliste par rapport à certains formats proposés par ailleurs et intégrés dans des notations de type XML, par exemple, (Tunrball et al 2008)

ou (Peters et al 2009), ou des éditeurs tels que Lilypond. Toutefois, nous verrons que ce type de structure peut faire l'objet d'abstractions et d'opérations assez puissantes qui permettent, en analyse, de rendre compte de la structure musicale et du traitement des thèmes dans une oeuvre. Employée en génération, cette structure permet aussi de produire une grande diversité de possibilités.

Une mélodie intègre une suite de notes, chaque note étant représentée par le constructeur de type `note` comme illustré ci-dessus. Cette mélodie a elle aussi une structure de traits en propre qui se décompose en deux parties : (1) les caractéristiques globales à cette mélodie, par exemple le mètre, l'armature, le tempo ainsi que le numéro de mesure où elle commence dans l'oeuvre et (2) la séquence de notes. A partir de ces caractéristiques globales, en considérant la séquence de notes, on peut identifier par exemple les temps forts de la mélodie (sur la base du mètre et éventuellement du tempo : un 3/8 lent n'aura pas la même distribution qu'un 3/8 rapide), l'ambitus de la mélodie ou bien encore les modulations à partir de l'armature.

Le constructeur de type est `melodie`. Une structure d'attributs pour une mélodie est, par exemple :

```
melodie: [nom = (chaîne de caractères libre)
mètre , armature, tempo, mesure-début,
liaison, nuance,
seq-notes = [ note: .... ]]
```

La mélodie donnée en annexe 3 (variation 10 de l'op 9 de Brahms), a la structure de traits typée suivante pour ses deux premières mesures :

```
melodie: [nom = 'var10',
mètre = 2/4 , nuance = piano,
armature = 2#, tempo = 'poco adagio',
mesure-début = 1,
seq-notes = [
note: [nom = fa,
rang = 1, % position dans la mélodie
hauteur = 5, durée = noire,
sous-durée = aucune, altération = #],
note: [nom = do, rang = 2,
hauteur = 5, durée = noire,
sous-durée = aucune, altération = #],
note: [nom = ré, rang = 3,
hauteur = 5, durée = noire,
sous-durée = pointée,
altération = aucune],
note: [nom = si, rang = 4,
hauteur = 5, durée = croche,
sous-durée = aucune,
altération = aucune ], etc.  ]]
```

On notera que lorsque l'armature est donnée, il est redondant d'indiquer les altérations pour chaque note, toutefois, en cas de modulation ou d'altération accidentelle, il est préférable de garder cette redondance. La séquence de notes est elle aussi réalisée indépendamment du mètre et du point de début de la mélodie. En cas de polyphonie, on ajoute au niveau de la mélodie la voix concernée. Les silences sont indiqués par la valeur 'silence' au lieu du nom de la note.

Enfin, le constructeur de type `melodie` permet de coder des informations qui portent sur des groupes de notes : liaisons de notes ou nuances par exemple. Si l'on veut décrire un crescendo de la note du rang 1 au rang 4 allant de la nuance *piano* à *mf*, l'attribut `nuance` reçoit alors un type plus complexe au lieu d'une valeur atomique :

```
melodie : [...
nuance = [ note:[rang =1, nuance = p]
           note:[rang =4, nuance = mf]
           intensite = crescendo] ... ]
```

### 2.3. De la mélodie à la polyphonie

Dans le cadre d'une polyphonie (les aspects harmoniques ne sont pas traités ici par manque de place), les constructions présentées ci-dessus s'étendent à une oeuvre entière, horizontalement en traitant la ligne mélodique et verticalement en autorisant l'intégration de plusieurs voix.

Tout d'abord, une voix dans une oeuvre peut être considérée comme une suite de mélodies, avec des numéros de mesure de début qui suivent le déroulement dans le temps. Nous définissons alors le constructeur de type `voix` comme une suite de mélodies. Des relations de discours relient ces mélodies, elles sont présentées ci-après. Si nous reprenons l'exemple en annexe, nous avons :

```
voix :[ ident = soprano,
        mètre = 2/4 , nuance = piano,
        armature = 2#,
        tempo = 'poco adagio',
seq-melodies:
  [melodie: [mesure-début = 1, etc.],
   melodie: [mesure-début = 4, etc.],
   ... ] ]
```

Les identificateurs tels que `soprano` peuvent être adaptés au type d'oeuvre. Le type `voix` hérite a priori des propriétés des types `mélodie` au niveau du mètre, de l'armature, de mesure-debut, etc. Toutefois, si une mélodie est dans un autre mètre ou ton, alors ceux-ci sont précisés à son niveau et l'héritage du niveau supérieur est bloqué. Ceci est une opération classique des structures de traits typés. Des contraintes d'intégrité ou de bonne formation portent sur la séquence de mélodies, par exemple, elles doivent s'enchaîner et non se superposer, ceci est contrôlé sur l'attribut `mesure-début`.

Une polyphonie se définit par un ensemble d'au moins deux voix superposées. Le type `polyphonie` hérite d'attributs globaux : le mètre, l'armature et le tempo (a priori), mais pas la nuance qui peut varier selon les voix :

```
polyphonie :[ ident = 'nom oeuvre',
              mètre = 2/4 , armature = 2#,
              tempo = 'poco adagio',
seq-voix: [voix: [ident= soprano, etc. ],
           voix: [ident= alto, etc. ],
           ... ] ]
```

Très globalement, et à titre d'exemple, on peut définir un canon à 2 voix à l'octave, avec un décalage de une mesure entre les voix en utilisant les attributs `mesure-debut`

avec les valeurs 1 et 2 pour les deux voix si la voix supérieure commence la première et hauteur qui aura les valeurs 5 et 4 respectivement :

```
polyphonie :[ ident = canon2, mètre = 2/4 ,
              armature = 2#, tempo = 'poco adagio',
              seq-voix:[voix: [ident= 1,
                               mesure-debut = 1, hauteur = 5, ...],
                        voix: [ident= 1, mesure-debut = 2,
                               hauteur = 4, ...], ...]]
```

On note que les deux voix ont le même identificateur, ce qui permet de décrire un canon à un haut niveau d'abstraction, adéquat sur le plan de l'analyse musicale. Le formalisme de la réentrance peut être utilisé à ce niveau.

La représentation d'une mélodie ou d'une oeuvre, tout en demeurant assez lourde, offre un niveau d'abstraction intéressant avec les structures de traits. Cette approche permet par exemple de bien mettre en lumière les niveaux de structuration et les éléments partagés. Ces structures peuvent être représentées en XML. Le lien avec des représentations de plus 'bas niveau' comme MIDI, ou plus abstraites (celle de Lilypond par exemple) peut être défini mais reste à élaborer.

### 2.4. Les opérations génériques sur les structures de traits

Plusieurs opérations peuvent être définies de façon abstraite sur une structure de traits, par exemple pour développer des variantes à un thème. Nous en présentons quelques une ici. Ces opérations sont essentiellement des transformations apportées sur la valeur de certains attributs. Pour ce faire, il est nécessaire que ces valeurs soient organisées selon un principe mathématique. Par ailleurs, il est possible de définir des opérations sur ces structures typées à des fins d'analyse (par exemple repérer des modulations) et non de génération.

Les opérations présentées ci-dessous peuvent aussi bien être utilisées en production (produire de nouvelles formes) qu'en reconnaissance (donc analyser des formes existantes et reconnaître les liens entre mélodies ou fragments de mélodies). Les règles décrites sont inversables. L'ensemble est codé en Dislog, un environnement de programmation en logique, développé sur notre plateforme <TextCoop> décrite très brièvement ci-dessous.

#### 2.4.1. La transformation par augmentation ou par diminution

Cette opération, courante en musique baroque, consiste à augmenter ou à diminuer de façon proportionnelle la durée de toutes les notes d'une mélodie. La plupart du temps, on trouve une augmentation d'un facteur 2 (par exemple le premier thème de la fugue en ut $\sharp$  mineur du premier livre du clavecin bien tempéré). L'idée est de donner de l'emphase au thème. Dans les techniques de variation, on trouve aussi des diminutions, souvent de 50%. Dans les variations op 9 de Brahms étudiées ici (annexe 3), la variation 2 prend la basse du thème original comme thème

(c'est aussi le cas de la variation 10 évoquée ci-dessus, ce qui est très rare), chaque mesure du thème initial (en 2/4) est réduite à un temps (dans une mesure en 9/8) ce qui fait une diminution 25%. Enfin, dans la variation 10, les voix intermédiaires reprennent le thème en diminution (25% aussi) avec des notes intermédiaires, dans le style d'un contre-sujet d'un thème fugué).

Si l'on veut doubler les durées, il suffit de passer à la durée immédiatement supérieure (une noire devient une blanche, une note pointée devient doublement pointée, etc.). Le formalisme que nous utilisons (Dislog) offre un langage de spécification de ces transformations qui est déclaratif et permet de se concentrer sur le phénomène à traiter et non sur le calcul. Une règle pour l'augmentation par 2 s'écrit comme suit :

```
note: [durée = X, sous-durée = aucune],
[precede(X1, X)] -->
note: [durée = X1, sous-durée = aucune].
```

Le prédicat `precede` définit X1, la durée immédiatement supérieure à X dans la suite ordonnée des durées. S'il y a une sous-durée, celle-ci est augmentée de la même façon.

#### 2.4.2. Les formes miroir

Les formes miroir sont fréquentes en musique baroque. Etant donnée une mélodie, la transformation consiste à inverser chaque intervalle : une tierce majeure ascendante devient une tierce majeure descendante. Les points de départ des deux formes (directe et miroir) peuvent être distincts. Des corrections peuvent être apportées par mutation si l'on veut préserver la tonalité. Le reste de la structure est préservé (les durées en particulier). Dans les polyphonies baroques, une forme miroir est superposée à sa forme directe (par exemple dans l'oeuvre d'orgue de Buxtehude et de Bach (ex. BWV 668, un choral figuré)). La variation 10 de l'op 9 de Brahms donnée en annexe commence par une mélodie au soprano et son miroir exact à la basse, le soprano commence par un fa $\sharp$  et la basse par un ré. Ce décalage permet de préserver la tonalité de ré majeur et d'assurer une harmonie stable.

La transformation se fait en plusieurs itérations en procédant par paire de notes. Un prédicat `miroir` calcule la note miroir. La transformation se code en Dislog comme suit (les noms de variables sont en majuscules) :

```
melodie: [nom = 'direct', ...,
seq-notes = [ ...
note : [rang = I, nom = N1,
hauteur = H1, alteration = A1],
note : [rang = J, nom = N2,
hauteur = H2, alteration = A2],
... ]],
[J = I+1,
miroir([N1, H1, A1, N2, H2, A2],
[MN1, MH1, MA1, MN2, MH2, MA2])]
-->
melodie: [nom = 'miroir', ...,
seq-notes = [ ...
note : [rang = I, nom = MN1,
```

```
hauteur = MH1, alteration = MA1],
note : [rang = J, nom = MN2,
hauteur = MH2, alteration = MA2],
... ]].
```

Le prédicat `miroir` produit le miroir d'un intervalle, à partir du premier vecteur donné en paramètre.

#### 2.4.3. Les formes inversées

Les formes inversées consistent à jouer une mélodie dans l'ordre inverse des notes. En général, les durées sont assez similaires (par exemple, des croches comme dans l'Offrande Musicale) pour éviter des problèmes de métrique. Une forme directe peut être superposée à sa forme inversée. Sur la base des structures de traits présentées ici, il suffit d'inverser les rangs, le rang 1 devenant le rang N s'il y a N notes dans la mélodie.

#### 2.4.4. Quelques autres transformations

A partir d'une structure de traits, il est possible d'opérer un très grand nombre de transformations plus élaborées que celles présentées ci-dessus. Nous en évoquons quelques exemples ci-dessous.

Etant donnée une mélodie représentée par une structure de traits, il est possible d'y insérer ou d'y ôter des notes supplémentaires, en particulier des notes de passage ou des broderies variées. L'opération se fait sur la base d'une grammaire locale qui définit (dans Dislog) la structure des broderies et autres ornements. Sur les notes existantes, cette grammaire va ajouter des notes et provoquer un décalage de rang et une réduction éventuelle des notes existantes en fonction de la longueur de la séquence que l'on veut insérer.

Il est aussi possible d'extraire un fragment d'une mélodie, entre les rangs I et J, pour en faire un élément autonome qui sera utilisé ailleurs. C'est par exemple le cas de fragments de thèmes utilisés dans des séquences de développement libre entre deux expositions de thèmes dans la forme sonate ou bien comme élément de divertissement entre deux expositions dans une fugue. Des variations de rythmes sont aussi simples à modéliser en modifiant la valeur de l'attribut durée.

L'alternance majeur-mineur peut aussi être traitée en (1) modifiant l'armature de la mélodie de façon adéquate et (2) en contrôlant les notes altérées dans la mélodie. De façon assez similaire, il est aussi possible de transposer une mélodie dans un autre ton.

Enfin, une analyse pas à pas de deux structures de traits correspondant à deux mélodies différentes, via les transformations évoquées ci-dessus, permet d'analyser les ressemblances et différences entre ces mélodies. Une analyse sémantique peut alors être opérée, par exemple (Barrington et al 2008).

### 3. POUVOIR D'EXPRESSION DE CE FORMALISME

Le système formel présenté ici, utilisé largement en syntaxe et sémantique formelle des langues, offre une grande puissance expressive. Une approche logique comme celle proposée dans le langage Dislog permet d'introduire de surcroît du raisonnement.

Nous n'avons pas à ce stade exploré toutes les possibilités expressives de ce système pour la musique, les exemples présentés ici sont demeurés très simples pour leur lisibilité. Il nous paraît acquis que ces structures permettent de représenter des constructions musicales bien plus complexes et diversifiées que celles présentées. De ce point de vue, il nous paraît possible de représenter les structures de la musique classique, mais aussi un certain nombre de structures 'discrètes' de la musique contemporaine. Par exemple, ce formalisme permet de décrire dans un type des **accords, des éléments d'une série, des clusters de notes, des réservoirs de sons** ou tout autre éléments global, ainsi que des formes de glissandi par exemple. La gestion temporelle des unités musicales peut être largement étendue, sans référence à un mètre a priori. Le modèle abstrait des structures de traits peut s'étendre à des types polymorphes ou des types sous-spécifiés, bien adaptés à certaines musiques comprenant une **composante aléatoire**.

Enfin, outre le caractère déclaratif des descriptions, la possibilité de décrire des transformations de façon déclarative et le recours à des contraintes de bonne formation ayant un caractère global en renforce la puissance aussi bien en analyse qu'en production.

Notre travail demeure essentiellement théorique. Il manque à celui-ci une interface élaborée comme celles développées dans PWGL et Open Music qui en donnent toute la facilité d'utilisation. Lilypond nous semble offrir essentiellement un langage d'édition. Il nous est un peu difficile d'opérer une comparaison sur la puissance ou la facilité d'utilisation des formalismes proposés dans chaque système du fait de leur différences en matière de cadre théorique. Le système de représentation que nous développons, basé sur des types, ainsi que les opérations 'logiques' et les transformations proposées nous semblent couvrir une large palette de phénomènes en écriture musicale. L'ensemble du système a été conçu pour permettre une analyse efficace du discours à un niveau d'abstraction adéquat. Sa proximité avec la langue nous semble être aussi un facteur intéressant, et pas seulement pour les musiques vocales, mais par exemple pour l'analyse du discours, des métaphores musicales ou bien des formes de la persuasion par la musique.

### 4. LE NIVEAU DISCURSIF

#### 4.1. Les théories du discours en linguistique

Le niveau discursif est le moins exploré en traitement automatique des langues et en linguistique. Il semble aussi

peu exploré en informatique musicale, alors qu'il est largement développé, mais de façon moins formalisée, en analyse musicale. Dans le domaine de la linguistique, la RST (Rhetorical Structure Theory (Mann et al 1988)) a été la théorie dominante pendant deux décennies. Elle a pour but de représenter les articulations des segments élémentaires (Elementary Discourse Units, EDUs) (Schauer 2006) d'un texte par le biais de relations de discours telles que : illustration, élaboration, justification, cause, conséquence, reformulation, etc. Si, à l'origine, seulement 22 relations ont été postulées, la littérature dénombre à présent environ 200 relations dont les contours sont parfois flous. La RST postule qu'une relation s'établit entre un noyau et un satellite (ou entre deux noyaux dans quelques cas). Par exemple, l'élément illustré (le noyau) est associé à son illustration (le satellite). Un texte est représenté par un graphe de telles relations. Une contrainte de bonne formation stipule que tous les segments de texte doivent être reliés les uns aux autres (Marcu 1997).

Plus récemment, la théorie pragma-dialectique (Van Eemeren et al 1984, 1992), associée à la théorie des actes de discours et à l'argumentation (Perelman 2012) (Moeschler 2002) a proposé un schéma d'analyse du discours basé sur l'interaction et l'argumentation. Cette théorie étudie les échanges selon une perspective descriptive (description des processus linguistiques) et normative (règles évaluant la validité ou la pertinence des énoncés). Cette théorie, en association avec les maximes de coopération de Grice, introduit 5 types d'actes illocutoires gérés par 10 règles du comportement dialogique et argumentatif dont la violation constitue un parallogisme. La théorie des actes de discours implicites (Walton et al 2008), (Reed et al 2007) développe ce point de vue dans des situations de dialogue naturel. Plusieurs relations sont introduites pour caractériser les relations entre locuteurs (apporter des preuves à un énoncé, contredire, etc.), ainsi que les relations entre les structures formelles des énoncés (inférence, contradiction, etc.).

Ces considérations, très brièvement exposées ici, peuvent paraître éloignées du discours musical. Ces théories se prêtent cependant très bien à la formalisation du discours musical via une transposition de leurs principes. De la RST, nous pouvons retenir le modèle noyau-satellite ainsi qu'un certain nombre de relations. Par exemple un thème (le noyau) peut faire l'objet de variations de différentes natures. Ces variations, considérées comme des satellites peuvent être caractérisées par des relations rhétoriques. De ce point de vue, la relation d'élaboration de la RST doit être considérée en musique comme une proto-relation qui doit être décomposée en plusieurs sous-relations. Les opérations d'augmentation ou de miroir peuvent aussi être considérées comme des relations rhétoriques.

Cependant, la RST modélise des relations qui demeurent statiques. Une oeuvre musicale ayant souvent une dimension argumentative ou rhétorique marquée (par exemple les sonates, les symphonies et les quatuors de Beethoven ou l'oeuvre d'orgue de Bach), il est important d'inscrire l'analyse du discours musical dans une perspective dyna-

mique de dialogue et d'argumentation.

Dans les pages qui suivent, et à travers des exemples, nous développons quelques éléments d'une approche d'analyse formelle du discours appliqué à la musique. Les segments du texte musical mis en relation que nous considérons ici correspondent au type *mélodie* décrit dans la section précédente. Cette notion demeure vague quant à sa segmentation automatique dans une partition, c'est aussi le cas des EDUs en langue. Toutefois, l'analyse musicale sait parfaitement isoler des segments musicaux autonomes qui interagissent avec d'autres (par exemple avec des notations telles que *A B A'*, *A B A B A'*, etc. ou les structures de type *da Capo*, *rondo*, etc.). Ce sont ces segments qui seront mis en relation au niveau du discours musical. Pour des segments à forte articulation (par exemple avec des modulations importantes ou des fragments bien marqués), il est possible d'introduire la notion de CEDU (Complex EDU, segment élémentaire complexe) qui permet de caractériser plus finement des parties de ce segment qui seraient développées par la suite.

Il est aussi important de noter que notre analyse peut être horizontale (ou linéaire) ou verticale : si une mélodie est considérée globalement avec son harmonisation, alors les relations de discours seront établies entre segments qui se suivent, directement ou pas. Dans le cas d'une polyphonie, les relations peuvent être établies entre voix, si celles-ci ont une véritable autonomie, comme c'est le cas dans nombre de madrigaux ou, plus tardivement, dans les derniers quatuors à cordes de Beethoven (dont l'op. 135) ou de Bartok.

Les unités du discours musical, tout comme celles du discours en langue, s'identifient et se délimitent sur la base de marques. En langue, il s'agit essentiellement de connecteurs, de classes de verbes, de modaux, de la ponctuation et de la typographie. En musique, il s'agira par exemple de la structure des mélodies, de l'harmonie (cadences), des silences, du rythme ou du mètre. Sur cette base, et sans le développer dans ce bref article, il est possible de définir des règles de segmentation d'une oeuvre en unités élémentaires (voir aussi (Bimbot et al. 2010)). Le formalisme pour la segmentation est le même que celui présenté pour les transformations.

Nous développons à présent deux exemples qui illustrent notre propos. Il s'agit tout d'abord de l'exposition initiale des deux thèmes de la sonate pour piano 'Appassionata' op. 57 de Beethoven où nous avons un traitement linéaire du discours musical (Harding 2010). Nous présentons ensuite l'exemple donné en annexe 3, la variation 10 des variations op. 9 de Brahms sur un thème de Schumann, où est développée une analyse davantage verticale.

Nous donnons ci-dessous un bref commentaire écrit de chaque analyse. L'analyse formelle est donnée en annexe sous forme d'un schéma, comme il est de coutume en analyse des actes de discours en pragma-dialectique. Dans ces représentations, les blocs carrés à droite décrivent un segment musical d'une voix donnée sous sa forme externe, un numéro de voix, en cas de polyphonie, y est associée, les numéros des mesures sont indiqués entre pa-

renthèses. Chaque segment est lié à sa représentation formelle (dans notre cas une structure de trait typée) via un arc qui décrit sa fonction illocutoire (par exemple : question, asserte, réponse, contextualise, soutien, harmonise). D'autres catégories sont en cours d'élaboration, en particulier celles liées à l'argumentation, où deux catégories émergent : support (par justification par exemple ou un autre opération rhétorique du même type) et attaque. Enfin, la force illocutoire n'est pas indiquée ici.

Les représentations formelles sont liées par des relations du type de celles présentées ci-dessus. Des relations peuvent co-apparaître (par exemple tierce inférieure + miroir). Au niveau des segments musicaux originaux, une distinction est opérée entre éléments thématiques, notés *A*, *B*, *A'*, *B'*, etc. et les éléments non-thématiques, mais qui peuvent faire des emprunts aux thèmes, qui sont appelés transitions et notés *Ti*. Si une séquence est complexe, elle peut être décomposée, par exemple, dans le cas qui suit, *A* est décomposée en *A1* et *A2*, car ces fragments sont appelés à évoluer indépendamment. Enfin, les transitions ont bien entendu une représentation formelle et souvent une fonction illocutoire.

#### 4.2. L'exposition de la sonate op. 57 de Beethoven

Cette oeuvre très connue de Beethoven fut composée entre 1804 et 1806, deux ans avant la cinquième symphonie. Ces deux oeuvres ont des liens rhétoriques importants. C'est l'époque du testament de Heiligenstad, début d'un long débat de Beethoven sur le destin, doit-on l'accepter tel qu'il est ou se révolter. Dans plusieurs de ses oeuvres il en résulte un questionnement suivi de réponses, de batailles gagnées ou perdues. L'aboutissement, positif, se trouve dans ses deux dernières sonates, ses 6 derniers quatuors et dans le 9ème symphonie.

Le premier thème, en *fa mineur*, appelé *A*, débute à la mesure 1 (mesure complète hors *anacrouse*). Il dure 4 mesures, mais il est constitué de deux fragments de 2 mesures chacun *A1* et *A2*. *A1* est construit sur une cellule rythmique typique de 3 notes qui est répétée de nombreuses fois et qui sert aussi de motif dans les transitions. *A* est reproduit en *A'*, en *sol♯ majeur*. *A* et *A'* ont la fonction illocutoire *question*, non seulement en raison du contexte de l'oeuvre, mais aussi et surtout en fonction de la forme du thème qui épouse l'intonation d'une expression interrogative en langue. Un tel type d'intonation est retrouvé à de multiples occasions dans l'oeuvre de Beethoven, par exemple, dans l'introduction lente de la sonate pathétique et de la 32ème sonate.

La transition *T1* entre *A* et *A'* est 'vide' (silences). *A2* est ensuite répété 2 fois. Suit une transition, *T3*, où apparaît un fragment nouveau dans le registre grave du piano : le thème du destin, commun avec la 5ème symphonie. Cette transition va contextualiser la question posée par *A* et *A'* et nous ramener en *fa mineur*, ton initial. Suit *A''* (mesures 16 à 23), une amplification de *A*, avec une force illocutoire supérieure (batteries d'accords au lieu d'arpèges, ambitus plus large). Suit une transition, *T5*, puis le second thème *B* et sa variante *B'*. *B* est construit sur la même cel-

lule rythmique que A, mais avec une forme miroir, et un passage au relatif majeur, lab majeur. Nous avons donc ici un miroir et une alternance vers le majeur, renversement de la question en une réponse, ici très lumineuse, chantée presque comme un hymne. Cette analyse est présentée graphiquement en annexe 1 de ce document.

#### 4.3. La variation 10 de l'op. 9 de Brahms

Les variations op. 9 sur un thème de Schumann sont dédiées à Clara Schumann, elle constitue en quelque sorte un témoignage de l'admiration et de l'affection du jeune Brahms pour le couple Schumann. Cette oeuvre date de 1854, elle coïncide avec l'internement de Schuman et reflète ces circonstances douloureuses. Cette série de variations, probablement l'une des plus belles oeuvres de Brahms, emprunte de nombreuses façons d'écrire à Schumann, et on y découvre les deux facettes de ce dernier, Eusébius et Florestan.

La variation 10 est la plus ample de ce cycle. Elle est la seule écrite en majeur (ré majeur), après 8 variations en fa $\sharp$  mineur et la neuvième en si mineur. Elle constitue une pause et un moment de détente dans cette oeuvre par ailleurs assez sombre. Dans cette variation, Brahms a voulu rendre hommage à Bach, celle-ci étant un vrai tour de force contrapuntique.

Cette variation a plusieurs éléments singuliers. Tout d'abord le thème n'est pas celui du thème initial, mais sa basse. Celle-ci est aussi utilisée en diminution comme thème de la variation 2. L'ambitus d'octave du thème traité dans cette variation a une valeur symbolique en musique baroque : l'octave représente allégoriquement le lien entre le ciel et la terre. Ici, avec les formes directe et miroir conjuguées, on a un double mouvement, du bas vers le haut (basse) et du haut vers le bas (soprano). La pièce est de plus peu modulante et donc très stable. Le thème (mesures 1 à 4) est d'abord exposé en même temps au soprano et à la basse en forme miroir tandis que les voix intermédiaires forment une sorte de réponse comme dans une fugue, réponse composée du thème, en diminution et à la tierce majeure supérieure. Les mesures 5 à 8 reproduisent à peu près le même schéma, en se terminant par une modulation en la majeur. De la mesure 9 à 12 (13), le thème et son miroir sont présentés au soprano et à l'alto, et en canon à une mesure, canon reproduit ensuite des mesures 13 à 16. A partir de la mesure 17, le canon se fait entre le soprano et la basse, mais de façon moins explicite pour cette dernière voix. L'oeuvre se termine par l'insertion du fameux thème de Clara, dans un passage plus modulant, où l'accompagnement en arpèges se fait à présent descendant dans le grave du piano.

Dans l'annexe 2 nous présentons des éléments de cette analyse dans le cadre de la pragma-dialectique.

## 5. BREF APERÇU DE <TEXTCOOP> ET DISLOG

Le formalisme des règles et ressources de <TextCoop>, DisLog (Saint-Dizier 2012), étend les possibilités expressives des approches à base d'expressions régulières et les adapte aux besoins de l'analyse de discours, en y intégrant des structures complexes (par exemple, les formes attribut-valeur) ainsi qu'un composant de raisonnement souvent utile dans l'analyse de telles structures. Notre approche s'appuie aussi sur les travaux, maintenant assez anciens, mais toujours actuels, des grammaires logiques (DCGs, XGs, MGs, etc.), qui s'appuient sur des modèles d'exécution inspirés des programmes logiques, dont Prolog. D'autres schémas de stratégie, par exemple à base de contraintes actives peuvent être associés. Cette approche nous paraît intéressante pour l'analyse du discours, en langue ou en musique, en raison de son caractère déclaratif marqué, de son indépendance relative aux stratégies de traitement, et aussi de son aptitude à intégrer naturellement des modules de raisonnement et des structures de contraintes puissantes (par exemple des structures de traits typés, des contraintes d'arbres).

Le formalisme que nous proposons ici permet de coder des règles d'analyse de structures discursives conçues par des linguistes ou des musiciens et codées manuellement aussi bien que des règles issues de mécanismes d'apprentissage à partir de textes annotés, qui produisent en sortie des formes contraintes. Compte-tenu de la taille réduite de ce document, nous ne pouvons présenter ici que quelques éléments simples de <TextCoop>.

Le langage DisLog qui s'exécute sur <TextCoop> comprend les symboles suivants, ils suivent en général la syntaxe de Prolog et des DCGs (grammaires à clauses définies) :

- des symboles pré-terminaux et non terminaux. Les pré-terminaux se dérivent directement en des entrées lexicales ou des expressions (caractéristiques de structures rhétoriques ou de domaines, par exemple). Ces marques d'annotations peuvent faire référence à des structures déjà identifiées. Les symboles non terminaux font appel à des grammaires, essentiellement à caractère local. Les règles ne s'appellent pas entrelées. Les liens entre structures sont réalisés par des opérations de liage sélectif. Les symboles non terminaux et préterminaux sont souvent associés à ou peuvent être directement des structures de traits attribut-valeur.
- des symboles terminaux indiqués entre crochets, cette possibilité est utile lorsqu'il y a peu de choix sur ces terminaux au sein d'une règle, dans le cas contraire, il est préférable de faire appel à un préterminal,
- des indications d'optionalité ou d'itérativité sur les symboles préterminaux et non terminaux,
- des symboles permettant d'exprimer la précedence linéaire,
- des 'gaps' qui représentent des séquences finies de termes qui ne présentent pas d'intérêt pour la règle

en cours de description. Les gaps peuvent être associés à des contraintes, en particulier des symboles terminaux ou non-terminaux qui ne doivent pas être ignorés.

- des appels à des prédicats qui introduisent des contraintes, des connaissances à intégrer ou des calculs divers.
- des fonctions d'assignation, explicites ou par défaut, d'étiquettes dédiées permettant d'étiqueter les structures reconnues avec d'éventuels attributs, calculés par les prédicats ci-dessus.

Les exemples donnés en section 2.4 montrent des emplois directs de Dislog sur des structures de traits typés. Une structure reconnue en partie gauche de règle est transformée, le résultat étant spécifié en partie droite de règle. Les appels à des prédicats Prolog sont intégrés juste avant le calcul du résultat. Pour l'heure, la plateforme TextCoop est dédiée au traitement de la langue, nous travaillons actuellement à une adaptation souple à l'analyse musicale.

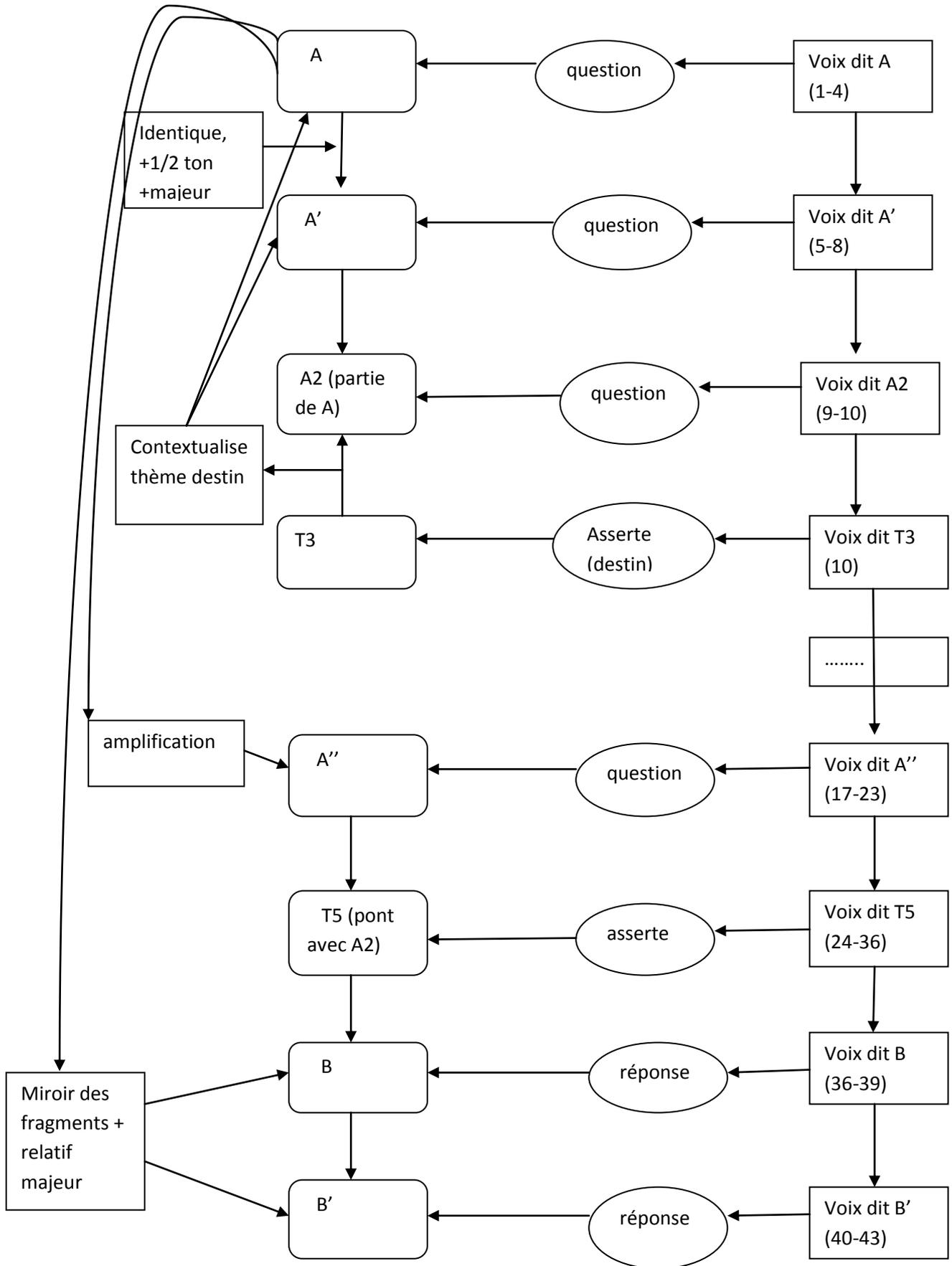
## 6. PERSPECTIVES

Dans ce bref article, nous avons présenté plusieurs formalismes de la linguistique et du traitement du langage naturel qui peuvent être transposés assez directement à l'analyse musicale et à la formalisation du discours musical. Nous avons tout d'abord introduit les structures de traits typées, permettant de représenter des mélodies aussi bien que des oeuvres polyphoniques. Nous avons aussi développé quelques opérations qui s'appliquent habituellement sur des thèmes en montrant leur niveau d'abstraction et d'adéquation descriptive. Enfin, nous avons montré que le modèle des actes de discours et de la pragmatique est adapté à la formalisation des articulations entre thèmes ou fragments thématiques dans le but de rendre compte de la structure et du fonctionnement rhétorique de l'ensemble d'une oeuvre.

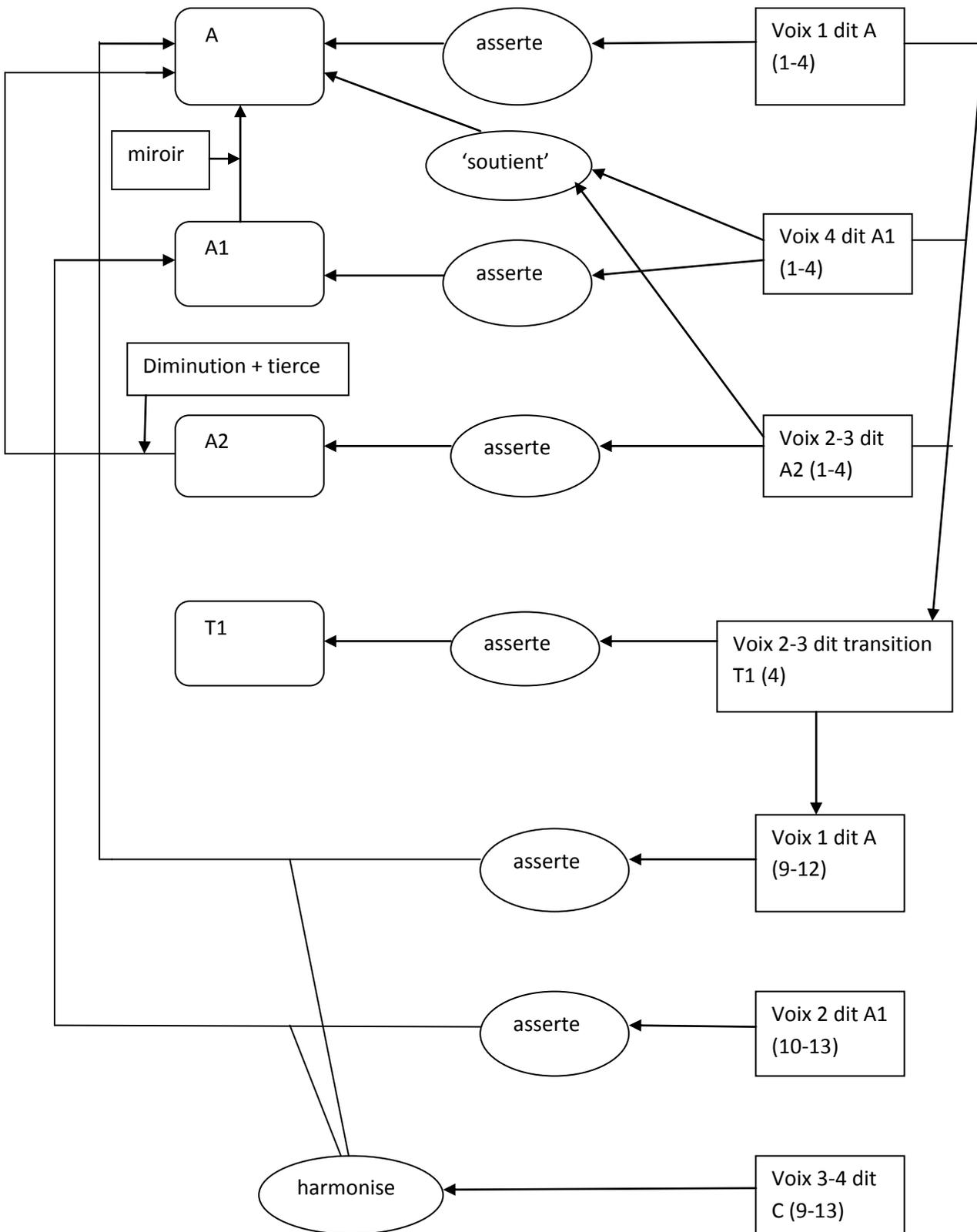
Cet article conserve une dimension prospective. Il est clair que chaque dimension doit être approfondie, validée et testée. Les représentations que nous avons introduites et illustrées par des oeuvres du 19ème siècle sont utilisables et extensibles aux oeuvres antérieures (par exemple le baroque tardif où la symbolique des nombres est importante et doit apparaître dans la représentation rhétorique) aussi bien qu'aux oeuvres plus récentes, où les notions des structures ont évolué à divers degrés (thème cyclique chez Franck ou Bartok, motifs en grappes chez Debussy, citations par exemple grégoriennes chez Tournemire, gestion des séries en musique sérielle, etc.). Enfin, comme rapidement évoqué, au niveau calcul, la plateforme <TextCoop> que nous avons développée pour l'analyse de la langue doit être adaptée, dès lors que les fonctionnalités utiles seront bien identifiées.

## 7. REFERENCES

- [1] Barrington, L., Yazdani, M., Turnbull, D., Lanckriet, G., Combining Feature Kernels for Semantic Music Retrieval, ISMIR, pages 614-619, 2008.
- [2] Bimbot, F., Le Blouch, O., Sargent, G., Vincent, E. Decomposition into autonomous and comparable blocks : a structural decomposition of music pieces, Irisa research Report, 2010.
- [3] Carpenter, R. The logic of typed feature structures, Cambridge university Press, 1992.
- [4] van Eemeren, F.H., Grootendorst, R. *Speech acts in argumentative discussions : A theoretical model for the analysis of discussions directed towards solving conflicts of opinion*, Foris Publications, 1984.
- [5] van Eemeren, F.H., Grootendorst, R. *Argumentation, communication, and fallacies : A pragma-dialectical perspective*, Lawrence Erlbaum Associates, 1992.
- [6] Katz, J., Pesetzky, D. The Identity Thesis for Language and Music, research report, MIT, 2009.
- [7] Harding, H. A. Analysis of Form in Beethoven's Sonatas, General Books, 2010.
- [8] Jackendoff, R. Parallel and non parallel between language and music, *Music Perception*, vol. 26-3, 2009.
- [9] Jackendoff, R., Lerdahl, F. The capacity for music : What is it, and what's special about it ?, *Cognition*, 2005.
- [10] Lerdahl, F., Jackendoff, R. A Generative Theory of Tonal Music, MIT Press, 1996.
- [11] Mann, W., Thompson, S. *Rhetorical Structure Theory : Towards a Functional Theory of Text Organisation*, TEXT 8 (3), pp. 243-281, 1988.
- [12] Marcu, D. The Rhetorical Parsing of Natural Language Texts, ACL, 1997.
- [13] Massin, J., Massin, B. *Beethoven*, Fayard, 1967.
- [14] Moeschler, J. Speech act theory and the analysis of conversation, in Vanderveken D. et al. (eds), *Essays in Speech Act Theory*, Amsterdam, John Benjamins, 2002.
- [15] Peeters, G., Deruty, E. Is Music Structure Annotation Multi-Dimensional ? A Proposal For Robust Local Music Annotation, LSAS, Graz (Austria), 2009.
- [16] Saint-Dizier, P. Processing Natural Language Arguments with the <TextCoop> Platform, *Journal of Argumentation and Computation*, vol 3-1, 2012.
- [17] Perelman, C. L'empire rhétorique : rhétorique et argumentation, Librairie philosophique Vrin, 2012.
- [18] Reed, C., Grasso, F., *Recent Advances in Computational Models of Natural Argument* International Journal of Intelligent Systems, 22(1), 2007.
- [19] Schauer, H. From elementary discourse units to complex ones, Sigdial workshop, ACL, 2006.
- [20] Turnbull, D., Barrington, L., Lanckriet, G. Five Approaches to Collecting Tags for Music, ISMIR, pages 225-230, 2008.
- [21] Walton, D., Reed, C., Macagno, F. *Argumentation Schemes*, Cambridge University Press, 2008.



**Annexe 1 : sonate op. 57 de Beethoven**



**Annexe 2 : var 10 de l'op. 9 de J. Brahms**

Var. 10

Poco Adagio

*p espress. dolce*

*pp e dolciss. l'accompagnamento*

*pp*

*col Pedale*

*p*

*un poco cresc.*

*dim.*

The musical score is written for piano and consists of six systems of two staves each. The key signature is two sharps (F# and C#) and the time signature is 2/4. The piece is marked 'Poco Adagio'. The first system begins with a piano (*p*) dynamic and the instruction 'espress. dolce'. The second system continues with a pianissimo (*pp*) dynamic and 'e dolciss. l'accompagnamento'. The third system features a 'col Pedale' instruction and includes sixteenth-note patterns in the bass staff. The fourth system includes the instruction 'un poco cresc.' and features a triplet in the bass staff. The fifth system ends with a 'dim.' (diminuendo) instruction. The score includes various musical notations such as slurs, ties, and fingerings.



# UN MODÈLE DE CONTRÔLE POUR LA SYNTHÈSE PAR FONCTIONS D'ONDES FORMANTIQUES AVEC OM-CHANT

*Raphaël Foulon, Jean Bresson*

IRCAM - UMR STMS, équipe Représentations Musicales  
1, place Igor Stravinsky, Paris F-75004

## RÉSUMÉ

Récemment, un système de contrôle du synthétiseur CHANT a été implémenté dans l'environnement de composition assistée par ordinateur OpenMusic. Ce synthétiseur est basé sur le modèle de production vocale, et utilise la technique des fonctions d'ondes formantiques (FOF). Nous présentons un ensemble d'outils développés dans le but de fournir un contrôle expressif de CHANT, permettant de formaliser et spécifier des phrases et effets vocaux de manière riche et efficace. Après un rapide état de l'art sur la synthèse vocale et le contrôle de CHANT, nous présentons la modélisation que nous avons utilisée pour synthétiser une phrase vocale, et détaillons l'implémentation de nos procédures de contrôle dans OpenMusic. Nous proposons ensuite quelques exemples d'applications, réalisés lors de la création de l'opéra *Re Orso* de Marco Stroppa, pour conclure avec quelques pistes pour étendre ce travail.

## 1. INTRODUCTION

La synthèse par fonctions d'ondes formantiques (FOF), et en particulier le synthétiseur CHANT, ont connu un essor dans les années 1980, pour ensuite être marginalisés, progressivement éclipsés par d'autres techniques de synthèse telles que la synthèse additive ou les systèmes à tables d'ondes. Basée sur le modèle de la production vocale, cette technique permet la production de timbres voisés de très haute qualité, et peut être aisément étendue à tout autre type de sons. Un système de contrôle de CHANT a été ré-implémenté dans l'environnement de CAO<sup>1</sup> OpenMusic, permettant de structurer les différents paramètres de la synthèse (fréquence fondamentale, ainsi que la fréquence centrale, l'amplitude et la largeur de bande de chaque formant) qui seront transmis au synthétiseur.

Le système CHANT a pour spécificité de présenter une logique de contrôle continue : contrairement à des protocoles plus courants comme CSound ou MIDI, celui-ci s'effectue via un ensemble de paramètres spécifiés à des instants donnés, qui seront interpolés temporellement avant d'être intégrés dans un processus de synthèse. Les protocoles de contrôle définis dans OpenMusic, basés sur la notion d'événements (éléments atomiques du contrôle, localisés et étendus dans le temps), donnent cependant au

1. Composition assistée par ordinateur.

contrôle un caractère discret, plus naturel dans les environnements symboliques et les processus compositionnels.

La contribution que nous présentons ici concerne des procédures de contrôle pour le synthétiseur CHANT, permettant de spécifier avec précision et expressivité les comportements du système lors de la transition entre le domaine discret (OpenMusic) et (pseudo) continu (CHANT). Il s'agit d'élaborer un ensemble de « règles », qui permettront de générer une suite de paramètres correspondant à un effet désiré (production d'une consonne, d'un vibrato...); chaque règle ayant pour but de formaliser un aspect spécifique du processus de synthèse, sur une certaine durée.

Cependant, notre but ne se réduit pas à la reproduction du jeu d'un chanteur : nous chercherons entre autres à nous inspirer des modes de production de la voix dans le but de « dépasser » les propriétés traditionnelles de la synthèse vocale.

## 2. ÉTAT DE L'ART

### 2.1. CHANT et les fonctions d'ondes formantiques

La synthèse par FOF s'inspire du modèle de production de la voix, et permet de reproduire – par exemple – le comportement des résonateurs de l'appareil vocal [8]. Cette méthode de synthèse consiste à générer un ou plusieurs trains de sinusoides amorties, à une fréquence donnée (fréquence fondamentale, ou  $f_0$ ) pouvant être paramétrée. Chaque train de sinusoides contribue au spectre sonore par un *formant*, correspondant à une zone spectrale riche en intensité. Dans le synthétiseur CHANT, implémenté par Xavier Rodet et son équipe [9], il est possible de contrôler cette technique à partir de nombreux paramètres, tels que les fréquences centrales des formants, leur amplitude, leur largeur de bande, temps d'excitation, phase, etc. Dans l'implémentation initiale de la synthèse par FOF, le synthétiseur était associé à l'environnement Formes [10], qui permettait de générer des « phrases » par l'évaluation de processus temporels déterminant les valeurs des paramètres de synthèse. Les processus ainsi créés avaient également pour vocation d'activer et de paramétrer un ensemble de règles codées dans le synthétiseur, permettant d'obtenir des vibratos, de corriger le spectre du signal vocal [1], ou encore de créer des « consonnes » entre les voyelles synthétisées [11].

## 2.2. Production de phonèmes

Parmi les techniques permettant de simuler la voix, et plus particulièrement la voix chantée, nous pouvons citer la synthèse par concaténation de diphones, implémentée dans le programme *Diphone Studio* [12], qui propose de construire des phrases vocales en effectuant des morphings entre plusieurs phonèmes. Une autre approche est l'utilisation de règles élaborées suite à l'étude du fonctionnement des comportements des signaux vocaux, qui permettent de les reproduire au sein d'un processus de synthèse. Le synthétiseur MUSSE (Music and Singing Synthesis Environment) en est un exemple, proposant d'effectuer une conversion texte vers chant, tout en introduisant de nombreuses règles (vibrato, consonnes, intonation) [13].

## 2.3. Contrôle de CHANT dans OpenMusic

Dans un précédent article, deux bibliothèques permettant le contrôle de la synthèse « chant » dans OpenMusic ont été présentées [5] ; la première, *chant-lib*, permettant de reproduire le comportement du synthétiseur à partir d'un modèle basé sur la concaténation d'évènements implémentés dans le langage Csound, et la seconde, *OM-Chant*, utilisant le mode de contrôle « continu » du synthétiseur. Plus récemment, *OM-Chant* a été étendue afin d'intégrer la notion d'évènements définis grâce à des objets musicaux de haut niveau tels que des matrices de paramètres [6]. Ces évènements de synthèse peuvent être des valeurs de fréquence fondamentale ( $f_0$ ), des matrices de paramètres de FOF, des générateurs de bruit ou des filtres formantiques.

Dans ce contexte, qui est le point de départ de notre travail, les comportements pouvant apparaître en cas d'évènements distants, contigus ou superposés restent à concevoir.

## 3. PRODUCTION D'UNE PHRASE VOCALE

Avant d'implémenter de nouvelles procédures de contrôle, nous avons effectué une étude des phénomènes de production vocale, à partir d'analyses temps-fréquence d'extraits audio. Les analyses ont été effectuées grâce au logiciel *Audiosculpt* [3], en utilisant la méthode LPC<sup>2</sup>. Les extraits audio analysés proviennent de sources variées : bases de données de phonèmes, enregistrements d'opéra, ou encore extraits sonores de la pièce *Chréode* de Barrière [2], produits grâce à une précédente implémentation de CHANT, qui témoignent d'une utilisation originale du synthétiseur.

### 3.1. Modèle utilisé

Nous partons d'une modélisation élémentaire de la production vocale, ne considérant que sa composante voisée.

2. Linear Predictive Coding.

Nous dégageons deux phénomènes dans ce modèle<sup>3</sup> : les états stables et les transitions. Le critère de discrimination utilisé pour définir les deux phénomènes est basé sur le comportement des formants. Le sonagramme visible sur la figure 1 présente un exemple sur lequel ces deux états sont identifiés.

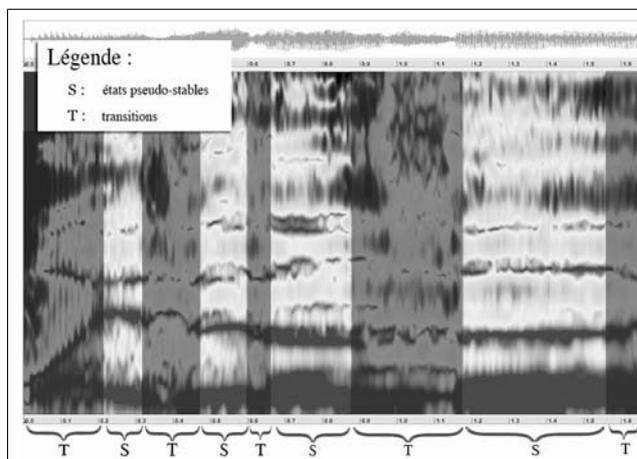


Figure 1. Analyse d'un extrait de voix chantée, identification des états stables (S) et des transitions (T).

Nous définissons un état stable (plus précisément, *pseudo-stable*) comme un état continu dans le temps, au sein duquel les paramètres formantiques n'évoluent pas ou peu : les variations ne doivent pas excéder 10% pour les fréquences centrales, 3dB pour les amplitudes et 20% pour les largeurs de bande. Dans le cadre de la voix parlée, les états stables peuvent être assimilés aux voyelles, ou dans le cas du chant lyrique, aux notes tenues. Cependant, les parties centrales de certaines consonnes voisées, telles que le 'm' ou le 'l', pourront également être modélisées par un état stable, du fait de la faible variation des formants en leur centre. Un exemple sera évoqué dans la partie 3.2.

Pendant l'étape de synthèse sonore, ces états pourront être produits par des FOF aux paramètres formantiques constants. Cependant, une certaine quantité de variations doit être introduite dans ces paramètres pour donner du « caractère » à la voix, de sorte que le rendu sonore ne paraisse pas artificiel. Deux solutions ont été abordées et testées avec succès : la variation de  $f_0$  (modulée aléatoirement ou par un vibrato<sup>4</sup>), et la variation aléatoire des paramètres de FOF.

Par opposition aux états stables, les transitions sont les parties du spectre dans lesquelles les formants présentent une évolution significative. Elles peuvent être assimilées aux consonnes (bien que cela soit réducteur lorsqu'on prend en compte les consonnes voisées), ou aux articulations en chant lyrique (staccato, legato, etc.).

Alors que les états stables peuvent être représentés par des valeurs de paramètres de FOF, les transitions seront

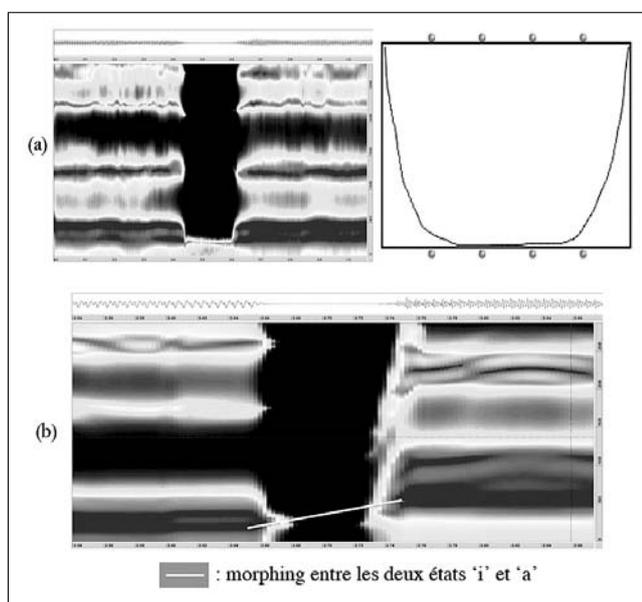
3. Ce modèle s'inspire des travaux effectués lors de l'implémentation de règles pour le contrôle de CHANT dans l'environnement Formes [11].

4. Cette solution est particulièrement adaptée au chant d'opéra.

définies à l'aide de ce que nous appellerons *trajets formantiques*. Ces trajets permettent de définir le comportement temporel des fréquences, amplitudes, largeurs de bande (voire d'autres paramètres, tels que le temps d'excitation) entre deux états stables.

Dans le but de modéliser des phénomènes vocaux (consonnes, articulations...), nous avons élaboré un certain nombre de trajets formantiques, à partir d'observations de signaux. Les phénomènes ainsi modélisés pourront être reproduits au sein d'un processus de synthèse.

Afin d'illustrer ce concept de trajet formantique, nous allons décrire la procédure de contrôle mise en place lors de la reproduction de la consonne 'b', dans la syllabe 'iba'. Pour simplification, seul le paramètre de la fréquence centrale du premier formant sera pris en compte ici.



**Figure 2.** (a) de gauche à droite : spectrogramme de la syllabe 'əbə', et trajet formantique s'appliquant à la fréquence centrale du premier formant dans la transition 'b'. (b) spectrogramme de la syllabe 'iba', obtenue par synthèse après application des différents trajets formantiques.

Le processus de reproduction de consonne peut être décrit ainsi : dans un premier temps, un trajet formantique est modélisé à partir de l'analyse d'un phonème, comme celui visible sur le spectrogramme de la figure 2 (a). Dans cet exemple, le trajet formantique est élaboré à partir d'analyses de consonnes 'b', et en particulier de l'observation du comportement de l'amplitude du premier formant (qui suit une courbe convexe). Puis, nous effectuons un *morphing* (par interpolation linéaire) à partir de deux états stables jouxtant la transition à créer (par exemple, les deux voyelles 'i' et 'a'). On applique enfin à ce morphing (par multiplication) le trajet formantique subi par l'amplitude du premier formant. Pour modéliser correctement la transition, il faut en réalité répéter ce processus pour chaque formant, et pour chacun de ses paramètres. Le sonagramme du résultat est en partie visible sur la figure 2(b). Le procédé est similaire pour la création d'une articulation type

staccato/legato : modéliser les trajets effectués par les amplitudes des différents formants est suffisante pour un obtenir un rendu réaliste.

Pour obtenir un contrôle plus fin sur le trajet formantique, nous introduisons un paramètre supplémentaire : la courbe de morphing, qui permet d'effectuer au début de la procédure une interpolation non linéaire entre deux états-stables, pour plus de réalisme. Nous appellerons par la suite l'ensemble (trajets formantiques / courbe de morphing) le *profil formantique* d'une transition.

Définir les différents phénomènes vocaux (syllabes, groupes phonème-articulation, etc.) par leur seul profil formantique fait la force de cette méthode : elle permet d'aborder le problème du contrôle de la synthèse de manière modulaire, contrairement aux techniques de concaténation de diphtonges, dans laquelle une consonne est définie de manière statique, en fonction du triphone<sup>5</sup> dans laquelle elle est présente. Notons cependant qu'une approximation a été faite dans ce modèle : il a été démontré que dans le cas de la voix parlée, la production d'une consonne peut dépendre de la voyelle précédant le triphone concerné [11].

### 3.2. Cas particuliers liés à la synthèse vocale

Comme nous l'avons présenté plus tôt, on ne peut classer les phonèmes<sup>6</sup> strictement en états stables et en transitions, et des cas particuliers sont à aborder.

L'exemple du 'm'<sup>7</sup> peut s'appliquer à d'autres phonèmes tels que le 'n' ou le 'l', ou même le 'ʒ' ou le 'z', si l'on ne considère que leurs composantes voisées. Une première manière de modéliser la consonne est d'utiliser une unique transition. Cependant, si le 'm' dure plus de 200ms, le trajet formantique se trouvera trop étiré dans le temps, et le rendu sonore pourra être jugé artificiel. Une autre modélisation, plus fine, consiste à modéliser la consonne grâce à deux transitions, qui formeront le début et la fin du 'm', et d'un état stable, qui formera sa partie centrale. Ainsi, on pourra par exemple obtenir un rendu de meilleure qualité pour un 'm' étendu dans le temps (de plusieurs secondes).

Par souci de concision, nous n'allons pas détailler d'autres cas particuliers, mais nous pouvons évoquer :

- les débuts et fins de phrases : ils pourront être représentés par des états stables ayant des paramètres d'amplitude nuls (silencieux), suivis d'une transition.
- le 'r' « roulé » : il s'agit d'un son pseudo-périodique, que l'on peut modéliser avec un pattern 'état stable-transition' répété, que l'on verra plus en détail dans la section 6.1.

## 4. IMPLÉMENTATION DANS OM-CHANT

La bibliothèque OM-Chant propose plusieurs objets, dotés d'attributs temporels tels qu'une date de début et une durée, représentant des événements de contrôle pour

5 . Un triphone est un triplet voyelle-consonne-voyelle.

6 . On parle ici de 'phonèmes' au sens donné par la phonétique.

7 . 'm' est une consonne occlusive nasale bilabiale voisée.

le synthétiseur CHANT. Ces objets permettent de formaliser les processus de synthèse, à partir d'événements de synthèse dissociés et contrôlant différents paramètres, tels que des valeurs de  $f_0$ , représentées par des objets *ch-f0*, et des paramètres formantiques (matrices de  $n$  formants), représentés par des objets *ch-FOF*. Ces différents objets permettent de régler les paramètres de synthèse et leur évolution dans le temps. Lorsque deux événements se suivent (qu'ils soient de type *ch-f0* ou *ch-FOF*, une interpolation linéaire est effectuée entre les valeurs des paramètres à la fin du premier et au début du second événement. Pour plus de précisions concernant le contrôle de CHANT dans l'environnement OpenMusic, on pourra se référer à [6].

#### 4.1. Génération des transitions

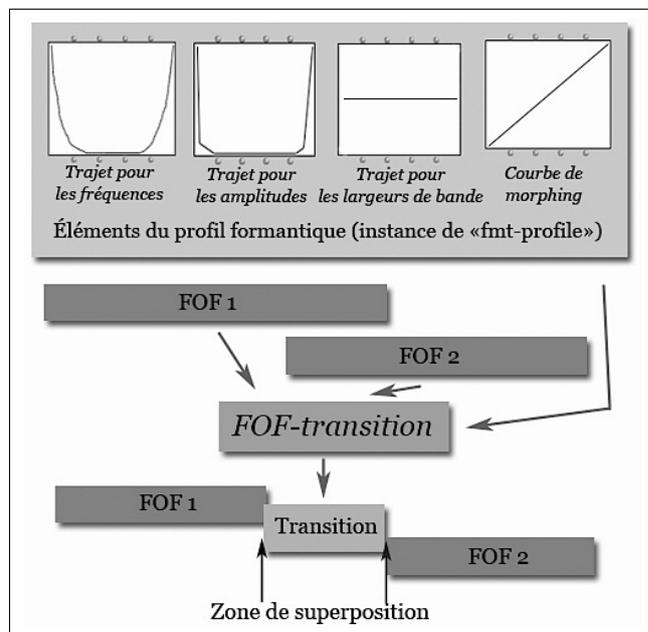
Nous avons implémenté une série d'outils permettant de traiter les cas d'événements distants, en les considérant comme des états stables, et générant les transitions qui les séparent. Dans la bibliothèque OM-Chant, la fonction *ch-transitions* permet de gérer et de transformer une séquence d'événements, en appliquant une fonction de manière itérative (cette fonction lui étant donnée en argument). C'est en se basant sur ce système qu'une première fonction, *FOF-transition*, a été implémentée. Cette fonction prend en entrée la définition du profil formantique de la transition désirée, ainsi que deux événements de type *ch-FOF*, les transforme si besoin est, et insère entre les deux un nouvel événement de transition (également du type *ch-FOF*). Ainsi, utilisée comme fonction-lambda avec *ch-transitions*, elle permet de gérer et transformer une séquence d'événements *ch-FOF* en insérant des événements transitoires.

La définition du profil formantique donné en paramètre se fait par l'intermédiaire de la classe *phoneme*. Celle-ci est également une matrice, permettant de stocker des courbes, correspondant respectivement aux trajets effectués par chacune des fréquences, amplitudes et largeurs de bande des formants (et extensible à d'autres paramètres). Est aussi stockée la courbe de morphing entre les deux états stables.

Sur la figure 3 se trouve un exemple de traitement d'une transition avec *FOF-transition*, sous forme schématisée.

Du point de vue algorithmique, *FOF-transition* effectuée dans un premier temps le « morphing » (ou interpolation) entre les paramètres des deux FOF passées en entrée, puis échantillonne les trajets formantiques avant de les multiplier au résultat.

Ce procédé pose le problème du taux d'échantillonnage : quelle granularité appliquer à l'événement de transition ? Nous avons introduit un paramètre optionnel permettant de régler la période d'échantillonnage de la transition. Par défaut, celle-ci est de 1ms. En effet, il s'est avéré qu'une fréquence d'échantillonnage de 1000Hz offrait un rendu de qualité dans le cas de consonnes et d'articulations communes de type staccato, legato, avec un temps de calcul négligeable (de l'ordre de quelques millisecondes pour le rendu d'une seconde de trajet formantique).



**Figure 3.** Illustration du fonctionnement de *FOF-transition*, qui prend en entrée deux événements « FOF », et insère une transition à l'endroit où ils se superposent, conformément au profil formantique de la consonne 'b' (simplifié, tous les formants ayant le même profil) passé en paramètre.

Une fonction analogue, *f0-transition*, permet de définir le trajet effectué par la fréquence fondamentale entre deux événements de type  $f_0$ . Le trajet est formalisé par une courbe, qui vient multiplier l'interpolation linéaire effectuée entre les deux objets *ch-f0* qui jouxtent la transition.

#### 4.2. Dictionnaire de transitions

La gestion des différentes définitions de profils formantiques est centralisée, via l'utilisation d'un dictionnaire. Il s'agit d'un patch OpenMusic associant à une liste d'identifiants (clés<sup>8</sup>), représentés par des chaînes de caractères, une liste d'objets de type *phoneme*. Il devient alors possible de définir un dictionnaire de transitions, qui peut être étendu à volonté par l'utilisateur.

Le dictionnaire peut aussi avoir un rôle créatif, au delà de celui de « base de données » : il est possible de faire la requête d'un profil formantique hybride, en sélectionnant deux clés correspondant à deux profils distincts, ainsi qu'une valeur de pondération permettant de calculer un profil intermédiaire entre ceux-ci.

#### 4.3. Transitions paramétrées

Afin d'introduire un aspect « dynamique » dans notre système de transitions, nous sommes partis du cas pratique suivant : comment créer une phrase contenant des articulations évoluant progressivement du *legato* au *staccato*, sans

8 . « Clés primaires », pour utiliser le vocabulaire des bases de données.

avoir à modéliser chaque articulation intermédiaire entre les deux types de jeu ? Nous avons alors mis en place un mécanisme permettant de donner un paramètre en entrée lors de la création d'un profil formantique, pour permettre à *phoneme* d'être instancié dynamiquement. Ainsi, il devient alors possible d'imaginer une transition paramétrée modélisant le staccato, le legato, ainsi que l'ensemble des profils intermédiaires entre ces deux articulations. Dans cet exemple, une valeur de 0 donnée en entrée créera un profil de staccato très net, 1 un profil de portamento, et les valeurs comprises dans ]0; 1[ un profil correspondant à une articulation intermédiaire. Un exemple d'application est donné dans la partie 6.2.

#### 4.4. Morphing formantique

En parallèle, nous avons implémenté une procédure permettant de générer des états hybrides entre stabilité et transition. Bien que cela n'entre pas directement dans le cadre de notre modélisation de départ, cette procédure permet d'étendre les possibilités compositionnelles offertes par le système.

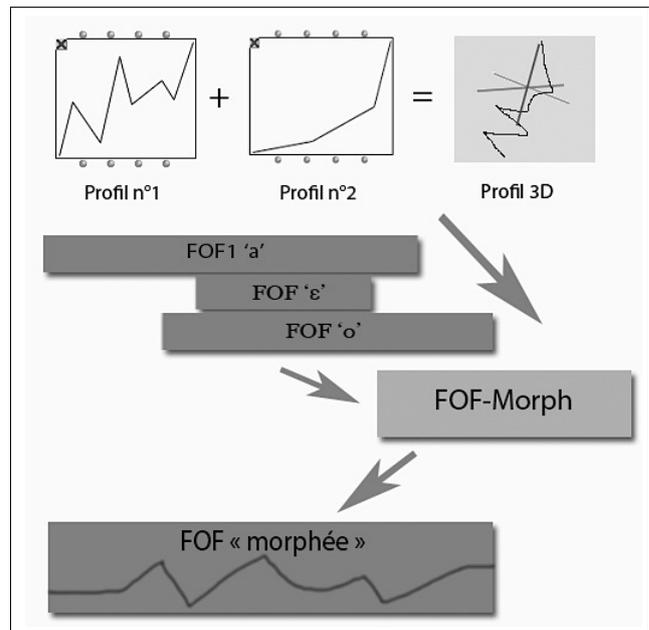
Nous appellerons cette procédure le « morphing formantique ». En effet, elle permet de prendre en entrée plusieurs états stables, et de pondérer leurs paramètres, afin d'obtenir un état hybride. La fonction *FOF-morph* a été implémentée selon ce principe et permet, à partir d'un ensemble d'événements FOF et d'une courbe, d'effectuer une pondération dans le temps entre les événements superposés, suivant un profil donné. En fonction du nombre d'événements, une courbe supplémentaire peut être spécifiée : le premier profil sera utilisé si deux FOF se superposent, le second si une troisième FOF est donnée (on parlera alors de « morphing 3D »). Une illustration de ce mécanisme est proposée sur la figure 4.

Cette procédure de morphing peut servir, entre autres, à modéliser des chants diphoniques.

### 5. ENVIRONNEMENT GRAPHIQUE

Pour proposer à l'utilisateur un environnement de composition plus efficace et intuitif, nous avons intégré ces procédures de contrôle aux *maquettes* OpenMusic. Les maquettes sont des interfaces graphiques permettant de représenter et d'agencer des structures musicales dans le temps. Elles sont composées d'une interface principale, comprenant une « timeline », sur laquelle il est possible de déposer ou de créer des patches ; ainsi qu'un patch interne d'évaluation [4]. Ce dernier permet de définir le processus produisant le résultat de la maquette (dans notre cas, un rendu sonore), après évaluation des patches contenus dans celle-ci.

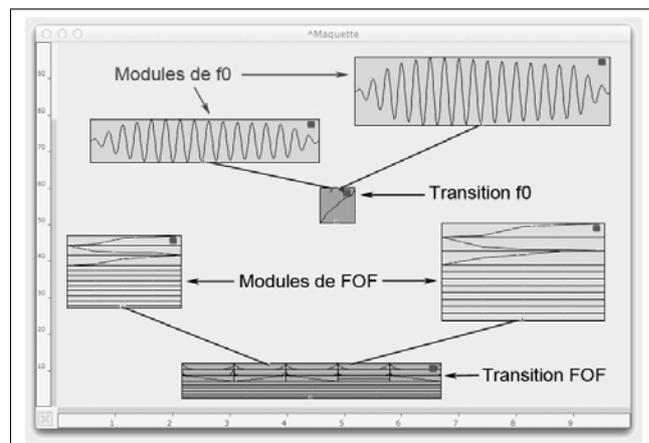
Grâce aux maquettes, il devient possible de représenter symboliquement les éléments de contrôle sur un espace 2D, et d'éditer graphiquement les phrases musicales. Concrètement, l'utilisateur peut créer et manipuler divers modules (patches) représentant des événements, qui sont classés dans différentes catégories :



**Figure 4.** Illustration du fonctionnement de *FOF-morph*, en effectuant une pondération entre trois états stables superposés, représentant les phonèmes 'ə', 'a' et 'o', en suivant deux profils (représentés en vue 3D).

- modules de FOF, qui représentent les états stables ;
- modules de f0 (avec ou sans vibrato) ;
- modules de transition (f0 et FOF) ;
- modules de morphing de FOF.

Un exemple d'utilisation est donné sur la figure 5. On peut observer sur cette figure que la phrase vocale a été construite en connectant différents modules : les modules de FOF sont connectés en entrée à un module de transition, de même que les modules de f0. Ainsi, les transitions sont générées automatiquement en fonction des états stables lors de l'évaluation de la maquette.



**Figure 5.** Maquette contenant deux modules de f0 avec vibrato et un module de transition f0 (en haut), ainsi que deux modules de FOF séparés par un module de transition (en bas). Les modules de transition ont été légèrement déplacés pour rendre visibles les connexions entre modules.

L'ensemble des objets produits par évaluation des différents modules sont de type homogène : le résultat d'un morphing ou d'une transition entre deux modules de FOF sera un objet du même type que celui produit par l'évaluation d'un module de FOF simple. Cela permet de définir des événements complexes tels que des morphings de morphings, ou des morphings de transitions...

Nous pouvons aussi noter que le système de base de données de phonèmes, qui vient enrichir le concept de dictionnaire de transitions évoqué en partie 4.2, permet de concevoir des « banques de chanteurs », qui contiennent les articulations ainsi que les phonèmes propres aux formants et au style de jeu d'un chanteur virtuel.

Pour accélérer la création d'une phrase musicale, nous avons implémenté plusieurs fonctions permettant de remplir automatiquement une maquette avec des modules de FOF et de f0. Ces fonctions prennent en entrée :

- une structure temps fréquence <sup>9</sup>, représentant les différentes notes jouées (on peut par exemple partir d'un objet *chord-seq* <sup>10</sup>).
- une suite de symboles, désignant les clés pointant vers des modules d'états stables et de transitions.
- une base de données de phonèmes, effectuant la relation clés/modules de FOF (instance de la classe *phoneme-DB*).
- une liste spécifiant les durées des différentes transitions.

Après analyse des entrées, et vérification de leur cohérence (rapport entre le nombre d'états stables et le nombre de notes, etc.), ces fonctions génèrent les modules qui viendront remplir une maquette cible. D'autres traitements particuliers sont effectués, comme dans le cas de suites de consonnes, où un état stable de courte durée (20 ms) est ajouté entre chaque « consonne ». Un exemple est visible sur la figure 6, et le résultat de son évaluation, sous forme de maquette, sur la figure 7.

Il devient alors possible de travailler à deux niveaux : à partir de structures musicales (*chord-seq*, « séquences de phonèmes »), en travaillant à partir de données musicales précises ; et directement à partir de la maquette, en agencant de manière intuitive des événements de contrôle dans le temps.

## 6. APPLICATIONS MUSICALES

L'introduction de ces nouvelles procédures de contrôle dans l'environnement OpenMusic a permis de créer des processus musicaux originaux. Nous allons détailler deux d'entre eux, qui ont été utilisés à l'occasion de la création de l'opéra *Re Orso* de Marco Stroppa. <sup>11</sup>

9. Structure sous la forme d'une liste d'éléments (fréquence, date, durée).

10. Dans OpenMusic, les instances de la classe *chord-seq* permettent de représenter des suites de notes ou d'accords.

11. Création en mai 2012 à l'Opéra Comique de Paris.

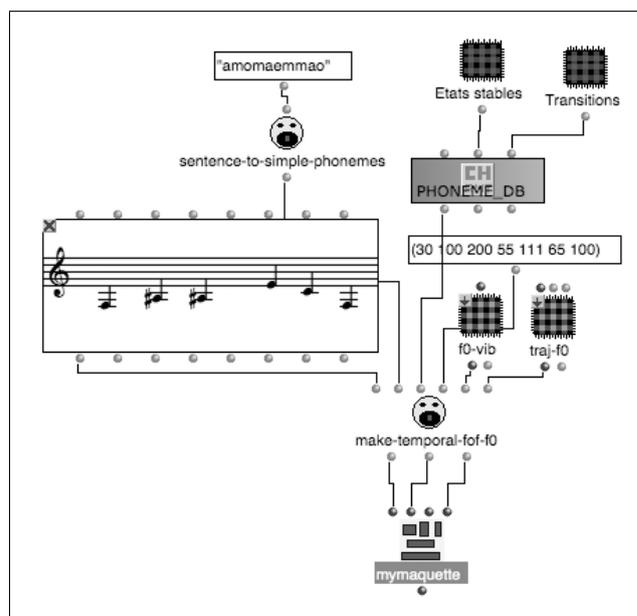


Figure 6. Patch illustrant le remplissage automatique d'une maquette à partir d'un objet *chord-seq*, d'une base de données de phonèmes, d'une chaîne de caractères et d'une liste de durées de transition.

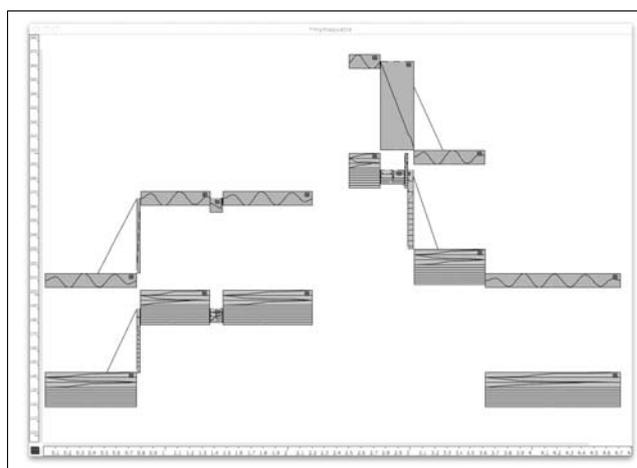


Figure 7. Maquette remplie automatiquement après évaluation du patch de la figure 6.

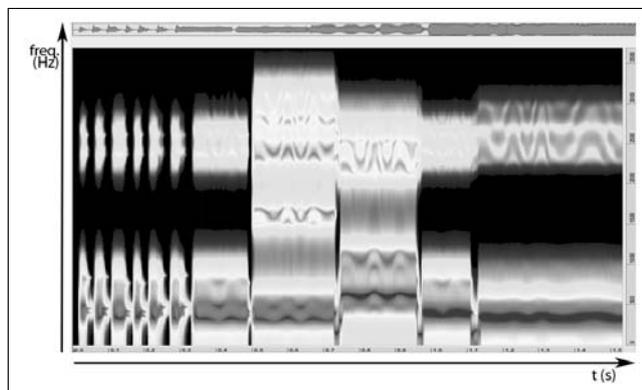
### 6.1. Roulement de 'r' à période variable

Contrairement au 'r' tel qu'on le prononce en français <sup>12</sup>, qui peut être synthétisé à l'aide de variations aléatoires agissant sur les différents paramètres formantiques, le 'r' roulé présente un motif périodique. En effet, en fonction de la morphologie de l'appareil vocal du chanteur, la langue vient frapper le palais à une certaine fréquence <sup>13</sup>. Au regard de cette particularité, nous avons choisi de modéliser le 'r' roulé à l'aide d'une suite de transitions, chacune correspondant à une période, soit à un claquement de

12. Il s'agit, en phonétique, d'une consonne *uvulaire*.

13. La moyenne se situe approximativement autour de 20Hz, cependant, il existe des individus capables de rouler les 'r' à des fréquences dépassant les 30Hz.

langue. Un ensemble d'états stables courts permettra de séparer les différentes transitions. Il devient alors possible de paramétrer la durée de chaque claquement, et donc d'effectuer des *accelerando* ou des *rallentendo* au niveau des roulements, chose impossible à effectuer pour un chanteur. Une analyse temps fréquence du résultat d'un tel processus est visible sur la figure 8.



**Figure 8.** Analyse temps fréquence d'un roulement de 'r', synthétisé par CHANT, dont la fréquence décroît au cours du temps

Ce processus de synthèse est un exemple de ce qu'offre cette implémentation de CHANT en termes de créativité : il devient possible de dépasser et de détourner le modèle de production de la voix.

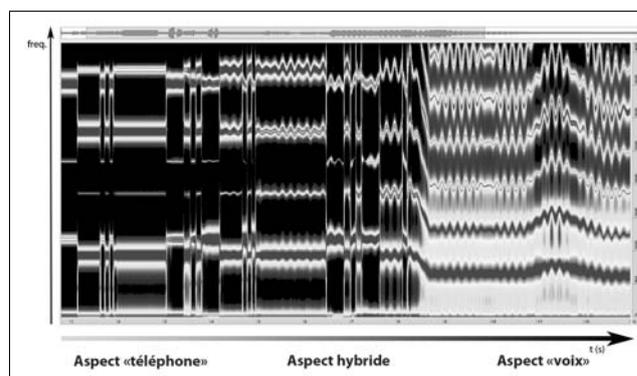
## 6.2. Morphings timbraux

Un autre exemple, tiré lui aussi de *Re Orso*, est celui du morphing entre un matériau « non-organique » et une voix. Nous avons ici voulu créer une phrase de quelques secondes au sein de laquelle une sonnerie de téléphone (sonnerie monophonique et monotimbrale) est transformée progressivement en une voix de chanteur lyrique. Pour cela, nous avons travaillé à partir de plusieurs éléments :

- les paramètres des états stables : en effectuant une interpolation entre les paramètres de formants reproduisant le timbre de téléphone, et ceux issus de l'analyse de la voix d'une chanteuse d'opéra <sup>14</sup>.
- la nature des articulations, en passant progressivement d'un staccato pur à un legato (en utilisant des transitions paramétrées).
- la durée des transitions, en partant de durées inférieures à 25 ms, pour aboutir à des durées supérieures à 50 ms.
- le vibrato, nul au début et dont l'amplitude augmente en fonction du temps (à noter que le vibrato suit une enveloppe sur chaque note).

Sur la figure 9 se trouve l'analyse spectrale du rendu sonore d'un tel processus, sur laquelle on peut observer l'évolution de ces paramètres.

14. Analyse d'un extrait de « La Reine de la Nuit », Mozart, interprétation de Erika Miklósa.



**Figure 9.** Analyse spectrale du rendu sonore du processus « téléphone vers voix », identification des trois « aspects » sonores (timbres).

## 7. CONCLUSION, PERSPECTIVES

Nous avons introduit dans l'environnement OpenMusic plusieurs méthodes permettant l'élaboration de procédures de contrôle originales du synthétiseur CHANT, en partant d'un modèle de base pour la voix chantée, que nous avons étendu afin de maximiser les capacités expressives du système. Deux mécanismes ont été mis en place : le premier permet de gérer une séquence d'événements en les traitant deux à deux ; un second permet de manipuler les événements de synthèse de manière intuitive grâce à une interface graphique. Ces outils sont intégrés dans OM-Chant et documentés dans le manuel utilisateur en ligne de la bibliothèque : <http://support.ircam.fr/docs/om-libraries/om-chant/>. De plus, des rendus sonores issus des processus décrits dans cet article sont disponibles : <http://repmus.ircam.fr/foulon/om-chant-exemples>.

Afin de terminer notre modèle de production vocale, il faudrait effectuer un travail en profondeur sur l'ajout de composantes bruitées. Ces fonctionnalités n'ont pas été proposées : nous nous sommes concentrés sur l'étude des sons voisés, qui constitue à elle seule un sujet de recherche très vaste. Cependant, les procédures proposées par notre système pourront intégrer le contrôle de la synthèse de sons non-voisés, notamment grâce aux modules de génération de bruit et de filtrage disponibles.

La synthèse par FOF est dotée d'un potentiel créatif très large, notamment de par la richesse des gammes de sons qu'elle est capable de produire. Si nous nous sommes souvent limités à la production de signaux inspirés de la voix chantée, des procédures de contrôle permettant de reproduire (et de détourner !) d'autres types de sons pourront à l'avenir être imaginées. Parmi eux, certains sons instrumentaux qui se prêtent bien à la synthèse par formants (percussions, cloches, cuivres...) ou encore d'autres sons issus de techniques originales, telles que la synthèse granulaire [7], ou les synthèses hybrides <sup>15</sup>.

15. Nous avons travaillé sur un mélange entre synthèse par FOF et synthèse additive, en modélisant un glas à l'aide de 171 formants, pour lui faire subir un morphing timbral de manière analogue qu'en partie 6.2.

## 8. REFERENCES

- [1] Baisnée P.-F. and the CHANT Group, « CHANT Manual », IRCAM, 1985.
- [2] Barrière J.-B., « Chréode, un chemin vers une nouvelle musique avec l'ordinateur », *IRCAM : une pensée musicale*, ed. Tod Machover, InterEditions, 1984.
- [3] Bogaards N., Röbel A., « An interface for analysis-driven sound processing », *AES 119th Convention*, New York, USA, 2004.
- [4] Bresson J., Agon C., « Temporal Control over Sound Synthesis Processes », *Proc. Sound and Music Computing Conference*, Marseille, France, 2006.
- [5] Bresson J., Michon, R., « Implémentations et contrôle du synthétiseur CHANT dans OpenMusic », *Actes des Journées d'Informatique Musicale*, Saint-Etienne, France, 2011.
- [6] Bresson J., Stroppa, M., « The Control of the CHANT Synthesizer in OpenMusic : Modelling Continuous Aspects in Sound Synthesis », *Proc. Int. Computer Music Conference*, Huddersfield, Royaume Uni, 2011.
- [7] Liuni M., Gentilucci M., « Multifog : a Multi-Level Control Device for FOG Granular Synthesis in Max/MSP », *Computer Music Modeling and Retrieval*, Malaga, Espagne, 2010.
- [8] Rodet X., « Time-domain Formant-wave Function Synthesis », *Computer Music Journal*, 8(3), Autumn 1984.
- [9] Rodet X., Potard Y., Barrière J.-B., « The CHANT Project : From the Synthesis of the Singing Voice to Synthesis in General », *Computer Music Journal*, The MIT Press, 8(3), 1984.
- [10] Rodet X., Cointe P., « Formes : Composition and Scheduling of Processes », *Computer Music Journal*, 8(3), Autumn 1984.
- [11] Rodet X., Depalle P., « High quality synthesis-by-rule of consonants », *Proc. International Computer Music Conference*, Vancouver, Canada, 1985.
- [12] Rodet X., Lefevre A., « The Diphone Program : New Features, New Synthesis Methods and Experience of Musical Use », *Proc. Int. Computer Music Conference*, Thessaloniki, Grèce, 1997.
- [13] Sundberg J., « The KTH synthesis of singing », *Advances in Cognitive Psychology*, Versita, 2006.

# GESTION DES FLUX DE MESSAGES ENTRE APPLICATIONS GRÂCE À UN PLUGIN DE SERVEUR OSC DANS LE PROCESSUS COMPOSITIONNEL ÉLECTROACOUSTIQUE ET MIXTE

Matthieu JACQUOT  
err0r500.music@gmail.com

## RÉSUMÉ

Cet article se propose d'étudier l'utilité dans la méthode de travail du compositeur de musique électroacoustique ou mixte d'un serveur OSC intégré sous forme de plugin à une station audionumérique standard (DAW), lui permettant ainsi de gérer les flux de messages entre applications de manière synchronisée avec le matériau sonore grâce, en particulier, au système d'automation. Le but est de lui permettre ainsi de tirer parti des points forts de chaque type d'application : montage, édition, synchronisation depuis une DAW classique ; traitement du signal grâce à un environnement de synthèse sonore (ESS)<sup>1</sup>.

## 1. INTRODUCTION

Bien que de nombreux ESS à la fois puissants et flexibles soient à la disposition des compositeurs électroacoustique (Max, PureData ou SuperCollider pour ne citer que les plus communs), leur manque fréquent d'outils pratiques pour la synchronisation avec un matériau sonore en cours d'élaboration peut rendre leur utilisation mal aisée pour le compositeur travaillant, au moins de manière préliminaire, sur média fixe<sup>2</sup>. Nombreux sont alors les compositeurs contraints d'adopter l'une des deux stratégies suivantes :

- effectuer les traitements sonores au sein de leur DAW grâce à des plugins standards puis les redéployer dans un environnement plus adapté au temps-réel (pour le cas de la musique mixte) en tentant d'approcher autant que possible la première version. Cela les empêche néanmoins de tirer pleinement profit de la puissance des ESS et augmente aussi considérablement le temps de développement puisqu'un même travail est à effectuer deux fois, dans deux environnements différents.

- travailler directement dans un ESS mais en étant alors obligé de réintégrer dans la DAW le matériau so-

<sup>1</sup>Il est à noter que ce plugin permet de synchroniser l'hôte avec quelque application capable de gérer la réception de messages OSC ce qui ouvre la possibilité de gérer aussi bien du son que des interfaces physiques (Monome), de l'image (Processing, Jitter, Gem) ou même des environnements 3D (Blender).

<sup>2</sup>L'intérêt porté au projet « rs.delos » de Roby Steinmetzer se proposant de recréer un séquenceur linéaire minimaliste directement dans Max semble corroborer cette assertion.[10]

nore traité. Ils sont aussi contraints de modifier en temps-réel les paramètres de traitement et d'enregistrer le rendu ce qui pose alors le problème de la reproductibilité du traitement. L'autre option pouvant être de procéder par séquences successives au cours de l'œuvre grâce à l'utilisation de « cues » successifs permettant de basculer d'une configuration à une autre. Bien que cette solution soit tout à fait justifiable sur un plan artistique et technique, il est néanmoins nécessaire que cela ne reste qu'une option et non une contrainte, ce qui reviendrait sinon à considérer que ces environnements limitent d'un côté tout autant qu'ils libèrent d'un autre.

Nous reviendrons tout d'abord sur les différentes solutions disponibles actuellement puis verrons en quoi l'utilisation d'un serveur OSC directement intégré à une DAW peut être une solution viable pour chacun de ces cas de figure.

## 2. PROBLÉMATIQUE ACTUELLE

### 2.1. Utilisation du MIDI

Historiquement, le protocole standard pour l'échange de messages entre applications fut le MIDI. Son évolution en « High-Definition Protocol » bien qu'en développement depuis 2005, n'est à l'heure actuelle (février 2013) pas totalement standardisé et ne peut donc pas être discuté ici [8]. Il existe néanmoins une alternative moderne : le protocole OSC développé par le CNMAT de l'Université de Californie à Berkeley. Ce protocole possède de nombreuses caractéristiques permettant de palier aux limitations historiques ou structurelles du protocole MIDI.

#### 2.1.1. Limitations liées au protocole

Les quelques avantages de l'OSC par rapport au MIDI n'étant pas sujets à débat [4][7] et nous intéressant dans la situation actuelle peuvent être rapidement énumérés :

- Le schéma de nommage, l'adresse, sous forme d'URL permet d'avoir un paramètre « lisible par un hu-

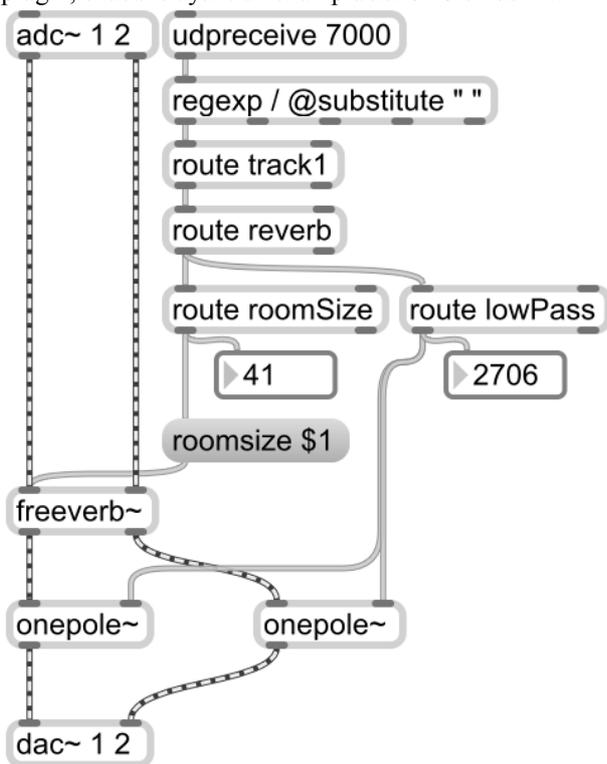
main » (*human-readable*) et pouvant donner une idée assez claire de son rôle ainsi que d'éviter les conflits involontaires entre les paramètres comme cela est aisé en assignant le même numéro de Control Changes MIDI à deux paramètres différents. Voici par exemple la « taille de la pièce » d'une réverbération appliquée sur la première piste :

Ex : /track1/reverb/roomSize

- Avec OSC, les valeurs transmises ont une précision sur 32-bit comparé au MIDI qui est sur 7-bit et peuvent ainsi s'adapter beaucoup plus facilement au type de paramètre contrôlé. Cela permet par exemple d'exprimer directement en Hertz la fréquence de coupure d'un filtre. Cette caractéristique reste néanmoins tributaire de la résolution des automatisations dans l'hôte<sup>3</sup>.

Ex : /track1/reverb/lowPassFreq 3128 (Hz)

- Nous voyons de plus, par ces deux simples exemples la possibilité d'utiliser des « motifs » dans l'adresse.[4] Le plugin de serveur OSC présenté tire bénéfice de cela en proposant un premier champ « main pattern » qui ici pourrait être /track1/reverb/ permettant ensuite de n'avoir qu'à ajouter la suite de l'adresse puis à spécifier si nécessaire les bornes des valeurs transmises. On peut alors envisager de multiplier les instances du plugin, chacune ayant un champ d'action bien défini.



**Figure 1.** Contrôle d'une réverbération et d'un filtre passe-bas en utilisant le protocole OSC sous Max4.

<sup>3</sup>A titre d'exemple, la précision de l'automatisation dans le logiciel Reaper de Cockos va jusqu'à 6 décimales.

- A cela il faut ajouter que l'OSC possède une précision temporelle de l'ordre de  $10^{-12}$ s alors que le MIDI se situe aux alentours des  $10^{-3}$ s.[4]

### 2.1.2. Limitations dues à l'implémentation logicielle

Il faut aussi noter que la prise en charge dans les DAW des messages MIDI (ici essentiellement des « Control Changes ») n'est pas toujours très adaptée à ce type d'utilisation car ils sont, pour des raisons historiques, plus souvent pensés comme une succession d'événements discrets plutôt que comme des flux et donc non dans l'optique d'une synchronisation avec le matériau sonore. De plus les logiciels font souvent la distinction entre piste audio ou MIDI ce qui signifie que la synchronisation doit s'effectuer entre deux pistes et non pas, comme c'est le cas pour les automatisations, comme des paramètres de la piste audio.

Il existe néanmoins des plugins permettant de transformer une enveloppe d'automatisation en message MIDI<sup>4</sup> mais cela n'affecte bien sûr pas les limitations liées au protocole lui-même et peut être la cause d'un nouveau problème.

En effet, router des messages MIDI en interne ne peut se faire que grâce à des logiciels conçus spécifiquement pour permettre l'interfaçage entre les applications au moyen de Bus virtuels. Néanmoins, la quantité de messages MIDI générée par les enveloppes peut rapidement se révéler problématique en conduisant à de faux positifs dans la détection de boucles (*feedback loops*), certains de ces logiciels ont en effet une manière très sommaire de les détecter en se contentant de monitorer le volume de données transmises et en fixant un simple seuil, bloquant l'envoi de tous les messages dès que celui-ci est franchi.<sup>5</sup>

### 2.1.3. Avantages du MIDI

Un avantage indéniable du MIDI comparé à l'OSC est son taux de pénétration du marché. En effet, la prise en charge de l'OSC au niveau des interfaces et des logiciels est bien plus rare que le MIDI qui reste jusqu'à aujourd'hui le protocole de communication standard. Il est néanmoins nécessaire de remarquer qu'un nombre croissant de logiciels et d'interfaces le supportent.

## 2.2. Utilisation d'une solution « Max4Live »

Une alternative récente pourrait être la solution propriétaire Max4Live développé conjointement par Ableton et Cycling '74 et intégrant dans une même solution logicielle la flexibilité de l'ESS Max avec celui de la DAW généraliste Live, c'est-à-dire justement ce que nous tentons d'effectuer. Nous tenterons néanmoins de porter un regard critique sur cette solution qui, mis à part le fait que l'ESS soit évidemment limité à Max et que la DAW elle-même ne convient pas forcément à tous les types

<sup>4</sup>Tel que ReaControlMIDI : <http://www.reaper.fm/reaplugs/>

<sup>5</sup>C'est le cas notamment de Loopbe1 développé par nerds.de

d'approches (par exemple, celles nécessitant un gros travail de montage de la source sonore traitée) :

### 2.2.1. *La spécificité de la plate-forme contraint le compositeur à s'adapter à elle*

L'avantage du serveur OSC est sa compatibilité avec n'importe quelle plate-forme supportant les plugins de type VST, AU ou RTAS c'est à dire quasiment toutes les DAW, y compris celle à laquelle le compositeur est déjà habitué et sur laquelle il a ses habitudes de travail. De plus il permet d'utiliser comme outil de traitement n'importe quelle application lui convenant, soit là encore par habitude, soit pour des raisons de spécificité du traitement à accomplir. Notons d'ailleurs aussi en faveur du serveur OSC qu'il permet d'utiliser plusieurs ESS simultanément. C'est donc ici l'environnement qui s'adapte au compositeur et non l'inverse.

### 2.2.2. *Limitation de la possibilité de transmissibilité du traitement*

Contrairement à l'approche électroacoustique dans laquelle le compositeur produit une œuvre « finie ». Le compositeur de musique mixte doit, au même titre que le compositeur de musique instrumentale, mettre en place un moyen pour des tiers de créer l'œuvre à sa place. Cela se fait sur deux plans : reproduire la partie instrumentale et le traitement sonore.

Si nous observons comment est gérée la partie instrumentale, nous nous rendons compte que les compositeurs et éditeurs ont naturellement opté pour le format le plus ouvert : un langage standardisé imprimé sur du papier ou diffusé via un format de fichier accessible à tous.

C'est pour la transmission du traitement sonore que le choix de l'ESS se révèle primordial et que l'ouverture rendue possible par VST2OSC prend toute son importance. En effet, un code source, qu'il soit sous forme de texte ou de schéma peut toujours faire sens pour celui qui le lit. A contrario, le format propriétaire adopté ici empêche légalement toute solution alternative d'avoir accès à ces informations, les fichiers de sessions étant par ailleurs sous forme binaire. Nous pouvons alors envisager deux conséquences :

- à court terme, car nous pouvons imaginer la situation à laquelle seraient confrontées les salles de concert en cas de prolifération de ce type de solution logicielle propriétaire du fait des coûts qu'elles engendrent.
- à long terme, car la maintenance d'un patch Max, PureData ou d'un code quel qu'il soit est déjà très problématique ; avec une solution fermée comme celle-ci, nous pouvons imaginer les problèmes auxquels se heurterait une tentative de reproduction d'ici plusieurs d'années. Pour s'en rendre compte il suffit tout simplement de lancer un regard sur les logiciels d'il y a 20 ans pour avoir un bon aperçu de l'ampleur de la tâche : pour pouvoir exploiter le « project »<sup>6</sup> il faudra retrouver une

version du logiciel compatible avec celui-ci ce qui nécessitera sûrement l'emploi d'une version du système d'exploitation compatible ce qui implique alors aussi de l'installer sur une machine que ce dernier est capable de gérer.

Pour clore le parallèle avec la partition, nous nous retrouvons dans la même problématique que si les éditeurs exigeaient des interprètes de posséder le logiciel utilisé pour la gravure de la partition, dans une version compatible, pour pouvoir y avoir accès.

### 2.2.3. *Nécessité de redéployer pour cette plate-forme spécifique des patches Max existants*

Bien que très proches, il est néanmoins nécessaire de redévelopper les patches Max pour les passer sous forme de « device » Max4Live. Cela peut se révéler assez fastidieux dans le cas de gros patches et ne présente comme unique avantage, par rapport à la solution proposée grâce au serveur OSC, la possibilité de faire des rendus « offline ».

## 2.3. ReWire

Les deux principaux problèmes liés à cette solution sont en premier lieu liés à la compatibilité. En effet, ReWire est une solution propriétaire de communication entre applications développé conjointement par Propellerhead et Steinberg. La licence restrictive de ce logiciel le rend de fait incompatible avec toute solution open-source ce qui signifie que, pour reprendre l'exemple des 3 ESS présentés ci-dessus, seul Max est compatible avec cette technologie. Les autres ne pourraient le devenir qu'en devenant propriétaires à leur tour. Ceci montre bien une grande réduction du panel d'applications à disposition du compositeur.

De plus la transmission des messages se fait en MIDI ce qui renvoie aux problèmes inhérents à celui-ci comme décrits ci-dessus (Cf 2.1).

## 2.4. OSCulator

OSCulator est un logiciel propriétaire, disponible uniquement sous OS X. Son but est, tout comme VST2OSC de permettre d'envoyer des messages OSC pour contrôler quelque logiciel susceptible de supporter le protocole. Un problème réside néanmoins dans le fait qu'il ne peut pas s'intégrer en tant que tel dans une DAW. Le seul moyen de le synchroniser avec un matériau sonore serait d'utiliser le MIDI pour faire l'interfaçage entre la DAW et OSCulator, ce qui nous renvoie aux mêmes problèmes que d'utiliser directement le protocole MIDI.

Il existe de nombreux autres logiciels permettant de contrôler des applications via des messages OSC mais tous sont pour l'instant pensés comme des moyens d'interfaçage geste-logiciel et ne permettent pas la reproductibilité du résultat.

<sup>6</sup>Nom donné aux sessions sous Ableton Live

### 3. SOLUTION PROPOSÉE

Le but est ici de trouver une solution de travail pour le compositeur qui soit à la fois intuitive, rapide, puissante et modulaire. Nous avons vu que jusqu'à présent les solutions étaient soit de consacrer beaucoup de temps par conviction au déploiement de solutions viables, soit d'accepter cet état de fait et de privilégier la rapidité au détriment de la pérennité.

#### 3.1. Gestion des messages

##### 3.1.1. Intégration

Le serveur OSC proposé, VST2OSC[6], se présente sous la forme d'un plugin<sup>7</sup>, s'intègre alors parfaitement à n'importe quelle DAW et permet donc de mettre à profit les outils standards de cette dernière, en particulier les enveloppes, pour moduler des paramètres d'un patch de manière à ce qu'ils soient à la fois aisément enregistrables, éditables, reproductibles et synchronisés avec le son. On peut aussi envisager de faire transiter les messages d'un contrôleur dans la DAW où ils seront enregistrés en tant qu'enveloppes du plugin, ce dernier transmettant les messages à l'ESS. La performance pouvant ensuite être reproduite et éditée à volonté.

L'avantage de cette méthode étant à la fois de lier le travail compositionnel à l'écriture du patch de concert qu'il permet par ailleurs de rendre plus progressif et modulaire en se concentrant tout d'abord sur de petits modules de traitement du son au fur et à mesure que leur utilisation s'impose au compositeur et de ne se poser la question de leur regroupement au sein d'un grand patch pour le concert qu'une fois ceux-ci développés, rendant ainsi le portage extrêmement rapide et fiable puisque seule la partie prise en charge par la DAW reste à intégrer.

##### 3.1.2. Détails de l'organisation du plugin



Figure 2. Interface graphique de VST2OSC

Bien que permettant de générer l'envoi des messages, l'interface graphique a été principalement conçue comme

<sup>7</sup> Le développement du plugin a été fait en C++ en utilisant les bibliothèques JUCE et OSCLib. Pour l'instant déployé uniquement sous forme de VST Windows, l'utilisation de JUCE permet néanmoins d'envisager aisément à l'avenir un déploiement cross-platform sous forme de VST mac, AU et RTAS.

un éditeur permettant de formater les messages à envoyer. L'utilisation effective se faisant par le biais des enveloppes de la piste.

Le plugin est composé de deux grandes parties :

- la première prend en charge les paramètres globaux :

Champ	Fonction	Valeur initiale
Destination IP	Adresse IP du client	∅ (string)
Destination Port	Port sur lequel le client écoute	7000 (int)
Main Pattern	Partie de l'adresse commune à tous les modules	∅ (string)

Table 1. Paramètres globaux

- La seconde, contient 10 modules indépendants, chacun possédant les paramètres suivants :

Champ	Fonction	Valeur initiale
End Pattern	Fin de l'adresse	∅ (string)
min	Valeur minimale des données (data) à envoyer	0. (float)
max	Valeur maximale des données (data) à envoyer	127. (float)
steps	Pas de la valeur à transmettre	1. (float)

Table 2. Paramètres des modules

Au moment de l'envoi des messages, l'adresse est créée en concaténant les deux chaînes « Main Pattern » et « End Pattern ».

Les données envoyées,  $f(x)$  sont quant à elles un simple mappage des valeurs transmises par la DAW au VST,  $x$  étant de manière standard compris dans l'intervalle  $[0,1]$  :

$$f(x) = \text{steps} \left\lfloor \frac{\text{min} + x(\text{max} - \text{min})}{\text{steps}} \right\rfloor \quad (1)$$

##### 3.1.3. Fonctionnement interne

Le fonctionnement du plugin reste très basique.

Chaque milliseconde le plugin va vérifier pour chacun des 10 modules si la valeur actuelle a été modifiée depuis le dernier envoi. Si c'est le cas, il génère le message envoyé en UDP vers le port et l'adresse IP fournis dans les champs correspondants ; dans le cas contraire il passe au module suivant.

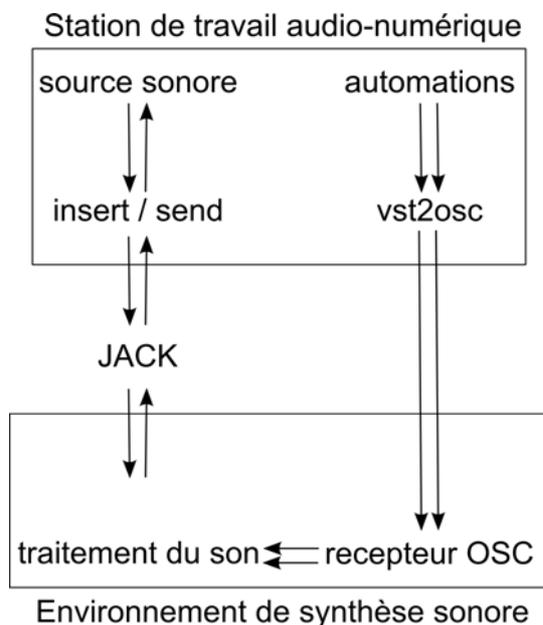
### 3.2. Gestion des flux audio

Bien que cet article se concentre sur les envois de message, une application courante pour un compositeur ferait vraisemblablement aussi intervenir de l'audio. Nous allons donc très brièvement voir une solution possible.

Au niveau de l'échange de données audio en temps-réel entre applications, les serveurs sons JACK (open-source pour linux, windows, mac) ou Soundflower (propriétaire pour mac) de Cycling '74 permettent déjà un travail de ce type grâce au routing bidirectionnel entre applications (contrairement au ReWire qui est unidirectionnel : du client ReWire vers l'hôte ReWire).[9]

Pour prendre l'exemple de JACK qui est la solution la plus ouverte et polyvalente : le nombre de connexions dont dispose chaque application peut être fixé arbitrairement et sont alors vues comme autant d'entrées/sorties standards pour chaque logiciel. Elles peuvent de plus être rappelées et connectées entre elles automatiquement grâce au système de baies de brassage (patchbays).[3]

Suivant les DAW il devient alors envisageable d'utiliser l'ESS comme un simple effet en insert appliqué directement dans chaque piste à traiter, ce qui évite par ailleurs d'avoir à utiliser 2 pistes : l'une pour l'envoi à l'ESS, l'autre pour monitorer ce qu'il renvoie. La seule contrainte étant de pouvoir router le signal vers une sortie vue comme physique par l'application et le récupérer de la même manière.



**Figure 3.** Les flux de données (audio + messages) de la solution proposée

## 4. EXEMPLES D'UTILISATION

### 4.1. Production

Pour imaginer l'utilité de ce système, il suffit de revenir à la genèse du plugin. En tant que compositeur l'auteur s'est trouvé face à un problème pour lequel le MIDI semblait être une solution peu pratique.

Après avoir codé un patch max permettant de générer la distorsion complexe d'une source sonore, il restait à trouver un moyen d'accéder aux 55 paramètres automatisables qu'il comptait.

L'utilisation de l'OSC semblait alors idéale puisqu'elle permettait grâce à un schéma de nommage logique de permettre une mémorisation rapide ainsi que d'éviter l'écueil potentiel d'un conflit dans les numéros de Control Change, surtout en cas de réutilisation ultérieure du patch.

### 4.2. Prototypage

L'auteur a récemment créé une installation multimedia interactive contrôlée par le public via une interface web distante[5]. Les modifications effectuées sur le site internet sont récupérées par l'ordinateur de l'installation puis transmises à un patch Max via un serveur OSC codé en Python.

Bien que VST2OSC ne soit pas utilisé lors des performances il permet un gain de temps considérable en court-circuitant cet interfaçage complexe et en permettant de tester l'aspect procédural, le traitement du son ainsi que la partie visuelle directement dans la DAW qui ne contenait que les extraits sonores nécessaires et 5 enveloppes d'automatisation correspondant aux 5 contrôleurs disponibles dans l'interface web.

## 5. CONCLUSION

### 5.1. Limitations

#### 5.1.1. Interceptions des frappes par l'hôte

De part sa fonction et son interface assez peu conventionnelles pour un VST, certains hôtes peuvent rendre problématique l'entrée de texte dans le plugin. Certains l'interceptent purement et simplement, d'autres ne donnent pas le « focus » total au VST et certaines combinaisons de frappes entrent alors en conflit avec des raccourcis claviers, déclenchant les actions correspondantes dans l'hôte.

Néanmoins, et même si la liste ne peut bien sûr pas être exhaustive, des solutions de contournement par le biais d'options disponibles dans les paramètres (Reaper) ou de fichiers de configuration à modifier (Live) ont pour l'instant toujours été trouvées pour permettre un parfait fonctionnement du VST.[6]

### 5.1.2. Latence

Contrairement à un traitement du son opéré directement au sein de l'hôte, le fait de devoir échanger des données audio et des messages induit une latence. Pour mettre en valeur ce problème nous allons réaliser un simple test en situation, reprenant le schéma de la figure 2 : un bruit blanc est généré par l'hôte puis routé vers et depuis Max grâce au plugin ReaInsert<sup>8</sup>.

Chaque 1000ms, VST2OSC envoie un message multipliant le signal dans Max par 0. ( $-\infty$  dB) puis par 1. (0 dB) après 500ms. Le son résultant est enregistré et fait ainsi apparaître la latence. Les résultats (Fig. 3) sont donnés avec « latence audio compensée »<sup>9</sup> La latence moyenne mesurée est d'approximativement 3ms, à noter aussi que ces résultats sont obtenus via une interface réseau « loopback » et sont donc susceptibles d'être supérieurs sur un réseau local.

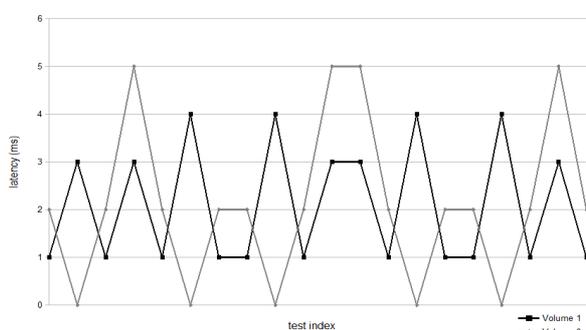


Figure 4. Test de latence

Il reste donc problématique d'utiliser ce type de configuration lors d'opérations nécessitant une grande précision temporelle mais cela semble néanmoins suffisant pour la plupart des utilisations musicales.

### 5.2. Améliorations

Dans le but de pouvoir multiplier les instances du plugin sans impacter la charge processeur, il est à envisager que seuls les modules pour lesquels un pattern a été défini ne soient vérifiés.

La taille du buffer dans lequel est stocké le message avant l'envoi est actuellement de 1024bit. Il pourrait être intéressant de laisser sa taille à la discrétion de l'utilisateur en cas par exemple de détection de perte de paquets. En effet, les messages sont transmis en UDP ; ceci permet par rapport au TCP de réduire la latence entre le serveur et le client mais la légèreté du protocole se fait au détriment de la fiabilité, rendant possible la perte d'information, la congestion du réseau et ne garanti pas l'ordre d'arrivée des paquets (deux paquets émis simultanément arriveront dans un ordre non prévisible)[2] ce qui peut se révéler très problématique. Le protocole

RUDP[1] pourrait être une solution intermédiaire, s'il venait un jour à être standardisé.

## 6. RÉFÉRENCES

- [1] Bova, T. Krivoruchka, T. "Reliable UDP Protocole", ietf.org, <http://tools.ietf.org/html/draft-ietf-sigtran-reliable-udp-00>, 1999.
- [2] Chuah, C.-N. "Transport Layer: UDP vs. TCP", UC, Davis, USA, [http://www.ece.ucdavis.edu/~chuah/classes/eec173A/eec189q-f04/lectures/L13\\_transport.pdf](http://www.ece.ucdavis.edu/~chuah/classes/eec173A/eec189q-f04/lectures/L13_transport.pdf), 2003.
- [3] Davis, P. "The JACK Audio Connection Kit", LAD Conference, Karlsruhe, RFA, [http://lac.linuxaudio.org/2003/zkm/slides/paul\\_davis-jack/why.html](http://lac.linuxaudio.org/2003/zkm/slides/paul_davis-jack/why.html), 2003.
- [4] Freed, A. Schmeder, A. Zbyszynski, M. "Open Sound Control - A flexible protocol for sensor networking", CNMAT, Berkley, USA, <http://opensoundcontrol.org/files/OSC-Demo.pdf>, 2011.
- [5] Jacquot, M. "--Manipulation--", peter1island.com, <http://peter1island.com/art/manipulation/>, 2013.
- [6] Jacquot, M. "vst2osc", Paris, France, <http://peter1island.com/technicalStuff/audio-plugin/vst2osc-rtas2osc.php>, 2013.
- [7] midi.org "MIDI Manufacturers Investigate HD Protocol", midi.org, <http://www.midi.org/aboutus/news/hd.php>, 2013.
- [8] midi.org "White Paper: Comparison of MIDI and OSC", midi.org, <http://www.midi.org/aboutmidi/midi-osc.php>, 2008.
- [9] propellerheads.com "ReWire - Technical information", propellerheads.com, [http://www.propellerheads.se/developer/index.cfm?fuseaction=get\\_article&article=rewiretechinfo](http://www.propellerheads.se/developer/index.cfm?fuseaction=get_article&article=rewiretechinfo), 2013.
- [10] Steinmetzer, R. "rs.delos, sort of a crossbreed of the pre-Max5 timeline object", arts.lu, [http://arts.lu/roby/index.php/site/maxmsp/rs\\_delos](http://arts.lu/roby/index.php/site/maxmsp/rs_delos), 2013.

NB : Chaque lien a été vérifié le 07 avril 2013.

<sup>8</sup>ReaInsert : <http://wiki.cockos.com/wiki/index.php/ReaInsert>

<sup>9</sup>Le plug-in ReaInsert permet de détecter et de compenser le délai dû à l'aller/retour entre applications.

# LIBTUILES : UN MOTEUR D'EXÉCUTION MULTI-ÉCHELLE DE PROCESSUS MUSICAUX HIÉRARCHISÉS

Florent Berthaut, David Janin et Myriam DeSainte-Catherine

Université de Bordeaux, LaBRI UMR 5800,

351, cours de la libération,

F-33405 Talence

{berthaut|janin|myriam}@labri.fr

## RÉSUMÉ

La *libTuiles* est une librairie expérimentale réalisée afin d'évaluer le potentiel multi-échelle et interactif de la modélisation par *tuilage* de processus musicaux hiérarchisés.

## 1. INTRODUCTION

### 1.1. Problématique

Il existe aujourd'hui de nombreux langages d'écriture et de traitements sonores, qu'ils soient textuels comme *Supercollider/chuck* [14] ou *Faust*[7], ou visuels comme *Max/Msp* et *PureData* [5]. L'écriture de pièces musicales reste cependant une tâche délicate car ces langages ne proposent que bien peu de métaphores ou de paradigmes de programmation permettant de décrire facilement (et de façon compositionnelle) le placement temporel des objets sonores à produire. Plus encore, si l'on souhaite intégrer une gestion d'évènements externes - déclencheurs de processus audio - le maintien de la cohérence temporelle et/ou rythmique de la pièce écrite peut devenir inextricable.

C'est pour remédier à cela que le séquenceur interactif *i-score* [2] propose explicitement un outil de spécification du placement des objets sonores les uns par rapport aux autres. En intégrant des points de contrôle explicites et un mécanisme de résolution de contraintes de placement, il permet une *écriture du temps* plus abstraite. Néanmoins, faute de mécanismes de contrôles dynamiques puissants - conditionnelle, boucle - l'applicabilité d'*i-score* reste encore limitée.

Le travail présenté ici, qui s'appuie sur une extension d'*i-score*, vise à expérimenter les possibilités d'intégration des outils de description programmatique de compositions musicales qui sont offerts par les langages évoqués ci-dessus, avec cette capacité d'*écriture (dynamique) du temps* apportée par *i-score*.

### 1.2. Caractéristiques fonctionnelles

Implémentation d'une proposition récente d'algèbre de synchronisation de signaux audio ou musicaux tuilés [4], la *libTuiles* apparaît avant tout comme un *outil de mixage*

multi-échelle et hiérarchique qui, en intégrant un module de lecture granulaire de flux audio, se révèle d'un usage particulièrement souple.

Couplée avec d'autres outils existants de l'informatique musicale tels que la librairie audio *libAudioStream* [6], le langage synchrone *Faust*[7] ou le séquenceur interactif *i-score*[2], la *libTuiles* vise à devenir le premier prototype de *moteur d'exécution* pour le T-calcul, une proposition de langage de programmation intégrant la programmation par *tuilage* (ou *tiled programming*) [11].

Inspiré par l'instrument *Drile* [3], un mécanisme d'exécution cyclique et hiérarchique associé à la *libTuiles* offre aussi, via un couplage avec le moteur temps-réel *JACK* et une interface, le *simpleTuilesLooper*, un outil de performance temps-réel qui nous permet d'expérimenter les métaphores et les concepts sous-jacents.

## 2. MODÈLES DE PROCESSUS MUSICAUX

Dans cette section, nous passons en revue quelques uns des concepts clés qui sous-tendent la *libTuile*. Il s'agit en particulier de comprendre la structure *comportementale* des objets musicaux qui sont modélisés et exécutés dans les systèmes musicaux interactifs.

### 2.1. Structure comportementale des objets musicaux

Il existe aujourd'hui de nombreux modèles de représentation des objets musicaux. Chacun répond à des contraintes d'usages spécifiques et apparaît donc comme une approximation des objets décrits.

Par exemple, la *partition musicale classique* est largement dédiée à une lecture rapide par les musiciens de phrases mélodiques, qui peuvent être multiples, et qu'ils devront *interpréter* et *synchroniser* entre elles.

Si on s'intéresse aux comportements décrits par ces partitions, cet aspect graphique de représentation musicale est d'un intérêt marginal. Cependant, on voit apparaître dans ces notations deux dimensions fondamentales qui participent à la structuration du comportement des objets musicaux décrits.

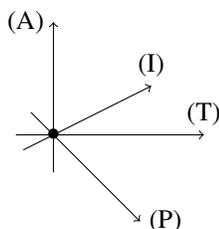
La première : la dimension *temporelle* (*T*), apparaît en effet dans cette écriture linéaire de la musique, faite de

successions de notes, d'accords, de mesures, etc. . . . La seconde, la dimension *parallèle* ( $P$ ), apparaît aux travers des voix multiples qui apparaissent sur ces partitions, qu'elles soient distribuées sur plusieurs portées ou, tout aussi bien, au sein d'une même portée, par exemple dans la réduction pour piano d'une pièce pour orchestre.

En analysant plus en détails ces partitions de musique classique, ou bien en annotant ces mêmes partitions de commentaires harmoniques tels que pour les *grilles d'accords* en jazz ou les *basses chiffrées* en musique baroque, on voit apparaître une troisième dimension tout aussi prépondérante. C'est l'*abstraction* qui rappelle que ces notations musicales n'offrent qu'une description approximative de la musique, le niveau d'approximation, c'est à dire d'abstraction, pouvant varier d'un style musical à l'autre.

Ce niveau d'abstraction peut aussi varier à l'intérieur même d'une partition donnée à travers les annotations stylistiques qui accompagnent les portées. Par exemple, des *regroupements* en accords, motifs, phrases, mouvements, etc., visent aussi à décrire des structures de plus en plus abstraites [13] qui rendent compte, parmi d'autres aspects, de l'intention du compositeur. . . .

Plus implicite, une quatrième dimension apparaît dès lors qu'on modélise un système musical interactif. Cette dimension de l'*Interaction* ( $I$ ) permet de rendre compte, par exemple, de la façon dont les musiciens peuvent/doivent adapter leur jeu au jeu d'un musicien soliste.



**Figure 1.** Structuration des objets musicaux

Autrement dit, un système interactif musical semble pouvoir être décrit comme une certaine fonction, partielle, de l'espace  $(T) \times (P) \times (A) \times (I)$  dans un certain espace de valeurs ( $V$ ).

En informatique, on modélise communément les systèmes réactifs, ouverts, par la représentation arborescente de leurs comportements possibles. Dans ce type de modèle, structuré dans le plan  $(T) \times (I)$ , chaque embranchement décrit, pour chaque événement reçu, les changements d'état possibles (et donc les changements de comportements possibles) du système modélisé. Ainsi, l'arrivée d'événements externes se traduit par des *alternatives* comportementales.

A contrario, la partition d'un quatuor, est décrite dans le plan  $(T) \times (P)$ . En faisant abstraction de la liberté d'interprétation des musiciens, il n'y a pas d'interaction. L'avancement dans la partition, la *timeline*, est linéaire.

Autre exemple, un *piano roll* se place typiquement dans le plan  $(T) \times (P)$ . Plus abstrait que la partition évoquée

ci-dessus, il ne décrit plus que les moments d'activation de telle ou telle piste.

Bien entendu, cette représentation spatiale de la structure des objets musicaux est parfaitement discutable. Notre approche vise à mieux comprendre la structure dynamique, au sens purement informatique, des objets musicaux mis en oeuvre dans un système informatique.

D'autres approches, plus orientées analyse musicale, font apparaître une structure bien plus complexe et plus riche. En particulier, nous ne disons rien ici de la structure de l'espace ( $V$ ), évoquées ci-dessus, des «valeurs musicales» possibles.

## 2.2. Une dimension temporelle multi-échelle

Une difficulté supplémentaire propre à la représentation des processus musicaux vient du fait que la description même de ces dimensions peut changer de nature en fonction du niveau d'abstraction choisi.

Cette nature a priori hétérogène des modèles de comportements musicaux est particulièrement flagrante dans le plan  $(T) \times (A)$ . En effet, dans ces représentations de processus musicaux, le temps est multi-échelle. On peut ainsi distinguer au moins quatre types d'échelles de temps de natures a priori différentes.

*Le temps logique (causal).* Les événements musicaux sont unitaires, positionnés les uns par rapport aux autres sur une échelle de temps logique, e.g. *avant, après, en même temps*.

Les intervalles de temps réels entre deux événements logiques peuvent varier de quelques secondes, par exemple pour des progressions harmoniques, à quelques minutes, pour des successions de thèmes, ou bien plus encore pour, par exemple, des séquences de mouvements voire de pièces musicales entières.

*Le temps symbolique (quantifié).* Les événements musicaux ont maintenant une durée et une position temporelle calculées sur une (ou des) échelle(s) de temps symbolique définie par référence à une (ou plusieurs) unité(s) de référence telle que la battue ou la mesure de l'ordre de la seconde.

Les concepts de positionnements temporels logiques évoqués ci-dessus échouent à décrire seuls une structuration temporelle plus complexe faite de superpositions partielles [8].

*Le temps réel asynchrone (placé).* Les événements musicaux sont maintenant positionnés, par exemple lors d'une performance, sur une échelle de temps réel. Ce placement temporel, perceptible par l'oreille humaine, est réalisé avec une précision de  $10^{-1}$ s à  $10^{-3}$ s.

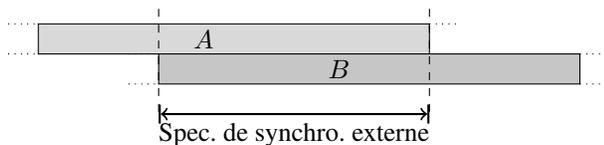
A cette échelle de temps, l'apparition des événements reste irrégulière. De nombreux phénomènes de tension et de résolution rythmique peuvent être obtenus, en musique savante comme en musique pop ou traditionnelle, par un positionnement temporel choisi *autour* de la pulsation abstraite. On est donc dans un flux événementiel asynchrone.

*Le temps synchrone («continu»).* Découpés en grains ou échantillons réguliers (dans le cas numérique) ou résul-

tant de processus continus (dans le cas acoustique) les sons musicaux obtenus sont maintenant réalisés (ou approximatés) sur une horloge régulière de période de  $10^{-1}$ s à  $10^{-5}$ s.

### 2.3. De la synchronisation aux tuilages

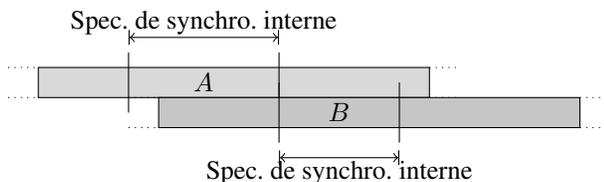
De pratique courante en traitement audio, la synchronisation consiste à organiser les processus musicaux dans le plan  $(T) \times (P)$ . Il s'agit de positionner dans le temps, l'un par rapport à l'autre, deux signaux, pour ensuite appliquer une opération de transformation (ou fusion), par exemple un *cross-fade*, sur les parties de signaux ainsi superposées.



La synchronisation de deux signaux requiert en général une analyse de ces signaux afin de déterminer le début et la fin de la zone de recouvrement. Par exemple, le séquenceur interactif *i-score* [2] permet, à l'aide de contraintes exprimées en logique de Allen[1], de spécifier les recouvrements autorisés.

Cette approche par *spécification externe*, qui s'appuie sur les signaux à mixer et l'analyse croisée de leurs superpositions possibles, n'est pas compositionnelle. Chaque synchronisation nécessite une analyse des signaux à mixer.

Le *tuilage* apparaît dès lors que ces informations de synchronisation sont intégrées aux signaux eux-mêmes [8] produisant ainsi des *signaux tuilés* [4]. Plus précisément, tout comme une tuile de toit, ou une séquence musicale avec barres de mesures, chaque signal peut être associé à un *intervalle de synchronisation*, qui spécifie comment le signal sera positionné lors d'une opération de synchronisation vis à vis des signaux (tuilés) voisins. On passe ainsi d'une spécification externe de la synchronisation à une *spécification interne* de la synchronisation.



Par opposition à cet intervalle de synchronisation, l'intervalle couvert par la totalité du signal est appelé *intervalle de réalisation*.

La synchronisation  $SEQ(A, B)$  de deux signaux tuilés  $A$  et  $B$  revient donc à positionner les fenêtres de synchronisation de ces signaux en séquence, sans autre analyse ou calcul. L'opération de fusion (paramétrable) est alors appliquée. Le signal obtenu reste tuilé. L'opération SEQ ainsi défini est compositionnelle. Plus encore, combinée à des fonctions de fusions variables, elle se révèle particulièrement expressive.

### 2.4. Une algèbre de synchronisation

Le produit de synchronisation  $SEQ(A, B)$  évoqué ci-dessus conduit donc à la définition d'une algèbre de synchronisation de processus audio ou musicaux qui est décrite dans [4].

La structure mathématique sous-jacente qui permet de décrire ces paramètres de tuilages se révèle particulièrement robuste. C'est un monoïde inversif [12]. Une étude théorique récente de la théorie des langages induite par ces structures démontre par ailleurs tout à la fois sa simplicité et sa puissance [10, 9].

On peut aussi dériver de la structure inversive du produit de synchronisation SEQ de nombreux autres opérateurs [4]. On dispose par exemple d'opérateurs de synchronisation à gauche FORK ou à droite JOIN.

L'ajout de deux opérateurs additionnels de resynchronisation RESYNC, agissant sur la position relative de la fenêtre de synchronisation, et d'expansion / contraction STRETCH, nous conduit alors à la définition d'une algèbre de synchronisation de signaux tuilés particulièrement souple d'utilisation [4].

Plus encore, le produit  $SEQ(A; B)$  peut être interprété de plusieurs façons selon le niveau d'abstraction choisi. En effet, sur l'échelle de temps logique, causal, le produit  $SEQ(A; B)$  peut être vu comme le positionnement séquentiel de l'évènement  $A$  suivi de l'évènement  $B$ . Au contraire, sur l'échelle de temps synchrone, ce même produit  $SEQ(A; B)$  décrit, sans ambiguïté, la *synchronisation temporelle* du signal  $A$  par rapport au signal  $B$ . Autrement dit, l'internalisation des paramètres de synchronisation conduit à un modèle de *signaux tuilés* qui est intrinsèquement *multi-échelle*.

Son intégration récente à un langage de programmation : le T-calcul [11] semble ainsi particulièrement prometteuse. Elle permet en effet, dans un formalisme unique, d'intégrer les deux échelles de temps extrêmes évoquées ci-dessus. On se convainc donc sans difficulté qu'il permet aussi de décrire les échelles de temps intermédiaires, du temps symbolique (quantifié) ou du temps réel asynchrone (placé).

## 3. IMPLÉMENTATION

Dans cette section, nous décrivons les composants logiciel de la libTuiles, notamment le mécanisme de commandes pour la communication entre threads asynchrones et synchrones.

### 3.1. LibTuiles : construction et exécution d'arbres de tuiles

LibTuiles est une bibliothèque C++ permettant de construire et d'exécuter des arbres de tuiles. Dans ces arbres, chaque tuile possède un identifiant unique sous forme d'un nombre entier non signé. La construction et la manipulation des arbres de tuiles s'effectuent à l'aide des méthodes suivantes :

**addLeaf**(*const float& d, unsigned int& id*) : créé une nouvelle tuile feuille de durée initiale *d* et assigne son identifiant à la variable *id*.

**addLoop**(*const unsigned int& id1, unsigned int& id*) : créé une nouvelle tuile résultant de l'application de l'opérateur LOOP sur la tuile *id1* et assigne son identifiant à la variable *id*.

**addSeq**(*const unsigned int& id1, const unsigned int& id2, unsigned int& id*), **addFork**(...) et **addJoin**(...) : créent tous trois une nouvelle tuile résultant de l'application respective des opérateurs SEQ, FORK et JOIN sur les tuiles *id1* et *id2* et assignent son identifiant à la variable *id*.

**setTuileLength**(*const unsigned int& id, const float& d*) : applique l'opérateur STRETCH sur la tuile *id* afin de redimensionner sa fenêtre de réalisation à la durée *d*.

**setTuileLeftOffset**(*const unsigned int& id, const float& lo*) : applique l'opérateur RESYNC sur la tuile *id* afin de modifier l'offset gauche de sa fenêtre de synchronisation.

**setTuileRightOffset**(*const unsigned int& id, const float& ro*) : applique l'opérateur RESYNC sur la tuile *id* afin de modifier l'offset droit de sa fenêtre de synchronisation.

**setBpm**(*const float& bpm*) : définit le tempo de lecture de l'arbre.

**setRoot**(*const unsigned int& id*) : définit la tuile *id* comme racine de l'arbre.

**play()** et **stop()** : lancent et arrêtent respectivement la lecture de l'arbre.

**removeTuile**(*const unsigned int& id*) : retire la tuile *id* de l'arbre.

**clear()** : supprime toutes les tuiles de l'arbre.

De manière interne, la construction et l'exécution des arbres s'effectuent dans un thread séparé, afin d'éviter des ralentissements dus à des calculs effectués dans le thread principal de l'application, e.g. le thread de l'interface graphique. Le mécanisme de communication entre les threads est décrit dans la section 3.3.

Lors de la lecture de l'arbre, la progression temporelle s'effectue dans la racine et est répercutée en descendant dans l'arborescence. Chaque opérateur fait en effet avancer la position dans ses enfants en fonction des paramètres de leurs intervalles de synchronisation et de réalisation. A chaque temps *t* est donc calculée la position dans chacune des tuiles. Il est de la même façon possible de connaître la position absolue de chacune des tuiles dans l'arbre. Puisque l'avancement temporel s'effectue pour chacun des nœuds de l'arbre de manière relative à son nœud parent, il est possible de modifier dynamiquement l'arbre pendant la lecture.

Des commandes d'activation et de désactivation sont envoyées depuis le thread de lecture en fonction de la position de la lecture dans la tuile, la tuile étant active entre 0 et la longueur de son intervalle de réalisation. Des commandes de durée sont aussi envoyées lorsque la taille de l'intervalle de réalisation est mise à jour ou lors des changements de tempo, ainsi que des commandes de position

absolue quand l'arbre est modifié. Ainsi, un moteur de synthèse/traitement synchrone, tel que celui décrit dans la section 3.2 peut recevoir toutes les informations nécessaires au positionnement temporel des différents processus associés aux tuiles.

Les propriétés des tuiles sont accessibles grâce à la fonction *getTuileProps(const unsigned int& id)* qui renvoie une structure associée à la tuile d'identifiant *id* (si elle existe) et composée des différentes propriétés de la tuile : taille d'intervalle de réalisation, offsets droit et gauche de l'intervalle de synchronisation, position absolue dans l'arbre. Ces propriétés étant susceptibles d'être modifiées lors de manipulations de l'arbre, ce mécanisme permet de mettre à jour leur représentations dans une interface graphique.

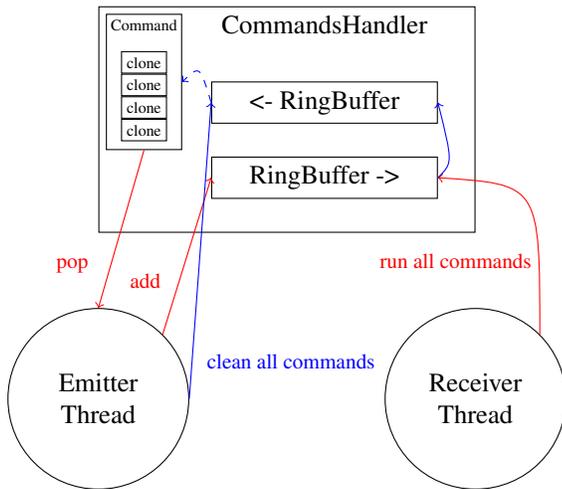
### 3.2. Moteur synchrone connecté à la libTuiles pour l'organisation temporelle de processus sonores

Le moteur d'exécution des tuiles, asynchrone, est couplé à un moteur de traitement/synthèse audio synchrone basé sur le serveur de son JACK. Ce moteur reçoit les commandes, décrites dans la section précédente, de configuration temporelle des processus associés aux tuiles par identifiant.

Les processus peuvent être de deux types. Les processus de lecture audio permettent de lire des fichiers son. Ils gèrent également l'étirement temporel, dû aux changements de tempo de l'arbre et à l'opérateur STRETCH, grâce à l'utilisation de la synthèse granulaire. A la vitesse de lecture initiale, les grains joués se chevauchent à moitié et le pas de déplacement dans le fichier son entre deux grains est égal à une moitié de grain. Lorsque la vitesse de lecture est ralentie ou que la durée des tuiles est allongée, le pas de déplacement est réduit et un offset aléatoire est ajouté afin de prévenir un effet métallique dû à la proximité des grains successifs. De plus le chevauchement entre grains est accentué afin de lisser le son. Lorsque la vitesse de lecture augmente, le pas de déplacement est augmenté et le chevauchement est également accentué afin de réduire les variations d'amplitude entre les grains successifs. Cette méthode de synthèse, malgré les différents artefacts et distorsions qu'elle génère comparativement à d'autres techniques d'étirement temporel, permet d'étirer les sons en temps-réel pour un coût en calcul très faible, ainsi que de se repositionner dans le son sans saut audible dans le signal.

Il est également possible d'associer des processus de traitement/synthèse sonore aux tuiles, par le biais d'effets FAUST. Il est ensuite possible de rediriger la sortie de n'importe quel processus vers ce traitement FAUST. A chaque fenêtre de rendu audio, le processus FAUST lit les échantillons sur la sortie du processus auquel il est connecté. Les traitements ne sont ainsi effectués que lorsque les processus se chevauchent temporellement, ce qui est contrôlé en manipulant les fenêtres de synchronisation des tuiles associées.

### 3.3. Architecture logicielle orientée objet pour la communication entre échelles temporelles



**Figure 2.** Architecture logicielle pour la transmission de commandes entre deux threads d'échelles temporelles différentes.

Un aspect important de l'architecture de la libTuiles est l'utilisation de Commandes, représentée sur la Figure 2. Ces composants logiciels permettent de respecter les contraintes dues à la connexion entre les échelles événementielles, temps-réel asynchrone et temps-réel synchrones, toutes prises en charge par des threads différents. En particulier, le thread temps-réel synchrone, donc géré par le serveur de son JACK, ne doit pas comporter d'allocations/désallocations mémoire ni de mécanismes de verrouillage. L'architecture proposée s'appuie sur plusieurs patrons de conception objet connus parmi lesquels le patron Prototype, le patron Abstract Factory et le patron Commande.

Une classe *CommandsHandler* gère la création et la manipulation d'instances de la classe *Command* ainsi que leur envoi d'un thread émetteur à un thread récepteur. Une instance de cette classe est donc partagée entre les classes gérant les threads. Des associations "chaîne de caractère - Commande" sont tout d'abord ajoutées à cette classe. Par exemple, le *CommandsHandler* du moteur synchrone possède les commandes *ActivateProcess* et *DeactivateProcess*, identifiées par leurs noms. Lors de l'ajout, une instance de chaque commande est créée comme prototype avec un certain nombre de clones possédant un pointeur vers leur prototype. Ainsi des classes associées à chaque message à faire passer d'un thread à l'autre peuvent être définies simplement en les faisant hériter de la classe *Command* et en redéfinissant leur fonction *run()* afin de manipuler une structure de données destinataire dans le thread récepteur, par exemple activer/désactiver un processus.

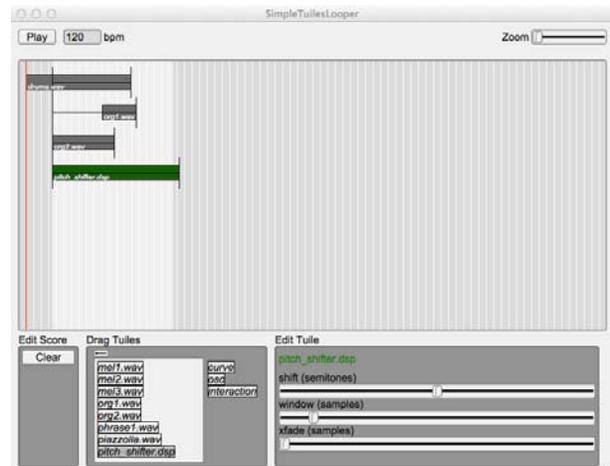
Lors de l'exécution, le thread émetteur récupère un pointeur vers une instance de la commande désirée en appelant une méthode *popCommand* du *CommandsHandler* avec le nom de la *Commande*. Cette instance est alors retirée des clones disponibles du prototype de cette *Com-*

*mande* et peut être ajustée avec différents paramètres tels que l'identifiant de la tuile, la nouvelle durée/position ... Aucune allocation mémoire n'est donc effectuée pendant l'exécution. Le pointeur vers cette commande est ensuite redonné au *CommandsHandler* et transite d'un thread à l'autre grâce à un *RingBuffer*, permettant ainsi d'éviter les verrouillages des threads.

Le thread récepteur appelle de son côté la méthode *runCommands* du *CommandsHandler* qui va lire les commandes depuis le *RingBuffer*, appeler leur fonction *run* et les renvoyer au thread émetteur par un deuxième *RingBuffer* en sens inverse. Finalement, le thread émetteur appelle la méthode *cleanCommands* du *CommandsHandler*. Lors de cette méthode, chaque pointeur revenant du thread récepteur est remis dans la liste des clones disponibles du prototype associé.

Cette architecture logicielle, tout en respectant les principes de programmation objet, permet de faire transiter des commandes entre plusieurs threads, sans allocations mémoire ni mécanismes de verrouillage. Elle est donc particulièrement adaptée aux systèmes mêlant différentes échelles temporelles dont certaines très sensibles aux retards comme les threads audio temps-réel.

## 4. EXPÉRIMENTATION INTERACTIVE



**Figure 3.** SimpleTuilesLooper permet d'expérimenter la composition temporelle de processus en s'appuyant sur la libTuiles

### 4.1. Le SimpleTuilesLooper

*SimpleTuilesLooper* est une application permettant de tester la composition temporelle avancée de processus notamment pour la performance live, en s'appuyant sur la libTuiles et le moteur synchrone décrits ci-dessus.

Cette application définit comme racine de l'arbre de tuiles une tuile *Loop* dont le premier enfant est une tuile feuille. Toutes les autres tuiles ajoutées à l'arbre sont synchronisées à cette première tuile feuille, dont la fenêtre de

synchronisation, dynamiquement modifiable, définit par conséquent la fenêtre de la tuile *Loop* et ainsi la zone de lecture en boucle dans l'arbre.

*SimpleTuilesLooper* permet de créer des tuiles à partir de fichiers audio et de fichiers dsp FAUST et de combiner ces tuiles pour construire un arbre à l'aide de la métaphore graphique du drag and drop. Les fichiers sont prélevés depuis un explorateur de fichiers et déposés sur la partition. Soit ils sont déposés sans contact avec d'autres tuiles et donc insérés par l'opérateur FORK avec la tuile racine, soit ils sont placés en composition séquentielle ou parallèle avec une tuile existante et insérés dans l'arbre avec les opérateurs associés.

Une boîte de dialogue permet de régler les paramètres des effets FAUST ainsi que la connexion des processus entre eux. L'arbre peut ensuite être lu et le tempo de lecture modifié. Un aperçu de l'interface du *SimpleTuilesLooper* est donné sur la figure 3.

#### 4.2. Interaction et placement temporel dynamique

Le *simpleTuilesLooper* permet donc d'exécuter en boucle des expressions de tuiles dont les définitions peuvent être éditées et modifiées de façon interactive, grâce au moteur d'exécution *libTuile*.

De plus, via l'utilisation de buffers d'entrées comme paramètres de transformation de tuiles, on peut facilement enrichir l'interaction offerte par le *simpleTuilesLooper*. Ce faisant, on retrouve la puissance d'expression des systèmes synchrones, tel que Pure data ou Max/Msp, qui sont naturellement paramétrés par de tels flux d'entrée.

Un mécanisme additionnel d'écriture interactive du temps est proposé. Une tuile d'interaction, *ad hoc*, permet d'associer à des événements *midi* le placement, en séquence ou en parallèle, de signaux tuilés.

### 5. CONCLUSION

La *libTuiles* et le *simpleTuilesLooper* présentés ici permettent donc d'expérimenter l'écriture dynamique du temps que permet la programmation par tuilage.

Les concepts et les développements associés sont aujourd'hui assez avancés pour une telle expérimentation.

Bien entendu, dans la perspective plus long terme de fournir un moteur d'exécution au *T-calcul* [11] et de produire une version 2.0 du logiciel *i-score* [2] qui intégrera des structures de contrôles plus puissantes, cette présentation n'est qu'un rapport d'étape. Il reste encore à détailler la sémantique opérationnelle du T-calcul [11] pour conduire à une implémentation complète de tout les mécanismes d'écriture du temps qu'il induit.

### 6. REFERENCES

- [1] J. Allen and G. Ferguson. Actions and events in interval temporal logic. In Oliviero Stock, editor, *Spatial and Temporal Reasoning*, pages 205–245. Springer Netherlands, 1997.
- [2] A. Allombert, M. Desainte-Catherine, and G. Assayag. Iscore : a system for writing interaction. In *Third International Conference on Digital Interactive Media in Entertainment and Arts (DIMEA 2008)*, pages 360–367. ACM, 2008.
- [3] F. Berthaut, M. Desainte-Catherine, and M. Hachet. Drile : an immersive environment for hierarchical live-looping. In *Proceedings of New Interfaces for Musical Expression (NIME10)*, pages 192–197, Sydney, Australia, 2010.
- [4] F. Berthaut, D. Janin, and B. Martin. Advanced synchronization of audio or symbolic musical patterns : an algebraic approach. *International Journal of Semantic Computing*, 6(4) :1–19, 2012.
- [5] Alessandro Cipriani and Maurizio Giri. *Electronic Music and Sound Design - Theory and Practice with Max/Msp*. Contemponet, 2010.
- [6] S. Letz et al. The LibAudioStream library, 2012. <http://libaudiostream.sourceforge.net/>.
- [7] D. Fober, Y. Orlarey, and S. Letz. FAUST architectures design and OSC support. In *14th Int. Conference on Digital Audio Effects (DAFx-11)*, pages 231–216. IRCAM, 2011.
- [8] D. Janin. Vers une modélisation combinatoire des structures rythmiques simples de la musique. *Revue Francophone d'Informatique Musicale (RFIM)*, 2, 2012.
- [9] D. Janin. Algebras, automata and logic for languages of labeled birooted trees. Technical Report RR-1467-13, LaBRI, Université de Bordeaux, 2013.
- [10] D. Janin. Overlapping tile automata. In *8th International Computer Science Symposium in Russia (CSR)*, LNCS (to appear). Springer-Verlag, 2013.
- [11] D. Janin, F. Berthaut, M. DeSainte-Catherine, Y. Orlarey, and S. Salvati. The T-calculus : towards a structured programming of (musical) time and space. Technical Report RR-1466-13, LaBRI, Université de Bordeaux, 2013.
- [12] M. V. Lawson. *Inverse Semigroups : The theory of partial symmetries*. World Scientific, 1998.
- [13] F. Lerdahl and R. Jackendoff. *A generative theory of tonal music*. MIT Press series on cognitive theory and mental representation. MIT Press, 1983.
- [14] S. Wilson. *The SuperCollider Book*. Cambridge : The MIT Press, 2011.

# SYSTÈME DE CAPTATION OPTIQUE POUR LA TRANSCRIPTION AUTOMATIQUE DE LA MUSIQUE DE CITHARE MALGACHE *MAROVANY*

*Dorian Cazau, Olivier Adam*  
Laboratoire d'Acoustique Musicale (LAM)  
UPMC, 11 rue de Lourmel - 75015 PARIS  
cazau@lam.jussieu.fr

*Marc Chemillier*  
Centre d'analyse et de mathématiques sociales  
EHESS, 198 avenue de France - 75013 Paris  
chemilli@ehess.fr

## RÉSUMÉ

Dans cet article, nous introduisons un système de captation optique original dédié à l'analyse musicale de la cithare malgache. Ce système a été appliqué à l'extraction automatique d'information musicale, assistée par des méthodes de pré-traitement pour la détection et la caractérisation acoustique des notes de musique. L'objectif visé est de réaliser une transcription systématique du répertoire musical de la cithare *marovany* en contexte de transe *tromba*, et de classifier les airs musicaux récurrents en terme de devises identificatoires. A travers ce travail nous recommandons pour remplir cet objectif l'utilisation d'un enregistreur multicanal à capteurs optiques synchronisés, en avançant les qualités technologiques suivantes : captation individuelle des cordes, fort rapport signal sur bruit (forte sensibilité au déplacement des cordes / insensibilité aux sources acoustiques externes), démarcation systématique des notes par des silences issues du contact doigt-corde. De telles caractéristiques simplifient grandement la tâche délicate de transcription automatique de musique polyphonique en milieu bruité.

## 1. INTRODUCTION

La *marovany* est une grande cithare montée sur une caisse rectangulaire en bois de récupération. Ses cordes métalliques en câbles de freins de moto peuvent mesurer jusqu'à 1 m 20, et sont tendues des deux côtés de la caisse. Elles sont ensuite clouées sur celle-ci, reposant à chaque extrémité sur un chevalet (en bois ou en métal), et sont surélevées par des tasseaux de bois placés à différents endroits suivant la hauteur que l'on veut donner à la corde. Elles forment, comme pour la cithare tubulaire (*valiha*), une échelle diatonique alternée. D'un point de vue musical, le répertoire joué à la cithare *marovany* se veut comme un flot mélodique arpégé. Bien qu'il n'y ait pas de véritable écriture verticale dans son répertoire, à l'exception de quelques accords (c'est-à-dire au minimum deux notes simultanées) ponctuels, les notes sont jouées très rapidement, sont rarement étouffées à l'intérieur d'une phrase musicale et les cordes résonnent par sympathie les unes avec les autres. Ces caractéristiques font qu'un certain nombre de notes se retrouvent très sou-

vent en simultanéité et confèrent à cette musique une complexité d'analyse digne d'une musique polyphonique. Le type de bois, les dimensions et le nombre de cordes utilisés ne sont pas fixés. On peut en effet trouver des cithares en bois léger ou dur, de longueur pouvant varier de 1 à 2 mètres de long, possédant 8, 10, 12 cordes sur des tasseaux allant de 2 cm de haut à 0.5 cm (des tasseaux plus étroits sont connus pour donner plus de son car les cordes sont plus près de la caisse, sonnante *mafo be*, « plus fort »). La cithare étudiée ici (photo à la figure 1) compte 13 cordes sur chacune des deux faces, avec une tessiture couvrant plus de deux octaves.

En plus de ses qualités purement musicales, l'étude de la cithare sur caisse *marovany* présente un intérêt majeur du point de vue anthropologique par le fait que sa musique est utilisée dans le culte de possession *tromba*. La transe est induite musicalement, en ce sens qu'elle est provoquée et entretenue par des stimuli musicaux. En contexte de *tromba*, le mode de jeu de la *marovany* est constitué essentiellement de motifs mélodiques, enchaînés et transformés progressivement à travers de multiples répétitions obsédantes. L'instrument est souvent accompagné d'un hochet dit *kantsa*, formé d'une boîte de conserve remplie de graines et clouée sur un manche en bois, fournissant le support rythmique de cette musique. Cette fonction sociale de la *marovany* détermine probablement son répertoire musical. En effet, les musiques de possession prennent souvent la forme de devises identificatoires [1], associant une formule musicale à une entité symbolique extramusicale. Une partie au moins du répertoire de la *marovany* pourrait ainsi se présenter comme une table d'associations et de correspondances entre formules musicales et certaines divinités [2]. Un autre aspect de sa fonction dans le culte de possession *tromba* est qu'elle participe à l'effervescence collective propice au déclenchement de la transe, en contribuant à faire « chauffer » l'atmosphère. Le moyen musical de cet échauffement est peut-être l'utilisation de formules musicales jouées selon une progression judicieusement contrôlée par le musicien. Afin de mieux comprendre ces deux aspects du lien de la cithare avec la transe, une analyse de son répertoire doit être réalisée, avec un inventaire systématique des formules musicales, si possible en relation avec des indices comportementaux, et des associations air/divinité. L'étude de ce répertoire spé-

cifique, fondée sur des critères musicaux, permettra d'apporter des éléments inédits à la question fascinante des relations entre la musique et la transe. Dans une perspective plus large, les études de la transe qui abordent la question de fond des mécanismes neurophysiologiques [3, 4], bénéficieraient des données musicales extrêmement précises et complètes fournies par nos capteurs.



**Figure 1.** Photo de la cithare malgache (en haut) et du dispositif de captation optique (en bas)

Les cithares malgaches, en particulier la *valiha*, considérée comme l'instrument national et de laquelle dérive la *marovany*, ont fait l'objet de nombreuses recherches [5, 6] d'ordre ethnomusicologique. Cependant, il n'existe pas à notre connaissance de projet d'analyses et de classifications systématiques du répertoire des airs musicaux de la cithare *marovany*, en fonction de leurs propriétés musicales (rythmiques, structurelles, modales et acoustiques) et extra-musicales (contenu sémantique et symbolique). Dans le but de fournir un aperçu significatif de la relation transe-musique à Madagascar, des investigations de terrain doivent être entreprises, avec l'aide de systèmes automatisés pour l'extraction et la caractérisation de patterns musicaux. Ces données traitées par l'informatique pourront ensuite servir à évaluer statistiquement certaines occurrences musicales en fonction de différentes séances de transe, et de dresser des corrélations entre des patterns musicaux et comportementaux au cours d'une séance de transe (par exemple, la manière dont un musicien injecte des indices de renouvellement musicaux pour relancer l'intérêt ou intensifier la danse). Une étude systématique de ces concordances permettrait d'établir le catalogue raisonné du répertoire commun aux joueurs de *marovany* en contexte de transe. L'automatisation de l'analyse de cette musique est rendue impérieuse par l'aspect laborieux que la tâche manuelle de transcription représente, sachant que des séances de transe peuvent durer plusieurs heures et qu'il n'existe aucun support écrit de musique. Aussi, la complexité de sa transcription (vitesse de jeu, caractère polyphonique, milieu bruité par des sources acoustiques) peut entraîner une grande variabilité dans les résultats, sans estimation possible de la qualité de la transcription. Durant les trances étudiées, ces transcriptions optiques sont accompagnées d'enregistrements audiovisuels

(des exemples de vidéo et d'extraits audio peuvent être trouvés sur la page web [7]), qui permettent l'analyse parallèle des indices comportementaux rapportés ci-dessus, mais ne constituent pas un support optimal pour la transcription musicale car comportant plusieurs sources musicales compétitives dans un environnement bruyant. Pour répondre à cette problématique, ce papier présente en section 2 un système de captation optique dédié à l'enregistrement in situ d'airs musicaux de la *marovany*. Ce système de captation a ensuite été intégré à une chaîne d'acquisition et de traitements visant à l'extraction automatique d'informations musicales, présentée en section 3.

## 2. SYSTÈME DE CAPTATION OPTIQUE

Plusieurs contraintes ont du être appréhendées dans le choix du système d'enregistrement. Des contraintes intrinsèques à la cithare *marovany* d'abord, incluant la rapidité de jeu, les différents modes d'attaque et d'étouffement des cordes, les segments occasionnels de polyphonie (accord) et les résonances par sympathie des cordes. L'accumulation de ces contraintes génèrent des signaux audio complexes à analyser. S'ajoutent à cela des contraintes externes, comme les autres sources acoustiques (principalement le hochet, les battements de l'assistance, les interjections vocales du possédé), les conditions environnementales (fortes humidité et chaleur) et techniques (batteries électriques rares et fragiles). Il est préférable d'éliminer les systèmes trop sensibles et pré-amplifiés gourmands en énergie (comme l'usage d'une alimentation fantôme 48V) qui pourraient se dégrader plus vite à cause de l'humidité et de la chaleur. L'ensemble de ces contraintes a été pris en compte dans le développement du système par captation optique présenté ci-dessous.

Les systèmes de captation optique ont déjà trouvé des applications dans le domaine de l'audio, comme la réalisation de mesures métrologiques de déplacements de cordes [8, 9] ou la MIDIfication<sup>1</sup> d'un piano à travers la technologie piano-bar développée par Moog [10, 11]. Notre système de captation, illustré à la figure 1, se rapproche de cette seconde application, bien que s'en démarquant aussi par la volonté d'intégrer aussi fidèlement que possible un grand nombre de paramètres physiques caractérisant le rendu acoustique de l'instrument. Le capteur optique choisi se présente sous la forme d'une fourche, illustré dans la zone agrandie de la figure 1. D'un côté, une DEL (Diode ElectroLuminescente) émet un faisceau lumineux de 0.5 mm de diamètre dans l'infrarouge. Son pic d'émission se situe autour de la longueur d'onde 940 nm. De l'autre, un phototransistor adapte son courant de sortie en fonction du flux lumineux qu'il reçoit. Son pic de sensibilité est à la longueur d'onde 850 nm. Ainsi, lorsque la corde, placée dans la fourche optique, passe au travers du rayon lumineux, elle bloque une partie du flux lumineux et module donc le courant de sortie du capteur. Pour obtenir une surface sensible supérieure au seul diamètre du

<sup>1</sup> . Néologisme signifiant la conversion in situ d'un instrument acoustique en son homologue MIDI.

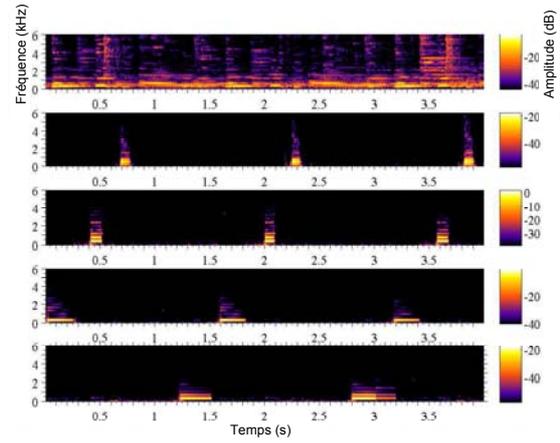
faisceau lumineux (0.5 mm), deux fourches optiques sont collées l'une à l'autre avec un léger décalage pour donner une surface couverte de 3.5 mm, puis mises en phases par des circuits de conditionnement électroniques. Le boîtier d'alimentation des fourches ne nécessite qu'une tension continue de 5V, et est isolé thermiquement, ce qui le rend bien adapté pour les conditions de terrain. Deux enregistreurs numériques portables ZOOM R16, permettant l'enregistrement de 2 x 8 pistes simultanément, servent enfin à l'acquisition des signaux optiques. Ceux-ci sont acquis par le logiciel Cubase (asservi par un ordinateur portable emmené sur le terrain). Un signal audio de référence est aussi enregistré simultanément avec un microphone Neumann KM 184 mt. La fréquence d'échantillonnage de tous les enregistrements est de 44.1 kHz, avec 16 bits pour la numérisation.

Chaque corde est ainsi équipée d'un tel capteur optique, qui sont tous fixés sur une barre commune, verticale et extérieure à l'instrument (voir photo 1). Deux contraintes ont été prises en compte pour le placement de ces capteurs. Il a fallu d'une part éviter que le système soit trop intrusif et altère la jouabilité de l'instrument<sup>2</sup>. Et d'autre part, le point de mesure de déplacement de la corde peut biaiser la mesure de son amplitude, puisque ce déplacement se compose d'une superposition de modes vibratoires, faits de ventres et de noeuds, le fait de placer un capteur en un noeud modal revient à supprimer la contribution énergétique de ce mode au déplacement global. Pour répondre à ces deux contraintes, la barre est installée au plus près du chevalet, pour qu'ainsi la zone de jeu soit moins dérangée, et que les capteurs soient placés approximativement sur la pente ascendante du ventre suivant directement le noeud du chevalet, qui est commun à tous les modes.

La figure 2 représente les spectrogrammes du signal audio et des signaux optiques (après pré-traitement, voir section 3) enregistrés sur un air *Sojerina*. Les signaux optiques offrent un fort rapport signal sur bruit, ainsi que des transitoires d'attaque raides. L'indépendance de chaque corde est bien respectée, chaque capteur détectant uniquement la vibration de la corde sur lequel il est placé. On peut aussi remarquer la bonne séparabilité des notes successives sur une corde, les blancs inter-notes correspondant aux instants de contact entre le doigt et la corde. En définitive un tel système de captation revient à décomposer un signal audio multi-sources complexe en ses composantes simples, ramenant plus particulièrement le traitement d'une séquence polyphonique en celui de plusieurs séquences monophoniques conjointes.

Nous présentons maintenant une courte étude comparative des enveloppes temporelles entre les signaux optiques et audio. Après avoir pré-traité les signaux (voir section 3), cinq descripteurs acoustiques classiques ont été calculés sur un ensemble de 30 paires de notes {audio;optique}, jouées isolément et laissées en résonance

2. Une première appréciation positive a déjà été donnée par un musicien malgache sur ce sujet [12], mais une prochaine version du système visera notamment une meilleure compacité.



**Figure 2.** Exemple de spectrogrammes du signal audio (graphe du haut) et de sa décomposition en quatre signaux optiques (graphes dessous), extrait d'un air *Midégana*

libre. Les paires ont été sélectionnées aléatoirement pour leur diversité. Les descripteurs traités sont définis comme suit :

**AT** , le temps d'attaque (en s) est défini par le temps nécessaire pour le signal d'atteindre 95 % de son énergie maximale  $E_{max}$

$$s(n = AT) = 0.95E_{max} \quad (1)$$

**D** , la durée physique du signal (en s) sera définie par le temps durant lequel l'énergie du signal est compris entre 5 % et 95 % de son énergie maximale  $E_{max}$

$$D = \{n/s(n) > 0.05E_{max} \ \& \ s(n) < 0.95E_{max}\} \quad (2)$$

**E** , le niveau énergétique rms (en Pa)d'un signal défini par

$$E(k) = \sqrt{\frac{1}{K} \sum_{k=1}^K |(x + kN)|^2} \quad (3)$$

calculé pour K fenêtres de N échantillons ;

**DH** , l'indicateur d'harmonicit  (adimensionn ) renseigne sur le contenu harmonique d'un signal. Il consiste   balayer la densit  spectrale X d'un signal avec un filtre en peigne Filt, dont la fr quence fondamentale varie dans une plage donn e (ici [140 ; 400] Hz , ce qui est grossi rement la tessiture de la cithare) [13]. Lorsque les vall es de ce filtre coïncident avec les pics du filtres, leur produit est minimal. Math matiquement, on a alors

$$DH = \min\left(\frac{E_{pond}}{E_{init}}\right) \quad (4)$$

avec  $E_{init} = \sum |X(k)|^2$  and  $E_{pond} = Filt(k, k_0) E_{init}$ , avec Filt d fini par  $Filt = 2(1 - |\cos(\frac{\pi F}{F_0})|)$ .

Descripteurs	At	AD	dB	DH
$E(\Delta_D)$	0.0129	0.4366	0.23	0.26
$\sigma(\Delta_D)$	0.0028	0.11	0.15	0.09

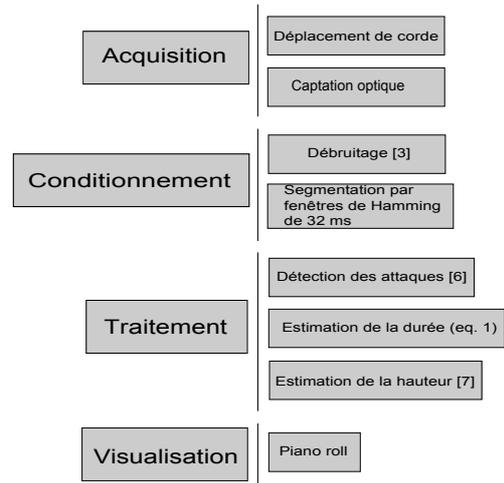
**Table 1.** Moyenne E et écart-type  $\sigma$  des différences acoustiques absolues entre des signaux de types optiques et audio pour les descripteurs acoustiques At, AD, dB, DH

Les descripteurs E et DH sont évalués relativement à une référence arbitraire respective au type optique ou audio. Les différences acoustiques absolues entre ces deux types de signaux sont simplement quantifiées avec l'opérateur  $\Delta_D = |D_{Audio} - D_{Optique}|$ , où D représente un descripteur acoustique donné. Le tableau 1 présente les résultats de cet opérateur, montrant que la distorsion d'information est minimale si on se restreint aux informations sur l'enveloppe temporelle. Par contre, le contenu spectral montre des différences plus significatives. Une tendance intéressante étant notamment que le contenu harmonique des signaux optiques est plus fort, ce qui peut s'expliquer par le fait qu'une mesure directe de déplacement de corde privilégie sa fréquence fondamentale et ses harmoniques dans le comportement vibratoire observé, en minimisant l'effet de couplage aux modes plus complexes de la table d'harmonie. La différence sur l'amplitude peut être significative si la vibration de la corde excite fortement certains modes de la table, générant des sons plus entretenus à travers un transfert d'énergie optimal des cordes vers la table.

### 3. APPLICATION À L'EXTRACTION AUTOMATIQUE D'INFORMATIONS MUSICALES

Généralement, la recherche d'information pour la transcription musicale peut se décomposer en plusieurs niveaux : les informations de bas niveau (la hauteur, l'attaque et la durée des notes), et les informations de haut niveau qui font appel à des notions plus globales de la musique (tempo, chiffrage de mesure, tonalité, reconnaissance d'instrument). Notre système de transcription comprend un dispositif d'acquisition de données (décrit ci-dessus) pour enregistrer la cithare et une partie logicielle intégrant un algorithme d'analyse qui déterminera la durée et la hauteur des notes jouées, qui sont les deux informations de bas niveau essentielles permettant d'écrire une partition, auxquelles on ajoutera leurs amplitudes. Ces informations seront ensuite regroupées pour ne faire qu'un fichier au format MIDI. Ce fichier pourra ainsi être lu et édité sur n'importe quel séquenceur audio ou logiciel de partition.

Une fois les signaux optiques correctement acquis, leur transcription ne présente pas de difficulté particulière. La figure 3 représente un diagramme-bloc des différentes fonctions constituant notre chaîne de traitement, de l'acquisition des signaux optiques à l'algorithme de détection et de caractérisation des notes de la *marovany*. Le pré-traitement des signaux consiste en les étapes suivantes. Pour des soucis de mémoire informatique, des séquences de 5s sont d'abord importées dans le logiciel Matlab. Un



**Figure 3.** Diagramme-bloc des fonctions de l'algorithme de détection et de caractérisation acoustique des notes

filtrage adaptatif [14] est appliqué pour optimiser le rapport signal sur bruit, celui-ci étant notamment détérioré par des bruits parasites d'origine électronique ou de résonance par sympathie des cordes<sup>3</sup>. En rentrant des segments de signaux dont les propriétés spectrales sont représentatives de ces bruits, l'algorithme [14] permet de les soustraire au signal utile en réduisant le résidu par moindres carrés. Une fenêtre de Hamming glissante de 512 échantillons (soit 11.6 ms, donnant une résolution temporelle de l'ordre de grandeur des temps d'attaque) avec une superposition de 256 échantillons balaye ensuite l'ensemble de la séquence. Les fronts d'attaque de chaque note sont détectés en utilisant une différence spectrale, avec prise en compte de l'incrément de phase, telle qu'introduit par Bello et al. [15] :

$$\hat{X}_{k,n} = |X_{k,n-1}| e^{j(2\phi_{k,n-1} - \phi_{k,n-2})} \quad (5)$$

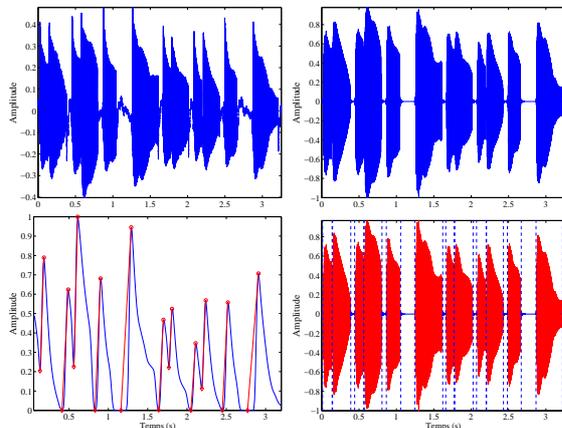
avec n l'indice de chaque fenêtre. Notre signal utile étant composé de sinusoïdes stationnaires, l'apparition d'un front d'attaque génère un pic dans l'erreur de prédiction définie par

$$r(n) = \sum_{n=1}^N |\hat{X}_{k,n} - X_{k,n}| \quad (6)$$

Un simple seuil est ensuite associé à ce résidu pour la validation d'une détection de notes ou non. A partir de cet onset détecté, le descripteur E (eq. 3) est calculé pour les fenêtres avoisinantes de façon à chercher le maximum local  $E_{max}(i)$  associé à la note i, en posant l'hypothèse que ce maximum est localisé près du front d'attaque, ce qui est vraisemblable pour les notes des instruments à cordes pincées. Ensuite, E est calculé sur l'ensemble des fenêtres qui suivent l'onset jusqu'à ce que sa valeur décroisse en-dessous de 5 % de  $E_{max}(i)$ , utilisée alors comme un seuil

<sup>3</sup>. La résonance par sympathie des cordes est en effet considérée dans notre situation de transcription bas niveau comme du bruit, bien que celui-ci prenne une place importante dans la constitution du timbre de l'instrument.

adaptatif d'énergie physique spécifique à la note  $i$ . Cette estimation nous permet de déduire la durée du signal (éq. 2) et l'amplitude par moyennage de l'énergie sur l'ensemble des fenêtres du signal. Une fois les notes détectées, la hauteur du signal à l'intérieur de cette fenêtre est enfin estimée par un algorithme robuste dérivé de la fonction d'auto-corrélation [16]. La figure 4 présente l'évolution de formes temporelles du signal à travers les différentes étapes de traitement.

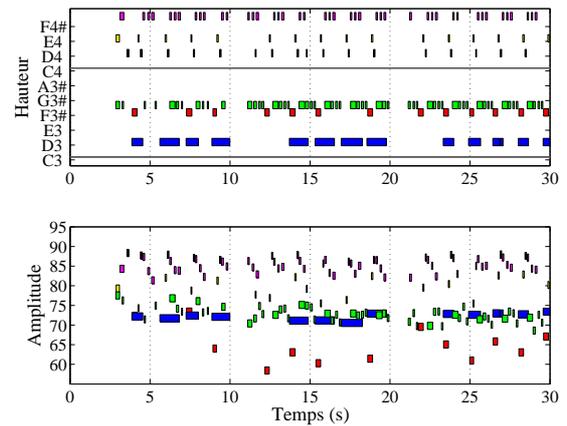


**Figure 4.** Evolution de formes temporelles du signal à travers les différentes étapes de traitement, avec de haut en bas et de gauche à droite : signal optique original, signal débruité, résidu  $r$  avec localisation des attaques, signal segmenté

Cet algorithme fut évalué sur des séquences préalablement étiquetées à la main, en reportant en plus de leur emplacement temporel absolu, leur durée, leur amplitude moyenne et leur hauteur. Une application brute de l'algorithme précédent aux séquences de type audio aboutit à des performances systématiquement inférieures à 55 % de détection de notes, alors que les séquences optiques donnent des résultats plus que satisfaisants (la tolérance fixée dans la détection des fronts d'attaque et de la durée étant respectivement de 32 ms et 0.5 s).

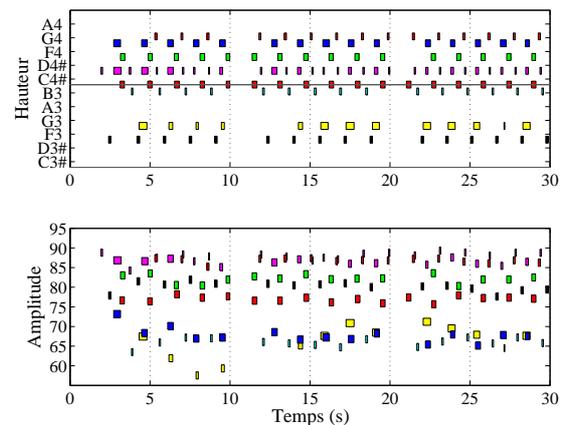
Nous proposons maintenant deux illustrations de cette méthode à travers les transcriptions automatiques de deux airs traditionnels joués à la cithare *marovany*, *Midegana* et *Sojerina* (les enregistrements sonores sont disponibles sur la page web [7]). Ces airs ont été enregistrés à Madagascar et joués par le musicien Velonjoro. Des problèmes techniques sur place ont rendu inutilisables plusieurs capteurs<sup>4</sup>, et chaque variante des airs a été partiellement reconstituée, avec plusieurs prises successives permettant d'enregistrer plusieurs séries de cordes. Malgré un recalage manuel et la superposition de prises distinctes, les signaux ont été jugés satisfaisants à l'oreille, la vitalité rythmique du jeu de Velonjoro étant bien préservée. Les figures 5 et 6 montrent les piano rolls de variantes de ces

4. Cette mission durant laquelle ces enregistrements ont été réalisés avait pour but de tester une version prototypique de notre système de captation [12]. La prochaine mission prévue pour Juin 2013 bénéficiera d'une version finalisée et entièrement opérationnelle de ce dispositif.



**Figure 5.** Piano roll résultant d'une transcription automatique d'un air *Midegana*

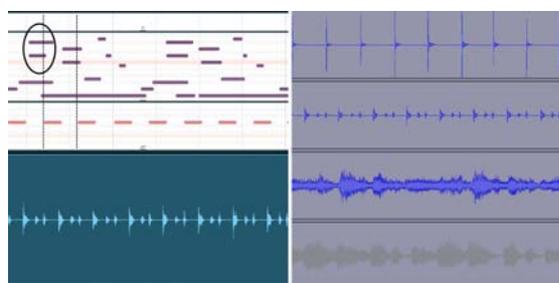
airs. Le piano roll est un moyen de représenter graphiquement un fichier MIDI. En ordonnée, il y a les différentes notes, représentées par des rectangles soit à travers leurs hauteurs (graphe du haut) ou leurs vélocités (graphe du bas) respectives, et en abscisse l'axe du temps. Chaque rectangle tracé sur le piano roll correspond à une note. Il est alors facile de visualiser les notes jouées au fil du temps.



**Figure 6.** Piano roll résultant d'une transcription automatique de *Sojerina*

La précision du système de transcription proposé est bien adaptée à la rapidité de jeu du musicien et rend bien compte de certaines subtilités des structures rythmiques caractéristiques du jeu de la *marovany*, comme on va maintenant l'expliquer brièvement. Le graphe gauche de la figure 7 superpose en fonction du temps la détection MIDI effectuée sur l'air *Sojerina* avec deux pistes correspondant à l'assise rythmique de cet air, les battements de main (ici représentés en clap MIDI) et la piste audio du hochet *kantsa*. Le graphe droit de la figure 7 montre la superposition des signaux audio de battements de main, de hochet, de l'original et de la transcription automatique MIDI (en grisé). En comparant l'audio de la cithare et celui de la dé-

tection MIDI, on remarque que les deux pistes sont bien synchronisées, bien qu'il manque des notes dans la partie MIDI, correspondant à des cordes manquantes (problème des capteurs défaillants présenté plus haut). On retrouve bien le caractère contramétrique de cette musique [17], la pulsation étant donnée par les battements de mains, on voit que l'accent du hochet est décalé par rapport à ce battement, en levée donc, c'est-à-dire à contre-temps. Plus précisément, il tombe sur la 2ème croche d'une subdivision ternaire de la pulsation (chaque pulsation se divise en trois croches). Dans la détection MIDI, il y a deux tierces qu'on remarque bien : RE-FA# et DO-MI, puis un arpège descendant SOL-RE-SI. Or ces deux tierces et le SOL sont plutôt placés sur l'accent du hochet, et non sur les frappements de mains, donc la cithare joue à contre-temps. Cette caractéristique est confirmée par la piste audio où l'on voit aussi bien dans l'original que dans la détection une intensité plus forte au moment des deux tierces, qui correspond aux accents du hochet.



**Figure 7.** A gauche, de haut en bas, superposition en fonction du temps la détection MIDI effectuée sur l'air *Sojerina* avec deux pistes correspondant à l'assise rythmique de cet air, les battements de main (ici représentés en clap MIDI) et la piste audio du hochet *kantsa*. Les deux barres verticales indiquent deux temps donnés par le clap MIDI, et la partie encerclée illustre la contramétrie du jeu, avec un placement entre à contre-temps. A droite, de haut en bas, la superposition des signaux audio de battements de main, de hochet, de l'original et de la transcription automatique MIDI (en grisé).

#### 4. CONCLUSION

Un système d'enregistrement par captation optique a été introduit dans ce travail et appliqué à la transcription automatique de la cithare *marovany*, instrument utilisé dans les séances de transe *tromba*. Parmi ces avantages technologiques, nous retiendrons l'acquisition d'un signal indépendant par corde, fort rapport signal sur bruit, forte sensibilité sur l'amplitude du mouvement. Il a été conçu à la fois pour répondre à des exigences immédiates d'études de terrain et pour assurer une méthode de transcription automatique robuste sur des projets de long terme tel qu'une classification systématique des airs de *marovany* en relation avec la transe.

La nécessité de retravailler sur des données audio se fera peut-être sentir pour récupérer des propriétés acous-

tiques plus complexes, incluant les propriétés vibratoires et acoustiques de l'instrument entier, comme le timbre de l'instrument, les modes d'ébranlement des cordes, l'intensité acoustique. Cependant, le système optique fournit un support fiable complémentaire pour les développements de la transcription à suivre, permettant la connaissance d'informations a priori susceptibles de simplifier des problèmes difficiles liés au travail direct sur l'audio, tels que le nombre de cordes vibrantes pour étudier les segments de polyphonie dans les morceaux.

Dans une perspective plus large, les applications de ce dispositif sont de deux ordres : anthropologique et artistique. Sur le plan anthropologique, les paramètres musicaux extraits de la classification du répertoire de la *marovany* pourraient conduire *in fine* à dégager une « théorie musicale » en la confrontant aux modèles heuristiques proposés par les musiciens autochtones, nous renseignant ainsi sur le rôle de la musique dans les mécanismes de déclenchement de la transe et sur le savoir-faire développé par les musiciens à cet effet.

A côté de cette application anthropologique, une autre perspective d'utilisation d'un tel système apparaît dans l'interaction musicale homme-machine. En effet, l'environnement informatique d'improvisation *ImproteK* [18] (développé à partir de l'environnement *OMax* [19] en partenariat avec l'IRCAM) permet de capter le jeu d'un musicien et d'improviser à partir des phrases enregistrées en ré-orchestrant les données MIDI avec les timbres générés par un échantillonneur. Des projets pourraient être mis en place pour que des musiciens traditionnels malgaches utilisent un tel système avec les données MIDI extraites par nos capteurs optiques. Des questions d'ordre plus esthétique (acceptabilité de formules musicales « dérivées » d'un catalogue connu, témoignage oral de cette capacité, intérêt musical dans l'amplification, la modification temps réel de paramètres musicaux, la ré-orchestration virtuelle d'un environnement musical) seront à considérer dans de futures recherches suivant cet axe.

#### 5. REMERCIEMENTS

Ce travail a été réalisé avec le soutien de l'ANR, au sein du Projet IMPROTECH ANR-09-SSOC-068. Il est issu en grande partie de deux projets de Master réalisés au LAM : *Transcription automatique de la cithare malgache* par Joachim Flocon-Cholet et *Etude de la cithare malgache en vue de la transcription automatique* par Mathilde Paul. Nous souhaitons aussi remercier Laurent Quartier pour la qualité de son apport technique à ce travail, ainsi que Velonjoro pour avoir participé à l'enregistrement de plusieurs variantes de *Midegana* et de *Sojerina*. La toolbox Matlab MIDItoolbox a été utilisée pour réaliser les pianos rolls présentés à la section 3.

#### 6. REFERENCES

- [1] Rouget, G. *La Musique et la transe. Esquisse d'une théorie générale des relations de la musique et de la*

- possession*, Paris, Gallimard, 1980.
- [2] Chemillier M. « Esthétique et rationalité dans les musiques de tradition orale », *Rapport de recherche au Ministère de la Culture*, 1998-2000.
- [3] Dianteill, E. et Hell, B. « Le possédé spectaculaire. Possession, théâtre et globalisation », *Gradhiva*, n°7, p 4-5, 2008.
- [4] Hell, B. « Négocier avec les esprits *tromba* à Mayotte. Retour sur le « théâtre vécu » de la possession », *Gradhiva*, n°7, p. 6-23, 2008.
- [5] Razafindrakoto, J. « Le timbre dans le répertoire de la valiha, cithare tubulaire de Madagascar », *Cahiers d'ethnomusicologie*, Vol. 12, p. 123-142, 1999
- [6] M. Domenichini, « Valiha », in « Stanley Sadie », ed. : *New Grove Dictionary of Musical Instruments*. London :Macmillan, Vol. 3, p. 705-706, 1984.
- [7] Chemillier, M. « Page web consacrée à la cithare *marovany* », <http://ehess.modelisationsavoirs.fr/marovany/index.html>, 2012-2013.
- [8] Seydoux, L. « Mesure de déplacement de cordes avec des fourches optiques », Mémoire de fin d'études, LAM (Laboratoire Acoustique Musicale), 2012.
- [9] Chabassier, J. « Modélisation et simulation numérique d'un piano par modèles physiques », Thèse de l'Ecole Polytechnique, 2012.
- [10] Mowat, W. « Bob Moog Piano Bar : MIDI Output device for acoustic pianos », *Sound On Sound*, mis en ligne sur [www.soundonsound.com/sos/mar05/...../articles/moogpianobar.html](http://www.soundonsound.com/sos/mar05/...../articles/moogpianobar.html), 2005.
- [11] Assayag G. and Bloch, G. *OMax. The Software Improviser*, Documentation version 2, IRCAM, May 2008.
- [12] Chemillier, M. « Compte-rendu de mission sur les capteurs optiques - Madagascar 29 juillet au 13 août 2012, Projet ANR IMPROTECH », mis en ligne sur <http://ehess.modelisationsavoirs.fr/marovany/capteurs/>, 2012.
- [13] Youngmoo, E. K. and Whitman, B. « Singer identification in popular music recordings using voice coding features », *ISMIR*, 2002.
- [14] Ephraim, Y. and Malah, D. « Speech enhancement using a minimum-mean square error log-spectral amplitude estimator », *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 33, p. 443-445, 1985.
- [15] Bello, J. P., Duxbury, C., Davies, M. and Sanders, M. B. « On the use of phase and energy for musical onset detection in the complex domain », *IEEE Signal Processing Letters*, Vol. 11, p. 553-556, 2004.
- [16] de Cheveigné, A. and Kawahara, H. « YIN, a fundamental frequency estimator for speech and music », *J. Acoust. Soc. Am.*, Vol. 111, p. 1917-1930, 2002.
- [17] Chemillier, M., Pouchelon, J., André, J. and Nika, J. « Le jazz, l'Afrique et la contramétricit  », à paraître, *Anthropologie et société*, 2013.
- [18] Nika, J. and Chemillier, M. « ImproteK : intégrer des contrôles harmoniques pour l'improvisation musicale dans la filiation d'OMax », *Actes des Journées d'Informatique Musicale*, p. 147-155, 2012.
- [19] Assayag, G., Bloch, G., Chemillier, M., Cont, A. and Dubnov, S. « The OMax Project Page », available at <http://omax.ircam.fr>.



# CRÉATION MUSICALE ASSISTÉE PAR ORDINATEUR EN MILIEU SCOLAIRE

## Penser des cadres et des outils pour une approche pratique par l'écoute et la manipulation

*Philippe Galleron*

CICM-EA 1572

Esthétique, musicologie et créations musicales  
Université Paris 8

[philippe.galleron@gmail.com](mailto:philippe.galleron@gmail.com)

*Alain Bonardi*

CICM-EA 1572

Esthétique, musicologie et créations musicales  
Université Paris 8

IRCAM

[alain.bonardi@ircam.fr](mailto:alain.bonardi@ircam.fr)

### RÉSUMÉ

Dans cet article nous présentons nos démarches entreprises en tant que chercheur et professeur pour développer et encourager la pratique de la musique assistée par ordinateur en milieu scolaire. Nous y exposons les expérimentations de dispositifs techniques et pédagogiques que nous avons créés ou auxquels nous avons participé. Pour chacune des expériences nous décrivons le contexte pédagogique, la conception technique et nous procédons à une évaluation.

Il ressort de ces expériences que le travail collaboratif de tous les acteurs de ces projets est fructueux et qu'il est nécessaire de le développer. En cela nous pensons que les technologies de l'information et de la communication ont un rôle important à jouer au niveau de la coparticipation à ces projets pour leur valorisation.

Convaincu que l'apprentissage se fait par la participation nous nous appuyons sur les concepts liés aux apprentissages informels et à la notion de « communauté de pratique » définie par Jean Lave et Étienne Wenger. Dès lors, nous proposons d'engager les élèves ainsi que les adultes participant aux projets dans une communauté rassemblée autour de la pratique musicale assistée par ordinateur.

### 1. INTRODUCTION

Il nous semble important de faire pratiquer la musique assistée par ordinateur (M.A.O.) en milieu scolaire. En effet du point de vue pédagogique il est intéressant de profiter de l'attrait que représentent pour les enfants les sons synthétiques et leur manipulation.

Du point de vue de l'apport culturel la pratique de la musique assistée par ordinateur les aidera à se construire des repères parmi les musiques auxquelles ils ont accès par les médias et qui font aujourd'hui fortement usage des outils informatiques. Du point de vue musical, la pratique permet aux élèves de découvrir les impacts majeurs que l'informatique a eu et continue d'avoir sur la création musicale contemporaine, par exemple sur le plan des sonorités, des façons de composer ou d'interpréter la musique.

C'est pourquoi la question de son développement en milieu scolaire représente à nos yeux un enjeu important pour l'enfant.

En France, dès les années 1970, les dispositifs technologiques tel que l'Upic du Centre d'Études de Mathématiques et Automatique Musicales, le Mélisson du Groupe de Musique Électroacoustique d'Albi et la SARL ARP Industrie (Art-Recherche-Pédagogie) et le Gmebogosse du Groupe de Musique Expérimentale de Bourges (G.M.E.B.) furent des expériences pionnières de musique assistée par ordinateur en milieu scolaire.

Actuellement il existe plusieurs expériences prenant des formes variées qui s'adressent à un public scolaire. Il peut être question d'agir en direct sur le son, de participer à des activités de composition, ou encore des ateliers de sensibilisation à l'écoute des principes des musiques électroacoustiques.

Les dispositifs technologiques tel que l'arc laser du BAO-PAO de Jean Haury et Jean Schmutz ou les joysticks de la « Méta-Mallette » du laboratoire PuceMuse s'inscrivent parmi les exemples de sensibilisation à la musique assistée par ordinateur se faisant à l'occasion de résidences artistiques ou de

l'accueil d'un dispositif technologique dans les conservatoires.

Parmi elles nous citerons par exemple le projet initié par Laetitia Pauget (musicienne en milieu scolaire), Alain Gerber (enseignant à l'école primaire) et Alain Basso (compositeur), qui existe depuis plus de dix ans dans une école primaire de Chambéry-le-Vieux. Il s'agit d'un projet de composition électroacoustique sur un logiciel de montage audio.

Nous pouvons évoquer également les démarches pédagogiques à destination des publics scolaires de Bernard Fort au Groupe Musiques Vivantes de Lyon qui est un centre de ressources pour les musiques électroacoustiques [5].

Cependant nous constatons que la musique assistée par ordinateur (M.A.O.) n'est pas une pratique courante et répandue à l'école.

Au cours de nos recherches, en tant que professeur de conservatoire intervenant en milieu scolaire et jeune chercheur nous avons dégagé trois démarches-type : (1) une écoute analytique et dirigée en vue de l'acquisition de ces principes (la spatialisation, la synthèse et le traitement du son) ; (2) une démarche de création et d'invention par proposition des élèves dans ces trois domaines mais dont la réalisation est faite par le professeur ; (3) une approche par la manipulation directe par l'élève des sonorités et des moyens de transformations qu'offre l'informatique musicale [11].

Comme nous défendons l'idée que l'apprentissage de la musique en milieu scolaire se fait par l'écoute et la pratique, il est donc indispensable que les élèves puissent manipuler les sonorités de l'univers musical qu'ils découvrent. Pour cela, il faut mettre à leur disposition les outils adéquats.

Pour être à la hauteur de ces enjeux, nous présenterons nos expériences personnelles et nos propositions dans ce domaine. Les trois expériences dont nous rendons compte ici furent réalisées en milieu scolaire. Elles répondent à une demande de projets artistiques, culturels de la part des enseignants de l'Éducation Nationale. Dans ce cadre les dispositifs doivent répondre à des contraintes très précises définies par les programmes pédagogiques, les ressources matérielles dont dispose l'école et l'implication des différents acteurs participant aux projets. Nous présenterons chacun des dispositifs en fonction de leur contexte et des enjeux qu'ils représentent. Ensuite nous ferons une brève description de la conception et de la réalisation technique du dispositif. Enfin nous ferons une auto-évaluation technique et pédagogique de la démarche entreprise.

## 1. EXPERIMENTATION DU TRACKING VIDÉO A L'ÉCOLE PRIMAIRE

### 1.1. Contexte et enjeux

Le projet pédagogique musical a été élaboré par quatre enseignants de CP d'une école élémentaire de Noisy-le-Sec. L'un des objectifs était de permettre aux élèves d'explorer leur ville par un biais artistique. Après une promenade dans la ville, les enfants ont fait des compte-rendus de ce qu'ils avaient vu. Leur maîtresse a attiré leur attention sur les couleurs qui les entouraient. De retour en classe, ils ont décrit les choses qu'ils avaient vues en nommant leur couleur. Ils ont ensuite écrit une poésie dont voici un extrait :

« C'est jaune comme les boîtes aux lettres de la poste  
C'est bleu comme l'eau de la piscine

C'est rouge comme les camions des pompiers

C'est vert comme la pelouse du Stade Huvier » (CP de l'école Carnot 2011).

Nos objectifs pédagogiques sur ce projet étaient de : (1) mettre les élèves en capacité de manipuler des sons issus des univers musicaux de l'électroacoustique ; (2) contribuer à mettre en relation corps et production sonore ; (3) établir des connexions entre diverses activités d'expression de soi développées à l'école, telles que la poésie, l'expression corporelle et les productions écrites.

Notre objectif d'un point de vue technique était de trouver un moyen pour que les enfants puissent déclencher eux-mêmes l'écoute de ces enregistrements lors d'un spectacle vivant. Comme ils travaillaient sur la reconnaissance des couleurs, nous avons décidé que l'objet qui déclencherait la phrase serait de la couleur de la phrase à interpréter. Concrètement si l'enfant présente un pavé bleu devant la caméra, il entendra la phrase : « C'est bleu comme l'eau de la piscine ». Quand il cache l'objet de couleur le son s'arrête et quand il réapparaît, la bande-son recommence depuis le début de la phrase. Pour cela nous avons conçu le dispositif « tracking couleur » (Figure 1).

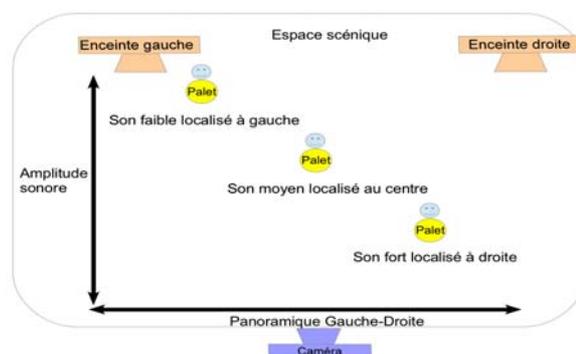


Figure 1. Dispositif « tracking couleur »

Nous souhaitons aussi que les enfants puissent appliquer un traitement sur ces enregistrements.

Nous avons donc décidé que l'élève pourrait faire varier l'amplitude suivant l'endroit où il se trouve sur scène. Plus il s'approche de l'avant-scène et plus le son devient fort. Plus il recule dans le fond de la scène et moins le son est fort. S'il se dirige côté « Jardin », le son s'amplifie par un effet panoramique sur l'enceinte de gauche. Inversement le son augmente à droite s'il se dirige côté « Cour ».

## 1.2. Conception et réalisation du dispositif

Nous avons commencé par enregistrer cette poésie phrase par phrase pour la conserver comme une trace sonore de notre exploration.

Nous avons choisi comme moyen l'analyse vidéo. L'avantage de ce dispositif est qu'il nécessite relativement peu de moyens matériels et que les caméras *Digital Video* sont assez répandues dans les écoles. D'autre part, l'utilisation de ressources de l'école comme les palets de couleur présente l'avantage de proposer aux enfants des objets connus d'eux, et dont la manipulation leur est acquise. Les contraintes du lieu semblent *a priori* assez faciles à reproduire pour des moments ponctuels tels que la restitution et pour les séances de travail spécifiques.

Nous avons utilisé comme environnement informatique un ordinateur portable *Macintosh* avec le logiciel de programmation visuelle *Max/Msp 5 de Cycling 74*.

La fiabilité du dispositif étant fortement liée aux variations de la lumière, nous avons éteint la lumière. L'obscurité a pour avantage de rendre invisibles les couleurs indésirables et supprime les problèmes liés aux ombres. Pour que la couleur soit visible, nous avons fixé une lumière directement sur l'objet. De cette manière, l'exposition de l'objet est toujours constante et de ce fait, sa couleur aussi.

## 1.3 Évaluation technique et pédagogique

La méthode est assez simple techniquement et demande du matériel assez répandu dans les écoles. Nous avons donc pu affiner les réglages de l'objet *jit.fitbounds*. Les couleurs sont bien reconnues et différenciées à l'exception du jaune qui n'est pas détecté systématiquement. C'est bien la limite principale de ce dispositif : il manque de fiabilité. Pour preuve, dans nos tous premiers essais, le jaune était la couleur la mieux identifiée. Il est donc difficile de prévoir le résultat tant qu'on n'a pas fait des tests dans les conditions réelles d'éclairage de la salle.

Ce manque de fiabilité pose plusieurs problèmes dans le cadre d'une séance pédagogique. En effet, l'objectif de renforcer l'apprentissage des couleurs n'est pas atteint pleinement, puisqu'une couleur n'est pas bien détectée. Cela a également pour conséquence de frustrer l'enfant qui a enregistré la phrase « C'est jaune comme ... », puisqu'il n'entendra pas le fruit de sa participation aux

séances d'enregistrement. Enfin, la condition d'avoir un dispositif rapide à installer (qui nous semble essentielle pour mener une activité régulière) n'est pas remplie. Il faut donc un espace dédié à cette activité pour que les réglages soient effectués patiemment et que les conditions soient toujours les mêmes : salle de spectacle ou salle de classe. Ce dispositif ne permet donc que des usages ponctuels.

Une autre limite propre à la captation vidéo reste le temps de latence pour déclencher un son. Cependant, au regard de l'utilisation musicale et esthétique que nous souhaitons en faire, cette réaction lente du dispositif ne pose pas de problème pour atteindre les objectifs de la séance. En effet, l'activité d'écoute et de jeu portera plus sur la contemplation sonore et la féerie et moins sur l'interprétation nécessitant une certaine virtuosité dans la vitesse d'exécution. De fait, les points forts du dispositif sont qu'il invite à l'écoute et à l'exploration attentive du phénomène sonore. En effet, le fait que l'élève ne soit relié à aucun fil et que le déclenchement du son soit opéré par un de ses jouets confère à l'activité un aspect magique. Ceci suscite une certaine curiosité de l'enfant qui va chercher à manipuler son jouet dans tous les coins de l'espace scénique en vue d'une réaction sonore inédite ou qu'il a anticipée. De plus, le fait que l'activité ait lieu dans la pénombre crée une atmosphère inhabituelle qui place l'enfant dans une situation d'éveil.

Les techniques de *tracking* vidéo comme la reconnaissance infrarouge et la reconnaissance de gestes devraient pouvoir compléter nos recherches pour optimiser ce dispositif.

## 2. CONTRÔLEUR MÉTA-MALETTE EN COLLÈGE

### 2.1. Contexte et enjeux

Le projet artistique a été mené par Jean-Philippe Dejussieu, directeur du conservatoire de Noisy-le-Sec qui, en tant que compositeur, a re-créé en lien avec le compositeur Gilbert Artman, deux pièces du répertoire du groupe *Urban Sax* rebaptisées *More Slowly Please* part 1-2 pour l'orchestre « Jeune Philharmonie de Seine Saint-Denis » et dans le cadre de concerts à l'église de Noisy-le-Sec.

Ce travail a été engagé avec les classes de Musique Assistée par Ordinateur et de formation musicale du conservatoire afin de constituer des ensembles utilisant le dispositif technologique « Méta-Mallette » du laboratoire PuceMuse. L'objectif était d'arriver à une interprétation *live* (en direct) de la pièce qui mélangeait l'orchestre symphonique et le dispositif d'informatique musicale.

Ce projet propose un parcours aux élèves du Collège et du conservatoire. Il comprend la découverte et l'accompagnement du parcours de création sonore (enregistrement, transformation, jeu en direct).

D'autre part, il propose la découverte de l'orchestre symphonique dans un répertoire des XX<sup>e</sup> et XXI<sup>e</sup> siècles, avec des œuvres des compositeurs américains Aaron

Copland et Léonard Bernstein, qui sont également au programme du concert final.

## 2.2. Conception et réalisation du dispositif

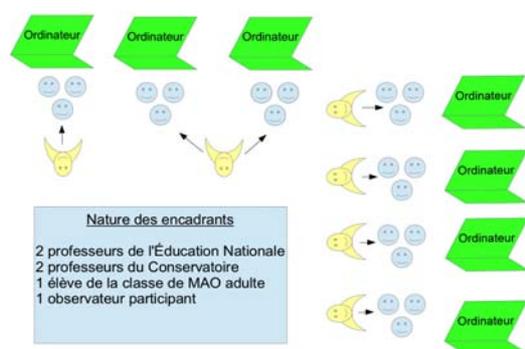
Le dispositif pédagogique proposait aux élèves un parcours composé d'ateliers de manipulations sur les joysticks de la « Méta-Mallette » et sur le logiciel *Audacity*. Il comprenait également des sorties culturelles sur le temps scolaire comme la visite des studios d'Urban Sax à Paris et la rencontre avec Gilbert Artman qui est à la fois le directeur de la compagnie et le compositeur des œuvres originales du projet. Le dispositif comprenait également des moments de manipulation des joysticks au conservatoire en dehors du temps scolaire. Enfin le projet a été ponctué de moments d'analyse critique avant et après la restitution devant un public.

Les familles des élèves ont également été associées en leur proposant de découvrir l'orchestre lors d'un concert et de suivre les répétitions avant le concert de la représentation finale.

Les activités de manipulation avaient pour objectifs, d'associer les élèves à l'interprétation de cette œuvre de musique mixte et d'initier chez eux une démarche de composition électroacoustique.

La pratique de la M.A.O. en milieu scolaire a donc été réalisée suivant deux axes : la préparation des sons et leur interprétation.

La préparation des sons a été réalisée sur les ordinateurs de la classe. Le logiciel *Audacity* a été installé à la demande de l'équipe pédagogique et avec l'accord du directeur d'établissement par le responsable informatique. Sur une douzaine de postes, sept ont été équipés. Le logiciel n'a pu en effet être installé que sur les ordinateurs possédant un système d'écoute. Les professeurs de la classe disposaient également d'un ordinateur équipé d'un vidéoprojecteur, mais qui n'a pas été exploité. La classe a été répartie par groupes de trois ou quatre élèves assis derrière les ordinateurs équipés. Les élèves ont manipulé à tour de rôle en utilisant le clavier et la souris. (Figure 2)



**Figure 2.** Séance sur *Audacity* en classe

Lors des activités sur *Audacity*, la consigne du professeur de M.A.O. était de préparer les

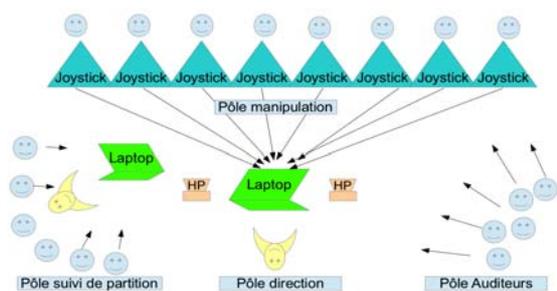
enregistrements déclenchés qui seraient ensuite joués par les joysticks. Cette tâche était guidée. Les élèves devaient faire un choix parmi les sons préparés par le professeur et déposés sur chaque ordinateur à partir d'une clef USB. Ces sons étaient composés de bruits divers : sonar, trompette, sons de chantiers... L'élève devait apprendre à ouvrir les fichiers sons dans un projet *Audacity*, puis écouter le son en intégralité puis choisir une portion de son de quatre secondes qui lui plaise. Une fois le son choisi, il lui restait à le couper et à l'isoler dans un nouveau projet *Audacity*. Il pouvait appliquer un traitement de son choix parmi les suivants : jouer avec le *pitch* (changement de la hauteur), inverser le sens de lecture ou ajouter un effet *reverb* (modifier l'impression d'espace). Une fois le résultat sonore jugé satisfaisant, il devait « nettoyer » le son destiné à la boucle, et appliquer deux fondus, en ouverture et en fermeture pour éviter les clics. Quand la boucle était prête, il fallait exporter le son au format audio, lui donner un nom et le déposer dans le dossier portant le numéro du poste de travail (numéro de l'ordinateur).

Une fois ce travail d'édition mené à son terme, la seconde phase fut une interprétation dirigée de ces sons édités. Les élèves opéraient des variations grâce aux capacités de contrôle qu'offraient les joysticks et le logiciel de la Méta-Mallette. La manipulation du logiciel Méta-Mallette était assurée par le professeur de M.A.O.

Le matériel nécessaire à la manipulation était fourni par le conservatoire, transporté dans la classe et installé à chaque intervention. Il était composé d'un ordinateur Macintosh, relié par un *hub* à huit joysticks. Deux enceintes servaient à la fois pour la diffusion du son et d'enceintes dites « de retour » pour les élèves. (Figure 3)

Parmi les vingt-quatre élèves de la classe, huit élèves manipulaient et les seize autres écoutaient. Le professeur de M.A.O. faisait effectuer trois rotations d'environ dix minutes sur le dispositif pour que chaque élève puisse manipuler les joysticks. Le réservoir de sons était changé à chaque nouveau groupe pour renouveler l'écoute. L'activité était entièrement menée par le professeur et par son élève adulte qui dirigeait la manipulation des joysticks (Figure 3).

Lors des dernières séances, les ateliers d'interprétation des sons travaillés se sont orientés vers l'interprétation de la pièce *More Slowly Please* qui était achevée. Le travail consistait alors à déclencher les boucles et les sons à des moments spécifiques indiqués sur la partition. Les élèves jouaient sur une bande réalisée par ordinateur qui simulait les instruments de l'orchestre.



**Figure 3.** Interprétation en classe

### 2.3 Évaluation technique et pédagogique

Sur le plan pédagogique, la volonté du compositeur était de proposer une analyse de la pratique créative et musicale. Cet objectif a été de notre point de vue réussi.

Cette analyse critique a été réalisée par les élèves et par le compositeur. Elle a permis notamment d'explicitier la démarche de l'artiste et le sens du projet au fur et à mesure du parcours. Concrètement, les élèves ont tenu un journal de bord du projet. Ce journal a permis d'en faire l'historique et de relier les différents éléments donnant accès à une lecture globale du projet afin d'éviter le morcellement des apprentissages. Ce travail permet aux élèves de relier les ateliers et les sorties entre eux, comme faisant partie d'un ensemble cohérent. Ce travail prend d'autant plus de sens pour l'élève qu'il est relié aux disciplines de la classe. Le journal de bord est en relation avec la matière « histoire des arts » et sert de sujet de discussion pour un travail d'expression linguistique.

Dans ce projet, plusieurs valorisations ont été proposées. Il y a eu, bien sûr, la création des œuvres musicales lors de deux concerts. Précédemment, un autre concert avait eu lieu au sein de l'établissement scolaire devant un public de professeurs et de professionnels de la musique. De plus après les concerts, chaque élève a reçu la partition d'orchestre des deux œuvres et un cd/dvd contenant des traces du travail, les matières sonores et un enregistrement des pièces réalisé lors des concerts.

Ce projet artistique a mis l'accent sur une articulation avec le travail de classe. Ce lien est très porteur de sens et il suscite l'engagement des participants, aussi bien des élèves que des professeurs de la classe.

Ce projet était en lien avec l'apprentissage des langues française, allemande et anglaise. Le compte-rendu des expériences et le journal de bord ont été rédigés par les élèves sous forme de résumé, de réflexion ou d'article sur le blog de l'école et de cette manière le lien avec les Technologies de l'Information et de la Communication (T.I.C.) qui sont également au programme.

Les dimensions musicales et culturelles du projet l'inscrivent naturellement dans l'enseignement de

l'histoire des arts : la découverte d'un orchestre philharmonique, de la musique contemporaine, de compositeurs importants, des différents timbres créés par ordinateur ou produits par des instruments acoustiques.

Sur le plan technique nous dresserons le bilan suivant : cette classe de collège possède ses propres ordinateurs, ce qui est un avantage indéniable par rapport aux écoles élémentaires que nous connaissons. Toutefois, certains ordinateurs ne peuvent pas être utilisés faute d'équipement multimédia complet. Les ordinateurs ne possédaient pas de casque ou d'enceinte de bonne qualité. Le travail sur le son était donc concentré sur les événements plutôt du type « émergence » qui sortaient des hauts-parleurs.

Le fait de n'avoir qu'un retour collectif sur les sons manipulés à l'aide des joysticks entraînait à notre sens une certaine confusion dans les sonorités. Il serait intéressant d'avoir une sorte d'enceinte-retour pour chaque joystick. Ce retour pourrait se faire avec un haut-parleur ou avec un casque.

En somme, si on améliorait la qualité sonore des productions alors on valoriserait cette pratique musicale.

En conclusion nous soulignons que la tenue d'un projet comportant autant d'articulations entre les éléments qui le composent est exemplaire. Cela permet à chaque acteur de trouver sa place dans le projet et de le porter dans son domaine de prédilection. Le bon déroulement des étapes du projet et de l'avancée du travail a été rendu possible par une bonne entente des acteurs et par des modes de concertation variés (réunion, mails, appels téléphoniques). Ce temps de concertation n'est pas, contre toute attente, une évidence dans ce type de projet. C'est pourquoi il nous semble important de prévoir un temps de concertation « institutionnalisé » pour s'assurer que tous les acteurs (élèves et adultes) puissent se retrouver en situation de participation pleine ou périphérique et ainsi se sentir pleinement investis.

La troisième expérience que nous avons analysée généralise la manipulation d'un contrôleur de jeu vidéo par les élèves avec l'introduction de la *wiiboard* que nous présentons.

## 3. LE CONTRÔLEUR *WIIBOARD* EN ACTIVITÉ PÉDAGOGIQUES.

### 3.1. Contexte et enjeux

Un enseignant d'une classe de CP a déposé un projet consistant à faire jouer ses élèves (non musiciens) au sein de l'orchestre lors d'un spectacle de fin d'année. Le directeur du conservatoire a proposé d'aménager une répétition avec l'orchestre symphonique amateur de la ville de Noisy-le-Sec en vue d'une restitution devant les parents. Ce spectacle n'était qu'un des éléments du projet qui était aussi très riche : présentation des instruments que les enfants avaient chez eux, rencontre avec les parents pour deux concerts de la jeune philharmonie, répétitions privées de l'orchestre ouvertes aux familles,

rencontre avec un chef d'orchestre, construction d'un lien avec les musiciens de l'orchestre.

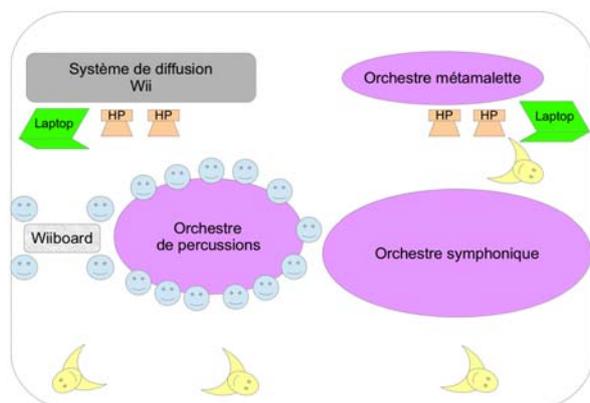
Le choix s'est porté au départ pour des questions matérielles sur la pièce *More Slowly Please* de Gilbert Artman et Jean-Philippe Dejussieu qui était en cours de re-création pour le projet dont nous avons déjà parlé. Nous avons pu rencontrer l'orchestre lors de trois répétitions et une générale le jour du concert.

Les élèves utilisaient de petites percussions pour jouer une partie non écrite dans l'œuvre originale. Les élèves ont appris des éléments de direction d'improvisation. A partir de leurs percussions et du dispositif technologique les enfants devaient produire des sons continus, des sons brefs et irréguliers (pointillisme) et moduler ces sons en jouant sur le volume et le tempo.

Ce projet était au départ indépendant du projet avec les collégiens. Cependant, à la suite d'une répétition avec l'orchestre et les joysticks, les élèves de CP ont apprécié l'ajout de sonorités électroniques. Grâce au dispositif de la *wiiboard*, nous avons pu proposer aux élèves de manipuler ces sons et d'intégrer des sons électroniques parmi les sonorités des percussions.

Nous avons alors utilisé un dispositif que nous avons baptisé « WiiSample ». Il permet à l'élève de faire varier l'intensité sonore de quatre sons électroacoustiques préchargés, en déplaçant ses pieds et donc le poids du corps. Ces variations sont réalisées en agissant sur un contrôleur de jeux vidéo appelé *Wiiboard*, capteur de pression liée au poids sur lequel l'utilisateur se tient debout.

Comme cette proposition a rencontré un vif succès nous avons dû veiller à ce que la *wiiboard* [...] ne soit pas un pôle d'attraction trop fort par rapport aux autres instruments. Nous ne souhaitons pas créer de déséquilibre. En effet, nous défendons dans ce projet l'intégration des sonorités électroacoustiques dans la musique comme un simple élargissement de la palette sonore pour faire de la musique. L'utilisation de la *wiiboard* devait devenir simple et évidente pour jouer cette pièce avec l'orchestre (Figure 4).



**Figure 4.** Usage d'une *wiiboard* en orchestre

### 3.2. Conception et réalisation du dispositif

L'objectif était donc de concevoir un dispositif permettant aux élèves de poursuivre les activités de production sonore en adéquation avec les gestes de direction appris. En l'occurrence un dispositif permettant de déclencher un son, de l'interrompre instantanément, de produire un son discontinu, de produire un son continu et à tout moment de contrôler l'intensité sonore de la production.

En amont, lors de nos séances pédagogiques, nous avons eu la chance de pouvoir travailler à partir de la banque de boucles sonores élaborées par le compositeur. Une fois ces boucles identifiées et reconnues par les élèves pour leurs éléments mélodiques répétitifs, nous les avons chargées dans le dispositif. Dès lors, nous avons consacré la *wiiboard* à deux usages différents.

Dans un premier temps, nous avons travaillé avec un support CD contenant un enregistrement témoin de l'orchestre philharmonique joué sur instrument midi. L'usage de la *wiiboard* a permis aux élèves d'intégrer en temps réel les boucles élaborées par le compositeur au fur et à mesure de la création. Ce travail pour les élèves comportait plusieurs activités simultanées : réalisation de mouvements lents et continus, réponse au geste de direction, adaptation du mouvement au changement de caractère de la musique.

Les élèves se déplaçaient sur l'enregistrement CD avec des mouvements lents inspirés des longues tenues de ce passage de l'œuvre. Au signal du chef d'orchestre, certains d'entre eux se déplaçaient vers la *wiiboard* pour déclencher une des quatre boucles disponibles à l'aide du contrôleur. Au déclenchement d'une boucle, le caractère rythmique et saccadé devait entraîner chez les enfants un changement d'attitude corporelle. Bien que le motif mélodique soit le même sur les quatre boucles, la nature des sons du point de vue du timbre et de l'énergie entraînait des mouvements corporels différents. Quand les élèves coupaient le son de la *wiiboard*, les autres reprenaient des mouvements lents mis en relation avec l'enregistrement du CD.

Dans un second temps, nous avons utilisé la *wiiboard* avec un nouveau réglage. A partir des mêmes boucles sonores nous avons sélectionné une très petite portion du son mise en boucle pour obtenir un son continu.

L'élève appelé sur le dispositif réalisait des sons continus en appliquant son poids sur l'instrument et discontinus en faisant par intermittence de petites pressions brèves avec ses pieds. Enfin, il avait la possibilité de jouer plus ou moins fort suivant le poids porté sur l'instrument.

C'est ce dernier type d'utilisation qui a été choisi par les CP lors de la restitution. Les boucles originales dans leur intégralité ont été jouées par des adolescents membres de l'orchestre à l'aide de joysticks, comme nous l'avons décrit dans la précédente expérience. Au déclenchement de ces boucles, les CP changeaient leur attitude corporelle et entonnaient une chanson dont les paroles étaient de leur création et en rapport avec l'esprit et la nature des sons (chanson sur le thème des robots).

### 3.3. Évaluation technique et pédagogique

L'expérience nous paraît comporter les éléments importants que nous avons tenté de développer dans notre recherche : le soutien d'un projet pédagogique et artistique par ses principaux acteurs, l'adaptabilité du dispositif aux contraintes matérielles et la constitution d'une pratique musicale partagée. Les éléments suivants nous semblent illustrer les situations que nous avons cherché à créer au travers de notre recherche.

Ainsi, la pièce a été rejouée après sa création dans une école élémentaire, avec pour principaux spectateurs les familles. Quelques musiciens du conservatoire ont également assisté au concert ainsi que le compositeur de l'œuvre, Gilbert Artman. Le directeur du conservatoire qui a recréé la pièce a présenté le travail collectif artistique et pédagogique en tant que compositeur et chef d'orchestre. Le directeur de l'école dont fait partie l'enseignante qui a proposé le projet était aussi présent. Le projet a donc été présenté dans toutes ses dimensions, en présence de tous les partenaires qui le portaient.

Nous tenons à préciser que les trois ensembles que sont l'orchestre, les élèves de CP et les joysticks n'en faisaient qu'un et que de nombreux échanges de regard entre ces trois groupes témoignaient d'une écoute mutuelle et d'une réalisation commune.

Du point de vue du dispositif, l'ordinateur était donc d'un côté du groupe et le contrôleur de l'autre. Ceci a permis une bonne mobilité des élèves lors de leurs déplacements, sans qu'ils soient gênés par des obstacles dus au câblage. De plus nous avons pu vérifier que le dispositif permet une utilisation collective au sein d'un orchestre de petites percussions. Nous sommes d'ailleurs satisfaits que l'utilisation du dispositif lors du concert par une seule partie des élèves n'ait pas entraîné de frustration ou de jalousie. A notre sens, ce résultat est dû à toute la préparation pédagogique en amont au cours des séances hebdomadaires durant lesquelles les enfants ont pu s'essayer au dispositif.

Du point de vue de la participation des acteurs du projet, l'expérience a révélé qu'au cours du projet le professeur est passé d'une participation périphérique (dans sa pratique musicale) à une participation pleine. En effet, le professeur s'est mis à diriger sa classe avec les éléments de langages de direction que nous avons appris aux élèves durant l'année. Cette situation donne à voir la réussite d'un partenariat entre un enseignant de l'Éducation Nationale et le conservatoire dont le musicien intervenant fait partie.

Lors de cette restitution nous avons clairement vu l'échange qu'il y a eu entre les élèves jouant avec la *wiiiboard* et les musiciens de l'orchestre. Ceci représente pour nous du point de vue de la mixité des pratiques musicales, l'intégration du dispositif au sein de l'orchestre. D'autre part, du point de vue de l'élève entrant dans ce type de pratique musicale, nous voulons y voir la reconnaissance explicite par les musiciens de l'orchestre de la place du nouvel arrivant dans le groupe.

## 4. PROLONGEMENTS ET PERSPECTIVES

### 4.1. Synthèse

A l'issue de nos expériences personnelles de Musique Assistée par Ordinateur en milieu scolaire et de nos propositions en la matière, nous avons tenté de dresser une synthèse des résultats obtenus dans cette étude.

En effet, la manipulation n'étant pas une fin en soi, nous avons relevé l'intérêt d'accompagner la pratique par des apports théoriques et culturels en tirant parti des documentations multimédia [5].

De plus, l'aspect de partage et de socialisation étant une dimension importante de la musique, nous avons évoqué que l'usage des technologies de la communication par le biais des sites et des blogs était un moyen particulièrement adapté dans ce contexte pour valoriser cette pratique musicale.

Enfin, dans le but de déterminer quels sont les potentiels d'usages des dispositifs technologiques dans un contexte pédagogique et de nous assurer de leur adaptabilité aux conditions du milieu scolaire nous avons établi une méthodologie sous la forme de tableaux [6].

À l'issue de nos recherches, nous avons pu constater qu'en milieu scolaire, les contraintes dues aux programmes, aux horaires et aux conditions matérielles sont très prégnantes. Pour tenter de les surmonter, nous avons fait appel aux notions issues de l'étude des apprentissages informels.

Enfin, l'étude croisée des avancées des outils technologiques et des apports des sciences de l'éducation nous a permis d'une part d'affiner l'élaboration de dispositifs techniques et d'autre part de faciliter leur insertion dans des pratiques musicales différentes. Ces croisements entre recherches et expériences et ces collaborations entre artistes, pédagogues et chercheurs nous ont paru particulièrement pertinents.

Certains projets présentés dans cet article nous ont même semblé être des modèles pour que la pratique musicale assistée par ordinateur se développe à l'école au même titre que les autres pratiques musicales. C'est pourquoi, pour conclure cet article, nous proposons des pistes de réflexions aux professionnels au travers des établissements qui les emploient.

### 4.2. L'usage des T.I.C. pour la pédagogie musicale

Concernant l'Éducation Nationale, il nous semble indispensable que l'Institution rattrape l'évolution de la société en termes d'habitudes musicales et de communication. Cela passe par l'accès et l'utilisation du multimédia, des blogs, des mails et des réseaux sociaux. En effet l'usage des T.I.C. permet actuellement une valorisation de cette pratique musicale qui nous apparaît comme essentielle pour la reconnaissance de l'activité. Pour cela, il faut continuer à équiper toute les écoles d'ordinateurs (comme dans certains départements

des Bouches-du-Rhône, d'Ille-et-Vilaine, de l'Oise et de la Corrèze dans les années 2000), ainsi que d'une connexion à internet. Cela induit la présence d'un personnel technique ultra compétent pour en assurer la maintenance [7].

Ce travail de mise en correspondance de la vie des élèves avec la vie de l'école permettra de favoriser le développement de communauté de pratique [12] dans lesquelles les adultes ont des raisons de s'investir et qui par là même participera à créer du sens pour les jeunes élèves. Cette vision tend vers une diminution de la ségrégation entre la musique à l'école comme une matière et la pratique musicale à l'extérieur comme une pratique culturelle et sociale.

S'agissant des conservatoires, nous considérons qu'on doit ouvrir un département M.A.O. en leur sein et pas seulement pour les musiques actuelles. En effet, la transversalité de la discipline par rapport aux autres activités instrumentales apparaît comme évidente. C'est selon nous une façon de décroquer les disciplines dans les écoles de musique et de changer les conceptions erronées entourant la Création Assistée par Ordinateur.

Nous touchons là à la prise de conscience de la part des professionnels concernés, à savoir les musiciens, les pédagogues et les chercheurs, des préoccupations de chacun.

### 4.3. Formation d'une communauté de pratique

Selon nous, cette connaissance des attentes de chacun peut être favorisé par la constitution ou la prise de conscience d'une communauté de pratique dans un engagement mutuel et dans le but d'une entreprise commune autour d'un répertoire partagé. Cela passe par une reconnaissance mutuelle des compétences de chacun. Cette communauté peut se fédérer comme nous l'avons étudié autour d'un projet commun. Comme semble le proposer le SCRIME qui avance « l'idée de monter des parcours musicaux avec les instruments actuels et reposant sur un réseau de compétences » [8].

En effet comme nous l'avons décrit, la définition d'un projet commun, qui peut être impulsé par la direction des Conservatoires, des laboratoires ou de l'Éducation Nationale, est l'occasion de discussions et d'échange en partenariat entre les professionnels. D'ailleurs cette dynamique existe déjà à différents échelons des politiques culturelle et d'éducation (projet culturel sur un territoire, projet de développement, projet d'établissement ou projet de classe...). De manière systématique, il manque, selon nous, un temps de concertation institutionnalisé ou tout du moins prévu dans le projet dans lequel chaque partenaire, dans un temps suffisant et dans un espace adéquat puisse faire état de ses avancées, de ses attentes et de ses difficultés.

De ce point de vue, nous considérons que les ressources humaines et le directeur d'une équipe doivent revoir en profondeur leur manière de faire travailler les personnels compétents en vue du développement de projets épanouissants pour les élèves et leur entourage,

ainsi que pour les professionnels qui les accompagnent dans leur développement personnel. C'est pourquoi la mise en place d'outils de travail collaboratifs via les T.I.C. nous semble une piste intéressante.

### 4.4. Cadre conceptuel et approche pratique

Dans nos recherches, nous nous intéressons donc à comprendre quel est l'apport des T.I.C. et des outils conceptuels de l'apprentissage en situation informelle dans la constitution de communautés de pratique. Parallèlement nous nous impliquerons dans cette communauté en tant que pédagogue. Nous opérerons ensuite une synthèse en établissant une grille de lecture de cette activité humaine de transmission et nous tenterons de modéliser le processus d'implication dans la communauté de pratique. Notre contribution consistera à élaborer des dispositifs pédagogiques prenant la forme d'un engagement dans une communauté de pratique, à construire des dispositifs techniques pour pratiquer la création musicale assistée par ordinateur et à encourager la mise en relation des professionnels de l'Éducation Nationale et du Conservatoire dans le but d'une entreprise commune. Nous nous appuyerons pour cela sur les travaux de Jean Lave et d'Étienne Wenger qui redéfinissent l'apprentissage comme une construction de compétence par l'individu en « s'engageant dans une pratique sociale » [2]. Dans cette théorie, l'individu apprend dans une situation et un contexte donnés. Cet apprentissage se fait de manière informelle et implique des « interactions entre expert et nouveaux venus dans une pratique culturelle » [2]. Ces interactions ont lieu au sein d'un groupe social défini qu'ils appellent « communauté de pratique » [12]. À l'intérieur de cette communauté de pratique, l'individu apprend en participant.

Nous analyserons les séances de musique en milieu scolaire en veillant à faire la distinction entre « apprentissage informel en tant que lieu (ou cadre) et [l']apprentissage informel en tant que méthode » [10].

À l'issue de ces analyses nous espérons proposer des dispositifs pédagogiques suffisamment riches dans lequel l'élève viendra nourrir lui-même son répertoire de pratique [9].

Pour concevoir des séances pédagogiques à l'intérieur de ce dispositif nous utiliserons les concepts d'« affordance » du dispositif et d'« engagement » dans l'activité qui sont issus également des théories des apprentissages informels.

Nous résumerons les affordances par ce que propose le dispositif et qui va favoriser ou non la participation et l'engagement de l'élève. En effet nous considérons que plus fort sera l'engagement de l'élève dans l'activité plus forte sera sa participation et par conséquent son apprentissage [4].

Nous étudierons la participation de l'élève sous l'angle « du mode d'engagement et d'investissement qui est tout autant défini par l'agent que par l'espace que laisse la communauté » [2] et selon la façon dont l'élève

procédera à une forme d'exploration curieuse comme elle est décrite par Nicolas Auray [1].

Les séances de musique que nous proposerons consisteront donc pour les élèves dans un premier temps à s'approprier cet outil technologique.

Les résultats obtenus nous permettront ainsi d'affiner l'élaboration des dispositifs techniques que nous proposons et de faciliter leur insertion dans une séance de production musicale.

Enfin en nous appuyant sur ces mêmes concepts nous essaierons de penser le rôle des participants (autres que les élèves) en étudiant ce que propose le dispositif comme engagement aux autres membres de la communauté de pratique, en particulier les professeurs, les musiciens intervenants et les autres acteurs que nous avons détaillés dans le chapitre concernant les contraintes humaines.

Ce dernier point, selon nous, nous renvoie au processus d'apprentissage inter-psychologique dans la pratique de coparticipation au travail [3]. Dans cette démarche « les individus décident de ce qui fait la valeur de ces propositions (*affordances*) à s'engager au travail, mais choisissent également une façon d'y participer, construisent des significations et incarnent ces pratiques. » [3].

Les résultats des recherches exposées dans cet article portaient sur des expériences menées en France. Dans le cadre d'un doctorat débuté en novembre 2013, nos travaux intégreront également un état de l'art des pratiques en milieu scolaire au niveau international.

Voici la façon dont nous essaierons de répondre aux multiples enjeux que représente à nos yeux l'apprentissage de la musique assistée par ordinateur en milieu scolaire.

## 5. RÉFÉRENCES

[1] AURAY Nicolas, 2011 : « Les technologies de l'information et le régime exploratoire », in van Andel, P., Bourcier, D., eds., *La sérendipité dans les arts, les sciences et la décision*, éditions Hermann.

[2] BERRY Vincent, 2008 : « Les communautés de pratique : note de synthèse », *Pratique de formation, analyse*, n° 54, juin, <http://www-ufr8.univ-paris8.fr/pfa/54presentation.htm>, consulté le 2 Juillet 2012.

[3] BILLET Stephen, 2008 : « Les pratiques participatives sur le lieu de travail : Apprentissages et remaniement de pratiques culturelle », *Pratique de formation, analyse*, n° 54, juin, <http://www-ufr8.univ-paris8.fr/pfa/54presentation.htm>, consulté le 2 Juillet 2012.

[4] BROUGERE Gilles, 2009a : « Vie quotidienne et apprentissages » in G. Brougère, A.-L. Ulmann (dir.), *Apprendre de la vie quotidienne*, Paris, PUF, p. 21-33.

[5] FORT Bernard et SAINT-MARTIN Dominique (Groupe Musique Vivante Lyon), 1995 : *Les sons de la nature*, Courlay, Jean-Marie Fuzeau, Coll. « Zoom sur ».

[6] GALLERON Philippe, 2012 : « Création musicale assistée par ordinateur : penser des cadres et

des outils pour une approche pratique par l'écoute et la manipulation », mémoire de master 2, Université Paris VIII.

[7] KAHNEBOUBI Medhi, 2009 : « Descriptions de quelques tentatives de dotations massives en ordinateurs portables en France », Science et Technologie de l'Information et de la Communication pour l'Éducation et la Formation, <http://sticef.univ-lemans.fr/>, consulté le 27 Janvier 2012.

[8] KURTAG György, DI SANTO Jean-Louis, DESSAINTE-CATHERINE Myriam, GUILLEM Philippe 2007 : « Pédagogie de l'électroacoustique du geste musical à la composition assistée par ordinateur », *In Actes des Journées d'Informatique Musicale JIM'07, Lyon (France)*.

[9] ROGOFF Barbara, PARADISE Ruth, ARAUZ Rebeca Mejia, CORREA-CHAVEZ Maricela, et ANGELILLO Cathy 2003 : « Firsthand learning through intent participation », in *Annual Review of Psychology*, University of California - Santa Cruz 2003.54:175-203, téléchargé depuis le site [arjournals.annualreviews.org](http://arjournals.annualreviews.org) le 12 Mars 2007, p.175-203.

[10] SCHUGURENSKY Daniel, 2007, « Vingt mille lieues sous les mers : les quatre défis de l'apprentissage informel », *Revue française de pédagogie*, n°160, juillet-août-septembre, p.13-27.

[11] VEITL Anne, 2001 : « Quelles ressources technologiques pour renouveler la pédagogie de la musique ? Présentation critique d'outils », Premier acte, agence conseil, Ministère de la Culture-DMDTS, Avril 2001

<http://www.culture.gouv.fr/culture/dmdts/rapportsPDF/veitl.pdf>, consulté le 18 Janvier 2012.

[12] WENGER Etienne, 2005 : *La théorie des communautés de pratique. Apprentissage, sens et identité*. Québec, Presses de l'Université de Laval.



# ETUDIER ET JOUER UNE ŒUVRE MUSICALE EN MANIPULANT DES OBJETS

*Fabrice Guédy*

Atelier des Feuillantines

[Fabrice.Guedy@feuillantines.com](mailto:Fabrice.Guedy@feuillantines.com)

*Frédéric Bevilacqua*

IRCAM

[Frederic.Bevilacqua@ircam.fr](mailto:Frederic.Bevilacqua@ircam.fr)

*Dominique Fober*

GRAME

[fober@grame.fr](mailto:fober@grame.fr)

*Loïc Robert*

Ecole Normale

Supérieure de Cachan  
[lrobert@ens-cachan.fr](mailto:lrobert@ens-cachan.fr)

## RÉSUMÉ

Cet article présente trois expériences musicales et pédagogiques menées à l'Atelier des Feuillantines. Elles s'inscrivent dans le cadre de divers projets nationaux comme « Interlude » [1] ou « Musique Lab 2 » [2][3][4]. Certains de ces projets visent à préfigurer ce que pourrait être l'écoute d'une œuvre au moyen de technologies de suivi de geste, dans des contextes allant de la pédagogie musicale aux installations grand public [5][6], d'autres sont explicitement axés sur le développement de nouveaux outils croisant pédagogie et création.

Les trois expériences décrites dans cet article appartiennent à deux grandes catégories. La première concerne l'utilisation d'un outil associant des gestes, une partition et un enregistrement, la seconde consiste à jouer une œuvre, en l'occurrence ici le canon par tons de l'Offrande Musicale de Bach, et une partie de la Rousserolle Effarvate d'Olivier Messiaen à partir de leurs constituants de base matérialisés par des objets dont la manipulation coïncide avec des traitements reproduisant le « mouvement » de la composition.

## 1. CONTEXTE

### 1.1. Contexte musical : l'Atelier des Feuillantines

L'Atelier des Feuillantines est à la fois un conservatoire et une école d'arts plastiques présentant l'originalité de faire participer les élèves à des expériences s'insérant dans leur cursus académique, souvent en collaboration avec l'Ircam. Les principales classes de musique sont : une classe de piano, deux classes de violon, une classe de violoncelle et de viole de gambe, et 6 classes de formation musicale (solfège). C'est essentiellement dans ces dernières que les élèves élaborent des projets d'installation ou de performance sur des œuvres ou des phénomènes musicaux souvent liés aux pièces qu'ils travaillent.

C'est également dans ce contexte que l'Atelier des Feuillantines a été partenaire de plusieurs projets nationaux et européens (Interlude, I-Maestro, Musique Lab 2, Phase), afin de permettre à des équipes de recherche de bénéficier d'un contexte d'expérimentation musicale sans avoir les inconvénients d'un concert ou d'un événement public à échéance fixe. Par ailleurs, l'atelier garantit le sérieux et la pertinence musicale des

expériences menées car elle font partie du cursus des élèves.

### 1.2. Contexte technologique

Les dispositifs mis en œuvre dans les expériences musicales décrites ci-dessous reposent sur des modules de captation du geste (MO), des technologies de suivi de geste [5][12], de traitement du signal (MuBu) [13], et d'affichage de partition (INScore) [9], (Figure 1).

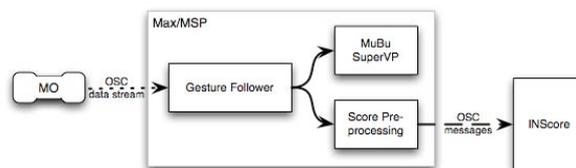


Figure 1 : Architecture technique

Les objets musicaux (MO) [1] transmettent des flots de données (accéléromètre 3D) au suivi de geste qui est implémenté dans Max/MSP. Ce dernier contrôle le module de traitement du signal (vocodeur de phase SuperVP et/ou des modules MuBu [13][12] ainsi qu'un module qui pilotent la partition avec des messages OSC.

Enfin le rendu graphique, sous forme de partition dynamique, est produit par INScore, un environnement pour le design de partitions augmentées [10]

## 2. EXPERIENCE 1 : L'APPRENTISSAGE DE LA BATTUE ET DE SES METAPHORES

Battre la mesure est une pratique qui commence assez tôt dans le cursus des élèves. Elle s'appuie sur un travail régulier concernant la maîtrise de la pulsation, durant lequel les élèves apprennent à se diriger mutuellement, soit dans des exercices de rythme qu'ils écrivent eux-mêmes, soit dans les œuvres qu'ils jouent sur leur instrument. Le travail sur le rubato d'un nocturne de Chopin peut passer par un exercice où les pianistes se dirigent entre eux et doivent restituer par un geste de chef d'orchestre les inflexions musicales de leur interprétation. Afin d'utiliser les bénéfices de ces exercices sur d'autres contenus musicaux que ceux produits directement par les élèves, un outil informatique a été élaboré permettant de suivre des gestes contrôlant des sons enregistrés provenant de sources variées. Ces gestes peuvent être quelconques,

même si la plupart des usages en solfège concernent la battue. Dans un premier temps, on enregistre un geste en écoutant un enregistrement sonore, au moyen d'un module de capteurs électroniques tenu dans la main qui « dirige ». On produit de cette manière des gestes dont les données sont alignées avec le signal audio.



Figure 2 : Le module de capteurs électroniques

Dans un second temps, on refait les mêmes gestes qui sont alors « suivis » : l'ordinateur va dilater ou contracter en temps réel l'enregistrement correspondant, en fonction de la vitesse du mouvement. De cette manière, il devient possible d'explorer plusieurs formes de ralentis ou d'accélération, de travailler les fins de phrases en s'affranchissant des contingences de l'instrument pour ne se focaliser que sur des phénomènes d'interprétation. A titre d'exemple nous leur avons demandé de segmenter une même phrase musicale de plusieurs manières différentes afin d'en faire ressortir les ambivalences, comme si l'on plaçait des virgules et des points dans une phrase sans ponctuation, pour en faire émerger une multiplicité de sens possibles dépendants de l'intonation adoptée.

L'outil permet par ailleurs d'afficher la partition et d'y visualiser un curseur qui suit l'enregistrement audio [7][8][9][10]. Celle-ci peut être intégrée sous deux formats possibles : soit à partir d'images scannées, soit à partir de fichiers MusicXml. Dans le premier cas, il est nécessaire d'indiquer la correspondance entre l'espace graphique des pages et l'espace temporel de l'enregistrement, afin de visualiser le suivi. Dans le second cas, la partition est représentée de manière symbolique, et le suivi se fait automatiquement grâce à la librairie Guido. Cette intégration de la partition permet un lien simple et direct avec quantité de situations pédagogiques liant théorie et pratique : mise en évidence d'éléments singuliers dans une polyphonie (sujet de fugue, thème de sonate, série, etc...), désignation par la main du professeur d'éléments de la partition projetée au tableau afin de soutenir une explication, lecture et suivi de partition rendu aisé par l'animation d'un curseur en temps réel, pour des partitions d'orchestre difficiles à lire.



Figure 3 : un élève tenant le module bâtit la mesure. Il voit un curseur se déplacer sur la partition de l'œuvre qu'il entend. Le professeur a donné comme prescription de mettre en évidence les entrées du sujet (Cours de solfège sur l'Art de la Fugue de Bach)

### 3. LES BENEFICES MUSICAUX DE CES EXPERIENCES

#### 3.1. Le travail de la souplesse

L'alignement entre le signal audio et les données gestuelles se fait de manière littérale, contrairement à la logique musicale d'un instrumentiste en situation de jeu: le temps du geste est mis en correspondance directe avec le temps du fichier son. La figure 3 montre cet alignement.

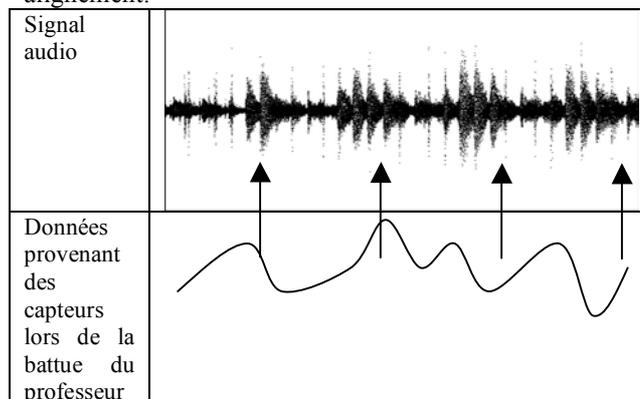
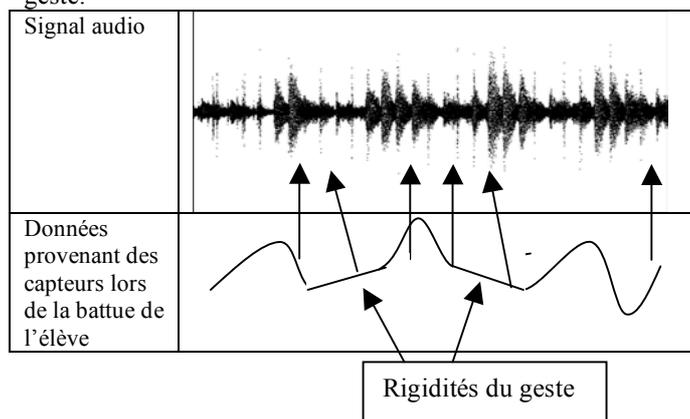


Figure 4 : Alignement littéral entre le temps gestuel et le temps du signal audio.

Si la battue est exécutée de manière plus rigide que dans l'enregistrement original, la tête de lecture va avoir tendance à « sauter » car il manquera une partie du geste:



**Figure 5 :** Dans le second cas, on voit que la tête de lecture de l'enregistrement « saute » et permet de rendre audible la rigidité du geste, de l'améliorer « à l'oreille ».

<p>En notation musicale, nous voyons que s'il y a une série de croches dans l'enregistrement, elles sont</p>	 <p>Signal audio original</p>
<p>transformées en croches pointées – double croches dans la restitution de l'élève, qui peut assouplir son geste en écoutant varier le rythme</p>	 <p>Signal déformé par le geste rigide</p>

**Figure 6 :** une déformation rythmique typique, due à la rigidité d'un mouvement

### 3.2. La direction augmentée

Au cours d'un stage de musique de chambre, nous avons utilisé ce même outil sur « la boîte à joujoux » de Debussy. Pendant que les élèves jouaient sous la direction d'un enseignant, ce dernier, grâce à un module de capteurs tenu dans sa main en même temps que sa baguette, contrôlait un enregistrement de l'œuvre. De cette manière, les élèves suivaient le même geste que celui suivi par l'ordinateur et entendaient leur propre son mélangé à celui de l'enregistrement. Les équilibres de l'orchestration, différente de celle utilisée durant ce stage, permettaient de soutenir les élèves, sans l'inconvénient qu'aurait eu un enregistrement figé.

### 3.3. Les cours d'analyse musicale

Ces technologies permettant de se déplacer de manière synchrone dans un espace musical complexe où se croisent l'enregistrement, la partition et le geste, nous avons pu simplifier des parties souvent ardues de certains cours d'analyse et de solfège. L'outil s'avère efficace pour repérer et expliciter des éléments morphologiques complexes de la partition sans être arrêtés par les difficultés de lecture.



**Figure 7 :** Déplacement du curseur sur la partition pendant qu'un élève dirige en tenant le module de capteurs. Il produit son analyse en manipulant l'outil.

## 4. EXPERIENCE 2 : LA RECONSTRUCTION PHYSIQUE D'UNE ŒUVRE MUSICALE

Un ensemble d'élèves de formation musicale s'est focalisé durant plusieurs mois sur l'étude de la transposition de façon théorique et pratique sur des

œuvres travaillées en classe d'instrument. Nous leur avons proposé comme application du cours de réaliser physiquement le canon par tons de l'Offrande Musicale de Bach. Ce canon se présente sous la forme d'une énigme constituée de deux portées contenant chacune deux clefs, et d'une phrase en latin. (Figure 7).



La seconde clef, l'indication de reprise et le signe indiquent que cette ligne peut être jouée une quinte plus haut décalée d'une mesure, de façon à créer une 3<sup>ème</sup> voix.

La phrase latine signifie « A chaque modulation qui grandit s'élève la gloire du Roi ». C'est un indice qui oriente la résolution de l'énigme vers l'idée d'une transposition ascendante

### Figure 8 : l'énigme

Pour mener ce projet, les élèves ont construit une sculpture permettant d'encaster les modules de capteurs utilisés avec l'outil précédent, à des hauteurs variables. Trois capteurs ont été utilisés simultanément, chacun contrôlant le degré de transposition de l'une des voix. Après la construction de la sculpture, un professeur de violon a procédé à l'enregistrement de chaque voix séparée, au violon et à l'alto. Nous avons choisi d'enregistrer chaque voix en fa# mineur, afin de minimiser les artefacts de l'algorithme de transposition, qui devait se faire sur une octave, entre do et do, pour chacune des voix. De cette manière, nous n'avions à transposer que d'une quarte augmentée ascendante et descendante, alors que si nous avions conservé la tonalité originale de do mineur pour enregistrer les 3 voix à transposer, nous aurions fait subir au signal des déformations importantes.



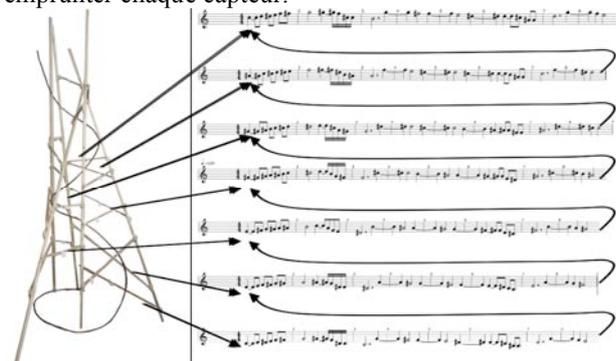
**Figure 9 :** Les séances d'enregistrement avec un professeur de violon de l'Atelier des Feuillantines

Un logiciel permettant de jouer en boucle chacune des trois voix enregistrée a été développé par Loïc Robert, élève de l'Atelier des Feuillantines et étudiant à l'Ecole Normale Supérieure de Cachan, dans le cadre de son stage de L3. Lorsque le module concernant l'une des voix est déplacé d'un étage à l'autre de la sculpture, la prochaine occurrence de la boucle se trouve transposée

d'un nombre de tons correspondant à la hauteur du capteur.



**Figure 10 :** vue d'ensemble et étiquetage des différentes positions afin de repérer le cheminement que doit emprunter chaque capteur.



**Figure 11 :** vue synoptique de la trajectoire de l'un des capteurs contrôlant l'une des voix.

Chaque flèche de la figure 7 correspond au déplacement du module associé à la voix de soprano. Lorsque le module est déplacé d'un étage à l'autre, la transposition correspondante est appliquée à la voix au prochain tour de boucle. Les trois voix de l'Offrande ont ainsi été contrôlées par trois personnes différentes devant déplacer les modules au bon moment et du bon degré de transposition.

L'objet a d'abord été utilisé en cours de solfège pour illustrer le concept de transposition. Dans ce cadre, il a été manipulé par les enseignants comme par les élèves dans des classes de 6 niveaux différents, allant de débutant à supérieur.

Après cette première phase, nous avons intégré la partie notation musicale de l'outil, de manière à pouvoir représenter chacune des voix séparément, avec son propre suivi, de manière à voir sur la partition les différentes transposition opérées en temps réel.



**Figure 12 :** Un cours de formation musicale utilisant le dispositif

#### 4.1. L'objet comme partition et comme instrument

Parce qu'il représente, dans l'espace en 3 dimensions, la structure de l'œuvre, nous pouvons dire que cet objet remplit la fonction d'une partition. Chaque barre sur lequel le module de capteurs est posé est muni d'une étiquette indiquant le chemin à suivre. Pour peu que soient écrites les hauteurs de départ sur la première étiquette le chemin va de soit. Cette partition est également un instrument car elle permet de jouer l'œuvre.

Une performance consistant à jouer l'œuvre sur une octave, lors de l'audition de fin d'année, a été insérée de manière banale au milieu d'œuvres jouées normalement. Cette audition a eu lieu à Paris au lycée Henri IV le 25 juin 2011. Nous avons choisi de la faire dans la chapelle de manière à pouvoir projeter la partition en train de se calculer sur la voûte surplombant l'installation. Une centaine de personnes, parents, élèves, enseignants ont assisté à cet événement, et souhaité utiliser eux-même l'instrument.



**Figure 13 :** Exécution de l'Offrande Musicale en public.

Une première démonstration a d'abord permis de faire comprendre au public le principe de composition, en manipulant un seul module, donc une seule voix, et en simulant une erreur de transposition.

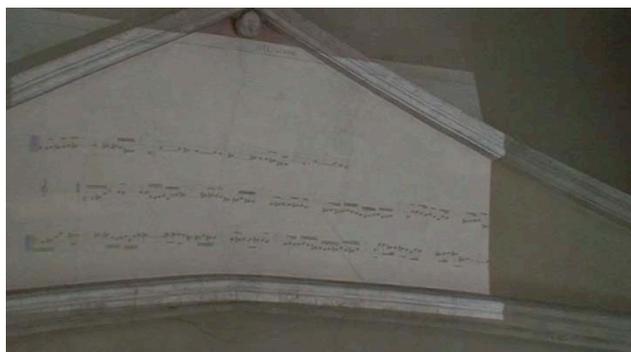


Figure 14 : La projection de la partition se calculant en temps réel à partir du geste, sur la voûte de la chapelle



Figure 15 : L'exécution de l'œuvre en manipulant l'objet par 3 exécutants

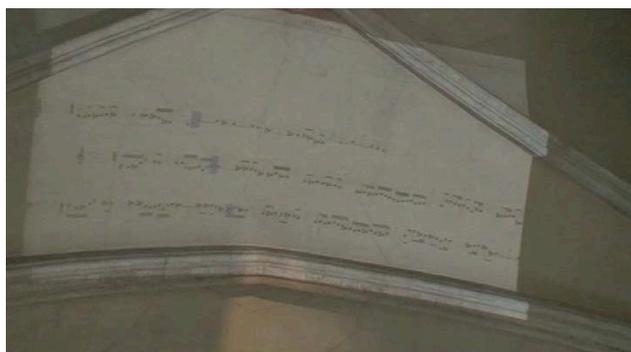


Figure 16 : Le mouvement des curseurs permettant au public de suivre chaque voix

L'exécution de l'œuvre en manipulant les modules de capteurs autour de l'objet-partition permet d'imaginer une écoute intermédiaire, ou l'auditeur assemble lui-même les constituants de l'œuvre. Par ailleurs, la manière dont son geste se saisit de ces éléments et les manipule peut impacter le rendu sonore de ces constituants, ce que nous sommes en train de mettre en œuvre dans un projet en cours.

## 5. EXPERIENCE 3 : LA RECONSTRUCTION DE LA ROUSSEROLLE EFFARVATTE D'OLIVIER MESSIAEN [11]

Composé entre 1956 et 1958, le "Catalogue d'oiseaux" est un ensemble de pièces pour piano dont le plan général, lorsqu'on aligne les tables des matières de

chaque livre, forme un palindrome sur le nombre d'oiseaux traités dans chaque livre :

1er Livre	I . Le Chocard des Alpes II . Le Lorient III . Le Merle Bleu
2ème Livre	IV . Le Traquet Stapazin
3ème Livre	V . La Chouette Hulotte VI . L'Alouette Lulu
4ème Livre	VII . La Rousserolle Effarvate
5ème Livre	VIII . L'Alouette Calandrelle IX . La Bouscarle
6ème Livre	X . Le Merle de roche
7ème Livre	XI . La Buse variable XII . Le Traquet rieur XIII . Le Courlis cendré

Figure 17 : le palindrome des tables des matières

Le palindrome est la suite numérique 3 1 2 1 2 1 3, et on verra plus tard que c'est précisément une succession dont le compositeur s'est servi pour construire certains rythmes rétrogradables, en en modifiant l'ordre. On remarque, par ailleurs, la position privilégiée de la Rousserolle au centre du cycle, formant l'axe du miroir, et cette propriété va engager Messiaen à multiplier les « effets miroir » dans l'œuvre, comme Bach le fait avec la pièce centrale du 1<sup>er</sup> livre du clavier bien tempéré, (Fa# majeur partageant l'octave en 2), pour des raisons similaires.

### 5.1. Modélisation de la main gauche

La composition de celle-ci se structure autour d'une série de 9 accords et de 7 durées qui bouclent, suivant une technique fréquente chez Messiaen. Le bouclage est toutefois plus subtil que dans la liturgie de cristal du quatuor pour la fin du temps, par exemple, car à chaque fois que l'on passe par la 6ème durée, celle-ci s'augmente d'une double croche, créant une prolifération organique dans la permutation symétrique des accords sur les durées.



Figure 18 : La série d'accords

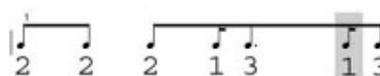


Figure 19 : La série de durées

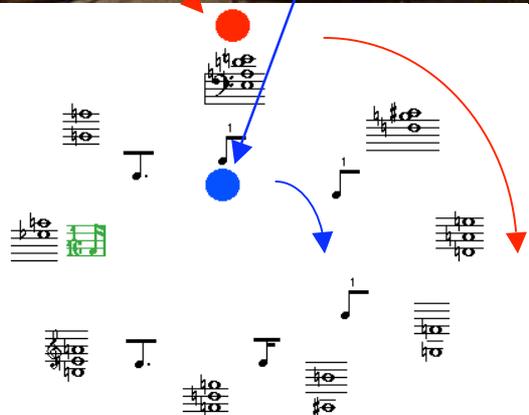
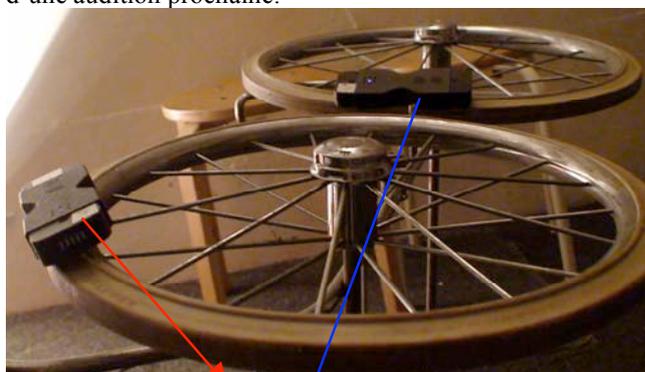
On remarque que les durées exprimées en nombre de double croches sont prises dans le palindrome de la table

des matières de la figure 12. Comme pour l'Offrande Musicale, nous avons construit, ou plutôt récupéré, un objet (ici des roues de landeau) nous permettant de visualiser et matérialiser le cercle des durées et celui des accords, puis de simuler la rotation de l'un sur l'autre. A chaque rotation, un accord correspondant à la position de la roue des hauteurs était émis, d'une durée correspondant à la roue des durées.

En réalisant physiquement le bouclage, ce type d'installation, contrairement à la précédente, permet de produire des variantes de l'œuvre en restant respectueux de celle-ci, comme si elle était un point singulier et optimal dans un espace de possibilités à explorer.

Un élève s'est servi du dispositif pour trouver rapidement des variantes qu'il a ensuite exécuté en cours.

La modélisation de la structure d'accentuation de la main droite a fait l'objet de plusieurs propositions d'élèves, et fera l'objet d'une installation et d'une exécution lors d'une audition prochaine.



L'expérience montre que cet usage, proche d'un instrument, sur un objet détourné, peut être enrichi par le jeu simultané avec un vrai instrument, ce que nous avons fait en faisant jouer la main droite de l'œuvre à un élève de piano, tandis que deux élèves manipulaient les roues. L'effet produit par la rotation n'était pas, dans ce cas, de jouer littéralement un accord d'une certaine durée, mais de jouer toute l'occurrence d'une série d'accords et d'une série de durées, en bouclant l'une sur l'autre comme dans l'Offrande Musicale.

## 6. CONCLUSION

L'unification des technologies de suivi de geste, de traitement du signal et de suivi de partition permet d'imaginer les outils d'une "classe augmentée" sans imposer une pédagogie particulière, de manière à amplifier les effets de méthodes existantes plutôt que de leur en substituer de nouvelles. C'est de cette manière que nous avons pu y intéresser un grand nombre d'enseignants. En outre, ce sont ces mêmes technologies sur lesquelles sont basés beaucoup d'outils de composition, créant un pont entre la création et la pédagogie.

L'absence d'environnement unifié et l'éparpillement des technologies qui implique, pour chaque situation, la présence de développeurs, a cependant beaucoup nuit à la reproduction de ces expériences par d'autres professeurs, qui pourtant auraient pu en enrichir le contenu et la portée. Cette situation provient de l'absence d'enseignants dans les phases d'élaboration des projets pédagogiques essentiellement conçus par des scientifiques. Les enseignants sont généralement appelés lors des phases de validation, souvent délicates.

## 7. REMERCIEMENTS

Les expériences décrites ont fait l'objet d'un soutien de l'Agence Nationale de la Recherche, Projet Interlude ANR-08-CORD-010, « Nouveaux paradigmes numériques pour l'exploration et l'interaction gestuelle expressive avec des contenus musicaux », ainsi que du Ministère de l'Education Nationale, et de la Direction de la Musique du Ministère de la Culture, projet Musique Lab 2, visant à créer un environnement d'aide à la pédagogie musicale pour les collèges et les conservatoires.

## 8. REFERENCES

- [1] N. Rasamimanana, F. Bevilacqua, N. Schnell, F. Guedy,, E. Come Maestracci, B. Zamborlin, JL. Frechin,U. Petrevski,« Modular Musical Objects Towards Embodied Control Of Digital Music », Tangible Embedded and Embodied Interaction, 2011
- [2] F. Guédy, J. Bresson, G. Assayag, Musique Lab 2 - Un environnement d'aide à la pédagogie musicale, Actes des Journées d'Informatique Musicale - JIM'07, Lyon, France, 2007.
- [3] J. Bresson, ML-Maquette / Musique Lab 2, Proc. International Computer Music Conference, New York / Stony Brook, NY, USA, 2010
- [4] J. Bresson, F. Guédy, G. Assayag, Musique Lab Maquette : approche interactive des processus compositionnels pour la pédagogie musicale, Sciences et Technologies de l'Information et de la Communication pour l'Education et la Formation, 13, 2006

- [5] F. Bevilacqua, F. Guédy, N. Schnell, E. Fléty, N. Leroy, " Wireless sensor interface and gesture-follower for music pedagogy", Proc. of the International Conference of New Interfaces for Musical Expression (NIME 07), p 124-129, 2007
- [6] Guédy, Fabrice, L'expérimentation musicale et pédagogique sur le geste de l'Atelier des Feuillantines avec l'équipe Interaction Musicale Temps Réel de l'Ircam, article disponible sur le site de la revue « l'Education Musicale » et sur celui de l'Ircam.
- [7] D. Fober, C. Daudin, S. Letz, and Y. Orlarey. Partitions musicales augmentées. In JIM, editor, Actes des Journées d'Informatique Musicale JIM2010, Rennes, pages 97–103, 2010
- [8] D. Fober, C. Daudin, S. Letz, and Y. Orlarey. Time synchronization in graphic domain - a new paradigm for augmented music scores. In ICMA, editor, Proceedings of the International Computer Music Conference, pages 458–461, 2010.
- [9] D. Fober, C. Daudin, Y. Orlarey, and S. Letz. Interlude - a framework for augmented music scores. In Proceedings of the Sound and Music Computing conference - SMC'10, pages 233–240, 2010.
- [10] D. Fober, Y. Orlarey, and S. Letz. Inscore – an environment for the design of live music scores. In Proceedings of the Linux Audio Conference – LAC 2012, 2012.
- [11] Guédy Fabrice – La Rousserolle Effarvate - Olivier Messiaen, Actes des JIM 2003.
- [12] F. Bevilacqua, N. Schnell, N. Rasamimanana, B. Zamborlin, F. Guedy, Online Gesture Analysis and Control of Audio Processing, Musical Robots and Interactive Multimodal Systems, Springer Tracts in Advanced Robotics, Volume 74, Springer Verlag, pp. 127-142, 2011.
- [13] N. Schnell, A. Röbel, D. Schwarz, G. Peeters, R. Borghesi, MuBu & Friends - Assembling Tools for Content Based Real-Time Interactive Audio Processing in Max/MSP, In Proceedings of the International Computer Music Conference (ICMC), Montreal, 2009



# PERSPECTIVES POUR L'APPRENTISSAGE INTERACTIF DU COUPLAGE GESTE–SON

*Jules Françoise*  
Ircam — UMR STMS  
IRCAM–CNRS–UPMC  
jules.francoise@ircam.fr

*Ianis Lallemand*  
Ircam — UMR STMS  
IRCAM–CNRS–UPMC  
ianis.lallemand@ircam.fr

*Thierry Artières*  
LIP6  
CNRS–UPMC  
thierry.artieres@lip6.fr

*Frédéric Bevilacqua*  
Ircam — UMR STMS  
IRCAM–CNRS–UPMC  
frederic.bevilacqua@ircam.fr

*Diemo Schwarz*  
Ircam — UMR STMS  
IRCAM–CNRS–UPMC  
diemo.schwarz@ircam.fr

## RÉSUMÉ

L'apprentissage de mappings du geste vers le son constitue aujourd'hui un enjeu de recherche majeur. Dans un travail précédent, nous avons proposé un modèle hiérarchique permettant de modéliser des structures temporelles à différentes échelles [8]. Nous nous intéressons ici à l'apprentissage de structures temporelles de plus haut niveau. Plus spécifiquement, nous nous proposons de formuler la problématique de l'*articulation* entre différents mappings geste–son dans un contexte d'apprentissage interactif. Ce champ de recherche émergent, aux croisements de l'apprentissage automatique et de l'interaction homme-machine, permet à notre sens de poser correctement la question de l'apprentissage « par démonstration ». Nous présentons d'abord successivement les cadres de l'apprentissage interactif et de la modélisation du couplage geste–son, puis les perspectives ouvertes par la réunion de ces problématiques, ainsi qu'une première extension de nos travaux précédents dans ce cadre.

## 1. INTRODUCTION : SYSTÈMES INTERACTIFS MUSICAUX BASÉS SUR LE GESTE

### 1.1. Contexte

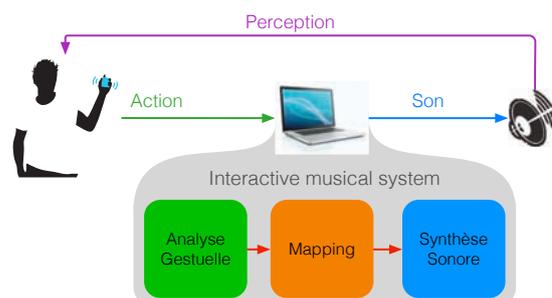
Nous considérons dans cet article des dispositifs interactifs musicaux impliquant une interaction gestuelle expressive avec des processus sonores. Ces systèmes sont fréquemment utilisés dans les domaines artistique et scientifique : installations sonores interactives, contrôle temps-réel pour la performance scénique ou nouvelles interfaces pour l'expression musicale<sup>1</sup>. De manière générale, ils s'articulent autour de trois composantes principales, représentées sur la figure 1 : les gestes de l'utilisateur<sup>2</sup> sont d'abord captés et analysés (1), puis couplés (2)

1. Voir en particulier la conférence internationale NIME (*New Interfaces for Musical Expression*) : <http://www.nime.org/>.

2. Dans cet article, nous nommons « utilisateur » toute personne interagissant avec un système interactif, sans considération pour les pro-

blèmes esthétiques que pourrait poser une telle notion d'« utilisation », comprise dans un sens littéral.

3. Cette approche se rapproche des notions de « mapping par l'écoute » développés dans la thèse de Baptiste Caramiaux [3], et du « *play-along mapping* » introduit par Fiebrink [6].



**Figure 1:** Schéma général d'un système interactif musical.

Souvent, ce mapping est construit par une formulation analytique, réalisée par exemple par la création de liens explicites entre paramètres gestuels et sonores.

#### 1.1.1. « Mapping par démonstration »

Nous considérons ici un cadre alternatif qui s'inscrit dans la continuité des recherches menées au sein de l'équipe Interactions Musicales Temps-Réel (IMTR) de l'Ircam. Appelée « mapping par démonstration », notre démarche consiste à définir le couplage geste–son par l'interaction elle-même en fournissant des exemples de réalisations concrètes du mapping<sup>3</sup>. Ceci induit une boucle d'interaction composée de deux phases :

blèmes esthétiques que pourrait poser une telle notion d'« utilisation », comprise dans un sens littéral.

3. Cette approche se rapproche des notions de « mapping par l'écoute » développés dans la thèse de Baptiste Caramiaux [3], et du « *play-along mapping* » introduit par Fiebrink [6].

1. **Apprentissage.** Dans un premier temps, l'utilisateur « joue » un ensemble d'exemples sonores : simultanément à l'écoute de ceux-ci, il réalise des gestes qu'il souhaite associer aux sons entendus, afin de fournir au système des exemples de réalisations du mapping à apprendre. Le système doit ensuite abstraire le couplage depuis les données d'exemple, par le biais de méthodes d'apprentissage automatique (cf. fig. 2a).
2. **Performance.** Dans un second temps, l'utilisateur peut réinterpréter les exemples sonores de départ en réalisant de nouveau les gestes fournis au système au cours de la première phase. Le couplage geste-son appris par le système permet de traduire les modulations des gestes réinterprétés en modulations sonores (cf. fig. 2b).

### 1.1.2. Travaux récents

Dans ce contexte de mapping par démonstration, nous nous sommes plus particulièrement intéressés à la réinterprétation des sons enregistrés. Les techniques musicales permettant d'aborder la question du *jeu* fondé sur un matériau sonore concret sont nombreuses : *cut-up*, compression ou dilatation temporelle, réalisation de boucles, ...

Nous avons récemment réalisé plusieurs systèmes utilisant le système de suivi de geste *Gesture Follower*<sup>4</sup> afin de combiner de manière expressive ces éléments de vocabulaire musical [1]. En particulier, cette technologie permet de capturer avec précision les dynamiques temporelles de gestes associés à des sons. Pendant la phase d'apprentissage, l'utilisateur effectue un geste aligné à un exemple sonore. Il peut ensuite exécuter de nouveau le geste qui, aligné en temps réel à la référence, permet la réinterprétation du son. Les variations temporelles gestuelles son alors traduites en variations sonores par des opérations de *time-stretching* utilisant un vocodeur de phase.

Cette approche a été récemment étendue par une modélisation hiérarchique des gestes, détaillée en 3.2. Ceci permet le développement de représentations gestuelles plus complexes, autorisant par exemple l'ajout de contraintes sur certaines phases du son (typiquement, les phases d'attaque peuvent ainsi être conservées)<sup>5</sup>.

### 1.1.3. Réalisations musicales

Ces approches développées dans l'équipe IMTR ont donné lieu à des réalisations à différentes échelles, du dispositif instrumental à l'installation interactive. Citons par exemple :

- Les *Modular Musical Objects* (MO)<sup>6</sup>, des interfaces gestuelles portables permettant de construire de nombreux scénarios d'interaction musicale par leur intégration dans des objets du quotidien [18].

4. Voir des exemples d'utilisation de *Gesture Follower* : [http://imtr.ircam.fr/imtr/Gesture\\_Follower](http://imtr.ircam.fr/imtr/Gesture_Follower)

5. Voir une démonstration du système : <http://vimeo.com/julesfrancoise/smc2012hierarchicalmapping>

6. [http://www.youtube.com/watch?v=Uhps\\_U2E9OM](http://www.youtube.com/watch?v=Uhps_U2E9OM)

- *Urban Musical Game*, un jeu sonore interactif utilisant un ballon augmenté<sup>7</sup>.

## 1.2. Enjeux scientifiques

De nombreux enjeux scientifiques résultent des problématiques posées par les modes d'expérience spécifiques aux systèmes interactifs musicaux. Ces enjeux sont liés à des questions de recherche difficiles dans le champ de l'apprentissage automatique.

### 1.2.1. Faible nombre d'exemples

Dans le cas d'un système interactif musical basé sur le couplage entre geste et son, les « règles » conditionnant la création des contenus sonores sont inférées uniquement à partir de données gestuelles et sonores fournies par l'utilisateur. Cette caractéristique limite nécessairement le nombre d'exemples d'apprentissage disponibles : en pratique, de tels systèmes doivent pouvoir être appris à partir d'un petit nombre d'exemples, voire d'un seul exemple (enjeu 1, section 1.2.4).

### 1.2.2. Apprentissage et interaction simultanés

Dans le cadre des systèmes interactifs musicaux de « mapping par démonstration » que nous considérons ici, les données d'apprentissage (initiales ou « d'adaptation », cf. 1.1.1) sont fournies par les gestes mêmes de l'utilisateur, c'est-à-dire par son interaction avec le dispositif. En d'autres termes, l'apprentissage ne peut pas être distingué du fonctionnement « réel » du système. Il se déroule simultanément à l'interaction des utilisateurs avec le dispositif (enjeu 2, section 1.2.4).

### 1.2.3. Adaptation de l'apprentissage à l'utilisateur

La prise en compte des attentes des utilisateurs lors de l'apprentissage constitue un critère fondamental pour des usages inscrits dans un domaine artistique. Il est en effet souvent très difficile de formuler explicitement des critères d'évaluation de la qualité « artistique » ou « musicale » des résultats produits par un système interactif, dans la mesure où les attentes à l'égard de ceux-ci sont fortement dépendantes du contexte et, par extension, du public. L'interaction avec le dispositif, en impliquant activement l'utilisateur, propose de reformuler la question de l'apprentissage d'un certain critère de qualité (et donc, incidemment, de la qualité de l'apprentissage) en termes d'apprentissage par l'interaction avec l'utilisateur (enjeu 3, section 1.2.4).

### 1.2.4. Résumé des enjeux scientifiques

Nous avons ainsi pu mettre en évidence trois enjeux scientifiques majeurs :

#### Enjeu 1 : Apprendre à partir de très peu d'exemples

7. <http://www.youtube.com/watch?v=jXGlvmrGBgY>

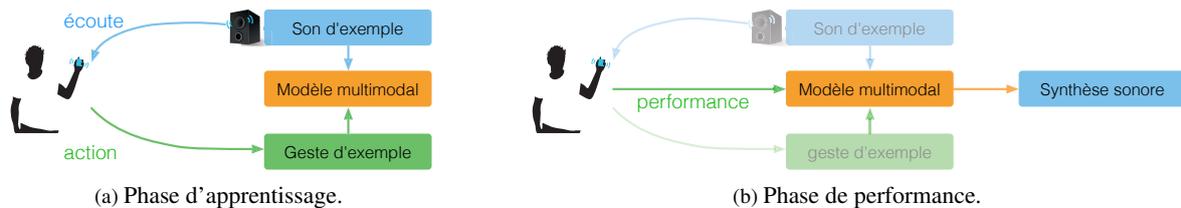


Figure 2: Apprentissage d'un modèle geste-son par « mapping par démonstration ».

Enjeu 2 : Apprendre à partir de données acquises pendant l'interaction

Enjeu 3 : Adapter interactivement l'apprentissage à l'utilisateur

Ces problématiques, prises indépendamment, constituent des champs disciplinaires vastes. Elles sont ici abordées dans la perspective transversale de l'apprentissage automatique interactif (*interactive machine learning*).

### 1.3. Problématique et aperçu

Situé dans la continuité de nos travaux actuels, cet article vise à proposer un ensemble de perspectives sur le problème de l'apprentissage du couplage geste-son à la lumière des développements récents du domaine de l'apprentissage automatique interactif. Ce champ de recherche émergent est introduit dans la section 2 ; nous y présentons l'état de l'art de la recherche actuelle, et établissons un lien avec les problématiques intrinsèques au contexte musical. Synthétisant des recherches récentes menées au sein de l'équipe IMTR, la section 3 présente un cadre de modélisation du couplage geste-son centré sur les aspects temporels. Enfin, la section 4 présente un ensemble de perspectives d'extension des modèles présentés dans la section 3 au cadre de l'apprentissage interactif.

## 2. APPRENTISSAGE INTERACTIF

### 2.1. Définition

L'apprentissage interactif est l'objet d'un intérêt récent dans le domaine de l'apprentissage automatique<sup>8</sup>. Il se positionne généralement à l'interface des disciplines de l'apprentissage automatique et de l'interaction Homme-machine, et se distingue du contexte de l'apprentissage automatique « traditionnel » en ce sens qu'il entend intégrer explicitement la nature *humaine* de l'utilisateur dans le processus d'apprentissage.

8. Voir par exemple les ateliers :

- <https://sites.google.com/site/iui2013imlw/>
- [http://www.cs.utexas.edu/~bradknox/AAAI-FSS-RLIHT12/RLIHT\\_\\_2012-AAAI\\_Fall\\_Symposium.html](http://www.cs.utexas.edu/~bradknox/AAAI-FSS-RLIHT12/RLIHT__2012-AAAI_Fall_Symposium.html)
- <http://research.microsoft.com/en-us/um/people/sumitb/adaiml09/>

### 2.2. Approche transversale des problématiques

Certaines exigences imposées par les contextes artistique et musical inscrivent de fait les enjeux scientifiques présentés en 1.2 dans une perspective transversale. Ces problématiques nous semblent pouvoir être étudiées de manière pertinente dans le cadre de l'apprentissage interactif, lui-même situé à l'interface entre plusieurs domaines de recherche.

#### 2.2.1. Exigence de généralisation

La question de la généralisation d'un apprentissage réalisé à partir de peu d'exemples se pose ici en termes spécifiques : il ne s'agit pas de construire à partir d'un seul exemple une représentation *fixe* des données d'entrée, permettant la classification ultérieure de données similaires. L'enjeu réside plutôt dans la capacité d'apprendre à générer des variations alternatives du matériau de départ, en interaction continue avec l'utilisateur. Le cadre d'apprentissage interactif mêle donc les problèmes de généralisation et d'apprentissage interactif, dans une perspective de *généralisation spécifique à un utilisateur* basée sur l'interaction.

#### 2.2.2. Exigence de simultanéité des phases d'apprentissage et d'interaction

Apprendre par et pendant l'interaction avec un utilisateur établit une identité entre la durée d'interaction avec le dispositif et le nombre d'exemples à partir desquels ce dernier peut être appris. Dans le cadre étudié ici, la question de l'apprentissage à partir de peu d'exemples croise donc celle de l'apprentissage à partir de données d'interaction, dans la perspective commune d'un apprentissage réalisé en une durée suffisamment courte (c'est-à-dire adaptée aux durées typiques d'interaction avec des dispositifs artistiques).

### 2.3. Façonnage interactif

Nous nous intéressons plus particulièrement à une classe spécifique de méthodes, désignées par le terme de « façonnage interactif » (*interactive shaping*).

Le façonnage interactif, récemment formalisé par W. Bradley Knox dans sa thèse de doctorat [14], constitue un cadre permettant de traiter des situations d'apprentissage

dans lesquelles un instructeur humain évalue le comportement d'un système informatique de manière interactive. Dans ce contexte, l'évaluation humaine se compose de « récompenses » scalaires, dont le caractère positif ou négatif correspond au caractère positif ou négatif de l'appréciation.

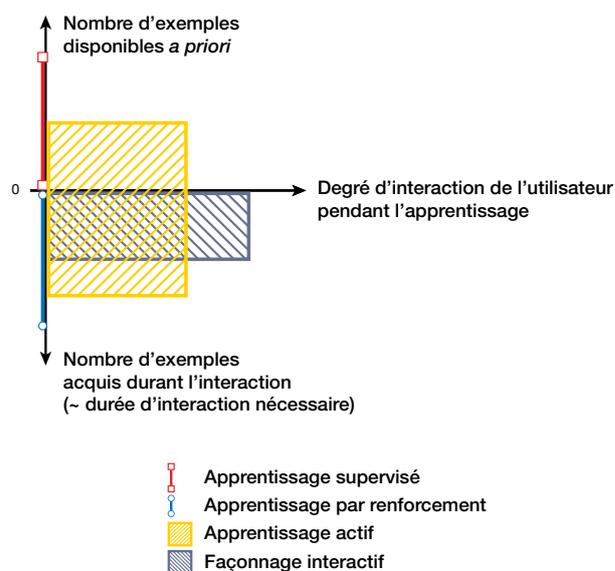
### 2.3.1. Relations avec l'apprentissage par renforcement

Cette notion de récompense s'inspire directement de celle utilisée dans le domaine de l'apprentissage par renforcement (*reinforcement learning*) [19]. Les méthodes d'apprentissage par renforcement fournissent un cadre permettant l'apprentissage d'un système à partir de récompenses attribuées par son *environnement*. Dans la mesure où le terme d'environnement se réfère à tout élément extérieur avec lequel le système est susceptible d'interagir, l'utilisateur d'un dispositif interactif geste-son peut être inclus dans cette notion. Cependant, la formulation « classique » de l'apprentissage par renforcement [17] ne prévoit pas que l'environnement récompense directement le système. Au sens propre, les récompenses ne sont en effet pas reçues par le système, mais calculées par le système à partir d'une spécification interne des issues possibles de son interaction avec l'environnement. Considéré comme l'« environnement » d'un système geste-son, l'utilisateur n'attribuerait donc pas de récompense de manière interactive : une spécification « objective » des attentes de celui-ci devrait être fournie préalablement au système. Comme évoqué en 1.2, cette contrainte est fortement problématique dans le domaine artistique, où la notion de qualité peut varier énormément selon le public et le contexte.

### 2.3.2. Intégration d'un instructeur humain : spécificités du façonnage interactif

Les approches de façonnage interactif [14] entendent contourner cette difficulté. À la différence des méthodes « classiques » d'apprentissage par renforcement, elles modélisent explicitement la nature humaine de l'instructeur et considèrent des récompenses attribuées de manière interactive. Elles sont ainsi particulièrement adaptées aux cas où une spécification explicite et *a priori* des attentes de l'utilisateur par rapport au système serait difficile, voire impossible (si l'on s'attend par exemple à ce que des utilisateurs différents formulent des attentes différentes).

La prise en compte de la dimension humaine de l'instructeur est à l'origine d'une autre spécificité des méthodes de façonnage interactif par rapport aux approches d'apprentissage par renforcement. Si la majeure partie de ces dernières calculent à tout moment une prédiction des récompenses attendues sur un horizon de temps futur, la plupart des méthodes de façonnage interactif sont « myopes » [14] : elles tiennent peu ou aucunement compte d'une prédiction de l'impact futur des récompenses attribuées. Lorsque l'instructeur humain attribue une récompense, le modèle compare cette récompense à celle qu'il attendait pour l'instant actuel uniquement. Les



**Figure 3:** Relations du façonnage interactif à d'autres champs de recherche.

méthodes de façonnage interactif entendent ainsi trouver un compromis entre leurs capacités de prévision et leur réactivité.

Ces deux spécificités rendent les approches de façonnage interactif particulièrement intéressantes au regard des exigences des contextes artistiques.

### 2.3.3. Résumé : intérêt du façonnage interactif pour l'apprentissage du couplage geste-son

La figure 3 propose une comparaison avec d'autres approches d'apprentissage automatique, comme l'apprentissage supervisé [16] et l'apprentissage par renforcement. Les approches de façonnage interactif y sont représentées comme impliquant le plus l'utilisateur, et requérant le moins d'exemples (tant *a priori* que collectés pendant l'interaction) pour leur apprentissage. Ces spécificités du façonnage interactif nous semblent particulièrement intéressantes dans le contexte de l'apprentissage du couplage geste-son, détaillé en 3. Nous verrons en particulier en 4.1 comment l'intégration explicite des réactions d'appréciation positive ou négative de l'utilisateur permettent de formuler de manière pertinente le problème de l'apprentissage de structures de haut niveau, difficile à poser dans ce cadre.

## 3. MODÉLISATION DU COUPLAGE GESTE-SON

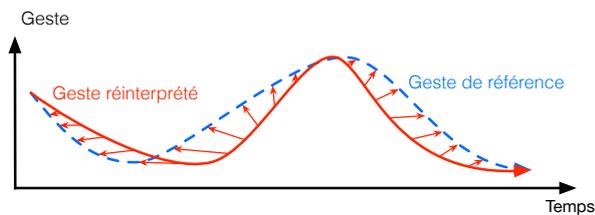
Cette partie vise à synthétiser des développements récents menés dans l'équipe IMTR dans le domaine de la modélisation statistique de mappings liant geste et son. Nous nous plaçons donc dans un cadre de « mapping par démonstration ». Cette situation implique la réalisation simultanée à l'écoute d'une performance gestuelle, qui rend prépondérante la dynamique temporelle du son. Nous nous intéressons donc ici à chaque exemple — sonore ou

gestuel — en tant que *morphologie temporelle* de courte durée (typiquement inférieure à quelques secondes). Dans la suite, nous désignons par « segment geste–son » la réunion de deux profils temporels gestuel et sonore.

Nous décrivons dans un premier temps une méthode d'apprentissage du couplage temporel entre geste et son, puis présentons une extension hiérarchique récente.

### 3.1. Modèles statistiques de segments geste–son

Dans *Gesture Follower* (voir 1.1.2), le système modélise un geste d'exemple par un modèle de Markov caché (*hidden Markov model* ou HMM) de topologie gauche-droite, qui permet d'encoder sa temporalité de façon continue et flexible. Un mapping temporel est alors formulé en alignant les morphologies gestuelle et sonore [2]. Le geste peut ensuite être rejoué en faisant varier expressivement son déroulement temporel. À chaque instant, un calcul d'alignement permet d'estimer l'avancement temporel à l'intérieur du geste, ce qui permet une resynthèse adaptative du son associé par des opérations d'étirement ou de compression temporelle basées sur des techniques de vocodeur de phase (cf. fig. 4).



**Figure 4:** Fonctionnement de *Gesture Follower*. Le geste de référence (bleu pointillé) est modélisé par un modèle de markov caché. En temps réel, une nouvelle interprétation du geste (rouge plein) peut être alignée sur la référence afin de rejouer dynamiquement le son associé.

### 3.2. Du segment à la séquence : extension hiérarchique

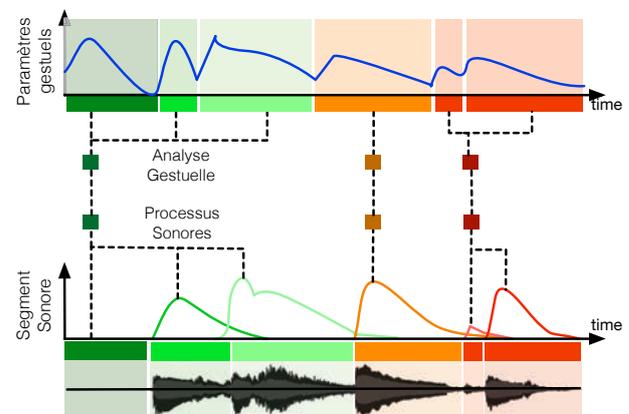
#### 3.2.1. Motivation

Le modèle présenté dans la section précédente fournit une modélisation fine de la structure temporelle de segments geste–son de courte durée. Cependant, afin d'élargir les possibilités d'interaction, ces segments nécessitent d'être combinés pour former des séquences plus complexes. Les séquences musicales peuvent souvent être décomposées de manière hiérarchique d'un point de vue temporel, par exemple en phases d'attaque/sustain, notes, phrase, etc. Il semble donc cohérent d'appliquer une décomposition similaire au couplage geste–son.

#### 3.2.2. Modélisation hiérarchique

Nous avons récemment proposé une approche hiérarchique pour la création de mappings temporels [8], représentée sur la figure 5. L'approche consiste à appliquer au

geste une segmentation multi-échelles, afin de structurer hiérarchiquement sa représentation temporelle (1). Cette structure est alors mise en relation avec une représentation analogue des processus sonores (2), permettant le développement de stratégies d'interaction complexes (3).



**Figure 5:** Représentation hiérarchique des relations geste–son. Chaque modalité est structurée hiérarchiquement par une segmentation multi-échelles, qui permet la création de mappings selon différentes échelles temporelles.

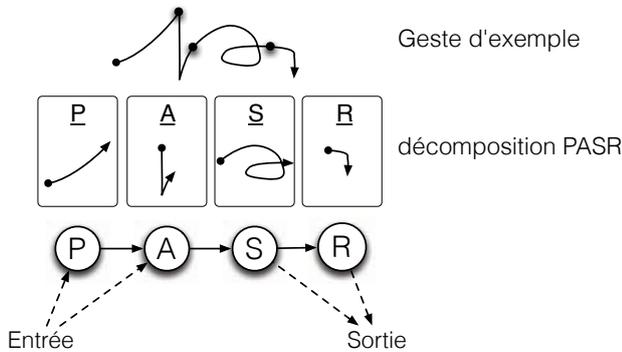
Formellement, le modèle introduit dans la section précédente peut être étendu à une structure multi-échelles, sous la forme d'un modèle de Markov caché hiérarchique [7]. Les segments geste–son appris sont alors liés les uns aux autres par une structure de haut niveau, dont les probabilités de transition peuvent être fixées par l'utilisateur.

#### 3.2.3. PASR : Préparation — Attaque — Sustain — Relâchement

Afin d'illustrer l'intérêt de telles structures multi-échelles, nous avons présenté dans des travaux précédents une décomposition gestuelle spécifique appliquée au contrôle de la synthèse sonore [8]. Nous reprenons ici cette représentation appelée PASR (par analogie avec la représentation sonore ADSR) qui décompose chaque geste en quatre phases. Une phase de préparation (**P**) est associée au geste d'anticipation précédant le début du son. Suivent des phases d'attaque (**A**) et sustain (**S**). Enfin, une phase de relâchement (**R**) décrit le geste de rétraction accompagnant la fin du son. La figure 6 détaille la segmentation d'un geste selon cette représentation et présente la topologie du modèle associé. Les flèches indiquent les transitions possibles : un utilisateur peut donc entrer dans un geste par les phases de préparation ou d'attaque. Le parcours se poursuit en suivant l'axe temporel, et le geste peut être achevé par une phase de sustain ou de relâchement.

Cette représentation offre plusieurs avantages. D'une part, elle rend possibles différents modes de jeu : les gestes peuvent être entièrement rejoués ou rapidement enchainés par des transitions des phases de sustain aux phases d'attaque. D'autre part, l'apport d'une structure de haut ni-

veau permet une précision accrue de la segmentation du geste en termes de reconnaissance et de précision temporelle. Enfin, elle permet l'imposition de contraintes sur certaines phases du son lors de la synthèse. dans une situation typique où le son est réinterprété temporellement en utilisant un vocodeur de phase (alignement du son à la performance gestuelle), il est possible d'imposer une conservation des transitoires sur les phases d'attaque, ce qui garantit une meilleure cohérence du rendu sonore.

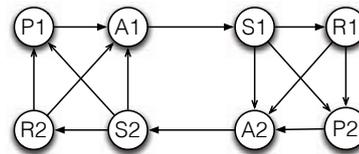


**Figure 6:** Décomposition PASR d'un geste. Les flèches entre les états **P**, **A**, **S** et **R** indiquent les possibilités de transition au sein du modèle.

### 3.2.4. Une question ouverte : la co-articulation

Dans la représentation hiérarchique actuelle, un geste est modélisé par une séquences de segments, une représentation qui semble pertinente d'un point de vue cognitif [11]. Cependant, le modèle est limité par deux aspects. Sa construction est soumise à une segmentation manuelle de l'utilisateur, qui ne peut définir des segments ou modèles de séquences que manuellement. D'autre part, le modèle ne laisse que peu d'espace à d'éventuelles variations à l'interface entre deux segments successifs. Ainsi, les phénomènes de *co-articulation* pouvant survenir entre des segments « significatifs » pendant la performance d'une nouvelle séquence seront difficilement caractérisés par le modèle actuel.

Considérons en effet le cas d'un système interactif composé de deux exemples de mapping geste-son, modélisés par deux représentations PASR. Dans la version actuelle du modèle, la seule « articulation » que l'on puisse concevoir entre ces deux représentations est détaillée dans la figure 7. Les états de fin (**S1**, **R1**) de la première représentation PASR1 sont connectés aux états initiaux (**P2**, **A2**) de la seconde représentation PASR2. Il est donc uniquement possible de réaliser ces deux gestes « à la suite », c'est-à-dire de naviguer de la fin d'un geste (au sortir de la phase de sustain **S1** ou de relâchement **R1**) vers le début de l'autre (entrée dans la phase de préparation **P2** ou d'attaque **A2**). Dans l'optique de dépasser de ces contraintes d'articulation, nous détaillons dans la section suivante un programme prospectif d'apprentissage de phases de *co-articulation* entre différents modèles PASR. Il s'agit de développer un système capable d'interpréter les gestes



**Figure 7:** Topologie d'une structure comprenant deux modèles de couplage PASR.

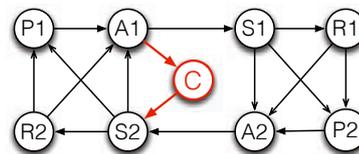
d'un utilisateur souhaitant augmenter le système avec des mouvements de co-articulation, en se basant uniquement sur ses gestes effectués durant l'interaction.

## 4. PROSPECTIVE : APPRENTISSAGE DU COUPLAGE GESTE-SON PAR FAÇONNAGE INTERACTIF

### 4.1. Problématique

Nous nous proposons d'exploiter le paradigme du façonnage interactif exposé dans la section 2.3 afin d'apprendre incrémentalement des modalités de co-articulation entre plusieurs représentations hiérarchiques de type PASR, telles qu'introduites en section 3.2. Cette approche consiste à interconnecter des modèles PASR appris pour différents exemples de couplage geste-son. Dans la suite de notre explication, nous considérons une situation d'exemple composée de deux modèles PASR. La topologie de la structure de haut niveau réunissant ces deux modèles est initialement la même que celle présentée dans la figure 7.

Comme expliqué en 3.2.4, ces possibilités de transition correspondent simplement à l'enchaînement de plusieurs actions distinctes. Nous nous proposons ici de laisser la possibilité à l'utilisateur d'étendre cette structure de base de manière incrémentale et interactive, en ajoutant une dimension de *co-articulation* entre paires de modèles PASR. Formellement, l'enjeu est de construire un « pont » supplémentaires entre deux modèles par l'ajout d'un état **C** connectant, par exemple, l'état **A1** du premier modèle à l'état **S2** du second (figure 8). L'état **C** correspond à une phase de haut niveau, au même titre que les phases de type **P**, **A**, **S** et **R**. Il génère un segment modélisant la co-articulation entre les phases **A1** et **S2**.



**Figure 8:** Ajout d'un état de *co-articulation* dans la topologie.

Cette inférence d'une nouvelle configuration d'une topologie de haut niveau relève d'une problématique d'*apprentissage de structure de modèles statistiques*. Il s'agit d'un domaine actif de l'apprentissage automatique [16] :

usuellement traité dans un cadre classique d'apprentissage supervisé où de nombreux exemples du matériau à modéliser sont disponibles, il constitue alors un problème extrêmement difficile. Nous envisageons ici de tirer parti de la dimension interactive intrinsèque à nos contextes applicatifs pour proposer une formulation de ce problème en termes d'*apprentissage interactif* : cette approche nous semble poser le problème de manière plus cohérente avec le cadre d'étude des systèmes interactifs musicaux.

#### 4.2. Proposition d'approche pour l'apprentissage incrémental et interactif de l'état de co-articulation C

Nous souhaitons permettre à l'utilisateur de « forcer » l'apprentissage d'une telle co-articulation par le seul moyen de l'interaction gestuelle. La difficulté de cette approche est double :

1. Il s'agit d'abord de détecter les gestes réalisés par l'utilisateur qui imposent effectivement au modèle une transition entre **A1** et **S2**. Ces gestes doivent être distingués efficacement de possibles variations interprétatives, s'éloignant du geste modélisé par la structure PASR de départ sans pour autant être investies d'une intention de modification de cette structure (i.e., l'utilisateur souhaite pouvoir rejouer différemment le mapping de départ, sans souhaiter toutefois articuler deux mappings distincts).
2. Une fois la phase de co-articulation détectée, celle-ci doit être apprise de manière incrémentale, à mesure que l'utilisateur réalise de nouveaux gestes dans lesquels elle apparaît (fournissant ainsi de nouveaux exemples d'apprentissages). Le système réalise en ce sens une adaptation de sa représentation interne du couplage (telle qu'évoquée en 1.1.1).

##### 4.2.1. Formalisation

Notre formulation entend traiter ces deux étapes simultanément, dans un cadre d'apprentissage par façonnement. Pour ce faire, une fois les différents modèles PASR appris séparément, nous initialisons un tableau de probabilités de transitions potentielles, entre les états du modèle PASR1 et l'état C, et entre l'état C et le modèle PASR2. Ces probabilités sont au départ égales à 0 ; de même, l'état C est initialement purement hypothétique, puisqu'aucun exemple de co-articulation n'a encore été fourni par l'utilisateur.

Durant la phase d'interaction, le système détecte les « erreurs », c'est à dire les transitions théoriquement « interdites » qui peuvent être empruntées sous la contrainte des mouvements de l'utilisateur. La détection de telles erreurs est un bon indicateur de l'introduction d'un mouvement de co-articulation entre deux segments.

Supposons qu'une co-articulation entre **A1** et **S2** soit ainsi détectée. On considère alors que l'utilisateur *récompense* la transition de **A1** vers C. La probabilité de cette

transition, initialement nulle, est mise à jour pour intégrer cette récompense. Parallèlement, la phase identifiée comme relevant de la co-articulation dans le geste de l'utilisateur permet d'apprendre un premier modèle du segment C. L'état C est pour l'instant construit « à l'écart » : il n'est pas encore considéré comme une co-articulation effective entre les modèles PASR, mais comme une co-articulation *potentielle*.

À mesure que l'utilisateur continue de récompenser cette co-articulation par son interaction avec le système, l'algorithme de façonnage interactif donne de plus en plus de poids au modèle « potentiel » intégrant l'état de co-articulation C. Lorsque ce poids atteint une valeur suffisante, l'état C est effectivement intégré comme lien de co-articulation entre les modèles PASR1 et PASR2. La modalité sonore correspondante est alors évaluée à partir de **A1** et **S2**, à l'aide d'un algorithme d'interpolation audio <sup>9</sup>.

##### 4.2.2. Implémentation

Nous disposons actuellement d'une implémentation des modèles de Markov hiérarchiques, sous la forme d'un objet Max/MSP. Celui-ci permet l'enregistrement de profils gestuels liés à des segments sonores, et intègre les algorithmes temps-réel de suivi. Le modèle peut donc être utilisé pour le contrôle sonore, par exemple en utilisant des techniques de vocodeur de phase. Nous développons actuellement en parallèle les deux composantes de la dimension d'apprentissage interactif : la méthode de détection de transitions relevant d'un aspect de co-articulation, ainsi que l'algorithme de façonnage.

#### 4.3. Perspectives

Les modèles actuels possèdent des limitations quand à l'estimation de la dimension sonore. En particulier, la modélisation de type suivi de geste fournit une représentation statistique du geste mais non du son. Ceci peut limiter les possibilités d'apprentissage des variations du mapping apparues sur différentes performances.

Pour dépasser cette limitation, nous étudions actuellement une extension multimodale des modèles de Markov cachés pour l'apprentissage du couplage geste-son. Cette approche s'inspire de travaux réalisés en animation d'avatars [10] et en inversion acoustique-articulatoire [12]. Cette fois, les processus gestuels et sonores sont modélisés conjointement. Un algorithme d'apprentissage multi-exemples permet d'entraîner le modèle sur plusieurs variations d'un même segment geste-son. Ce modèle multimodal peut ensuite être inversé pour le contrôle sonore : une nouvelle performance gestuelle permet la synthèse des paramètres sonores en temps réel. De par la modélisation statistique conjointe, l'estimation des paramètres de contrôle sonore est réalisée de manière probabiliste, ce qui induit une meilleure fluidité de la synthèse. Ce modèle

<sup>9</sup> . Rappelons que dans le cadre présenté dans cette section, seule la dimension gestuelle est modélisée de manière probabiliste. L'estimation de la modalité sonore constitue une question ouverte, qui fera l'objet d'un travail ultérieur (esquissé en 4.3).

fournit une perspective intéressante pour l'estimation de l'aspect sonore de segments de co-articulation gestuelle. D'une part, la capacité d'apprentissage multi-exemples permet une meilleure approximation des mouvements de co-articulation gestuelle. D'autre part, le modèle pourrait utiliser le contexte de la co-articulation gestuelle comme premier estimateur de la co-articulation sonore correspondante. Par apprentissage incrémental, ce modèle pourrait ensuite être affiné par les exemples successifs fournis par l'utilisateur.

## 5. CONCLUSION

Cet article se place dans la continuité de nos recherches sur l'interaction gestuelle avec des processus sonores. Nous considérons spécifiquement le problème de l'apprentissage de mappings entre geste et son présentant une structure temporelle hiérarchique. Si les modèles actuels permettent de modéliser le mapping par un ensemble de segments associant geste et son, leurs relations dans une structure de haut niveau doivent pour l'instant être fixées *a priori*. Au regard des récents développements dans le champ de l'apprentissage interactif, nous présentons une perspective pour l'apprentissage de telles structures de haut niveau. Nous détaillons en particulier une application visant à apprendre et à intégrer incrémentalement des gestes de co-articulation au sein d'une structure de transition prédéfinie.

## 6. REMERCIEMENTS

Ces recherches sont soutenues par le projet ANR LEGOS (11 BS02 012). Nous remercions tous les membres de l'équipe IMTR.

## 7. REFERENCES

- [1] Bevilacqua, F., Zamborlin, B., Sypniewski, A., Schnell, N., Guédy, F., et Rasamimanana, N. "Continuous realtime gesture following and recognition", *Gesture in Embodied Communication and Human-Computer Interaction*, 73–84, 2010.
- [2] Bevilacqua, F., Schnell, N., Rasamimanana, N., Zamborlin, B., et Guédy, F. "Online Gesture Analysis and Control of Audio Processing", *Musical Robots and Interactive Multimodal Systems*, 127–142. Springer, 2011.
- [3] Caramiaux, B. *Etudes sur la relation geste-son en performance musicale*, Thèse de doctorat, Ircam – Université Pierre et Marie Curie, 2012.
- [4] Cohn, D., Atlas, L., et Ladner, R. "Improving generalization with active learning", *Machine Learning*, 15(2), 201–221, 1994.
- [5] Cont, A., Dubnov, S., et Assayag, G. "Anticipatory model of musical style imitation using collaborative and competitive reinforcement learning", *Anticipatory Behavior in Adaptive Learning Systems*, Springer, New York, 2007.
- [6] Fiebrink, R., Cook, P. R., et Trueman, D. "Play-along mapping of musical controllers", *Proceedings of the International Computer Music Conference (ICMC)*, 2009.
- [7] François, J. *Realtime Segmentation and Recognition of Gestures using Hierarchical Markov Models*, Mémoire de Master, Université Pierre et Marie Curie – Ircam, 2011.
- [8] François, J., Caramiaux, B., et Bevilacqua, F. "A Hierarchical Approach for the Design of Gesture-to-Sound Mappings", *Proceedings of the 9th Sound and Music Conference (SMC)*, Copenhagen, Danemark, 2012.
- [9] Franklin, J. A., et Manfredi, V. U. "Nonlinear Credit Assignment for Musical Sequences", *Second International Workshop on Intelligent Systems Design and Applications (ISDA)*, Atlanta, USA, 2002.
- [10] Fu, S., Gutierrez-Osuna, R., Esposito, A., Kaku-manu, P. K., et Garcia, O. N. "Audio/visual mapping with cross-modal hidden Markov models", *Multimedia, IEEE Transactions on*, 7(2), 243–252, 2005.
- [11] Godøy, R. I., Jensenius, A. R., et Nymoen, K. "Chunking in music by co-articulation." *Acta Acustica united with Acustica*, 96(4), 690–700, 2010.
- [12] Hofer, G. *Speech-driven animation using multi-modal hidden Markov models*. Thèse de doctorat, University of Edinburgh, 2009.
- [13] Hunt, A. et Kirk, R. "Mapping Strategies for Musical Performance." *Trends in Gestural Control of Music*, 231–258, 2000.
- [14] Knox, W. B. *Learning from Human-Generated Reward*, Thèse de doctorat, Austin, TX, 2012.
- [15] Le Groux, S., et Verschure, P. F. "Towards Adaptive Music Generation by Reinforcement", *Proceedings of the 7th Sound and Music Conference (SMC)*, Barcelona, Spain, 2010.
- [16] Murphy, K. P. *Machine Learning : a Probabilistic Perspective*, MIT Press, Cambridge, MA, 2012.
- [17] Puterman, Martin L. *Markov Decision Processes : Discrete Stochastic Dynamic Programming*, John Wiley & Sons, Inc., Hoboken, NJ, 1994.
- [18] Rasamimanana, N., Bevilacqua, F., Schnell, N., Fléty, E. et Zamborlin, B. "Modular Musical Objects Towards Embodied Control Of Digital Music Real Time Musical Interactions", *Proceedings of the fifth international conference on Tangible, embedded, and embodied interaction*, 2011.
- [19] Sutton, R. et Barto, A. *Reinforcement Learning : An Introduction*, MIT Press, Cambridge, MA, 1998.
- [20] Wanderley, Marcelo M. et Battier, Marc. *Trends in Gestural Control of Music*, Ircam - Centre Pompidou, 2000.

# POUR UNE ECRITURE MULTIMEDIA INTERACTIVE DANS LA COMPOSITION MUSICALE

Sabina Covarrubias Acosta  
Université Paris 8  
Vincennes Saint-Denis  
sabinacovarrubias@gmail.com

## RÉSUMÉ

L'objectif principal de ce texte vise à formaliser la notion d'*écriture multimédia interactive* (EMI), dans le domaine de la composition musicale.

Dans un premier temps, nous exposerons la nécessité d'utiliser l'EMI pour noter des éléments composant une œuvre musicale qui ne sont pas susceptibles d'être saisis au moyen du système occidental d'écriture musicale.<sup>1</sup>

Dans un deuxième temps, nous montrerons les avantages de l'utilisation de ce type d'écriture dans l'ethnomusicologie, et comment les mêmes techniques d'écriture peuvent être transposées au domaine de la composition.

Ensuite, nous proposerons quatre fonctions de l'EMI établies selon les caractéristiques de l'objet représenté et selon le processus de création et de mise en scène de l'œuvre.

Finalement, nous proposerons la création d'un support d'écriture pour mettre en œuvre l'EMI : une version de *partition multimédia interactive*.

## 1. INTRODUCTION

Dans le domaine de la création des nouvelles œuvres musicales, nous constatons que le *système occidental de notation musicale* est limité au moment de saisir et de représenter de manière précise et efficace plusieurs éléments dont nous avons besoin dans nos compositions. Il peut s'agir soit d'éléments de musiques de tradition orale, soit de divers moyens de production sonore.

En accord avec Scaldaferrì [11] et Will [13] à propos de la notation des éléments extraits de musiques de tradition orale, nous constatons que beaucoup de problèmes apparaissent au moment de noter une musique qui est caractérisée par l'absence d'écriture.

Le polymorphisme est l'un de ces aspects. Dans la musique de tradition orale, chaque version d'une même musique peut être différente. Or, tout ce qui est sujet à variations en cours de performance ne peut être fixé par l'écriture musicale. En outre, on ne parvient pas à noter ou transcrire de manière précise les caractéristiques du timbre vocal sur une partition, ni les spécifications de mouvements corporels, ni les techniques d'exécution

<sup>1</sup> Nous reprenons à Philippe Michel cette désignation de la musique "classique" dans ses extensions historiques antérieures et contemporaines.[9]

instrumentale et vocale qui rendent possible la production du son.

En ce qui concerne les nouveaux moyens de production sonore, nous nous confrontons à la même problématique. Nous expérimentons des difficultés pour noter avec précision des trajectoires<sup>2</sup> et d'autres sons produits au moyen de nouvelles techniques d'exécution instrumentale et vocale.<sup>3</sup>

D'autre part, même si nous avons la possibilité de créer de nouveaux graphismes pour saisir ces aspects, et faire usage des textes et des images descriptifs qui accompagnent la partition, cette solution ne nous convainc pas. En effet, ces graphismes ne peuvent décrire de manière précise les aspects qu'il nous intéresse de noter. Ces textes restent encore imprécis et leur compréhension demanderait trop de temps.

Ce sont malgré tout des aspects que nous voulons inclure dans nos compositions. De là, nous posons la question suivante : *Dans le cadre des nouvelles compositions, comment sauvegarder des éléments qui ne sont pas susceptibles d'être saisis précisément au moyen du système occidental d'écriture musicale puis transmis rapidement à l'interprète?* Pour apporter une solution à ce problème nous proposons l'utilisation de l'EMI dans la composition musicale. Cette écriture (qui peut être de type descriptif ou prescriptif)<sup>4</sup>, est mise en œuvre au moyen d'un nouveau support d'écriture, à savoir une version de *partition multimédia interactive*.

## 2. ANTECEDENTS DE L'UTILISATION DE L'ECRITURE MULTIMEDIA

### 2.1 Utilisation des ethnomusicologues

Dans les modélisations<sup>5</sup> animées créées par les ethnomusicologues au moyen de l'écriture multimédia, telles que *Les Clés d'écoute*<sup>6</sup>, certains des aspects représentés nous intéressent :

<sup>2</sup> Par exemple, les déplacements des musiciens sur scène ou des mouvements spécifiques d'un instrument ou d'un contrôleur de son.

<sup>3</sup> La recréation de nouveaux sons au moyen des nouvelles techniques d'exécution vocale, créées par nous-mêmes, reste imprécise quand nous nous limitons à utiliser uniquement la partition et les instructions sur papier.

<sup>4</sup> Pour plus d'informations voir la section 4.2 de ce document.

<sup>5</sup> Nous reprenons à Simha Arom la notion de modèle. [3]

<sup>6</sup> Les *clés d'écoute* sont des animations musicales interactives. Elles s'appuient sur la notion d'« oreille culturelle » ou « oreille locale » (termes introduits par Bernard Lortat-Jacob). Cela consiste à mettre en évidence une manière d'écouter, de faire et de penser la musique

### 2.1.1 Des trajectoires dans l'espace.

Le travail de Dana Rappoport, *L'antiphonie dans une ronde funéraire toraja (Indonésie)*<sup>7</sup>, est un modèle animé qui révèle un principe d'antiphonie vocale. Il fait usage de trajectoires animées pour faire comprendre un partage possible des syllabes dans l'espace.

### 2.1.2 Timbre et technique d'exécution vocale.

La clé d'écoute consacrée au chant diphonique<sup>8</sup> montre l'aspect physiologique de la technique vocale de ce chant. Ce film fait usage de radiographies animées pour mettre en évidence le rôle de la langue dans la division de la cavité buccale en deux parties, ainsi que la sélection des harmoniques à chanter par un déplacement entre l'avant et l'arrière de la langue. L'animation des radiographies est synchronisée avec le son du chant et une transcription animée afin de révéler la correspondance entre la position de la langue et le son produit.

### 2.1.3 Fonctionnement d'un système musical de tradition orale.

La compréhension globale du fonctionnement de la variation dans un système musical modulaire est rendue possible dans le cédérom réalisé par l'équipe de Simha Arom sur la musique des Pygmées Aka. [4]

## 2.2 Avantages de l'écriture multimédia dans l'ethnomusicologie.

Parmi les avantages techniques qu'offre l'écriture multimédia dans l'ethnomusicologie, Chemillier [5] remarque le soulignement de certaines parties de la transcription, la sonorisation de l'image en synchronisation avec le son, l'animation d'éléments de l'image et l'interactivité avec le récepteur.

Il constate également que « l'adoption des nouvelles formes d'écriture facilite la compréhension et permet une lecture plus rapide » [5]

---

culturellement déterminée. L'enjeu est d'imaginer des modes de représentation des musiques adaptés aux paramètres culturellement pertinents (sonagrammes, transcriptions, etc.), et de tirer parti du concept d'interactivité. [7]

<sup>7</sup> Modèle créé par Dana Rappoport. Disponible sur <http://ehess.modelisationsavoirs.fr/ethnomus/index.html> Clé No2

<sup>8</sup> Travail de Tran Quang Hai. Disponible sur <http://ehess.modelisationsavoirs.fr/ethnomus/index.html> Clé No 5

## 3. POUR UNE ECRITURE MULTIMEDIA INTERACTIVE DANS LA COMPOSITION MUSICALE

### 3.1. La nécessité d'adopter l'EMI dans la composition musicale.

L'utilisation des nouveaux moyens d'écriture peut ouvrir de nouveaux horizons quant à la transformation substantielle du langage musical. En effet, les limites de l'écriture musicale peuvent conditionner le processus de composition et par conséquent le langage musical. En d'autres termes, l'écriture peut conditionner les caractéristiques du langage musical. Les études anthropologiques des dernières décennies dans ce domaine ont montré que l'écriture n'est pas une activité neutre et mécanique mais qu'elle implique des parcours cognitifs et comportementaux précis, ainsi que des technologies particulières de communication et de transmission. Une des conséquences les plus importantes de ce fait est que l'écriture « exerce une action en retour sur les données qu'elle entendait symboliser, c'est-à-dire sur le langage lui-même » Goody [8]. Nous trouvons un exemple de l'écriture influençant le langage dans la musique avec le passage d'une conception horizontale à une conception verticale au XVI<sup>e</sup> siècle. En effet, c'est à cette époque que l'écriture joue un rôle décisif dans le passage à une pensée par accords et à tonalité moderne. C'est là qu'apparaît l'usage systématique de la partition au sens moderne. Molino [10]. La transition à une conception verticale de la musique n'aurait pas été possible sans l'écriture. Ainsi, on constate la manière dont l'écriture a influencé les caractéristiques du langage. En remarquant le lien entre écriture, processus de composition et langage musical, il est logique de penser que si les processus de composition et l'écriture sont inséparables, une transformation des technologies d'écriture et des moyens d'écriture aurait aussi une conséquence sur le langage musical; cela implique forcément un changement dans les processus de composition et la pensée musicale du compositeur. En outre, des aspects multimédia pourraient être utiles pour la transmission des aspects musicaux jamais traités auparavant dans la composition et, en conséquence, favoriseraient un changement dans la pensée musicale du compositeur lors de la création.

### 3.2 L'écriture multimédia, de l'ethnomusicologie au domaine de la composition.

Nous proposons de transposer l'écriture multimédia utilisée par les ethnomusicologues aux musiques nouvelles. En effet, nous soutenons que si cette écriture a fonctionné pour l'ethnomusicologie, elle fonctionnera aussi pour la composition musicale. Ce sont les mêmes éléments que nous voulons représenter: la description et la sauvegarde des trajectoires, des timbres, des techniques d'exécution et la description du fonctionnement d'un système musical.

En nous appuyant sur la proposition de Chemillier «*Pour une écriture multimédia de l'ethnomusicologie*», [5] nous proposons une *écriture multimédia interactive dans la composition*. Par conséquent, nous proposons l'utilisation d'une partition *étendue*. Toujours en accord avec ce même auteur : "Il s'agit non pas de remplacer les partitions mais plutôt "d'adopter de nouvelles formes d'écriture qui facilitent la compréhension et qui permettent une lecture plus rapide". [5] A cela, nous ajoutons : qui permettent au compositeur d'écrire les éléments qui ne sont pas susceptibles d'être saisis par l'écriture musicale.

## 4. LES FONCTIONS DE L'EMI

### 4.1 La fonction prescriptive et descriptive

A la lumière de notre étude, la distinction proposée par Charles Seeger entre la fonction descriptive et la fonction prescriptive de la notation musicale [12] peut être transposée à l'écriture multimédia interactive. Si les éléments du support d'écriture mettent en valeur les instructions à suivre, il s'agit d'une *écriture multimédia interactive du type prescriptif*. En revanche, si les éléments du support d'écriture permettent de donner une description pour faire comprendre le fonctionnement d'un événement<sup>9</sup>, il s'agit d'une *écriture multimédia interactive du type descriptif*.

Par exemple, l'écriture multimédia prescriptive pourrait montrer comment jouer une pièce de musique, en donnant des instructions qui favorisent l'imitation. C'est le cas de la série de cours par vidéo réalisée par David Courney, qui enseigne comment jouer des instruments indiens<sup>10</sup>. Ici, le spectateur n'obtient pas l'information nécessaire qui lui permettra de développer ses propres versions de la musique. En ignorant le fonctionnement d'un système, l'élève se limite à suivre une prescription (des instructions).

Dans le cas de la fonction descriptive, nous prendrons à nouveau comme exemple le CD ROM développé par l'équipe de Simha Arom [4] sur la musique des pygmées Aka. La partie consacrée à l'analyse musicale montre les représentations animées des structures musicales. Au moyen des applications interactives l'utilisateur peut comprendre le fonctionnement d'une musique constituée par modules. Ainsi, il peut effectuer plusieurs versions d'une même musique en changeant de lieu les modules qui la constituent.

### 4.2 Les deux fonctions de l'EMI lors de la mise en scène.

Nous remarquons deux fonctions de l'EMI lors de la mise en scène d'une œuvre que nous appellerons la fonction *auxiliaire* et la fonction *substantielle*.

<sup>9</sup> Par exemple, le fonctionnement d'un système, technique d'exécution, des mouvements corporels ou sur scène, etc.

<sup>10</sup> Diffusés sur le site : <http://chandrankantha.com> et aussi sur <http://www.youtube.com/watch?v=L3ehdj0B-QA>

#### 4.2.1 La fonction auxiliaire.

Dans cette catégorie, l'EMI n'est pas indispensable pour compléter la création de l'œuvre, celle-ci pourrait être écrite en sa totalité au moyen du système occidental de notation musicale. Pourtant, l'utilisation de l'EMI simplifiera la lecture, rendant ainsi plus rapide le processus de mise en scène. En effet, en accord avec Rappoport et Chemillier, nous observons que l'adoption des nouvelles formes d'écriture, telle que l'écriture multimédia, facilite la compréhension du texte et permet une lecture plus rapide. [6] Il s'agit donc d'une alternative en contraste avec les explications et descriptions données sur papier au début d'une œuvre. Ces explications sont données sur une ou plusieurs pages, sous la forme de textes en paroles ou en schémas et requièrent parfois une quantité considérable de temps pour être comprises. L'EMI peut alors compléter l'écriture musicale de tradition occidentale. Parmi les aspects qui nous intéressent, cette utilisation de l'EMI est auxiliaire pour :

- Enseigner aux interprètes différents types de techniques d'exécution instrumentale ou vocale, qu'elles soient nouvelles ou tirées de musiques de tradition orale.
- Faire comprendre aux interprètes les règles d'un système d'improvisation, de variation ou d'ornementation.
- Favoriser la précision lors d'une synchronisation en temps réel entre un événement visuel (vidéo, théâtre) ou sonore préalablement réalisé et une exécution en temps réel.
- Faire comprendre aux interprètes des trajectoires de mouvement lors de l'exécution de l'œuvre: Déplacements sur scène, mouvements corporels, et mouvements des contrôleurs du son.
- Guider la participation de musiciens de tradition orale dans le contexte d'une œuvre de musique nouvelle différente de leur tradition musicale: Le cas où le musicien de tradition orale joue ou chante selon les règles de son système musical ou, le cas où le musicien exécute une musique différente de celle de sa tradition.

#### 4.2.2 La fonction substantielle

Dans cette catégorie, l'utilisation de l'EMI est indispensable tant pour la création de l'œuvre que pour sa mise en scène.

Parmi les possibilités qu'offre cette utilisation de l'EMI nous voulons souligner :

- La création de versions différentes d'une œuvre polymorphe à chaque utilisation de la partition interactive.<sup>11</sup>

<sup>11</sup> C'est le cas de la partition interactive proposée par Allombert et al.[1]

- La direction précise de la performance des musiciens placés à grande distance les uns des autres

## 5. CONCLUSIONS

La formalisation de la notion d'écriture multimédia interactive dans le domaine de la composition musicale peut donner lieu à l'appropriation de celle-ci par les compositeurs qui visent à noter des éléments qui ne sont pas susceptibles d'être saisis au moyen de l'écriture musicale de tradition occidentale.

En outre, pour que l'EMI puisse être mise en œuvre, il reste à développer un support d'écriture multimédia<sup>12</sup> dont les fonctions principales sont :

- Transmettre aux interprètes des aspects non susceptibles d'être notés par le système occidental d'écriture musicale.
- Permettre une lecture rapide de l'œuvre.
- Faire comprendre le fonctionnement de différents systèmes d'improvisation, de variation ou d'ornementation, qu'ils soient nouveaux ou provenant de musiques de tradition orale.
- Simplifier ou rendre possible l'apprentissage des nouvelles techniques d'exécution instrumentale et vocale ainsi que des techniques provenant de la musique de tradition orale.
- Décrire des trajectoires spécifiques dans l'espace.
- Représenter des mouvements corporels qui rendront possible la production du son, qu'ils soient nouveaux ou provenant de musiques de tradition orale.
- Rendre possible la synchronisation précise d'un événement visuel préalablement réalisé et d'une exécution en temps réel, où l'interprète se trouve à une grande distance de l'événement visuel.
- La création d'œuvres polymorphes à chaque interaction de l'utilisateur.

## 6. REFERENCES

- [1] ALLOMBERT, A., DESAINTE-CATHERINE, M., « Interactive scores: A model for specifying temporal relations between interactive and static events », *Journal of New Music Research*, Décembre 2005, vol. 34, n° 4, p. 361-374

<sup>12</sup> Dans le passé, des partitions multimédia ont été proposées. Leurs caractéristiques spécifiques ne s'ajustent pas à nos besoins. C'est la raison qui nous pousse à développer notre version de *partition multimédia interactive*. Parmi les partitions multimédia qui ont été développées nous trouvons: *Interlude*, développé par l'équipe de D. Foer, *iAnalyse* développé par P. Couprie et *Iscore* développé par A.Allombert, M. Desainte-Catherine et G.Assayag. Pour plus d'information sur ces projets voir : <http://interlude.ircam.fr/>, <http://logiciels.pierrecouprie.fr/>, A. Allombert, M. Desainte-Catherine, G. Assayag [ 2]

- [2] ALLOMBERT A., DESAINTE-CATHERINE, M., ASSAYAG G. « Iscore : A system for writing interaction », Proc. of the 3rd Digital Interactive Media in Entertainment & Art (DIMEA 2008), Athènes, 2008
- [3] AROM, S. « Modélisation et modèles dans les musiques de tradition orale », *Analyse Musicale*, N° 22, Février 1991, p. 91
- [4] AROM, S. et al, « Les Pygmées, peuple et musique » [cédérom], Montparnasse Multimédia, CNRS Audiovisuel, OSTROM, 1998.
- [5] CHEMILLER, M. « Pour une écriture multimédia de l'ethnomusicologie ». *Cahiers de Musique Traditionnelle* 16, 2003, p. 59-72
- [6] CHEMILLIER, M., RAPPOPORT, D. « Pourquoi présenter des modèles musicaux sur Internet ? », Actes de la table ronde Sémantique et Archéologie : aspects expérimentaux. Renouvements méthodologiques dans les bibliothèques numériques et les publications scientifiques, organisée par l'Ecole française d'Athènes, Athènes, 18 et 19 novembre 2000, édités par Andrea Iacovella, *Bulletin de Correspondance Hellénique*. p. 3
- [7] CREM, « Clés d'écoute » [en ligne] 2001, [consulté le 8 février 2013]. Disponible sur Internet <<http://ehess.modelisationsavoirs.fr/ethnomus/index.html>>
- [8] GOODY, J. « The Interference between the Written and the Oral », *New York, Cambridge University Press*.1987 p.630
- [9] MICHEL, P. *Problèmes de perception formelle dans la musique occidentale du XXe siècle*, thèse de doctorat, sous la direction de Mme le Professeur Eveline Andreani, Université Paris 8, 1997.
- [10] MOLINO, Jean. « Qu'est-ce que l'oralité musicale? » *Musiques, Une encyclopédie pour le XXe siècle*. Tome V. L'unité de la Musique. Sous la direction de Jean-Jaques Nattiez. Cité de la Musique, Paris, 2007 p. 477-527
- [11] SCALDAFERRI, Nicola « Pourquoi écrire les musiques non écrites ? Esquisse d'une anthropologie de l'écriture de la musique » *Musiques, une encyclopédie pour le XXIe siècle*, Tome V. L'unité de la Musique. Sous la direction de Jean-Jaques Nattiez. Cité de la Musique, Paris, 2007, p.627-666
- [12] SEEGER, Charles, 1958 : « Prescriptive and Descriptive Music Writing » *The Musical Quarterly*, vol. XLIV, no2, 1958 p.184-195; repris

in Id., *Studies in Musicology* (1935-1875), Berkeley-Los Angeles University of California Press, 1977, p.168-181; trad. Fr., " Notation prescriptive en notation descriptive", *Analyse musicale*, no 24, 1991, p. 6-12.

- [13] WILL, U.: « La baguette magique d'ethnomusicologie. Re-penser la notation et l'analyse de la musique ». *Cahiers de Musiques Traditionnelles*, Vol. XII. Georg, Genève, p.9-34.



# PROGRAMMATION ÉVÉNEMENTIELLE DE PARTITIONS MUSICALES INTERACTIVES.

D. Fober, S. Letz, Y. Orlarey

{fober,letz,orlarey}@grame.fr

Grame - Centre national de création musicale

## RÉSUMÉ

INSCORE est un environnement pour la conception de partitions musicales interactives qui intègre un système original d'interaction basé sur des événements et sur un langage de script permettant d'associer des messages arbitraires à ces événements. Initialement conçu pour être piloté via OSC, la version textuelle des messages OSC s'est rapidement constituée en format de stockage, puis étendue en un langage de script permettant une plus grande souplesse dans la description des partitions et des interactions avec ces partitions. Cet article présente ce langage de script et illustre notamment ses capacités à décrire des interactions sous forme événementielle, tout en restant dans l'espace temporel.

## 1. INTRODUCTION

INSCORE est un environnement pour la conception de partitions musicales qui propose une approche de la notation musicale étendue à des objets graphiques arbitraires [4], incluant la représentation de l'interprétation musicale. Cette approche nouvelle de la notation permet également la représentation des relations temporelles entre les objets de la partition [3].

La conception d'INSCORE fait réponse à une carence des outils informatiques actuels pour la notation de la musique, qui n'ont pas évolué en proportion des nouvelles formes de création musicale (voir par exemple [8] [5]). En particulier, il y a un fossé significatif entre les musiques interactives et la manière statique de les noter. Les technologies d'aujourd'hui permettent le calcul et l'interaction avec la musique en temps réel, mais la dimension symbolique de la notation est généralement exclue du processus d'interaction excepté dans des travaux très récents [6].

Conçu pour être piloté par des messages OSC [9], INSCORE se prête naturellement à une utilisation interactive. Cette approche de la programmation de partition par messages est également déclinée en un langage de script, basé sur une extension des messages OSC, et fournissant des primitives d'interaction reposant sur des notions d'événements. Ces événements sont similaires à ceux typiquement disponibles pour la gestion des interfaces utilisateurs (par exemple en Javascript via le DOM [7]), avec une extension dans le domaine temporel.

Cet article présente tout d'abord deux exemples de partitions interactives, mises en oeuvre dans des créations récentes. Il présente ensuite le système de messages et les événements d'interaction, qui permettent à la fois de décrire la partition et d'interagir avec elle. Des exemples d'usages viennent enfin illustrer les capacités d'expression du système.

## 2. PARTITIONS INTERACTIVES

Les musiques interactives sont aujourd'hui l'objet d'intérêts artistiques et scientifiques convergents. L'interaction soulève des problèmes à la fois pour la composition, la description et pour l'exécution des oeuvres. Ces problèmes sont abordés dans les aspects temporels de la partition interactive [1] ou du contrôle [2], et sont liés au calcul de l'oeuvre.

Pour la notation de l'oeuvre interactive, deux créations récentes ont mis en oeuvre INSCORE pour créer des partitions dynamiques avec des approches originales, qui témoignent également des besoins de la création contemporaine. Il s'agit de *Calder's Violin* et de *Alien Lands*.

### 2.1. Calder's Violin

*Calder's violin*, composé par Richard Hoadley, a été créé à Cambridge en Octobre 2011. La pièce est définie comme "*musique automatique pour violon et ordinateur*" et présente dynamiquement au musicien, de la notation musicale symbolique générée de manière algorithmique (figure 1). Cette partition est alors jouée par le musicien en parallèle des sons générés par l'ordinateur ou par d'autres musiciens. Les technologies utilisées reposent sur SuperCollider pour l'environnement de programmation audio et sur INSCORE pour la notation. Pour plus de détails se référer à [6].

### 2.2. Alien Lands

*Alien Lands* est un ensemble de pièces pour percussions et quatuor à corde, composées par Sandeep Bhagwati. Dans leur version interactive, ces pièces ont été données à Montreal en Février 2011. L'utilisation d'INSCORE relève de quatre catégories :

- partition traditionnelle avec tourne de page automatique,



Figure 1. Calder's Violin : exemple de partition.

- partition avec choix automatiques réalisés par l'ordinateur : ordre des mesures, sélection de lignes,
- partition complexe automatique, comprenant des éléments générés de manière algorithmique (figure 2),
- partition complexe interactive, où les éléments algorithmiques sont générés à la demande des musiciens.

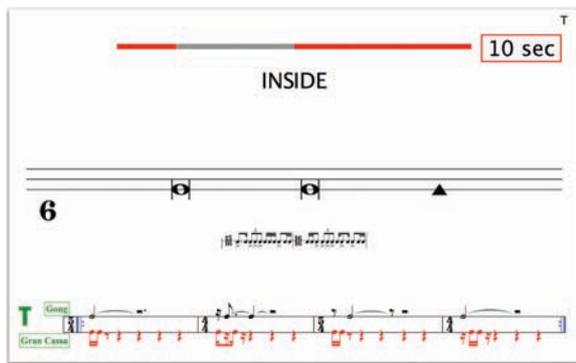


Figure 2. Alien Lands : une partition complexe automatique.

### 3. DESIGN DE PARTITION PAR MESSAGES

Le principe de base pour la description d'une partition, consiste à envoyer des messages OSC au système pour créer les différents composants de la partition et pour contrôler leurs attributs, aussi bien dans l'espace graphique que temporel.

#### 3.1. Format général des messages

Le format général des messages INSCORE est illustré figure 3. Il s'agit d'une spécialisation des messages OSC qui peut être vue comme *orientée objet*, où l'adresse désigne l'objet cible du message, *method* désigne une méthode de l'objet cible et *params*, les paramètres de la méthode. Un message INSCORE peut donc être vu comme l'appel d'une méthode d'un objet de la partition.

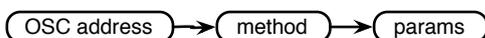


Figure 3. Format général des messages INScore.

Le système inclut des messages pour gérer les attributs graphiques des composants de la partition (position, couleur, échelle, rotations, effets...), pour gérer leurs attributs temporels (date, durée), pour exprimer les relations entre espaces graphiques et temporels, pour synchroniser des objets, pour représenter des signaux et pour gérer des événements d'interaction.

**Exemple 1** Change la position *x* de l'objet *obj*. L'adresse décrit la hiérarchie des objets : *obj* est contenu dans une scène nommée *scene* qui est incluse dans l'application d'adresse ITL.

```
/ITL/scene/obj x -0.5
```

#### 3.2. Le langage de script

Bien que prévu pour être émis sous forme de paquets sur un réseau, les messages OSC peuvent s'exprimer sous forme textuelle. C'est cette forme textuelle qui constitue le format de sauvegarde des partitions. Elle a été rapidement étendue pour en faire un langage de script.

##### 3.2.1. Adresses étendues

Les adresses OSC sont étendues pour permettre l'émission de messages aussi bien vers INSCORE qu'à destination d'une machine et/ou application externes (figure 4). Cela permet d'initialiser à la fois la partition et les ressources externes qui peuvent y être associées.



Figure 4. Extension du schéma d'adressage.

**Exemple 2** Initialisation d'une partition et d'une application externe à l'écoute du port 12000, sur une machine nommée *host.adomain.net*. Dans les scripts, le point virgule (;) est utilisé comme terminaison de message.

```
/ITL/scene/score set gmnf 'myscore.gmn';
host.adomain.net:12000/run 1;
```

##### 3.2.2. Variables

Des variables ont été introduites pour permettre le partage de paramètres entre messages. Une variable associe un identificateur et une liste de paramètres ou une liste de messages (figure 5). Les variables peuvent être utilisées dans les paramètres des messages sous la forme  $\$$ identificateur.

**Exemple 3** Déclaration de variables et utilisation comme paramètre de couleur. Le caractère '!' marque le début d'un commentaire.

```
color = 200 200 200;
colorwithalpha = $color 100;
msgsvr= (! une liste de messages
  localhost:7001/world "Hello world",
  localhost:7001/world "how are you?" );
/ITL/scene/obj color $colorwithalpha;
```

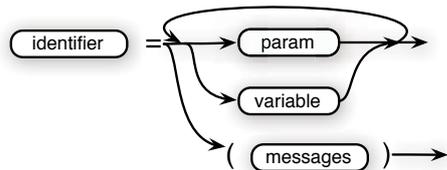


Figure 5. Variables.

Domaine graphique	Domaine temporel
mouseDown	timeEnter
mouseUp	timeLeave
mouseenter	durEnter
mouseleave	durLeave
mousemove	

Table 1. Principaux événements du système.

### 3.2.3. Langages

Les scripts supportent également l'inclusion de langages de programmation comme javascript (par défaut) ou lua. Les sections correspondantes sont indiquées de manière similaire à html (figure 6). Le code est évalué au moment de la lecture du script et le résultat attendu de l'évaluation est un ensemble de messages INSCORE qui vont remplacer le script correspondant.

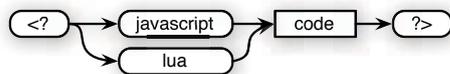


Figure 6. Langages.

principalement les événements du domaine graphique et permettent d'accéder à la position de la souris au moment de l'événement dans différents référentiels (\$x \$y \$sx \$sy) ou encore à la date correspondant à cette position (\$date).

**Exemple 4** Un objet qui suit les clics de la souris. La virgule (,) est utilisée comme séparateur de message dans les listes de messages.

```
/ITL/scene/obj watch mouseDown (
  /ITL/scene/obj x '$sx',
  /ITL/scene/obj y '$sy' );
```

## 4. INTERACTION ÉVÉNEMENTIELLE

Le processus d'interaction événementielle repose sur l'association de messages à des événements du système. Ces messages sont émis lorsque l'événement auquel ils sont associés se produit. Le format général des messages pour créer de telles associations est décrit en figure 7.

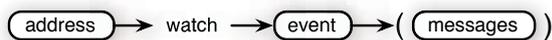


Figure 7. Format général d'un message d'interaction.

### 4.1. Typologie des événements

Les événements définis par le système sont d'une part des événements typiques d'interface utilisateur (tels que clics de souris par exemple) ainsi que des événements définis dans le domaine temporel (table 1)

Dans le domaine temporel, un événement est déclenché quand un objet entre ou quitte une zone temporelle (timeEnter, timeLeave) définie par deux dates, ou quand la durée d'un objet entre ou quitte une zone temporelle (durEnter, durLeave) définie par deux durées.

### 4.2. Variables contextuelles

Lorsqu'un événement se produit, les messages associés sont évalués, notamment parce qu'ils peuvent faire référence à des *variables contextuelles*. Ces variables concernent

## 5. CAS D'USAGE

### 5.1. Tourne de page

Une application simple des événements temporels consiste à décrire des tournes de page automatique. Un objet surveille les zones temporelles correspondants aux différentes pages et rappelle ces pages lorsque qu'il entre dans ces zones. A noter que cet objet peut aussi bien être un curseur qui se déplace sur la partition.

```
/ITL/scene/obj watch timeEnter 0 12
  (/ITL/scene/score page 1);
/ITL/scene/obj watch timeEnter 12 24
  (/ITL/scene/score page 2);
etc.
```

### 5.2. Séquences d'interactions

Les messages d'interaction décrits en figure 7 acceptent des messages arbitraires du système en tant que messages associés à un événement. Il est donc possible d'associer un message d'interaction à un événement et ainsi, de décrire des séquences d'interactions.

**Exemple 5** Description d'une séquence d'interactions basées sur le clic de souris : le premier clic change la couleur de l'objet, le second modifie l'échelle, le troisième fait une rotation, le quatrième modifie également l'échelle. La profondeur de ces associations n'est pas limitée.

```

/ITL/scene/obj watch mouseDown (
  /ITL/scene/obj color 100 100 255,
  /ITL/scene/obj watch mouseDown (
    /ITL/scene/obj scale 1.4,
    /ITL/scene/obj watch mouseDown (
      /ITL/scene/obj angle 45. ,
      /ITL/scene/obj watch mouseDown (
        /ITL/scene/obj scale 0.8 ))) );

```

### 5.3. Séquences dans le domaine temporel

La séquence d'interactions décrites ci-dessus (section 5.2) peut se décliner dans le domaine temporel en associant les changements d'état à des événements temporels et en déplaçant l'objet dans le temps. Avec cette approche, il est à la fois possible de changer l'ordre des événements, mais également de contrôler le déroulement de ces événements dans le temps.

Ce type de description permet de combiner dans une même expression, des approches événementielles, des accès non-séquentiel et du contrôle temporel.

**Exemple 6** Description d'une séquence d'interactions par association à des événements temporels. Les événements temporels sont déclenchés par l'entrée de l'objet dans des zones temporelles consécutives, et dont la durée est la ronde.

```

/ITL/scene/obj watch timeEnter 1 2
  (/ITL/scene/obj color 100 100 255);
/ITL/scene/obj watch timeEnter 2 3
  (/ITL/scene/obj scale 1.4);
/ITL/scene/obj watch timeEnter 3 4
  (/ITL/scene/obj angle 45.);
/ITL/scene/obj watch timeEnter 4 5
  (/ITL/scene/obj scale 0.8);

```

### 5.4. Structuration de l'espace temps

Dans le cadre d'expériences pédagogiques réalisées aux Ateliers des Feuillantines, la reconstruction du canon par ton de l'Offrande Musicale de J.S. Bach a été réalisée sous forme d'un objet à la fois *partition et instrument*, et INSCORE a été utilisé pour calculer dynamiquement la représentation symbolique de la musique jouée<sup>1</sup>. L'objet réalisé représente la structure de l'oeuvre en trois dimensions (figure 8). Chaque barre, sur lequel un capteur peut être posé, est muni d'une étiquette indiquant le chemin à suivre, et correspond à une transposition du canon.

Du point de vue de la représentation symbolique, les différentes transpositions peuplent l'espace temporel à des dates différentes, de telle sorte que sur l'instrument, l'association d'une position à une date permet de rappeler la transposition correspondante via les événements temporels d'INSCORE.

1. <http://tinyurl.com/ah3jftu>



**Figure 8.** L'objet *partition / instrument* de l'Offrande musicale.

## 6. CONCLUSION

L'association de messages à des événements se révèle être un mécanisme simple, puissant et homogène pour la description de partitions dynamiques. Il est désormais possible de décrire l'ensemble des comportements d'une partition dans le langage de script d'INSCORE. L'utilisation d'applications externes pour piloter ces comportements peut dès lors se limiter au déplacement d'objets dans le temps en associant à ces objets, des comportements liés à des espaces temporels.

### Remerciements

Cette recherche a été menée dans le cadre du projet INEDIT qui est soutenu par l'Agence Nationale pour la Recherche [ANR-12-CORD-009-03].

## 7. REFERENCES

- [1] A. Allombert, "Aspects temporels d'un système de partitions musicales interactives pour la composition et l'exécution," Ph.D. dissertation, Université Bordeaux 1, Oct. 2009. [Online]. Available : <http://tel.archives-ouvertes.fr/tel-00516350>
- [2] A. Cont, "Antescofo : Anticipatory synchronization and control of interactive parameters in computer music." in *Proceedings of International Computer Music Conference*, ICMA, Ed., 2008.
- [3] D. Fober, C. Daudin, S. Letz, and Y. Orlarey, "Time synchronization in graphic domain - a new paradigm for augmented music scores," in *Proceedings of the International Computer Music Conference*, ICMA, Ed., 2010, pp. 458–461.
- [4] D. Fober, C. Daudin, Y. Orlarey, and S. Letz, "Interlude - a framework for augmented music scores," in *Proceedings of the Sound and Music Computing conference - SMC'10*, 2010, pp. 233–240.

- [5] J. Freeman, "Bringing instrumental musicians into interactive music systems through notation," *Leonardo Music Journal*, vol. 21, no. 15-16, 2011.
- [6] R. Hoadley, "Calder's violin : Real-time notation and performance through musically expressive algorithms," in *Proceedings of International Computer Music Conference*, ICMA, Ed., 2012, pp. 188–193.
- [7] B. Höhrmann, P. Le Hégarret, and T. Pixley, "Document object model (dom) level 3 events specification," World Wide Web Consortium, Working Draft WD-DOM-Level-3-Events-20071221, December 2007.
- [8] T. Magnusson, "Algorithms as scores : Coding live music," *Leonardo Music Journal*, vol. 21, pp. 19–23, 2011.
- [9] M. Wright, *Open Sound Control 1.0 Specification*, 2002. [Online]. Available : [http://opensoundcontrol.org/spec-1\\_0](http://opensoundcontrol.org/spec-1_0)



# CARTES ET TABLEAUX INTERACTIFS : NOUVEAUX ENJEUX POUR L'ANALYSE DES MUSIQUES ELECTROACOUSTIQUES

*Pierre Couprie*

OMF-MINT Université Paris-Sorbonne

MTIRC De Montfort University

pierre.couprie@paris-sorbonne.fr

## RÉSUMÉ

L'analyse musicale utilise différents types de représentations sous la forme de tableaux ou de cartes. Elles permettent de mettre en évidence des relations syntagmatiques (structures) ou paradigmatiques (relations entre des objets segmentés). Si ces tableaux et cartes sont couramment utilisés dans le champ de la musique écrite, cette activité est moins courante dans le domaine de la musique électroacoustique. Dans ce cas, le modèle acousmographie est généralement utilisé. Pourtant, il existe plusieurs logiciels (Open Music, Musique Lab 2, OMax, IAA, EAnalysis) permettant de réaliser ces types de représentations et apportant des fonctions d'interactivités adaptées à la musicologie et au corpus actuel des musiques électroacoustiques.

## 1. INTRODUCTION

La réalisation de tableaux et de cartes est une activité assez courante dans le domaine de l'analyse [1]. Présenter des unités segmentées sous une forme différente qu'un alignement temporel simple permet souvent d'en saisir la complexité ou de dégager de nouvelles relations. C'est une activité moins courante dans le domaine de l'analyse des musiques électroacoustiques, probablement due à la difficulté de représenter des unités sonores complexes.

Depuis de nombreuses années, la représentation de type acousmographie est devenue une véritable référence dans le domaine de l'analyse de l'électroacoustique [2]. Permettant de symboliser des unités sur un axe temporel, elle convient parfaitement à un corpus historique de musique de sons fixés qui occupe les musicologues depuis plus de 30 ans. A partir du milieu des années 90, de nouvelles formes de créations électroacoustiques remplacent progressivement le concert traditionnel en intégrant l'interactivité, l'improvisation, d'autres formes artistiques comme la vidéo, la poésie sonore ou des pratiques sociales comme le jeu. Ces formes de création ne sont pas nouvelles mais elles ont bénéficié des progrès de l'informatique portable pour s'imposer progressivement dans de nombreux centres de création.

Dans cet article, je vais d'abord présenter ces nouvelles pratiques artistiques et les enjeux qui les accompagnent dans le domaine musicologique. Je montrerai l'utilisation des tableaux et des cartes

navigables pour augmenter les possibilités de l'acousmographie traditionnelle. La seconde partie de l'article exposera différents exemples de tableaux et de cartes réalisés à partir de logiciels adaptés ou conçus pour l'analyse musicale.

## 2. L'EVOLUTION DU MODELE ACOUSMOGRAPHIE

### 2.1. Le corpus actuel des musiques électroacoustiques

Le corpus actuel des musiques électroacoustiques se recentre depuis une quinzaine d'années sur des pratiques intégrant l'interactivité, l'improvisation ou la fusion avec les autres arts. Ces formes de création ne sont pas nouvelles mais elles se sont généralisées au point de devenir majoritaires dans certains centres de création. Pour le musicologue, elles bouleversent profondément un certain nombre de notions comme celle d'œuvre, d'instrument ou de partition.

Ainsi, la notion d'instrument doit être augmentée à l'ensemble des machines utilisées par le musicien. Dans bien des cas, l'instrument augmenté se superpose aussi en partie à la partition. En effet, cette dernière ne contient plus seulement les signes destinés aux musiciens mais aussi les algorithmes permettant de générer la partie électronique. Analyser une œuvre interactive devient une activité complexe dans laquelle l'objet même de l'étude est complexe à circonscrire. Dans certains cas, une captation de la performance en vue d'une étude musicologique devient même quasiment impossible à réaliser. A l'instar de Jean Baudrillard [3] déclarant la fin du réel dans l'image numérique, de telles performances posent la question de la définition même de l'objet d'analyse. Doit-on analyser l'œuvre et ses sources (partition, sons préenregistrés, *patches*, etc.) qui sont les modèles traditionnels de la performance ou cette dernière qui devient, comme pour l'image numérique, le modèle de l'œuvre ?

De la captation de l'œuvre à son étude en allant jusqu'aux problématiques d'archivage, il est aisé de constater que le métier de musicologue est en pleine mutation. La généralisation du numérique dans la création en est probablement une des causes, mais c'est aussi la profonde mutation et l'hybridation des pratiques artistiques qui obligent le chercheur à intégrer le

numérique et manipuler des objets complexes dans son travail d'analyse.

## 2.2. Les limites du modèle acousmographie

Le logiciel Acousmographe existe depuis plus de 20 ans. Il a été un formidable catalyseur de travaux d'analyse sur la musique électroacoustique. Ce modèle peut être défini comme la symbolisation d'unités segmentées alignées sur un axe temporel de gauche à droite et sur un axe vertical représentant un paramètre d'analyse. La forme, la couleur, la texture et la taille du symbole sont aussi associées à un ensemble de paramètres ou à une classification des unités analysées. Cette symbolisation est en réalité une extension du sonagramme qui est déjà lui-même une représentation du spectre réel. Ce type de représentation est aussi basé sur le modèle du magnétophone, le sonagramme représentant la bande magnétique se déroulant de droite à gauche. Ce sonagramme augmenté de symboles est devenu une référence en matière d'outil d'aide à l'analyse musicale. Il convient parfaitement à l'analyse d'œuvres de support ou d'œuvres mixte suivant un déroulement temporel prédéterminé. Dans le cas des corpus actuels le modèle acousmographie présente d'évidentes limites.

La première limite est la forme même de la représentation. Les dimensions habituelles construites sur l'axe temporel ne correspondent plus au travail du compositeur utilisant les technologies numériques. Elles sont basées sur les technologies analogiques dans lesquelles les voies de mixage sont superposées à l'aide de magnétophones dont le déroulement temporel est synchrone ou fonctionne sur le même mode. Les interfaces actuelles de création mélangent des données temporelles (par exemple les fragments audio préenregistrés), des processus (par exemple les opérations de transformation audio) et des états séquentiels d'interactions entre le musicien et son instrument (par exemple les données des contrôleurs MIDI ou OSC). Ces nouveaux objets n'ont plus de bords, possèdent des dimensions temporelles variables, floues ou inexistantes. Cette complexité de dispositif devient impossible à symboliser dans une seule dimension temporelle.

La deuxième limite est celle du format des données à analyser. Le dispositif électroacoustique produit des données audio, MIDI/OSC, un ensemble de fichiers variés correspondant à des listes de données et il est souvent intéressant d'avoir un enregistrement vidéo de la performance. L'ensemble de ces données ne rentre plus dans le modèle acousmographie.

Enfin, la troisième limite est celle du lien avec les autres logiciels. Le musicologue ne travaille plus avec un seul outil, certains logiciels seront particulièrement adaptés pour appliquer des descripteurs audio, d'autres vont permettre d'analyser les fichiers audio et vidéo d'une manière spécifique, les logiciels utilisés pour la création de l'œuvre sont aussi des outils importants pour le musicologue. Il devient donc essentiel d'utiliser des

logiciels permettant les échanges de données ou l'interaction à partir de différents protocoles.

## 2.3. Les enjeux d'un nouveau modèle pour l'analyse des musiques électroacoustiques

En étudiant les pratiques actuelles de l'analyse musicale, nous voyons les musicologues complexifier leur démarche. Ainsi, l'étude des *patches* du compositeur devient essentielle, la collaboration du musicologue à des recherches dans le domaine de l'archivage [4] est un enjeu majeur de l'étude des œuvres interactives ou les propositions de représentations faites par les logiciels de création sont une source d'information très riche pour le musicologue.

Au-delà de ces nouvelles données, un des enjeux qui semblent les plus problématiques est celui des flux de travail. En effet, le musicologue doit optimiser son usage des différents logiciels. Mais ce n'est possible que si ces derniers permettent l'interopérabilité. En effet, l'absence de format d'échange musicologique qui permettrait de faciliter le transfert et l'analyse des données musicales est un des grands enjeux des futurs logiciels.

Le modèle acousmographie doit donc être augmenté afin de prendre en compte les corpus actuels et les nouveaux enjeux de la musicologie dans le cadre des humanités numériques.

## 2.4. Les tableaux et les cartes en analyse musicale

Les tableaux et cartes utilisés en analyse musicale permettent de mettre en évidence certaines relations et groupements dans les objets analysés. La mise en tableau ou carte commence par une étape de segmentation, les unités ainsi isolées sont ensuite réorganisées visuellement sous la forme de tableau (fragment de partitions ou de sonagrammes alignés) ou de carte. Dans ce dernier cas, la position graphique des unités est plus libre et permet de mettre en évidence des relations plus complexes ou sur plusieurs niveaux. Ces relations peuvent être d'ordre paradigmatique (équivalence entre les objets) ou syntagmatique (regroupement des objets en structures) [5]. Ils permettent surtout de s'abstraire de l'axe temporel traditionnellement utilisé dans la partition ou la représentation graphique. Ils sont couramment utilisés dans trois types d'activités analytiques : la classification, la mise en tableau paradigmatique et l'analyse génétique.

La classification d'objets segmentés (unités, fonctions ou élément de structure) permet de faire émerger des catégories. Généralement ces catégories sont ensuite organisées afin de faire apparaître des catégories d'ordre supérieures.

La réalisation de tableaux paradigmatiques [6] permet de classer les objets en fonctions de leur ordre d'apparition et de leur équivalence. Ce type de tableau met en évidence le travail de développement du matériau musical.

L'analyse génétique utilise des cartes représentant les modifications du matériau source avant et pendant son intégration dans l'œuvre. Cette activité révèle les choix réalisés par l'artiste pendant le processus de création.

Ces trois exemples montrent la richesse de l'utilisation des tableaux et cartes dans l'analyse musicale.

### 3. QUELQUES EXEMPLES DE LOGICIELS

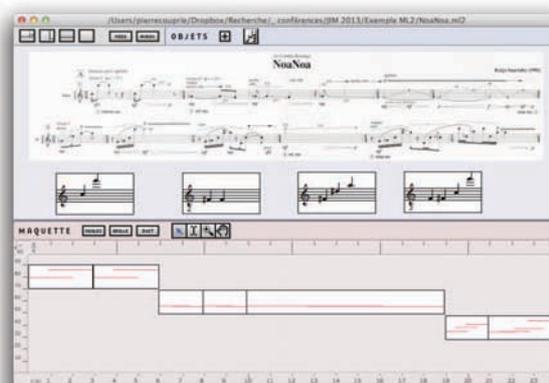
Je vais maintenant présenter quelques exemples de logiciels permettant de créer et de manipuler ces tableaux et cartes. Certains outils sont avant tout dédiés à la création, d'autres ont été spécifiquement développés pour l'analyse musicale.

#### 3.1. La maquette d'Open Music et de Musique Lab 2

Open Music est un environnement de composition développé par l'Ircam. Il permet de travailler sur des processus créatifs et de manipuler des données MIDI, audio ou simplement textuel. Ce logiciel possède un outil nommé Maquette permettant de positionner des objets (notes, fichiers audio) ou des opérations à effectuer sur ces objets sur un axe temporel. Open Music permet ainsi de résoudre une des limites de l'Acousmographe à savoir la représentation de processus musicaux.

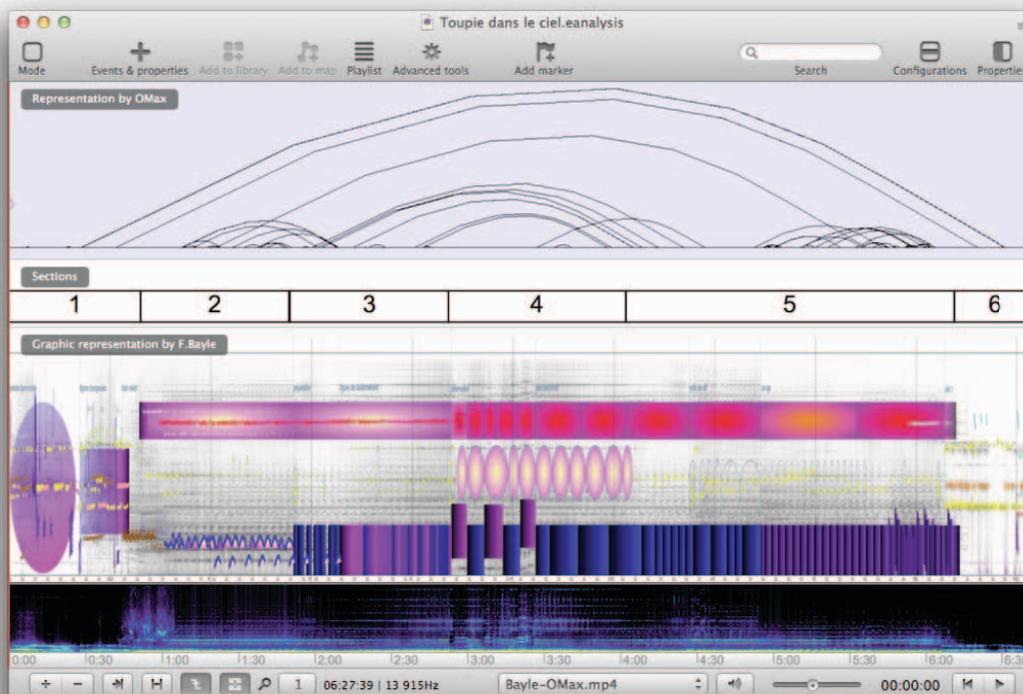
Le logiciel Musique Lab Maquette [7] est une version d'Open Music réalisé dans une visée pédagogique. Comme avec Open Music, le chercheur peut manipuler

des fragments d'œuvres et d'explorer les processus de création. Il devient ainsi très simple d'extraire des fragments de l'œuvre [8] afin de les manipuler séparément et réaliser différents types de tableaux (figure 1) ou modéliser la structure d'une œuvre.



**Figure 1.** Le début de *NoaNoa* de Kaija Saariaho, représentation paradigmatique dans Musique Lab 2.

Ce type de représentation offre aussi la possibilité de construire des parcours d'écoute dans l'œuvre. La maquette apparaît donc comme un outil simple et puissant pour manipuler temporellement les objets analysés.



**Figure 2.** *Toupie dans le ciel* de François Bayle, représentations assemblées avec EAnalysis : générée par OMax (haut), par le compositeur (milieu), sonagramme (bas)

### 3.2. La visualisation de la structure dans OMax

OMax est un logiciel d'improvisation MIDI ou audio qui intègre, depuis la version 4, une interface de visualisation [9] représentant en temps réel le modèle en court d'apprentissage. La segmentation réalisée par OMax est basée sur un ensemble de descripteurs audio. Le logiciel repère ensuite les structures des fragments segmentés en obéissant à un oracle des facteurs. La représentation affiche les liens entre les différentes structures qui permettront au logiciel de produire une improvisation. En s'arrêtant avant la phase d'improvisation, il est possible de générer la représentation d'une œuvre (improvisée ou non) résultant de l'analyse réalisée par OMax. Dans le cas de la figure 2, la représentation en quatre arches correspond parfaitement à celle du compositeur :

- une grande arche reliant le début à la fin (la section 1 à la fin de la section 5 et à la section 6) ;
- deux arches situées au premier et au troisième quart permettant d'isoler la deuxième et une partie de la cinquième section de la pièce ;
- une arche centrale mettant en évidence la relation entre la troisième et le début de la cinquième section.

La figure 2 représente la juxtaposition de l'analyse d'une œuvre acousmatique (*Toupie dans le ciel* de François Bayle) réalisée par le compositeur [10] et la représentation réalisée par OMax. Ce type de représentation met en évidence les relations entre les différentes sections de l'œuvre. La correspondance évidente avec l'analyse du compositeur lui-même permet de valider ce type de représentation pour l'analyse musicale.

Cette représentation apporte une direction de recherche dans les opérations de navigation au sein d'un même fichier. Le fait de relier des positions temporelles espacées et de produire, lors de la phase d'improvisation, une lecture différente de l'œuvre présente d'évidentes possibilités pour l'analyse musicale. Appliqué à un outil de lecture, OMax permettrait de multiplier les possibilités de navigation au sein d'une ou plusieurs œuvres.

### 3.3. Interactive Aural Analysis (IAA)

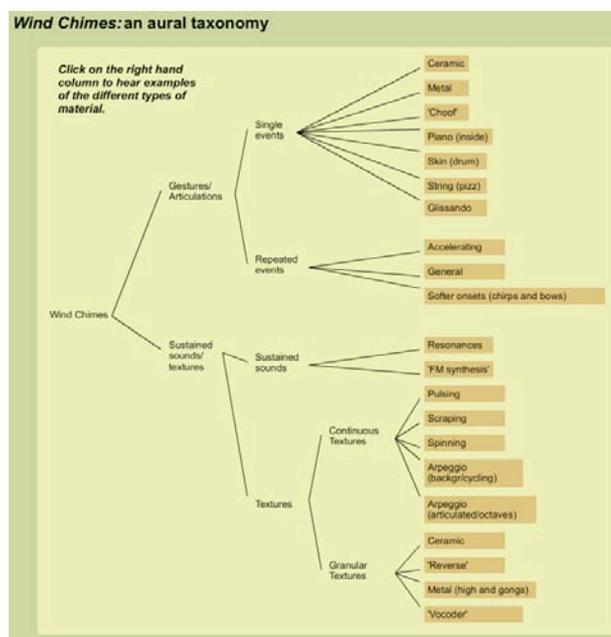
IAA [11] est une activité d'analyse musicale développée par Michael Clarke à l'université d'Huddersfield. Une des réalisations consiste à publier des analyses sous la forme de logiciels interactifs. Après trois réalisations fixes, une nouvelle version du logiciel en cours de développement permettra à l'utilisateur de créer ses propres analyses.

Le principe de ce logiciel est d'utiliser plusieurs types de représentations interactives pour analyser et présenter l'analyse d'une œuvre.

La classification (*aural taxonomy*) permet de visualiser les relations entre les sons sous la forme d'une carte (figure 3). Cette carte reprend les matériaux

d'origine de l'œuvre et produit un premier classement. Ce classement permettra ensuite de segmenter l'œuvre.

Le tableau paradigmatique (*aural paradigmatic chart*) permet de visualiser les unités segmentées sous une forme paradigmatique (figure 4). Nous y retrouvons les catégories de la figure 3 (axe vertical) et le résultat de la segmentation (axe horizontal).



**Figure 3.** Classification interactive des sons de *Wind Chimes* de Denis Smalley par Michael Clarke.

L'interactivité de cette carte et ce tableau est particulièrement intéressante pour l'analyse des musiques électroacoustiques : l'utilisateur peut cliquer sur les objets afin d'écouter les sons, il peut même utiliser des fonctions de filtrage dans le cas du deuxième exemple. Michael Clarke nous montre ici une voie très prometteuse dans l'usage du numérique en analyse musicale.

### 3.4. EAnalysis

EAnalysis [12] est un logiciel d'aide à l'analyse des musiques électroacoustiques ou non écrites du groupe de recherche MTIRC<sup>1</sup> de l'université De Montfort de Leicester. Le développement de ce logiciel m'a permis d'expérimenter et de mettre en pratique un certain nombre de recherches prospectives effectuées depuis 2007 sur le logiciel iAnalyse Studio<sup>2</sup> [13] et dans le domaine de la représentation graphique d'analyse

<sup>1</sup> *Music, Technology and Innovation Research Centre*: <http://mti.dmu.ac.uk>.

<sup>2</sup> Depuis début 2013, iAnalyse est devenu iAnalyse Studio, suite de logiciels d'aide à l'analyse musicale, et est disponible gratuitement sur : <http://ianalyse.pierrecouprie.fr>.

musicale. EAnalysis a hérité de certaines fonctions de iAnalysis Studio mais, si le second logiciel est dédié à l'aide à l'analyse de la musique écrite, le premier est

particulièrement utile pour l'analyse des musiques électroacoustiques.



Figure 4. Tableau paradigmatique interactif de *Wind Chimes* de Denis Smalley par Michael Clarke.

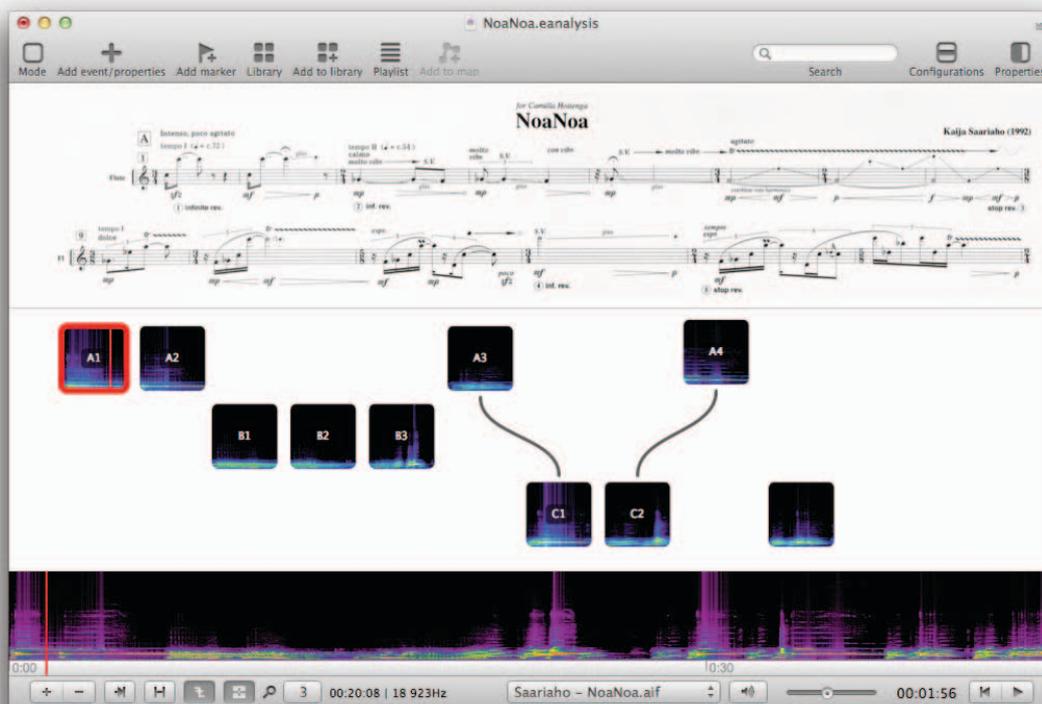


Figure 5. *NoaNoa* de Kaija Saariaho, représentation réalisée avec EAnalysis : partition (haut), tableau paradigmatique (milieu), sonagramme (bas).

L'objectif d'EAnalysis est de proposer au musicologue un ensemble d'outils graphiques pour l'aider dans son travail d'analyse. Depuis la présentation aux dernières Journées d'Informatique Musicale de Mons, le logiciel a beaucoup évolué, notamment dans le domaine de la navigation à l'intérieur d'un ou plusieurs fichiers audio et/ou vidéo et dans la réalisation de plusieurs modèles de représentations. Un des aboutissements tient dans la possibilité de réaliser des cartes de fragments d'une ou plusieurs œuvres (figure 5). Le musicologue peut ainsi sélectionner des extraits, les ajouter dans la carte et les organiser en fonction des critères qu'il a choisis. Ces extraits sont symbolisés par des fragments de sonagrammes, de forme d'onde ou simplement des graphiques et peuvent être reliés entre eux. L'ensemble est manipulable en lecture dans les deux sens. Ainsi, le musicologue peut visualiser directement sur la carte l'enchaînement des fragments en lisant l'œuvre (un curseur apparaît sur chaque fragments) ou, à l'inverse, visualiser la position des fragments dans l'œuvre en sélectionnant les éléments de la carte.

Différents types de cartes peuvent être réalisés avec EAnalysis : typologie, tableau paradigmatique, arbres. génératif, création de parcours d'écoute, etc. La force de ce logiciel tient dans le fait de pouvoir juxtaposer ou superposer plusieurs types de représentations et ainsi mettre en relation un ou plusieurs enregistrements audiovisuels, des images (partitions, photographies, etc.), des représentations analytiques ou des données provenant d'autres logiciels.

#### 4. CONCLUSION

Cet article m'a permis de démontrer les possibilités offertes par différents logiciels dans le domaine de l'analyse musicale. D'une représentation traditionnelle de type acousmographie nous sommes maintenant en mesure de l'enrichir avec un ensemble d'autres représentations sous la forme de tableaux et de cartes interactives.

Ces nouveaux outils répondent parfaitement aux enjeux posés par le corpus actuel des musiques électroacoustiques qui nécessite de pouvoir s'extraire de la dimension temporelle et de manipuler facilement des objets segmentés.

#### 5. REFERENCES

- [1] Donin, N., Goldman, J. "Charting the Score in a Multimedia Context: the Case of Paradigmatic Analysis", Music Theory Online, 14-4, USA, 2008.
- [2] Delalande, F. "Pratiques et objectifs des transcriptions des musiques électroacoustiques", L'analyse musicale, une pratique et son histoire, Genève, 2009.
- [3] Baudrillard, J. *Le crime parfait*, Paris, Galilée, 1995.
- [4] Bonardi, A. "Approches pratiques de la préservation/virtualisation des œuvres interactives mixtes : *En Echo* de Manoury", JIM, Saint-Etienne, 2011.
- [5] Chouvel, JM. *Analyse musicale Sémiologie et cognition des formes temporelles*, L'Harmattan, Paris, 2006.
- [6] Ruwet, N. *Langage, musique, poésie*. Seuil, Paris, 1972.
- [7] Bresson, J., Guédy, F., Assayag, G. "Musique Lab Maquette : une approche interactive des processus compositionnels pour la pédagogie musicale", Revue des sciences et technologie de l'information et de la communication pour l'éducation et la formation, 13, Le Mans, 2006.
- [8] Donin, N. Goldszmidt, S. Theureau, J. "Instrumenter les opérations d'écoute analytique ? Un bilan du projet 'écoutes signées' (2003-2006)", JIM, Rennes, 2010.
- [9] Lévy, B., Bloch, G., Assayag, G. "OMaxist Dialectics : Capturing, Visualizing and Expanding Improvisations", NIME, Ann Arbor, 2012.
- [10] Bayle, F. *Erosphère*, Paris, Magison, 2009.
- [11] Clarke, M. "Analysing Electroacoustic Music: an Interactive Aural Approach", Music Analysis, USA, 2012.
- [12] Couprie, P. "EAnalysis : aide à l'analyse de la musique électroacoustique", JIM, Mons, 2012.
- [13] Couprie, P. "Analyser la musique électroacoustique avec le logiciel iAnalyse", EMS, Paris, 2008.

## EXTENSION DU MUSICAL VERS L'INTERMÉDIALITÉ: ANALYSES DES APPROCHES DE THIERRY DE MEY ET DE GEORGES APERGHIS

*Felipe Merker Castellani*

Université de Campinas (UNICAMP)  
Fondation pour la recherche de São Paulo  
felipemerkercastellani@yahoo.com.br

*Anne Sèdes*

CICM - EA 1572 - Université Paris 8  
MSH Paris Nord  
sedes.anne@gmail.com

### RÉSUMÉ

Cette étude est une première approche en vue d'établir une méthodologie pour analyser l'extension du musical vers l'intermédialité. Nous examinons la manière dont Thierry De Mey et Georges Aperghis opèrent la convergence entre la musique et d'autres moyens d'expression artistique.

Sur le plan méthodologique, nous examinons les aspects poétiques et poétiques accessibles dans le discours des compositeurs, à l'examen des partitions et des archives audiovisuelles.

Chez Thierry De Mey, il paraît que la convergence entre musique et autres media est conduite par l'usage de la notion de mouvement et par la maîtrise du temps musical à diverses échelles, ce qui est facilité par les moyens du numérique.

Chez Georges Aperghis, du moins dans ses œuvres récentes mettant en scène les nouvelles technologies numériques, c'est la composition du dispositif scénique, qui, pensé comme une partition, offre les conditions de la convergence entre l'écriture musicale et l'écriture théâtrale.

### 1. INTRODUCTION

Cet article est le fruit du séjour de Felipe Merker au CICM<sup>1</sup>, l'hiver dernier, dans le cadre de sa recherche doctorale ayant pour sujet "La création musicale en relation avec d'autres pratiques artistiques"<sup>2</sup>. Cette recherche se développe sur deux axes qui interagissent entre eux : l'investigation musicologique et la pratique artistique; l'étude présente s'inscrit dans ce premier axe. Elle résulte d'un dialogue avec Anne Sèdes sur l'extension du musical et de la mixité vers l'intermédialité et sur les méthodologies adéquates en musicologie alliant recherche et création<sup>3</sup>.

On cherche à examiner dans cette étude l'approche de l'intermédialité par des musiciens, dans la continuité de l'approche de la mixité et de la convergence entre

moyens, telle que posée par Horacio Vaggione<sup>4</sup> [21]. On va tenter d'examiner la convergence entre la musique et d'autres média, en recherchant d'une certaine façon, l'extension de l'approche musicale du compositeur à l'intermédialité. C'est à dire, on va proposer une approche méthodologique qui montre la mise en œuvre de la convergence basée sur une approche musicale.

On a choisi de traiter le cas de Thierry De Mey et celui de Georges Aperghis car ils illustrent particulièrement une telle démarche.

### 2. APPROCHE MÉTHODOLOGIQUE

Pour cette étude, les sources principales émanent des entretiens qu'a eu Felipe Merker<sup>5</sup> avec les compositeurs en question, des partitions, des archives audiovisuelles et des écrits des compositeurs, voire de leurs commentateurs.

L'approche méthodologique proposée vise l'examen de la convergence et les interactions entre les différents moyens/media qui sont mis en œuvre par les compositeurs. Nous cherchons à examiner la construction de la 'musicalité', dans son extension intermediale, ce qui revient à nous interroger sur qui est 'musicable' dans le maniement des différents media de l'art. On comprend ici le medium dans le sens d'un moyen d'expression accessible à l'artiste, et pas spécialement au sens des (nouveaux) media propres au numérique.

Evidemment, le numérique est facilitateur de convergence, mais on examine ici moins les aspects techniques et technologiques des media que le potentiel artistique et spécialement musical qu'ils portent en eux, et qui ne peut émerger qu'à travers l'usage qu'en font les créateurs dans un contexte culturel donné. Car c'est bien la création artistique qui reste au centre de nos recherches, même si elle se développe en interaction avec l'apparition et l'usage de nouvelles technologies, d'appareils, comme aurait dit Walter Benjamin [4].

<sup>1</sup> Centre de recherche en Informatique et Création Musicale, université Paris 8, MSH Paris nord (<http://cicm.mshparisnord.org/>).

<sup>2</sup> Felipe Merker Castellani est doctorant de l'Université de Campinas, sous la direction de Silvio Ferraz, il est boursier de la Fondation pour la recherche de São Paulo (FAPESP).

<sup>3</sup> A. Sèdes a développé cette réflexion sur l'extension du musical à divers media dans le cadre de son séminaire "composition et recherche" à l'université Paris 8, et dans le cadre du programme "musiques mixtes" de la MSH Paris Nord (2011/2012).

<sup>4</sup> « Il est souhaitable de chercher des passerelles très fines afin de faire interagir les deux sources d'une façon très rapprochée, au niveau du résultat sonore, mais aussi au niveau du processus de composition lui-même, en travaillant à partir de la même situation musicale, c'est à dire en assumant le postulat d'une possible convergence entre les mondes instrumentale et électroacoustique, en créant une vectorisation commune » [21].

<sup>5</sup> Nous remercions Georges Aperghis, Thierry De Mey, Émilie Morin et Jean Geoffroy.

Par exemple, pour le musicien électroacousticien, qu'est-ce qui est musicable dans l'emploi de la vidéo numérique comme moyen d'expression accessible<sup>6</sup>? On parle ici de moyens d'expression musicale, pas de moyens de communication (M. Mc Luhan) [20] ou de transmission (R. Debray) [8].

On examinera les niveaux, que nous nommerons poétiques et poïétiques<sup>7</sup> dans le discours des compositeurs, dans leurs œuvres, et autant que possible en ayant recours aux archives vidéo, et aux partitions.

Le niveau poétique concerne l'environnement conceptuel et esthétique du compositeur, ses inspirations, ses désirs, ses intuitions, ses idées, et la façon dont il les verbalise en évoquant ses projets.

Le niveau poïétique concerne les "manières de faire" du compositeur, l'espace conceptuel, l'espace composable pour reprendre une proposition d'H. Vaggione, [21], reprise par Carvalho [5] et les opérations qu'il effectue dans cet espace pour faire œuvre, ici en faisant converger le musical et d'autres media.

A partir du discours du compositeur, nous pouvons déjà identifier le champ problématique et opératoire qui constitue la 'facture' des pièces analysées ; nous pouvons également tenter de retrouver dans les œuvres les traces des opérations constitutives de l'œuvre.

### 3. THIERRY DE MEY, LE MOUVEMENT COMME MOYEN DE CONVERGENCE ENTRE MUSIQUE, VIDÉO ET DANSE

Nous commencerons par examiner la relation entre la musique et la danse dans les films de Thierry De Mey réalisés en collaboration avec la chorégraphe Anne Teresa De Keersmaecker. Ensuite, nous travaillons sur l'écriture de gestes développée par le compositeur depuis des années 1980.

<sup>6</sup> On revoit ici au texte de Marc Billon et Anne Sèdes "Musiques visuelles, composition (musicale) audiovisuelle", également présenté aux JIM 2013.

<sup>7</sup> On emprunte à Paul Valéry l'usage des termes "poétique" et "poïétique". On précise, pour notre part, l'usage de ces deux termes pour des raisons de facilité méthodologique : pouvoir poser d'une part l'étape du projet et de son inspiration (sa poétique), et d'autre part, examiner sa mise en œuvre avec les moyens choisis (sa poïétique) (espace composable, construction de l'instrument). Rappelons que Valéry divise en deux groupes les études sur l'art (ce qu'on appelle à l'époque "l'esthétique"), d'une part le groupe qui concerne l'esthétique, qui se rapporte à l'étude des sensations, et d'autre part, le groupe qui concerne la poétique, et qui se rapporte à l'étude des productions. Pour Valéry, la poïétique est « tout ce qui concerne la production des œuvres ; et une idée générale de l'action humaine complète, depuis ses racines psychiques et physiologiques, jusqu'à ses entreprises sur la matière ou sur les individus, permettrait de subdiviser ce second groupe, que je nommerais *Poétique*, ou plutôt *Poïétique*. D'une part, l'étude de l'invention et de la composition, le rôle du hasard, celui de la réflexion, celui de l'imitation ; celui de la culture et du milieu ; d'autre part, l'examen et l'analyse des techniques, procédés, instruments, matériaux, moyens et supports d'action » [22]. Évidemment, en musicologie, on peut difficilement employer le terme de "poïétique" sans renvoyer à la sémiologie de la musique de Jean-Jacques Nattiez, mais n'ayant pas recours ici à la sémiologie, nous retournons à la source originale, sous la plume de Paul Valéry.

Chez Thierry De Mey, c'est la notion de mouvement qui va lui permettre de faire converger les media. Celle-ci se manifeste de plusieurs façons dans son travail : comme des 'catalogues' de figures musicales - c'est-à-dire des mouvements mélodiques et rythmiques, comme des topologies - des formes géométriques que les danseurs suivent sur le sol de l'espace scénique, ou encore comme la trajectoire de la caméra cinématographique.

« L'interface entre les différentes disciplines est le mouvement. Puisque, si tu mets la musique, la danse, le cinéma, le point commun entre ces disciplines est qu'ils sont des arts du mouvement. C'est-à-dire: ils ne sont pas uniquement des arts du temps, ils ne sont pas uniquement des arts de l'espace, ils ne sont pas uniquement des arts d'une unique sensation: ce n'est pas uniquement l'écoute, ce n'est pas uniquement la vue. Ces sont des arts où l'élément synthétique, l'élément qui fait la synthèse de tout ça c'est le mouvement [...]. Si je prends ça du côté musical, c'est une conception du rythme, le rythme n'est pas uniquement une combinaison des longueurs, des temps forts, des temps faibles, etc. Mais le rythme est la tension entre les éléments. Le rythme est entre les choses » [7].

Du côté de la relation entre la musique et la danse, Thierry De Mey recourt entre autres à la construction d'un «vocabulaire» et d'un système d'équivalences [13]. C'est le cas du film *Rosas danst Rosas* (1983), où la musique et la danse vont toujours ensemble, chaque événement musical correspondant à un mouvement déterminé de la danse. Cette équivalence rigoureuse n'empêche pas les contrepoints - entre la musique et la danse - ni le déroulement temporel. Au contraire, elle les construit. L'automatisme de la musique répétitive et de la coïncidence absolue met en place deux tensions : l'une basée sur l'extrême précision entre les mouvements des quatre danseuses, entre leurs phases et déphasages et l'autre s'installant entre cette précision et la répétition d'un répertoire gestuel particulier, composé de « gestes courts et typés de la séduction – la main qui lisse les cheveux, qui recouvre l'épaule dénudée ; les bras tendus entre les cuisses ; la tête jetée en arrière avec émoi ; l'expiration à bout de souffle » [19]. Nous pouvons dire que mettre ensemble ces deux différents mondes, celui de la construction formelle de la pièce, de la permutation et de la combinatoire continue des structures musicales et de la danse, et l'autre celui des gestes courts, quotidiens et presque connotés, constitue cette dernière tension, mais aussi donne une particularité, un style propre à l'œuvre.

Un mouvement d'un autre ordre complète les interactions entre musique – danse – film, c'est le chemin tracé par les caméras cinématographiques sur le bâtiment moderniste où se réalise le film [9]. Celles-ci vont donner à la temporalité musicale (et aussi corporelle) un caractère spatial, en créant des perspectives, des jeux de figure et fond, ainsi que des polyphonies et des amplifications des éléments de la chorégraphie et de la danse. Par exemple, le grand parcours sur les vitres du bâtiment où se superposent les

danseuses, les différentes séries de mouvements, les reflets et la lumière des différents points de vue choisis.

Dans *Tippeke*<sup>8</sup>, film et musique pour violoncelle et électronique, la manière choisie pour construire la convergence est différente, il y a un élément extérieur à la musique et à la danse qui organise et concentre les relations entre les deux media. Le texte, la comptine flamande, raconté par Anne Teresa De Keersmaeker opérera cette convergence.

«La protagoniste, Anne Teresa De Keersmaeker raconte/chante/danse une vieille comptine flamande où il est question d'un petit garçon qui ne veut pas rentrer à la maison. L'histoire déplie un enchaînement récursif de neuf menaces successives : *un p'tit chien qui ne veut pas mordre Tippeke, un bâton qui ne veut pas frapper le p'tit chien, un feu qui ne veut pas brûler le bâton, etc...* ; pour arriver à la chute en cascade de dominos où l'on repart en ordre inverse : *oui le p'tit chat veut bien bouffer la souris, et la souris de sauter sur la corde pour la ronger, et la corde de sauter sur la vache pour l'attacher, etc.* » [15].

La danse est travaillée à partir de chaque élément du texte - le chien, le bâton, le feu, etc. - qui correspond à un mouvement déterminé, et qui est exécuté simultanément à la narration. Chaque élément correspond aussi à un son multiphonique<sup>9</sup> du violoncelle (**Figure 1**) et chaque caractère du texte - affirmatif, interrogatif, négatif - une manière de les exécuter. Les consonnes sont ponctuées par des modes de jeu percussifs du violoncelle, qui détachent aussi quelques mouvements de la danse.

Au-delà de la convergence entre texte - écriture instrumentale / texte - danse, ce que Thierry De Mey appellerait des relations en triangle [9], il en existe d'autres, qui sont produites par le travail avec des technologies numériques. Par exemple, les données fréquentielles de la voix et les sons multiphoniques sont utilisés pour filtrer les ambiances sonores du film; les ambiances sont aussi manipulées en fonction de procédés cinématographiques, comme les glissandi qui accompagnent l'éloignement de la caméra. Le compositeur a fait appel aussi à des processus de composition assistée par ordinateur, pour analyser et générer le matériel mélodique du violoncelle à partir des multiphoniques.

Donc, nous pouvons parler d'au moins deux niveaux de convergence, le premier généré par la relation en triangle qui va permettre l'interaction entre l'écriture instrumentale et la création chorégraphique. Ici sont

mis en place les différentes formes de lecture et d'appropriation du texte, chacun des deux moyens reste avec sa propre spécificité ; cependant la présence de l'élément étranger aux deux concentre et organise les opérations constitutives au niveau temporel.

**Figure 1.** Extrait de la partition de *Tippeke*, les multiphoniques sont les notes en forme de diamant.

L'autre niveau construisant une convergence de type morphologique, se base sur le son de la voix d'Anne Teresa De Keersmaeker et sur celui des multiphoniques du violoncelle même qui interagit avec la bande sonore du film. C'est la matière sonore qui va converger et se mélanger. L'utilisation du traitement du son par ordinateur est liée aussi à la question de la morphologie sonore et son interaction avec le monde de l'écriture instrumentale.

*Water* (2002), pour six percussionnistes, sons échantillonnés, violoncelle et électronique, a été écrit originalement pour le spectacle *April Me* d'Anne Teresa De Keersmaeker et fait aussi partie du projet cinématographique *Counter Phrases*. C'est un autre exemple où un élément synthétique fait la liaison entre la musique et la danse. Ce sont alors les structurations temporelles et rythmiques, qui à la fois génèrent la musique et la danse.

D'après la note de programme de l'œuvre, l'idée est de construire une superposition de canons lesquels suivent une logique pendulaire. Dans cette logique, nous avons des structures rythmiques qui se "contractent où se dilatent en leurs extrémités ou en leurs centres", elles accélèrent ou décelèrent à ces points temporels. Après superposition progressive, nous observons un mouvement complexe de phase et déphasage qui conduit vers « l'unisson » total à la fin de la pièce [11].

Cette logique opératoire est utilisée pour la musique et pour la danse, donc il s'établit un parallélisme temporel entre les deux. Parallèlement aux couches de pendules rythmiques, à travers la répétition des phrases chorégraphiques, les danseurs forment leurs propres jeux de phase et déphasage. La localisation dans l'écran, la quantité de danseurs qui constituent chaque groupe

<sup>8</sup> La musique de *Tippeke* est une commande de l'IRCAM et a été réalisée avec la collaboration de François Deppe, assistant à la composition et Serge Lemouton, assistant musical. La musique et le film ont fait partie de le spectacle *Woud - Three Movements to the music of Berg, Schönberg and Wagner*, une chorégraphie de Anne Teresa De Keersmaeker interprétée par la Compagnie Rosas.

<sup>9</sup> Normalement la technique du multiphonique dans le violoncelle consiste en effleurer la corde avec la pression nécessaire pour produire un son harmonique, mais sur une position intermédiaire. Dans *Tippeke*, ce terme fait référence à l'utilisation de deux points de effleurage dans une unique corde, un avec la pression ordinaire, l'autre avec la pression de harmonique, produisant un son composé par deux différents hauteurs proéminents.

répétant le même répertoire de mouvements et la superposition de différents groupes donnent une dynamique globale très vivante qui produit un effet de différentes temporalités distribuées dans l'espace. La disposition du film en triptyque (avec trois écrans simultanés) met en évidence cet aspect polyphonique et effectue aussi des ampliatiions de points de vue déterminés dans la trajectoire formelle du film.

On observe aussi un échange constant de données dans certaines opérations musicales, le compositeur ayant utilisé les données fréquentielles des sons échantillonnés comme information temporelle. Par exemple, un son qui a comme fréquence fondamentale 220 Hz est répété en accélération jusqu'à 220 fois par seconde et est rejoué aussi en parcours inverse, comme un son continuo, répété 220 fois par seconde, en décélération jusqu'à une pulsation [7]. Ces sons échantillonnés sont des instruments qui sont joués par les percussionnistes, par exemple les cloches de vache. D'un côté, nous avons une approche granulaire avec des processus sur le temps multi-échelle et d'un autre côté le choix d'utiliser le son instrumental comme une recherche d'intégration entre mondes instrumental et électroacoustique.

Nous pouvons remarquer que la plupart des opérations effectuées par De Mey dans les œuvres mentionnées ci-dessus concernent l'élaboration de structures temporelles, lesquelles servent à la fois la musique et la danse. Nous observons trois manières de les construire, à travers la simultanéité de phrases musicales et de la danse (*Rosas danst Rosas*), de la présence d'un élément extérieur (*Tippeke*) et d'une forme organisatrice (*Water*) qui conduisent les deux moyens, ces opérations créent les interactions globales et ponctuelles et réalisent la convergence entre ces deux media.

Thierry De Mey développe une autre approche de la convergence entre différents moyens à travers ses pièces qui associent gestes et musique, comme *Hands* (1983), *Musique de tables* (1987), *Unknownness* (1996), *Silence must be!* (2002) et *Light Music* (2004). Ce parcours réunit diverses préoccupations présentes dans la poétique du compositeur, premièrement nous pouvons parler de la recherche d'une écriture exploitant les mouvements de l'interprète. Si nous regardons le cas de *Musique de tables*, par exemple, le compositeur élabore un vocabulaire de gestes qui est exploité pendant toute la pièce avec des variations et combinaisons diverses; ainsi sont créés des contrepoints et des complémentarités. Ces gestes sont nommés dans la partition. Par exemple, le 'revers' est la dénomination de l'action de la frappe de la table par le revers de la main. Dans ce cas on peut souligner que presque tous les gestes produisent des sons, eux même produisant des figures spatiales comme les cercles et les balançoires.

Dans *Silence must be!* la continuité entre mouvement et son est bouleversée. Dans la partition il y a beaucoup d'indications d'actions de caractère théâtral que l'interprète exécute à la scène, par exemple la première indication de la pièce est : "le chef fait face au public, il

est éclairé de manière à mettre les mouvements de ses mains en évidence. Il se prend le pouls au poignet gauche au moyen du pouce de la main droite, la main gauche tournée paume vers le haut" [12]. Ensuite, l'interprète exécute les gestes de chef en silence et seulement à la deuxième partie est établi la relation entre ces gestes et le son.

*Light Music*<sup>10</sup> est le point culminant de cette démarche, nous pouvons y observer une intéressante interaction entre le travail de l'écriture, le développement du dispositif audiovisuel et l'interprétation. Le projet a été développé par Thierry De Mey en collaboration avec Jean Geoffroy, l'interprète, et Christophe Lebreton, concepteur du dispositif interactif.

Chaque partie de la pièce va exploiter un type différent d'écriture musicale et d'interaction entre les gestes de l'interprète et le résultat sonore et visuel. Par exemple, la première partie est composée des gestes flous qui tracent des chemins dans le sens horizontal et vertical de la scène avec des ponctuations d'autres gestes très courts, le résultat sonore suit les gestes, avec l'alternance entre sons longs et sons courts ; le visuel va seulement reproduire la trajectoire de ces gestes. La deuxième partie est composée de beaucoup de gestes empruntés à *Silence Must be!* alternés par des trajectoires floues qui dialoguent avec la première partie. Dans la section finale s'impose une continuité entre sons et gestes, l'interprète va déclencher des sons percussifs avec les mêmes gestes silencieux du début de cette partie. La troisième partie est entièrement marquée par une sorte de peinture de lumière, l'écriture indique des formes que l'interprète doit dessiner dans le vide : des parenthèses, des serpents, des guirlandes, des labyrinthes, etc. Chaque forme dessinée correspond aussi à une forme sonore spécifique, qui est manipulée en fonction d'où commence chaque action dessinée, sa vitesse, son parcours horizontal et/ou verticale, etc.

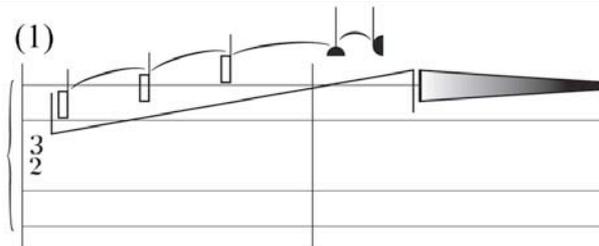
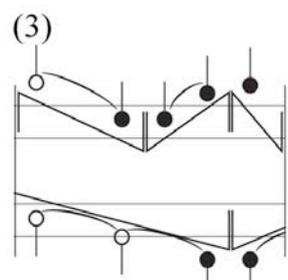
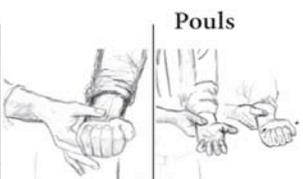
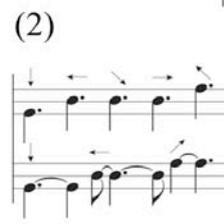
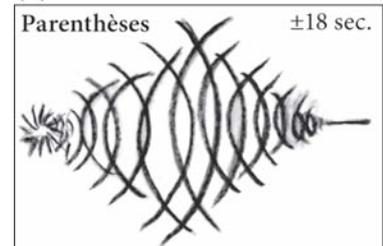
Dans la table suivante (**Table 1**, page suivante), nous présentons les interactions entre les gestes prédominants dans chaque partie de la pièce et ses résultats sonores et visuels. Nous choisissons dans ce cas de partir des aspects perceptifs dans la partition et de l'enregistrement vidéo de la pièce. Donc, nous parlons des interactions perceptibles entre gestes et images et gestes et sons.

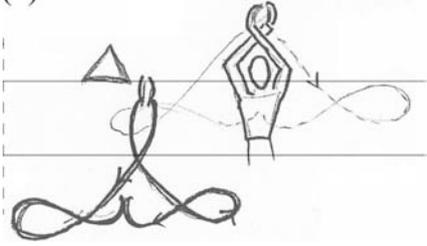
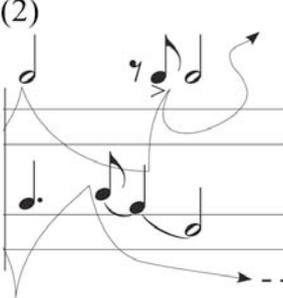
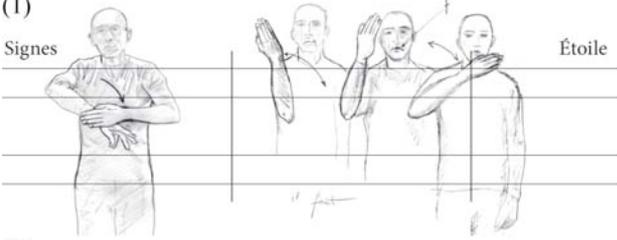
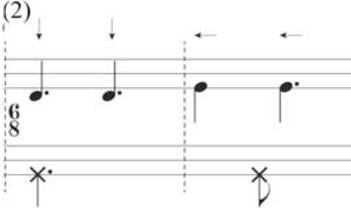
Si nous pensons en terme d'espace composable, en tant que l'ensemble des variables, des opérations, de traitements et aussi des contraintes avec lesquelles on compose, que nous indique cet exemple ? Nous commençons l'exposé sur De Mey en disant qu'il travaille avec la construction d'un vocabulaire et d'un système d'équivalences; si nous regardons la table suivante, nous constatons que chaque partie explore un vocabulaire spécifique de gestes, qui procure des résultats sonores et visuels en relation avec le dispositif interactif. En regardant en détail chaque partie nous comprenons qu'il existe plusieurs niveaux de relation

<sup>10</sup> *Light Music* est une coproduction entre Charleroi-Danses et GRAME.

entre les composants de la pièce : la seule causalité du geste provoque un son ; le *mapping* entre un ou plusieurs variables gestuelles contrôle des paramètres sonores ; la similarité visuelle du geste et ses conséquences sonores, où la ressemblance entre les deux est donnée préalablement par l'écriture ; le même traitement entre les images et le son; et la réutilisation du matériau d'autres œuvres. On peut noter qu'un niveau n'exclut pas l'autre, ils peuvent apparaître ensembles dans un même contexte, c'est justement la superposition et la variation de ces différents niveaux qui garantissent un caractère dynamique au déroulement temporel de *Light Music*, par exemple l'accumulation

graduelle de boucles dans la partie IV parallèle à l'utilisation des échantillons percussifs, ou les différents formes de *mapping* dans les dessins de la partie III. D'une façon générale les opérations qui habitent cet espace composable sont de l'ordre d'un travail sur la simultanéité et la causalité entre les trois moyens. Cela consiste à connecter les causes gestuelles et leurs conséquences sonores et visuelles dans différentes formes possibles, et aussi à les faire évoluer et se transformer ensembles dans le temps, à établir des interactions.

<p>(1)</p>  <p>(2)</p>  <p>(3)</p> 	<p><b>Partie I</b></p> <p><b>Mouvement [partition] :</b> combinaison entre mouvement de rotation des mains avec déplacement vertical et horizontal (1) ; rotations très rapides sans déplacement horizontal et vertical (2) ; même position de main avec déplacements rapides horizontaux et/ou verticaux (3).</p> <p><b>Interaction entre gestes/sons [dispositif] :</b> la vitesse dans l'axe horizontal et vertical opère les changements de paramètres sonores, comme l'altération de la fréquence plus préminente dans les filtres et résonateurs ; opposition sonore entre les gestes flous et lents et entre les gestes courts et ponctuels.</p> <p><b>Interaction entre images/gestes [dispositif] :</b> le dispositif reproduit seulement les gestes effectués dans la lumière.</p>
<p>(1)</p> <p>En silence, hors tempo Prendre pulsation cœur sur poignet gauche</p>  <p>Pouls</p> <p>(2)</p>  <p>(3)</p> 	<p><b>Partie II</b></p> <p><b>Mouvement [partition] :</b> gestes scéniques (1) et de chef (2) empruntés de <i>Silence must be !</i> ; dessins de formes spécifiques dans le vide : la mer (3) et la sphère.</p> <p><b>Interaction entre gestes/sons [dispositif] :</b> gestes silencieux (sans interaction) ; dans le dessins la vitesse opère les changements de paramètres des filtrages ; les gestes de chef déclenchent les échantillons de sons percussifs.</p> <p><b>Interaction entre images/gestes [dispositif] :</b> le dispositif reproduit les gestes effectuées dans la lumière et laisse ses traces.</p>
<p>(1)</p> <p>Parenthèses ±18 sec.</p> 	<p><b>Partie III</b></p> <p><b>Mouvement [partition] :</b> plusieurs dessins de formes spécifiques : parenthèses (1), fil tiré (2), spirales, etc.</p> <p><b>Interaction entre gestes/sons [dispositif] :</b> l'interaction est effectuée par plusieurs paramètres groupés: vitesse, taille et direction. Par exemple : la vitesse de la 'double hélice' modifie l'intensité sonore, chaque point d'arrêt</p>

<p>(2)</p> <p>Fil tiré impro ±40 sec.</p> 	<p>dans la 'descente rauque' génère un déclic, dans quelques gestes c'est établie la relation de l'axe horizontal (haut – bas) et l'espace fréquentiel (aigu – grave), comme dans les 'labyrinthes', etc.</p> <p><b>Interaction entre images/gestes [dispositif] :</b> le dispositif suit la trajectoire des mains de l'interprète et laisse ses traces pour quelques secondes, générant une sorte de peinture de lumière.</p>
<p>(1)</p>  <p>(2)</p> 	<p><b>Partie IV</b></p> <p><b>Mouvement [partition] :</b> dessins (1) et mélange entre des gestes mécaniques et des trajectoires spatiales indiquées par flèches, qui donnent une caractéristique plus 'plastique' à ces gestes (2).</p> <p><b>Interaction entre gestes/sons [dispositif] :</b> dans le dessin, il y a un déclenchement des enchantions de textures sonores qui sont répétés en boucle et se superposent avec les autres; les gestes mécaniques déclenchent d'autres échantillons, principalement de sons percussifs.</p> <p><b>Interaction entre images/gestes [dispositif] :</b> la trajectoire de la peinture de lumière des dessins est répétée en boucle, chaque nouveau dessin est superposé avec les dessins antérieurs ; dans les gestes mécaniques le dispositif reproduit le gestes exécuté dans la lumière; à la fin de la partie l'interprète intervient avec son corps entier dans l'image.</p>
<p>(1)</p> <p>Signes</p>  <p>Étoile</p> <p>(2)</p> 	<p><b>Partie V</b></p> <p><b>Mouvement [partition] :</b> écriture dans le vide du mot 'SILENCE'; langage de signes de deux phrases : « il faut avoir un chaos à l'intérieur de soi pour enfanter une étoile qui danse » F. Nietzsche (1) et "silence must be" ; gestes de chef (2).</p> <p><b>Interaction entre gestes/sons [dispositif] :</b> changement de la bande sonore synchronisée avec le langage de signes (indiqué dans la partition), les gestes du chef déclenchent les échantillons de sons percussifs.</p> <p><b>Interaction entre images/gestes [dispositif] :</b> dans toute cette partie le dispositif produit la peinture de lumière des gestes.</p>

**Table 1.** Différentes interactions présentes dans *Light Music*.

En soulignant l'usage de la notion de mouvement comme moyen principal de convergence, on peut observer que l'extension intermédiaire du musical dans les œuvres de Thierry De Mey, s'effectue en premier lieu dans une convergence avec la danse, ceci en appliquant le même genre d'opérations entre les deux moyens, lesquels font partie d'une pensée temporelle qui va permettre au compositeur d'élargir son champ de travail vers ces extensions intermédiaires. La vidéo va aussi contribuer à cette démarche, amplifiant des convergences et créant des points de vue particuliers

dans la trame des polyphonies de mouvements, qu'ils soient dansés, ou musicaux. Le numérique contribue encore à cette convergence en permettant des interactions entre les différentes morphologies sonores, entre l'écriture instrumentale et l'électronique et aussi entre les gestes, les sons et les images comme nous remarquons dans *Light Music*.

Nous poursuivons notre questionnement en analysant la démarche de Georges Aperghis, et en nous focalisant sur les dispositifs utilisés dans ses dernières œuvres.

#### 4. GEORGES APERGHIS, LE DISPOSITIF SCÉNIQUE COMME MOYEN DE CONVERGENCE INTERMÉDIALE

Quant à l'extension du musical à un autre moyen d'expression, on doit pour commencer souligner qu'Aperghis a depuis longtemps étendu le musical au domaine du théâtre, en faisant converger l'écriture musicale et écriture théâtrale. L'intégration des technologies numériques sur scène, pour ainsi dire leur "mise en scène", l'amène à recourir au dispositif scénique de type multimédia.

La poétique d'Aperghis évoque déjà des convergences et interactions entre des éléments hétérogènes. Le compositeur va concevoir le théâtre musical comme une véritable polyphonie ; les composants du spectacle (costumes, lumière, décors, etc.) qui dans l'opéra ou le théâtre ancien ont été pensés en fonction d'une « lecture musicale d'un livret » proposant des situations dramatiques linéaires – du type cause et effet – sont libérés pour être organisés d'une façon autre. Le récit est aussi bouleversé, il n'y a que des fragments de textes ou de « contrepoints d'histoire diverses (qui se tissent entre elles) » [2].

« J'ai envie de déconnecter la cause et les effets. De dire que l'action 2 n'est pas le résultat de l'action 1, que ça n'a rien à avoir. Dans le théâtre traditionnel européen, si quelqu'un fait une action, c'est parce que c'est une réponse à l'action précédente, c'est ce qu'on appelle la psychologie, le personnage se comporte d'une façon psychologique, donc il y a une relation d'enchaînement psychologique [...]. Ce qui m'intéresse c'est d'arriver à déconnecter au maximum tous les composants et les remettre ensemble autrement, grâce au dispositif » [6].

Dans cet effort de déconnecter et de libérer des composants du théâtre musical, Aperghis va aussi émanciper le corps des interprètes à travers sa propre écriture, qui deviendra génératrice de comportements physiques. C'est là une caractéristique très marquante de ses pièces vocales. Nous pouvons citer les pièces comme *Récitations* (1978) ou *Quatorze jactations* (2001), qui sont très éloignées dans le parcours du compositeur, mais qui à la fois concrétisent cette démarche dans laquelle l'exploitation de divers modes de jeu vocaux construit une écriture incarnée. On entend dans ces pièces des rythmes créés par constantes alternances entre les différents comportements vocaux : entre le chant, les murmures, le rire, la voix inhalée parlant de plusieurs manières différentes. Nous pouvons dire qu'il existe des interactions entre cette écriture et les textes, c'est qu'Aperghis va appeler le « système des masques » [16], le compositeur va appliquer des modes de jeu vocaux, mais aussi gestuels, dans un texte compréhensible pour bouleverser sa compréhension. Il y a donc il y a une confrontation entre le texte et ces « masques » sonores et gestuels.

Nous proposons d'aborder la question de l'extension du musical dans trois œuvres d'Aperghis : *Machinations* (2000), *Avis de tempête* (2004) et *Luna Park* (2011)<sup>11</sup>.

Dans ces œuvres récentes, Aperghis propose un dispositif scénique qui concentre en soi même le son, les images, le texte et les actions des interprètes. On pourrait assimiler ces dispositifs à des espaces scéniques composés, porteurs du potentiel de l'œuvre. Aperghis les assimile pour ainsi à des partitions.

« Un dispositif qui permet de structurer, de distribuer, le son, le comportement [des interprètes], des images et le texte. Ça c'est vraiment un dispositif qui à la fois est un dispositif musical, parce que c'est la musique qui fait tout ça. Donc, ce qui s'appelle la partition est en même temps un dispositif scénique qui correspond à la partition. C'est une espèce de double. Un dispositif scénique qui permet à la partition de fonctionner, pas seulement au niveau sonore, mais aussi visuellement. Donc, ce dispositif là est une chose très importante pour moi. Il faut trouver la chose la plus économique, la plus simple, même si ça va apparaître très compliqué, pour moi c'est la plus simple, qui distribue tout ça. Qui permet à tout ça d'exister » [6].

*Machinations* (2000), pour quatre voix de femmes, électronique et vidéo, met en place toutes les questions abordées antérieurement par Aperghis. Le dispositif scénique, développé par Daniel Lévy, est composé de quatre tables avec des lumières et caméras, qui enregistrent les mouvements de mains et des objets qui sont posés sur elles et les projettent dans les quatre écrans derrière les interprètes (**Figure 2**, page suivante).

Olivier Halévy [18] parle de quatre différents types d'articulation entre les différents moyens de *Machinations*: amplifications, quand les images concrétisent le contenu du texte ; déformations, quand une des interprètes reprend le matériau de l'autre d'une façon différente ; dissociation, quand « l'objet manipulé est méconnaissable dans l'image projetée à cause du changement d'échelle et d'angle de prise de vue »; explicitation, « lorsque les images et les phonèmes rendent visibles certains aspects implicites des textes logiques ». Ce sont là des exemples des interactions entre les comportements des interprètes, les images et le texte, qui convergent grâce au dispositif. On peut noter aussi que *Machinations* est la première pièce dans laquelle Aperghis va utiliser l'électronique, qui a été développée par Olivier Pasquet et Tom Mays et a été pensée comme un agent transformateur. Le compositeur va parler du dispositif informatique musical comme d'une sorte de « parasite » [3], lorsqu'il casse, transforme et multiplie les voix. Ceci est renforcé par la présence des assistants musicaux sur scène qui interviennent sur des situations présentées, comme des 'machinistes', qui suspendent le flux musical et même le

<sup>11</sup> *Machinations* est une coproduction de l'IRCAM-Centre Pompidou, Westdeutscher Rundfunk Köln, *Luna Park* est une production de l'IRCAM-Centre Pompidou et *Avis de tempête* est une production de l'Opéra de Lille, coproduction Lille 2004, Opéra de Nancy et de Lorraine, IRCAM-Centre Pompidou.

sens des événements et les guide selon leur propre volonté.

En écoutant et en regardant *Machinations* nous constatons que la convergence entre le sonore, le visuel et le corporel se produit à travers une conduite musicale. Chaque action sur les tables doit avoir sa propre enveloppe temporelle, c'est à dire, commence et s'arrête sur des points spécifiques avec une vitesse et intensité déterminées. Ainsi sont construits des contrepoints entre le flux de phonèmes et les actions – des simultanités, oppositions et complémentarités – et aussi entre chaque

action projetée dans les écrans. Le texte est traité à partir d'une logique similaire, par exemple il y a un processus d'ornementation d'une voyelle par des consonnes, et même si un texte intelligible est mis en place il peut être dirigé vers le musical, avec des accents sur des mots spécifiques, avec des variations sur le mode de jeu vocal de chaque mot, etc. C'est justement en composant cette polyphonie sonore, visuelle, corporelle et aussi de sens qu'Aperghis opère une extension du musical.



Figure 2. Dispositif scénique de *Machinations* (photographie de Agnès Fin, IRCAM).

*Avis de tempête*, opéra, présente un dispositif scénique composé de sept écrans mobiles, qui rappellent des cerfs-volants suspendus sur la scène et une tour centrale avec des lumières et des caméras. Ceci a été développé par Peter Missotten et Kurt d'Haeseleer (de Filmfabriek). Au-delà de ces caméras, l'actrice/danseuse Johanne Saunier porte deux mini-caméras sans fil dans ses mains. Au centre nous avons l'ensemble instrumental et les chanteurs au fond.

Le dispositif est une « machine d'observation » [17] qui capte le mouvement des interprètes sur scène et les distribue sur les écrans. Comme chaque caméra enregistre un point de vue spécifique, les images ne sont pas identiques, les mouvements sont toujours autour du centre et en chaque écran est projeté une partie de ce parcours. Cette trajectoire répétitive donne aux images un caractère de « système autonome, vivant, qui a sa propre logique ». Quant à Johanne Saunier, elle manipule des caméras de façon à produire des images toujours en mouvement. Les images des interprètes sont mixées avec d'autres, et aussi entre elles-mêmes, créent des nouvelles possibilités de celles réalisées sur la scène.

La partie musicale électronique a été composée en collaboration avec Sébastien Roux avant la partie

instrumentale, ce sont des séquences qui sont déclenchées pendant le spectacle qui ponctuent et contrastent avec la partie de l'ensemble et des voix. Les claviéristes vont aussi diffuser les sons de la partie électronique.

Nous avons des interactions dans *Avis de tempête* entre images et position spatiale des interprètes, entre les mouvements de la danseuse et les images, entre l'électronique et l'écriture instrumentale. Le livret est pour ainsi dire un assemblage de fragments divers, comme pour *Machinations*. En fait, dans *Machinations* et *Luna Park* aussi, il n'y a pas une idée totale de l'œuvre préconçue, mais plutôt un ensemble de fragments indépendants de musique, de texte et d'images qui sont agencées pendant le processus de montage du spectacle. Par exemple, dans *Machinations* la partition définitive sur support papier est construite après la création de la pièce en tant que dispositif, c'est la trace visible résultant de la trajectoire de l'œuvre. Dans cette partition on retrouve le texte, les indications de caractère pour la voix, et les *cues* de l'électronique, il n'y a beaucoup de détails de la création de la pièce. Se pose alors une question: comment faire pour noter cette musique, qui est composée d'un dispositif d'éléments sonores, visuels, textuels et aussi des comportements

physiques. Ou du moins, comment faire pour exploiter la pièce par le moyen de la partition traditionnelle.

Dans *Luna Park*, le dispositif scénique développé aussi par Daniel Lévy est composé par quatre compartiments dans lesquels les interprètes sont

confinés, chacun étant équipé de caméras de surveillance, des lumières et d'un écran de façade ; derrière les quatre compartiments il y a également un grand écran (**Figure 3**). Nous avons aussi deux capteurs de mouvement dans les mains de Richard Dubelski.

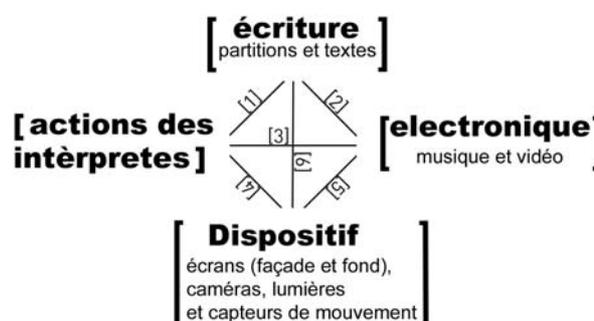


**Figure 3.** Dispositif scénique de *Luna Park* (photographie de Sylvia Gomes, IRCAM).

Le diagramme suivant (**Figure 4**, page suivante) présente le réseau d'interactions entre les composants de *Luna Park*, nous pouvons citer quelques exemples : (1) soit à travers les partitions, soit à travers les textes, l'écriture est génératrice des actions des interprètes, il y a l'inverse aussi, Johanne Saunier va construire ses mouvements à partir des caractéristiques des textes, par exemple, une séquence de mouvements qui suit des répétitions des mots ou des phrases ; (2) ces actions sont conçues pour se mettre en place dans les compartiments du dispositif ; (3) l'électronique et l'écriture interagissent par le traitement en direct des flûtes et de la voix et par les capteurs de mouvement qui déclenchent des échantillons sonores, manipulent les paramètres de la synthèse vocale et de la transformation prosodique [1] ; (4) ces capteurs et le traitement en direct permettent également l'interaction entre des actions et l'électronique musicale, et les caméras créent l'interaction des interprètes avec la vidéo ; (5) la disposition des caméras permet de construire plusieurs points de vue des actions et des corps des interprètes, la qualité des tissus des écrans créent des superpositions entre des images projetées et les actions sur la scène, le grand écran au fond crée des perspectives ; (6) les lumières du dispositif transforment des images des interprètes enregistrées en direct.

Comme nous le disions plus haut, dans la création de *Luna Park* la plupart des éléments sont agencés dans le processus de montage du spectacle, il n'y a pas une globalité préconçue de la pièce. Donc on observe qu'il n'y a pas une hiérarchisation des interactions et une logique temporelle de conception de l'œuvre. C'est ce dispositif qui établit les possibilités de relation entre

tous les composants, entre les images, les sons et les actions et donc la convergence musicale entre les media. Georges Aperghis déjà a effectué une extension du musical à partir de son propre travail d'écriture, en rencontrant une manière de faire converger la musique et le théâtre, soit pour incarner l'écriture musicale de comportements physiques, soit pour obtenir la musicalité d'une écriture textuelle. Avec l'électronique et le dispositif scénique multimédia, il peut étendre sa démarche musicale en créant de nouvelles possibilités d'interaction entre les éléments scéniques présents dans les polyphonies de ses œuvres. Ainsi des images peuvent suivre, ou conduire, des actions des interprètes; des caméras et des écrans peuvent spatialiser des comportements physiques musicaux sur scène; l'électronique peut multiplier et transformer des voix et des instruments, etc.



**Figure 4.** Différentes interactions présentes dans *Luna Park*

## 5. CONCLUSION

Du point de vue du musicien, le recours à l'intermedialité peut passer par l'extension de l'approche musicale à d'autres domaines d'expression artistique : vidéo, danse, théâtre, arts de la scène ; autant de moyens que l'introduction du numérique aura permis de faire converger avec l'écriture musicale. On a essayé de montrer ici l'approche de deux compositeurs qui ont, l'un avec la danse, l'autre avec le théâtre, trouvé des voies singulières pour arriver à des formes propres, des formes mixtes, étendues à d'autres media. En interaction avec la danse, De Mey a tiré de la notion de mouvement les moyens de faire converger de façon musicale d'autres media donnés par le numérique (recours à l'informatique musicale, à la vidéonumérique). Au-delà de la première étape du théâtre musical, Aperghis, en définissant des dispositifs scéniques-partition intermediales, a su renouveler la notion même de partition musicale, de l'écriture à l'interprétation de l'œuvre, là encore en intégrant les moyens offerts par les technologies numériques de la scène.

Sur un plan méthodologique, cette étude est une première approche. On devrait pouvoir aller encore plus loin dans l'examen des manières de faire des compositeurs, et de leurs environnements conceptuels et espaces composites.

Comme on l'a dit en introduction, nous nous intéressons aux approches recherche/création et comment ces deux activités s'alimentent l'une l'autre, pour avancer sur le terrain de la composition musicale. En quoi examiner les manières de faire d'un De Mey ou d'un Aperghis nous est-elle utile dans la recherche comme dans la création ? Elle nous permet de faire ressortir quelques notions qui contribuent à faire émerger une épistémologie de la composition musicale, comprise comme l'ensemble des concepts qui alimentent une communauté de pensée dans un style donné. Ici la communauté des compositeurs qui manient la mixité en l'étendant à d'autres media, soit déjà existants, la danse, le théâtre, et/ou se déployant avec le développement expérimental puis massif du numérique.

Qu'est-ce qui dans l'approche des compositeurs peut faire converger musicalement plusieurs media ? Qu'est-ce qui est musicable ? De Mey identifie et traite dans cette perspective le mouvement comme moyen, manie la multitemporalité, la rythmicité, la mise en espace. Aperghis procède musicalement à l'articulation des divers media dans le dispositif scénique, et c'est ce qui constitue la partition de l'œuvre. Dans le contexte de l'intermedialité, comment se construit le musical ? Voilà une question simple à laquelle la musicologie, irriguée par l'examen de la création musicale, devrait pouvoir répondre.

## 6. REFERENCES

- [1] APERGHIS, G., BELLER, G. « Contrôle gestuel de la synthèse concaténative en temps réel dans Luna Park », *Rapport de recherche et développement*, IRCAM, Paris, 2011. Disponible en ligne.
- [2] APERGHIS, G. « Quelques réflexions sur le théâtre musical », in. GINDT, A. *Georges Aperghis: le corps musical*. Actes Sud, Paris, 1990, pp. 61.
- [3] APERGHIS, G. *Machinations*. Paris, IRCAM-Centre Pompidou, 2002, pp. 8 (note de programme).
- [4] BENJAMIN, W. « L'oeuvre d'art à l'époque de sa reproduction mécanisée », in. *Ecrits Français*, Folio Essais, Paris, 1991, pp. 147-248.
- [5] CARVALHO, G. « Formaliser la forme », Journée d'Informatique Musicale, Paris, 2005. Disponible en ligne.
- [6] CASTELLANI, F. M. *Entretien avec Georges Aperghis*. Paris, 2012.
- [7] CASTELLANI, F. M. *Entretien avec Thierry De Mey*. Enghien-les-Bains, 2012.
- [8] DEBRAY, R. *Introduction à la médiologie*. Paris, PUF, 2000.
- [9] DE MEY, T. « How to know the dancer from the dance... », in. LAUXEROIS, J., SZENDY, P. *De la différence des arts*. L'Harmattan/ IRCAM, Paris, 1997.
- [10] DE MEY, T. *Light Music*. 2007 (partition).
- [11] DE MEY, T. *Pièces électroniques: scies, bonds, empreintes...* Bruxelles: Sub Rosa, 2004 (livret du CD).
- [12] DE MEY, T. *Silence Must be!*. 2002 (partition).
- [13] DE MEY, T. *Tools for movement composition: memo of workshop*. Bruxelles, 2007.
- [14] DE MEY, T. *Tippeke*. 1996 (partition).
- [15] DE MEY, T. *Un compositeur, une œuvre: Thierry De Mey, Tippeke, film et musique, pour violoncelle et électronique*. IRCAM, Paris, 1998 (note de programme).
- [16] GINDT, A. *Georges Aperghis: le corps musical*. Actes Sud, Paris, 1990, pp. 89.
- [17] D'HAESELEER, K. « Dispositif scénique: rencontre avec Kurt D'Haeseleer », in. HOUDART, C. *Avis de Tempête: Georges Aperghis*. Intervalles, Paris, 2007, pp. 27.
- [18] HALÉVY, O. « Le dispositif et l'événement: les nouvelles technologies sur la scène théâtrale », *Musica Falsa*, No. 18, MF, Paris, 2003, pp. 28.
- [19] PLOUVIER, J-L. « Fragments fibonnaciens: Anne Teresa De Keersmaecker et la musique », in. ADOLPHE, J-M., JANSEN, S., LUYTEN, A. et al. *Rosas: Anne Teresa de Keersmaecker*. La Renaissance du Livre/ Rosas, Tournai/Bruxelles, 2002, pp. 279.
- [20] MCLUHAN, M. *Understanding media*. Mc Graw-Hill Book Company, New York, 1964.
- [21] VAGGIONE, H. « Perspectives de l'électroacoustique », *Chimère*, n° 40, 2002, pp. 57-67. Disponible en ligne.
- [22] VALÉRY, P. « Discours prononcé au deuxième congrès international d'esthétique et de science de l'art » (1937), in. *Variété* (4), Gallimard, Paris, 1939, pp. 235-265.

# LES ORIGINES DU NOM DE RIM (REALISATEUR EN INFORMATIQUE MUSICALE)

Laura Zattra

Università di Padova, Dipartimento dei Beni Culturali  
Laboratoire STMS (Ircam-CNRS-UPMC), équipe Analyse des pratiques musicales  
[laura\\_zattra@yahoo.it](mailto:laura_zattra@yahoo.it)

## RÉSUMÉ

La qualification de Réalisateur en Informatique Musical (RIM) – le plus proche collaborateur d'un compositeur accueilli en résidence dans un laboratoire de recherche – est aujourd'hui familière mais, au contraire, ce nom est assez récent. La présence de ce professionnel, par contre, est vérifiée depuis l'ouverture des centres de recherche à la production de musique. C'est l'IRCAM (Institut de Recherche et Coordination Acoustique/Musique) qui, dès ses premières années d'activité (l'Institut ouvre en 1977), semble avoir officialisé la fonction spécifique de ce métier. Il nous a semblé important d'étudier les étapes de la formation de ce nom, au cours des décennies.

L'enquête historique a été poursuivie en focalisant l'attention sur les documents publiés et inédits. La recherche bibliographique a permis de répertorier la littérature consacré à l'argument à l'Ircam (CRI, Centre de Ressources de l'Ircam). La recherche dans les archives papier a exploré les projets et rapports d'activité ainsi que les compte rendus de réunions (Direction artistique, Direction générale, Conseil d'administration) dans la période 1976-2012. Les données recueillies ont permis d'éclaircir l'évolution du nom : scientifique (ingénieur / chercheur) ; tuteur ; assistant / tuteur ; assistant musical ; Réalisateur en informatique musicale.<sup>1</sup>

## 1. INTRODUCTION

La naissance du métier de RIM remonte à une période où la *computer music*, après avoir fait la démonstration de la validité esthétique de l'utilisation de l'ordinateur en musique, devient l'objet d'une démarche de recherche systématique. C'est à partir des années 1970 que les studios de recherche musicale offrent aux compositeurs la possibilité de travailler avec les informaticiens pendant le temps même de leur activité créatrice. L'assistant montre au compositeur les capacités de tel ou tel outil, application, effet sonore (si

le compositeur ne connaît pas suffisamment le programme ou s'il n'a pas une idée précise de ce qu'il veut obtenir), lui communique les résultats les plus récents de la recherche informatique, traduit ses idées artistiques dans un langage admissible par l'ordinateur et réalise les opérations techniques, permettant la concrétisation de ces idées et leur finalisation au sein d'une œuvre musicale.

Au cours des années 1960 et 1970, dans la plupart des centres européens et américains – qu'ils fussent dévolus à la musique analogique (Cologne, Milan) ou à la *computer music* (CSC de Padoue, CCRMA de Stanford) –, la collaboration était généralement informelle (les assistants étaient souvent des chercheurs informaticiens, des techniciens du centre ou des compositeurs-informaticiens autonome qui décidaient d'offrir leur compétence aux autres compositeurs). C'est l'IRCAM (Institut de Recherche et Coordination Acoustique/Musique) qui, dès ses premières années d'activité (l'Institut ouvre en 1977), semble avoir officialisé cette fonction spécifique. Pour P.-M. Menger, qui y mène une enquête sociologique au début des années 1980, la « création [y apparaît] comme action collective : le compositeur, l'ingénieur et le tuteur » [4 : p.129].

Au-delà de ces caractérisations très générales, on n'en sait guère plus à ce jour sur ce métier. Malgré sa présence continue dans l'histoire de la production de *computer music*, les témoignages, et a fortiori les études musicologiques, restent excessivement rares. Notre contribution veut répondre à ce manque. Les résultats présentés dans cet article font partie d'un projet plus vaste mené à l'Ircam qui a suivi deux démarches parallèles et complémentaires. L'une, historique, que nous présentons partiellement ici, a eu pour but de tracer l'histoire du nom désignant cette activité ; l'autre, socio-ethnographique, a consisté à entamer une enquête sur le développement de la conscience professionnelle du métier.

La qualification de RIM est aujourd'hui familière mais, au contraire, ce nom est assez récent. Il nous a semblé alors nécessaire d'étudier les étapes de la formation de cet appellation, au cours des décennies, des dernières années 1970 jusqu'à nos jours. On a reconstruit cette évolution d'après les documents publiés et les documents internes conservés à l'Ircam

<sup>1</sup> La recherche a été conduite dans le cadre du projet de recherche « Contribution à l'histoire des interactions arts/science: l'émergence de la réalisation en informatique musicale » développé auprès du Laboratoire STMS (Ircam-CNRS-UPMC), équipe Analyse des pratiques musicales, comme chercheuse invitée CNRS (INS2I).

(CRI, Centre de Ressources de l'Ircam). La recherche dans les archives papier a exploré les projets et rapports d'activité ainsi que certains compte rendus de réunions (Direction artistique, Direction générale, Conseil d'administration) dans la période 1976-2012.

Une recherche de cet ordre n'est pas seulement un exercice philologique ; elle peut aider plus en général l'historiographie de la musique du XXe siècle, en mettant en lumière un phénomène sociohistorique – l'apparition d'une pratique musicale et d'une catégorie socioprofessionnelle nouvelle – dont la signification excède en fait largement le cas particulier de la musique et touche aux enjeux scientifiques et sociétaux les plus actuels relativement à l'articulation entre recherche et scientifique et création artistique.

Dans l'article, on citera les documents inédits en note de bas de page. Les documents publiés par contre feront partie de la liste des références.

## 2. ANNEES SOIXANTE-DIX

### 2.1. Les scientifiques

L'ouverture de l'Ircam entraîna la nécessité d'une collaboration étroite entre les ingénieurs/chercheurs et les compositeurs.<sup>2</sup> Il est bien évident, cependant, qu'au début ni les premiers ni les deuxièmes n'avaient l'idée de donner un titre spécial pour un travail accessoire et qui n'était qu'une des formes du métier d'ingénieur/chercheur. Un document qui date de la fin de 1976, souligne la force intégratrice de la collaboration compositeurs/scientifiques :

«L'I.R.C.A.M. pourquoi ? / Depuis une dizaine d'années, d'importantes découvertes dans les domaines de l'électroacoustique et de l'informatique ont profondément modifié la fonction des compositeurs de musique ; [...] Cette révolution, dont les conséquences sont encore embryonnaires mais n'ont pas fini de s'étendre, doit être maîtrisée. Tel est l'objet de l'I.R.C.A.M., qui se propose : - d'inventorier systématiquement les possibilités

<sup>2</sup> A partir de décembre 1976 l'Ircam devient une association indépendante liée au Centre Georges Pompidou par une convention et dotée d'autonomie financière. L'Ircam commence son histoire à la fin de l'année 1969, quand le Président de la République Georges Pompidou annonce la construction d'un centre culturel sur le site du Plateau Beaubourg. Quelques mois après, le choix est également fait par le Président lui-même de confier à Pierre Boulez la responsabilité du département musical. L'activité est organisée en cinq départements: instruments et voix, électroacoustique, département ordinateur, diagonal (coordination entre les départements) et pédagogie. En 1975 apparaît le livre *La musique en projet*, sous-titré *Le premier livre de l'IRCAM*. Direction Pierre Boulez [1], une série de textes pour la conférence de presse de mars 1974 (Vinko Globokar, Luciano Berio, Jean-Claude Risset, Gerald Bennett, Michel Decoust et un texte de Pierre Boulez). L'avant-propos explique que le centre n'est pas encore né, puisque le Centre Beaubourg n'ouvrira ses portes qu'en 1976. Mais la production sonore et la création musicale ont déjà commencé, par conséquent « il nous paraît important que le public soit associé dès le départ aux réflexions et aux premières recherches qui justifient la création d'un Institut comme l'IRCAM » [Ivi, 8].

nouvelles qu'offrent aux compositeurs et interprètes les techniques scientifiques récentes de production de sons nouveaux ; - de mettre les compositeurs, que leur formation n'a pas préparés à utiliser ces nouvelles ressources, en mesure *d'appréhender la démarche des scientifiques qui en assurent le maniement*, et par un *travail en commun* de l'influencer en vue *d'en tirer le meilleur profit pour la création musicale* ; de diffuser, dans un public de spécialistes et de non-spécialistes.... » (italique par nous).<sup>3</sup>

Dans la période 1977-1979 l'Ircam vivait sa première phase historique [2]. L'identité du centre n'était pas si claire et créait des difficultés (« l'IRCAM ne s'affirme ni comme un institut scientifique ni comme une institution musicale » [3 : 129] et « l'incertitude de la distinction entre recherche et création impose ici très concrètement sa loi » [3: 130].

IRCAM is basically a research institute, not an institute for the production of music. Hopefully, a lot of music will come out of the research : yet Boulez has made it clear that, while research at IRCAM should keep in close contact with the creators and even with the public, it should not be constrained by the pressure of production needs [5].

Nonobstant cette incertitude, le rapport de Gerald Bennett sur l'activité de l'Ircam en 1978 atteste l'importance de la collaboration au cœur des activités, même s'il n'utilise aucun mot pour déterminer un rôle 'officiel' des scientifiques [6]. Bennett se limite à déclarer que « les programmes pour l'oeuvre de [Balz] Truempy [*Wellenspiele*, pour instruments et synthétiseur, oeuvre commandée par l'IRCAM] ont été écrits par G[iuseppe] di Giugno et Jean Kott. La conception des instruments virtuels utilisés sur la machine a été réalisée par Balz Truempy et Neil Rolnick, qui participait également à l'exécution de l'oeuvre » [ibidem]. Dans le rapport sur l'activité de l'Ircam en 1979 les « outils [ont été] imaginés puis construits par les chercheurs, ingénieurs et techniciens de l'IRCAM » [7 : 1].

Ce manque d'un titre spécial est confirmé par les déclarations que m'ont fait les interviewés pendant cette recherche ainsi que par les sources : au début les départements avaient une relative autonomie de gestion et de fonctionnement et les personnes qui aidaient les compositeurs décidaient volontairement de le faire mais, on peut dire, au-delà de leur fonction officielle.<sup>4</sup>

<sup>3</sup> Documents dactylographiés (9 pages) daté 7 octobre 1976, archives internes Ircam. Un texte identique fut envoyé au ministre de l'Intérieur en juin 1977 et eut pour titre *L'I.R.C.A.M. Ses objectifs – son statut – ses activités*.

<sup>4</sup> Serge Lemouton (27 juin 2012) et Andrew Gerzso (19 octobre 2012), l'article [8] publié en 1974 et [2].

### 3. ANNEES QUATRE-VINGT

#### 3.1. Le tuteur

A la fin de 1979, les départs presque simultanés des chefs des départements G. Bennett L. Berio, V. Globokar, M. Decoust et J.-C. Risset, imposent une réorganisation du centre opérée par Pierre Boulez. En novembre, Boulez écrit :

Le moment est venu pour l'Ircam de faire le point. Nous abordons, en effet, après deux ans de fonctionnement, une nouvelle phase de notre action [...]. Deux actions s'imposent à cet effet : 1) Développer et réorienter notre pédagogie : si les stages en groupe ont constitué une bonne initiation, ils doivent être suivis et prolongés par une formation individuelle donnée aux meilleurs, *grâce à une assistance personnalisée* [...]. 2) Développer la production musicale à partir de nos moyens techniques : il faut, à cette fin, multiplier la présence à l'Ircam, de compositeurs extérieurs invités. Ils [...] utiliseront le potentiel de l'Ircam en y apportant leurs propres ressources et inspirations. Ils nous fourniront un apport intellectuel nouveau, mettant en circulation des idées et points de vue qui ne seront pas exclusivement 'de la maison' (inédit, p. 1-2 ; note italique).<sup>5</sup>

Pour la première fois, on parle de cette collaboration en termes d'"assistance personnalisée" aux compositeurs les plus doués. Mais l'idée d'un vrai rôle, d'un métier, dans la chaîne productrice est encore prématurée.

Pendant les premières années 1980, l'Ircam est organisée en trois composantes : recherche, pédagogie, création/diffusion. Les compositeurs reçoivent une aide dans la partie pédagogie, les scientifiques travaillent à leur tour dans la partie recherche et création (« dans leur immense majorité les compositeurs n'ont pas reçu de formation scientifique [...]. Il faut leur faciliter l'apprentissage de ces nouvelles techniques pour permettre l'exploitation des possibilités extraordinaires qui leur sont offerts »).<sup>6</sup>

Le nom « tuteur » apparaît à la fin de 1982. On a retrouvé une note de Pierre Boulez relative à l'organisation du secteur artistique de l'Ircam qui expliquait, le 15 octobre, que la cellule production musicale était supprimée mais que :

[I]es tuteurs seront régulièrement conviés au comité artistique afin de rendre compte des projets dont ils sont responsables et de faire toute proposition qui

leur paraît souhaitable pour le bon déroulement de leur travail.<sup>7</sup>

Les premiers témoignages d'une définition de tuteur qui indique nettement sa fonction, sont datés au début de 1983: le tuteur assure le travail d'enseignement et de 'guidage' et par ailleurs il a lui-même une part active dans la recherche musicale et la documentation afférente.<sup>8</sup> Il est un lien entre la recherche et son application à la pédagogie et à la production musicale.<sup>9</sup> C'est dans cette période que les sources montrent l'idée qu'il s'agit d'un métier qui cherche à se séparer des autres, d'une véritable profession (« poste de tuteur ») : lorsque un des membres du comité d'administration demande (20 mai 1983) si la direction de la musique entend équiper les conservatoires nationaux de matériel 4X, un autre souligne que subsiste le problème de la pédagogie. Pierre Boulez relève alors que ce problème est soulevé chaque année par l'Ircam qui sollicite des créations de postes pour des tuteurs.<sup>10</sup> Une autre source contemporaine de même veine confirme cet état d'esprit, car l'on parle de « contrats des tuteurs supplémentaires ».<sup>11</sup>

À partir de cette période et jusqu'à 1988, le terme tuteur triomphe. En même temps, au cours des années plusieurs résolutions des comités de l'Ircam avaient souligné de plus en plus le problème de ce rôle et de cette profession émergente, qui nécessitait d'être définie avec précision. On a retrouvé l'ordre du jour du comité de coordination du 13 avril 1988 : « Le comité de coordination du 13 avril sera essentiellement consacré à définir le rôle des tuteurs dans les perspectives nouvelles de Production et de pédagogie de la maison », <sup>12</sup> et le compte rendu de la réunion où, au tout premier point, on lit le titre : « 1. Le problème des tuteurs » :

C'est un problème qui se pose depuis fort longtemps et pour lequel aucune solution n'a pu encore être trouvée, que ce soit au niveau de leur statut, de leur emploi du temps, de la distinction entre leur travail de tuteur et leur volonté compositionnelle.<sup>13</sup>

Pour la première fois, on souligne également l'importance de leur activité de composition, pour

<sup>7</sup> Document *DIFFUSION GENERALE Note de Pierre Boulez relative à l'organisation secteur artistique de l'IRCAM, Le 15 octobre 1982*, archives internes Ircam.

<sup>8</sup> *L'ircam – Bilan et perspectives* (3 mars 1983), archives internes Ircam.

<sup>9</sup> Procès-verbal du conseil d'administration de l'ircam, séance du 25 avril 1984 : p. 3, point b) et c), archives internes Ircam.

<sup>10</sup> *Comité d'Administration du 20 mai 1983*, p. 3, archives internes Ircam.

<sup>11</sup> Dossier *Structures / Création bureau de Production Juillet 83* : document (3 pages) « Diffusion générale », daté 5 juillet 1983, signé Pierre Boulez, archives internes Ircam.

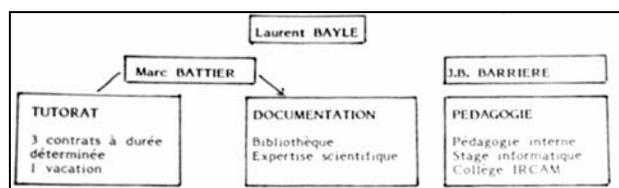
<sup>12</sup> Ordre du jour Comité de coordination du 13 avril 1988, archives internes Ircam.

<sup>13</sup> Compte rendu du Comité de Coordination du mercredi 13 avril 1988, archives internes Ircam.

<sup>5</sup> Pierre Boulez, *De nouvelles structures pour l'Ircam*, 13 novembre 1979, doc. inédit (dactylographié, 9 pages), archives internes Ircam.

<sup>6</sup> Document *NOTE pour Monsieur Groshens. Texte réunis le 20 mai 1981 pour Groshens, CC PB BM* (ce texte est présent dans plusieurs documents de l'époque), archives internes Ircam.

« comprendre les compositeurs », mais en dehors de leur temps de tutorat. Le compte rendu présente la vision d'Andrew Gerzso qui les définit idéalement comme des instrumentistes, virtuoses de leur instrument (Synthétiseur, ordinateur...), avec des connaissances techniques très poussées (ivi, p.3). Leur mission est désormais très claire : la réalisation des œuvres (information, initiation du compositeur ; élaboration du planning des studios ; suivi musical pendant la réalisation) ; la documentation (documentation active des œuvres ; coordination de la documentation des développements pour le secteur scientifique) : la pédagogie (rôle actif dans le stage informatique ; rôle actif et de coordination dans la pédagogie interne, parallèle à la documentation ; rôle administratif dans le Collège Ircam) (ibidem).



**Figure 1.** Tuteurs à l'Ircam en 1988 (Source : *Compte rendu du Comité de Coordination du mercredi 13 avril 1988*, p. 4, inédit, Archives de l'Ircam)

### 3.2. Vers l'assistant musical

Dans un document non daté intitulé *Le tutorat à l'Ircam* mais, d'après le contexte, certainement très proche du compte rendu du 13 avril 1988, on discute, à la suite des discussions menées en Comité de Coordination et de consultations auprès de ses membres, pour permettre à l'Administration d'établir un Statut du tuteur mieux adapté et en accord avec la législation de l'époque.<sup>14</sup> Ce qui est digne d'être noté, c'est que dans cette source on distingue tout d'abord entre le tuteur-compositeur et le tuteur-non compositeur mais aussitôt après (page 2), on introduit pour la première fois un nouveau nom qui pourrait être la solution pour un nouveau statut :

Assistant musical ou Assistant-compositeur. Cette notion permet une assimilation, par exemple avec l'assistant-metteur en scène de théâtre (ivi).

Mais la question du passage au nom d'Assistant musical était loin encore d'être définitivement résolue. Le terme tuteur persiste pendant la période du passage de relai entre le directeur Pierre Boulez et Laurent Bayle (en juin 1990 Bayle est nommé directeur adjoint jusqu'à décembre 1991 ; il sera le nouveau directeur à partir de janvier 1992). Notons qu'en 1990

<sup>14</sup> *Le tutorat à l'Ircam*, document non daté, fin des années 1980, archives internes Ircam.

Laurent Bayle, à propos des ressources humaines et des modes de répartition, parle encore de la présence de tuteurs consacrés à la création musicale, ayant un rôle d'encadrement des compositeurs et leur apportant un soutien technique en les guidant dans leurs choix.<sup>15</sup> Dans un long passage du *Rapport d'activité de l'Ircam 1989*, cité dans son intégralité car ce qui nous intéresse ici c'est la définition très claire des fonction de cette figure professionnelle, nous lisons que le travail du tuteur est divisé généralement en trois phases :<sup>16</sup>

dans la première, le compositeur expose ses idées compositionnelles initiales. Elles sont ensuite travaillées, formalisées en collaboration étroite avec les tuteurs. Il s'agit donc d'une phase d'exploration d'une durée variable, consistant à essayer un certain nombre de développements, tester les idées compositionnelles et à adapter et/ou écrire des programmes. Le projet compositionnel et l'environnement technique qui lui est associé y prennent peu à peu la forme quasi-définitive. A partir de ce environnement, le compositeur peut ensuite entreprendre la seconde phase de travail, phase dans laquelle le tuteur n'intervient que très modérément et qui consiste à écrire la partition. La dernière concerne la réalisation de l'œuvre et se déroule nécessairement à l'Institut. La fonction des tuteurs est très importante dans cette activité. Ils assurent tout d'abord la formation technique des compositeurs dans le cadre d'un stage d'été de six semaines. Ce stage ne se limite pas simplement à une initiation aux différentes techniques d'analyse, de traitement, de synthèse du son ou encore de composition assistée par ordinateur ; il propose une véritable introduction à l'ensemble des recherches scientifiques développées à l'IRCAM. Les tuteurs assistent ensuite les compositeurs dans l'élaboration du projet et l'aident à traduire en réalité technologique ses idées compositionnelles. Ils contribuent ainsi plus fondamentalement à la liaison 'composition recherche' essentielle dans la perspective de l'IRCAM. Pour cette raison, les tuteurs sont tous des compositeurs ayant une connaissance confirmée des développements techniques existant à l'intérieur et à l'extérieur de l'Institut [9 : 46-47].

## 4. ANNEES QUATRE-VINGT-DIX

### 4.1. L'assistant musical

Assistant : le glissement vers cette forme, commencé vers la fin de la décennie, est achevé entre 1991 et 1992. Deux sources contemporaines nous permettent de l'établir. Dans le *Projet d'activité 1992* (non daté mais

<sup>15</sup> CA [Comité d'Administration] 9 janvier 1990, p. 4, archives internes Ircam.

<sup>16</sup> [9] et copie présente dans l'archive : *Rapport d'activité 1989*, chapitre 'Les activités musicales', sous-chapitre 'La création musicale', pp. 46-47, archives internes Ircam. Nota Bene: la compilation de l'ensemble des textes composant ce *Rapport d'activité* avait été coordonnée par Marc Battier, lui-même concerné au premier chef par ce passage

présenté dans le Conseil d'administration du 9 janvier 1992 et vraisemblablement composé vers les derniers jours de 1991), le conseil d'administration annonce que les activités de recherche musicale et de production musicale seront fusionnées en une seule entité appelée 'Création', afin de susciter une meilleure interaction entre les compositeurs et les assistants musicaux.<sup>17</sup> Dans le *Rapport d'activité 1991*, présenté au conseil d'administration en juin 1992, on indique la liste des œuvres achevées, avec le nom, à côté, de l'assistant musical intervenu.<sup>18</sup> Si les documents officiels de l'Ircam utilisent la forme nouvelle, les collaborateurs de l'Ircam tardent à se rendre compte qu'il y avait eu ce changement et persévèrent dans l'utilisation du titre de tuteur (plusieurs lettres du début des années 1990 conservées dans les archives de l'Ircam).<sup>19</sup> De toute façon, ce terme d'assistant musical finit par remplacer celui de tuteur.

Dans le *Projet d'activité 1994*, nous lisons des patronymes qui seront présents à l'Ircam jusqu'à nos jours.<sup>20</sup> Toutes les œuvres listées (opéras et spectacles musicaux ; ballets, films ; œuvres pour ensemble ; œuvres pour solistes ou formation musique de chambre) avaient reçu l'aide d'un ou plusieurs assistants musicaux : Pierre Charvet, Eric Daubresse, Christophe de Coudenhove, Thomas Hummel, Serge Lemouton, Cort Lippe, Leslie Stuck. Le responsable de la création musicale était à l'époque Alain Jacquinot, le coordinateur des assistants musicaux Jean-Baptiste Barrière, le responsable de la production et de la pédagogie Andrew Gerzso.

Le métier d'assistant musical portera ce titre jusqu'à toute la première moitié des années 2000, on peut dire encore mieux jusqu'à 2007. L'archive de l'Ircam garde plusieurs textes probant à ce sujet. Le *Rapport d'activité 1997* (toujours préparé par Marc Battier), présente une réflexion sur son rôle « particulièrement remarquable [...], dont la fonction constitue un lien essentiel entre équipes de recherche et compositeurs » (p. 94).<sup>21</sup>

Même pour des créateurs expérimentés en informatique ou en technique de studio, l'assistant est un interlocuteur précieux qui, loin d'influencer l'écriture ou la sonorité d'une œuvre, la maintient sur les rails en vue de son exécution. L'assistant propose des solutions informatiques aux questions posées par le compositeur et coordonne leur mise en œuvre. Après la création, il est également le garant de la survie de la pièce, par l'important travail de

documentation technique effectué a posteriori (ibidem).

Dans le même texte (p. 205) on constate que les assistants se divisent en Assistant musicaux création (Eric Daubresse, Serge Lemouton, Tom Mays, Gilbert Nouno, Leslie Stuck, Frédéric Voisin, Carl Harrison-Faïa, Ipke Starke) et Assistants musicaux pédagogie. Les premiers sont consacrés à la collaboration avec les compositeurs. Plus tard, à partir de l'année 2000, on remarque une nouvelle ramification du titre d'assistant (les œuvres produites sont aussi très variées). Assistants Musicaux Création : (Harrison-Faïa, Lemouton, Nouno, Olivier Pasquet, Manuel Poletti, Voisin), Assistants Informatique Musicale (Denis Lorrain), Pédagogie (Mikhail Malt, Benjamin Thigpen, Hans Tutchku, Cyrille Brissot), animateur d'ateliers en informatique musicale (Jean Lochard).<sup>22</sup>

## 4.2. Vers le RIM

Le titre d'assistant musical a eu une longévité remarquable par rapport aux précédents et a duré plus de 15 ans. Cependant, les textes que nous avons consultés rappellent un certain malaise à l'égard d'une appellation ainsi que d'une définition particulièrement difficile. Dans un passage du procès verbal de la réunion du conseil d'administration du 11 décembre 2001,<sup>23</sup> Pierre Boulez qui était présent pose la question sur le rapport entre l'assistant et le compositeur : « où on est-on aujourd'hui ? Le compositeur est-t-il suffisamment avancé pour ne plus avoir besoins d'assistants ? [...] Son statut lui convient-il [...] ? » (p. 8). Bernard Stiegler, quelques semaines avant de prendre officiellement ses fonctions de directeur du centre (en janvier 2002), exprime une réflexion sur l'identité de cette figure, qui a accumulé du savoir-faire et est devenue un type d'expert au statut encore ambigu, même si la situation actuelle a démontré sa nécessité. Bien que les compositeurs puissent travailler à domicile avec leurs propres outils informatiques, ils souhaitent venir tout de même à l'Ircam en raison de l'apport des assistants musicaux (p. 8-9). Bernard Stiegler ajoute que

[l]es assistants musicaux apportent énormément en termes de ressources aux compositeurs et aux chercheurs, mais également à la pédagogie et cette spécificité va continuer à se développer. Leur rôle s'inscrit tout à fait dans ce qui se développe depuis quelques années dans le domaine de la conception d'instruments numériques sous la forme d'une sorte de médiateur entre concepteurs industriels ou technologiques et utilisateurs, avec des personnes

<sup>17</sup> *Projet d'activités 1992*, archives internes Ircam.

<sup>18</sup> [10] et copie conservé dans le dossier *Ircam Conseil d'administration du 25 juin 1992 + PV signé*, archives internes Ircam.

<sup>19</sup> Voir par exemple la lettre de Philippe Schœller à Gilbert Paris datée 24 février 1992 ou la lettre datée 17 mars 1992 de Curtis Roads au directeur Bayle, archives internes Ircam.

<sup>20</sup> *Projet d'activité 1994*, archives internes Ircam.

<sup>21</sup> [11] et *Rapport d'activité 1997*, sous la direction de Marc Battier, archives internes Ircam.

<sup>22</sup> *Projet d'activité 2000*, archives internes Ircam.

<sup>23</sup> *Conseil d'administration de l'Ircam. Procès verbal de la réunion du 11 décembre 2001*, archives internes Ircam.

ayant des profils intermédiaires. [...] [C]ette organisation de l'Ircam [...] est aussi une grande partie de son savoir-faire [...] et de la capitalisation de ses connaissances. Il faut néanmoins mieux la contractualiser [...]. (Ibidem).

A ces mots, Laurent Bayle ajoute le fait qu'il y a déjà eu beaucoup d'amélioration du statut, qu'il s'agit certainement d'un thème délicat sur lequel ils avaient fait plusieurs études juridiques, car c'est la question des droits d'auteurs, entre autres, qui est mise en cause (« une question qui se pose au niveau mondial avec le développement des nouvelles technologies », ibidem). Chaque nouvelle appellation devra être conforme au règlement des droits d'auteurs, que ce soit avec le compositeur ou son éditeur (ibidem). La question n'est pas encore résolue à ce jour.

## 5. DE 2007 A AUJOURD'HUI

### 5.1. Le Réalisateur en Informatique Musicale

D'où vient ce mot nouveau ? Nous savons que les sources inédites et surtout les projets et les rapports d'activité de l'Ircam, utilisent le titre d'assistant musical jusqu'à 2007. Ce qui nous a frappé dans nos recherches c'est ce fait, qui évidemment révélait une mutation de perspective, d'un changement d'appellation dans le titre d'un colloque qui a eu lieu en 2007. Le *Projet d'activité 2007* (sans date mais rédigé l'on suppose entre la fin 2006 et les premières semaines de 2007) présente le projet d'organiser un colloque sur ce thème : *Journée Le métier d'assistant musical – samedi 23 juin 2007* (figure 2).

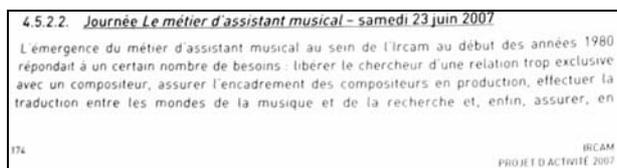


Figure 2. *Projet d'activité 2007*, p. 174 (Ircam, Médiathèque).

Or, on sait que le colloque eut lieu le 22-23 juin 2007, mais avec un titre différent : *Journée professionnelles sur le métier de Réalisateur en Informatique Musicale* et en anglais : *Professional-level meeting on the profession of computer music designer* (Organisateur : Andrew Gerzso).<sup>24</sup> Les sources

<sup>24</sup> Lien direct au programme du colloque : <http://archiprod-externe.ircam.fr/notes2prog/LO40429-01-wm.pdf>. Andrew Gerzso avait écrit dans le programme : « Le réalisateur en informatique musicale est-il une exclusivité de l'Ircam ? Vraisemblablement pas, puisque aujourd'hui dans tous les milieux où des créateurs travaillent avec des nouvelles technologies pour le son ou la musique — danse, théâtre, image & arts plastiques, musique —, on trouve, sous d'autres étiquettes — régisseur, ingénieur, etc. —, des professionnels qui maîtrisent des concepts, des technologies et des pratiques similaires » (ivi, p. 5). Une partie de ma recherche future sera destinée à la transcription de

papier ne viennent donc pas combler les manques de l'information sur cette métamorphose. Ne pouvant concevoir pourquoi le titre était changé de cette manière, on a cherché à comprendre d'où venait l'idée de la nouvelle appellation. Grâce à l'aimable efficacité de Serge Lemouton (assistant musical / réalisateur en informatique musicale, comme il s'était défini lui-même dans le curriculum vitae apparu dans le programme du colloque) et à son archive informatique personnelle j'ai pu avoir connaissance d'un courrier électronique très précieux qu'il avait sauvegardé. L'appellation "Réalisateur en informatique musicale" est apparue pour la première fois dès 1997, sur proposition de Leslie Stuck, dans un compte rendu d'une réunion des "assistants musicaux" (26 May 1997) et envoyé, assez ironiquement pour notre recherche sur le nom, à une liste (en anglais) appelée "tutors" :

#### REUNION DES ASSISTANTS A L'EXCELSIOR 23 MAI 1997

On a discuté le statut d'assistant en préparation pour le dernier rendez-vous avec LB [Laurent Bayle, directeur de l'Ircam à l'époque]. On voudrait commencer par insister un peu sur les points de la dernière réunion: les postes, les bios, etc. Chacun devrait emmener sa bio à la prochaine réunion, d'ailleurs.

\* Le Conseiller et Réalisateur de l'Informatique Musicale (CRIM)

Tous présents étaient d'accord que le problème essentiel est: pour bien aider un compositeur, il faut être compositeur, qui veut dire que, malgré tout, on va composer. Quand on travaille pour l'IRCAM, on devrait donner notre meilleur pour l'IRCAM; quand on compose, on devrait donner notre meilleur pour notre musique.

[...] Leslie Stuck.

Comme Serge Lemouton l'a souligné, après cette date le nom RIM a été officialisé (début 2005) par André Santelli et Bernard Stiegler. A peu près à la même époque, les RIM de l'Ircam ont obtenu que leur nom et leur biographie soient systématiquement inscrits au programme des concerts.<sup>25</sup> D'autres étapes importantes dans la reconnaissance du métier, outre les deux *Journées professionnelles sur le métier de réalisateur en informatique musicale*, ont été l'ouverture d'une formation de niveau Master 2 RIM à l'université Jean Monnet de Saint-Etienne depuis la rentrée de septembre 2011, sous la direction de Laurent Pottier (ex assistant Ircam), un article publié par des

l'enregistrement de ce colloque déposé à l'Ircam (Médiathèque, consultation sur place).

<sup>25</sup> Serge Lemouton, *Vingt ans de pratique de la Réalisation en Informatique Musicale. Enjeux, perspectives et état des lieux d'un métier en devenir*, Dossier de demande de validation des acquis de l'expérience en vue de l'obtention totale du diplôme : Master 2 - Musique et informatique, 2012, p. 67. La thèse de Lemouton présente aussi un magnifique schéma qui montre les « générations de réalisateurs en informatique à l'Ircam (Production) » (p. 67), ainsi qu'une réflexion sur la question salariale, du statut et du contrat (p. 68-69).

RIM aux JIM – Journées d'Informatique Musicale en 2002 [12] et la réalisation d'une *Fiche descriptive du métier AFPA: Référentiel Emploi Activités Compétences des Réalisateurs en Informatique Musicale de l'Ircam* (Paris, décembre 2006).<sup>26</sup>

## 6. CONCLUSIONS

La recherche dans les archives papier de l'Ircam a eu pour but de reconstruire les origines du nom de RIM. Nous avons exploré les projets et les rapports d'activité ainsi que certains compte rendus de réunions (Direction artistique, Direction générale, Conseil d'administration). Les données recueillies ont permis d'éclaircir un certain nombre de points ainsi que d'ouvrir l'étude à d'autres lignes qui restent encore vagues. La recherche sur l'évolution de l'appellation présentée ici a permis de tracer la transition du nom ainsi qu'une certaine conscience de la part des protagonistes. Si à l'époque de l'ouverture de l'Ircam on utilisait le terme générique de scientifique, on est passé ensuite au nom de tuteur, et encore à l'alternance entre Assistant et Tuteur, ensuite d'Assistant musical et enfin de Réalisateur en Informatique Musicale. La recherche future veut approfondir une approche comparative des sources, notamment inédites, avec les articles techniques publiés par les RIM (afin de préciser les contours du métier tel que ses acteurs se le sont représentés). Enfin, une prospection sur l'histoire des assistants musicaux en Europe et aux Etats-Unis sera nécessaire.

## 7. REFERENCES

- [1] *La musique en projet. Le premier livre de l'Ircam*, Paris, Gallimard, 1975.
- [2] Born, G. *Rationalizing Culture. Ircam, Boulez and the Institutionalization of the Musical Avant-Garde*. University of California Press, 1995.
- [3] Veitl, A. *Politique de la musique contemporaine. Le compositeur - la recherche musicale et l'état en France de 1958 à 1991*. Paris, L'Harmattan, 1997.
- [4] Menger, P.-M. *Les laboratoires de la création musicale. Acteurs, organisations et politique de la recherche musicale*. Paris, La documentation française, 1989.
- [5] Risset, J.C. *The Development of Digital Techniques. A Turning Point For Electronic Music?*, Rapports IRCAM 9/78. Paris, Centre Georges Pompidou, 1978.
- [6] Bennett, G. *La recherche à l'Ircam en 1978*, Rapports Ircam 19/79. Paris, Centre Georges Pompidou, 1979 (disponible en ligne la version
- anglaise : <http://www.gdbennett.net/>, dernière consultation le 18 février 2013).
- [7] *La recherche à l'Ircam en 1979*. Rapports IRCAM 29/80, Paris, Centre Georges Pompidou, 1980.
- [8] A. R. [?], *Le 'petit Beaubourg' de Pierre Boulez*. Le Monde, 9 mars 1974.
- [9] *Rapport d'activité 1989*. Paris, Centre Georges Pompidou, 1990.
- [10] *Rapport d'activité 1991*. Paris, Centre Georges Pompidou, 1992.
- [11] *Rapport d'activité 1997*, sous la direction de Marc Battier. Paris, Centre Georges Pompidou, 1998.
- [12] Poletti, M. – Mays, T. – Faia, C. "Assistant musical ou producteur ? Esquisse d'un nouveau métier". Proceedings of the JIM *Journées d'Informatique Musicale, 9e édition, Marseille, 29 - 31 mai 2002*, 2002 ([http://jim.afim-asso.org/jim2002/articles/L29\\_Mays.pdf](http://jim.afim-asso.org/jim2002/articles/L29_Mays.pdf)).

<sup>26</sup> Ibidem.



# MODULATEURS EN ANNEAU ET SAXOPHONE : LE DISPOSITIF D'ECRITURE MIXTE ET L'INTERPRETATION PARTICIPATIVE DANS L'ŒUVRE *LE PATCH BIEN TEMPERE II*

*Tom Mays*

CICM – Université de Paris VIII  
CNSMD de Paris  
[contact@tommys.net](mailto:contact@tommys.net)

*Pedro Bittencourt*

CICM – Université de Paris VIII  
EM Universidade Federal do Rio de Janeiro  
[contact@pedrobittencourt.info](mailto:contact@pedrobittencourt.info)

## RÉSUMÉ

*Le Patch Bien Tempéré II* pour saxophone et traitement en temps réel est la deuxième œuvre d'une suite de pièces du compositeur Tom Mays explorant l'écriture mixte pour instrument seul et différents traitements en temps réel de base. Dans le cadre d'une collaboration proche, aussi appelée *interprétation participative*, une pièce mixte a été créée avec le saxophoniste Pedro Bittencourt. Le saxophone (au choix) est au cœur des traitements en temps réel et l'exécution instrumentale agit finement sur l'électronique, qui par son comportement influence constamment le jeu du saxophoniste, résultant en un instrument hybride et étendu. Le son du saxophone est traité uniquement par quatre modulateurs en anneau implémentés dans le logiciel Max/MSP. Un suivi d'amplitude permet à la fois de déployer les modulateurs progressivement et de créer des glissandi de modulation en fonction de l'intensité du saxophone. Après une discussion sur la notion de *interprétation participative*, nous analysons la pièce en exposant les techniques d'écriture instrumentale et électronique pour et avec les modulateurs en anneau. Nous listons les adaptations spécifiques effectuées durant le processus de collaboration.

## 1. INTRODUCTION

Dans la musique de nos jours, le compositeur opère souvent la partie électronique. L'équilibre entre les sources sonores et la nature de leurs interactions représente des éléments clés pour qu'une interprétation musicale soit satisfaisante. Nous considérons que dans ce cas le responsable de l'électronique (fût-il le compositeur ou un technicien) partage l'interprétation avec l'instrumentiste. Dans le cadre de l'interprétation participative [BITTENCOURT, 2013], compositeur et interprète collaborent pendant toutes les étapes du processus créateur, pour construire ensemble l'interprétation musicale à travers plusieurs échanges.

Dans cet article nous proposons une discussion sur la notion de *interprétation participative* suivie d'une analyse de la pièce *Le Patch Bien Tempéré II*, accompagnée d'une exposition des techniques d'écriture instrumentale avec les modulateurs en anneau. Nous présentons les échanges entre compositeur et

instrumentiste déterminants dans le processus de création et d'interprétation musicale de la pièce, et nous listons les adaptations faites durant l'élaboration de l'œuvre.

## 2. DE L'INTERPRETATION PARTICIPATIVE

Nous proposons que l'interprétation musicale puisse être une série d'interactions partagées entre l'instrumentiste et le compositeur dans certaines conditions du processus créateur de musique mixte.

L'interprétation participative [BITTENCOURT, 2013] est un processus dynamique et créatif, fruit de multiples interactions partagées entre instrumentistes et compositeurs à long terme. Elle s'étend depuis la conception des pièces (avant l'écriture de la partition), pendant les adaptations faites au fur et à mesure, jusqu'à la création en concert et l'enregistrement d'une version en studio. On réalise des tests, des improvisations et toutes les expériences nécessaires avec l'instrument et les outils électroniques demandées par les compositeurs. Le but est d'« optimiser » les nouvelles pièces et avoir une version satisfaisante (c'est à dire, approuvée par le compositeur). De nouvelles techniques étendues peuvent être requises par le compositeur et développées par l'instrumentiste. Des erreurs peuvent se produire et être utilisées pour faire découvrir de nouvelles sonorités inouïes. Pour exploiter les possibilités expressives de l'alliage de l'instrument acoustique avec l'électronique dans différents contextes musicaux, certains matériaux sonores d'autres pièces peuvent encore être repris et articulés différemment dans de nouveaux contextes. L'instrumentiste peut ainsi développer une plus grande familiarité avec la pièce selon les directives du compositeur et dans l'autre sens le compositeur peut se familiariser avec les singularités de l'instrumentiste.

De cette façon l'interprétation musicale d'une nouvelle pièce mixte peut être construite graduellement et ensemble, nourrie par tous les échanges de compétences du compositeur et de l'instrumentiste. Le participatif englobe cet « apprentissage mutuel » lié au contexte d'optimisation de chaque pièce. La découverte et l'apprentissage de nouvelles sonorités (mixtes, électroniques et instrumentales) peut émerger pendant ce travail dynamique qui se reconfigure selon son propre déroulement. Le compositeur y est aussi considéré comme un interprète, non seulement parce qu'il opère l'électronique en concert (cela pourrait être aussi réalisé

par quelqu'un d'autre, comme un technicien son avec les instructions nécessaires) mais aussi parce qu'il a participé aux choix basés sur les échanges dans le processus créateur, avant, pendant et après l'écriture de la partition et la partie électronique. Ce vécu est le différentiel de ce que nous considérons comme la construction de l'interprétation musicale avec l'instrumentiste.

Dans nos collaborations actuelles, le compositeur demeure le seul auteur des œuvres, et c'est toujours lui qui fait le choix final. C'est lui qui détermine l'ordre des étapes de travail, le rôle de l'instrumentiste étant toujours de le soutenir. Pour employer un terme de la systémique, [DURAND, 1979] nous décrivons un système à apprentissage où les décisions finales sont faites par le compositeur. Cela justifie que l'instrumentiste ne soit pas un « co-auteur » des pièces musicales.

Dans cette approche, compositeur et instrumentiste cheminent ensemble. Dans l'interprétation participative il n'y a pas un cadre de travail prédéfini. Il n'y a pas un schéma à suivre avec des étapes temporellement établies ou séparées par tâches. Les étapes ne sont pas nécessairement effectuées dans l'ordre présenté et peuvent se répéter à n'importe quel moment du processus créateur pour s'achever selon les besoins. Les musiciens découvrent ensemble ces procédés tout en s'influençant mutuellement. Notre propre objet d'étude est modifié par notre action. Nous y trouvons quelques convergences avec des propos des sciences cognitives [VARELA, 1989], des sciences de l'information et de la communication [WATZLAWICK, 1972] et de la recherche-action intégrale et systémique [MORIN, A., 2010] qui sont encore en développement.

Plus qu'être « fidèle » à l'œuvre, l'instrumentiste aide à sa construction et à la concrétisation des premières versions en concert et en studio. L'instrumentiste se met à disposition du compositeur pour expérimenter et potentialiser ses idées musicales. L'interprétation participative émerge de l'apprentissage issu de ces échanges de compétences des musiciens<sup>1</sup>.

Il faut rappeler que l'interprétation participative telle que nous le proposons est plutôt réalisée avec des œuvres mixtes qui n'ont pas encore été créées pour qu'il soit possible d'avoir assez d'échanges et surtout qu'elles soient mutuelles. Cela n'empêche pas que des œuvres déjà écrites ou même créées soient entièrement exclues de l'interprétation participative. Selon l'ouverture accordée par le compositeur, l'instrumentiste peut construire avec lui une nouvelle interprétation dans ce cadre participatif, ou même une nouvelle version, en retravaillant des aspects qu'une écoute postérieure puisse faire apparaître.

### 3. L'ŒUVRE ET L'ÉCRITURE

*Le Patch Bien Tempéré II* (2012) est une œuvre d'environ huit à neuf minutes pour saxophone au choix (sopranino, soprano, alto, ténor, baryton, basse) et traitement en temps réel. C'est la deuxième d'une suite

de pièces du compositeur Tom Mays explorant l'écriture mixte pour instrument seul et traitements de base [MAYS, 2010]. Une collaboration proche avec le saxophoniste Pedro Bittencourt a permis de créer une pièce où l'instrument (sax au choix : alto, ténor, barytone ou soprano) est au cœur des traitements et où l'interprétation agit finement sur l'électronique et résulte en un instrument hybride et étendu. Le son du saxophone est traité uniquement par quatre modulateurs en anneau implémentés dans Max/MSP. Un suivi d'amplitude permet à la fois de déployer les modulateurs progressivement et de créer des glissandi de modulation en fonction de l'intensité. Le traitement en temps réel de la pièce est implémenté dans l'environnement de création *Tapemovie* [Mays/Rubiano 2010], lui-même construit avec Max/MSP de Cycling74.

#### 3.1. Les objectifs de la pièce

La composition de cette pièce a été motivée par plusieurs objectifs :

1. Faire une pièce compatible avec tous les saxophones, à la discrétion de l'instrumentiste – lisant la même partition et utilisant le même patch informatique en prévoyant les transpositions nécessaires.
2. Donner à entendre que l'instrument et le traitement se mêlent parfaitement pour ne faire qu'un seul son – ni acoustique, ni électronique, mais un hybride des deux. L'instrument acoustique sera réellement modifié. On n'entend pas le son séparé de l'instrument, ni l'électronique seule.
3. Essayer d'explorer complètement les possibilités musicales des modulateurs en anneau en les employant UNIQUEMENT, sans le moindre délai ou filtrage.
4. Intégrer le contrôle des paramètres de traitement au jeu instrumental par le biais d'un suivi d'amplitude, impliquant ainsi plus fortement l'écoute de l'instrumentiste qui doit entendre pleinement la modification de son instrument par l'électronique pour atteindre l'expressivité souhaitée.
5. Faire en sorte que la pièce soit simple à monter et à exécuter et que les événements se déclenchent automatiquement par le son de l'instrument lui-même.

Pour réaliser le premier objectif, il fallait gérer les fréquences de modulation en fonction d'une transposition donnée. Ces fréquences sont donc exprimées en notes transposées et la transposition selon le type de saxophone choisi est appliquée globalement (voir 3.2.2). Dans l'écriture pour le saxophone, il fallait éviter les modes de jeu trop idiomatiques d'un saxophone spécifique et chercher la compatibilité. Les multiphoniques, par exemple, sont très différents selon le membre de la famille (soprano, alto, ténor, ou baryton). Par contre, des sons éoliens sans le bec et un léger

<sup>1</sup> Nous n'y excluons pas les techniciens son. Ils sont parfois indispensables pour la bonne réalisation des pièces.

glissando sur un Fa grave sont possible à réaliser avec tous les saxophones.

Le deuxième objectif est atteint par la nature même des modulateurs en anneau qui altèrent la structure harmonique du son en éliminant le son d'origine (voir 3.2.2). Ceci permet d'avoir un niveau sonore important, car il y a peu de risques d'effet Larsen, et par conséquent nous avons un mélange très efficace avec l'instrument en direct.

Le troisième objectif a comme conséquence une partie électronique tellement « accrochée » à l'instrument qu'elle ne peut pas le soutenir ou le prolonger temporellement – quand l'instrument s'arrête, le traitement s'arrête. Ainsi la pièce ressemble plus à une pièce pour un instrument seul qu'à une pièce de musique mixte. Dans la version actuelle de la partition, il n'y a plus souvent d'indications sur le comportement de l'électronique.

Le quatrième objectif soulève l'intérêt de faire une implémentation informatique des modulateurs en anneau, le réel avantage étant dans le contrôle des paramètres. Dans *Le patch bien tempéré II* un suivi d'amplitude du saxophone sert à la fois à déployer progressivement les quatre modulateurs et également à faire glisser les fréquences. Les nuances et changements d'intensité dans la pièce deviennent alors très importants à l'égard de l'expressivité du jeu. Le saxophoniste contrôle l'électronique et les transformations de l'électronique suggèrent le phrasé et le contrôle des nuances du saxophoniste.

Le cinquième objectif est très important pour la vie de l'œuvre. Avec une possibilité pour l'instrumentiste de répéter seul chez lui en ayant un minimum de connaissance technique et d'équipement et un minimum de souci d'exécution, la pièce aurait plus de chance d'être jouée par des saxophonistes de différents niveaux et sur toute la famille d'instruments (du soprano au basse).

### 3.2. Le dispositif

*Le patch bien tempéré II* a été réalisé dans l'environnement *tapemovie*<sup>2</sup> qui est lui-même un patch Max/MSP modulable et scriptable [MAYS, RUBIANO, 2010]. Avant de continuer sur l'exposition du dispositif, nous allons voir un bref historique des modulateurs en anneau et un aperçu de son fonctionnement sur le plan théorique.

#### 3.2.1. Les modulateurs en anneau – historique et théorie

Bien que l'historique et la théorie de la modulation en anneau soient bien connus du public averti de la musique électronique et de l'informatique musicale, nous préférons faire ici un récapitulatif complet pour que même des débutants puissent suivre dans un contexte pédagogique. Certains peuvent s'il le souhaitent passer

directement à la suite – Les modulateurs en anneau – implémentation dans la pièce.

La modulation en anneau est sûrement le traitement le plus perceptible et le moins dangereux à faire en temps réel de tous. Cette transformation est perceptible car le spectre et donc le timbre du son se trouvent modifiés de façon significative et cette *séparation* du son traité du son d'origine augmente très fortement sa perceptibilité. Le même décalage du spectre qui rend la modulation en anneau perceptible diminue énormément les risques d'effet Larsen, car les fréquences qui sortent des haut-parleurs ne sont pas les mêmes que celles qui rentrent via le microphone. Ainsi on empêche le *renforcement* qui provoque l'effet Larsen.

La modulation en anneau existe depuis les années 1930 dans des applications de téléphonie, mais ne sert dans la musique que depuis les années 50, surtout en Allemagne, au Japon et en Italie. Vers le début des années 60 cependant, la moitié des studios de musique électroacoustique au monde possédait un modulateur en anneau.

Karlheinz Stockhausen, travaillant dans les studios du WDR Cologne, a employé des modulateurs en anneau dans plusieurs pièces – notamment *Mixtur* (pour orchestre et 4 générateurs de sinus et modulateurs en anneau), *Mikrophonie II* (pour 12 chanteurs, Orgue Hammond, comme porteuse, et 4 modulateurs en anneau) et *Mantra* (pour deux pianos, deux générateurs de sinus et 2 modulateurs en anneau) [DAVIES, 1976].

Tous ces modulateurs en anneau historiques étaient bien sûr des appareils analogiques. Aujourd'hui il est beaucoup plus *simple* et *portable* de faire des modulateurs en anneau avec des moyens informatiques, mais le son n'est pas tout à fait pareil. Le modulateur en anneau numérique manque le comportement non-linéaire et les harmoniques supplémentaires d'un modulateur en anneau analogique [PARKER, 2011]. Cependant le dispositif numérique permet un contrôle précis et varié qui n'est pas facilement reproductible avec les moyens analogiques.

Pour faire une modulation en anneau en informatique (dans Max/MSP<sup>3</sup> ou PureData<sup>4</sup>, par exemple) on prend un signal complexe (entrée micro par exemple) et on le multiplie avec un oscillateur sinusoïdal *bipolaire* (oscillation positif et négatif autour de « 0 »). Le résultat est la somme ET la différence de toutes les fréquences du signal complexe (porteuse) avec la fréquence de l'oscillateur.

Si un son complexe avec un spectre harmonique de 1000 Hz, 2000 Hz, 3000 Hz et 4000 Hz est modulé à 250 Hz, le résultat est :

diff	somme
750	1250
1750	2250
2750	3250
3750	4250

<sup>3</sup> Max/MSP de Cycling 74 : [www.cycling74.com](http://www.cycling74.com)

<sup>4</sup> PureData de Miller Puckette : <http://fr.flossmanuals.net/puredata/>

<sup>2</sup> Le site de *tapemovie* : <http://tapemovie.org/>

La porteuse et la modulante disparaissent et il n'y a plus de rapport harmonique entre les partiels. Le son est devenu *inharmonique* et sa structure spectrale est perturbée [DAVIES, 1976] [OBERHEIM, 2008].

Si nous modulons un son harmonique avec une fréquence égale à la fondamentale de ce son, on retrouve un son quasi identique à l'original. Par exemple, modulons 100-200-300-400 Hz avec 100 Hz et nous avons :

diff	somme
0	200
100	300
200	400
300	500

Le 0 Hz va disparaître de la perception car 0 Hz n'est pas audible et le reste va faire un son parfaitement aligné avec le spectre du son d'origine, avec un partiel de 500 Hz en plus. Dans le résultat, le 100Hz sera un peu moins fort que le 100 Hz dans le son d'origine car c'est le 2e harmonique décalée plus bas. Mais globalement nous allons avoir un timbre très proche.

Cela voudrait dire que pour différents *rappports* entre la hauteur du son d'origine et la fréquence de modulation, on a différents timbres qui sont plus ou moins *harmoniques* ou *inharmoniques*. Les changements de timbres s'effectuent soit en changeant la fréquence de modulation sur une hauteur constante, soit en changeant la hauteur sur une fréquence de modulation constante.

Si la fréquence de modulation est très basse, entre 0 Hz et 20 Hz, le résultat est un son avec des battements - un effet de *lfo*. Les battements sont d'une fréquence 2 fois plus que la fréquence de modulation car c'est le résultat d'un écart entre la somme et la différence [DAVIES, 1976].

Si la fréquence de modulation est au dessus de la fondamentale de la note jouée, la *différence* crée une fréquence qui descend en dessous de 0 Hz. Dans ce cas la fréquence se trouve repliée autour de zéro et redevient positive, subissant juste une inversion de phase.

Le *Frequency Shifter* (transposition de fréquence) est proche de la modulation en anneau, donnant uniquement la somme ou la différence, mais pas les deux en même temps. Il le fait en décalant la phase de la porteuse de 90 degrés avant de la moduler avec un signal sinusoïdal également décalé de 90 degrés. Le rapport de phase du résultat fait qu'il n'y a qu'une bande de côté qui reste, l'autre étant annulé [BODE, 1984 p. 6]. Nous pouvons facilement recombinaison un shifting « haut » et un shifting « bas » pour reconstruire les deux côtés de la modulation en anneau.

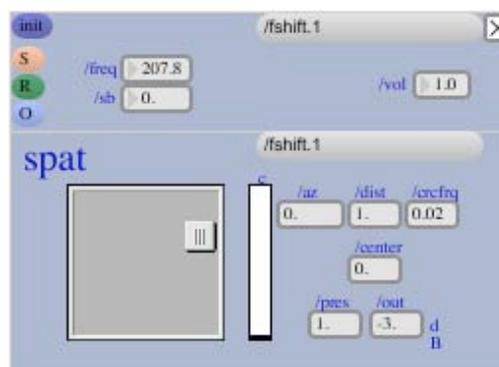
Le *Frequency Shifter* (transposition de fréquence) est proche de la modulation en anneau, donnant uniquement la somme OU la différence, mais pas les deux en même temps. Il le fait en décalant la phase de la porteuse de 90 degrés avant de la moduler avec un signal sinusoïdal également décalé de 90 degrés. Le rapport de phase du résultat fait qu'il n'y a qu'une bande de côté qui reste, l'autre étant annulé [BODE, 1984 p. 6]. Nous pouvons facilement recombinaison un shifting « haut » et un shifting

« bas » pour reconstruire les deux côtés de la modulation en anneau.

### 3.2.2. Les modulateurs en anneau – implémentation dans la pièce

La modulation en anneau dans *tapemovie* se fait par le biais d'un *frequency shifter* réalisé avec l'objet Max/MSP *freqshift~* qui sépare la somme et la différence en deux sorties distinctes. Pour cette pièce, la somme et la différence sont toujours additionnées, rendant ainsi les *frequency shifters* équivalents à la modulation d'anneau.

Dans *tapemovie*, le module de traitement utilisant l'objet *freqshift~* s'appelle *fshift*. Son éditeur montre ces paramètres (voir **Figure 1**).



**Figure 1.** L'éditeur du *fshift* dans *tapemovie*.

Le paramètre *freq* indique la fréquence de la modulation. Le paramètre *sb* est pour le *side band* (0 pour somme + différence, 1 pour somme, -1 pour différence). Puis, *vol* c'est le volume, contrôlable entre 0 et 1. Dans la partie basse de l'éditeur se trouvent les paramètres de spatialisation. L'espace et le mouvement du son ne sont pas des paramètres centraux à la composition, mais ce *spat* permet d'appliquer une légère rotation de chaque modulateur dans quatre haut-parleurs pour une meilleure mise en espace et une valorisation de l'écoute.

La gestion des fréquences de modulation se fait en note MIDI transposée pour être en rapport avec la transposition choisie selon l'instrument. Si nous demandons la fréquence de modulation qui correspond à la note 65 pour aller avec le Fa du saxophone alto, nous envoyons la variable *pitch* et cette note sera additionnée avec la valeur globale de transposition pour le saxophone « -9 », la transposition du sax alto. Ensuite, la variable *pitch* sera envoyée vers un convertisseur *note vers fréquence* avant d'arriver au module *fshift* (voir **Figure 2**).

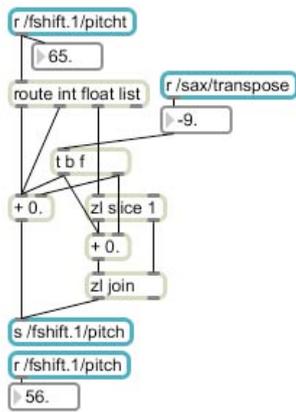


Figure 2. La prise en compte de la transposition pour le *pitch*.

Nous choisissons la transposition du saxophone dans un menu (voir Figure 3), et nous pouvons entrer une valeur de référence d'intonation si le saxophoniste préfère s'accorder au La = 442 plutôt que 440, par exemple (voir Figure 4). Ces paramètres modifient le rapport entre la hauteur transposée et la hauteur réelle.

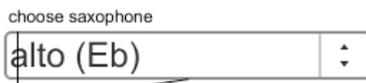


Figure 3. Choix du saxophone.



Figure 4. Référence d'intonation

### 3.2.3. Le suivi d'amplitude

L'enveloppe d'amplitude du saxophone est suivie continuellement par le module de suivi d'amplitude dans *tapemovie*, *afol* (voir Figure 5).

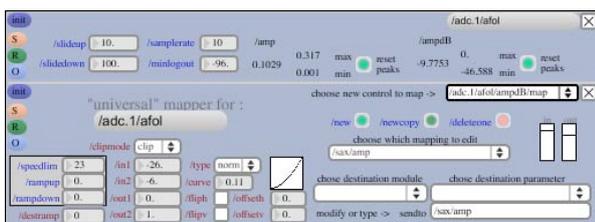


Figure 5. Module *afol* (suiveur d'amplitude).

Nous voyons également la partie du *mapping* qui crée un paramètre */sax/amp* entre -26 et -6 dB, par une courbe légèrement exponentielle. Nous voyons l'indicateur du paramètre */sax/amp* dans la Figure 6.

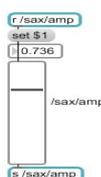


Figure 6. Paramètre */sax/amp*.

Le paramètre */sax/amp* est utilisé ensuite pour ouvrir progressivement les modulateurs en anneau par 4 niveaux d'amplitude entre *pp* et *ff* (voir Figure 7)

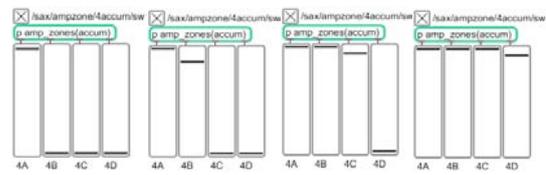


Figure 7. Ouverture progressive des quatre modulateurs selon le niveau de */sax/amp* (*pp*, *mp*, *f*, *ff*).

### 3.2.4. L'avancement automatique des événements

Jusqu'à l'heure de l'écriture de ce document, les événements ont été déclenchés manuellement par le compositeur<sup>5</sup>. L'avancement automatique n'a pas encore été implémenté. Il le sera prochainement, en utilisant principalement le suivi d'amplitude pour détecter les silences entre les phrases et la détection de hauteur pour certains notes spécifiques qui indiquent l'arrivée à des endroits spécifiques de la partition.

## 3.3. Stratégies d'écriture pour modulateurs

Le travail musical avec modulation en anneau et/ou *frequency shifter* tourne autour de trois axes qui correspondent à trois zones de fréquences :

1. Création de *battements* avec des fréquences de modulation sub-audio – de 0 Hz à 20 Hz, mais les valeurs les plus marquées sont surtout de 0 à 10 Hz.
2. Création d'un instrument *préparé* où chaque note de l'instrument sonne avec un timbre différent. Dans ce cas on *accorde* la modulation à une note quelque part au milieu de la tessiture du jeu.
3. Un effet de *scintillement* ou une sorte de *pédale aigue* en utilisant une fréquence de modulation bien supérieure aux notes jouées par l'instrument.

Pour les deux premières techniques, le son transformé est tellement lié à l'instrument acoustique que nous avons l'impression que l'instrument lui-même a été altéré. Par exemple, si nous utilisons la modulation pour créer l'impression des *multiphoniques* sur l'instrument, il est très difficile pour l'auditeur d'entendre que c'est un traitement et non un vrai multiphonique. Le même phénomène s'applique pour les battements qu'on pense fait par l'instrumentiste.

<sup>5</sup> La création du *Patch bien tempéré II* a eu lieu le 27 novembre 2012 dans le cadre les Journées Musiques Mixtes III à l'Université Paul Valéry (Montpellier). Un enregistrement en studio a été fait le 5 février 2013 au ZKM (Karlsruhe) pour le CD *ENLARGE YOUR SAX* du saxophoniste Pedro Bittencourt (à paraître).

Le troisième technique se détache de l'instrument car le scintillement aigu est bien au dessus de la tessiture de l'instrument et ne change pas de façon significative avec chaque note. Dans cette zone de fréquences, nous avons plutôt un « effet » qu'une réelle transformation de l'instrument.

En travaillant musicalement avec un instrument acoustique et la modulation en anneau, il faut penser aux possibilités de ces trois « axes » de travail, en trouvant des moyens de changer la fréquence de modulation pour varier le résultat musical - de préférence lié aux phrases musicales.

On peut également mettre plusieurs modulateurs en parallèle pour densifier le résultat ou n'utiliser qu'un côté d'un *frequency shifter* pour épurer la densité.

On peut aussi modifier les paramètres de la modulation ou le niveau d'écoute de la modulation avec des détections de paramètres du son joué comme l'enveloppe d'amplitude ou la hauteur.

### 3.4. L'analyse et l'écriture

La pièce est composée en trois sections, chaque étape dans l'évolution d'une *histoire* entre le saxophone et les modulateurs en anneau, pour essayer de *s'entendre* et de construire quelque chose ensemble. Les trois sections sont assez différentes, mais le sens de résolution et le retour au rythme *respiré* de la troisième fait penser plutôt à une forme A-B-A'. Chaque section développe un rapport spécifique entre le saxophone et les modulateurs en anneau, avec un caractère bien défini et un rôle précis dans l'*histoire*.

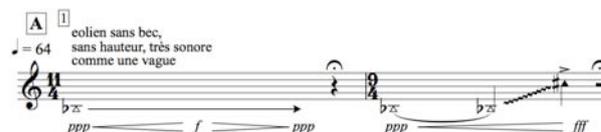
#### 3.4.1. Section A – « unisson »

Le principe central de la première section est l'**unisson** et l'utilisation des modulateurs en anneau pour *tirer* sur cet unisson et leur permettre de s'écarter et de dévier vers d'autres hauteurs et timbres. Le saxophone semble « coincé » sur une seule note (Fa écrit), comme s'il n'arrivait pas s'en sortir et c'est grâce aux modulateurs en anneau qu'il a finalement la possibilité de la quitter et partir ailleurs.

La première utilisation des modulateurs en anneau a comme but de créer des battements. On peut créer des battements soit en employant une fréquence de modulation dans la zone sub-audio (< 20 Hz), soit en « s'accordant » à moins de 20 Hz de la note modulée [DAVIES, 1976, pp.4-5].

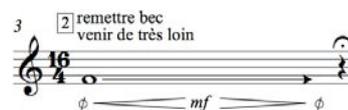
Le saxophoniste commence à jouer la pièce sans le bec<sup>6</sup>, pour produire deux longs sons de souffle (sons éoliens sans hauteurs déterminées) qui subissent un effet de battement grâce aux fréquences sub-audio des quatre modulateurs en anneau. Un crescendo/decrescendo fait apparaître et disparaître progressivement les modulateurs

en anneau, tout en modifiant légèrement les fréquences de modulation. Pour être plus précis, le modulateur 1 change entre 0.5 Hz à 2.6 Hz durant le crescendo/decrescendo, le 2 change entre 0.77 Hz à 4.3 Hz, le 3 entre 1.1 Hz et 6.1 Hz, et le 4 entre 1.4 Hz et 8.1 Hz. Le résultat est que les battements apparaissent un par un en accélérant pendant le crescendo, puis disparaissent un par un, en ralentissant pendant le decrescendo. Après un point d'orgue, un deuxième souffle éolien fait un crescendo, ouvrant et modifiant les modulateurs en anneau avec les mêmes paramètres que le premier souffle, mais cette fois-ci il termine subitement sur un triple forte. (voir **Figure 8**).



**Figure 8.** Les premières deux mesures de la partie pour saxophone.

A la mesure 3 (*événement 2*, comme indiqué par le numéro encadré) le saxophoniste remet le bec et commence à jouer le Fa grave (note écrite). Les crescendos/ decrescendos continuent à faire apparaître un par un les modulateurs en anneau, mais le rapport entre l'intensité et la fréquence de modulation va changer un peu (voir **Figure 9**).



**Figure 9.** Mesure 3 : l'arrivée du Fa.

Le premier modulateur que nous entendons dès le premier son du saxophone se met autour de l'unisson de la note Fa (note MIDI 65) alors que les 3 autres restent dans la zone de sub-audio. Comme nous l'avons rappelé plus haut dans la partie Le dispositif, une modulation en anneau à l'unisson rehausse tout le spectre par la valeur de la fondamentale [DAVIES, 1976, p. 3] créant ainsi un spectre qui est très proche de l'original à l'exception de la disparition de la fondamentale – comblée finalement par le son direct du saxophone. Le résultat est la quasi non-existence du traitement – un simple son amplifié. Pourtant, le premier modulateur en anneau se décale légèrement en suivant l'enveloppe d'amplitude du saxophone et varie en conséquence entre un dixième de demi-ton plus bas et un dixième de demi-ton plus haut (entre note MIDI 64.9 et 65.1). Pour rappel, toutes ces valeurs sont relatives à la transposition de l'instrument – pour le sax alto (Eb) : -9 demi-tons, pour le sax ténor (Sib) : -14, etc.

A la mesure 4, le saxophone refait un crescendo/decrescendo sur Fa, mais cette fois-ci introduit de très légers glissandi microtonaux d'un huitième de ton plus haut (voir **Figure 10**). Le modulateur 1 augmente un petit peu l'ambitus de

<sup>6</sup> Dans la première version, il était proposé de jouer cette partie de sons éoliens avec le bec, mais cela ne fonctionnait pas bien. En enlevant le bec, plus de fréquences peuvent être utilisées (bruits blancs).

réponse à l'amplitude, glissant entre plus et moins un quart de ton du Fa (entre note MIDI 64.5 et 65.5).

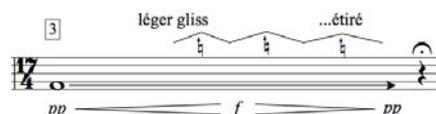


Figure 10. Mesure 4 : léger glissando.

A la mesure 5, l'ambitus du modulateur augmente jusqu'à plus et moins un ton (note MIDI 64 à 66) et les battements deviennent encore plus importants. Un saut d'octave du saxophone provoque un effet de *bisbigliando* car la somme et la différence de la modulation en anneau une octave au dessus du son modulé « remet » les harmoniques manquants après le saut d'octave, mais à des intensités différentes. Dans le **Table 1** nous voyons qu'un spectre harmonique avec une fondamentale hypothétique de 200 Hz modulé par 100 Hz devient un spectre d'harmoniques impaires à une fondamentale de 100 Hz.

Modulation	Spectre porteur	somme et différence	Résultat avec son acoustique
	800		800
		600 (2 fois)	700
	600		600
		500 (2 fois)	500
	400		400
		300 (2 fois)	300
	200		200
100		100	100

Table 1. Modulation une octave en dessus de la note jouée.

Quand nous remixons le son acoustique et le résultat de la modulation, nous retrouvons un spectre complet à 100 Hz avec des harmoniques doublés et par conséquent plus fortes. Dans le cas donc du saut d'octave de la mesure 7, nous entendons plutôt un changement de timbre (voir **Figure 11**).



Figure 11. Saut d'octave dans les mesures 5 à 7.

A la mesure 8, le saxophone revient sur le Fa grave et ne fait que des changements d'intensité. Le modulateur 1 augmente encore son ambitus en réponse à l'amplitude du saxophone, atteignant ainsi plus ou moins un ton entier (note 63 à note 67). Cette différence de hauteur génère un son résultat d'environ un ton de perturbation, sortant ainsi de la zone des battements et entrant dans la zone du décalage de spectre et l'inharmonicité dans le timbre. Cela permet de passer de nouveau à des glissandi du saxophone à la mesure 11 (comme à la mesure 4), cette fois-ci d'un quart de ton. A mesure 11, l'ambitus du changement de modulation du modulateur 1 garde la même taille mais monte un demi-ton (note 64 à note 68),

tirant ainsi tout le résultat légèrement vers le haut (voir **Figure 12**).

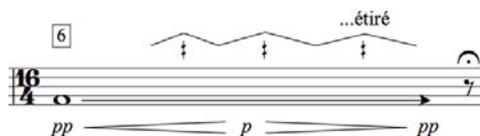


Figure 12. Les glissandi de la mesure 11.

Aux mesures 12 à 13, le saxophone va faire un vrai changement de hauteur de Fa vers Mi. Le modulateur 1 va réagir à l'amplitude du saxophone en *baissant* sa fréquence avec la *montée* d'intensité. Quand la phrase commence *piano*, le modulateur 1 est accordé en unisson avec le Fa (65), mais au fur et à mesure du crescendo, la fréquence baisse pour s'accorder sur un Mib (note 63), montant ainsi le nombre de battements. Au moment du *forte*, quand le saxophone descend à Mi, le modulateur est plus proche de la note jouée et les battements diminuent. Le decrescendo sur le Mi fait remonter le modulateur 1 à la fréquence du Fa, et les battements remontent en même temps que la disparition du son (voir **Figure 13**).



Figure 13. Mesures 12 à 13 et le passage au Mi.

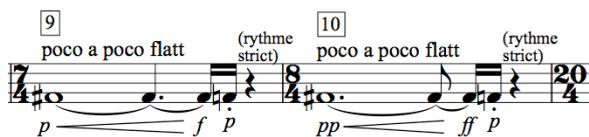
Aux mesures 14 à 16 il se passe la même chose mais une octave plus haut et vers une note un demi-ton au dessus. Le modulateur 1 s'accorde parfaitement sur le Fa que joue le saxophone et le crescendo pousse vers le haut, augmentant la vitesse des battements. Quand le saxophone passe au Fa# il est de nouveau accordé avec le modulateur, mais les changements d'intensité qui suivent jouent encore sur le rapprochement et l'éloignement et par conséquent sur le décalage et la vitesse des battements (voir **Figure 14**).



Figure 14. Mesures 14 à 15 et le passage au Fa#.

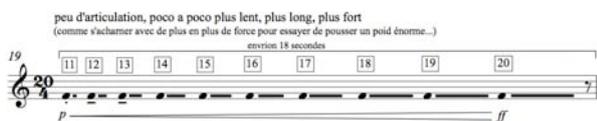
A la mesure 17, les modulateurs 2 à 4 qui étaient jusqu'alors toujours sur des fréquences sub-audio pour faire des battements commencent à s'accorder autour du Fa grave (note 65) avec des ambitus de glissando de l'ordre d'un dixième de ton. Le saxophone joue le Fa# grave (après avoir réussi à jouer le Fa# plus aigu précédemment) et se met en rapport de forts battements avec les quatre modulateurs le temps de son crescendo là où les perturbations sont encore plus intense grâce au *flatterzung*. L'arrivée sur Fa bécarre en *subito piano* conclut la phrase. A la mesure 18, il arrive la même

chose avec un peu plus de battements dus à un ambitus de glissandi des modulateurs un peu plus grand (voir **Figure 15**).



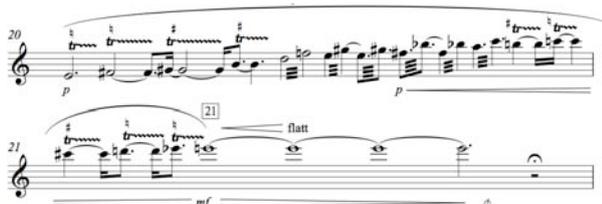
**Figure 15.** Mesures 17 et 18.

A la mesure 19 la tension va monter jusqu'au point culminant où le saxophone répète la note Fa avec de plus en plus d'insistance et de résistance – les notes de plus en plus intenses et de plus en plus longues. A chaque note, un nouvel événement modifie les notes (et donc les fréquences) des modulateurs en anneau pour créer des *accords* avec des inharmoniques de plus en plus riches et étalés dans le spectre (voir **Figure 16**).



**Figure 16.** Mesure 19 et la montée des *accords* inharmoniques.

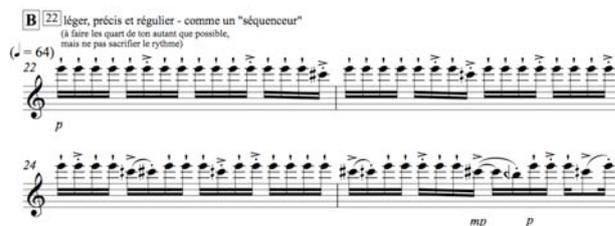
Étant passé par un « mur du son » métaphorique, le saxophone commence à la mesure 20 sur une montée en trilles et trémolos vers le Mi aigu. Les modulateurs restent dans l'état du dernier *accord*. Au moment de l'arrivée un événement se déclenche qui provoque un glissando des quatre modulateurs lentement vers l'unisson du Mi aigu, pendant le long decrescendo du saxophone. Un *flatterzung* souligne les perturbations du glissando, mais on retrouve la « pureté » du Mi aigu à la fin du decrescendo (voir **Figure 17**).



**Figure 17.** Trilles, trémolos et le nouvel unisson.

### 3.4.2. Section B – « expansion »

La partie B introduit la pulsation et l'expansion de la mélodie. Elle commence sur le Mi aigu et s'ouvre ensuite progressivement vers des notes plus graves et de plus en plus étalées dans le spectre. Les 4 modulateurs en anneau sont accordés à l'unisson avec le Mi aigu, puis progressivement sur d'autres notes de la progression. Après une longue expansion, la tessiture se resserre dans le grave vers la fin et se recentre sur le Fa# grave pour préparer la section C (voir **Figure 18**).



**Figure 18.** Le début de la section B – mes. 22 à 25.

La fonction des modulateurs en anneau dans cette section n'est plus de créer des battements, perturbations ou glissandi comme dans la section précédente, mais de créer une *préparation* du saxophone, soit un timbre différent à chaque nouvelle note, et de fournir un contexte microtonal qui motive l'écriture en quart de tons pour le saxophone.

### 3.4.3. Section C (A') – « résolution »

La troisième section est le moment d'apaisement où la mélodie semble flotter à travers le champs de modulation en profitant de son soutien et en le guidant.

Les modulateurs en anneau ne réagissent plus aux changements d'amplitude du saxophone. Ils sont accordés au début au Fa# grave (note 66) à l'événement 27, au Fa grave (note 65) à l'événement 28, au Mi (note 64) à l'événement 29, au Fa à l'événement 30, à nouveau au Mi (note 64) à l'événement 31, au Fa à l'événement 32, au Fa grave et Mi aigu à l'événement 33 (notes 65 et 88), puis un peu en dessus de Mi aigu (entre 87.75 et 87.9) à l'événement 34. Ces derniers réglages arrivent au moment où le saxophone tient une note Fa aigu (note 89) qui devient graduellement un son éolien. La différence avec la modulation produit une belle fréquence grave pour le long decrescendo de la fin (voir **Figure 19**).



**Figure 19.** Les dernières mesures de la pièce.

### 3.4.4. Les hauteurs

Les hauteurs de la pièce ont été déterminées en utilisant le calcul de la modulation en anneau – la somme et la différence entre une fréquence de modulation et la porteuse (saxophone).

Par exemple, si nous jouons toutes les notes du saxophone avec différentes « notes » de modulation en anneau, nous produisons des gammes en micro-intervalles assez riches. Sur la **Figure 20** nous voyons la gamme produite sur une fréquence de modulation égale à la fréquence de la note Fa# grave (note 66).



**Figure 20.** Gamme sur modulation par Fa#.

#### 4. LES ADAPTATIONS

Après avoir présenté en détail l'analyse musicale et technique du *Le patch Bien Tempéré II* et pour revenir aux aspects pratiques de l'*interprétation participative*, nous listons maintenant les adaptations spécifiques survenues lors des séances de travail entre l'instrumentiste et le compositeur :

- 1 (indication de la partition) — jeu sans le bec pour avoir plus de volume avec le son éolien et générer les battements avec les modulations en anneau ;
- 3 — laisser la clé Eb enfoncée et ouvrir doucement et à peine la clé Tf (moins de ¼ de ton), tout en compensant avec l'embouchure quand nécessaire ;
- 6 — la même chose que 3, mais il faut ouvrir un peu plus la clé Tf pour que l'intervalle soit un peu plus grand que la première fois, autour d'un ¼ de ton;
- 11 à 20 – choix de la notation proportionnelle ;
- 21 – il faut être attentif à mener le decrescendo final et la justesse du Mi en fonction de l'arrivée de l'électronique à l'unisson ;
- 22 à 26 – plusieurs nuances ont été modifiées après l'écoute de l'enregistrement de la création. Tous les ¼ de ton sont faisables au sax alto, à deux exceptions : 1) sol ¼ #, qui est toujours trop bas (avec clé Tf) et parfois difficile mécaniquement ;
- 26 — Si grave ¼ # difficile. Remarque : sur d'autres saxophones de la famille (sopranino, soprano, ténor, baryton, basse), cela peut se comporter différemment, mais nous ne l'avons pas encore testé.
- 27 à 34 — on remarque ici l'utilisation des sons éoliens intégrés à la fin des notes et les vibratos lents d'intensité. C'est comme un rappel des souffles et du « léger glissandi, comme un vibrato lent » de la première partie.

#### 5. CONCLUSIONS

La modulation en anneau est un des « classiques » parmi les traitements audio qui existent. Bien que l'on puisse dire que les implémentations numériques sonnent un peu moins bien que leurs aînés « analogiques », la souplesse avec le contrôle et l'intégration d'autres paramètres dans un dispositif en temps réel donne de réels avantages.

Dans le *Patch Bien Tempéré II*, nous avons vu en détail des applications musicales de la modulation en anneau – battements, coloration par note, création de sons supplémentaires graves ou aigus. Nous avons vu également l'intégration avec le jeu du saxophoniste, avant tout par son niveau d'intensité, mais aussi les changements dans son jeu en fonction du retour du son.

Le saxophoniste construit son phrasé en fonction des transformations souhaitées par les modulateurs. Les remarques du compositeur ont demandé une écoute très attentive des battements et autres transformations pour arriver à des résultats satisfaisants en rapport avec l'électronique. Les informations sur le comportement de l'électronique ne sont pas encore disponibles sur la partition actuelle et pourraient être rajoutées pour clarifier les effets souhaités. L'interprétation participative qui comporte de nombreuses échanges et adaptations pendant les sessions de travail, répétitions et enregistrements a été fondamentale pour que de cette pièce puisse être développée, interprétée, pour ouvrir des perspectives de versions futures avec d'autres instruments de la famille du sax et d'autres saxophonistes désirant jouer la pièce.

#### 6. REFERENCES

- [1] BARKATI, K. [2009] « La polysémie du temps réel et du temps différé », *Actes des Journées d'Informatique Musicale*, Grenoble, 2009. Disponible sous forme électronique : <http://acroe.imag.fr/jim09/index.php/descrip/conf/schedConf/actes>, consulté le 27.02.13
- [2] BITTENCOURT, P. S. [2013] « Interpretação participativa na música mista contemporânea », *Revista Interfaces* n°18, Vol. I/2013, Centro de Letras e Artes UFRJ, éd. 7 Letras, Rio de Janeiro, à paraître, disponible en forme électronique: <http://www.cla.ufrj.br/index.php/periodicos/revista-interfaces>
- [3] BODE, H, [1984] « History of electronic sound modification », *Journal of the Audio Engineering Society*, vol 32, no. 10. pp. 730-738. Disponible sous forme électronique : [http://cec.sonus.ca/econtact/13\\_4/bode\\_history.html](http://cec.sonus.ca/econtact/13_4/bode_history.html), consulté le 27.02.13
- [4] DAHAN, K., LALIBERTÉ, M. [2008] « Réflexions autour de la question d'interprétation de la musique électroacoustique », *Actes des Journées d'Informatique Musicale (JIM)*, Albi, 2008. Disponible sous forme électronique : <http://gmea.net/jim08/index.php/Articles>, consulté le 27.02.13
- [5] DAVIES, H. [1976] « A Simple Ring Modulator », in *MUSICS - an impromptu experivisation arts magazine*, numéro 6, publication indépendante, Londres. Disponible sous forme électronique : [http://electro-music.com/forum/phpbb-files/hughdaviesringmod\\_140.pdf](http://electro-music.com/forum/phpbb-files/hughdaviesringmod_140.pdf), consulté le 27.02.13
- [6] DURAND, D. [1979] *La systématique*, Collection Encyclopédique Que sais-je ?, dixième édition mise à jour [2006], éd. Presses Universitaires Françaises, Paris

- [7] LIPPE, C. [2002] « Real-Time interaction among composers, performers and computer systems », in *Information Processing Society of Japan SIG Notes* Volume 2002, Number 123, pp. 1-6, Disponible sous forme électronique : <http://www.music.buffalo.edu/faculty/lippe/lippepublications.shtml>, consulté le 27.02.13
- [8] MAYS, T [2010] *TRAITEMENTS TEMPS REEL ET ECRITURE : Vers un lexique musical de divers traitements sonores de base*, mémoire de Master II, Université de Paris VIII, p. 82, Paris
- [9] MAYS, T., RUBIANO, R., [2010] « Tapemovie : Un environnement logiciel pour la création temps réel intermédia », *Actes des Journées d'Informatique Musicale 2010*, Rennes. Disponible en forme électronique : <http://jim10.afim-asso.org/actes/81mays.pdf>, consulté le 27.02.13
- [10] MORIN, A. [2010] *Cheminer ensemble dans la réalité complexe. La recherche-action intégrale et systémique*, éd. L'Harmattan, Paris
- [11] OBERHEIM, T [2008] Lecture/Interveiw pour le Red Bull Music Academy, Barcelona. Transcript disponible sous forme électronique : [http://www.redbullmusicacademy.com/lectures/to-m-oberheim-polyphonic-one-love?template=RBMA\\_Lecture%2Ftranscript](http://www.redbullmusicacademy.com/lectures/to-m-oberheim-polyphonic-one-love?template=RBMA_Lecture%2Ftranscript), consulté le 27.02.13
- [12] PARKER, Julian [2011] « A Simple Digital Model of the Diode-Based Ring-Modulator », *Proceedings of the 14th Internatinal Conference on Digital Audio Effects*, Paris. Disponible en forme électronique: [http://www.academia.edu/1047370/A\\_Simple\\_Digital\\_Model\\_of\\_the\\_Diode-Based\\_Ring-Modulator](http://www.academia.edu/1047370/A_Simple_Digital_Model_of_the_Diode-Based_Ring-Modulator), consulté le 27.02.13
- [13] VARELA, F. [1989] *Invitation aux sciences cognitives*, éd. Du Seuil, Paris
- [14] WATZLAWICK, P., BEAVIN, J.H., JACKSON, D.D. [1972] *Une logique de la communication*, éd. du Seuil, Paris

# FAUST2ANDROID : UNE ARCHITECTURE FAUST POUR ANDROID

*Romain Michon*  
CCRMA  
Department of Music  
Stanford University, CA 94305  
USA  
rmichon@ccrma.stanford.edu

## RÉSUMÉ

*faust2android* est un programme qui permet de compiler un code FAUST en une application pour terminal Android. Les tâches de traitement du signal ainsi que l'accès aux ressources audios se font de manière native en C++ grâce au *Native Development Toolkit (NDK)* d'Android. L'interface utilisateur et les autres éléments de l'application sont programmés en JAVA.

*faust2android* fait partie d'un projet plus large qui vise à implémenter un environnement d'utilisation d'objets FAUST sous Android : FAUSTDROID.

*faust2android*<sup>4</sup> est un programme qui permet de générer une application Android à partir d'un code FAUST[3]. Il utilise un script qui place le code C++ généré par le compilateur de FAUST dans une application Android « modèle » au contenu dynamique.

*faust2android* a été implémenté dans le cadre d'un projet plus large : FAUSTDROID. Cette application va permettre de mettre à disposition du musicien un environnement dans lequel il sera possible de télécharger de manière simple des objets FAUST contenus dans un catalogue en-ligne, de les connecter entre eux et de les utiliser en temps réel sur un terminal Android.

## 1. INTRODUCTION

Alors que les iPads et les iPhones ont été au cours des dernières années des plates-formes privilégiées pour le développement d'applications pour le traitement du signal en temps réel, les terminaux Androids ont été délaissés, principalement à cause de l'importante latence audio générée par leur système d'exploitation. Cependant, de récents développements<sup>1</sup> montrent que Google ainsi que d'autres entreprises<sup>2</sup> semblent s'intéresser à résoudre ce problème.

Une autre avancée dans le domaine de l'audio pour Android concerne la possibilité de pouvoir connecter une interface audio externe sur un terminal Android. Bien qu'il soit encore loin d'être simple d'accéder dans une application Android à une ressource audio USB externe, nous avons été en mesure d'utiliser une interface Behringer GUITAR LINK UCG102<sup>3</sup> sur une tablette Google Nexus 7 sans augmenter la latence audio et en la réduisant même de quelques millisecondes.

Ces différentes observations montrent que l'utilisation fiable de terminaux Android pour le traitement du signal en temps réel devrait devenir possible dans un futur proche. Cela a grandement motivé les travaux présentés dans cet article.

1. <http://developer.android.com/about/versions/jelly-bean.html>.

2. <http://www.sonomawireworks.com/>.

3. <http://www.behringer.com/EN/Products/UCG102.aspx>.

## 2. ANDROID ET TRAITEMENT DU SIGNAL EN TEMPS RÉEL

### 2.1. La question de la latence

Android a toujours été connu par la communauté de développeurs de logiciels comme une plateforme inutilisable pour l'audio en temps réel à cause de l'importance de la latence qui la caractérisait. Cependant, ce problème a été récemment traité par Google dans la dernière version de son système d'exploitation (Jelly Bean 4.2). Les témoignages de développeurs sur cette amélioration sont nombreux. Par exemple, Victor Lazzarini explique dans un billet de décembre 2012 posté sur son blog<sup>5</sup> qu'il a été en mesure d'obtenir des latences inférieures à 100ms pour de simples traitements de son et inférieures à 120ms pour des synthétiseurs utilisant une interface graphique (« Touch to Sounds Latency »).

Nous avons pu obtenir avec *faust2android* des résultats similaires pour le traitement du son et meilleurs pour la synthèse avec contrôle via une interface graphique (environ 30ms). Bien qu'encore loin d'être suffisante pour une utilisation dans un contexte musical, ces performances surpassent de loin celles des précédentes versions d'Android qui offraient des latences supérieures

4. *faust2android* est disponible dans le dossier */tools/* du référentiel de FAUST : <http://sourceforge.net/projects/faudiostream/>. Il peut également être utilisé directement dans le compilateur en ligne de FAUST : <http://faust.gramme.fr/index.php/online-examples>.

5. <http://audioprogramming.wordpress.com/category/android/>.

à 300ms.

## 2.2. C ou JAVA ?

La plupart des applications Android est programmée en *JAVA* et le *SDK* de Android contient une API pour l'audio en temps-réel. Ainsi, une classe pour le traitement du signal peut être implémentée directement en *JAVA* ce qui permet de simplifier grandement l'architecture générale de l'application. De plus, *FAUST2* permet maintenant de générer du code *JAVA* au lieu de *C++* ce qui est un argument supplémentaire pour le choix de *JAVA*.

Plusieurs tests lors desquels des codes *JAVA* générés par *FAUST* ont été manuellement copiés dans une application Android puis exécutés sur un téléphone Samsung Galaxy S2 et une tablette Google Nexus 7 ont été menés. Tandis que les résultats de ces expériences ont grandement variés d'un terminal à un autre (nous n'avons par exemple pas été en mesure d'accéder aux ressources pour l'acquisition audio sur le Galaxy S2), ils ont été de manière générale assez décevants, principalement à cause de l'instabilité des applications et de l'importance de la latence.

A l'inverse, Victor Lazzarini a décrit sur son blog dans un billet de mars 2012 <sup>6</sup> une technique permettant d'accéder aux ressources audios temps-réel d'un terminal Android de manière native en utilisant *OpenSL ES* <sup>7</sup> et le *NDK* <sup>8</sup>. Après plusieurs tests, cette technique s'est avérée être beaucoup plus efficace et stable que celle n'utilisant que du code *JAVA* et a été utilisée pour implémenter *faust2android*.

## 2.3. Audio en temps-réel avec faust2android

Comme cela a été mentionné dans la partie 2.2, le *NDK* de Android rend possible l'utilisation de classes *C++* dans une application *JAVA* en les compilant sous la forme d'une *shared library*. Cette tâche est grandement simplifiée par *SWIG* <sup>9</sup> qui crée les fichiers d'interface nécessaires entre les deux langages de manière presque automatisée.

Dans une application générée par *faust2android*, les différentes tâches sont réparties de la manière suivante entre *JAVA* et *C++* :

JAVA	C++
- application Android	- traitement du signal
- interface utilisateur dynamique	(calcul d'une Frame audio)
- envois des valeurs des différents paramètres DSP à chaque Frame	- informations sur les paramètres DSP et sur l'interface utilisateur
	- gestion des ressources audios

6. <http://audioprograming.wordpress.com>.

7. *Open Sound Library for Embedded Systems* : <http://www.khronos.org/opensles/>.

8. *Native Development Toolkit* : <http://developer.android.com/tools/sdk/ndk/>.

9. *Simplified Wrapper and Interface Generator* : <http://www.swig.org/>.

La gestion des ressources audios en *C++* est menée à bien à l'aide de l'API <sup>10</sup> mise en place par Victor Lazzarini qui fournit des fonctions de haut niveau à *OpenSL ES* sur Android.

Pour résumer, une application générée par *faust2android* effectue les tâches de traitement du signal nativement ce qui est beaucoup plus efficace et stable que si cela était fait directement en *JAVA*.

## 3. IMPLÉMENTATION DE FAUST2ANDROID

### 3.1. Générer le code

A la différence d'autres architectures *FAUST*, *faust2android* ne peut pas générer un simple fichier *C++* contenant l'ensemble des éléments nécessaires à la création d'une application complète. En effet, comme cela a été expliqué précédemment, une application *faust2android* contient des fichiers *JAVA*, *C++* et *XML*.

*faust2android* utilise un simple script *Bash* pour mener à bien les différentes actions permettant de créer l'application. Dans un premier temps, ce dernier appelle le compilateur de *FAUST* qui produit du code *C++*. Ce code est alors copié dans un fichier d'architecture qui permet de l'interfacer avec une application Android « modèle » dont le contenu totalement dynamique s'adapte en fonction des informations contenues dans le code généré par *FAUST*.

*faust2android* peut également créer de manière optionnelle un projet éclipse si l'utilisateur souhaite modifier manuellement le contenu de l'application.

Enfin, le cross-compilateur d'Android est appelé par le script pour générer le fichier binaire de l'application. Une option permet de charger ce dernier sur une tablette ou un téléphone connecté en USB à la machine qui effectue ces opérations. Un résumé du fonctionnement de *faust2android* peut être vu dans la figure 1.

### 3.2. Interface utilisateur

Bien que la diversité des *widgets* fournis par défaut avec le *SDK* de Android soit assez limitée, ils sont actuellement utilisés pour construire les différents contrôleurs de paramètres d'une application générée par *faust2android*. En effet, alors qu'une architecture *FAUST* standard autorise la création de potentiomètres horizontaux et verticaux, de potards, de cadrans digitaux, de boutons et de groupes verticaux et horizontaux, à la date de la rédaction de cet article, *faust2android* ne permet la mise en place que de potentiomètres horizontaux, de cadrans digitaux, de boutons et de groupes horizontaux et verticaux.

Par conséquent, si un potentiomètre vertical ou un potard est déclaré dans un code *FAUST*, il est automatiquement converti en potentiomètre horizontal.

10. <https://bitbucket.org/victorlazzarini/android-audiotest>.

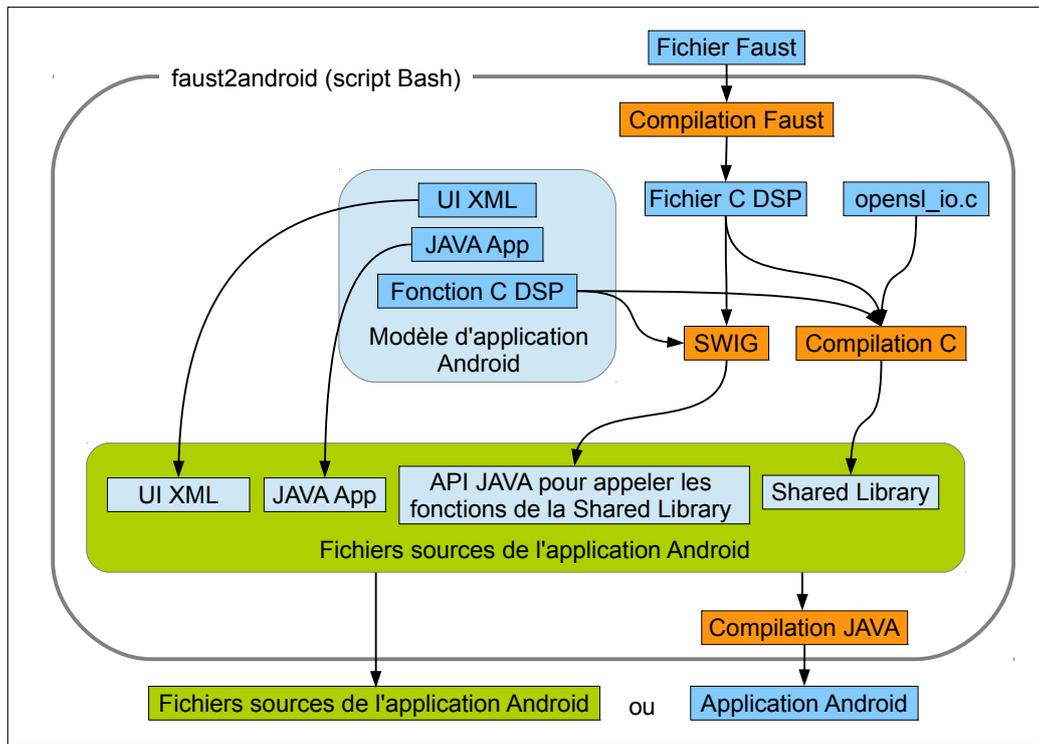


Figure 1. Résumé du fonctionnement de *faust2android*.

La figure 2 présente un exemple d'interface utilisateur générée à partir du code suivant qui implémente un simple synthétiseur FM :

```
import("music.lib");
import("filter.lib");

freqMod = hslider("v:Modulator/Frequency",
    440, 20, 15000, 1);
modIndex = hslider("v:Modulator/Modulation
    Index", 0, 0, 1000, 0.1);
freq = nentry("v:Carrier/Frequency", 440,
    20, 8000, 1);
vol = hslider("v:General Parameters/Volume
    ", 0, -96, 0, 0.1) : db2linear;
bal = hslider("v:General Parameters/
    Balance", 0.5, 0, 1, 0.1);
gate = button("gate");

process = osc(freqMod)*modIndex+freq : osc
    * gate * vol <: *(bal),*(1-bal);
```

#### 4. PROJETS FUTURS

*faust2android* a été développé dans le cadre d'un projet plus large visant à combler le vide entre les musiciens et la communauté FAUST de développeur *OpenSource* qui travaille entre autre à l'implémentation d'algorithmes de traitement du signal.

Ce système sera basé sur une application Android appelée FAUSTDROID où un utilisateur pourra facilement accéder un catalogue en-ligne à partir duquel il sera possible de télécharger des objets FAUST et de les utiliser

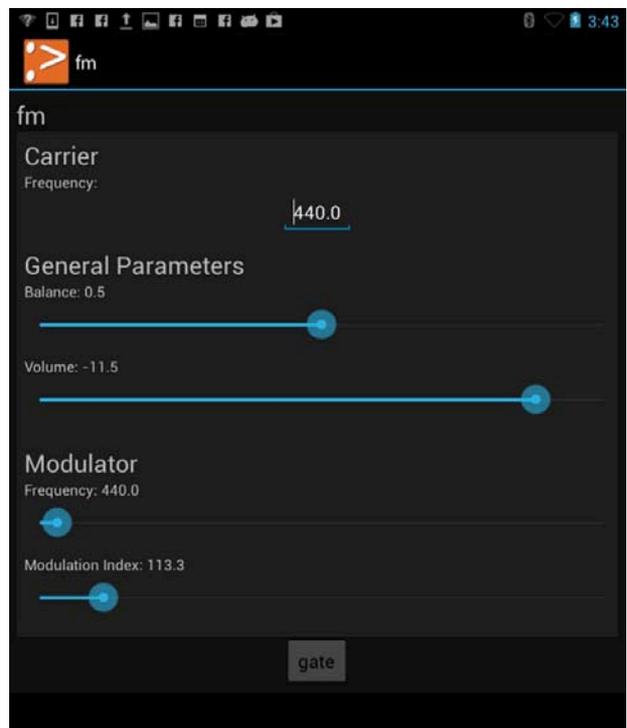


Figure 2. Exemple d'interface graphique d'une application Android générée par *faust2android*.

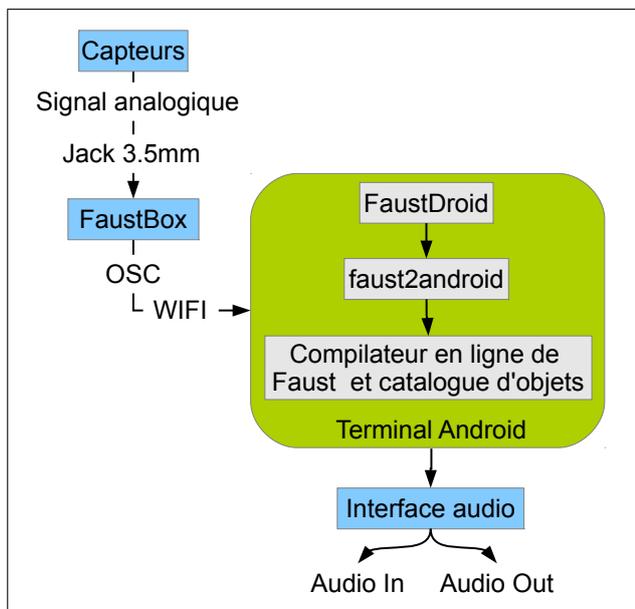


Figure 3. Résumé du projet FAUSTDROID.

en les connectant entre-eux et en les contrôlant via une interface graphique aussi simple que possible.

Ce projet est résumé par le diagramme présenté dans la figure 3.

#### 4.1. FaustDroid

De récents développements menés à GRAME<sup>11</sup> autour du compilateur en ligne de FAUST permettent l'utilisation d'une API RESTFull pour compiler du code FAUST à distance pour différentes architectures et systèmes d'exploitation<sup>12</sup>. Le compilateur en ligne de FAUST met également déjà à disposition un catalogue d'objets pouvant être édité par n'importe quel internaute pour ajouter ou modifier des éléments.

FAUSTDROID offrira la possibilité d'accéder à ce catalogue de la manière la plus simple possible, de compiler certains de ces objets avec *faust2android* et d'intégrer l'application résultante dans un espace de travail où plusieurs objets pourront être assemblés et connectés.

#### 4.2. La FaustBox

Un autre élément important du projet FAUSTDROID est la FAUSTBOX. Cet objet sera basé sur un Arduino Uno et un Arduino Wifi Shield<sup>13</sup> (cf. figure 4) et sera alimenté par une batterie. Elle permettra l'utilisation de capteurs pour contrôler les différents paramètres d'un objet FAUST fonctionnant dans FAUSTDROID.

La FAUSTBOX enverra des messages OSC au terminal Android via le WIFI pour communiquer avec FAUSTDROID.

11. <http://grame.fr>.

12. [git://faudiostream.git.sourceforge.net/](http://git://faudiostream.git.sourceforge.net/)  
[gitroot/faudiostream/FaustWeb](http://gitroot/faudiostream/FaustWeb).

13. [www.arduino.cc](http://www.arduino.cc).



Figure 4. Arduino Uno with a Wifi Shield.

Enfin, chaque capteur sera alimenté en électricité et enverra son signal analogique via une prise Jack stéréo 3.5mm qu'il sera possible de connecter à la FAUSTBOX. Dans la mesure où les signaux des capteurs auront tous la même échelle, il sera possible de les interchanger très facilement.

## 5. CONCLUSION

Tandis que des dizaines d'applications pour le traitement du signal en temps différé sont actuellement disponibles sur le *Google Play Store*<sup>14</sup>, seules quelques unes d'entre elles offrent des possibilités en temps réel. Il est aisé d'en déduire que l'origine de ce manque est due à la mauvaise réputation de Android concernant la latence audio.

Cependant, comme cela a été expliqué dans cet article, il semblerait que Google déploie beaucoup d'effort pour trouver une solution à ce problème comme le prouve la plus récente version d'Android.

Plus le nombre de personnes montrant de l'intérêt pour ce sujet sera important, plus Google s'attachera à fournir des solutions efficaces pour réduire la latence de son système d'exploitation afin de le rendre utilisable pour le traitement du signal en temps réel à des fins musicales.

*faust2android* est juste une pierre de plus à l'édifice et nous ne pouvons qu'espérer que d'autres travaux dans un futur proche contribueront à convaincre Google de réduire encore plus la latence audio d'Android.

## 6. REFERENCES

- [1] Michon, R. ; Orlarey, Y., « The Faust online compiler : a web-based IDE for the Faust programming language », *Proceedings of the Linux Audio Conference (LAC)*, CCRMA – Université Stanford, USA, 2012.

14. <https://play.google.com/store>.

- [2] Michon, R. ; Smith, J. « Faust-STK : a set of linear and nonlinear physical models for the Faust programming language », *Proceedings of the Conference on Digital Audio Effects (DAFx-11)*, IRCAM, Paris, France, 2011.
- [3] Orlarey, Y. ; Fober, D. ; Letz, S. « An algebra for block diagram languages », *Proceedings of the International Computer Music Conference (ICMA)*, Gothenburg, Suède, 2002.



# COMMENT EMBARQUER LE COMPILATEUR FAUST DANS VOS APPLICATIONS ?

*Stéphane Letz*  
GRAME  
letz@grame.fr

*Dominique Fober*  
GRAME  
fober@grame.fr

*Yann Orlarey*  
GRAME  
orlarey@grame.fr

## RÉSUMÉ

Le compilateur Faust [1] est désormais disponible sous la forme d'une librairie nommée *libfaust*. Associée à la technologie LLVM, cette librairie peut être embarquée dans n'importe quelle application ou plugin audio, leur permettant ainsi de compiler et d'exécuter dynamiquement du code Faust de manière native, aussi efficacement que du code compilé traditionnel.

L'article présente la librairie *libfaust*, l'infrastructure de compilation LLVM, et deux applications de cette technologie : *faustgen* un plugin Max/MSP permettant d'éditer, de compiler et d'exécuter du code Faust depuis Max, et *FaustNode*, une extension de la WebAudio API permettant de compiler et d'exécuter du code Faust depuis un navigateur Web.

## 1. INTRODUCTION

Faust (Functional Audio Stream) est un langage de programmation spécifiquement conçu pour les applications musicales. Il permet de développer rapidement des applications efficaces et fiables pour les principales plateformes audios (plugins VST, Max/MSP, Puredata, applications OSX, iPhone, Linux, etc.).

A la différence d'autres environnements audio, Faust est un langage compilé. Un programme Faust décrit un processeur de signaux sous la forme d'une fonction qui prend un ensemble de signaux en entrée et produit un ensemble de signaux en sortie. C'est le calcul des signaux de sortie (en fonction des signaux d'entrée) qui est effectivement compilé, sous la forme d'une fonction *compute*, à appeler à chaque cycle audio avec les buffers audio d'entrée/sortie adaptés.

La version actuelle de Faust traduit le code DSP sous la forme d'une classe C++, à intégrer ensuite dans un fichier d'architecture. Celui-ci connecte le coeur de calcul généré par le compilateur aux entrées/sorties audio et à une interface de contrôle (qui peut être une interface utilisateur classique avec boutons, curseurs, zones de texte ou créée à partir d'une interface de pilotage utilisant le protocole OSC par exemple) [2]. Le code résultant est alors compilé avec un compilateur C/C++ classique, pour obtenir au final une application ou un plug-in opérationnel.

La branche de développement *faust2* permet de générer du code dans différents langages cibles, et en particulier

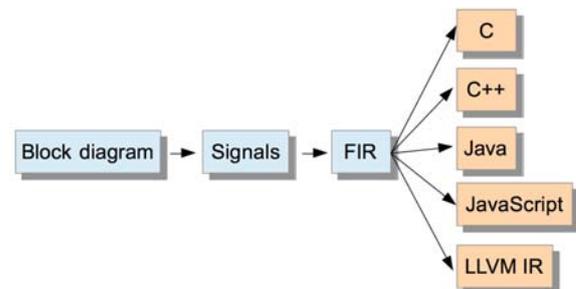


Figure 1. Etapes internes de la compilation dans Faust

du code intermédiaire LLVM qui peut ensuite être compilé dynamiquement sous la forme de code exécutable. Cette technologie permet ainsi d'embarquer une chaîne de compilation complète dans une application, du code DSP vers le code exécutable.

Nous montrerons comment cette nouvelle fonctionnalité a été mise en oeuvre dans *faustgen*, un objet externe pour Max/MSP, ainsi que dans *FaustNode*, une intégration de Faust dans la Web Audio API. Nous expliciterons ensuite l'API définie par *libfaust* et comment l'utiliser pour embarquer le compilateur Faust.

## 2. LA CHAÎNE DE COMPILATION

### 2.1. FIR : Faust Imperative Representation

Le compilation du code DSP source se passe en plusieurs étapes : le programme est d'abord évalué afin de construire le *bloc-diagramme* de traitement des signaux qu'il représente. Des signaux symboliques sont ensuite propagés dans ce bloc-diagramme pour établir ce que ce bloc-diagramme calcule, c'est à dire comment les signaux de sortie s'expriment en fonction des signaux d'entrée (Figure 1). C'est cette représentation interne des *signaux* qui est compilée directement dans une classe C++ dans la version actuelle de Faust.

Dans *faust2*, un langage intermédiaire nommé FIR (Faust Imperative Representation) a été défini pour décrire les calculs réalisés sur les échantillons de manière générique. Ce langage contient les opérations pour lire et écrire des variables et des tableaux, effectuer des opérations arith-

métriques et définir les structures de contrôle nécessaires (boucle *for*, boucle *while*, structure de test *if*...). Le langage des signaux est maintenant compilé dans ce langage FIR intermédiaire.

Le langage FIR est ensuite traduit en différents langages cible : C, C++, Java, JavaScript ou LLVM. Ce dernier langage cible est particulièrement intéressant dans une optique de compilation dynamique à l'intérieur même des applications ou plug-ins.

## 2.2. LLVM

LLVM (Low Level Virtual Machine) est une infrastructure de compilation composée de différents modules pour analyser des langages sources, générer du code intermédiaire, faire l'édition des liens avec du code défini de manière externe ou dans des bibliothèques, et finalement générer du code exécutable (soit de manière statique en vue de produire un fichier exécutable, soit de manière dynamique avec un compilateur à la volée) (Figure 2).

Le code de cette nouvelle architecture de compilation écrite en C++ est plus moderne et plus flexible que celle construite autour de GCC. Il est donc plus facile à utiliser dans des projets variés. LLVM est ainsi devenue la chaîne de compilation par défaut disponible sur OSX (sous la forme du compilateur *clang*, du débogueur *lldb* etc.).

Le code intermédiaire LLVM (LLVM IR [4]) est un langage bas niveau typé, qui représente les instructions sous la forme SSA<sup>1</sup>. Pour générer ce code, un backend FIR vers LLVM IR a été ajouté dans *faust2*. Il produit au final un module LLVM, qui regroupe les déclarations de structures de données (objet DSP et interface utilisateur...) ainsi que l'ensemble des fonctions pour le manipuler (*init*, *buildUserInterface*, *compute*...).

Le code LLVM IR peut ensuite être optimisé à l'aide de différentes passes d'optimisation, et enfin compilé à la volée en code natif exécutable dépendant de l'architecture sous-jacente (i386, x86\_64, PPC...). Les fonctions en code natif vont être générées et pourront être appelées avec les paramètres adaptés.

Par ailleurs du code externe défini par exemple dans un fichier source C ou C++, peut être compilé en code LLVM IR en utilisant le compilateur *clang*, et assemblé avec le module LLVM généré pour un source Faust donné. Il est ainsi possible d'ajouter le code du Work Stealing Scheduler [3] de manière très efficace, grâce en particulier à la phase de "Link Time Optimisation" qui est disponible dans l'infrastructure LLVM.

## 2.3. La chaîne de compilation complète

La chaîne de compilation complète part du code DSP source, compilé en code intermédiaire LLVM grâce au backend LLVM, et produit au final du code exécutable à l'aide d'un *compilateur à la volée* (JIT) (Figure 3).

1. Static Single Assignment : où chaque variable est assignée une seule fois

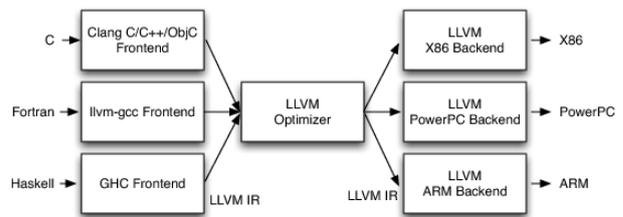


Figure 2. Schéma des composants LLVM

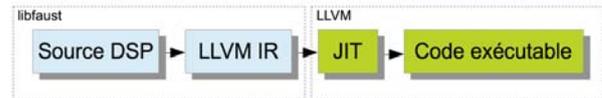


Figure 3. Chaîne de compilation complète

## 3. L'OBJET *FAUSTGEN* POUR MAX/MSP

Jusqu'à présent, l'utilisation d'objets externes Faust dans Max/MSP nécessitait de produire un objet compilé en statique avec le fichier d'architecture adapté et compilé avec un compilateur C. L'exécutable résultant devait ensuite être chargé dans l'environnement Max/MSP. Cette technique rendait difficile la modification dynamique du code source du programme Faust, puisque celui devait être recompilé et le programme Max/MSP relancé pour pouvoir utiliser la nouvelle version.

### 3.1. L'objet *faustgen*

C'est un objet externe pour Max/MSP qui embarque le compilateur Faust. S'appuyant sur *libfaust* et LLVM, *faustgen* permet de compiler et d'exécuter dynamiquement du code Faust. Celui-ci est édité dans un éditeur intégré, et l'exécutable natif est compilé lorsque l'éditeur est fermé. Il est alors inséré dans le graphe de calcul audio.

Les options de compilation classiques de Faust<sup>2</sup> peuvent être passées en paramètre. Les noms et les caractéristiques des paramètres de contrôle sont imprimés dans la console de Max/MSP, à charge ensuite à l'utilisateur de créer et connecter les éléments graphiques ( curseurs, boutons, zones numériques...) nécessaires pour contrôler le processus audio.

Chaque objet *faustgen* contient potentiellement un algorithme DSP différent. Pour partager le même algorithme, l'objet *faustgen* doit être nommé et il pourra ensuite être dupliqué pour créer autant d'instances différentes. Si le programme est modifié, toutes les instances de celui-ci seront mises à jour (Figure 4).

Le programme source au format texte ainsi que les options de compilation utilisées sont sauvegardés dans le patch. Pour optimiser les temps de re-chargement en économisant une partie des étapes de la compilation, une ver-

2. par exemple *-vec*, *-lv 1*, *-vs 32*, ou *-sch* pour le mode Work Stealing Scheduler

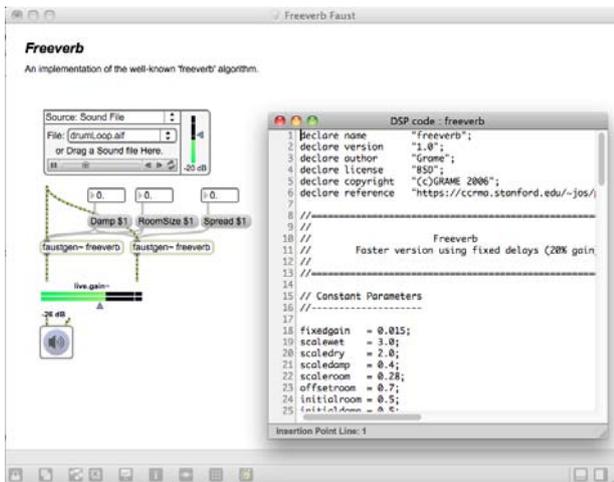


Figure 4. Une utilisation de *faustgen* pour programmer *freeverb*

sion binaire (au format LLVM bitcode) du code LLVM intermédiaire est également sauvegardée.

#### 4. FAUSTNODE

La programmation audio dans les navigateurs était jusqu'à récemment assez primitive, et essentiellement possible grâce à l'utilisation de technologies externes comme QuickTime ou Flash. L'introduction de l'élément audio dans HTML 5 a été un progrès important, en permettant la gestion de flux en "streaming". Néanmoins, pour pouvoir décrire des applications audio plus complexes, l'introduction de fonctionnalités nouvelles était nécessaire.

Web Audio est une API [5] JavaScript de haut niveau pour le traitement et la génération de son dans les applications web en HTML5. Le processus de calcul est décrit sous la forme d'un graphe de noeuds audio "natifs" (typiquement implémentés en C/C++ pour des raisons d'efficacité) et dont le rendu est ensuite effectué en temps-réel par l'implémentation sous-jacente. Des noeuds de calculs audio écrits en JavaScript (JavaScriptNode) peuvent également être insérés dans le graphe.

##### 4.1. Génération directe de code JavaScript

Nous avons expérimenté la génération directe de code JavaScript en implémentant un backend JavaScript dans *faust2*. Le code est ensuite enrobé dans un fichier d'architecture adapté qui permet de l'utiliser comme un JavaScriptNode, en l'insérant ensuite dans le graphe de noeuds audio de Web Audio.

Cette méthode produit des noeuds de calculs audio opérationnels, mais dont le code est trop lent pour en envisager une réelle utilisation dans un contexte temps-réel (avec par ailleurs les problèmes de performances liés au côté "dynamique" de JavaScript, et en particulier sa gestion mémoire avec "ramasse-miettes" qui potentiellement peut interrompre le thread audio).



Figure 5. Un exemple de *karplus* fonctionnant dans un navigateur avec une interface graphique décrite en SVG à partie de la description JSON

##### 4.2. Le noeud natif FaustNode

L'autre possibilité que nous avons exploré est la création d'un noeud C++ natif *FaustNode*. Ce noeud permet à la chaîne de compilation décrite précédemment d'être intégrée dans les navigateurs, et à des programmes Faust arbitraires d'être compilés dynamiquement et exécutés à vitesse native.

Cette intégration dans la Web Audio API a été testée en modifiant le moteur open-source WebKit<sup>3</sup>. Un nouveau noeud natif (qui s'insère dans une hiérarchie de classes C++) nommé *FaustNode* a été développé. Il reçoit en paramètre le code source Faust sous la forme d'une chaîne de caractères. Le source DSP est compilé et le code exécutable natif est produit. Il est ensuite inséré dans le graphe de calcul audio. Un système de cache est utilisé pour éviter la recompilation inutile d'un même source DSP.

La description de l'interface utilisateur peut être produite sous un format JSON (JavaScript Object Notation), un format de données textuel générique qui permet de représenter de l'information structurée, et sera ensuite analysée pour créer une interface utilisateur fonctionnelle (Figure 5).

#### 5. INTERFACE DE PROGRAMMATION

La librairie *libfaust* est distribuée avec un fichier d'entête. Pour être intégrée à une application, elle doit être ajoutée à l'édition des liens ainsi que toutes les librairies statiques LLVM. Elle propose l'interface de programmation suivante : une *fabrique* d'objets DSP peut être construite

3. WebKit : <http://www.webkit.org>

à partir d'un programme Faust source (lu à partir d'un fichier ou d'une chaîne de caractères) à l'aide de la fonction `createDSPFactory()` (par analogie, cette fabrique correspond à la définition de la classe C++ obtenue avec le mode de compilation traditionnel de Faust).

Une fois la fabrique produite, des instances de l'objet DSP vont être construites avec la fonction `createDSPInstance()` (par analogie, cette opération correspond à la création d'objets C++ avec l'opérateur `new` appelé sur la classe `mydsp` obtenue avec le mode de compilation traditionnel de Faust). Une fabrique peut également être construite en rechargeant une description du module LLVM (au format bitcode ou textuel) sauvegardée au préalable.

L'objet créé possède l'interface classique d'un processeur de signaux Faust :

```
class llvm_dsp : public dsp {
public:
    int getNumInputs ();
    int getNumOutputs ();

    void classInit(int rate);
    void instanceInit(int rate);
    void init(int rate);

    void buildUserInterface(UI* interface);
    void compute(int count,
                 float** input,
                 float** output);
```

Il peut donc être utilisé à l'intérieur d'un des nombreux fichiers d'architecture de manière similaire à un objet dont le code C++ serait produit avec la méthode classique. Enfin les objets et la fabrique sont détruits avec les fonctions `deleteDSPInstance()` et `deleteDSPFactory()`.

L'utilisation interne de la technologie LLVM reste donc cachée dans l'interface de programmation décrite, ceci pour faciliter la tâche des développeurs. Une interface plus bas niveau permettant d'accéder au module LLVM est également disponible.

## 6. CONCLUSION

Grace à *libfaust* et à la technologie LLVM, le compilateur Faust peut désormais être embarqué dans les applications et plug-ins, pour compiler et exécuter dynamiquement du code DSP. Deux exemples d'utilisation ont été présentés. D'autres développements sont en cours, visant à intégrer cette technologie dans d'autres environnements musicaux comme *SuperCollider*, *Csound* etc.

### Remerciements

Cette recherche a été menée dans le cadre du projet IN-EDIT qui est soutenu par l'Agence Nationale pour la Recherche [ANR-12-CORD-0009].

## 7. REFERENCES

- [1] Y. Orlarey and D. Fober and S. Letz "FAUST : an Efficient Functional Approach to DSP Programming", *New Computational Paradigms for Computer Music*, Editions DELATOUR FRANCE, 2009.
- [2] D. Fober and Y. Orlarey and S. Letz "FAUST Architectures Design and OSC Support", *IRCAM, (Ed.) : Proc. of the 14th Int. Conference on Digital Audio Effects (DAFx-11)*, Page(s) : 231–216, 2011
- [3] S. Letz and D. Fober and Y. Orlarey "Work Stealing Scheduler for Automatic Parallelization in Faust", *Linux Audio Conference*, 2010
- [4] Low Level Virtual Machine "Language Reference Manual", [llvm.org/docs/LangRef.html](http://llvm.org/docs/LangRef.html)
- [5] Web Audio Application Programming Interface "W3C Editor's Draft", [dvcs.w3.org/hg/audio/raw-file/tip/webaudio/specification.html](http://dvcs.w3.org/hg/audio/raw-file/tip/webaudio/specification.html)

# GASPR GRAPHICAL AUDIO SPATIALIZATION PROGRAM WITH REAL TIME INTERACTIONS

*Thierry Dilger*  
www.sonabilis.com  
thierry.dilger@gmail.com

## RÉSUMÉ

Il existe un paradigme dominant qui relie la lecture de fichiers sonores à un système de diffusion multicanal. Il est constitué d'une représentation « vue de dessus » (ou 3D) de la pièce d'écoute avec les enceintes placées virtuellement dans cet espace. Le principal inconvénient de ce paradigme est le manque d'ordre et d'harmonie de la visualisation des trajectoires donnant des systèmes extrêmement complexes. De plus, il est très difficile d'avoir une vue d'ensemble du processus de spatialisation d'une œuvre.

Le logiciel GASPR a été développé avec cette priorité pour permettre au compositeur d'avoir une représentation graphique de la distribution des sons dans l'espace et le temps. Il se base sur un détournement de technologies utilisées dans la création de jeux vidéo et possède un environnement programmable permettant la création d'œuvres comportementales complexes. Il est en outre possible d'utiliser toute sorte de capteurs pour le piloter en live, créer des compositions participatives...

GASPR s'appuie sur le codage couleur RGB (rouge, vert, bleu) travaillant sur 3 axes qui sont : le temps relatif (x), le système de diffusion (y) et l'intensité de chaque sons (z). Ce paradigme ouvre de nouvelles perspectives de créations s'affranchissant des limites des systèmes actuels.

## 1. INTRODUCTION

C'est en septembre 2011 que la création du logiciel GASPR a débutée. Elle s'est rapidement imposée comme une nécessité pour pouvoir répondre à la création d'un grand projet en multidiffusion : la Fête des lumières 2011 à Lyon (Figure 1). Ce projet est situé au cœur de la ville sur la place Gaillaton. L'espace à occuper est un grand rectangle de 40 mètres de long et 10 mètres de large sur lequel le Shadow\_Collectif [1] a décidé de proposer un travail sur le thème de la canopée. En charge de la création sonore, j'imagine très vite des nuées de singes qui se balancent de lianes en lianes d'un bout à l'autre de la place. Je dispose alors 8 enceintes regroupées par paire sur la carte de la place et je m'aperçois que je ne dispose pas de logiciel qui permet de réaliser cette spatialisation...

Les seuls outils à ma disposition travaillent uniquement sur des systèmes de diffusion circulaire (sweet spot), l'espace est ici longiligne. De plus, je ne souhaite pas travailler avec une acoustique virtuelle, préférant utiliser les caractéristiques acoustiques in situ. La création peut être appréciée aussi bien à l'intérieur des paires d'enceintes qu'à l'extérieur, en marchant ou en s'arrêtant.

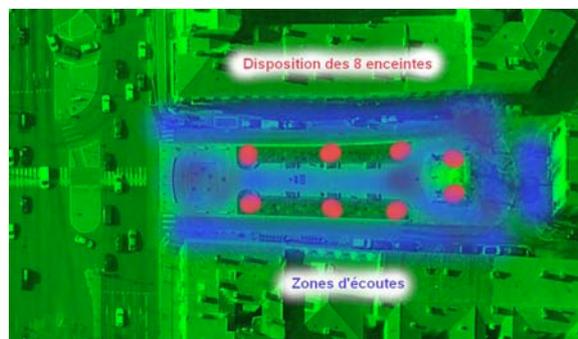


Figure 1. Projet « Canopée » 2011

Je décide d'utiliser une technique de mapping RGB, déjà expérimenté dans le projet « Sound Island » [2], pour la société Virtools, qui permet de travailler sur de grandes envergures et surtout d'obtenir un feedback immédiat de la composition sonore dans un espace virtuel (Figure 2).

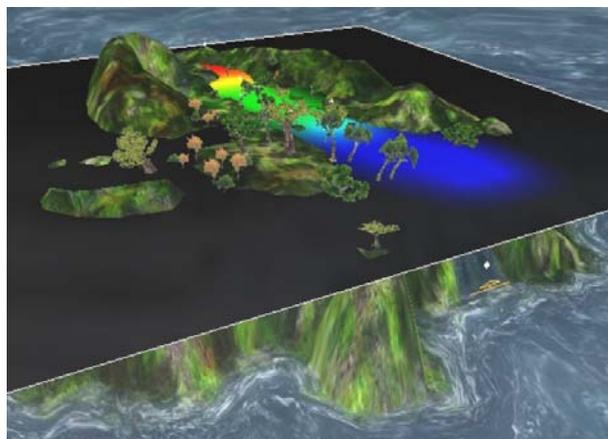


Figure 2. Mapping RGB du projet « Sound Island »

## 2. DES APPROCHES EXISTANTES

Dans le projet « Sound Island », la spatialisation est représentée de manière géographique : une tache de couleur, à un endroit donné, correspond à un son. L'idée d'utiliser un mapping RGB permet de mélanger jusqu'à 3 sons et de pouvoir extraire séparément les informations d'intensité. Un « tracking » permanent de la position du promeneur est nécessaire (équivalent aux coordonnées GPS) en liaison avec un tableau contenant les différentes zones sonores. Le tout est programmé sous Virtools et est disponible, par défaut, dans les ressources du logiciel.

D'autres logiciels travaillent aussi avec une représentation visuelle de la composition spatiale comme l'Holophon du GMEM [3]. Ici on travaille avec deux fenêtres, une représentant les trajectoires dans un espace virtuel de diffusion et une représentant le montage des sons sur une « timeline » (Figure 3).

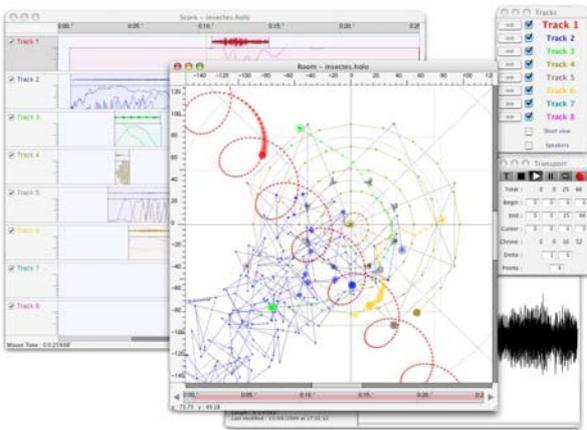


Figure 3. L'Holophon du GMEM

Le logiciel Iannix [4] permet de représenter les trajectoires des sons dans l'espace, ici dans un espace virtuel 3D (Figure 4).

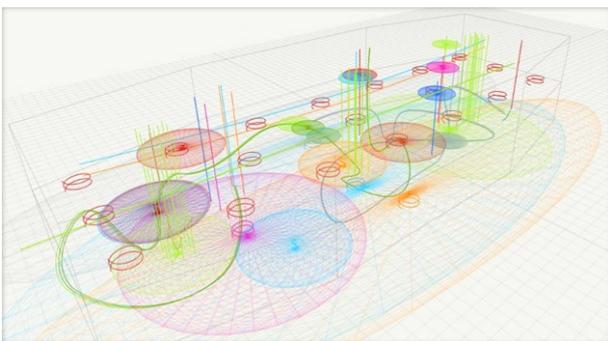


Figure 4. Iannix pour la World Expo 2012

L'avantage de ces propositions est la qualité du rendu visuel, très esthétique et très rapide à appréhender. Un autre avantage est le fait de pouvoir placer un auditeur virtuel dans la scène et ainsi de profiter de multiples types de rendu avec une même composition : stéréo, 5.1, ambisonic, WFS, binaural...

L'inconvénient principal de ces approches est de considérer la dimension graphique (ou géographique) avant la dimension temporelle. Repenser notre rapport avec une représentation de l'espace/temps, d'une structure de composition, sont les fondements de la création du logiciel GASPR.

## 3. UN NOUVEAU PARADIGME

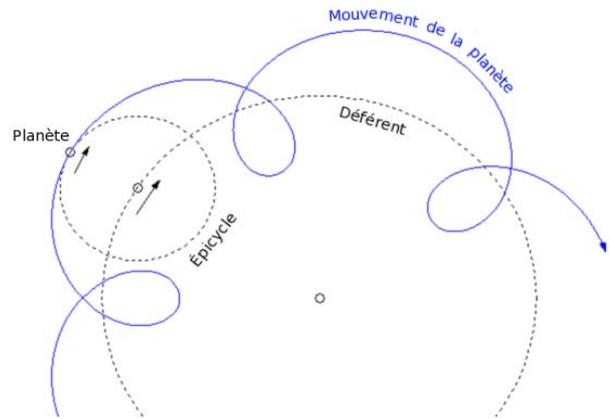


Figure 5. Trajectoire elliptique d'une planète dans un modèle géocentrique

Nicolas Copernic [5], en son temps, arrivait à expliquer le mouvement des planètes d'après un autre référentiel que la terre. Ce nouveau référentiel (le soleil) permettait, à la fois de simplifier la représentation des trajectoires et dans le même temps de faire des mesures plus précises (Figure 5).

L'idée majeure de Copernic est de repenser les rapports fondamentaux qui relient un point d'observation avec des phénomènes perceptifs en acceptant l'idée de s'écarter du modèle intuitif. Le logiciel GASPR travaille avec une représentation abstraite de l'espace. En cela il se rapproche du paradigme utilisé pour les « Spacemaps » du logiciel « CueStation » [6] qui utilise des cartes de relation entre les enceintes dissociées de la géographie du lieu (Figure 6).

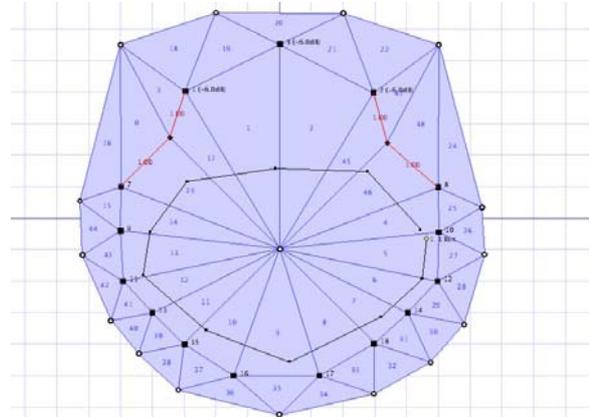


Figure 6. « CueStation » de Meyer Sound

La proposition développée s'appuie sur deux technologies issues des outils de développement des jeux vidéos : un moteur comportemental et graphique (GameMaker / DirectX) et un moteur sonore (FMOD). Il y a ici un « mapping » des informations visuelles pour commander les informations sonores. En d'autres termes, la performance du moteur graphique a une influence directe sur la qualité de restitution de la composition sonore ("A new visual paradigm for surround sound mixing" [7]). La spatialisation est traitée en panning d'intensité et permet de faire circuler un son sur un maximum de 8 sorties. Le champs d'application privilégié de GASPR étant les installations sonores (il est possible aussi de l'adapter en situation de concert / performance).

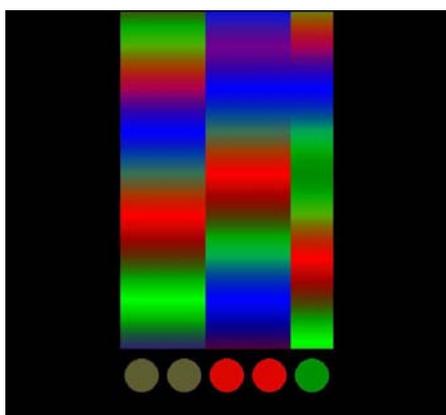


Figure 7. Un mixage spatial dans GASPR

### 3.1. Le graphisme au service du son

Créer une composition dans GASPR revient à assembler des éléments de couleurs et de formes différentes sur une « timeline » 2D (Figure 7). Les couleurs utilisées sont le rouge, le vert et le bleu (système RGB) et tous les mélanges qui en résultent. Les formes peuvent être courtes ou longues et peuvent aussi utiliser la transparence. **L'intensité de chaque couleur commande directement l'intensité d'un son.** En mélangeant plusieurs couleurs entre elles, on obtient un mélange d'intensité de plusieurs sons et donc un mixage. Tout comme pour « Sound Island », il y a un « tracking » en temps réel d'une position sur cette carte qui extrait les informations de couleurs. Le logiciel étant prévu pour 8 sorties, il y a 1 « tracking RGB » par enceinte (Figure 8).

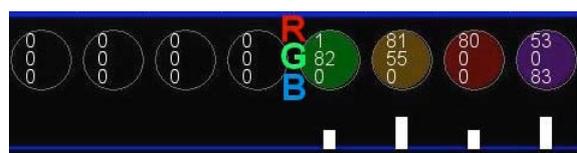


Figure 8. « Tracking RGB » pour 8 enceintes, ici les enceintes 5, 6, 7 et 8 reçoivent du son

Un travail proche (dans sa technique), est celui de Memo Akten avec « Simple Harmonic Motion study #2a » [8]. Une structure géométrique, en 3 dimensions, se déplace autour d'une ligne médiane. Quand les noeuds de cette structure rencontrent cette ligne alors un son est produit. C'est bien le moteur graphique qui dirige la création sonore. Je cite aussi l'importance du modèle physique qui permet la transformation fluide de la structure, ce qui a un impact direct sur le rythme de la création engendrée (Figure 9). La représentation visuelle dans GASPR est aussi dynamique : Les blocs de couleurs se déplacent, s'allongent, se mélangent, disparaissent... au rythme du « framerate » de l'application.

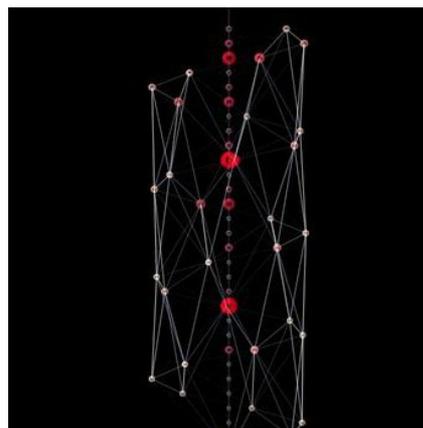


Figure 9. Simple Harmonic Motion study #2a

### 3.2. La liberté de placement des enceintes

C'est l'axe vertical de la « timeline » qui nous renseigne sur les différentes enceintes utilisées. La dernière version de GASPR fonctionne avec 8 enceintes. Ici le positionnement des enceintes est libre car il n'est pas attaché à une représentation intuitive de l'espace. Chaque enceinte possède « sa piste » que l'on remplit avec les sons. Cela a l'inconvénient de figurer la disposition de manière abstraite mais a le gros avantage de pouvoir s'adapter à de nombreuses situations de diffusion comme par exemple : 8 enceintes en ligne, 5 enceintes dans une pièce et 3 dans une autre, utilisation de la hauteur, d'enceintes aux caractéristiques différentes...

Dans l'illustration suivante (Figure 10), un son de métro est assigné aux blocs verts reliés aux enceintes 1,3,5,4,2 (de gauche à droite et de haut en bas). Ils se superposent dans le temps (axe horizontal) pour créer une transition douce lors du passage d'une enceinte à une autre.



Figure 10. Distribution d'un son dans les enceintes

Voici ici le même mouvement représenté dans l'espace de diffusion (Figure 11).



Figure 11. Mouvement d'un objet sonore

Le schéma suivant (Figure 12) résume ce paradigme visuel en prenant comme exemple le système de diffusion 5.1 (le caisson de grave n'est pas utilisé, il n'y a pas de correspondances avec le schéma précédent). La barre de lecture se déplaçant de gauche à droite :

- Un bloc bleu dans l'enceinte 3 (centrale), qui joue un son uniquement dans cette enceinte.
- Le même bloc bleu dans les enceintes 2,3 et 4. Ce même son est joué dans les enceintes frontales.
- Le même bloc bleu (son) sur les 5 enceintes.
- Un bloc bleu sur l'enceinte 4 et un bloc vert sur l'enceinte 2. Cela correspond à la lecture d'un fichier stéréo (2 canaux).
- Le même fichier stéréo plus le mélange des 2 canaux dans l'enceinte centrale.
- Un bloc bleu en escalier correspond à un son qui saute d'une enceinte à l'autre en tournant.
- Un bloc bleu et vert en escalier avec transparence correspond au mouvement fluide d'un son stéréo qui tourne autour de nous.

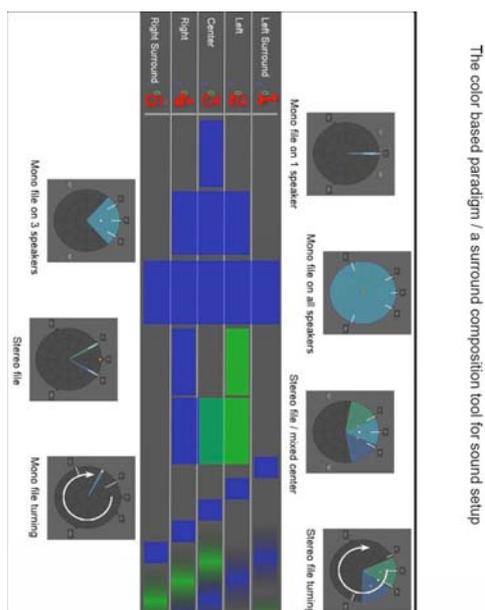


Figure 12. Paradigme visuel de GASPR

### 3.3. La gestion du temps absolu et relatif

C'est à l'intérieur de la barre de lecture que se trouve les 8 « tracking RGB » temps réel. C'est l'axe horizontal de la « timeline » qui représente la composante temporelle de la création. Ici le temps peut être absolu : la lecture s'effectue d'un point A vers un point B, à vitesse constante et sans s'arrêter. Il peut être aussi relatif car la taille d'un bloc de couleur ne correspond pas à la durée d'un son. En effet, **les sons sont tous joués en boucle** qu'ils soient sur la barre de lecture ou pas. **La barre de lecture commande uniquement le volume de chaque son** (extraction des valeurs RGB). Un même son peut très bien être utilisé deux fois dans une création : une fois avec un bloc court (petite apparition) et une fois avec un bloc long (grande apparition). Ainsi même si la lecture s'arrête au milieu d'un bloc, le son continue d'être entendu (avec l'intensité liée à l'opacité du bloc). On peut par conséquent, avec une même composition faire entendre des versions courtes et des versions très longues.

### 3.4. La création comportementale

GASPR est un environnement dynamique, chaque bloc de son possède une vie propre. Ils sont considérés comme des « sprites » de jeu vidéo. Ainsi ils peuvent recevoir des « scripts » de comportement comme la génération aléatoire, le « pathfinding », l'intelligence artificielle... Ils peuvent naître, mourir, se rencontrer, se détester... Ils se déplacent au gré de leur comportement, programmé, au préalable, par le compositeur. Un aspect important dans ce type de création, est de la laisser vivre un certain temps (lancer la lecture avec les comportements) pour pouvoir l'apprécier et éventuellement opérer des modifications. Les comportements peuvent faire appels à des modèles aléatoires, combinatoires, physiques... A l'instar de la relation entre spatialisation et gravité dans le logiciel « Spatium » [9] (Figure 13).

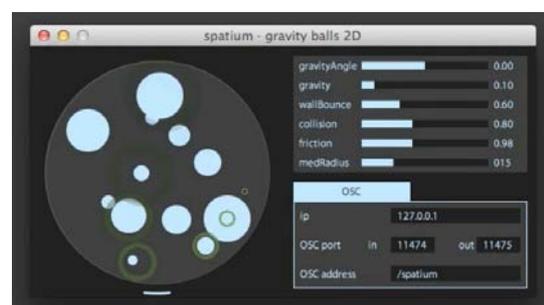


Figure 13. Éditeur de gravité dans « Spatium »

L'écoute d'une œuvre avec GASPR est comme une promenade en forêt : on peut choisir son point d'observation pour essayer de surprendre une création, on peut revenir sur ses pas et croiser quelque chose qui n'était pas présent la première fois...

## 4. FONCTIONNEMENT TECHNIQUE

GASPR est un logiciel de composition spatial comportemental « stand alone » pour Windows XP, 7 ou 8 (32 bits et 64 bits). Le nombre de blocs que l'on importe dans GASPR n'est pas limité et le moteur sonore (FMOD) peut facilement gérer 200 blocs en mixage temps réel avec un petit processeur (1 GHz). Le poids de l'application est inférieur à 4 Mo (sans les sons). Les sons utilisés sont des fichiers WAV mono jusqu'en 24/96 non compressés (en fonction de la carte son). FMOD s'appuyant sur la couche DirectX de Windows, toutes les cartes sons 8 sorties reconnues par Windows fonctionnent avec GASPR dans la configuration « surround 7.1 ». Tous les contrôleurs de jeux sont reconnus nativement (pour un contrôle temps réel). L'export d'une composition est un exécutable (.exe) avec un dossier contenant les sources. A noter qu'il est possible d'encrypter les sources pour les protéger (cryptage FMOD utilisé dans les jeux vidéos).

### 4.1. SoundObject

Ce sont les briques élémentaires de toutes compositions dans GASPR. Il n'est possible de mélanger que trois couleurs en même temps, donc de ne mixer que 3 sons en simultané. Cette limite a été étendue à 9 sons en utilisant simultanément 3 « timeline » synchrones. Il ne faut pas oublier que l'on raisonne ici par enceinte : 9 sons par enceinte, à un temps donné, ce qui correspond à 72 sons en mixage dynamique sur 8 enceintes. Il est aussi possible, si plusieurs sons partagent le même comportement (ou trajectoire spatiale) de les regrouper au sein d'un même bloc de couleur. Les trois couleurs possèdent un système de hiérarchie, ainsi les « SoundObject » rouge sont toujours au dessus des verts qui sont au dessus des bleus. Cela signifie que si 3 blocs de couleurs différentes et d'intensités maximales se rencontrent, le « son bleu » a un coefficient de 1, le vert de 2 et le rouge de 3. Le codage d'intensité étant réalisé sur 8 bit (256 niveaux). A l'écoute ce mélange de couleur produit un fondu doux d'un son en fonction de la présence d'un autre son : c'est un effet de « ducking » automatique.

### 4.2. SoundStructure

Une « SoundStructure » est un ensemble de « SoundObject » qui partage un même comportement. Dans un contexte social, un « SoundObject » est un individu et une « SoundStructure » une population (Figure 14). Par exemple un groupe de 4 blocs distribués sur 4 enceintes ayant un mouvement d'aller retour correspond à une « SoundStructure ». Cela peut être relié à des sons de vent qui apparaîtront et disparaîtront de manière progressive dans les enceintes. Les « SoundStructure » peuvent aussi être utiles pour lire des fichiers multicanaux (jusqu'à 8 canaux) ou réaliser des « prémix » de certaines matières.

### 4.3. MixMap

La « MixMap » représente la distribution de tous les sons dans les 8 sorties. Elles sont représentées simultanément sous 2 angles. Dans l'illustration suivante (Figure 14), la « timeline » 1 se situe dans l'espace en haut à gauche de l'interface (représentation verticale, gros plan de la distribution des sons). Cette même « timeline » se retrouve sur une bande plus bas, sur toute la largeur de l'interface : c'est la vue globale. De la même manière, la barre de lecture apparaît deux fois : une grande ligne horizontale pour la vue gros plan, une petite ligne vertical pour la vue générale. Il y a actuellement 3 « MixMap » synchrones dans GASPR donc 3 « timeline ». Il est possible de faire la correspondance entre une « MixMap » et une partition musicale. Dans ce cas on ne lit pas des notes qui changent de hauteur mais des sons qui changent de position spatiale. Les « timeline » possèdent des « clefs » (comme la clef de sol) qui ne sont pas représentées mais induites et reliées à la position spécifique des enceintes dans l'installation. Il est possible de décoder la partition uniquement si on possède la clef. Étant donné qu'il y a 3 « timeline », il est possible d'utiliser plusieurs clefs simultanément : par exemple une disposition en diagonale des blocs sur une « timeline » peut donner un mouvement circulaire à l'écoute, sur une autre « timeline » la même diagonale peut donner, par exemple, un mouvement de saut entre enceintes opposées.

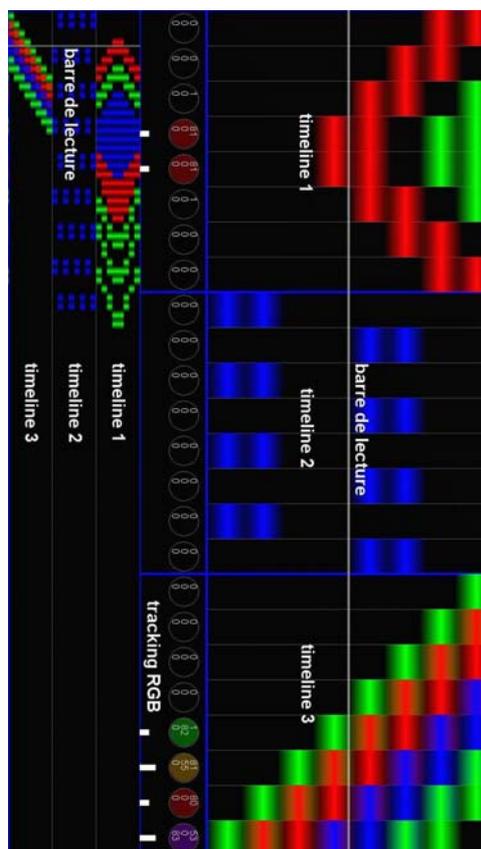


Figure 14. Fenêtre générale du logiciel GASPR

## 5. ECRITURE SPATIALE INTERACTIVE

L'environnement d'écriture dans GASPR repose sur le paradigme choisi. Il s'agit de faire une représentation de la distribution des sons dans les enceintes en liaison avec l'intensité de couleur de chaque bloc. C'est avec la « MixMap » que le compositeur écrit les rapports son / espace / temps / interaction. Il y a aujourd'hui un éditeur intégré dans GASPR pour pouvoir placer ses blocs, les dupliquer, les allonger, changer leurs opacités... En lecture, comme les « SoundObject » sont dynamiques, on peut anticiper l'arrivée de tel ou tel son dans la composition, vérifier et corriger les mouvements. Avec un peu d'exercice, il est possible de lire une « MixMap » comme une partition et de « chanter les positions »...

### 5.1. Temps déterminé

C'est la situation d'écriture la plus simple où le compositeur connaît à l'avance la durée de la création. Il travaille pour un format précis par exemple une pièce de 3 minutes en 5.1. Ici la barre de lecture se déplace continuellement et rencontre les « SoundStructure » qui peuplent cette création. Ces « SoundStructure » peuvent être figées (œuvre linéaire) ou bien dynamique. Dans ce dernier cas, chaque lecture de la création donnera lieu à une expérience d'écoute différente. Le temps de la création peut être aussi relié à une horloge absolue (l'heure réelle) pour opérer des mouvements. Ainsi dans le projet « Canopée » [10] pour la Fête des lumières, un impératif était que la création soit plus silencieuse à partir de 19 heures pour ne pas déranger les riverains (changement dynamique de l'opacité des blocs à 19 heures).

### 5.2. Temps indéterminé

L'écriture dans un temps indéterminé est la plus contraignante pour le créateur car il doit créer suffisamment de « SoundStructure » et de comportements pour permettre de proposer à l'auditeur une expérience riche. De plus le compositeur ne peut pas entendre la totalité de sa création, seulement des occurrences qui reflètent son travail. Ici l'œuvre n'a pas de durée définie et il est possible de l'apprécier par écoutes successives. On peut comparer ce type de création à un paysage sonore réel qui ne possède ni début ni fin. L'idéal est d'utiliser des fichiers sonores très longs car même si une « SoundStructure » revient, il n'y aura que peu de chances pour retomber au même endroit dans le fichier son (pour rappel tous les fichiers sont lus en boucle en permanence). C'est aussi le terrain pour créer des compositions émergentes issues de règles de comportement qui se combinent dans le temps jusqu'à fabriquer un objet que le compositeur même ne peut pas anticiper. C'est le cas de l'installation " Mutasound Paris #1 " [11] qui retrace la vie sonore délirante d'une place parisienne sur 8 enceintes.

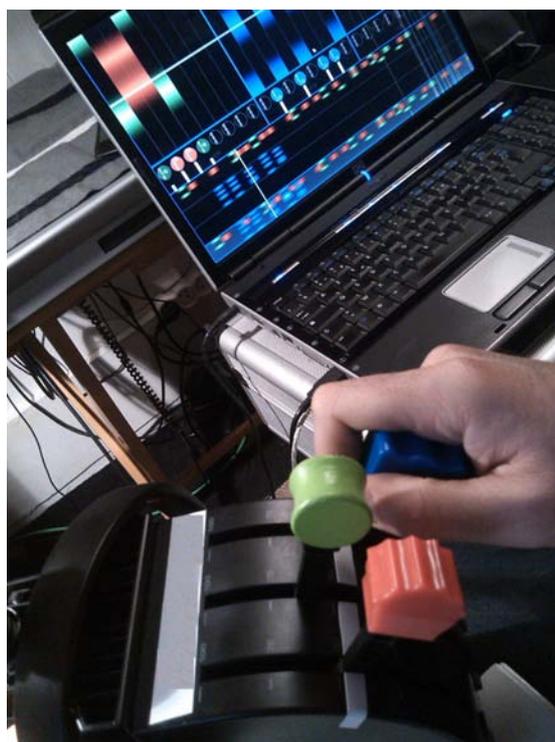


Figure 15. Contrôle direct du logiciel GASPR

### 5.3. Contrôle direct

Le contrôle direct de la création est typiquement relié aux performances, concerts... En pilotant GASPR avec des contrôleurs (capteurs ARDUINO, contrôleurs jeu vidéo, Kinect...), il devient possible d'avoir un rapport instrumental (Figure 15). On lance des objets sonores avec des trajectoires déjà définies, on modifie ces trajectoires en temps réel, on change les volumes (opacité dynamique), on change la vitesse de déplacement de certains objets, on influe sur leur comportement... En utilisant plusieurs capteurs simultanés, on peut faire réagir l'œuvre de manière plus vivante et créer l'interaction avec un groupe de personnes, une situation adaptée à une installation artistique.

### 5.4. Contrôle indirect

Le but d'un contrôle indirect sur l'œuvre est de ne pas casser radicalement l'équilibre sonore qui s'est installé dans le cas où certains objets sonores possèdent une vie autonome. Le compositeur peut réguler ces mouvements grâce notamment à la hiérarchie qui existe entre les couleurs. Cette hiérarchie est très utile dans des œuvres complexes, et permet de faire émerger des sons au milieu d'une grande activité comportementale (les sons rouge). Il est possible aussi de déclencher une « SoundStructure » spéciale, qui n'a pour but que d'influer sur le comportement d'autres « SoundStructure ». On est ici dans un rapport de diplomatie et pas un rapport d'ordre.

## 6. VERS UN NOUVEAU FORMAT

Repenser un paradigme de création sonore dans l'espace permet de répondre à des besoins très spécifiques et dans le même temps ouvre la porte vers de nouveaux horizons. La composition spatiale interactive est un espace d'expression immense pour les artistes. Malheureusement, ils sont confrontés à des difficultés de diffusion dans lequel GASPR peut apporter une proposition. Créer une œuvre dans ce format est en fait un portage de sons déjà travaillés (voir prémixés) auquel on ajoute des comportements spatiaux. On peut imaginer de faire une version GASPR d'œuvres sonores déjà existantes (avec la limite des 8 enceintes) et même des versions inédites qui intègrent la dimension comportementale (sur des œuvres qui n'en avaient pas). La possibilité d'interagir avec l'œuvre est aussi un élément intéressant pour le public. Aujourd'hui, il existe une version PC avec un export .exe de la création plus un répertoire avec les sons (cryptés). Il suffit d'avoir un système d'écoute multicanal (par exemple 5.1) configuré dans Windows pour en profiter (fonctionne sans installation sous XP, 7 et 8 en 32 et 64bits, processeur simple cœur de 1GHz minimum). Le logiciel contrôle le « framerate » (FPS) et adapte automatiquement sa vitesse de travail en fonction, ce qui garantit une lecture identique sur des ordinateurs de puissance différente. Une option en test actuellement, est de laisser le droit au public de pouvoir remplacer les sons pour faire sa propre création. Cependant il ne peut pas changer les règles d'interaction et la composition d'une « MixMap ». Le compositeur reste présent dans la mise en scène spatiale et comportementale. Les auditeurs de demain seront eux aussi, en quelque sorte, des créateurs sonores.

Pour les futurs versions de GASPR, des ouvertures sont envisagées comme une option d'écoute en binaural (pour les créations 5.1), la compatibilité avec les interfaces tactiles ou encore l'intégration du protocole TCP/IP pour synchroniser plusieurs GASPR, dépasser les 8 sorties et créer des œuvres synchrones sur plusieurs sites...

## 7. RÉFÉRENCES

- [1] Shadow\_Collectif, Collectif parisien, créé en octobre 2010, regroupe 4 artistes plasticiens, créateurs visuels, créateurs vidéos, créateurs sonores. <http://shadowcollectif.wordpress.com/>
- [2] Projet « Sound Island », Thierry Dilger, 2002. [http://www.virttools.co.kr/dc/minisite/MiniSite/html/e03\\_DEMO\\_view\\_SoundIsland.htm](http://www.virttools.co.kr/dc/minisite/MiniSite/html/e03_DEMO_view_SoundIsland.htm)
- [3] Logiciel Holophon du GMEM, outils de manipulation temps réel de trajectoires sonores. [http://dvlpt.gmem.free.fr/web/static.php?page=Holophon\\_main](http://dvlpt.gmem.free.fr/web/static.php?page=Holophon_main)
- [4] Iannix : séquenceur graphique open-source. <http://www.iannix.org/fr/>
- [5] Nicolas Copernic, astronome, père de l'héliocentrisme. [http://fr.wikipedia.org/wiki/Nicolas\\_Copernic](http://fr.wikipedia.org/wiki/Nicolas_Copernic)
- [6] Matrix3's CueStation, logiciel de spatialisation pour spectacle live. <http://www.meyersound.com/products/matrix3/>
- [7] Thierry Dilger, « A new visual paradigm for surround sound mixing » poster for the ICSA first international convention on spatial audio (Detmold, Germany november 2011) and engineering brief presentation for the 132 AES convention (Budapest, Hungary, april 2012).
- [8] Memo Atken, créateur graphique. <http://www.msavisuals.com/>
- [9] Spatium, outil de spatialisation sonore. <http://spatium.ruipenha.pt/>
- [10] Thierry Dilger, projet « Canopée » pour le Fête des lumières, Lyon, 2011 <http://www.sonabilis.com/canopee.html>
- [11] Thierry Dilger, « Mutasound Paris #1 » presentation de l'installation pour The Global Composition, World Conference on Soundscape and Acoustic Ecology (Dieburg, Germany july 2012).



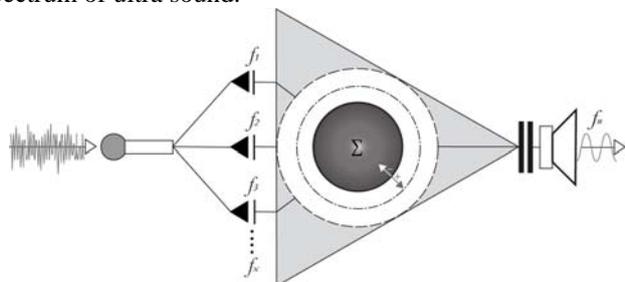
# SONAPTICON – SPACE AS AN ACOUSTIC NEURONAL NETWORK

*Tim Otto Roth*

Imachination Projects,  
Oppenau/ Cologne, Germany  
tor@imachination.net

## RÉSUMÉ

The Sonaption connects acoustics and neuroscience with space changing the Sound Dome with its 43 loudspeakers at the Center for Arts and Media ZKM Karlsruhe literally into a musical mindstorm. [1] The core of the system are the so called audio neurons communicating not with electric currents but with sound waves. This unique translation of neuronal systems' electrical communication into acoustic communication is combined with the deceleration of biological neuronal processes that take place on the timescale of milliseconds. In this way, in the Sonaption neuronal communication mechanisms become both a sensual and a comprehensible experience. At the concert like Premiere in the context of the ZKM's IMATRONIC festival and the symposium Neuroaesthetics [2] in November 2012 three solo piccolo players revealed by a simple combinatoric play [3] the different facets of the Sonaption, to be it a microtonal sound wash or a pseudo repetitive pattern in the inaudible sound spectrum of ultra sound.



**Figure 1.** Scheme of an audio neurone.

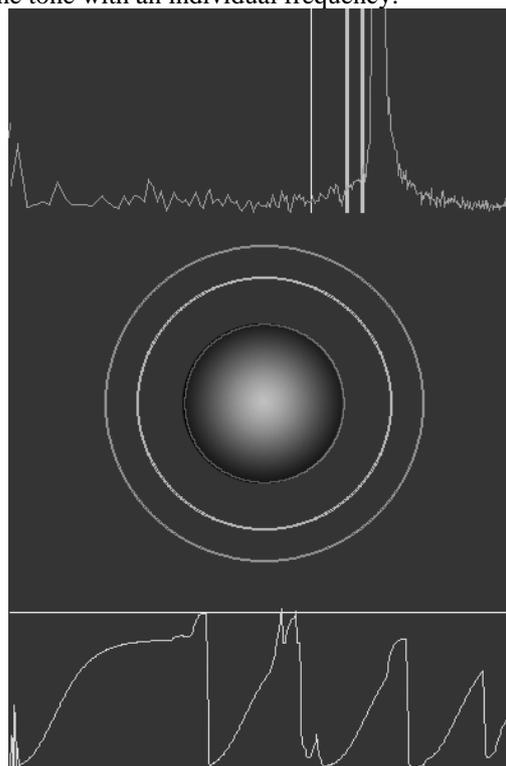
## 1. INTRODUCTION – ACOUSTIC FEEDBACK

Biological neurons are interconnected by organic wires and communicate by short electric impulses, so-called 'spikes'. In fig.1 you see a simple scheme of how a neuron works: The incoming impulses are summed up and change the membrane potential of a neuron. There are two different types of incoming spikes: if the spike increases the membrane potential, the sending neuron is called excitatory, if it decreases the potential, the sender is called inhibitory.

In the absence of external input, the membrane potential decays to a resting potential which is marked by the inner dashed ring. A neuron sends out an electric impulse, it 'fires', when the incoming impulses from

other neurons force the potential over a certain threshold, here indicated by the outer dashed ring. After having fired the neuron is not excitable for a certain period of time and the potential goes back to the resting potential.

Most people do know feedback in acoustic systems as a cascading effect of a microphone in combination with a loudspeaker. The Sonaption uses a similar but different scheme. The basic units of the Sonaption are audio neurons interconnected not by wires but by sound transmitted in space. An audio neuron registers impulses (*spikes*) from its acoustic environment by a microphone and fires its own impulse by a loudspeaker (fig.1). The clue of that system is that every neuron gets assigned a sine tone with an individual frequency.



**Figure 2:** Visualization of the Python based audio-neuron showing the frequency tracking (top), the current membrane potential (circle in the middle) and the history plot of the membrane potential (bottom).

Looking to a sound spectrum you see that most sounds do consist of a broad band of frequencies. But sine tones oscillating at a single frequency show up

characteristic peaks in a spectrum (see top of fig. 2). This is how you can simply connect the audio neurons: Every audio neuron registers a specific set of frequencies assigned to other neurons. These frequencies are marked as lines in two colours in the spectrum. If an audio neuron registers a significant steep peak at one of these lines it interprets this as impulse of a connected neuron and the line flashes. Depending on the colour assigned to a line an impulse has an excitatory or inhibitory effect.

## 2. DECELERATION OF BIOLOGICAL PROCESSES

Unlike in *in vivo* neural tissue, where neuronal interactions occur on the timescale of milliseconds, the Sonapticon allows to decelerate all biological processes by the factor 10-200. This generates an environment where spectators can identify and understand intuitively the causal relationship between integration and initiation of spikes on a single neuron level. Thereby, the Sonapticon creates an unprecedented sensual experience of the basic computational principles that underlie all mental processes, be it the initiation of movement, the creation of simple thoughts, or even the emergence of consciousness.

This simple interaction scheme is the basis of all nervous activity. It is still an enigma how these interactions form something like a simple thought in our brain. With latest microscopic techniques you can observe *in vivo* in small probes of neuronal tissues how firing neurons self-organize and create synchronized activity patterns clip. Unlike in *in vivo* neural tissue, where neuronal interactions occur on the timescale of milliseconds, the *in silico* technique of the Sonapticon allows to decelerate all biological processes. This generates an environment where spectators can identify and understand intuitively the causal relationship between integration and initiation of spikes on a single neuron level.

Consequently the Sonapticon is a system of decelerated synthetic neurons which allows a real time interaction with neuronal dynamics and control of the biological parameters. The Sonapticon combines two basic components:

- digital *in silico* methods, with latest mathematical models of biological neurons (e. g. the conductivity based model by Alain Destexhe) [4]

and

- an empirical environment (the Sound Dome) using acoustics as an analogue space of experimentation and interaction. Here delay effects come into play due to the propagation speed of sound resulting in dynamics which differ essentially from cable wired systems.

Principally every computing device with a microphone and loudspeaker can function as an audio neuron to be it a laptop, a tablet computer or a smart phone. In August 2012 a first performance did take

place at the Bernstein Centre for Computational Neuroscience on the Campus of the Charité in Berlin. The visitors installed a little python based software (see screenshot fig. 2) to be it for PC or Mac and changed their laptops with small external speakers into acoustic neurons. Step by step a network of twenty acoustic neurons was built up, exploring the changing dynamics by adding further neurons.

Above all the Sonapticon invites to interact with the system in all different forms. An instrument like the singing saw producing clear sine peaks resulted to be a perfect device to explore the system's resonances.

## 3. THE SONAPTICON AT THE ZKM'S SOUND DOME

Thanks to the invitation by the Institute of Music and Acoustics at the Center for Art and Media, ZKM Karlsruhe we – the author and the neuro-mathematician Benjamin Staude – could explore together the Sonapticon as guest composers from 2010-12 in a very sophisticated environment. In collaboration with the sound engineer Holger Stenschke from ZKM we were changing the Sound Dome into an acoustic neuronal network. The Sound Dome consists of 43 loudspeakers building an impressing hemisphere of 15 meters in diameter. The loudspeakers get joined via microphones. Intuitive understanding of the acoustic feedback through sine tones in the Sound Dome space is complemented visually through LED lights mounted on the speakers flashing up in a blue or yellow colour in case the neuron is firing. Above all a video projection on the floor shows a visualization of all the active neurons with their changing membrane potentials.



**Figure 3:** Impression from the premiere with three piccolo players standing around the floor projection and flashing loudspeakers in the background.

In the ZKM's Sound Dome the audio analysis uses a frame work based on MaxMsp programmed by Holger Stenschke. The adaptation of the neuron model and the composition frame work was realized in Python by Benjamin Staude. For the visualization Tim Otto Roth used Gem in combination with Puredata.

#### 4. SPACE

First of all the Sonapticon creates in the Sound Dome a completely new experience emerging the visitor into a surrounding acoustic atmosphere: You feel to stand in the middle of an organic process of self organizing tones.

Above all the use of sine tones assigned to specific loudspeaker in the space creates a particular effect. If you are walking around in the installation the sine tone patterns recombine differently in the visitors ears depending on the locally dominant tones. If the system runs very fast you even get the effect of standing waves changing even if you slightly turn the head.



**Figure 3:** Impression from the audio visual spatial setting in the Sound Dome (Klangdom) at the ZKM Karlsruhe.

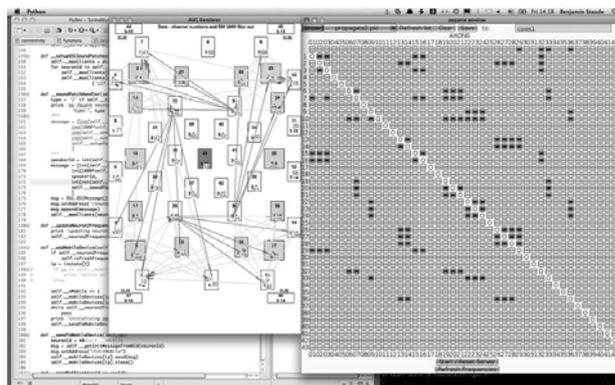
Finally this controlled environment allows to explore systematically the characteristics of the Sonapticon. For instance with regards to the size of the Sound Dome and the speed how sound waves propagate in space delay effects come into play which appear also in neuronal tissues: So here it makes a difference where and when a sound is played in the environment – an invitation for musicians to play with the environment.

#### 5. A NEW INSTRUMENT

Thus, we transformed with the Sonapticon the Sound Dome into a new instrument for which we have developed a special method of composing in and with the Sound Dome space: choosing scales up to the eighth tone range, giving tones a location by assigning frequencies to an audio neuron, setting the connections between the audio neurons and finally varying the tempi by playing with the system's update rate.

The piece at the concert premiere comprised eight movements starting with a single neuron. Step by step the number of neurons the three piccolo players interacted with was increased. Above all the composition used different tonalities as a Western chromatic, an Indian Shruti or a microtonal eighth tone scale. In the last movements the frequencies were

transposed subsequently by perfect fifths ending up in a finale with inaudible ultra sound.



**Figure 4.** Composition matrix with connection map at the right and the resulting network at the left.

Finally the models themselves are a playfield for composition. Parameters can be changed globally, but also the characteristics of individual neurons can be varied. Modifying the ratio of neurons and interneurons has a significant effect. Until now this was studied only rudimentary, so it is a future task for the Sonapticon to explore also with different models acoustically the phenomenon of synaptic plasticity.

#### 6. REFERENCES

- [1] An extended documentation of the project including two video introductions is provided by the project page: <http://www.pixelsex.org/sonapticon>.
- [2] Record of Tim Otto Roth's and Benjamin Staude's talk on the Neuroaesthetics symposium at ZKM Karlsruhe: <https://itunes.apple.com/de/itunes-u/neuroaesthetics-symposium/id601309561?mt=10>.
- [3] Only one tone is played by each player in a certain sequence, so together the three soloists play all possible tone combinations. This time based combinatory play is inspired by Tom Johnson's counting pieces. See: <http://www.editions75.com>.
- [4] Destexhe, Alain: Self-sustained asynchronous irregular states and Up-Down states in thalamic, cortical and thalamocortical networks of nonlinear integrate-and-fire neurons, in: Journal of Computational Neuroscience 27 (Dec 2009) Nr. 3, p. 493-506. Destexhe's model has been adapted by adding a refractory period.
- [5] Barry, Robert: Artist collaborates with neuroscientist to build 'audio-neurons', interview on [wired.co.uk](http://www.wired.co.uk) (10 December 2012), <http://www.wired.co.uk/news/archive/2012-12/10/tim-otto-roth>.



# SONS SOLIDIENS ET ESPACES AURAUX COMPOSITES

*Otso Lähdeoja*  
Matralab  
Concordia University  
Otso.lahdeoja@free.fr

## RÉSUMÉ

Cet article se propose comme une exploration des possibilités offertes par le son solidien pour la création musicale et intermédiatique. La technologie du son solidien étant basée sur le haut-parleur électrodynamique, elle n'est pas en soi porteuse de nouveauté. L'intérêt que nous souhaitons exposer réside plutôt dans l'appropriation de cette technologie pour la création musicale. Dans ce domaine le son solidien a encore été très peu utilisé, et nous y percevons un potentiel considérable. Tout en tenant compte des contraintes techniques et acoustiques du son solidien, il présente la possibilité de rendre des objets ou un espace acoustiquement actifs et ainsi ramener l'espace physique dans le domaine de l'espace composable. L'utilisation de multiples sources sonores – tels qu'instruments acoustiques, haut-parleurs et son solidien – nous paraît constituer une stratégie fertile, donnant lieu à des travaux artistiques expérimentaux.

## 1. SONS SOLIDIENS ET SURFACES VIBRANTES

Cette recherche prend comme point de départ l'utilisation du son solidien comme moyen original pour créer des espaces sonores composites. Par son solidien on entend ici l'induction de vibrations dans un objet ou une surface solide. La surface vibrante transfère son énergie dans l'air, résultant dans du son aérien coloré par les caractéristiques acoustiques de l'objet émetteur, d'une manière analogue au haut-parleur. Aussi, en entrant en contact physique avec la surface vibrante, on acquiert une perception haptique – audiotactile – du son. Dans notre optique, la notion du son solidien est étroitement associée à celle d'un objet vibrant et produisant du son aérien.

Tout solide peut entrer en vibration et irradier du son aérien, bien que l'énergie nécessaire pour l'induction du mouvement dépend de l'inertie et de l'impédance acoustique du matériau. Cette polyvalence du son solidien nous semble fertile pour la création artistique ; un large éventail d'objets et de structures peut être transformé en source sonore, produisant des effets acoustiques et perceptifs variés. On peut penser à des petits objets quotidiens vibrants – par exemple des luminaires (utilisés dans les performances du groupe « luminosonique » de Concordia University Matralab [11]), au mobilier sonore, jusqu'à des grandes surfaces telles que les fenêtres ou des murs entiers. A chaque

échelle et matériau, la perception sonore sera différente. Il s'agit d'une « scénographie » ou une « architecture » sonores où un ensemble d'objets ou de structures devient un « espace instrumental » [9]. Il serait tentant d'utiliser le terme « objet sonore » pour désigner ces structures vibrantes, mais la confusion avec l'héritage Schaefferien du terme est trop présente. Ici nous proposons d'employer le terme « objets ou structures acoustiquement actives ».

L'induction du son dans des solides est obtenue par des transducteurs acoustiques – ou « vibreurs » - qui sont essentiellement des haut-parleurs où la membrane a été remplacée par une masse. L'inertie de la masse mouvante transfère son énergie au solide et le fait vibrer. Il existe sur le marché un large panel de transducteurs aux caractéristiques et tailles variées. Les plus petits sont de la taille d'une pièce de monnaie avec des puissances d'ordre de quelques Watts, les plus puissants sont capables de faire entrer en vibration un mur en brique.

En soi ces transducteurs acoustiques n'ont rien de nouveau – le haut-parleur électrodynamique est bien connu depuis longtemps. Toutefois, cette technologie continue à être rarement utilisée en musique, bien que dans le domaine des arts plastiques sonores leur utilisation se répand actuellement [10] applications ont vu le jour dans le cinéma et le home-cinéma, ainsi que dans des recherches de sensation auditive pour malentendants par la perception audiotactile. Notre recherche est animée par l'idée que le son solidien peut apporter une dimension fraîche et intéressante dans la pratique des musiques électro-acoustiques et mixtes.

### 1.1. Transducteurs de sons solidiens - défi de la qualité sonore

Le son aérien irradiant d'un objet vibrant est modelé par les caractéristiques acoustiques de celui-ci. Le signal irradié est coloré par la réponse fréquentielle du support et ses modes de résonance. Aussi, la chaîne électroacoustique comporte le transducteur, et - pour une part à importance moindre ici – l'amplificateur. Déjà, les transducteurs en soi ne sont pas des appareils de haute-fidélité, mais en y ajoutant le régime vibratoire complexe d'une surface solide, les sons résultant est une version lourdement filtrée de l'original, sans commune mesure avec les systèmes de haut-parleurs perfectionnés. Ce problème de qualité sonore constitue le plus grand frein

pour une utilisation plus répandue des sons solidiens en musique.

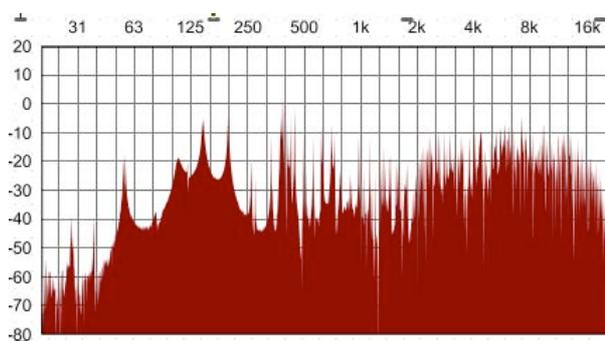
A l'Université de Mons, nous avons conduit des essais préliminaires pour comparer les caractéristiques acoustiques de différents transducteurs sur des plaques de verre. Ce matériau a été choisi pour son universalité, sa réponse plutôt satisfaisante au son solidien, et sa maniabilité. Une méthode de réponse impulsionnelle (RI) par balayage exponentiel de sinusoides a été utilisée. Nos tests visaient à mettre en évidence des différences entre transducteurs afin de comparer des produits sur le marché. Les résultats peuvent être consultés sur le site de l'institut Numediart [8]. D'une manière générale, les tests ont mis en évidence le type de filtrage qu'induit le couplage surface solide – transducteur au signal audio. La figure 1. montre un courbe de RI typique de plaque de verre, caractérisé par un filtrage extrême, marqué par les modes de résonance du solide. Le profil de la RI est propre à chaque solide et à chaque emplacement de transducteur sur celui-ci, mais l'ordre de complexité de la courbe est similaire d'un cas à l'autre.

## 1.2 Amélioration de la qualité sonore des systèmes de son solidien

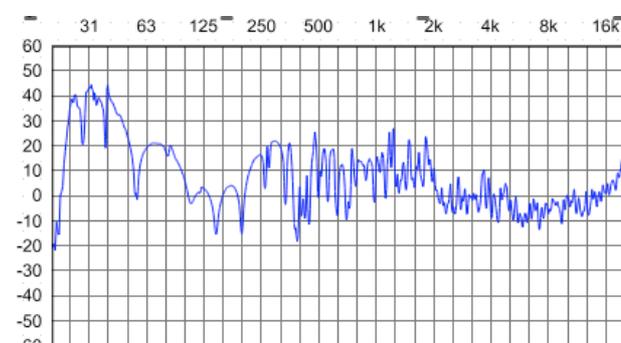
Afin de pallier à cette transformation problématique du signal, nous avons implémenté un système de filtrage numérique par l'inverse de la RI mesurée en amont de l'induction du son dans le solide. L'idée étant de contrer l'effet de filtrage par un filtrage inverse afin d'obtenir une sonorité plus proche de l'original. En guise d'exemple, la figure 2. représente l'inverse de la RI de la figure 1, avec l'application d'un léger lissage avant inversion. Notre outil a recours à la librairie d'objets *HissTools Impulse Response Toolbox* (HIRT) pour max/MSP, développée par l'équipe de University of Huddersfield [6]. Cette librairie offre une implémentation de la mesure de RI par différentes méthodes ainsi qu'un filtrage multicanal par convolution, économe en puissance de calcul. Notre outil réunit les étapes de a) mesure de RI, b) inversion et lissage de la courbe, c) filtrage par convolution, en un patch simple, relativement rapide à utiliser. Les RI des sons solidiens que nous souhaitons corriger sont complexes, et le filtrage par convolution via la transformée de Fourier apporte une solution adéquate en proposant un grand nombre de bandes fréquence (selon la taille de fenêtrage *fft*). Nous visons un outil automatisé permettant de mesurer la RI de la chaîne électroacoustique<sup>1</sup> et en déduire un filtre de correction en « un clic », le but étant de produire des filtres correctifs pour chaque transducteur utilisé. Cet outil est en cours de développement et il sera utilisé pour la création d'une pièce pour violoncelle et espace acoustiquement actif en

<sup>1</sup> Les variables de la chaîne électroacoustique traitée ici : ordinateur, interface audio, transducteur, l'objet solide et l'emplacement de transducteur sur celui-ci, microphone et son placement, caractéristiques acoustiques de l'espace.

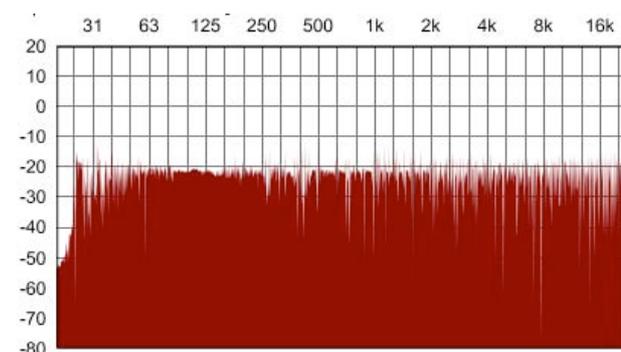
septembre 2013. Dans nos essais actuels, nous employons 20 canaux simultanés de filtrage par convolution sans problèmes liés à la puissance de calcul de l'ordinateur.



**Figure 1.** Courbe de réponse impulsionnelle typique d'un transducteur de son solidien sur une plaque de verre.



**Figure 2.** L'inverse de la courbe présentée dans la figure 1., avec un léger taux de lissage (iraverage~ *smoothamount* 0.02).



**Figure 3.** La réponse impulsionnelle résultante de l'application du filtre de la figure 2. sur le signal de la figure 1, filtrage par convolution.

Les résultats obtenus par le conditionnement du signal induit dans un solide sont encourageants – une nette amélioration de l'équilibre spectral du son est obtenue. On arrive à atténuer les défauts perceptifs évidents typiques des matériaux (par ex. résonance métallique, son sourd du carton ou de l'aggloméré). La figure 3. montre la réitération de la mesure de la figure 1. Avec l'application du filtre inverse sur le signal induit dans

l'objet. On remarque un aplanissement des modes de résonance les plus saillants visibles dans la figure 1. Toutefois, le caractère chaotique et complexe du filtre n'est pas complètement effacé, surtout dans l'aigu du spectre. Selon nos expériences, les paramètres clés pour arriver à des résultats satisfaisants sont le type et le placement du microphone pour la mesure de la RI, ainsi que le facteur de lissage de la courbe obtenue avant son inversion pour créer le filtre. Toutefois, les caractéristiques physiques de la surface telles que ses dimensions, sa densité, sa profondeur et sa rigidité, posent des limites évidentes aux performances sonores du matériau. Le son solidien reste tributaire du matériau qui le porte. Le travail avec le choix des matériaux eux-mêmes s'avère donc de première importance, et il peut être envisagé comme un choix compositionnel. L'utilisation du son solidien nécessite des compromis et des stratégies artistiques claires afin d'obtenir une qualité sonore satisfaisante. Nos recherches artistiques pour la construction des espaces sonores incorporant du son solidien sont détaillées ci-dessous (voir 3.1 et 4).

### 1.3 Orchestre d'objets acoustiquement actifs

L'induction du son dans des solides nous permet de concevoir des espaces peuplés d'objets acoustiquement actifs ou délimités par des structures architecturales sonores. Certains éléments du mobilier, les fenêtres et les murs non porteurs, trouvés *in situ*, ont des caractéristiques acoustiques satisfaisantes pour devenir sources sonores. On peut aussi fabriquer des éléments scénographiques pourvus d'une RI optimale. Il est ainsi possible de constituer non seulement un orchestre d'objets acoustiquement actifs à l'image d'un acousmonium, mais aussi une architecture sonore à large échelle.

Cette extension de l'espace sonore a pour résultat l'introduction de la construction de l'espace et de son mobilier dans le processus compositionnel. Le compositeur accède à la possibilité de travailler avec les notions de scénographie aurale et de l'architecture aurale en choisissant les matériaux des objets sonores, ainsi que leurs taille, forme et emplacement. L'espace physique de l'œuvre devient support sonore et ainsi une dimension de l'*espace composable* [5].

## 2. ESPACES AURAUX COMPOSITES

Au-delà des du son solidien, il nous semble intéressant d'étendre les recherches vers des *espaces aurales composites*, peuplés de différents types de sources sonores. Nous avons identifié cinq sources différentes :

- 1) Instruments acoustiques
- 2) Son solidien - surfaces vibrants d'objets acoustiquement actifs
- 3) Spatialisation - ensembles de haut-parleurs créant des sources sonores virtuelles ; stéréophonie, Ambisonie, VBAP (et d'autres systèmes de même type)

- 4) Haut-parleurs monophoniques (source sonore de type point localisé)
- 5) Haut-parleurs omnidirectionnels

Chaque type de source sonore possède une manière d'« habiter » l'espace, sa « signature » sonore. A l'oreille, on peut intuitivement distinguer entre ces quatre sources même dans le noir<sup>2</sup>. Les processus en jeu dans cette identification sont celles de l'analyse du champs sonore, constamment employés par notre cognition pour la construction d'une image cohérente du monde environnant. Selon Bregman [4], l'appareil audio-perceptif a la capacité d'organiser les signaux perçus en des « flux aurales » (*auditory stream*), distinguant les identités et les localisations des événements/sources sonores. Les processus exacts sont complexes et encore mal connus, faisant l'objet d'importants travaux dans le domaine de la psychoacoustique à l'heure actuelle. Dans le cadre de cette étude, nous nommons la « signature » d'un type de source sonore son *empreinte aurale*, en lien avec le vocabulaire d'identité aurale développé par Barry Blesser et Linda-Ruth Salter [3].

De notre point de vue explicitement centré sur la composition musicale la capacité de reconnaissance des « flux aurales » est essentielle, car elle permet d'identifier le type de source sonore et créer une image mentale – une topographie – de l'espace sonore.

## 3. COMPOSITION DES ESPACES AURAUX

L'articulation perçue de l'espace selon les différentes sources sonores ouvre la voie à la construction d'événements qui s'y inscrivent, donc à une écriture. Un discours musical peut ainsi s'inscrire d'abord dans un effet de distance dans un champs ambisonique, y dessiner une trajectoire avant de venir se placer dans une table à côté du public (son solidien), ou dans un haut-parleur monophonique, ou encore un instrument acoustique. Tout ce développement est perceptible – donc potentiellement porteur de sens et de possibilité de narration. Le compositeur acquiert ainsi une « dimension » supplémentaire pour créer son discours. Il s'agit d'une extension de l'idée de la « mélodie de timbres » (*Klangfarbenmelodie*) vers une « mélodie d'empreintes aurales » de l'espace sonore.

Les pièces mixtes combinant instruments acoustiques et haut-parleurs sont un exemple bien connu de l'utilisation des empreintes aurales. Par exemple, la pièce *Alien Lands* de Sandeep Bhagwati comprend une diffusion ambisonique, un quatuor à cordes ainsi qu'un quatuor de percussions [2]. La localisation du son y joue un rôle essentiel, ainsi que le type de source (acoustique –

<sup>2</sup> Ce constat est le résultat de nos tests perceptifs initiaux – il n'a pas de valeur générale et cette question nécessite un approfondissement conséquent.

ambisonie). A notre avis, il est cependant possible de continuer l'exploration plus loin, afin d'inclure les deux autres types de source sonore identifiés plus haut (voir 2.).

L'empreinte aurale des sources sonores peut être traitée comme un paramètre de composition au même titre que le timbre, la durée et la hauteur. Toutefois, la notion de l'empreinte aurale correspond à un agrégat combinant la localisation ainsi que la couleur caractéristique d'une source. L'écriture des « empreintes aurales » se confond avec l'écriture spatiale

### 3.1. Ecritures possibles

Afin de donner une idée plus précise des possibilités d'écriture qu'offre un espace aural composite, nous présentons ici quelques exemples concrets issus de nos travaux de composition en cours.

- Contrepoints : un contrepoint d'empreintes aurales peut prendre la forme d'un dialogue de sources sonores. Par exemple, un instrument acoustique peut entrer en dialogue avec un objet ou un mur, relayé à son tour par une source virtuelle ambisonique. Plusieurs sons peuvent ainsi peupler et se mouvoir dans l'espace sonore, créant des rapports entre eux. Un vaste champs de stratégies d'écritures contrapuntiques est possible, selon la quantité de sons et des sources sonores, ainsi que leur nature.

- Scènes auditives : avec un espace aural composite, il est possible de constituer des scènes auditives – soit entièrement artificielles, soit simulant une scène qu'on trouve dans le monde réel. On peut penser à une simulation d'une pluie, actualisée par un champs ambisonique immersif, ainsi que des surfaces sonores posées par terre, produisant le son des impacts des gouttes d'eau. On peut aussi créer des champs sonores éloignés de nos références habituelles, selon une distinction de son ambiant – flux sonore qui s'en détache. Par exemple, on peut produire un champs sonore uniforme avec l'ambisonie, et y placer des flux sonores saillants issus des instruments, haut-parleurs monophoniques ou encore des objets acoustiquement actifs (son solidien).

- Proche-loin : il est possible de créer des effets dramatiques de distance et de proximité avec un espace aural composite. Un mur entier vibrant (son solidien) constitue une source sonore plane, à l'opposé de la nature ponctuelle d'un haut-parleur. Il peut ainsi produire un front d'onde acoustique (quasi) plane, donnant l'illusion d'un son provenant d'un lointain infini – ce qui est problématique avec les techniques « 3D » actuelles telles que l'ambisonie [1]. D'un autre côté, le son solidien conduit au contact avec le corps du spectateur donne une perception audiotactile, porteuse d'une impression de proximité et d'intimité. Le son d'une pièce musicale peut être acheminé *dans* les corps des spectateurs.

De nombreuses autres stratégies sont possibles avec un espace aural composite, avec d'intéressantes possibilités expressives. Il est à noter ici qu'à nos yeux, les avantages d'un tel espace ne se trouvent pas tellement dans la reproduction/simulation de champs auditifs existants dans le monde réel, mais bien dans la construction d'un espace sonore nouveau, porteur de sens poétique et de dramaturgie musicale.

## 4. TRAVAUX ARTISTIQUES

Notre recherche s'inscrit dans une méthodologie de recherche-crédation. Le développement technologique et théorique y est amené en un dialogue constructif avec la création artistique. Deux projets de création sont actuellement en cours au Matralab de Concordia University, et un troisième a été complété lors d'une résidence en 2012 à Maubeuge.

Un premier projet est une pièce instrumentale pour violoncelle et architecture sonore qui explore la notion de l'instrument augmenté dans un sens figuré. L'idée est de faire dialoguer un instrument soliste avec l'espace qui l'entoure, équipé de transducteurs de son solidien. L'image est celle d'un zoom effectué dans la caisse de résonance du violoncelle qui prend des proportions gigantesques de la taille d'une pièce entière. Le spectateur entre *dans* l'instrument, ou bien l'instrument s'agrandit jusqu'à comprendre l'espace entière.

Une deuxième pièce en développement est un spectacle de danse, où le *kinesis* du corps humain est mis en rapport avec le mouvement du son dans un espace aural constitué d'une diffusion ambisonique et d'une scénographie constitué d'objets et de surfaces vibrantes. La pièce effectuera une recherche dans les percepts que peuvent produire diverses juxtapositions et relations entre le corps humain et l'espace aural composite. La pièce est chorégraphié par Satu Tuomisto et produite par JoJo – Centre de Danse d'Oulu en Finlande et verra sa première en mars 2014.

La troisième pièce est une installation d'art plastique sonore nommé *Chordophones*, coproduite avec l'Agglomération de Maubeuge Val-de-Sambre. Il s'agit d'un ensemble d'instruments à corde suspendus du plafond, équipés de transducteurs de son solidien. Une composition électro-acoustique est distribuée dans les instruments, leur donnant un son – une voix – à chacun. Il en résulte un espace acoustique et visuel où le son acoustique se confond avec le son électronique, bousculant les repères habituels de perception. Une documentation audio-visuelle de la pièce est disponible sur le site <http://otsola.org/>.



**Figure 4.** *Chordophones* – installation pour instruments à cordes et transducteurs de son solide

## 5. DEVELOPPEMENTS FUTURS

Cet article constitue un état de lieux d'un travail en cours. Un ensemble d'outils informatiques est en développement, des stratégies pour optimiser la qualité du son solide sont explorées. Un travail continu de création artistique expérimentale fournit un cadre de test pour la recherche. A l'heure actuelle, nous souhaiterions aller plus loin dans les essais perceptifs des espaces auras composites, afin de mieux saisir leurs possibilités pour la création de percepts spatiaux et timbraux. Aussi, la terminologie déjà esquissée dans cet article nécessite un approfondissement, qui n'est possible que par la pratique de ces espaces et la compréhension concrète de ce qui y est saillant. Beaucoup reste à faire, dans les trois domaines que sont la technologie, la théorie et la création. Cependant, il nous a semblé utile de communiquer les résultats actuels dans ce rapport de travail intermédiaire.

Une question centrale qui fera l'objet d'efforts supplémentaires est celle de l'écriture de l'espace aural composite. Nous sommes mus par la vision de Xenakis d'une « partition cinétique » [7] qui servirait à fixer les mouvements des sons dans l'espace, sous une forme qui se prête à être organisée par le compositeur. Cet espace est un objet composable complexe, combinant les percepts d'espace et de « flux auditif ». Sa formalisation n'est pas aisée, et par conséquent sa manipulation ne l'est non plus. Nous souhaiterions développer un outil d'écriture informatique générique, adaptable à différentes combinaisons de sources sonores. Il s'agirait d'une notation basée sur un axe temporel, permettant l'écriture précise des trajectoires des sons entre les différentes sources et localisations dans l'espace. Cet outil d'écriture serait directement lié à un *patch* de distribution dynamique de signaux audio entre les sources sonores.

## 6. CONCLUSION

Dans cet article nous avons présenté l'état de nos travaux dans l'utilisation du son solide en musique et en art intermédiaire. Les problèmes liés à la nature et aux technologies du son solide ont été discutés. Le filtrage par une RI inversée a été proposé comme solution pour contrer l'effet de coloration du signal par le comportement acoustique des solides. L'intérêt de la création des espaces combinant des sources sonores différentes a été argumenté et le terme « espace aural composite » introduite pour les désigner. Cet espace permet d'étendre le geste de la composition vers l'espace physique et les empreintes aurales des sources sonores. Enfin, un ensemble de travaux artistiques a été présenté, proposant des pistes pour l'appropriation des technologies et concepts discutés.

## Remerciements

L'auteur souhaite remercier le Fonds Québécois pour la Nature et la Technologie ainsi que le Matralab de Concordia University pour leur soutien.

## 7. REFERENCES

- [1] Colafrancesco, Julien, L'ambisonie d'ordre supérieur et son appropriation par les musiciens : présentation de la bibliothèque max/msp hoa.lib, *actes des Journées d'Informatique Musicale*, Mons, 2012.
- [2] Bhagwati, Sandeep, *Alien Lands*, pour quatuor à cordes, quatuor de percussion et diffusion électroacoustique, 2011  
<http://matralab.hexagram.ca/projects/?title=AlienLands>
- [3] Blesser, Barry et Salter, Linda-Ruth, *Spaces speak, are you listening ?*, the MIT Press, Massachusetts, 2007.
- [4] Bregman, Albert S., *Auditory scene analysis*, the MIT Press, Massachusetts, 1990.
- [5] Carvalho, G., Formaliser la forme, *actes des Journées d'Informatique Musicale*, CICM, Université Paris 8 – MSH Paris Nord, 2005.
- [6] Harker, Alexander et Tremblay, Pierre Alexandre The HISSTools Impulse Response Toolbox: Convolution for the Masses. *Proceedings of the International Computer Music Conference*. 2012
- [7] Kanach, Sharon (Ed.), *Music and Architecture*, by Iannis Xenakis, Pendragon press, New York, pp. 144-145, 2008.
- [8] Lähdeoja, Otso et Reboursière, Loïc, Augmented window: structure-borne sound drivers for sound-emitting solid objects and surfaces, *QPSR of the numediart research program*, Vol. 4, No. 4, Décembre 2011

[9] Leitner, Bernhard, *Sound Space Manifesto*, New York, 1977

<http://www.bernhardleitner.at/texts>

[10] Pook, Lynn, *Stimuline*, installation audio-tactile.

<http://www.lynnpook.de/english/stimuline/index.htm>

[11] Stein, Julien, Stein, Max et Basanta, Adam, Music for 12 lamps,; <http://www.julianstein.net/>

# ATELIER AUTOUR DU SÉQUENCEUR I-SCORE

Myriam Desainte-Catherine, Nicolas  
Hincker

Pascal Baltazar

Théo de la Hogue

LaBRI / SCRIME  
[myriam@labri.fr](mailto:myriam@labri.fr)  
[nicolash24@hotmail.com](mailto:nicolash24@hotmail.com)

Les Baltazars / l'Arboretum  
[pascal@baltazars.org](mailto:pascal@baltazars.org)

GMEA CNCM  
[theo@gmea.net](mailto:theo@gmea.net)

## RÉSUMÉ

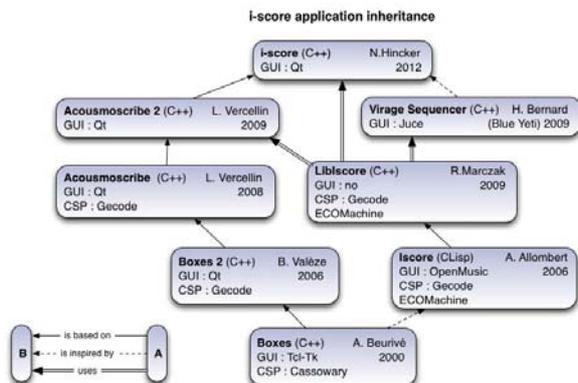
Dans le cadre de Journées d'Informatique Musicale, le LaBRI propose un atelier autour du logiciel d'écriture intermedia i-score. Ce logiciel est le fruit d'une collaboration entre l'Arboretum et le LaBRI à travers les développements réalisés par Nicolas Hincker en dialogue avec Pascal Baltazar et sous la supervision scientifique de Myriam Desainte-Catherine.

L'accueil des Baltazars en résidence de création a permis de confronter l'ergonomie d'i-score à une utilisation poussée en situation de production. Ces retours concrets ont ainsi pu guider une relation de travail art-science pour le design d'une interface renouvelée.

Enfin c'est aussi l'occasion de présenter le futur de ce logiciel dont les fonctionnalités vont s'étendre aux écritures d'un temps non linéaire contraint par des relations logiques.

## 1. HÉRITAGE

Le logiciel i-score est basé sur plus de 10 années de recherche au LaBRI. Son moteur et son interface ont tout deux profité de multiples recherches et développements qui sont résumés dans le schéma ci-dessous :



## 2. APPROCHE

Le logiciel i-score est un séquenceur intermedia pour l'écriture de scénarios interactifs. À la manière d'un chef d'orchestre, l'exécution d'un scénario i-score permet de contrôler d'autres logiciels satellites compatibles avec les protocoles OSC ou Minuit (tels que Max ou Pure Data) selon un temps souple et interactif respectant tout de même certaines contraintes de structure.

L'utilisateur peut disposer des boîtes dans le temps, les relier entre elles ou à des points de déclenchement. Leur contenu peut être édité par une liste de commandes en début ou en fin de boîtes, ainsi que par des automations. Celles-ci seront alors émises aux logiciels satellites lors de l'exécution.

Le déroulement d'un scénario se fait d'une manière similaire à n'importe quel séquenceur, à ceci près qu'il est ici possible de retarder ou d'anticiper l'exécution des boîtes reliées à des points de déclenchement tout en laissant inchangée l'exécution d'autres boîtes non soumises à ces événements interactifs. Enfin l'exécution peut être ralentie ou accélérée afin d'adapter le déroulé au temps vivant d'une représentation.

## 3. INTERFACE

L'interface d'i-score se décompose en deux espaces distincts : la « *timeline* », où les éléments temporels sont disposés et reliés, et l'« *explorateur* », qui permet de visualiser sous forme arborescente l'ensemble des paramètres des logiciels distants et de les sélectionner afin de les affecter aux boîtes temporelles.

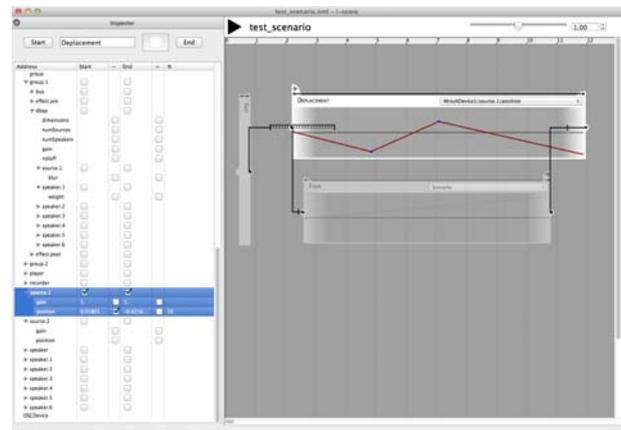


Figure 2: aperçu de l'interface d'i-score avec l'explorateur (à gauche) et la timeline (à droite)

En comparaison avec les interfaces précédentes, l'ensemble des informations présentées aux utilisateurs a été allégé et réorganisé de manière à réduire les actions accompagnant la réalisation d'une tâche d'écriture.

Concernant l'édition de la *timeline*, les tâches de création de boîtes temporelles et de mise en relation respectent en première intention les accès traditionnels qu'offrent d'autres séquenceurs : les durées et les dates d'exécution sont fixes *a priori* mais l'utilisateur peut ensuite accéder simplement aux fonctionnalités avancées de mise en interactivité du scénario.

Les tâches de sélection de paramètres dans l'explorateur sont aussi au cœur du travail en cours. La grande quantité de paramètres que peut offrir une application satellite confronte les utilisateurs à un temps de recherche trop long lors de la réalisation de mémoires ou d'automations.

Nous envisageons une solution qui s'appuie sur la possibilité déjà existante de créer des boîtes hiérarchiques. Il s'agira à terme de réduire l'espace des paramètres affichés à une sélection liée à un sous-scénario. Ainsi le choix de regrouper des éléments temporels au sein d'un même sous-scénario conduira nécessairement à focaliser l'écriture sur un sous-ensemble de paramètres plus simple à appréhender.

#### 4. UN FUTUR LOGIQUE

Jusqu' alors les avancées menées par le LaBRI ont montré l'utilité pour des régisseurs numériques dans le spectacle vivant de pouvoir écrire et contrôler un temps souple et contraint.

En outre, l'essor rapide des expériences interactives ou vidéo-ludique dans notre quotidien montre qu'il est aussi nécessaire de pouvoir écrire et contrôler un temps non linéaire à destination d'utilisateurs multiples interagissant simultanément. Que ce soit dans les domaines de la muséographie, du jeu vidéo, des web documentaires, du spectacle vivant. L'écriture des sons, des images, des lumières ou des mécanismes tend non plus à les assembler indépendamment mais cherche à les penser ensemble et en interaction.

Pendant 3 ans, le projet OSSIA (Open Scenario System for Interactive Application), financé par l'Agence Nationale de la Recherche dans le cadre de l'appel à projet CONTINT, cherchera à constituer un ensemble d'outils logiciel pour la scénarisation dans le domaine de l'interaction numérique.

Dans le cadre de ce projet, le LaBRI envisage deux extensions de son modèle pour permettre l'écriture de scénarios ouverts : les branchements temporels conditionnels et les boucles. Les branchements vont permettre de finir plusieurs chemins d'exécution selon des événements reçus en temps réel. Les boucles vont permettre d'exécuter plusieurs fois un même objet temporel, par exemple tant qu'une condition est vérifiée.

#### 5. RÉFÉRENCES

Allombert, A., Desainte-Catherine, M., Toro, M., « Modeling temporal constraints for a system of interactive score » in *Constraint Programming in Music*, G. Assayag and Charlotte Truchet, Eds., chapter 1, pp. 1-23. Wiley, 2011.

Allombert, A., Marczak, R., Desainte-Catherine, M., Baltazar, P., Garnier, L., « VIRAGE : Designing an interactive intermedia sequencer from users requirements and the background », *International Computer Music Conference*, New York, USA, June 1-5 2010, International Computer Music Association (ICMA), 2010.

Baltazar, P., Allombert, A., Marczak, R., Couturier, J.-M., Roy, M., Sèdes, A., Desainte-Catherine, M. « Virage : Une réflexion pluridisciplinaire autour du temps

dans la création numérique », *Actes des 14<sup>e</sup> Journées d'Informatique Musicale*, Grenoble, 2009.

Allombert, A., Desainte-Catherine, M., Assayag, G., « Iscore : Writing the Interaction », in *Proceedings of the 3rd Digital Interactive Media in Entertainment and Art (DIMEA)*, Athens, Greece, September 2008.

Allombert, A., Desainte-Catherine, M., Assayag, G., « De Boxes à Iscore : Vers une Ecriture de l'Interaction », in *Proceedings of the 13th Journées d'Informatique Musicale (JIM)*, Albi, France, Mars 2008.

Toro, M., Desainte-Catherine, M., « Concurrent Constraint Conditional Branching Interactive Scores », in *Proc. of Sound and Music Computing (SMC) '10*. Barcelona, Spain, July 2010.

Toro, M., Desainte-Catherine, M., Baltazar, P., « A Model for Interactive Scores with Temporal Constraints and Conditional Branching », in *Proc. of Journées d'informatique musicale (JIM)*, pp31--38, Rennes, France, May 2010.

# DEVELOPMENT OF A MORPHING TOOL FOR THE "COSMOSF" SYNTHESIZER

*Sinan Bökesoy*  
 sinan@sonic-lab.com  
 15/6 Dogan bey sok.  
 Buyukada-Istanbul

## SUMMARY

The purpose of this article is to explain the development of a sound morphing mechanism for the *Cosmosf* advanced stochastic synthesizer<sup>1</sup> in its compositional workflow. A sonic morphing process defines a metaphor of exploring the sonic shape and timbral evolution on intermediate levels between two or more sonic materials. There is no robust methodology to bring exact solutions in morphing processes however they offer a rich terrain of possibilities which could be handled within various approaches to bring interesting perceptual phenomena. Themes regarding the parameter space handling of *Cosmosf*, the 3D visual interface for the morphing operation, automation and innovative live control methods will be introduced here.

## 1. INTRODUCTION

The term <morphologie> corresponds scientifically to study of the form and structure. A short definition of sound morphing would be applying gradual transformation from one sound as the departure to another sound as the destination. The transfer of features of one sound onto another sound in order to generate cross synthesis of hybrid timbres [1][2][3]. It is a network of methodologies and organization of diverse essays for bringing heuristic results. These solutions are naturally surrounded with restrictions and artifacts regarding the analysis methodology applied for the morphing process. [4][5][6] The answer to the question of what the best and most affordable process is not clear. The purpose here is not presenting a totalized argument about morphing of sonic materials, and revealing just the constraints of the existing morphing methods but rather showing the variety that can be created within a structured matrix of divergent processes. Here, the morphing process should not be regarded as a strict mimicking process, but as a synthesis tool with specific ease to use which can even establish a shortcut to instant creativity to the composer.

### 1.1. Synthesis by analysis and criteria of morphing

<sup>1</sup> *Cosmosf*<sup>TM</sup> is designed & programmed by Sinan Bokesoy, available at [www.sonic-lab.com](http://www.sonic-lab.com)

Sound morphing can be performed by blending the parameters of the consistent structures which do generate the source and destination sounds. An important aspect of such a hybridization process is to understand that there are several possible ways of combining and cross synthesizing things [7]. Generically the definition above assumes that the matters presented as source and destination would be concrete sound samples. The morphing between audio sample materials through analysis is a top down processing approach. Through analysis we focus to obtain several feature based audio descriptive parameters which deliver the evolution of perceptually meaningful components over time.[8] However this information lacks in projecting the physical structural facts other than statistical measurements based on the analysis of perceptual parameters.[3] STFT assumes that all the sounds are built with accumulation of sinusoidal partials, it delivers useful base to extract perceptual parameters but does not inform us about the intrinsic parameters of the structural model of the sound. Cross synthesis which create useful results practically makes use of the extracted feature space and is therefore limited within the analysis process by proposing an incomplete representation of the source and destination materials. Nevertheless, the topic spectro-morphology has gained a lot of access among composers and researchers to bring a good variety of results (Figure 1). [3][5][9][15]



**Figure 1.** Morphology of Wishart described on *Sonic Art* pp27.

Though the sample materials are fixed mediums, the morphing process is not necessarily between equal structures, since we are not necessarily under the hood of a unified synthesis structure which has generated these sounds. Morphing process between physical models of sounds and cross synthesis by analysis works well when the destination has similar structure in comparison to the source structure of the sound with adequate feature representations [5].

To gain more success in morphing between sample sounds, a useful restriction the 'correspondence' is a looked-after quality even regarding the digital representation (recording quality). When morphing sounds, addressing the correspondence between model parameters is important. Be it an appropriate sinewave additive model or source-filter models, the morphing process becomes smoother between musical instruments that guarantees temporal and spectral correspondence.[7] More specifically, correspondence between the analysis frames of the sounds and between spectral envelope parameters representing each frame. If there is a structural correlation between source and destination, then the parameter space handling will be the most effective for applying a gradual morphing process.

The gradual transformation of the morphing process between a rigid source and destination sample sound is expected to establish linear connections and vector translations of perceptual parameters. But some application areas of the morphing process such as when interpolating the parameters of physical models [4] the morphing process does not lead to perceptually linear results. Additional techniques for warping the interpolation is needed. The gradual morphing process itself becomes a fertile medium, however as a synthesis tool the non-linear aspects are even offering other unvisited palette of possibilities.

### 1.2. Morphing within a synthesis model

Wishart[15] categorizes the sounds of coherent structure, but constantly changing in their present state and still keeping their global timbre space in terms of possible variations as multiplex sonic objects. The individual components exhibit a process of change gradually in timbre space exhibiting dynamic morphology. What methods can be proposed to morph effectively in between such sonic objects rather than top-down analysis approach using the concrete samples ? One observation is that the classification of such sounds and of their morphing process is not relevant, it does not count to the typology of the sounds as the aim is to generate unheard outcome. [14].

If we can communicate directly with the synthesis model at the level of a bottom-up sonic construction process, and apply the morphing process under the hood of the synthesis model by observing the process of change of the individual components, the result will deliver the necessary correspondence and smoothness in gradual transformations.

We can define the current state of the parameter space as a preset of sound, reminiscent to hardware synthesizers user interfaces. The parameter morphing process incorporates handling of these structural parameters with a high level interpolation mechanism accessible easily by the user, who will scan this sonic

space characteristic to the sound synthesis model. The translations of these gestures are linear and recallable operations in form of parameter space transformations. The weight of each preset destination is being calculated by euclidian distance calculation. Morphing process through the handling of the intrinsic parameter space is interesting as it offers to the user accessing the intermediate states directly on the structural level not on the concrete audio sample material.

The next chapter will introduce the *Cosmosf* synthesizer (Figure 2), which does create multiplex sonic results evolving in a multiple timescaled process within a bottom-up construction mechanism, likewise a good example of implementation of a morphing tool in such category.

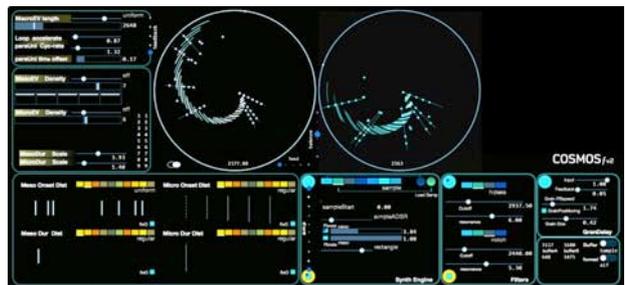


Figure 2. A part of *Cosmosf* app. Interface.

## 2. COSMOSF

*Cosmosf* is a multi time-scale complex event generation system (Figure 3), where the events are organized to create a sonic phenomena within controlled complexity. The event generation system on each time scale is driven by a stochastic mechanism inspired by Xenakis's formalized music tools [10] but by making benefit of modern computing power.

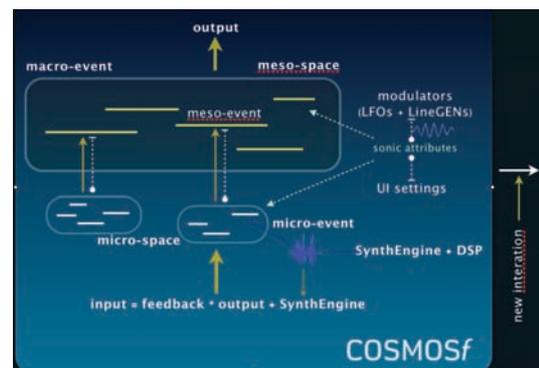
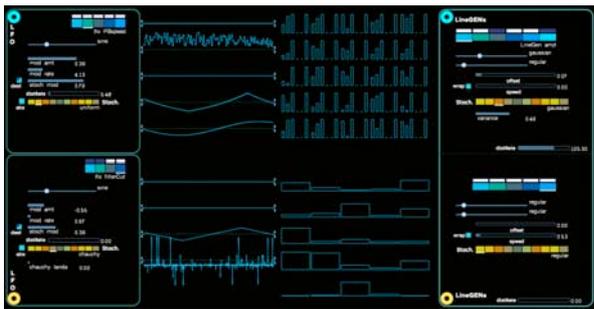


Figure 3. The basic flowchart of *Cosmosf*.

- The system creates the event space by first defining a macro space for distributing its events. Inside the macro space, meso events of certain density are distributed. (Figure 3)

These meso events also contain a micro space with certain density of micro events. This top down event distribution approach is being accompanied with a bottom-up sonic construction, where the micro events are containing the sonic data and together their micro space deliver the sonic data to each meso event.

- Various modulation sources generated by stochastic functions, LFO's and LineGEN's can be applied on micro event or meso event levels. (5 for each micro event, and 4 for each meso event). Maximum event density is 9 for meso events and 9 for micro events, which is in fact only a limitation of computer processing power. (Figure 4)



**Figure 4.** The multiscale modulation sources in *Cosmosf*, offering complex manipulation of various parameter destinations.

- Recursive audio transfer from macro space output to the micro event entry can be applied. (Figure 5) The output buffers can capture the audio output of a macro space, freeze the captured material and deliver the content to the micro events for further processing. The length of the macro space, meso events, the micro events and their onset times can be determined with various stochastic methods offering a dynamic behavior within controlled degrees of complexity.



**Figure 5.** The output buffers of *Cosmosf* deliver the macro sound back to the micro level. They can continuously feed back the signal or be frozen to deliver a static material.

- The application also lets its clock speed being altered in continuum precisely between extreme settings apart from the macro space duration and event distribution process. Time is an independent dimension in the hands of the composer for creating passages from micro to macro levels in continuous gestures.

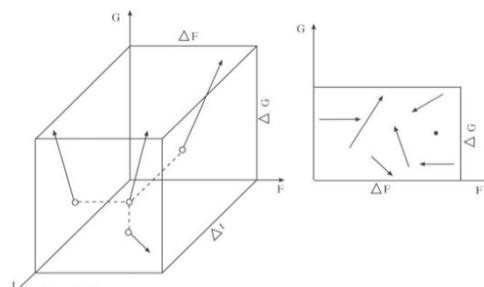
Thanks to its coding in C++, *Cosmosf* delivers particular accuracy of handling immense amount of data within a special interface design. The various facts of the model and application has been stated before in relevant resources in depth. [11][12]

### 3. MORPHING IN COSMOSF

Production music being created by the manipulation of concrete materials relies on organizing sonic material manually in order to create desired structures with temporal aspects and morphologies but with hand-made interaction. *Cosmosf* sonic universe incorporates a dynamic morphology, which is practically hard to obtain with such methodologies. In fact, the purpose of *Cosmosf* is to deliver to the composer the computing power to create generative processes, which is not possible to achieve by hand. The design of the morphing process in *Cosmosf* permits the composer to focus away from low level parameter selection to higher-level control structures offered by *Cosmosf* user interface.

#### 3.1. Structural facts if *Cosmosf* on the morphing

An instance of the morphing process points to a certain state between two or more sounds/preset parameters which can be represented as a multi-dimensional vector in the parameter space. Vectorization of sound as within the definition of sonic entities and their accessibility problem in the timbre space has been investigated before [9][10]. Xenakis has proposed as a unifying concept the design process for the transformation of sound. He takes instances of sound as a quantifiable and controllable entity (Figure 6) by proposing the vector definition of sonic entities with offering 4 independent dimensions e.g. the  $Er(c,h,g,u)$ : the timbre or instrument family, the pitch, the density and the duration depending to an axis of lexicographic time.



**Figure 6.** Grains in the timbre space distributed in a limited volume (Xenakis, *Formalized Music*, Fig. ii-12, pp55)

The function  $F(Er,t)$  defines the temporal behavior of the pointed sonic entity in this multidimensional timbre space.

### 3.2. Incomplete representation

The vector dimensions of the morphing pointer in *Cosmosf* do include all the parameters available on the user interface. We can categorize them as parameters effecting directly the perceptual attributes, the event distribution process and signal processing/modulation parameters. According to cybernetics approach; a system subjected to complete control is necessarily incompletely represented, which is a problem of the construction of our knowledge about the system under observation [17]. Among the perceptual attributes, which *Cosmosf* cannot address precisely, the pitch is a parameter where its perceptive phenomena is dependent on the temporal scale of the sonic entity carrying this information. The micro-time scale for the definition of pitch is inadequate. The organization of micro sonic entities horizontally and vertically, can create an illusion of perceptive phenomena by destructing the perception of pitch and constructing emergent perceptive properties. (integration/segregation and sequential organization) [13].

On *Cosmosf*, pitch can be defined as a sample playback speed on the level of micro events and also at the level of the additive synthesis partials with defined frequency value. However on higher levels the projection of this dynamic system driving a dynamic morphology, events do break the connection with the departure pitch value of sonic entities (if perceivable) and the various local pitch values perceived on different levels. Finally pitch on the level macro is not a dimension of the morphing vector.

Another perceptual parameter, amplitude/intensity can be set on each level of event distribution. But again, like the pitch it does not exist as a macro level dimension of the morphing vector.

In a synthesis model with a bottom up construction of audio and recursive treatment of sound signals, the process of change against time becomes itself an act of the sonic design. The morphing vector interacts with the event distribution mechanism and modulation sources for the following time factors :

- *onset and duration times on micro, meso, macro level*
- *speed of micro and meso event playback* (where time change here will be expressed as pitch manipulation)
- *speed of modulation sources* : extreme values define various additional phenomena such as frequency or amplitude modulation or spectral change with secondary phenomena.

But the clock speed of the application has been kept as an independent control of the structural mechanism and therefore of the control of the morphing pointer vector. The final form of sound becomes a function of  $E(\text{clk}(t), M(t))$ , where  $\text{clk}(t)$  is the clock-rate over time and  $M(t)$  is the morphing vector over time.

### 3.3. The morphing process in recursive audio mode

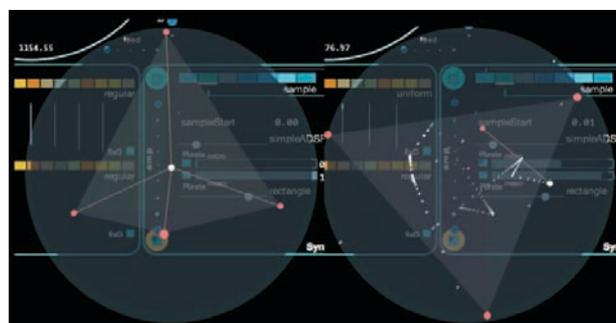
The recursive sonic process implemented in *Cosmosf* offers a nonlinear sonic state for the morphing pointer which points to a linearly related dimensional parameter space. A non-linear state cannot be restored by manipulating the controllable parameter space without any memory function calling to the previous sonic states. The end material is not fixed and pointing on its parameter space in a steady parameter space signifies only the departure point of the process as the initial parameters. Therefore naming this process as 'morphing' between defined destinations is a blurred definition, as the destinations are changing continuously.

## 4. CONTROL OF THE MORPHING MECHANISM

*Cosmosf* V2.2 morphing tool introduces a spheric space where the 4 presets of *Cosmosf* is being located on the surface of the sphere by representing the corners of a tetrahedron. (Figure 7) A spherical space by its polar coordinates is defined like;

$$\begin{aligned}x &= r * \sin(\text{theta}) * \cos(\text{phi}) \\y &= r * \sin(\text{theta}) * \sin(\text{phi}) \\z &= r * \cos(\text{theta})\end{aligned}$$

where the  $x,y,z$  are the cartesian coordinates retrieved from the spherical coordinates, radius  $r$ , inclination  $\text{theta}$  and azimuth  $\text{phi}$  with their range by definition :  $r [0,1]$ ,  $\text{theta} [0,2\text{PI}]$  and  $\text{phi} [0,\text{PI}]$ .



**Figure 7.** The 3D morphing sphere of *Cosmosf*. The tetrahedron corners can be seen as red spheres representing each assigned preset. The morphing pointer is the white sphere and the lines connected to it represent the weight of each preset. The interface shows the last 150 positions of the pointer by fading it to gray.

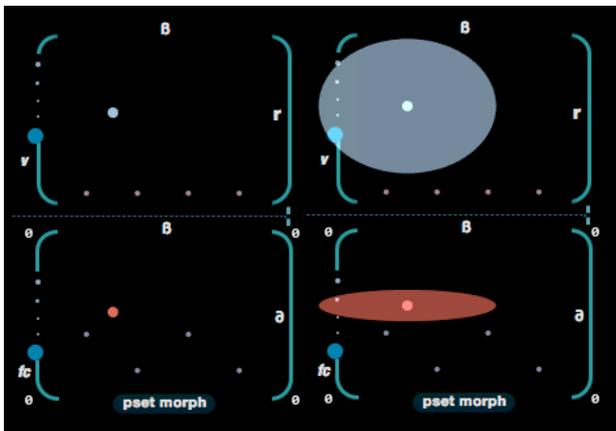
By changing the polar coordinate angles  $\text{phi}$  and  $\text{theta}$ , and the *radius*, one can access to every point inside this morphing sphere. The user can load different presets saved in the preset library to assign to each corner of the tetrahedron. The weight of each preset parameter destination is being calculated by an euclidian distance calculation, using the morphing pointer position and the defined tetrahedron corner coordinates.

The C++ arrays *morphSlotINT[ii][i]*, *morphSlotFLT[ii][i]* and *morphSlotSTR[ii][i]* are holding the integer, float and string valued parameters of the 4 presets.

For instance to calculate the weighted float valued parameters according to the morphing pointer this routine will be used;

```
for (int ii = 0; ii<500; ii++) {
    morphSlotFLT[ii][4]=
    morphSlotFLT[ii][presetA]*dist[0]+
    morphSlotFLT[ii][presetB]*dist[1]+
    morphSlotFLT[ii][presetC]*dist[2]+
    morphSlotFLT[ii][presetD]*dist[3];
}
```

where the *dist[0]*, *dist[1]*, *dist[2]* and *dist[3]* are the normalized distances of the pointer from the tetrahedron corners as delivering the weight for each preset parameter. Their sum is always equal to '1'.



**Figure 8.** Adjustment of polar coordinates on *Cosmosf* user interface. The red and blue circle can be dragged by mouse precisely. The ellipse around them represent the stochastic distribution range affecting each circle. With the vertical sliders one can filter the movements of the morphing pointer and also change the speed of the stochastic distributions.

#### 4.1. Manipulation of the morphing pointer

The morphing pointer can be moved with the mouse by adjusting the polar coordinates *phi*, *theta* and *radius* as shown on the Figure 8. Adjusting the '*phi*' or '*theta*' alone lets us move the morphing pointer on circles of longitude and latitudes reminiscent to geographical positioning with radial distance '*r*'.

Adjusting the radial distance '*r*' alone will move the pointer on a line through the center of the morphing sphere. Also *Cosmosf* offers the user some mathematical functions calculated and updated in audio-rate, which can serve as perfect paths moving the morphing pointer along with by only adjusting the speed of the movement.

All these interactions offer interesting phenomena of accessing the polar coordinates.

#### 4.2. Stochastic morphing

The use of stochastic distribution functions on the spheric parameters lets us to move the morphing pointer randomly in respective regions of the sphere. *Cosmosf* has many probability distribution functions integrated in its mechanism such as in order to define the polar coordinates for the morphing state we can apply them directly on the parameters *phi*, *theta* and *radius*. Likewise a rich indeterministic approach will be introduced where it can simulate the probability that the sonic state points in the clouds of probability functions formed in 3D visual space of the morphing sphere (Figure 9).



**Figure 9.** Random modulation of the polar coordinates. On the left all the three are uniformly modulated, on the right the '*phi*' angle modulation is dominant, and radius *r* is static.

The user defines the range of the distribution functions applied on the relevant polar coordinate. (Figure 8) It is shown as the ellipse size drawn around the polar coordinate value which itself becomes the center of the ellipse on the user interface. According to these settings the stochastic distribution functions modulate the position of the morphing pointer in 1D (lines, or arcs or circles) 2D (filled circles or arcs) and 3D (cones, and regions of the spherical space) volume elements or surface fields. It becomes a continuous gesture traversing the space. (Figure 10)



**Figure 10.** Random modulation can be restricted to form interesting phenomena of the morphing pointer movement in 3D space.

The control of the morphing pointer can be achieved also by external software such the graphical event composition applications [16] via the OSC protocol and therefore additional manipulation processes can be created. However this kind of control has a limitation of being in control rate and not in audio rate resolution.

*Cosmosf* breaks this limitation as it does calculate the weight of each parameter for the morphing tool in audio rate. Audio rate means, that the application does the calculations and updates inside the audio routines, 44100 times a second. Its update routine interacts with the UI and also with the OSC incoming data only in control rate. The audio-routine delivers a *phasor* with lets us to maintain it as the master clock for the application. With its integrated mathematical functions, traversing in the vector space of parameters happens with highest precision and controllable change from order to disorder thanks to its stochastic features.

Controlling the process time independently for the different levels of the morphing tool lets us manipulate the following aspects independently (Figure 8);

- *the rate of the stochastic distribution functions which modulate the position of the morphing pointer*
- *the degree of the interpolation between discrete destinations inside the morphing sphere.*
- *the rate/speed of the mathematical functions for moving the morphing pointer along their path.*

Additionally we can still control the rate of these operations independent of the *Cosmosf* cycle length, hence we could slow or speed up the *Cosmosf* cycle, where we could keep the speed of the motion of the morphing pointer.

Likewise *Cosmosf* gains another tool to create sonic emergence and perceptual phenomena as to happen due to this kind of time warping and automation possibilities. The morphing pointer becomes a tool for dynamic stochastic synthesis as the resolution of the operation equals to the sample rate, hence drawing the waveform itself as the result.

The parameter weighting of the morphing pointer applies its value depending on the micro, meso or macro level operations of the relevant parameter. *Cosmosf* is listening to all these weighted parameter calculations in audio-rate but the *Cosmosf* event generation or the synthesis engine reacts to these changes in micro event time or meso event time. For instance; if the meso event density has been changed, then the change is applied on the next meso space calculation. If the macro cell length has been changed, it will take its effect at the beginning of the next cycle. However the synthesis engine and modulation sources listen to the parameters at audio-rate resolution.

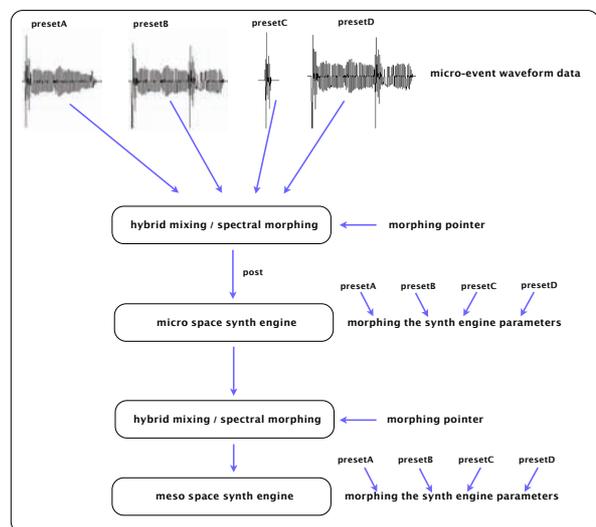
### 4.3. Behavior of the morphing tool on the synthesis engine

There are several methods in applying the result of the morphing tool. These address the organization of the synthesis engine. (Figure 11) For instance;

- The morphing tool can apply the morph between each presets micro event sample data in two different ways:

1. By applying an amplitude balance among the waveform data assigned to the micro events can be achieved between the 4 presets.
2. A spectral balance among the waveform data assigned to the micro events can be achieved.

The latter applies a basic spectral additive balancing process between concrete waveform data without doing the perceptual analysis process. The idea is to obtain intermediary results.



**Figure 11.** The morphing acting on various levels of the synthesis engine in *Cosmosf*. On the parameter space and also on the waveform data by known morphing methods. The multi-scale approach is another door to emergent sonic design within this scheme.

- The amplitude balance of the 4 weighted micro event outputs can be obtained as “post” processed by the synthesis engine, and “pre” to the meso events. This is a hybrid mixing process of the 4 presets at the “pre” meso event level without balancing/morphing their micro-event waveform data. The meso event synthesis parameters, all the distribution and modulation parameters will be morphed.
- The above, plus the weighted parameter morphing of the micro event preset data. This is a hybrid mixing of the morphed presets with according to the weighted parameters among all the preset data.

In summary, the morphing tool tries to apply the morphing process on various levels on the defined input data from the user. This way, a rich palette of intermediate results can be achieved in real time.

#### 4.4. Always there is room for improvements

The morphing tool has a potential to balance and create new parameter space among the chosen presets by combining different strategies such as hybrid feature mixing, morphing and operating on different time scales according to the structural levels micro, meso or macro. The audio rate operation and modulation by stochastic functions, automation with mathematical functions and the manipulation of the clock rate, all offer unique results under a unified synthesis engine tool. Regarding the criteria of the morphing processes introduced in the introduction section, the morphing tool being implemented for *Cosmosf* has still room for many improvements, and primary goals are now developing;

- *the ability of the user to choose features with different weights in order to influence Cosmosf behavior.*
- *specification of different distance calculation strategies for the parameter space or a selection of it.*
- *a surround sound support benefiting from the 3D morphing space.*

### 5. CONTROLLING THE MORPHING PROCESS WITH EXTERNAL HARDWARE

As *Cosmosf* parameter space can be controlled with OSC compatible hardware, the aim here is to provide the morphing process as a performance tool, acting on the software parameters at once with global simple gestures in realtime. Since a while, 3D infrared depth sensors / time of flight cameras has arrived as end user products. With the help of such modern control tools which can offer a 3D touch free control environment, one can develop a perfect platform for the control of the *Cosmosf* morphing tool which has a 3D visual interface. (Actually mouse/touch pad operation within a 3D environment is an indirect medium.)

#### 5.1. *Cosmosf* and Kinect<sup>2</sup> / SigmaNIL<sup>3</sup> library

The SigmaNIL vision framework operates with any depth sensing camera (Kinect, AsusXTion Pro Live etc.) and provides tools for finger level precision hand shape recognition, hand gesture recognition and hand skeleton tracking by analyzing the depth image data coming from

the sensor device. Using the hand shape recognition module of SigmaNIL we can recognize the following shapes to start modulating the morphing parameters of *Cosmosf*. Performance of these shapes is the initial step to put the overall state of the application in live editing mode. (Figure12)



**Figure 12.** Any shape in similarity with the above finger compositions will be recognized as the initiation of the editing mode.

While in editing mode, each distance between the finger tips, thumb-index-middle, are mapped to the assigned morphing parameter of *Cosmosf*. The distance between the finger tips are being calculated by using the finger tip joints provided by the SigmaNIL hand skeleton module.

All handshapes and hand skeleton targets used for this editing process require a certain training process by using the embedded SigmaNIL tools. Different persons have applied these shapes in a repetitive session in front of the camera for a certain duration. (Figure 13)



**Figure 13.** Depth field image retrieved from theKinect sensor and the semantic mapping of the for the hand shape by SigmaNIL.

These depth images are the input for the training process, doing the analysis and finding out the best feature candidates defining the hand shapes based on the hand skeleton data for these specific targets. (Figure 14)



**Figure 14.** Hand skeleton construction and distance calculation according to the joint data by SigmaNIL.

<sup>2</sup> Kinect™ sensor , Microsoft Corporation 2010.

<sup>3</sup> SigmaNIL™ Framework by SigmaRD Computer Vision & Graphics.

To exit the editing mode with the current state of the parameters, one possible solution is that the user turns the hand by keeping the relevant finger distances around the wrist so that the thumb tip position appears higher on the 'y' axis relative to the other finger tips.

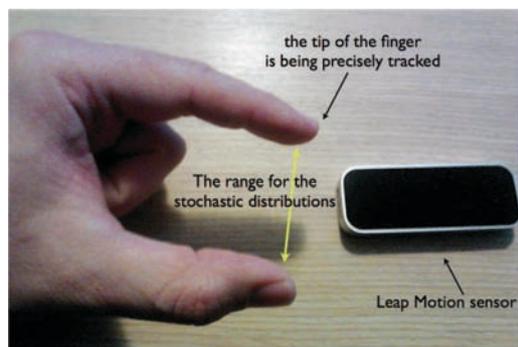
The analysis speed is limited by the camera frame rate. In order to increase the precision of the editing process, the user should hold the hand close to the camera in order to capture the hand in full sight e.g. 640X480 resolution. The precision is also depending on the distance from the camera, since the infrared beam dispersion angle is a factor of the strength and precision of the 3D scanning and does operate better in near field.

The mapping the relevant data to *Cosmosf* morphing space happens via OSC transfer at the moment.

## 5.2. *Cosmosf* and Leap Motion Controller

At the time of writing this article, we had the hands on a developer version of the Leap Motion<sup>4</sup> unit. It is a desktop unit, which offers the tracking of all the finger position data along with the orientation and 3D positioning on the air. The SDK of Leap Motion allows to implement the captured high-level streaming data in a C++ environment easily.

Via the Leap Motion Controller, the morphing pointer can be controlled by moving the tip of a finger in 3D space. The finger tip position is being mapped to *Cosmosf* spherical morphing environment and converted to polar coordinates. (Figure 12)



**Figure 15.** The Leap Motion unit and an example gesture controlling the range of distribution functions modulating a polar coordinate of the morphing tool.

The application can track the number of hands and fingers on the scene. The following gesture controls have been implemented.

- *Single Hand - 1 Finger pointing* : Morphing Pointer position data mapping in 3D space.
- *Single Hand - 2 Finger range definition* : defines one polar angle stochastic distribution range
- *Double Hand - 2 Finger range definition* : defines polar angles stochastic distribution range for each hand
- *Double Hand - 2 Finger radius definition* : The distance between hands define the radius stochastic distribution range.

This simple and short code is an example of extraction of the positioning data and processing it.

```
if( leap.isFrameNew() && simpleHands.size() ){
    if (simpleHands.size() == 2){
        for(int i = 0; i < simpleHands.size(); i++){
            if(simpleHands[i].fingers.size() == 2){
                for(int j = 0; j < 2; j++){
                    int id = simpleHands[i].fingers[j].id;
                    ofVec3f pt = simpleHands[i].fingers[j].pos;
                    pp[i][j] = pt;
                    hand1distance = pp[0][0]-pp[0][1];
                    hand2distance = pp[1][0]-pp[1][1];
                    hand1med = (pp[0][0]+pp[0][1])*0.5;
                    hand2med = (pp[1][0]+pp[1][1])*0.5;
                }}
            }
```

*Finger 1* and *Finger 2* vectors are subtracted and the length of the resulting vector is calculated on each hand, giving the range data for the relevant gesture. Also the middle space point between *Finger1* and *Finger2* is assigned as a position data to calculate the distance between two hands.

These are the most simple implementation schemes. Many further gesture analysis and teaching methods can be applied for controlling additional parameters of *Cosmosf* based on the knowledge of previous research on the field. As a first impression, the control in a 3D touch free environment works robust and precise due to the impressive frame rate of the data delivered from Leap Motion. It encourages us to continue the development of the 3D user interface.

## 6. REFERENCES

- [1] M. Caetano and X. Rodet, "Automatic timbral morphing of musical instrument sounds by high-level descriptors", in Proc. ICMC, 2010.
- [2] —, "Sound morphing by feature interpolation", in Proc. ICASSP, 2011.

<sup>4</sup> Leap Motion Inc. 2012, [www.leapmotion.com](http://www.leapmotion.com)

- [3] J. M. Grey and J. W. Gordon, "Perceptual effects of spectral modifications on musical timbres", *Journal Acoust. Soc. Am.*, vol. 63, no. 5, pp. 1493–1500, 1978.
- [4] T. Hikichi and N. Osaka, "Sound timbre interpolation based on physical modelling", *Acoustics Sci Technology*, vol. 22, no. 2, pp. 101–111, 2001.
- [5] N. Osaka, "Timbre interpolation of sounds using a sinusoidal model", in *Proc. ICMC*, 1995.
- [6] E. Tellman, L. Haken, and B. Holloway, "Morphing between timbres with different numbers of features" *Journal Audio Engineering. Society*, vol. 43, no. 9, pp. 678– 689, 1995.
- [7] M. Caetano, N. Osaka, "A Formal Evaluation Framework for Sound Morphing", in *Proc. ICMC*, 2012.
- [8] Bresson, J., *Representation et Manipulation de Données d'Analyse sonore pour la Composition Musicale ; the perceptual organization of sound*. Rapport de stage IRCAM – Ecole Supérieure en Sciences Informatique – Université de Nice – Sophia Antipolis, 2003.
- [9] Grey, J. M. « Multidimensional Perceptual Scaling of Musical Timbre » in: *Journal of Acoustical Society America*, 1977.
- [10] Xenakis, I. *Formalized Music*. Pendragon Press, Hillsdale, NY, 1992.
- [11] Bokesoy, S. "Feedback Implementation within a Complex Event Generation System for Synthesizing Sonic Structures", *Proc. of Digital Audio Effects (DAFX'06)*, Montreal, Canada, pp. 199-203., 2006.
- [12] Bokesoy, S. « Presenting CosmosF as a Case Study of Audio Application Design in Openframeworks » in : *Proceedings of the International Computer Music Conference ICMC'12*, Lubiana, pp. 197-203, 2012.
- [13] Bregman, Albert. *Auditory Scene Analysis ; the perceptual organization of sound.*, Cambridge, MA, USA : MIT Press, 1990.
- [14] Schaeffer, Pierre. *Traité des objets musicaux*. Paris: Seuil, 1977.
- [15] Wishart, Trevor. *On Sonic Art*, ISBN 371865461, 1987.
- [16] Coduys, T., and Ferry, G. 'IanniX aesthetical/symbolic visualisations for hypermedia composition', *Proc. International Conference Sound and Music Computing*. Paris, 2004.
- [17] Wiener, Norbert. *Cybernetics (Control and Communication in the Animal and the Machine)*, MIT Press, Cambridge MA, 1948, 1961.



## **ZOOPHONIE : NOTES SUR UNE INSTALLATION SONORE ET SES DEVELOPPEMENTS.**

*Guillaume Loizillon*  
Université Paris 8  
guillaume.loizillon@univ-  
paris8.fr

### **RÉSUMÉ**

Cet article présente les différents aspects de la conception, de la réalisation et du développement d'une installation sonore conçue pour des lieux extérieurs paysagés ou laissés en friche.

Cette installation se déploie selon deux modalités.

- Celle d'un patch Max msp global gérant une composition sonore automatisée et sa multi diffusion.
- Celle de *boîtes* autonomes, petites constructions sonnantes assemblées autour d'objets d'usage quotidien.

### **1. INTRODUCTION**

Zoophonie est une installation destinée à prendre place dans des environnements extérieurs : parcs, jardins publics, espaces naturels ou urbains. Les supports de diffusion qui la composent, haut-parleurs ou *boîtes*, prennent place dans les arbres, les buissons, les massifs, les voies de promenades ou le mobilier urbain. Une version plus resserrée consiste à travailler sur un unique arbre, de grande taille. Le dispositif de l'installation se structure autour de deux grandes familles d'éléments :

- Un système de haut-parleurs de diverses factures en connexion avec un ordinateur, une carte audio multicanaux, et un programme modulaire écrit avec le logiciel Max msp.
- Des *boîtes* autonomes, construites autour d'objets variés, cabanes à oiseaux, cages, paniers ou contenants de diverses

factures. Dans chaque objet est installé un petit système de diffusion sonore à base de lecteur mp3 et de haut-parleurs amplifiés, tels ceux destinés aux ordinateurs personnels. Ces objets prennent place de manière visible dans l'installation. Ils permettent d'entretenir un rapport causal imagé au déploiement des sons.

Le jeu sur la figuration constitue l'élément clé de l'installation. Il fait passerelle entre un travail sonore qui travaille la question du mimétisme et ses frontières et sa contextualisation environnementale. Il est ainsi envisagé une finalisation modulaire de l'installation, uniquement composée de ces *boîtes*. Le traitement de l'espace ne s'envisage ainsi pas uniquement d'un point de vue topologique et vectoriel. La (ré)introduction d'éléments sonores plausiblement pensables dans un lieu donné peut se voir comme une pratique d'assemblage à *même le réel* qui ouvre vers des formulations de récits et de fictions.

### **2. REALISATION DU PROJET**

Après différentes approches, *Zoophonie* a été installée lors d'une manifestation culturelle organisée par la Ville d'Aubervilliers, en Seine Saint Denis, au cours de l'été 2012. Le placement s'est effectué dans le *Square Aimé Césaire*, en bordure du canal Saint Denis, dans la proche banlieue parisienne. Le projet du *jardin zoophonique* était de faire entendre aux promeneurs une faune imaginaire qui



distinctes, et dont l'une est elle même modulée. La fréquence de la porteuse est également modulable par une enveloppe qui sert aussi à donner un profil dynamique à l'événement sonore. On peut reconnaître ici un mode de construction se rapprochant du principe d'*algorithme*, tel que le DX7 de la firme Yamaha le proposait dans les années 80. L'enveloppe peut être déclenchée manuellement ou bien avec un métronome aléatoire : le tempo nominal est affecté d'un coefficient de déviation créant une nouvelle valeur à chaque battement

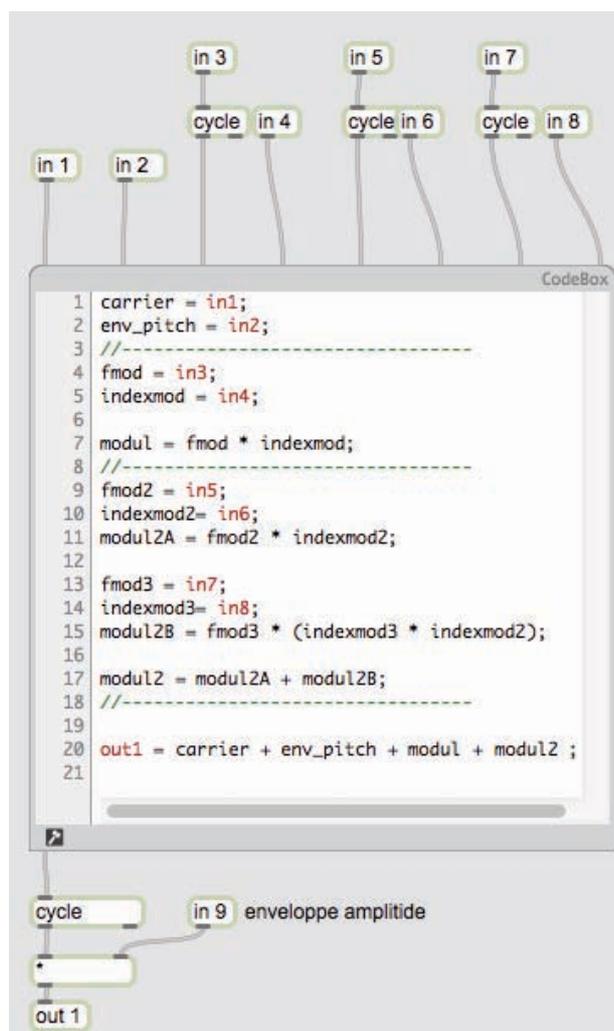


Figure 2 : moteur de synthèse FM (GEN)

Le noyau du module, proprement chargé d'effectuer la synthèse FM, utilise le sous-ensemble de programmation GEN, disponible depuis la version 6 de Max msp. GEN offre la possibilité d'accéder à un niveau bas de

programmation et d'écrire avec précision des opérations sur le signal. Il s'agit ici de préciser, avec le langage interne, le schéma des enchaînements de modulation. La sortie de la *codebox* est ensuite directement connectée à l'oscillateur porteur.

D'autres de ces modules générateurs, mis au point pour différentes évocations animalières ou naturelles, explorent des configurations bâties selon des schémas différents, appliquant des modulations à une plusieurs porteuses utilisant des formes d'ondes autres que sinusoïdales. Ces modules ont d'abord été écrits selon le système traditionnel de construction des patches dans Max msp. (Objets connectés entre eux sur un mode graphique). L'ensemble de ces éléments de synthèse est en cours de réécriture avec GEN.

### 3.2. Filtrage spectral dynamique.

Avec cet élément, on entre en contact avec des typologies sonores plus abstraites, n'évoquant pas de manière manifeste des cris ou des chants d'animaux. Le filtrage dynamique sur une grande quantité de bandes mobiles permet des rendus sonores granuleux et mobiles pouvant figurer des envols, des déplacements de masse ou encore des images sonores de nuées ou de brouillard.

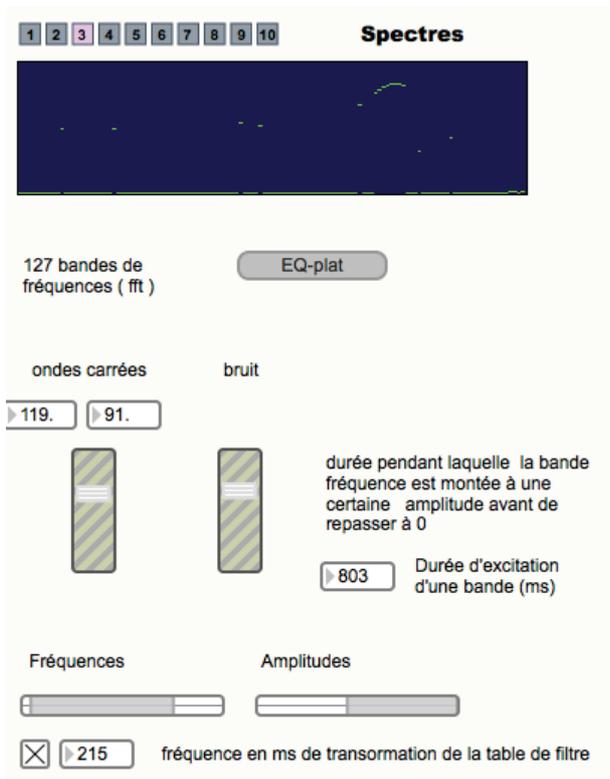


Figure 3 : interface de module de filtre spectral.

Ce module spectral utilise un filtre à 127 bandes construit autour d'une FFT. Le filtre s'applique à deux formes d'ondes rectangulaires et à un bruit blanc (canal gauche, canal droit). L'automate fonctionne cadencé par une horloge qui détermine le choix aléatoire d'une bande de fréquences à amplifier. La bande est sélectionnée sur l'ensemble du spectre ou dans un ambitus déterminé. Elle est alors amenée à ressortir un temps donné puis à s'atténuer dans un second.

L'action en temps réel est également possible en agissant directement sur la graphique du filtre.

### 3.3. Assemblage du dispositif.

Chaque module constitutif de *Zoophonie* permet de mémoriser l'ajustement spécifique de l'ensemble des paramètres. Pour la première réalisation publique de l'installation il a été stocké environ une dizaine de réglages pour chacun des modules, ainsi que mis en place un système de fondu d'enchaîné pour une

transition souple entre chacun des réglages. Les modules sonores ont été ensuite regroupés dans un patch général, qui permet un déroulement automatique ainsi que l'intervention en temps réel. Cette construction modulaire alimente un système de huit haut-parleurs ou zones de diffusion.

## 4. LES BOÎTES AUTONOMES.

Cet aspect du travail de création est lié à l'installation qu'il complète. Il constitue aussi un point de départ vers d'autres types de projets et de réalisations. Partant des modules de synthèse de l'installation ou encore, explorant d'autres modes de synthèse sonore, des séquences d'une dizaine de minutes, ou plus, sont composées et chargées dans des lecteurs mp3 de poche. Ceux-ci sont ensuite connectés à un haut-parleur amplifié et insérés à l'intérieur des différents objets dont la forme actualise une fonction de micro récit. Ainsi, la dimension sonnante, alliée à l'objet déterminé dans un usage habituel, constituent un montage concret, comme ont pu l'explorer les artistes de la poésie concrète. Dans cette perspective de l'objet sensible comme langage, l'assemblage à pour objectif de développer un imaginaire de la représentation des sons qui s'articulerait autour de la question de la fiction.

Les *boîtes* sont destinées à se placer dans différents points de l'installation. Elles sont visibles, devinées ou incitées à être découvertes. Dans le cadre strict du projet *Zoophonie*, ces objets sont majoritairement des petites maison d'oiseaux ou des cages, suspendus dans des arbres. Le développement de ces *objets pièges* constitue un élément central pour l'évolution de cette installation et de son extension vers d'autres domaines de figurations sonores. Le projet consiste à créer une collection importante de *boîtes*, allant jusqu'à 20 et peut-être au-delà. Un développement technique plus avancé conduit à une réflexion plus poussée sur le mode d'alimentation en sons des *boîtes*. Sans nécessairement abandonner l'idée d'un fichier son, fixé et intégré à l'objet, il est possible de

penser une système, via liaison HF, permettant de faire entendre du son provenant directement des modules de synthèse. La difficulté de cette proposition augmente si il s'agit d'adresser un grand nombre de boîtes, démultipliant ainsi les fréquences H.F.

#### 4.1. Lo-fi et espace sonore.

L'introduction de microsystèmes audio dans des objets, conduit de manière quasi obligée à l'usage de sons aux fréquences distribuées dans le milieu et le haut du spectre. Les *boîtes* de *Zoophonie* adoptent des systèmes de haut-parleurs amplifiés, du type de ceux qui sont destinés à être connectés à la sortie d'un ordinateur portable ou d'un téléphone mobile pour en améliorer l'écoute. Ces systèmes n'atteignent que très exceptionnellement ce qu'il est convenu d'appeler la haute fidélité (Hi-Fi)<sup>3</sup>. On sait que les termes de Hi-Fi et Lo-Fi visent aussi à quantifier l'impression générale de la qualité d'écoute de l'environnement, indépendamment des appareils. Dans cette acception, développée par R.Murray Schaffer dans les années 70, le passage de la Hi-fi à la Lo-fi cartographie l'espace sonore selon l'importance de la saturation qui y règne et la capacité d'une plus ou moins bonne discrimination des sons par l'oreille. La limitation technique des performances sonores des petits appareils de l'audio portable est intéressante à contourner et à exploiter selon cette perspective environnementale. La démultiplication de ces objets, dispersés dans un lieu hétérogène et sans contour spatial précis permet de bâtir une forme de spatialisation diffuse et englobante. Ainsi, la question de la Lo-fi, devient constitutive de la forme de spatialisation spécifique des *objets sonnants*. Compte tenu des zones spectrales explorées et des effets mimétiques qu'ils activent, même

<sup>3</sup> La Hi-Fi désigne originellement une tentative de normalisation de la qualité de reproduction sonore des appareils par les industriels. L'objectif mythique est d'obtenir une qualité de reproduction la plus proche possible de la source. Cette idée de norme est de nos jours largement dépassée

s'ajoutant à un réel sonore parfois déjà assez dense, les sons et les objets de *zoophonie* sont destinés à agir comme agents d'introduction de précisions narratives.

L'utilisation de telles typologies de timbres et le caractère isolé des sources favorisent des localisations précises. La mise scène de ces sons par les *boîtes*, qu'elles soient figuratives ou simplement volumes géométriques, invite le spectateur au déplacement et contribue à une dynamique de la spatialisation construite sur la multiplicité des points d'écoute possibles. L'espace peut aussi se rendre manifeste sur le plan horizontal et sur le plan vertical. L'installation cantonnée aux limites d'un arbre ou d'une façade, par exemple, permet d'expérimenter cette sensation d'une profondeur verticale



Figure 4 : Un baladeur mp3 connecté à un haut-parleur de surface : à intégrer dans une *boîte*.

Sur le terrain de la composition électroacoustique, il demeure évident que l'utilisation de tels systèmes de diffusion produit des contraintes sur la nature des choix sonores et sur le travail de composition. Outre la bande passante limitée, la linéarité des haut-parleurs est tout à fait relative, et l'introduction de ceux-ci dans des objets modifie encore le rendu. Cela conduit à faire le mixage et le *mastering* des séquences directement sur ce système d'écoute, voir même sur la boîte déjà

assemblée. Dans le cadre de *zoophonie*, qui fonctionne majoritairement avec des fréquences médium ou hautes, les résultats offerts par ces équipements sont en adéquation avec le sens du projet.

Cependant, l'utilisation des basses fréquences (et même medium basse) reste délicat et complexe dans ce cadre de petits *objets sonnants*. L'usage de haut-parleurs de contact, qui peuvent mettre en vibration des objets est une piste de recherche et de travail pour la confection d'autres modèles de *boîtes* d'où le son rayonnerait des surfaces mêmes des objets.

[4]Licht, Alan. *Sound art, beyond music, between categories*, Rizzoli, New-York, 2007.

#### 4.2.Conclusion provisoire.

*Zoophonie* désire rester un projet ouvert et aux devenirs multiples, que sa réalisation s'oriente vers l'usage d'un système intégré de modules de synthèses et diffusion multicanaux ou vers la dispersion de petits objets autonomes. L'évocation et l'imagerie animalière ne sont alors que des points de départ d'une fiction possible. Ce rapport au réel peut aisément déborder ce seul cadre de départ afin de s'orienter dans une conception plus élargie de l'imaginaire sonore, quand celui-ci se relie à un environnement public pour y construire la déambulation.

Sur le terrain de la forme et de l'apparence, on peut également penser à des développements spécifiques pour ces objets dont l'une consisterait, outre l'utilisation de *Ready-mades*, à mettre en place une collaboration, pour la construction de ceux-ci, avec des artistes plasticiens intéressés par ce projet.

### 5. REFERENCES

- [1]Duchamp, Marcel. *Duchamp du signe*, Champs Flammarion, Paris, 1994
- [2]Donguy, Jacques, *Une génération. Poésie concrète, poésie sonore, poésie visuelle*, Henri Veyrier Paris, 1985
- [3]Schaffer, R Murray. *Le paysage sonore*, JC Lattès, Paris 1979.

# MUSIQUES VISUELLES, COMPOSITION (MUSICALE) AUDIOVISUELLE

*Marc Billon*  
Université Paris 8  
marcbillon@club-internet.fr

*Anne Sèdes*  
CICM-EA 1572  
Université Paris 8  
MSH Paris Nord  
anne.sedes@univ-paris8.fr

## RÉSUMÉ

Ce texte présente le projet *pulse plus pixel*, composition musicale audiovisuelle de Marc Billon. Ayant situé l'origine des musiques visuelles dans l'abstraction picturale du début du siècle dernier et déclaré un intérêt spécifique pour le courant du cinétisme optique, on montre comment on va faire converger musicalement le son et l'image en s'appuyant sur les moyens de l'écriture polyrythmique audiovisuelle. On décrit les moyens et méthodes employés par le compositeur. Des extraits audiovisuels en ligne illustrent le propos au fil de l'article. La pièce pourra être diffusée dans le cadre des JIM 2013.

## 1. INTRODUCTION

Avec l'avènement massif du numérique et son potentiel de convergence dans l'écriture intermédiaire, les musiques visuelles, longtemps marginalisées du côté du cinéma expérimental au cours du vingtième siècle, suscitent aujourd'hui un nouvel intérêt. Après avoir rappelé quelques points historiques pris en héritage et évoqué la question de la convergence musicale entre media, dont la pratique des musiques mixtes nous a fourni une première approche, on va se placer du point de vue du compositeur pour aborder les manières de faire de la composition musicale audiovisuelle, à travers le projet *pulse plus pixel* de Marc Billon, qui sera montré dans son intégralité dans le cadre des JIM 2013<sup>1</sup>.

## 2. MUSIQUES VISUELLES, HERITAGE, CINÉTISME OPTIQUE

### 2.1 Héritage de l'abstraction picturale

Les musiques visuelles trouvent leurs sources aux origines de l'abstraction picturale [1]. Pour exemple, on citera entre bien d'autres toiles *Amorpha : fugue à deux couleurs* de Frantisek Kupka (1912), ou *Fugue en rouge* (1921) ou *En rythme* (1930), de Paul Klee. Dès 1913, Léopold Survage portait un projet cinématographique à partir de ses *Rythmes colorés*. Les travaux des cinéastes Oskar Fischinger, Viking Eggeling, Walter Ruttmann, Stefan et Franciszka Themerson, Mary Ellen Bute, les

<sup>1</sup> L'ensemble des exemples vidéo présentés dans cet article est disponible en ligne sur Vimeo. La qualité des documents audiovisuels est évidemment assez modeste comparée à la projection originale en salle. Les URL sont à copier dans votre navigateur.

frères Withney ont également contribué à cette pratique artistique.

Avec le développement des moyens numériques, alliant image et son, on assiste dans les années 2000 à une effervescence nouvelle des musiques visuelles (Ryoji Ikeda, Carsten Nicolai, etc). Dans le domaine académique, la revue *Organized Sound* a récemment consacré un de ses numéros au sujet [2], ce qu'avait également fait le *Computer Music Journal* dès 2005 [3]. Rappelons également les travaux du Groupe de travail de l'AFIM sur la visualisation du son entre 2006 et 2008[4].

### 2.2 Mouvement, cinétisme optique

Pour le travail qui suit, on va s'intéresser spécialement au cinétisme optique. Ce courant émerge en avril 1955, avec l'exposition *Le mouvement*, galerie Denise René, présentant Duchamp, Vasarely, Soto, Calder, Agam, Jacobsen, Tinguely, Bury. Lors de cette exposition, un film de Robert Breer et de Pontus Hulten sera réalisé. A cette occasion, le "manifeste jaune" sera édité, posant que le nouveau medium associe désormais les dimensions du mouvement et du temps. Le cinéma expérimental y est convoqué, avec pour référence les recherches de Norman Mac Laren, Len Lye, Oskar Fischinger, V. Eggeling, Laszlo Moholy-Nagy, Hans Richter, etc...

L'optocinétisme se singularise par sa recherche d'instabilité perceptive au niveau visuel, à travers l'étude du mouvement en faisant usage de figures alternant fréquemment le noir et blanc.

*"Does not optics, even in the form of illusion, belong to Kinetics ? Does not aggressing the retina in fact make it vibrate ? Now, the maximum contrast is black and white. How fertile the experience of phototechnique had been to me ! Negatives-positives, transparencies, photograms, the hallucinatory play of black and white. Here was the starting-point of my discovery : the same composition resolved in white and black automatically gave me a second composition resolved in black and white."* [5]

L'usage du noir et blanc, avec des alternances de lignes blanches et de lignes noires, facilite les illusions

optiques, bien connues sur le plan scientifique (figure radiale de Mackay). [6] Ces illusions sont également présentes dans le cinématisme optique et dans l'*op art*. On citera par exemple les œuvres de Vasarely "*Beta*", "*Gamma*" (1958- 1965), "*Bora II*" 1964, Bridget Riley "*Fall*" (1963), "*Cercle d'accélération*" (1961), Yvaral "*Imprégnation rétinienne n° 2573*" (1965), Soto : "*Spirale*" (1955).

On a ainsi le sentiment que l'œil devient moteur. "*Il y a là l'intuition majeure, suivie par de nombreux représentants de l'optocinétisme, d'une corrélation étroite entre la vitesse de la perception et la vivacité de la réponse motrice, entre l'exacerbation de la stimulation visuelle et la puissance de la décharge kinesthésique.*" [7]

L'œuvre optocinétique est ainsi appréciée en fonction de sa réponse esthétique immédiate, ce qui la rend très proche de la perception musicale.

### 3. CONVERGENCE MUSICALE ENTRE MEDIA

Dans un texte développant les perspectives de l'électroacoustique à l'heure de la numérisation, Horacio Vaggione évoquait à propos de la mixité musicale la convergence entre moyens acoustiques et électroacoustiques.

« *Il est souhaitable de chercher des passerelles très fines afin de faire interagir les deux sources d'une façon très rapprochée, au niveau du résultat sonore, mais aussi au niveau du processus de composition lui-même, en travaillant à partir de la même situation musicale, c'est à dire en assumant le postulat d'une possible convergence entre les mondes instrumental et électroacoustique, en créant une vectorisation commune.* "[8]

Cette "vectorisation commune" serait basée sur le maniement de la multiplicité des échelles temporelles offert par les moyens compositionnels à disposition et rendu accessible grâce aux moyens du numérique.

Du point de vue du compositeur, la convergence concerne finalement la construction de la "musicalité", si l'on peut dire, ce qui fait musique, ce qui peut converger musicalement dans l'intermédialité.

A l'heure du numérique, électroacoustique et vidéo convergent. En ce sens, un environnement de programmation tel que MAX et sa bibliothèque Jitter, proposant aussi bien l'audio numérique que la vidéo numérique, est révélateur. On pourrait évoquer une extension intermédiaire de la mixité musicale vers les musiques visuelles<sup>2</sup>. Des questions émergent alors:

comment construit-on la convergence musicale entre divers media ? Qu'est-ce qui est musical, pour ainsi dire "musicable" dans le domaine de l'image ?

On va tenter d'esquisser quelques réponses en présentant l'approche de Marc Billon, pour son projet de composition audiovisuelle *pulse plus pixel*.

### 4. RYTHME SONORE, RYTHME VISUEL, POLYRYTHMIE AUDIOVISUELLE

Si l'on compose la musique avec les moyens audio numérique et vidéo numérique, la convergence va pouvoir se construire à partir d'opérations à divers niveaux temporels. En ce sens, une approche du rythme appliquée à la composition audiovisuelle peut être intéressante. Rappelons qu'au niveau rythmique, tout comme pour le phénomène sonore, la microtemporalité offerte par les techniques numériques nous permet de descendre bien en dessous des échelles musicales habituelles héritées du jeu instrumental traditionnel. Sur le seul plan des musiques électroacoustiques, l'exploration de la dimension rythmique offre encore d'immenses perspectives. Le rythme n'est d'ailleurs pas seulement musical, c'est également un moyen pour produire du mouvement visuel, comme les arts plastiques nous l'ont montré avec par exemple Julio Le Parc, "*Séquence quantitative*", toile, 1959-1971, ou bien avec les cinéastes expérimentaux. On citera Paul Sharits, *Ray Gun Virus* réalisé à partir de photogrammes peints (1966), Tony Conrad, *the Flicker* (1966), Peter Kubelka, *Arnulf Rainer*, 1958-1960, avec pour tous ces films l'usage du *Flicker* (clignotement).

La composition *pulse plus pixel* vise à la construction d'une organisation musicale où coexistent différentes temporalités sonores et visuelles dans une recherche d'instabilité perceptive.

Exemple vidéo : *Instabilité 1*  
<https://vimeo.com/60022339>

Exemple vidéo : *Instabilité 2*  
<https://vimeo.com/60041227>

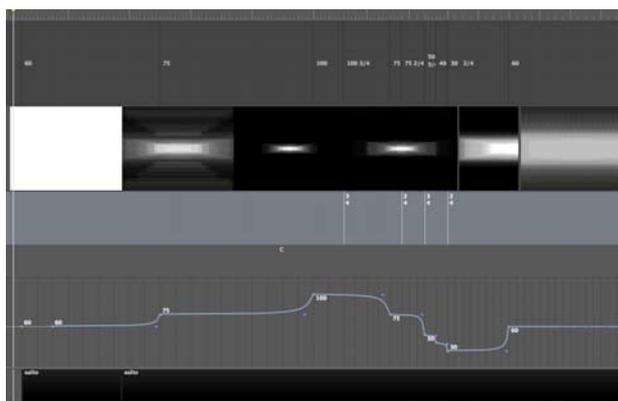
Dans cette pièce, le rythme sonore se construit par la superposition de figures de longueurs et de vitesses différentes. Les opérations traditionnelles telles que l'amplification, la diminution, le monnayage sont utilisées. Le polyrythme est obtenu en superposant des cellules de différentes pulsations.

Exemple vidéo : *Cellule polyrythmique*  
<https://vimeo.com/60021615>

<sup>2</sup> Sèdes, A., "Musique visuelle, extension de la mixité", conférence donnée dans le cadre des journées d'étude sur *l'extension intermedia et la pérennité dans le cadre des musiques mixtes : image, improvisation, préservation et "recasting"*, CITAR, Université

Catholique Portugaise, Porto, 21 avril 2012. Programme "musiques mixtes" de la MSH Paris Nord.

Afin d'obtenir des effets d'accélération et de ralentissements, il est fait usage de la modulation de tempo.



**Figure 1.** Modulation de tempo dans le logiciel de montage son.

Exemple vidéo : *Esquisse 1*  
<https://vimeo.com/60021614>

Le rythme visuel est généré de deux manières. Il peut être assujéti aux rythmes sonores dans Max lors de la création de patterns audiovisuels, ou bien construit dans le logiciel de montage image.

Exemple vidéo : *rythmepattern*  
<https://vimeo.com/60021892>

Il est possible alors d'intervenir à l'image près, ce qui offre des perspectives intéressantes de création rythmique. La relation s'établit sur la durée d'une séquence d'images en mouvement qui constitue un plan visuel et du rythme sonore associé sur un segment temporel donné. Chaque impulsion sonore possède la durée d'une *Frame* soit 1/30 de seconde considéré comme valeur isochrone de la cellule rythmique.<sup>3</sup>

Exemple vidéo : *imageparimage 1*  
<https://vimeo.com/60021616>

Cependant le temps de perception de l'image n'est pas identique au temps de perception du son. Pour construire le rythme visuel, il faut tenir compte de ces différences pour permettre un certain confort perceptif ; sans quoi les figures rythmiques de l'image sont difficilement identifiées au niveau microtemporel et donnent une impression de clignotement continu lorsque le rythme visuel occupe l'ensemble du plan de vision. Une option consiste à inclure les rythmes visuels microtemporels dans un plan fixe en suivant une trajectoire, ascendante ou descendante comme dans l'exemple qui suit.

<sup>3</sup> Le format de travail était fixé à 30 fps pour des questions de compatibilité entre les logiciels utilisés.

Exemple vidéo : *imageparimage 2*  
<https://vimeo.com/60021617>

La logique du mouvement peut ainsi être anticipée. Dans ce travail les effets de jaillissements restent dévolus à des fins d'articulations rythmiques. Ces deux manières complémentaires d'aborder le rythme visuel donnent une certaine liberté créative afin d'échapper au parallélisme stricte et constant de l'image et du son. La possibilité d'une polyrythmie audiovisuelle est soudain à portée de main, des exemples sont en cours de réalisation, pour le projet *pulse plus pixel*. On assume également des propositions hétérythmiques où l'œil et l'oreille reconstruisent le rythme.

Exemple vidéo : *imageparimage 1*  
<https://vimeo.com/60021616>

## 5. MOYENS ET METHODES POUR *PULSE PLUS PIXEL*

### 5.1 Les outils numériques employés

L'approche compositionnelle utilise ici les moyens du temps réel pour des techniques de traitement, de transformation et de mixage de l'image, de contrôle de l'image par le son, mais aussi les moyens du temps différé, qui permettent des techniques de traitement du son et de montage image / son comparables aux techniques de composition habituelles en studio électroacoustique. Le montage étant finalement un moyen de construction du rythme à divers niveaux temporels.

L'environnement de programmation Max/Jitter est utilisé comme un générateur audiovisuel qui permet de constituer un catalogue de sons et d'images fixes ou animées.



**Figures 2, 3, 4.** Exemple d'images fixes créées dans Jitter

La station audionumérique de type *DAW* est dédiée au montage et aux traitements du son avec l'option de visionner les séquences vidéo déjà réalisées. Un logiciel de montage image / son permet un travail à l'image près sur les séquences vidéo et un travail sur les pistes audio en parallèle. L'échange de fichiers entre les plates-formes se fait aisément.

Ces trois outils logiciels constituent le *set-up* complet pour la production et l'édition de séquences audiovisuelles.

## 5.2 Génération et transformation du son

Les trains d'impulsions sont créés dans Max, ou dans la station audionumérique par l'utilisation de *plug-in* standard.

Chaque impulsion se situe dans un intervalle temporel variable compris entre 10 à 33,333 ms de manière à correspondre au temps de l'image en mouvement au sein du film (30 images /s), du fait du format de travail utilisé.

Ces impulsions sont regroupées dans plusieurs catalogues différents en fonction de leurs caractéristiques spectrales. Lors de la création de ces banques sonores, une attention particulière est portée aux bandes de fréquences utilisées de manière à faciliter la mise en relation avec l'image, lors de la construction du *mapping*. L'espace fréquentiel est segmenté sur trois bandes étroites dans l'extrême grave, le médium et l'extrême aigu.

Les transformations sonores consistent au niveau spectral à effectuer des corrections d'égalisation, utiliser la modulation en anneau, ainsi que des filtres passe-bande ou en peigne. Au niveau dynamique, le signal est traité par de la compression/ limitation/*gate* ; le niveau spatial quand à lui fait usage d'unités de réverbération dédiées aux traitements de l'image stéréophonique.

Dans ce travail, la génération du son précède systématiquement la création des visuels. Partant du sonore vers le visuel, la création d'images incite à retravailler la forme musicale sur le plan de la cohérence rythmique, ce qui revient à produire un rythme audiovisuel. Finalement, il y a un échange permanent entre les deux approches temps réel ou temps différé.

## 5.3 Transformations de l'image

Les transformations sur l'image sont obtenues essentiellement par l'utilisation de la bibliothèque de *shaders* dans Jitter. En effet ces scripts<sup>4</sup> fournissent une batterie de filtres proposant des traitements variés : facteur d'agrandissement, réallocation spatiale du plan... Ils assurent aussi un contrôle sur la lumière, les ombrages, la luminance ou radiance. Leur utilisation permet de recourir aux ressources GPU de la carte graphique et de réaliser ainsi des économies sur le travail du CPU.

L'usage des scripts de *shaders* correspond à la dernière étape de traitements opérés sur les visuels avant le montage final.

Exemple vidéo : *shadertransformation*  
<https://vimeo.com/60021895>

<sup>4</sup> fichiers *.jxs* présents dans le dossier "Jitter shaders" de l'application.

## 5.4 Construction de l'image en mouvement - Comment le son génère l'image

- *image par image*

La mise en relation image/son dans le cas d'une construction image par image repose sur le rapport créé entre le mouvement de la figure du plan et le rythme sonore perçu. Il ne s'agit pas dans ce cas d'une translation terme à terme, mais plutôt d'obtenir une trajectoire visuelle structurée et en relation avec le rythme sonore. On peut alors composer un rythme audiovisuel proposant une variété de visuels en lien avec la structure rythmique de la musique.

Exemple vidéo : *imageparimage 1*  
<https://vimeo.com/60021616>

- *réalisation de séquences audiovisuelles dans Max*  
 L'environnement Max/Jitter permet la lecture et la transformation de l'image en temps réel à partir de données caractéristiques du son. La mesure de l'amplitude du signal sonore contrôlera par exemple le facteur d'agrandissement de l'image. L'usage des descripteurs des caractéristiques psychoacoustiques du son (*loudness, noisiness, brightness, centroïde,...*) permet de connecter ces articulations dynamiques à certaines propriétés de l'image comme l'intensité lumineuse, la saturation, la luminance (exemple vidéo *mapping*) ou encore certaines variables des filtres offerts par Jitter (exemple vidéo *esquisse1*).

Exemple vidéo : *mapping*  
<https://vimeo.com/60024755>

Exemple vidéo : *esquisse1*  
<https://vimeo.com/60021614>

Le *mapping* des paramètres variables du son et de l'image constitue ainsi une façon simple mais efficace de *transduire* l'énergie sonore au visuel. Toutes ces séquences audiovisuelles sont porteuses des premières structures rythmiques de l'image et du son avant transformations, montage et mixage.

## 6. CONCLUSION

Ce travail expérimental est un *work in progress*. L'usage d'une écriture polyrythmique indique à lui seul l'approche musicale.

On remarquera également que toutes les opérations compositionnelles décrites peuvent se comparer de très près avec les opérations habituelles de la composition musicale électroacoustique en studio: approche multi-échelle, transformations, montage, mixage, développement de banques de sons et de banques d'image. On pourrait presque dire que l'on manipule des objets aux propriétés comparables, selon qu'il s'agit d'objets développés en milieu sonore ou en milieu visuel, grâce aux moyens numériques.

A un autre niveau, on compose ici une forme de convergence basée sur le musical, au moyen d'une polyrythmie audiovisuelle.

On vise finalement l'expérience musicale, encore, là où la perception des multiples échelles du temps s'incarne dans l'intersensorialité de l'écoute musicale, convoquant ici aussi bien la vision que l'audition.

## 7. REFERENCES

- [1] Rousseau, P. « *Arabesques*, le formalisme musical dans le début de l'abstraction », Catalogue de l'exposition *Aux origines de l'abstraction 1800-1914*, Musée d'Orsay, réunion des musées nationaux, Paris, 2003, pp. 230-245.
- [2] Organized Sound, « Composing Motion : A visual music retrospective », Cambridge, vol 17, issue 2, aout 2012.
- [3] Computer Music Journal, « visuel music » vol. 29, n°4, Winter 2005, MIT Press.
- [4] Sedes, A. « Rapport du groupe de travail AFIM sur la visualisation du son », Paris 2008. [www.afim-asso.org/IMG/pdf/rapportGTVvisualisation.pdf](http://www.afim-asso.org/IMG/pdf/rapportGTVvisualisation.pdf)
- [5] Vasarely, V. : « White and black », Catalogue de l'exposition *Le mouvement*, édition Denis René. Paris 1955, p. 42.
- [6] Gregory, R. L. « L'œil et le cerveau », De Boeck Université, collection neurosciences et cognition, 2000, p.254.
- [7] Pierre, A., « Accélération optiques », Catalogue de l'exposition *L'œil Moteur. Art optique et cinétique, 1950-1975*, édition des musées de Strasbourg, 2005, p. 34.
- [8] Vaggione, H., « Perspectives de l'électro-acoustique », in revue *Chimère* n° 40, Paris, 2002, pp. 57-67. Disponible en ligne.



## AVANT-PREMIÈRE DE « LA GALERIE D'ONDES »:

Une installation sonore et visuelle pour explorer l'intersensorialité des infrasons

Alexis Story CRAWSHAW  
CICM – EA 1572

Université Paris 8, MSH Paris Nord

storyalexis@yahoo.com

Robert H. LAMP  
Monterey Peninsula College

newmountainstudio@hotmail.com

Nicolas FDIDA  
ONERA

nicolas.fdida@onera.fr

### RÉSUMÉ

Cette installation collaborative entre compositeur et artiste cherche à explorer le potentiel expressif de la musique infrasonore sur un plan intersensoriel, à travers des terrains sonore, visuel et parfois vibrotactile. L'installation est caractérisée par la musique infrasonore qui fonctionne en tant que sculpture sonore pour créer des ondes stationnaires. Les sculptures visuelles, sensibles à ces ondes acoustiques, créeront une manifestation intersensorielle en interaction avec ces infrasons. Également, le positionnement initial des objets visuels dans la salle ainsi que la possibilité pour les spectateurs d'en déplacer certains encourageront le public à explorer les points d'écoute et l'influence intersensorielle divers selon l'emplacement des maxima et minima de l'amplitude des ondes dans l'espace.

### 1. INTRODUCTION

Ce projet d'installation sonore et visuelle est un effort collaboratif entre la compositrice Alexis Crawshaw, doctorante à l'Université de Paris 8, l'artiste plasticien Bob Lamp, instructeur au Monterey Peninsula College et le physicien Nicolas Fdida, Docteur en Physique à ONERA. Il bénéficie du soutien du Conseil Scientifique de la Maison des Sciences de l'Homme (MSH) Paris Nord.

Cette maquette présentée dans le cadre des JIM 2013 (Hall d'exposition de l'université de Paris 8 du 13 au 15 mai 2013) est la maquette en avant-première pour une installation d'une plus grande échelle dont le titre provisoire est « La Galerie d'Ondes ».

Le projet de « La Galerie d'Ondes » sert à développer l'aspect pratique de la méthodologie pour la thèse d'Alexis Crawshaw, « Les musiques électroacoustiques infrasonores, extrêmement graves et intersensorielles: approche théorique et pratique ». Dans cette thèse, Crawshaw émet l'hypothèse principale qu'on peut développer la musique infrasonore sur un plan intersensoriel, afin de faire connaître les propriétés musicales des infrasons. Également, elle suppose qu'on peut développer une expressivité de cette musique. Crawshaw prend une approche de recherche et création qui s'articule autour d'un développement de la théorie et d'une démonstration par la pratique. Ainsi, ce projet se place au sein de cette démonstration artistique.

Étant donné la nature intersensorielle des infrasons et leurs impacts sur des structures architecturales et le corps humain, on constate qu'il est possible de développer ces caractéristiques sur un terrain artistique. À l'aide d'objets visuels issus de médias différents (en dessin et en sculpture, ayant des aspects mécaniques, vibrotactiles, ou de la nature cymatique), on cherche diverses manières de manifester les infrasons au domaine visuel et, aussi, comment les infrasons peuvent faire répondre ces œuvres visuelles sur les plans sonore et vibrotactile. Concernant ce dernier point des réponses, on cherche à créer des vibrations sympathiques, à activer des résonances et à jouer avec le contrepoint intersensoriel.

Au niveau des contributions possibles à la communauté artistique, cette collaboration interdisciplinaire sert à aider à l'écoute des sons extrêmement bas. Elle est une illustration intersensorielle de la présence des infrasons ainsi qu'une manifestation intersensorielle de leurs propriétés musicales.

L'intérêt de sa réalisation se pose également dans le développement d'un cadre d'œuvres musicales qui exploitent les infrasons et leurs propriétés intersensorielles.

### 2. L'INFRASON

#### 2.1. Définition

Dans le cadre de ce projet, on définit les infrasons en tant que des oscillations acoustiques qui sont inférieures ou égales à 20 Hz (le seuil entre la hauteur et le rythme étant entre 16 et 20 Hz), qui souvent se manifestent perceptivement par des caractéristiques vibrotactiles [4].

On perçoit les infrasons par leur mélange sensoriel unique (l'audition, la pression et le toucher). Leur influence sur le toucher est due particulièrement à la catégorie vibrotactile (la perception de vibration), parfois à la proprioception (la perception de l'orientation du corps dans l'espace) et parfois aux sens viscéraux (la perception interne des viscères). La vision peut aussi être affectée par des effets infrasonores [3].

#### 2.2. Seuil de perception

La fausse idée couramment répandue est que

l'être humain ne peut pas percevoir les fréquences inférieures à 20 Hz. Même si les infrasons ne sont plus des hauteurs, on peut toujours détecter leurs oscillations rythmiques avec un niveau de pression sonore suffisamment fort. En effet, le seuil d'audition est une question de rapport entre le niveau de pression sonore et la fréquence. Pour les fréquences inférieures à 100 Hz, il faut toujours tenir compte que plus la fréquence est faible, plus le niveau de pression sonore doit être fort pour la détecter.

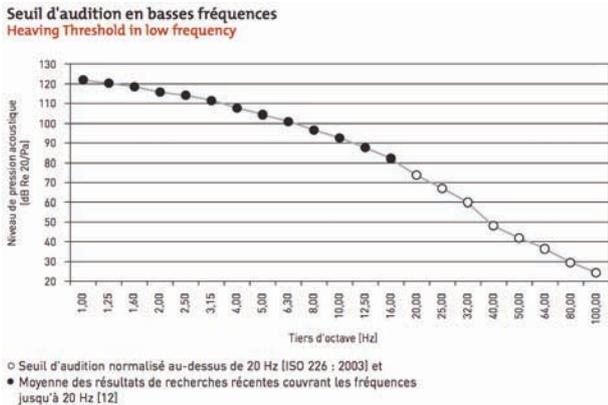


Figure 1. Seuil d'audition en basses fréquences d'après plusieurs auteurs (Chatillon, 2006) [2].

Voici un graphique du seuil de l'audition d'après les données de plusieurs auteurs, cité par Chatillon, 2006. Il spécifie que le seuil d'audition des sujets les plus sensibles se trouve à plus de 10 dB en dessous de la moyenne [2]. Il y a également des études menées qui suggèrent qu'il est possible de se sensibiliser aux infrasons [8], [12].

Un autre facteur contribuant au mythe de leur imperceptibilité est que la plupart des caissons de basse ne sont pas optimisés pour reproduire ces fréquences à l'amplitude qu'il faut pour qu'elles soient détectables par l'oreille humaine. A part dans une configuration sonore typique pour un concert électroacoustique, une projection cinématographique, un home cinéma ou un studio, on a très rarement l'occasion de les entendre diffusés dans un contexte artistique. Ainsi, pour ce projet, ou pour d'autres qui emploient des infrasons, il faut tenir compte des limitations des dispositifs de diffusion utilisés.

### 2.3. La problématique concernant les limites de perception et de diffusion

En tant que compositeur, il faut considérer des techniques de composition (particulièrement concernant la psychoacoustique), des moyens d'amplification et des moyens de diffusion alternatifs (et de modalité sensorielle) qui peuvent servir dans le cas d'une sensibilisation aux infrasons. On doit envisager ces limitations en tant que contraintes créatives et en tant qu'une invitation pour une recherche artistique.

Étant donné que le caisson de basse utilisé diffuse des infrasons entre 15-20 Hz autour du seuil

d'audition, nos objets visuels complètent ce spectre d'amplitude limitée par une approche de renforcement intersensoriel. Les réactions de ces objets aux ondes acoustiques créent une redondance des données sonores qui se manifestent aussi dans la modalité visuelle. Ainsi, l'écoute des infrasons dans cet environnement est assistée par des repères visuels, voire des œuvres visuelles qui apportent leur propre esthétique à l'ensemble de l'installation. En effet, notre projet était motivé par ce problème des limites. En présentant ce projet, on offre une solution artistique pour le traiter.

### 2.4. Sur les risques possibles des infrasons

Pour ce projet, on a pris en compte des recherches sur les impacts des infrasons sur le corps humain, particulièrement concernant les risques possibles sur la santé. Une étude rigoureuse de cette littérature a été effectuée dans le rapport de Master 1 de Crawshaw, « L'infrason en Art ». Tout d'abord, il faut remarquer qu'au niveau de l'audition, le niveau de pression sonore employé pour cette installation reste loin du seuil d'altération permanente voire temporaire de l'ouïe. En plus, on suit les conseils de Chatillon indiqués dans ce graphique :

#### Valeurs limites proposées pour l'exposition aux infrasons aériens Suggested limit values for exposure to airborne infrasound

INFRASONS CONTINUS	
Calcul de l'exposition	<ul style="list-style-type: none"> <li>Utilisation de la pondération G</li> <li>Sommation des énergies reçues entre les tiers d'octave compris entre 1 Hz et 100 Hz</li> </ul>
Valeur limite d'exposition en dB(G) sur une durée de 8 heures	102 dB(G)
Si la durée d'exposition est diminuée par un facteur 2	Augmentation de la valeur limite de + 3 dB
INFRASONS IMPULSIONNELS	
Calcul de l'exposition	Pas de pondération
Valeur limite d'exposition	145 dB(Lin)

Figure 2. Valeurs limites proposées pour l'exposition aux infrasons aériens (Chatillon, 2006) [2].

On reste en-dessous de 114 dB pour permettre à une exposition saine de 30 mn aux infrasons (ou 2 fois la durée de la composition musicale). En tout cas, notre dispositif de diffusion ne diffuse les infrasons qu'autour de leur seuil de perception.

Il y a aussi des individus qui rapportent des gênes occasionnelles et psychosomatiques des infrasons autour du seuil de perception. D'après nos recherches concernant l'état de l'art infrasonore, les études

cognitives concernant l'émotion musicale et les recherches scientifiques sur les réponses aux infrasons, il est évident que le contexte d'écoute est très important à cette réception des infrasons. On crée un environnement artistique où des spectateurs sont libres d'entrer ou de partir comme il leur plaît. Ce n'est pas une installation imposée dans un espace public très fréquenté. Également, les participants peuvent se promener dans l'espace pour accentuer certains effets et en diminuer d'autres. Ainsi, les spectateurs sont des véritables participants dans cet environnement. La volonté qu'ils exercent dans l'installation— en choisissant d'y entrer, d'y rester, de l'explorer et de rentrer en interaction avec— va fortement influencer leur attitude aux basses fréquences utilisées.

### 3. « LA GALERIE D'ONDES »

#### 3.1. Dispositifs

##### 3.1.1. Sonores

Cette avant-première utilise un caisson de basse UMS-1P de la marque Meyer Sound comme dispositif de diffusion. Il a deux membranes de 10'. Il a une réponse fréquentielle de 29 Hz – 135 Hz (à  $\pm 4$  dB) et un spectre opérationnel de 25 Hz – 160 Hz, selon ses spécifications techniques [9]. D'après des tests menés par Crawshaw, cette machine produit des résultats sonores exploitables jusqu'à 15/16 Hz, voire 12 Hz à une distance de moins d'un mètre depuis la membrane, dans une salle assez silencieuse.

La carte son est une Presonus Inspire 1394 qui n'a pas de contre-indication concernant la diffusion des fréquences en dessous de 20 Hz depuis les sorties dédiées aux haut-parleurs [11].

##### 3.1.2. Visuels

Les éléments visuels pour ce projet sont de trois variétés. Du premier type, on a 3 sculptures à membrane. Ces objets, décorés avec des dessins en Sharpie™ (des stylos-feutres), consistent en un papier avec une surface en BoPET (« Biaxially-oriented polyethylene terephthalate ») avec un substrat de BoPET et une couche fine peinte de la pulpe de papier blanc. Deux de ces objets ont une toile qui sert comme membrane, cadrée par et suspendues sur des tiges en acier. Le premier, nommé « Sensitive Dependence », a un passe-câble au milieu de la toile avec deux bâtons très fins, comme des aiguilles, fixés sur chaque côté. Des limiteurs en aciers réduisent le mouvement des aiguilles.

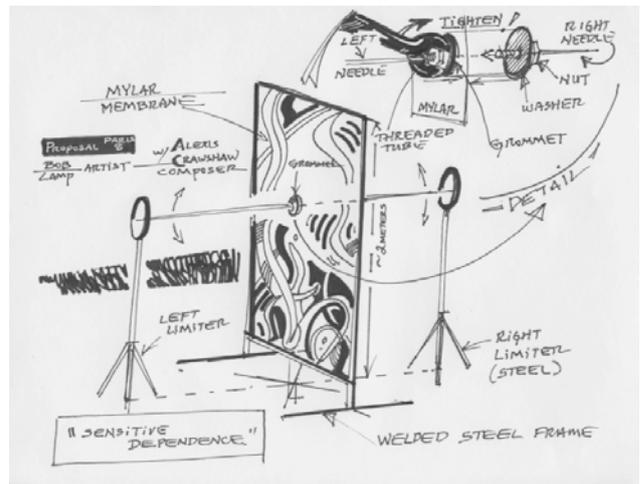


Figure 3. Esquisse d'objet « Sensitive Dependence », (Lamp, 2013) [5].

Le deuxième, « Mirror, Mirror », a une surface en BoPET très réfléchissante au dos. Il a un manche et des roulettes pour être déplacé.

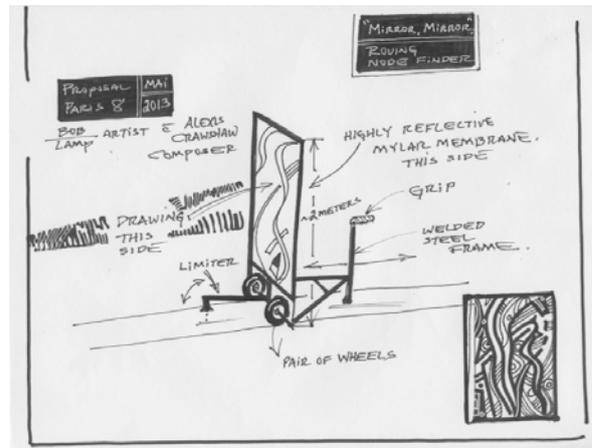
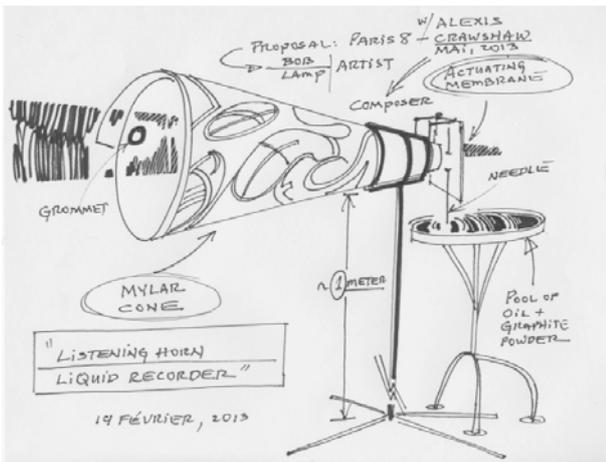


Figure 4. Esquisse d'objet « Mirror, Mirror », (Lamp, 2013) [6].

La troisième sculpture à membrane, « Listening Horn Liquid Recorder », consiste en une corne en papier BoPET roulé, fixé par un support en métal. Le petit bout de cette corne a une membrane avec une aiguille attachée. L'aiguille est posée dans une assiette remplie d'un fluide.



**Figure 5.** Esquisse d'objet « Listening Horn Liquid Recorder », (Lamp, 2013) [7].

Quant au deuxième type d'objet visuel, un bol profond en cuivre est rempli avec « oobleck », un fluide non-newtonien à base de farine de maïs et d'eau. Ce bol est percé avec des cordes élastiques pour le fixer au-dessus d'une des deux membranes de 10' du caisson de basse.

Finalement, on a des dessins adhésifs en BoPET avec Sharpie™ des chiffres, disposés au sol et qui mesure la largeur de la salle (la longueur de l'installation).

### 3.2. Conception

#### 3.2.1. Emploi de l'espace

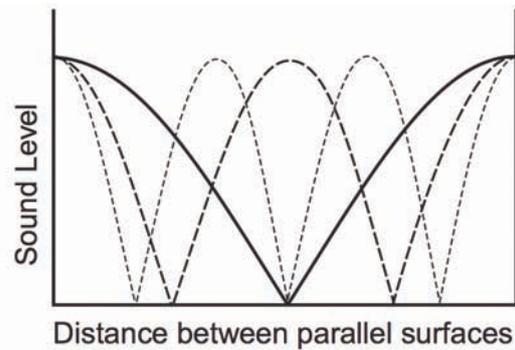
Comme on l'a évoqué plus tôt, l'espace physique de l'installation est un des éléments le plus intégral à ce projet. L'exposition occupera le dernier tiers de la longueur du hall d'entrée de l'université de Paris 8, un espace qui fait 20,14 m x 15,10 m x 2,70 m. Les fréquences des ondes utilisées seront calibrées pour activer des ondes stationnaires principalement à la largeur et parfois à la hauteur de la salle. Comme on ne veut pas déranger les autres projets qui seront exposés ailleurs dans la salle, on évitera les résonances de la longueur.

On peut calculer les modes propres de résonance de la cette salle avec la formule

$$f = \frac{c}{2L} \quad (1)$$

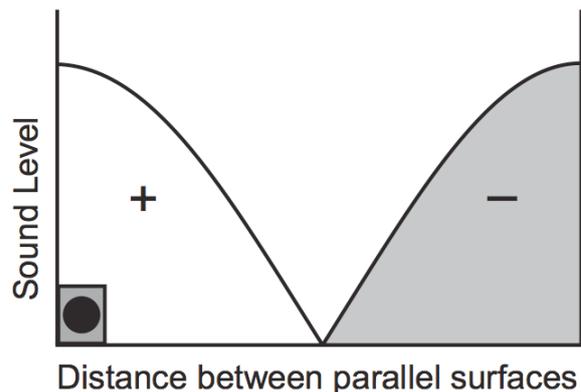
où  $f$  représente la fréquence en Hz,  $c$  représente la vitesse de son dans l'air et  $L$  représente la longueur de la salle ( $2L$  donne la longueur de l'onde sonore en mètres  $\lambda$ ) [1]. Ainsi, étant donné une température ambiante de 20°C, ces modes propres sont 8,37 Hz, 11,16 Hz et 62,44 Hz pour la longueur, la largeur et la hauteur, respectivement.

Selon les harmoniques de ses fréquences, on peut créer des motifs des nœuds et ventres comme ceux-ci :



**Figure 6.** Ce graphique indique le niveau de pression sonore dans une salle pour la fréquence fondamentale (—), la première harmonique (- - -) et la deuxième harmonique (- . - .) (Toole, 2008) [13].

Le caisson de basse sera posé à côté du mur pour maximiser les effets d'une onde stationnaire, comme indiqué dans ce graphique-ci :



**Figure 7.** Le niveau de pression sonore et la polarité du mode propre de résonance dans une salle selon l'emplacement d'un caisson de basse contre un des murs (Toole, 2008) [13].

#### 3.2.2. Composition musicale

La contribution musicale est électroacoustique et consiste en un seul canal, diffusé par le caisson de basse Meyer Sound. Divisée en 4 mouvements distincts, sa forme globale bascule entre la composition musicale et la sculpture sonore. Deux de ses mouvements ressemblent à une composition dans l'esprit d'une évolution temporelle comprise d'une recherche des sonorités variées au spectre infrasonore. Elle ressemble aussi à une sculpture sonore au sens qu'il y a deux transformations en une texture plus réduite comprise pour la plupart d'ondes sinusoïdales. Ces textures durent plusieurs minutes, et restent morphologiquement statiques lors le spectateur est immobile dans l'espace. Cependant, c'est pendant ces moments « statiques » que l'influence sur l'écoute selon le déplacement du spectateur sera la plus évidente.

Le contenu de ces mouvements à texture dynamique est un travail et un approfondissement des outils et techniques explorés dans la recherche de

Master de Crawshaw. Étant donné les limites de diffusion du caisson de basse, elle exploite des techniques dans le domaine des hauteurs (fréquences supérieures à 20 Hz), en utilisant des battements monauraux, des pulsations isochroniques, des harmoniques et des balayages descendants pour sensibiliser l'oreille à la présence des fréquences infrasonores. Elle se désigne à une palette de moyens réduite avec laquelle elle crée des infrasons : la synthèse FM, la synthèse additive et la synthèse granulaire avec le logiciel Max 6 et la synthèse par modélisation physique avec le logiciel Tassman 4. Ainsi, à travers la composition seule, il y a aussi un aspect d'aide et parfois de renforcement dans le domaine des hauteurs pour assister la perception des infrasons. Un travail compositionnel à travers des synthèses familières démontre le potentiel musical des infrasons.

### 3.2.3. Objets visuels

À part leur propre intérêt esthétique, les objets visuels fonctionnent en tant qu'un renforcement de la partie sonore : une traduction et une amplification.

Les sculptures du premier groupe sont conçues pour recevoir des ondes acoustiques sur leurs membranes en BoPET et elles sont équipées avec un appareil différent pour amplifier visuellement le moindre des mouvements de ces surfaces. L'objet « Sensitive Dependence » a des bâtons attachés à la membrane, l'objet interactif « Mirror, Mirror » a une surface réfléchissante pour créer des variations de la propre image de l'utilisateur (comme des miroirs de carnivals) et l'objet « Listening Horn Liquid Recorder » a une assiette de liquide pour créer des perturbations dans l'esprit de l'art cymatique et un enregistreur phonographique. En plus, en étant interactif et déplaçable, l'objet « Mirror, Mirror » fournit à l'utilisateur une aide visuelle qu'il peut déplacer pour chercher les nœuds et les ventres des ondes sonores.

Le bol en cuivre posé sur le caisson de basse crée des sculptures en « oobleck », provoquées en réponse aux ondes acoustiques. Cet objet présente un exemple de l'art cymatique du genre non-newtonien. C'est une version plus esthétique et élaborée que l'exemple qui se trouve ici [10]. La forme du bol est conçue pour contenir le fluide et assurer la propreté de la membrane du dispositif de diffusion.

Les chiffres dessinés sur le sol portent des indications de distance, en mètres, à travers l'installation. Pour le spectateur, ils servent de repère gradué de l'emplacement des nœuds et ventres des ondes dans la salle. Également, cet élément fonctionne avec l'appareil errant, « Mirror, Mirror ».

Finalement, tous les graphiques des dessins en Sharpie™ dans l'exposition servent à représenter les notions des ondes en propagation, des résonances et des perturbations.

## 4. CONCLUSION

Cette première maquette est un *Work in progress*. Une version plus élaborée est prévue pour le printemps de 2014, pour correspondre avec l'ouverture du nouveau bâtiment de la MSH Paris Nord. D'autres versions sont en cours d'étude.

En alliant la musique et les arts plastiques, ce projet offre une convergence intersensorielle unique pour faire découvrir la musique infrasonore.

Même s'il est encore trop tôt pour tirer des conclusions sur le plan théorique, ce projet démontre que la création artistique est un terrain d'expérimentation de premier ordre pour la recherche scientifique, particulièrement de la perception intersensorielle.

## 5. REFERENCES

- [1] Candel, S. *A tutorial on acoustics*. Cours de l'École Centrale Paris, 1998.
- [2] Chatillon, J. "Limites d'exposition aux infrasons et aux ultrasons : Étude bibliographique". *Hygiène et sécurité du travail : Cahiers de notes documentaires*. 2e trimestre 2006-203, pg. 67-77, INRS, 2006.
- [3] Crawshaw, A. (2011). "L'infrason en Art". Mémoire 1. Université de Paris 8, Saint-Denis, France, 2011.
- [4] Crawshaw, A. 2012. "Le potentiel de la musique infrasonore: Avec quelques applications intersensorielles". Mémoire 2. Université Paris 8, Saint-Denis, France 2012.
- [5] Lamp, R. Esquisse d'objet « Sensitive Dependence ». 2013.
- [6] Lamp, R. Esquisse d'objet « Mirror, Mirror ». 2013.
- [7] Lamp, R. Esquisse d'objet « Listening Horn Liquid Recorder ». 2013.
- [8] Leventhall, G. "A review of Published Research on Low Frequency Noise and its Effects". Report for Defra, 2003. Source: <http://www.scribd.com/doc/56798241/12/General-Review-of-Effects-of-Low-Frequency-Noise-on-Health1>
- [9] Meyer Sound Laboratories, Inc. "UMS-1P Self-Powered Subwoofer: Operating Instructions". Meyer Sound Laboratories, Inc., Berkeley, CA, États-Unis, 2000.
- [10] Non-newtonian fluid on a speaker cone. (2008). Extrait le 20 septembre 2013 de <http://www.youtube.com/watch?v=3zoTKXXNQIU>
- [11] Presonus. Extrait le 7 février 2013 de <http://www.presonus.com/products/Inspire-1394/techspecs>.
- [12] Salt, A.N. & Huller, T.E. "Responses of the ear to low frequency sounds, infrasound and wind

turbines". *Hearing Research* 268 (2010) 12-21  
doi:10.1016/j.heares.2010.06.007.

- [13] Toole, F. *Sound Reproduction: Loudspeakers and rooms*. Focal Press, Burlington, MA, États-Unis, 2008.

## LA BIBLIOTHEQUE HOA, BILAN ET PERSPECTIVES

*Julien Colafrancesco*  
CICM - EA1572, Université  
Paris 8, MSH Paris Nord,  
Labex Arts H2H  
jcolafrancesco@gmail.com

*Pierre Guillot*  
CICM - EA1572, Université  
Paris 8, MSH Paris Nord,  
Labex Arts H2H  
guillotpierre6@gmail.com

*Eliott Paris*  
CICM - EA1572, Université  
Paris 8, MSH Paris Nord,  
Labex Arts H2H  
eliottparis@gmail.com

*Anne Sèdes*  
CICM - EA1572, Université  
Paris 8, MSH Paris Nord,  
Labex Arts H2H  
anne.sedes@univ-paris8.fr

*Alain Bonardi*  
CICM - EA1572, Université  
Paris 8, MSH Paris Nord,  
Labex Arts H2H  
alain.bonardi@ircam.fr

### RÉSUMÉ

Cet article présente l'état actuel de la bibliothèque HOA en cours de développement, suite à un premier article paru dans les actes des Jim 2012. Nous présentons en détail l'ensemble des objets. Nous précisons l'apport de la décomposition en ondes planes dans le contexte ambisonique ainsi que l'usage de la synthèse binaurale pour une ambisonie virtuelle. Enfin nous abordons la prise en main de la bibliothèque par les musiciens.

### 1. INTRODUCTION

La bibliothèque HOA<sup>1</sup> (High Order Ambisonics) propose un ensemble de classes C++ et d'objets Max/MSP destinés à l'ambisonie d'ordre supérieur. Cette bibliothèque est un livrable issu du projet de recherche "La spatialisation du son par les musiciens pour les musiciens" développé au CICM dans le cadre du LABEX Arts H2H de l'université de Paris 8. Ce projet regroupe en équipe d'étudiants et chercheurs confirmés dans un cadre de recherche et création, associant l'informatique et la création musicale. La bibliothèque HOA permet de synthétiser, contrôler et transformer des champs sonores pour des environnements de travail accessibles aux musiciens. Grâce à son architecture modulaire, cette bibliothèque facilite la compréhension et l'appropriation de concepts clefs liés à l'ambisonie. L'utilisation des harmoniques sphériques (ou plus exactement circulaires, étant donnés les dispositifs matériels utilisés : des haut-parleurs sur un plan à 2 dimensions) et la création de traitements audio numériques dans le domaine correspondant y sont ainsi facilités.

Dans le cadre de cet article nous présenterons l'état de la bibliothèque et des ses développements, en

revenant sur une première publication parue dans le cadre des JIM 2012 qui annonçait le projet, en présentant les modules et les traitements offerts dans la bibliothèque, l'interopérabilité avec la décomposition en ondes planes, la synthèse binaurale associée au modèle ambisonique et la prise en main pas les musiciens. Nous concluons sur le futur de la bibliothèque.

### 2. LA BIBLIOTHEQUE DEPUIS SA PRESENTATION AUX JIM 2012

Depuis la première ébauche de la bibliothèque présentée aux JIM 2012 à Mons, une première phase d'évaluation nous a amenés à préciser notre approche.

#### 2.1. Retrait de la compensation de la distance

Les prémices de notre travail présentées en 2012 faisaient mention d'un module de compensation de champs proches (module NFC). Cette méthode proposée par Jérôme Daniel [1] occasionne cependant en pratique une très grande amplification en basse fréquence à mesure que la source sonore virtuelle se rapproche de l'auditeur et que l'ordre de décomposition ambisonique augmente. Des solutions à ce problème sont en cours d'étude [2][3][4] Les différentes propositions se basent sur un même concept, l'association d'un ordre de décomposition à une bande de fréquence spécifique. Plus précisément, il est question d'associer des bandes de fréquence de plus en plus haute à mesure que l'ordre augmente. Ces solutions occasionnent cependant des artefacts sonores très perceptibles qui rendent difficiles une exploitation de la technologie à des fins musicales. Cette année, lors de notre travail de refonte de la bibliothèque, nous avons décidé de retirer le module NFC de la version publique livrée.

<sup>1</sup> La bibliothèque est téléchargeable sur le site <http://www.mshparisnord.fr/hoalibrary/>.

## 2.2. Orientation vers la 2D

Le choix de proposer une bibliothèque ambisonique pour des systèmes de restitution à deux dimensions a été adopté pour plusieurs raisons.

Nous souhaitons avant tout offrir un ensemble d'outils facilitant l'appropriation de l'ambisonie par les musiciens, ce choix implique de prendre en compte les systèmes de restitution les plus répandus qui se situent sur un plan 2D (tels que les systèmes quadriphoniques et octophoniques) et qui ne sont pas toujours idéalement répartis dans l'espace<sup>2</sup>. Rappelons encore que le nombre de haut-parleurs dépend de l'ordre de décomposition. Pour une restitution bidimensionnelle, le nombre minimum de haut-parleurs croît de manière linéaire en fonction de l'ordre<sup>3</sup> (nombre minimum de haut-parleurs =  $2 * \text{ordre} + 1$ ), pour des systèmes à trois dimensions, le nombre de haut-parleurs augmente de manière quadratique (nombre minimum de haut-parleurs =  $(\text{ordre} + 1)^2$ ). Ainsi, les systèmes circulaires nous permettent de conserver un meilleur rapport entre le nombre de haut-parleurs et l'ordre de décomposition, ou le nombre d'harmoniques. Cette restriction et le fait que certaines opérations varient de manière significative à des ordres élevés<sup>4</sup>, a confirmé notre choix d'une bibliothèque adaptée à des systèmes de restitution à deux dimensions afin de nous permettre d'explorer et de développer ces opérations à des ordres qui nous aurait été inaccessibles en 3D.

Cependant, nous n'excluons pas la possibilité d'une déclinaison future de la bibliothèque pour des systèmes de restitution à trois dimensions en admettant que ce qui est applicable en 2D, l'est aussi en 3D ; nous pouvons espérer que nos propositions resteront pertinentes.

### 3. VERSION 1.1 DE LA BIBLIOTHEQUE HOA POUR MAXMSP

Des bibliothèques de spatialisation telles que, entre autres, le Spat<sup>5</sup> de l'IRCAM, celles de l'ICST<sup>6</sup>, de Graham Wakefield<sup>7</sup> ou encore de l'université de York<sup>8</sup> utilisent également l'ambisonie afin de spatialiser les

<sup>2</sup> Pour de plus amples explications sur les contraintes liées aux systèmes de restitutions et au positionnement de l'auditoire se référer à la documentation de la bibliothèque HOA : *L'ambisonie d'ordre supérieur et Les optimisations*, Guillot Pierre, 2012-2013, téléchargeable sur le site <http://www.mshparisnord.fr/hoalibrary/>.

<sup>3</sup> En 2D, pour un ordre 3 de décomposition, il faut  $2*3+1=7$  haut-parleurs, les usages standard amèneront en général à en utiliser 8.

<sup>4</sup> Ordre fractionnaires et synthèse granulaire par exemple.

<sup>5</sup> <http://forumnet.ircam.fr/product/spat/>

<sup>6</sup> <http://www.icst.net/research/downloads/>

<sup>7</sup> <http://www.mat.ucsb.edu/~wakefield/software.html>

<sup>8</sup> <http://www.york.ac.uk/music/mrc/software/objects/>

sons dans l'espace. La bibliothèque HOA propose d'offrir de nouvelles perspectives créatives aux musiciens par son approche particulière de l'ambisonie et du domaine des harmoniques circulaires, libérée d'une recherche de réalisme héritée des modèles acoustiques traditionnels. Nous revenons sur les nombreux traitements et les nouvelles interfaces désormais disponibles au sein de la bibliothèque<sup>9</sup>.

### 3.1. Modularité et accessibilité

Le principal enjeu de cette bibliothèque est de faciliter l'emploi et la création de traitements originaux dans le domaine des harmoniques sphériques pour le musicien. Il nous a dès lors paru indispensable de mettre l'accent sur sa modularité et son accessibilité.

Le développement du code en C++ et l'organisation modulaire des classes permettent la portabilité de notre bibliothèque sur de multiples plateformes. Un premier déploiement a été effectué à destination de l'environnement Max/MSP.

Effectuer un traitement dans le domaine des harmoniques sphériques se résume le plus souvent à appliquer ce traitement en parallèle sur différents canaux. Hors, les spécificités d'un environnement de programmation graphique tel que Max/MSP ne facilitent pas ce type de mise en place. Nous nous retrouvons à devoir effectuer un grand nombre de manipulations telles que l'instanciation et la connexion d'objets. Afin de factoriser ces interventions, nous avons développé l'objet *hoa.plugin~*. Un même traitement créé sous forme de patch peut être appliqué à l'ensemble des canaux ambisoniques et variabilisé selon un ordre et un indice. De plus, *hoa.connect* permet de faciliter les connections entre les objets de la chaîne de traitement tout en associant des couleurs aux liaisons en fonction de ce qu'elles représentent (une harmonique positive, une harmonique négative ou un signal dans le domaine des ondes planes).

Afin de rendre accessible la bibliothèque et ses traitements aux utilisateurs ne disposant que d'un casque ou d'un couple de haut-parleurs, nous offrons un rendu binaural et stéréophonique grâce aux objets *hoa.binaural~* et *hoa.stereo~*.

### 3.2. Principaux traitements

*Hoa.encoder~* et *hoa.decoder~* permettent l'encodage et le décodage ambisonique. Avoir séparé ces deux processus a permis d'offrir à l'utilisateur la possibilité d'exercer tout type de traitement lié à l'approche l'ambisonique, ce qui répond au principal enjeu de cette bibliothèque.

Trois types d'optimisation de décodage sont rendus possibles grâce à l'objet *hoa.optim~* :

<sup>9</sup> Se référer à la documentation contenue au sein de la bibliothèque HOA pour Max/MSP et *Les champs sonores*, Guillot Pierre, 2012-2013, téléchargeable sur le site <http://www.mshparisnord.fr/hoalibrary/>

- un décodage basic (présent par défaut, c'est-à-dire sans optimisation, et optimal pour un auditeur idéalement placé au centre du système circulaire de restitution sonore) ;

- un décodage maxRe (pour une meilleure restitution du champ sonore pour des positions d'écoute excentrées) [5].

- un décodage dit inPhase permettant de répondre à la situation critique où l'auditoire se trouve extrêmement proche des haut-parleurs ou même à l'extérieur du cercle de haut-parleurs [6].

L'objet *hoa.rotate~* applique une rotation de tout l'ensemble du champ sonore.

Les objets *hoa.projector~* et *hoa.recomposer~* élargissent ou resserrent la scène sonore en un point grâce au traitement *fisheye*.

L'objet *hoa.wider~* simule des ordres de décomposition fractionnaires afin de varier la résolution angulaire des lobes harmoniques, et de pouvoir passer linéairement d'une source ponctuelle à un champ sonore omnidirectionnel.

L'objet *hoa.halo~* applique un filtrage spatial du champ sonore afin d'en révéler ou cacher certaines parties.

L'objet *hoa.plug~* permet la mise en œuvre simplifiée de traitements sonores directement au sein de Max/MSP. Cet objet peut être utilisé à différents endroits de la chaîne de traitement du signal (sans encodage, avant, ou après un encodage ambisonique) suivant l'effet escompté.

Le patch *hoa.decorrelation~* permet de décorréler des canaux ambisoniques grâce à des lignes à retard et d'offrir un champ sonore diffus.

Le patch *hoa.grain~* est une application dans le domaine des harmoniques sphériques de la synthèse granulaire quasi-synchrone [Roads, 1996], [Roads, 2001]. Les différents paramètres sont modulés en fonction de chaque harmonique afin de synthétiser un champ sonore diffus à la texture spécifique.

Le patch *hoa.mirror~* propose de révéler ou cacher une image miroir du champ sonore en jouant sur le poids des harmoniques négatives.

Le patch *hoa.mixer~*, aide à mieux comprendre la contribution de chaque harmonique sphérique en laissant à l'utilisateur le soin de régler indépendamment le gain de chacune d'elles.

Le patch *hoa.closer~* reprend l'approche ambisonique de l'objet *ambipan~*<sup>10</sup> afin de jouer sur la distance d'une source.

Deux types de réverbération dans le domaine des harmoniques sphériques sont proposés :

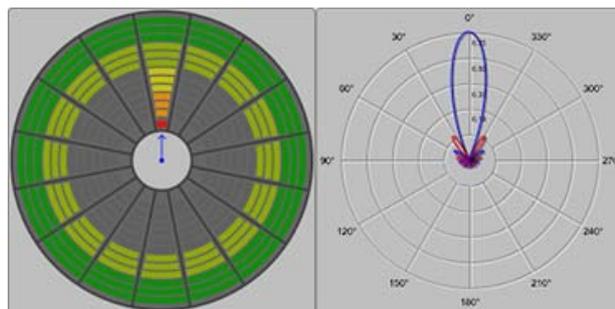
Le patch *hoa.reverberation~* utilise un réseau de lignes à retard avec réinjection adapté au domaine des harmoniques sphériques grâce à *hoa.plug~*.

L'objet *hoa.convolve~* permet d'obtenir une réverbération par convolution dans le domaine ambisonique en conjonction avec *hoa.irconverter~* qui

permet de transposer des réponses impulsionnelles dans le domaine des harmoniques sphériques.

### 3.3. Interfaces graphiques

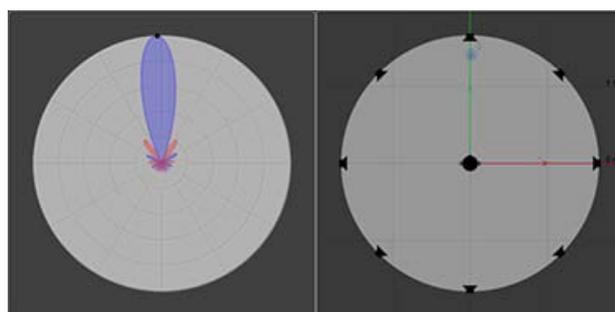
La bibliothèque HOA propose quatre interfaces graphiques. Les deux premières concernent la visualisation du signal (*hoa.meter~* et *hoa.scope~*) [Figure 1]. Les secondes (*hoa.control* et *hoa.map*) ont pour but d'offrir un contrôle graphique sur certains paramètres ambisoniques et permettent de mieux s'approprier les concepts liés à l'ambisonie.



**Figure 1.** Représentation des contributions de seize haut-parleurs pour un dirac encodé à 0° à l'ordre 7 via l'objet *hoa.meter~* (gauche) et des harmoniques correspondantes *hoa.scope~* (droite).

L'objet *hoa.scope~* permet, à la manière d'un *phase scope*, de visualiser l'amplitude et la phase des harmoniques.

L'objet *hoa.meter~* offre une alternative à la représentation traditionnelle du VU-mètres, peu pertinente dans le contexte ambisonique par la visualisation des contributions de plusieurs haut-parleurs positionnés de manière circulaire.



**Figure 2.** Représentation via l'objet *hoa.control* (gauche) de l'encodage d'un dirac à 0° à l'ordre 7 et de sa projection sur un plan cartésien via l'objet *hoa.map* (droite).

L'objet *hoa.control* aide à la compréhension et à la visualisation des harmoniques sphériques tout en permettant leur manipulation (ordre d'encodage, ordre fractionnaire, types d'optimisation et azimuth).

<sup>10</sup> Objet disponible sur le site du CICM: <http://cicm.mshparisnord.org/>

L'interface *hoa.map* permet la spatialisation de sources sonores ponctuelles sur un plan 2D. Les sources peuvent être manipulées, indépendamment ou de manière groupée. Les données d'azimut de chaque source peuvent être directement traitées par l'encodeur ambisonique, les données de rayon, quant à elles, peuvent être utilisées grâce à une mise à l'échelle avec l'amplitude de la source, pour contrôler la distance via *hoa.closer~* ou encore une réverbération via *hoa.convolve~*.

#### 4. DECOMPOSITION EN ONDES PLANES

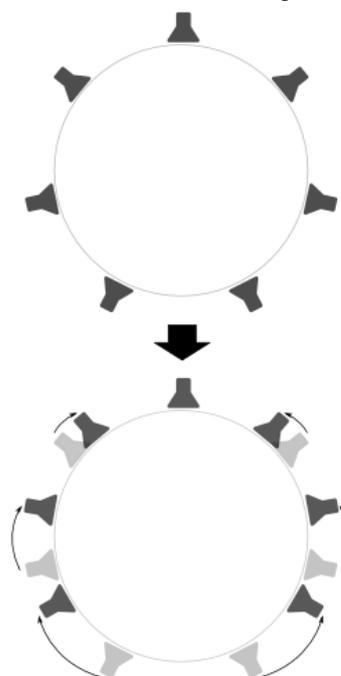
L'onde plane est un modèle d'onde fréquemment utilisé en physique. Une onde plane dispose de fronts d'ondes assimilables à des plans (ou, en 2D, à des lignes), de même amplitude et se propageant selon une même direction. En acoustique, nous associons souvent ce modèle à une source sonore éloignée. En effet, le front d'onde d'une source sonore ponctuelle omnidirectionnelle peut être associé à la surface d'une sphère dont la courbure et la variation d'amplitude diminuent à mesure que le rayon augmente. A partir d'une certaine distance entre la source sonore et le point de mesure, et donc à partir d'un certain rayon, le front d'onde est assez peu courbé et varie assez peu en amplitude pour être associé à celui d'une onde plane.

Alternativement à la décomposition en harmoniques sphériques, il existe d'autres manières de décomposer un champ sonore. La décomposition en ondes planes est une de ces opérations, elle nous permet de représenter un champ sonore comme une somme infinie d'ondes planes.

L'ambisonie, et plus particulièrement son étape de décodage, utilise implicitement une version discrète de cette décomposition. En effet, il s'agit de restituer un champ sonore par le biais d'un ensemble de haut-parleurs approximatifs comme étant à onde plane. Les émetteurs ne sont donc représentés que par leur direction, la distance les séparant de l'auditeur n'étant pas prise en compte. Les décodages par projection ou pseudo-inverse [7] peuvent donc être vus comme autant d'outils nous permettant de passer d'une représentation du champ sonore aux harmoniques sphériques à une représentation en onde plane.

Si dans le cas de l'ambisonie ce processus est associé à l'étape de restitution sur haut-parleur, il est possible d'utiliser cette représentation à d'autres fins. On pourrait ainsi concevoir cette dernière comme une représentation de passage nous permettant de mettre en œuvre des traitements directionnels. Parmi les applications possibles, notons le filtrage d'un champ sonore afin de mettre en avant ou de retirer l'un de ses objets. L'efficacité de ce type d'opération étant dépendante de la finesse de notre représentation (i.e. du nombre de fonctions de base considérées) ainsi que de la validité avec laquelle il est possible d'associer le modèle de l'onde plane à l'onde générée par la source sonore traitée.

Dans le cadre de ce projet nous avons tenté de répondre à l'une des problématiques soulevées par Jérôme Daniel dans sa thèse. Ce dernier décrit en effet une opération initialement proposée par Gerzon [8] : la *distorsion de la perspective*<sup>11</sup>. Il s'agit d'une opération appliquée à une représentation ambisonique permettant d'élargir ou de resserrer la scène frontale. Pour mieux visualiser ce dont il s'agit, nous pouvons noter une certaine similitude entre ce type de traitement et l'application en optique, d'un effet de type *fish-eye*. La distorsion de la perspective est obtenue par Gerzon via une certaine forme de transformation de Lorentz, l'avantage de ce mode opératoire étant la préservation des propriétés de l'onde plane. En d'autres termes, une onde plane à qui l'on fera subir une distorsion de la perspective restera une onde plane. Dans sa thèse, Jérôme Daniel souligne cependant que cette propriété n'est vraie qu'à l'ordre 1, que son extension aux ordres supérieurs n'est malheureusement pas vérifiée.



**Figure 3.** Déplacement des haut-parleurs virtuels dans le but d'obtenir une contraction de la scène frontale.

Malgré cette limitation, la distorsion de la perspective nous a semblé être une manipulation disposant d'un certain potentiel musical. Nous nous sommes donc intéressés à la mise au point d'une stratégie nous permettant d'arriver à un résultat perceptivement équivalent dans le cadre d'une représentation ambisonique d'ordre supérieurs. Nous parlons d'une

<sup>11</sup> Il est possible de se questionner concernant la légitimité avec laquelle nous pouvons employer le terme « *perspective* », ce dernier n'étant pas vraiment adapté à la « géométrie » de l'écoute dans un contexte ambisonique.

équivalence sur le plan perceptif car l'extension aux ordres supérieurs nous oblige à faire fi de la conservation des propriétés de l'onde plane. En cela notre démarche pourrait être vue de manière analogue à l'élaboration d'un vocoder de phase, ce dernier nous permet d'étirer ou de compresser un son, mais il ne conserve pas les propriétés d'une simple impulsion de Dirac. La validité d'un time-stretcher se situe ailleurs, dans sa capacité à conserver certains indices pertinents sur le plan perceptif comme par exemple la durée des attaques d'un instrument de musique. C'est dans cette démarche que nous avons étudié l'extension d'un effet de distorsion de la perspective aux ordres supérieurs.

La solution que nous proposons se base sur l'interopérabilité entre la décomposition en harmoniques sphériques et la décomposition en ondes planes mentionnée précédemment. L'opération s'articule en trois étapes, la représentation ambisonique d'ordre supérieur est tout d'abord transposée dans le domaine des ondes planes. Afin de nous représenter plus aisément cette opération, nous pouvons imaginer que chacune de ces ondes est associée à un haut parleur virtuel, chaque haut-parleur virtuel étant disposé de manière homogène sur le pourtour du cercle de restitution. La deuxième étape consiste à déplacer les haut-parleurs de manière à obtenir la distorsion de perspective désirée. Par exemple, si nous désirons un rétrécissement de la scène frontale, il nous faudra déplacer les haut-parleurs d'une manière analogue à celle illustrée en [Figure 3]. La dernière étape consiste à retourner dans le domaine des harmoniques sphériques où nous pouvons disposer d'une représentation ambisonique d'ordre supérieur ayant subi une distorsion de la perspective.

Dans le contexte de la bibliothèque HOA, la mise en place de ce genre de pratique a permis via la création d'une série d'objets facilitant les aller-retour entre les deux domaines cités précédemment. La classe "projection" permet ainsi de projeter une décomposition en harmoniques sphériques sur une base arbitraire d'ondes planes et la classe "recomposition" altère dans un premier temps la direction d'incidence des différentes ondes planes avant de retourner dans le domaine des harmoniques sphériques.

## 5. AMBISONIE ET BINAURALE

La bibliothèque HOA permet d'étendre les techniques ambisoniques afin d'offrir une restitution sonore au casque grâce à la synthèse binaurale [9] [10]. Cette proposition, déjà présente dans certaines bibliothèques de spatialisation du son tel que le Spat<sup>12</sup> ou Harpex<sup>13</sup>, est ici revisitée afin d'offrir un rendu pour les ordres supérieurs. L'originalité de l'approche présentée réside dans son optimisation grâce aux bibliothèques de calculs

matriciels qui permettent d'offrir une restitution préservant la linéarité des variations de phase via l'utilisation de filtres FIR (Finite Impulse Response) [11] tout en restant accessible sur des ordinateurs personnels<sup>14</sup>.

La mise en œuvre d'une ambisonie virtuelle par convolution dans le domaine temporel peut sembler être un choix non optimal comparée aux approches par FFT [12] et par filtres IIR [9]. Pour de petites réponses impulsionnelles, et sur des machines favorisant la parallélisation et la vectorisation des calculs, effectuer une convolution directement peut cependant être un choix judicieux [12]. Obtenir une restitution binaurale à partir d'un champ sonore représenté dans le domaine des harmoniques sphériques revient, dans un premier temps, à restituer ce champ sonore pour un nombre défini de haut-parleurs virtuels. En fonction de l'angle d'incidence du haut-parleur dont il provient, chaque signal est par la suite filtré par un couple de HRIR (Head Related Impulse Response).

Afin d'obtenir une restitution binaurale de qualité, un grand nombre de haut-parleur doit être considéré, ceci implique de nombreuses convolutions devant être effectuées en parallèle. La solution proposée par la bibliothèque HOA réside dans sa représentation matricielle des calculs.

Une nouvelle matrice de décodage est en effet proposée. Ce nouveau décodage peut être vu comme une projection du décodage multicanal classique sur uniquement deux canaux intégrant les informations relatives aux différentes HRIR.

En termes de calculs, nous passons d'un grand nombre de FFT au profit d'une seule multiplication matricielle. Dans les bibliothèques que nous avons utilisées, cette opération est hautement vectorisée, parallélisée et bénéficie d'algorithmes de FMF (Fast Matrix Factorization). L'ensemble nous permet d'optimiser les calculs afin d'être accessible sur la majorité des ordinateurs personnels d'aujourd'hui.

Comme chaque haut parleur nécessite l'utilisation d'un couple d'HRIR distinct, l'ordre de décomposition est contraint par le nombre de réponses impulsionnelles. Dans la bibliothèque HOA, nous utilisons les HRIR du laboratoire CIPIC<sup>15</sup>, qui comporte 178 réponses impulsionnelles horizontales pour chaque oreille, l'ordre maximal est donc de 35 et nous tirons parti de la bibliothèque de calcul matriciel MKL<sup>16</sup> pour optimiser l'ensemble des calculs matriciels.

Élargir les moyens de restitution par le rendu binaural permet de spatialiser une multitude de sources sonores sans pour autant augmenter le nombre de filtres binauraux et offre aux utilisateurs un accès au casque à l'ensemble des traitements ambisoniques proposés par la

<sup>12</sup> <http://www.ircam.fr/>

<sup>13</sup> <http://www.harpex.net/>

<sup>14</sup> L'utilisation du CPU (processeur Intel Core i7 à 2,3 GHz) est d'environ 23% pour un ordre 35, 12% pour un ordre 7 et 9% pour un ordre 1.

<sup>15</sup> <http://interface.cipic.ucdavis.edu/>

<sup>16</sup> <http://software.intel.com/en-us/intel-mkl>

bibliothèque HOA sans la nécessité donc de posséder un système composé d'une multitude de haut-parleurs.

## 6. PRISE EN MAIN PAR LES MUSICIENS

Dans le cadre du projet de recherche « la spatialisation du son par les musiciens pour les musiciens », il était établi que l'observation des retours d'usage se ferait au cours de situations réelles de création musicale, conduites au sein de l'équipe. Deux projets de pièces mixtes ont vu le jour dans ce contexte, ainsi qu'une installation sonore interactive. Par ailleurs, des évaluations dans des cadres pédagogiques ont pu avoir lieu.

### 6.1. Anne Sédès, Immersion, pour violoncelle et électronique live

Il s'agit d'une pièce<sup>17</sup> pour violoncelle, et électronique live avec Max/MSP comportant une pédale midi et 8 haut-parleurs. Nos avons donc été amenés à travailler à l'ordre 3, ce qui permet d'obtenir le maniement de 7 harmoniques.

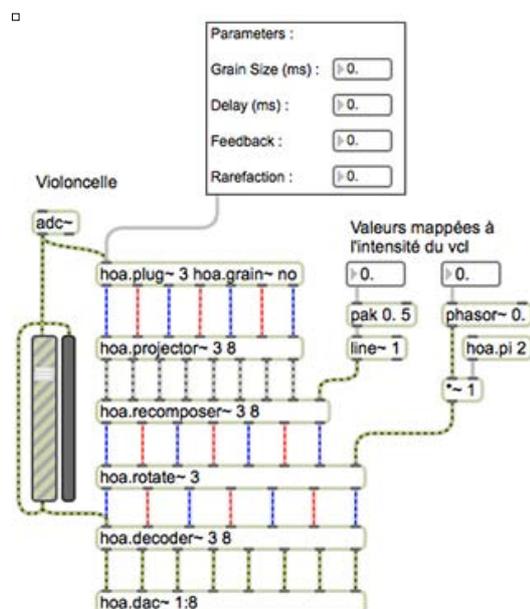


Figure 4. Le module de spatialisation avec HOA dans le patch de la pièce Immersion.

L'enjeu de la pièce est la mixité, ou comment faire converger une écriture instrumentale dédiée au potentiel sonore du violoncelle avec le traitement spatial du son instrumental en temps réel, grâce aux moyens de HOA, en s'appuyant sur le maniement de quelques variables.

Parmi l'ensemble des traitements offerts dans HAO, nous avons fait le choix de la granulation, en utilisant *hoa.grain~* dans *hoa.plug~*. D'une part, c'était un des premiers traitements disponibles dans la bibliothèque. D'autre part, la granulation a un potentiel expressif que nous sommes habitués à manier (accès à des variations de timbre, de rythme, à du délai non répétitif, à de l'aléatoire, à des effets de spatialités décorrélables selon les canaux etc.). Nous avons également fait le choix de la variation de la rotation globale et de la contraction/dilatation du champ sonore, en nous inspirant de l'emploi habituel de trajectoires de sources sonores, mais ici appliqué à l'ensemble du champ sonore [Figure 4].

Les paramètres variables de *hoa.grain~*, qui est basé sur une technique de ligne à retard granulaire, sont les suivants : le choix de l'harmonique, la taille du grain, le retard, la réinjection et la raréfaction.

Nous avons construit au fil de l'écriture instrumentale une sorte de canon, introduisant par changements de programme à la pédale midi les successions de couches granulaires propres à chaque harmonique, les changements de variables étant soit contenus dans un tableau *patternstorage*, soit mappés à l'intensité du jeu instrumental (via l'objet *peakamp~*).

Nous avons également utilisé des techniques de *Fisheye* (*hoa.recomposer~*) et de rotation (*hoa.rotate~*), connectées au suivi d'amplitude (*peakamp~*) que nous avons appliquées à l'ensemble des couches produites, afin d'obtenir des jeux de mouvements spatiaux globaux,

<sup>17</sup> La pièce a été composée au studio de la MSH Paris Nord entre Juillet et novembre, avec quelques séances d'essai avec le violoncelliste Guilherme Carvalho, puis avec des enregistrements des séances. L'écriture instrumentale a été fixée à l'automne. Une première présentation a eu lieu à l'université Paul Valéry-Montpellier 3, dans le cadre d'un concert-lecture, le 27 novembre 2012. La pièce sera rejouée dans le cadre des JIM 2013 le 13 mai prochain. Cette pièce a également une fonction pédagogique, car elle a permis de fournir des exemples didactiques dans le cadre de l'atelier de composition du département de musique de l'université Paris 8.

en contrepoint avec la source instrumentale directe diffusée au centre. Afin d'obtenir une cohérence entre le jeu du violoncelle en direct et son traitement spatial, le son direct du violoncelle a en effet été connecté vers l'entrée du 1er harmonique, pour que le violoncelle seul soit diffusé en mono, indépendamment des traitements de spatialisation.

Dans cette pièce, Nous avons donc pu tester la prise en main de la bibliothèque, sans trop de difficultés, et faire des retours de détail aux développeurs. Nous avons au fond, servi de beta-testeurs. Nous avons ainsi pu évaluer la grande stabilité de la bibliothèque.

Sur le plan de la musicalité des traitements, nous avons fait le choix de la granularité, pour des raisons expressives et parce qu'il permettait facilement de construire des couches différenciables, autrement dit, de la décorrélation temporelle [13], correspondant chacune à un harmonique. En quelque sorte, nous avons considéré chaque harmonique circulaire comme une couche, comme une voix, pour une polyphonie spatiale. Le jeu avec les rotations et l'effet *fisheye* nous a permis d'introduire une véritable dynamique du mouvement sans tomber dans des effets systématiques de trajectoires circulaires d'une source sonore. Nous avons testé au passage l'usage de la stéréo et de l'ambisonie pour des séances domestiques de préparation du travail en amont du studio. Avec une approche de la mixité, basée sur la variation de peu de traitements du son instrumental avec HOA (la granulation et ses variables, la rotation et l'effet *fisheye*), nous avons donc réussi à construire une pièce dont le rendu nous donne musicalement satisfaction.

## 6.2. Alain Bonardi, Pianotronics 1, pièce pour piano et live électronique

Cette pièce pour piano et électronique temps réel, diffusée sur un dispositif octophonique, est en cours de composition; elle sera créée lors des Journées d'Informatique Musicale en mai 2013. Nous travaillons sur la variabilité de l'espace sonore, en explorant ses continuités et discontinuités selon deux axes :

- la variabilité de présence du piano : les situations sonores vont d'un piano au premier plan en solo en jouant sur son timbre à un piano absent en son direct mais présent dans l'électronique, en passant par des situations intermédiaires. La spatialisation par HOA contribuera largement à ces variations de présence.
- la variabilité du champ électroacoustique par des micro-polyphonies sur les harmoniques circulaires, à partir essentiellement d'un sampler aux sorties distribuées sur ces dernières.

## 6.3. Transduction, installation interactive

Dans le cadre du projet, une installation sonore interactive est également développée. En utilisant des interfaces de contrôle et de jeu de type de Kinect, iPad

ou Wii-balance, nous invitons un public non initié à jouer avec des contenus spatialisés. Une première proposition a été présentée à Savantes banlieues 2012 (fêtes de la science<sup>18</sup> à Villetaneuse). Une deuxième proposition aura lieu lors des présentes JIM et dans le cadre du festival Futurs en Seine 2013. Elle inclura cette fois-ci la création d'un interacteur spécifique sur IPAD, proposé par Manuel Deneu, dans le cadre d'un partenariat avec l'ENS Louis Lumière.

Ce développement artistique sous forme d'installation, qui cherche à transduire l'énergie gestuelle du public visiteur à l'écoute du dispositif de spatialisation est pour l'équipe un moyen de tester, "à taille réelle" et dans des situations standard de production artistique, le rendu en cours de développement de la bibliothèque HOA. Il va accompagner spécialement les développements à venir sur 2013, concernant le développement des interfaces de contrôle et d'écriture de la spatialisation du son avec HOA.

## 6.4. Prise en main par de jeunes musiciens

La récente résidence d'Alain Bonardi au Conservatoire à Rayonnement Régional de Bayonne fin janvier 2013 abordait les thèmes de la synthèse et de la spatialisation. La librairie HOA a été présentée à cinq jeunes compositeurs de la classe d'électroacoustique (professeur : Patrick Defossez) qui ont pu faire des premiers tests sur un dispositif octophonique. Ils ont mis en œuvre l'encodage, le décodage et la gestion d'effets sous forme de plug-ins (lignes à retard, réverbérations). Les points positifs immédiatement relevés sont :

- la possibilité d'appropriation rapide des outils d'ambisonie grâce aux patches d'aide immédiatement opérationnels;
- la facilité de mise en œuvre d'un effet dans *hoa.plug~*.

Ayant tout d'abord utilisé l'objet *ambipan~* des CICM Tools (ambisonie d'ordre 1), ils ont pu comparer les deux approches. Les retours se focalisent sur deux points:

- alors que l'objet *ambipan~* proposait d'emblée la saisie de l'azimut et de la distance à l'origine (coordonnées polaires), ce dernier paramètre est absent de la librairie HOA : l'objet *hoa.encoder~* ne permet de manipuler que l'azimut. Il est certes possible de coder une atténuation liée à la distance, mais cette dimension n'était alors pas donnée d'emblée au musicien.
- les harmoniques circulaires ne sont pas évidentes à comprendre et à appréhender dans l'espace sonore.

En fin de résidence, une courte restitution a été organisée, donnant lieu à des esquisses de quelques minutes, essentiellement à base d'improvisations instrumentales traitées par l'ordinateur (la plupart de ces

<sup>18</sup> <http://savantebanlieue.plainecommune.fr/>

jeunes compositeurs étant également musiciens improvisateurs). Désormais, ils développent ces esquisses avant une prochaine présentation fin mai ou début juin devant un jury. Les points négatifs étaient que les étudiants ne trouvaient pas dans HOA les facilités offertes dans l'objet *ambipan* en terme de contrôle du rayon de la source. C'est suite à cette remarque de musiciens en situation de création que *hao.closer* a été ajouté à la bibliothèque.

HOA est également utilisée dans le cadre de l'atelier de composition instrumentale et électronique du département de musique de l'université de Paris 8, sous la direction de J.-M. Lopez Lopez, A. Sèdes et A. Bonardi. Les travaux de composition mixte des étudiants seront présentés en concert en juin 2013. Là encore, les retours d'utilisation sont étudiés de très près.

## 7. CONCLUSION

Grâce à la mise en œuvre de ce projet, nous avons pu avancer sur la possibilité de traitement au niveau des harmoniques circulaires, la possible interopérabilité entre décomposition en ondes planes et décomposition en harmoniques sphériques ainsi que l'ambisonie virtuelle (par le biais de la synthèse binaurale). Nous avons tenté de mettre le potentiel de l'ambisonie à la disposition des musiciens électroacousticiens et des développeurs en informatique musicale.

Les présents résultats nous incitent à poursuivre nos recherches en développant les techniques de convolution et de décorrélation appliquées au domaine des harmoniques sphériques afin d'offrir aux musiciens électroacousticiens de nouveaux traitements. Par ailleurs, il nous semble également utile d'élargir nos systèmes de restitution pour permettre l'utilisation de configurations irrégulières de haut-parleurs auxquelles les musiciens sont souvent confrontés. Également, nous n'excluons pas d'examiner la combinaison d'autres techniques de spatialisation telle que la WFS avec l'ambisonie dans le cadre d'HOA.

Notons que des perspectives de portage de la bibliothèque sont à l'ordre du jour dans la communauté de l'informatique musicale. Ainsi, une étude de faisabilité du déploiement sur FAUST est en cours avec le GRAME.

Pour 2013, l'effort va porter sur le développement d'interfaces de contrôle et d'écriture pour la spatialisation du son avec HOA. Cet effort engage une réflexion sur les interfaces de contrôle et d'écriture de la mise en espace du son selon diverses échelles temporelles. C'est la raison pour laquelle notre équipe porte un intérêt pour les projets OSSIA et ISCORE et commence sur ce plan un dialogue avec l'équipe du GMEA à Albi.

Notons enfin que le retour d'utilisation en situation de création artistique a permis d'évaluer en taille réelle et en situation professionnelle de production la pertinence conceptuelle des outils sur le plan de leur utilisation musicale, autant que la stabilité technique des logiciels offerts. Du point de vue du musicien, elle a permis de développer une pensée musicale en interaction avec les outils en cours de développement. Nous ne pouvons ici que nous féliciter d'un travail dont la méthodologie a bénéficié d'une dynamique d'équipe, alliant recherche et création.

## 8. REFERENCES

- [1] Daniel, J. "Spatial Sound Encoding Including Near Field Effect : Introducing Distance Coding Filters and a Viable New Ambisonic Format", AES 23rd International Conference, Copenhagen, Danemark, 2003.
- [2] Daniel, J. Moreau, S. "Spatial encoding and decoding of focused virtual sound sources", Proc. Of the 1st Ambisonics Symposium, Graz, Austria, 2009.
- [3] Ahrens, J. Spors, S. "Further study of sound field coding with high order ambisonics" Proc. of the 116th Audio Engineering Society Convention, Berlin, Germany, 2004.
- [4] Favrot, S. Buchholz, J. "Reproduction of nearby sound sources using hier-order ambisonics : Implementation and evaluation", Proc. of the 36th German Annual Conference on Acoustics, Berlin, Germany, 2010.
- [5] Gerzon, M. A. "General metatheorie of auditory localisation", Preprint 3306 of the 92nd Audio Engineering Society Convention, Vienne, Autriche, 1992.
- [6] Malham, D. "Experience with large area 3D ambisonic sound systems", Proceedings of the Institute of Acoustics, York, Royaume-Uni, 1992.
- [7] Colafrancesco, J. "L'ambisonie d'ordre supérieur et son appropriation par les musiciens", Actes des Journées d'Informatique Musicale, Mons, Belgique, 2012.
- [8] Gerzon, M. A. "Ambisonic Decoder for HDTV". Proc. of the 92nd AES Convention, Vienne, Autriche, 1992.
- [9] Larcher, V. "Techniques de spatialisation des sons pour la réalité virtuelle", Thèse de doctorat, Université Paris 6, 2001.
- [10] Noisternig, M. Musil, T. Sontacchi, A. Höldrich, R. "3D Binaural Sound Reproduction using a Virtual Ambisonic Approach", International Conference On Virtual Environments, Human-Computer Interfaces, And Measurement Systems, Lugano, Suisse, 2003.
- [11] Smith, J.O. "Partial Fraction Expansion", in *Introduction to Digital Filters with Audio Applications*,

[http://ccrma.stanford.edu/~jos/filters/Partial\\_Fraction\\_Expansion.html](http://ccrma.stanford.edu/~jos/filters/Partial_Fraction_Expansion.html), 2007, Livre en ligne, accès 2013.

- [12] Smith, J.O. "Fourier Theorems for the DFT", in *Mathematics of the Discrete Fourier Transform (DFT) with Audio Applications, Second Edition*, [http://ccrma.stanford.edu/~jos/mdft/Fourier\\_Theorems\\_DFT.html](http://ccrma.stanford.edu/~jos/mdft/Fourier_Theorems_DFT.html), 2007, Livre en ligne, accès 2013.
- [13] Vaggione, H. « Décorrélation microtemporelle, morphologies et figurations spatiales, actes des Journées d'informatique musicale », Actes des Journées d'Informatique Musicale, Marseille, France, 2002.



## ATELIER : OUTILS POUR L'ANALYSE DE LA MUSIQUE ELECTROACOUSTIQUE

*Alain Bonardi*  
CICM - EA1572  
Université Paris 8  
Ircam  
alain.bonardi@orange.fr

*Bruno Bossis*  
APP Université Rennes 2  
OMF-MINT Université Paris-  
Sorbonne  
bruno.bossis@univ-rennes2.fr

*Pierre Couprie*  
OMF-MINT Université Paris-  
Sorbonne  
MTIRC De Montfort  
University  
pierre.couprie@paris-  
sorbonne.fr

*Frédéric Dufeu*  
CeReNeM  
University of Huddersfield  
f.dufeu@hud.ac.uk

*Mikhail Malt*  
Ircam  
OMF-MINT Université Paris-  
Sorbonne  
mikhail.malt@ircam.fr

*Laurent Pottier*  
CIERC - EA3068 Université  
Jean Monnet Saint-Etienne  
laurent.pottier@univ-st-  
etienne.fr

### RÉSUMÉ

Le groupe de recherche sur l'analyse des musiques électroacoustiques se réunit sous l'égide de la SFAM. Il regroupe six chercheurs et a pour objectif de produire un état de l'art sur l'usage des technologies numériques dans l'analyse de l'électroacoustique.

#### 1. LE GROUPE DE RECHERCHE

Le groupe de recherche sur l'analyse des musiques électroacoustiques est issu de la Société Française d'Analyse Musicale<sup>1</sup>. Ce groupe réunit six chercheurs appartenant à différentes institutions et partageant des axes de recherches proches : l'analyse musicale, l'étude des musiques électroacoustiques et le développement des pratiques du numérique dans la musicologie.

Le groupe se réunit régulièrement afin de partager différentes recherches et d'esquisser de futures directions. La première journée d'étude a été accueillie par le CIERC de l'Université Jean Monnet de Saint-Etienne et a permis de faire le point sur les recherches menées par les membres du groupe et par Georges Bloch, invité pour l'occasion. Les présentations portaient sur l'usage des descripteurs audio, les étapes de segmentation/description/typologie, l'usage de différents logiciels dans l'analyse musicale (OMax, EAnalysis, Max) et l'analyse génétique à partir de sources informatiques (Max et SuperCollider).

Après cette première journée, le groupe se dirige vers une deuxième phase : l'analyse de six œuvres à travers le prisme de différentes approches.

#### 2. L'ATELIER

L'atelier sera l'occasion de faire découvrir les outils utilisés par les membres du groupe autour de l'analyse de *Entwurzelt* de Hans Tutschku pour six chanteurs et électronique. Il se déroulera en 2 parties :

- une introduction générale sur le sujet et sur les aspects musicaux abordés, suivie d'une courte présentation générale des outils ;
- une séance pratique sur l'usage des différents outils.

Lors de la séance pratique, les participants pourront installer et manipuler les logiciels. Un lien de téléchargement sera proposé lors de l'ouverture des JIM. Les logiciels utilisés durant l'atelier sont :

- Audiosculpt (Macintosh) : <http://forumnet.ircam.fr/product/audiosculpt/>
- EAnalysis (Macintosh) : <http://eanalysis.pierrecouprie.fr>
- Open Music (Macintosh & Windows) : <http://forumnet.ircam.fr/shop/fr/forumnet/43-openmusic.html>
- Sonic Visualiser (Macintosh, Windows & Linux) : <http://www.sonicvisualiser.org>
- tIAAls (Macintosh & Windows) : <http://www.hud.ac.uk/research/researchcentres/tacem/>
- Vamp libxtract plugin pour Sonic Visualiser (Macintosh & Windows) : <http://www.vamp-plugins.org/download.html>
- Zsa.descriptors (bibliothèque d'analyse de descripteurs audio, en temps réel, pour Max, Macintosh & Windows) : [http://www.e-j.com/?page\\_id=499&lang=fr](http://www.e-j.com/?page_id=499&lang=fr)

<sup>1</sup> SFAM : <http://www.sfam.org>.

### 3. REFERENCES

- [1] Bonardi, A., "Approches pratiques de la préservation/virtualisation des œuvres interactives mixtes : En Echo de Manoury", JIM, Saint-Etienne, 2011.
- [2] Bossis, B., "De l'amplification à l'autonomie de l'instrument : ruptures électroniques et conséquences esthétiques", *Analyse Musicale*, Paris, 2009.
- [3] Couprie, P. "EAnalysis : aide à l'analyse de la musique électroacoustique", JIM, Mons, 2012.
- [4] Dufeu, F. "L'instrument numérique comme objet d'analyse des musiques mixte", JIM, Rennes, 2010.
- [5] Malt, M., Jourdan, E., "Le 'BSTD' – Une représentation graphique de la brillance et de l'écart type spectral, comme possible représentation de l'évolution du timbre sonore" L'analyse musicale aujourd'hui, Crise ou (r)évolution ?, Strasbourg, 2009.
- [6] Pottier, L., "Turenas (1972) de John Chowning, vers une version interactive", Musimediane, Paris, 2011.

# MULTI-DIMENSIONAL CONTROLLERS, EXPRESSIVITY, SOFTWARE: AN INTEGRATION PROBLEM

*Jean-Baptiste Thiebaut*  
ROLI

jean-baptiste@weareroli.com

*Sean Soraghan*  
ROLI

sean@weareroli.com

*Lauren Ianni*  
ROLI

ljianni@post.harvard.edu

## RÉSUMÉ

This paper identifies a problem faced by many manufacturers of multi-dimensional controllers (MDCs) for musical control. Namely, the difficulty faced by these manufacturers when trying to integrate their hardware with 3rd party audio software. It presents this problem through the explanation of a new MDC - the Seaboard - and the difficulty being faced with the integration of the Seaboard into audio software environments. A list of parameters is put forward that may facilitate the categorisation of gestures in new MDC systems. These parameters are suggested in order to aid in the development of a new MDC framework that focuses on gestural control and relevant musical output. A solution to this specific protocol is beyond the scope of this paper ; we're setting the context for gathering academic and industrial feedback to design an appropriate protocol for MDC communication and integration.

## 1. INTRODUCTION

Multi-Dimensional Controllers (MDCs) have the potential to facilitate creative expression in music as well as in other fields such as music therapy and social interaction. We distinguish discrete controllers - whose mapping requirements are addressed by MIDI - from continuous controllers, whose richer, higher definition data, is inadequately handled in existing musical data protocols. This paper does not propose a specific protocol, rather it presents a higher level framework for the categorisation of new sets of gestures, that could be implemented in existing protocols to enhance musical creation and facilitate software compatibility and lower the barriers of adoption.

There has been an increase in the development of MDCs, yet the method of encoding abstract musical data in computers remains grounded in keyboard-based discrete controllers. Indeed, musical expression that uses controllers is currently limited by the boundaries of the keyboard paradigm, where the note begins when the movement finishes and focuses mainly on discrete pitches and durations. In the field of music technology, professionals push the boundaries of their design, but their output is compromised by the limitations of discrete and low resolution encoding protocols, such as MIDI. On the other hand, high definition protocols, such as OSC, are too open to allow versatile and easy integration in commercial sound software. While it is becoming increasingly easy to get access

to sophisticated sound design tools, it is significantly harder for musicians to personalise and be creative with those sounds, due to the limited interaction available. In order to increase the playability of sound software, controllers require a change in form factor and new ways to encode data.

## 2. THE SEABOARD

The Seaboard is a new tangible keyboard instrument, which facilitates intuitive music creation [6]. The intent of the product design is to deliver an integrated music-creation device, which merges traditional keyboard design with modular technology. The Seaboard is the first application of the patent-pending SEA (Sensory, Elastic, Adaptive) technology.



**Figure 1.** The Seaboard

The Seaboard is an MDC which enables both continuous and discrete control through note-by-note real-time continuous polyphonic control of pitch, amplitude, and timbral variation. In contrast to previous, limited keyboard and sound production interfaces, the Seaboard diversifies and maximizes the degree and types of musical expressivity through its ability to simultaneously control continuous and discrete aspects of sound. The Seaboard is a continuous action interface (CAI), a system which registers spatial or gestural movement in time to enable more complex inputs based on continuous movement. The Seaboard's continuous functions include glissando/slide effects, timbral, and dynamic variations in real time. Its continuous control includes the ability to gather and map rich sets of data in a variety of ways. The Seaboard is also a discrete control interface (DCI), a system with inputs

which can be distinguished in time. The Seaboard's discrete functions include inputs to generate the notes of the chromatic scale. The Seaboard's discrete controls include analog (usually switch-based) controls that simulate a mechanical action.

## 2.1. Integration Challenges

One of the most significant challenges during the development of the Seaboard has been the translation of the rich output data produced by the Seaboard into a structure that third party audio software programmes can interpret. The majority of existing hardware controllers and software programmes make use of the MIDI protocol. Attempts to integrate the Seaboard with the MIDI protocol have helped to highlight some specific issues that exist with MIDI.

When polyphonic synthesizers were first being developed during the late 1970s, one major issue was the inability for different manufacturers' machines to communicate. MIDI was introduced as a solution to this problem. Although the introduction of MIDI was a major breakthrough in computer music, issues arise from the fact that it is based on the physical keyboard interface.

### 2.1.1. Limited Note Control

On a traditional keyboard interface, a note begins when the movement finishes, when the hammer hits the strings, and ends when the key is released. The users only have control over the beginning and end of a note. They have no control over the parameters of that note over the duration that it plays. The MIDI protocol extends this notion. In MIDI, there are 'note on' and 'note off' values. The volume, pitch and timbre of these notes are determined by the channel in which the notes originate.

### 2.1.2. Pitch Bending

When synthesizers were first developed, the idea of pitch-bend manifested as a controllable wheel, that globally altered the pitch output. Altering the position of this wheel would change the pitch of every note that was played on the synthesiser. MIDI was developed as a way of facilitating the communication between such synthesizers. As a result, pitch-bend, was ported to the MIDI protocol as a global variable. MIDI encodes pitch bend globally, to every note, rather than on a note by note basis. MIDI offers 16 channels, that may be used for different voicings (e.g. piano, bass, guitar). On each of these channels, there may be a certain number of notes in play at any one time. The major problem is that pitch bend is applied to the channel rather than the note. Therefore, within a given channel, you may only bend the whole group of notes, not individual notes.

## 2.2. Integrating the Seaboard

The Seaboard must produce MIDI output if it is to be easily integrated into the major existing 3rd party software programs. However, the MIDI protocol is not sufficient to represent the kinds of interaction that take place on the Seaboard. For example, using MIDI with one channel per instrument, the act of increasing the pressure or varying the location of a singular touch on the Seaboard will in fact alter the volume and pitch of all current notes being played.

This is indicative of a widespread general problem in computer music today. Namely, that it is relatively simple to create new MDCs, but it is a lot harder to integrate such systems as controllers of audio software. The next section focuses on other existing MDCs and the challenges of integration manifest for them.

## 3. OTHER MDCS

This section discusses some other existing MDCs, detailing their functionalities, the challenges they face, and their significance in the context of integration with audio software. The list is non exhaustive, and rather characterises different approaches for MDCs described later as acting, stationary and hybrid controllers.

### 3.1. The Haken Continuum

The Haken Continuum (c. 1998) is a continuous surface keyboard that offers real-time control over pitch, amplitude, and timbre [4]. The Continuum features a photoelastic playing surface lit by a single-frequency polarized light source from the underside, combining the pitch sensing and polyphonic surface abilities of the Dynamic Keyboard and the Pitch Extractor.



Figure 2. The Haken Continuum

The primary hardware feature of the Continuum is the continuous ribbon. In terms of MIDI integration, the Continuum faces the same problems as the Seaboard - namely that pitch bend and volume control are encoded in MIDI as global, channel parameters. The Haken team ended up offering MIDI and OSC, and developed a complex mapping software, the EaganMatrix. The Haken also has embedded sound. The difficulty of integrating their rich data with existing systems pushed them to develop an entire system in isolation from other mainstream music hardware and software.

### 3.2. Sound Beam

The Sound Beam is a single dimension controller that converts a continuous beam (a distance) into MIDI data [8]. The Sound Beam mapping, which converts an ultrasound beam into MIDI data, is a good example of an expressive conversion of continuous data to MIDI. However, the mapping is closed source, hardware dependent, and the result is monophonic since users control only one dimension, through interaction with a single beam of ultrasound. Most MDCs offer a number of dimensions, but extending beyond one immediately poses integration problems with software.

### 3.3. SoundGrasp

Sound Grasp is a gestural interface for the performance of live music that aims to facilitate musical control without the need for direct machine interaction [7]. Specifically, the system is designed to enable quick, performative recording and manipulation of audio samples. The interface takes the form of a wearable glove through which the user may perform one of eight gestures that are recognised by the system. These gestures are mapped to specific audio sampling processes such as record, play, reverb and, filter control.



**Figure 3.** Glove Positions Used in the SoundGrasp System (taken from [7])

Within the Sound Grasp system, a group of specific gestures are well-defined, and they are mapped appropriately to useful musical outputs. The system makes use of a custom and purpose-built audio processing unit for the recording and manipulation of audio. In essence, the Sound Grasp represents an example of a very well-developed framework for the mapping of gestural data to relevant musical output, but which has been developed within the sole context of audio sampling and recording.

### 3.4. Current Progress in Music Information Protocols

Progress has been made towards a more adaptable encoding protocol for musical data. For example, HD-MIDI, announced in 2005, is a proposed extension to MIDI that will bring many improvements. For example, included in the specifications for HD-MIDI is the addition of new note parameters. HD-MIDI will also bring the ability to bend individual notes [3].

The VST3 (Virtual Studio Technology) from Steinberg, launched in 1996, integrates virtual effect processors with instruments in a digital audio environment. VST3 facilitates more than one MIDI input and output at a time [2].

CopperLAN is a communication system that facilitates the interfacing and linking of professional audio and musical equipment. It features 'full auto-setup, plug and play, universal remote editing, total setup recall, streamlined and unified user interface methods, clear and concrete identification of devices' [1]. However, its major drawback is its cost, which is a barrier to adoption.

## 4. TOWARDS AN MDC FRAMEWORK

The development of the MDC framework requires the categorisation of abstract MDC gestures, and the identification of appropriate mappings from these gestures to relevant musical output. Intuitive gestures, drawing on predictable reactions experienced in the real world, should be the fundamentals of this framework. For example, in traditional musical practice, an increase in force generally results in an increase in volume, and a rapid oscillation of position generally results in a vibrato effect. The MDC framework should attempt to match high energy gestures to high energy changes in sound, and subtle movements to subtle changes in sound. MDCs not only offer control of pitch and duration, but facilitate unprecedented control over dimensions of sound such as timbre and spatialisation.

### 4.1. MDC Types

The MDC framework should take into consideration non-haptic as well as haptic controllers. As well as haptic and non-haptic, an important distinction can be made between what we call acting MDCs, stationary MDCs, and hybrid MDCs. Acting MDCs involve a device that physically moves, and outputs information about its position and orientation. A large portion of acting MDCs are non-haptic (e.g. MYO). Stationary MDCs are those that output information about a user's interaction with a device. For example, the Kinect outputs information about the user's location in 3D space, and the Seaboard outputs information about the user's touches on the surface of a physical device. An example of a hybrid MDC is a smartphone, that outputs information about the user's touches on the screen, as well as information about the orientation of the device.

The distinction between acting, stationary and hybrid is separate from the distinction between haptic and non-

haptic. Any device can be either haptic or non-haptic, regardless of whether it is a stationary, acting, or hybrid MDC.

#### 4.2. Data Flow

Figure 4 shows the typical data flow of a (haptic) MDC system. Firstly, in the 'Musician/Hardware' block, physical interaction takes place between user and system and this is recorded by the system's sensors. Next, in the 'MDC : Software & Firmware' block, the sensor data is collated together to extract 'events,' which are then processed and mapped to output data. In the 'Digital Music Production Software' block, data is fed into a 3rd party software sound engine, and sound output is generated (the 'Music' block). The rest of this section will discuss the 'MDC : Software & Firmware' block in more detail.

#### 4.3. A Distinction of Gestural Parameters

The first step in developing a MDC framework that facilitates the integration of hardware MDCs with sound software will be the identification of a set of abstract parameters that can be grouped together to define any conceivable MDC.

Given the set of distinctions between MDC devices set out earlier in this section, we present a preliminary suggested list of parameters, grouped under 'event' (for haptic MDCs), 'object' (for non-haptic MDCs) and 'device' (for acting and hybrid MDCs). Firstly, these different parameters will be presented and then a general explanation of how gestures can be processed using these parameters will be given.

##### 4.3.1. Event Parameters

In the case of a haptic device, user interaction with the device will usually consist of the user manipulating the device's sensors over time, through some physical, gesture-based interaction. Individual values for each sensor are stored, giving rise to data with the following structure : Sensor 1 [Value], Sensor 2 [Value], ... ,Sensor N [Value]. The data at each sensor, when taken collectively, gives rise to an instantaneous 'frame' of *events*. Each of the events in a frame have values for the following parameters :

- ID Tag
- Onset Timestamp
- Onset Magnitude/Size
- Onset X-Location
- Onset Y-Location
- Onset Z-Location
- Event Magnitude/Size
- Event X-Location
- Event Y-Location
- Event Z-Location

The ID tag is used so that specific events can be tracked over time, frame-by-frame. The 'onset' parameters relate to the state of an event during its initial frame (its initialisation). The 'event' parameters define the instantaneous state of an event, from frame to frame. The continued comparison, between some present frame and the previous frame, gives rise to specific gestures. For example, the euclidian distance in 3D space between an event's current location and its location a number of frames previously, defines a movement in a certain direction, at a certain velocity.

##### 4.3.2. Object Parameters

In non-haptic devices, interaction typically takes place through the manipulation of identifier *objects* (e.g. user limbs/wii remote) in 3D space. Much in the same way that *events* have specific instantaneous parameters, objects have similar parameters :

- Object ID
- Object X-Location
- Object Y-Location
- Object Z-Location
- Object Orientation
- Object *State*

Objects are tracked, frame-by-frame, in a similar way that *events* are tracked. The orientation parameter is used to track rotary movements of specific objects. Relations between an object's current and previous orientation gives indication of yaw, pitch and roll gestures.

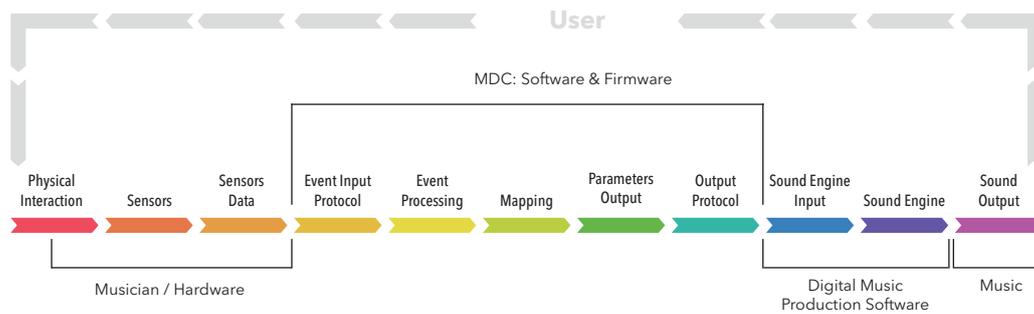
An object's 'state' refers to its configuration. Interactions with non-haptic, object-tracking systems, often include object-specific gestures, where objects go from being in one state/configuration to another. For example, an open hand may close into a fist, which would be recognised by the system as a meaningful gesture. The specific possible configurations of objects will depend on the specific system, and the identifier objects that are used. For example, in the SoundGrasp system the objects are the user's hands, and the hands can be in any of 8 specific 'states' (see Figure 3).

##### 4.3.3. Device Parameters

For acting and hybrid MDCs, parameters are required that describe the location and orientation of the device itself. These can be defined as :

- Device Orientation
- Device X-Location
- Device Y-Location
- Device Z-Location

Device-specific gestures can be tracked by tracking the differences in these parameters from frame to frame.



**Figure 4.** Data Flow in a Stationary Haptic MDC

#### 4.4. Mapping

Having set out a list of parameters that may be used by different types of MDCs to track gestures as they occur, the next key step in facilitating effective gestural control of sound is to define effective mapping strategies such that gestures produce relevant, intuitive sound output. The design of an appropriate and effective mapping strategy is one of the most important issues in the development of new MDCs. This issue of mapping is influenced by three key factors : real-time constraints, dimensionality, and customisation.

##### 4.4.1. Real Time Constraints

MDCs need to work in real-time, with minimal latency between user action and resulting sound output. In order to achieve minimal latency, the number of previous frames kept in memory must be small - usually only 1 or 2. This means that as gestures occur, systems only have access to instantaneous moment-by-moment segments of them.

This is an important restriction, given the ‘temporal precision’ with which musical performers can perform gestures. Schmeder et. al (2010) give an explanation of the term temporal precision [9]. Temporal precision refers to the smallest time interval over which humans have control in terms of musical gestural performance. Using the example of the ‘flam’ technique from drumming, Schmeder et. al report this smallest time interval to be 1 millisecond.

Therefore, in order for gestures to produce meaningful sound output, systems must make use of relevant mapping strategies between event parameters and resulting output data. For example, by mapping instantaneous x-location to pitch, smooth movements over the x-axis will effectively translate as smooth glides in pitch.

##### 4.4.2. Dimensionality

Mapping strategies can exist as one-to-one connections, one-to-many connections, or many-to-many connections. Hunt et. al (2002) performed an experiment to evaluate the effects that different types of dimensionality mappings

had on users [5]. They found that users learned the performative limits of one-to-one mappings very quickly, but that this lead them to perceive a sense of restrictiveness in the system. Conversely, many-to-many mappings were harder to learn, but this gave users a sense of performative potential that they could spend time mastering.

##### 4.4.3. Mappings and Customisation

Although parameters must be mapped to relevant sound outputs, there is a question as to whether or not all mappings should be permanent.

Permanent mappings help to establish MDCs with their own performative identity, and can help foster intimate relationships between user and system. In the realm of musical instruments, it is through the struggle with an instrument’s inherent restrictions (and permanent mappings) that true creative innovation occurs.

However, one of the great advantages of new MDC systems is the level of personalisation and customisation they facilitate. Since there is no physical connection between gesture and sound output, users could be free to map *any* gesture to *any* sound output. Indeed, this could potentially open up a whole new paradigm for user creativity with digital musical instruments. As well as exploring the new performative potential offered by digital musical interfaces, users could start creating and sharing new performative possibilities, by experimenting with user-defined mappings. An analogy in the realm of traditional musical instruments is the use of a violin bow to bow the strings of an electric guitar. Allowing users to experiment with new mappings in the digital domain should be facilitated at the front end of MDCs, and shouldn’t require an understanding of the whole protocol chain. Detaching the mapping discovery from the constraints of low level physical and digital inputs and outputs is integral to our proposition.

#### 4.5. Output Data - Beyond MIDI and OSC

There are specifications available that could be used to encode the MDC state and gestural controls in terms of musical parameters. For example, MIDI or OSC. The main problems with MIDI have been described many times

before, but it appears useful to rephrase them here in light of the MDC requirements. Firstly, MIDI is based on the piano paradigm in which notes exist as discrete events with on/off values. Secondly, MIDI encodes pitch bend and volume as global parameters, rather than on a note by note basis.

OSC addresses the limitations of MIDI in many ways : it is high definition, low latency, flexible and multi-dimensional. Schmeder et al. (2010) suggest to use ontology-oriented description of events, and suggest that the naming of these events should originate from the application designers [9]. But overall, the lack of a prescribed ontology for OSC messages, and the requirements at the physical level (Ethernet), appear to be obstacles to a mainstream adoption. We argue that, for a music protocol to gain adoption, it has to offer a common definition of discrete and continuous events, notes, volumes, pitches and durations and also timbre or spatialisation.

#### 4.5.1. Note Events

The note events consist of the data that is visible to the sound engine. In contrast to the *gestural event* parameters described earlier, the *note event* parameters are a result of the mapping and possible customisation happening at the hardware and firmware levels. For the MDC framework to overcome MIDI shortcomings, it has to offer polyphonic control over individual note parameters, during the life-cycle of notes. Within each individual note event, changes in timbre, pitch and volume need to be available in real time and high definition to the sound engine. The definition of these note events will progress through consultations with audio software developers, artists and pro users.

Given the groundwork already covered by the MIDI protocol, an important question in the development of the MDC framework is whether the MIDI specifications should be extended to allow for richer data, or whether a whole new protocol should be developed. More fundamentally, we argue that the MDC framework, possibly in the form of an API, should aim at transcending the transport layer (USB MIDI, TCP/IP or Bluetooth) to focus on facilitating the end to end communication with comprehensive and meaningful gestures to produce rich, expressive, polyphonic sounds.

#### 4.6. Framework Requirements

The success of this new software protocol design relies on collaboration between its three primary user groups : artists (end users), hardware manufacturers, and audio software developers.

In order to gain support and participation for this new framework within (a) hardware manufacturing, (b) software developers and (c) artist communities, an open approach must be taken towards IP and cost. Developers and manufacturers will only be willing to get on board with the framework if they can implement it openly and freely within their products. The inability to integrate novel inventive MDCs with existing software packages is a major

problem, and the only way to reach a widespread, sustainable solution is to make that solution open-source, adaptable, and understandable.

The end users of MDCs and audio software will come into varying levels of contact with the framework, depending on their interest in configuration, personalisation and experimentation. The framework, therefore, needs to be easily readable and usable to end users of music hardware and software.

## 5. CONCLUSION

This paper has outlined the difficulties in integrating multi-dimensional controllers (MDCs) with audio software programmes, through the specific example of the development of the Seaboard and its integration using the MIDI protocol. A proposal has been put forward for a new MDC framework that focuses on gestural control and relevant musical output. The key issues involved in the development of this framework have been discussed.

The goal of this proposal is to state ambitious goals rallying the communities of new instrument makers, audio software developers and musicians eager to bridge the existing gaps between hardware and software worlds, and produce more meaningful, intuitive controls for sound production. The proposed MDC protocol would provide a common language for developers and manufacturers to quantify, discuss, share and integrate their gestural controls and mappings. The development of this new protocol will be best facilitated by open discussion and participation by those currently invested in the development of MDC systems.

## 6. ADDITIONAL AUTHORS

Roland Lamb  
ROLI  
roland@weareroli.com

## 7. REFERENCES

- [1] <http://www.copperlan.org/index.php>.
- [2] [Http://www.steinberg.net/en/company/technologies/v](http://www.steinberg.net/en/company/technologies/v).
- [3] C. Anderton. MIDI : The Next 30 Years. *Electronic Musician*, page 12, Feb. 2013.
- [4] L. Haken, R. Abdullah, and M. Smart. The continuum : a continuous music keyboard. In *International Computer Music Conference*, page 81, 1992.
- [5] A. Hunt, M. M. Wanderley, S. S. West, and M. Paradis. The importance of parameter mapping in electronic instrument design. In *New Interfaces for Musical Expression (NIME)*, pages 1–6, 2002.
- [6] R. Lamb and A. N. Robertson. Seaboard : a new piano keyboard-related interface combining discrete and continuous control. In *New Interfaces for Musical Expression (NIME)*, pages 503–506, 2011.

- [7] T. Mitchell and I. Heap. SoundGrasp : A Gestural Interface for the Performance of Live Music. In *New Interfaces for Musical Expression (NIME)*, number June, pages 465–468, 2011.
- [8] S. Paul and D. Ramsey. Music therapy in physical medicine and rehabilitation. *Australian Occupational Therapy Journal*, (47) :111–118, 2000.
- [9] A. Schmeder, A. Freed, D. Wessel, A. Street, and B. Ca. Best Practices for Open Sound Control. In *Linux Audio Conference*, 1997.



## AN EVOLVING COLLABORATION : PERFORMER AND COMPOSER APPROACHES TO CREATING VISUAL MUSIC

*Nicole Canham*  
CICM - EA 1572  
University of Queensland, Australia;  
MSH-Paris Nord  
n.canham@uq.edu.au

*Carlos Lopez Charles*  
CICM - EA 1572  
Université Paris VIII, France;  
MSH-Paris Nord  
carloslopezcharles@gmail.com

### ABSTRACT

This article aims to identify the key factors that contributed to an evolving collaborative practice between a performer (Canham) and a composer (Lopez Charles). Three new works for clarinet, electronics and video, presented as case studies, form the basis of this study. Journals, artifacts including sound recordings of experiments, rehearsals, performance documentation and joint reflection will allow the researchers to describe and reflect upon the evolution of their collaborative practice as it unfolded.

Nicole Canham, clarinetist, is a PhD candidate at Queensland University (Australia), researcher in residence at the Maison des Sciences de l'Homme, Paris Nord, working in the area of qualitative research methods applied to understanding aspects of musical practice. Carlos Lopez is a PhD candidate at Paris VIII University, composer and musicologist who specialises in composition with new technologies.

### 1. INTRODUCTION

Dannenberg suggests that the use of computers in music composition “leads to new ways of thinking about music composition and performance” [5]. In this project, two participant/researchers set out to document and observe their own collaborative process in the context of the development of a new work for clarinet, electronics and video which was based around interactive components facilitated by the use of computer technology. The aim was to track and identify key factors – including new ways of thinking – that emerged during the process of creating three visual music works.

### 2. THE STUDY OF COLLABORATION

Collaboration has been a subject of study in numerous fields including business [14], the arts and sciences, [13] in improvisational contexts in music and theatre [18] and in the field of performer/composer collaboration [see for example 6, 7, 11 and 12]. Performer/composer studies have addressed issues of the nature of collaboration – which are often framed in terms of ‘process’ – and the outcomes of the collaborative work environment – which

Hayden refers to as ‘output’ [11]. Analysis and observations about collaboration, subsequently, can be made from a variety of angles which include detailed descriptions of processes involved [13], or a focus on the employment of a collaborative approach with a certain outcome or output in mind.

John-Steiner suggests that “construction of knowledge is embedded in the cultural and historical milieu in which it arises [13]”. Her view provides an interesting challenge in the study of performer/composer collaborations in so far as these roles have often been perceived to be quite separate. In the western art music tradition, it has been argued that this segregation of roles is a product of the view of the score-as-object, an idea attributed to German music critic E.T.A Hoffman [8]. Hoffman’s notion of ‘Werktreue’, first proposed in the early 1800s, introduced a concept of musical works in which all other aspects of the presentation should be subservient to the score: “musical activities, be they of composition, performance, reception, evaluation, or analysis, should no longer be guided by extra-musical considerations of a religious, social, or scientific sort. They should now be guided by the works themselves [8]”.

The nineteenth century view of composition and performance has endured in two key ways. Both the direction and evolution of practice have been influenced. One of ways in which this influence can be perceived is in the limits the *werktreue* concept has placed on the musicological discourse [8]. A preoccupation with the score has shaped the scope and focus of scholarship, with many other aspects of the compositional and music making process being overlooked. This has included a narrow view of the role the performer [19] and, until recently, limited contribution from performers in academic research [4]. A second consequence of a score-based approach has been the evolution of practice to reflect this emphasis, one example of which is the lengthy period in which improvisation for classically trained performers was not in vogue [19].

In recent years, however, many of the limitations of a nineteenth century view of composition and performance have been addressed or challenged [2, 8, 9, 19, 22]. One advantage of the new directions in thinking and practice

proposed has been the examination of a range of creative and artistic roles and relationships. Increasing research based upon a reconsideration of the work of the composer, the view of the performer and the opportunities offered by the addition of computer technology in the composition and performance workspace provide examples of a shifting dialogue.

### 3. CHALLENGING TRADITIONAL ROLES

A common theme in recent research into the activities of composers and performers has been the unmasking of processes that were not previously viewed as central to a discussion about creating new works. Studies of collaboration have been one area of scholarship in which more nuanced pictures of performer and composer have been proposed. These studies also often reflect consideration of other elements influencing collaborative environments, including the use of digital technology.

With regard to composers, Hayden and Windsor [11] highlight the inaccuracy of the commonly held view of the composer as a lone genius when examined in the context of the creation of new works. Rather, they suggest that composers have always had some degree of interaction with performers, conductors or publishers, but that this has often been neglected in the discourse.

Performers, according to Crispin, [4] negotiate a reality in which many aspects of their practice are deliberately kept half-concealed: the years of training and discipline required in order to develop the connection between performer and instrument is often hidden in the desire to make performances appear effortless. Crispin suggests this reveals a preoccupation with the ideal at the expense of developing a “deep understanding of live performance in all its guises [4]”.

Digital technology has also been recognized as a key compositional and collaborative tool in the twenty first century. Dannenberg, commenting on the rapid growth in capabilities of digital technology and the opportunity these present in creative contexts, suggests that “the biggest challenges ahead are artistic rather than technological. One of the attractions of this pursuit is that there are relatively few precedents and no established theory [5]”.

In the absence of a single established theory, a range of ideas have been proposed in understanding the role of digital technology in artistic collaborations. Within the domain of music making, some studies have focused on the way technology has been used as an instrument for artistic practice – Partti’s study of the ‘digital musician’ is one example [16]. Digital technology has also been harnessed to facilitate the sharing of information among musicians through web-based communities of practice. [17] Literature concerning the design of computer-supported co-operative work systems (CSCW) addresses the influence of social factors on systems design. [3] One

aspect of CSCW which is pertinent to this study is the notion of tailorability. [21] Although Dannenberg [5] emphasizes that creating visual music using technology is essentially an artistic challenge, questions of tailorability - how system design decisions are to be made, and who is to make them - remain an issue. These three elements of performer, composer and digital technology are focal points of this investigation of aspects of collaboration.

### 4. RESEARCH AIMS

#### 4.1. Research questions

The study was built around a central question: What were the key factors that shaped this collaboration?

These factors included observation and consideration of the facilitators and constraints inherent in the study, which included:

- Practical and creative matters, including issues of systems design;
- The nature of the working process, identification of patterns of decision making, evaluation and discussion; and,
- Discussion of the similarities and differences between performer and composer objectives as the collaboration progressed.

In acknowledging that the project united artists with different backgrounds and approaches, there were a series of sub-questions for both researchers which related directly to their own questions about their role in a collaborative work environment.

#### 4.2. Composer-related questions:

How can different compositional approaches open up space for contrasting ways of collaborating with a performer? How can these approaches influence the relationships between sound and image in the creation of visual music?

#### 4.3. Performer-related questions:

What makes a performer of an acoustic instrument necessary in the electroacoustic and visual framework?

To what degree can a performer make connections between visual and musical elements if they have no visual score or cannot watch the video during the performance?

To what degree does or should the performer have a role in the creative (compositional) process?

## 5. APPROACH

In this study, the authors and co-participants brought different areas of expertise to the research. Nicole Canham has commissioned numerous new works in her professional career as a performer, and has collaborated with a wide range of artists in diverse settings. Canham's doctoral research draws upon qualitative research methods, in particular that of case study. Carlos Lopez Charles is a composer whose work has focused mainly on electro-acoustic composition and computer programming. His current Phd study in the area of visual music is concerned with the compositional techniques in this field. The study presents three different examples of composer-performer collaboration which took place over a two-year period. During this time three new works for clarinet, video and electro-acoustic music were created. The creation of each work, presented as a case study, reflects approaches to collaboration shaped by a variety of factors including the participants' varied expertise. The aim of presenting two cases of earlier collaborative works prior to the current project is two-fold: as a form of triangulation, and to provide a framework for evaluating to what degree the collaborative process evolved between our first and our most recent attempt.

Willis [23] suggests that what we believe to be the "nature of truth (ontology) and what it means to know (epistemology)" (p.10) forms the basis of the ways in which research can be conducted and understood. Given the research aims, which are focused on gaining a greater understanding of multiple perspectives of the same situation, the ontological position of this study is that there are multiple versions of reality rather than a single truth, locating this study within the constructivist paradigm [1, 10, 24]. Reality, then, is constructed through the accounts and reflections of the researchers who were co-participants in the project [10, 23, 24]. The research design reflects a close connection between the authors' researcher roles and their roles as participants as they worked together to co-construct an account of their evolving collaborative process.

### 5.1. Methodology and Methods

Collaboration has been studied using a range of qualitative research methods, which have included interview [13], observation [12], video interaction analysis [18] and reflective journal [11]. Case study is a research methodology that is often used in research based upon a constructivist paradigm, and has also been applied in studies of collaboration [10, 11]. One of the advantages of case study is that it enables the researcher to employ a variety of different methods in the gathering of data [25], which includes those methods stated above and other artifacts. In the case of this study, artifacts included meeting notes, task-based activities, scores with annotations, email correspondence, and recordings of rehearsals and performances.

### 5.2. Case Study

There are several different types of case study that researchers utilize, which Stake describes as intrinsic, instrumental and collective [20]. Willis [25] offers two further sub-categories which are perhaps more reflective of the ways in which the data is expected to be used: "descriptive case study [25]" is concerned with detailed descriptions of cases. "Interpretive case study" takes the descriptive case study work further, aiming to use the data generated to make statements about contexts beyond those being documented and described. "The focus is on understanding the intricacies of a particular situation, setting, organizations, culture, or individual, but that local understanding may be related to prevailing theories or models [24]".

Instrumental case study, like Interpretive case study, enables the researcher to develop detailed accounts of individual cases, but the purpose of the study is to use the case to provide insight about an issue.

This instrumental case study describes the creation of three new works in order to understand the nature of each collaborative context on an individual basis. These cases will then be considered together with a view to developing insights about the outcomes of an evolving collaborative process. Three examples of the creation of different works are used to illustrate ways in which creative and technical approaches to visual music from composer and performer perspectives evolved.

## 6. CASE STUDY 1 – *NOT ALONE*, FOR CLARINET, LIVE ELECTRONICS AND VIDEO

*Nicole*: Reading back over our emails at the time *Not Alone* was being written, I notice that the conversation had a lot to do with practical concerns. I liked the piece a great deal on first hearing and nothing in the score was technically problematic, so my feedback focused on page turns, breathing and some articulation issues. *Not Alone* was my first foray into working with live electronics in a surround sound environment. Perhaps as a result of this, the score Carlos provided me was highly detailed to ensure that the clarinet input would make the delay effect respond as envisaged.

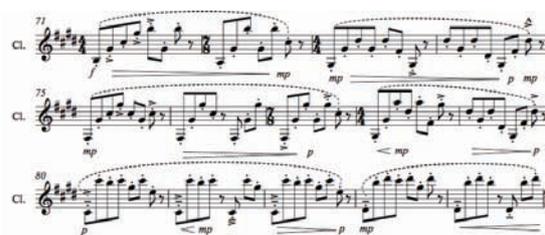


Figure 1. Score excerpt, *Not Alone*

The video that accompanied *Not Alone* was created by a colleague after the music had been composed, almost as a creative response to the music.

In a way, it seems as though we were trying to superimpose different ways of working over the top of each other. I commissioned *Not Alone* as part of a multimedia installation/performance piece which involved a larger team of Australian and Mexican artists. *Not Alone* was a collaboration-within-a-collaboration in that sense, as I was also developing a cohesive 60-minute long program. I wrote to Carlos about this on 4 May, 2011, about three weeks before the premiere:

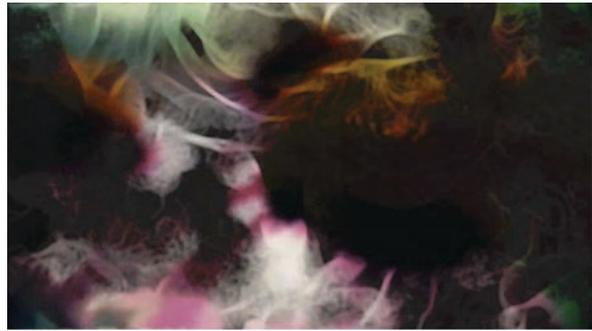
“I included a quote about time for each piece that inspires me on the musical interpretation side of things. In terms of the order, I've suggested it considering three different perspectives:

1. Ease of change of instruments, because I'm changing as we go;
2. Key [tonality] relationships between pieces so there isn't a jarring sensation going from one piece into the next;
3. Narrative possibilities of music in that order.”

Conceptually, bringing in the time quotes as my own interpretive aid had to do with the ‘time’ theme of the larger program. It was, however, quite different from the musical frame of reference that Carlos had first discussed with me. He was interested in delay effects used by electric guitarists in the 1980s, and was keen to find a way to transfer this idea into a piece for clarinet. Carlos was also left to consider all the technical matters that I didn't understand, or perhaps hadn't thought through coming from a background of mostly acoustic performance. Carlos' system design had a reactive function to performer and composer input: my sound was processed with a delay feedback effect and Carlos varied the volume and the spatialisation of the delayed clarinet sounds. In performance, I felt as though my playing and his live diffusion of the sound were similar to making chamber music, although I never experienced the full perspective of the spatialised performance.

## 7. CASE STUDY 2 – LAS FLORES Y LAS NUBES, FOR CLARINET, ELECTRONICS AND VIDEO

*Carlos:* My goal for this piece was to have a higher degree of integration between video and music than in *Not Alone*. Early discussions with Nicole focused on the technical aspects of how I wanted to approach this challenge. I also made a point of working with Nicole using improvisation as a basis for testing and developing ideas. We had two improvisation/recording sessions before a period of time apart.



**Figure 2.** Still image from *Las flores y las nubes*

Although I intended to take a less conventional approach with this piece, I still made most of the compositional decisions on my own. I composed the video first and then the electroacoustic tape part, trying to make an audiovisual piece that could stand on its own and over which I could then add a clarinet part. To make the video, I used a custom-made program that allowed me to vary the size, movement and color of up to 12,000 points. Using this idea as a model for the composition of the tape, I created a patch to control different variables of a flow of short (15 ms. to 200 ms.) electronic sounds produced by a very sophisticated FM-based synthesizer.

At this stage, I was consciously avoiding the incorporation of clarinet sounds into the tape because I felt that approach (which is quite common) to be too standard. The Max/MSP patch that I made for the clarinet was programmed so that it changed its variables automatically at specific moments of the piece which meant that synchronization of parts was critical. For this reason, I included a click track that would enable the performer to synchronize with everything else very precisely. We didn't see each other after our improvisation sessions until 3 weeks before the premiere of the piece, so the clarinet part was completed by me and then discussed and revised in our final rehearsal period. I sent a very detailed score, along with a patch, a video and a click track to Nicole six weeks prior to our planned rehearsals. As it turned out, none of the work from our improvisation sessions was included in the finished score.

In our first rehearsal, I realized that the clarinet's sound did not blend as well as I wanted with the electronic sounds that I had chosen. I decided to integrate some transformed clarinet sounds into the tape part that would allow for a tighter connection between tape and clarinet. Adding these sounds, using a reverb effect and making minor changes to the score helped to achieve a higher degree of integration between the clarinet and the electronic part. We eventually used improvisation for the ending of the piece.

*Nicole:* *Las Flores*, like *Not Alone*, was premiered within a larger multimedia performance. The context of the premiere also reflected some of the broader technical

and programming challenges that we were negotiating. This included practicalities of how the video was to be projected, the projection surface (black), the style of projection (via Matrox) resulting in a wrap-around style image, and the low level of light necessary in order to have the videos look most vibrant.

The technical and programming concerns surrounding the performance as a whole also impacted staging options. Although I performed almost the whole program from memory, I had to perform *Las Flores* with music in a fixed space so that I could access the click track to ensure that my part and the electronic part were together in performance. In order to accommodate the video requirements of the piece, I was obscured from view by some large pieces of black foil blocking the light coming from the scone on my music stand. Transitions from one piece to the next were also critical in this performance. Carlos had left the ending of *Las Flores* quite open which gave us the opportunity to weave together a number of elements. This included some improvisation in performance on my part for the ending of the piece, which also incorporated mechanical instruments made by an artist-colleague. Whilst I felt sonically embedded in this piece, I didn't visually feel a part of it and it was difficult for me to be seen by the audience.

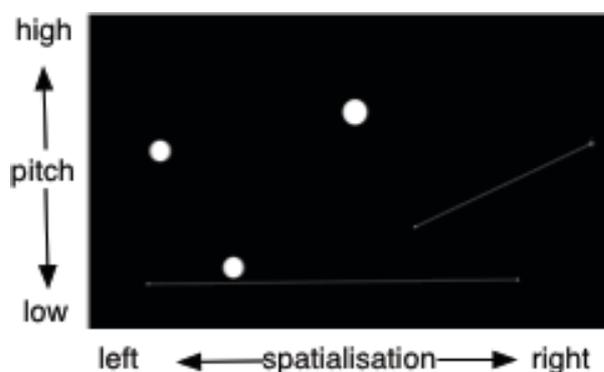
### 8. CASE STUDY 3 – POINTS, LINES, PLANE, FOR CLARINET, LIVE ELECTRONICS AND LIVE VIDEO

*Nicole & Carlos:* We had originally planned to re-work *Las Flores*, and so we set aside some time for discussion and experimentation. This included individual and joint reflection on what we felt worked well, and what might be improved based upon our experiences of both *Not Alone* and *Las Flores*. Feedback from Carlos concerned finding a better blend with clarinet and tape. Feedback from Nicole was concerned with having a better visual connection between performer and video, and exploring more meaningful ways of connecting with the visual element as a performer. Feeling able to make connections between the performative aspect of the visual component and other visual elements, namely the performer as visual component, was impeded by the necessity of performance with score and click track. We were also keen to set some new challenges and parameters for this collaboration. We wanted all the elements to be able to interact, as a basis for creating a collaborative environment where we would challenge our traditional roles as performer and composer. The goal was to develop a working process which might enable us to foster a collaborative environment where roles felt equal and to some degree interchangeable.

*Nicole:* One of our conversations stands out to me – I had been discussing with Carlos some of the problems I find with using a click track and he suggested, “what if

we get rid of the click track?” We then moved on to considering how we might structure the score and this led to the thought... “what if we get rid of the score?” These two questions for me, and the possibilities opened up by leaving both click track and score behind, were the beginning of a completely different approach, and a new piece.

*Carlos:* Reflecting on my discussions with Nicole led me to realize that the issues that we were trying to address (in particular those of *music-video* and *performer-video*) could be solved by establishing interactions between electronic sounds, instrumental sounds and moving images through the use of the computer. Inspired by Josh Nimoy's “Bouncing Balls”, I created an interactive system based on the idea of balls that make electronic sounds each time they bounce against lines drawn by the user. I programmed a patch in Processing to control the moving images and used Serge Le Mouton's *sampler~* object in Max/MSP to play the audio samples each time a ball bounced. When bouncing against a line, the horizontal coordinate of a ball is mapped to a sound's position in a two-channel (left right) stereo field, while its vertical coordinate is mapped to its pitch (high to low). The communication between Max and Processing was established through the use of OSC messages.



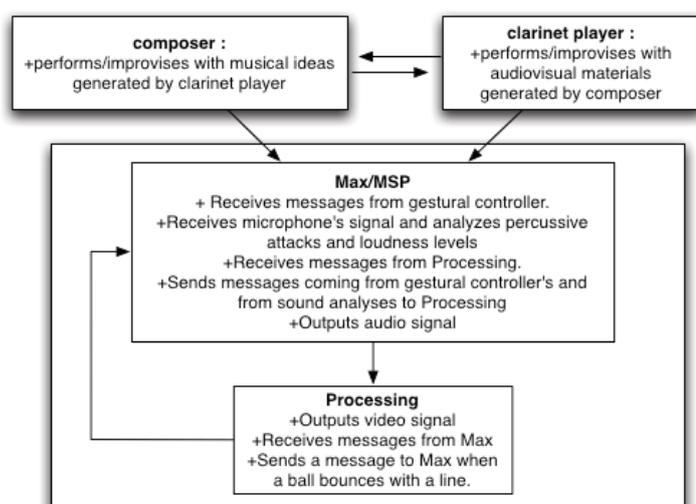
**Figure 3.** Mapping of the balls' coordinates when colliding with a line

In a way, this system was already offering an answer to the sound-image integration in the piece, but the relationship between the performer and the visuals still had to be addressed. In order to do this, I used Miller Puckette's *bonk~* Max/MSP object to detect the clarinet's percussive attacks and Tristan Jehan's *analyzer~* object to track its loudness.

*Nicole:* Our first session working with Carlos' new system gave me a lot of ideas in terms of how the clarinet and the video might relate in an interactive setting, particularly with the use of my instrument as controller. Visually, I was very taken with being able to trigger pitched, bouncing balls with my clarinet sounds. A secondary task, and one that came out of these sessions, was then to consider how I might sonically relate to the sound being produced by the balls once I

had triggered them. This included my own ‘homework’, which was to develop techniques that in some ways mimic or approach the sound that the balls produce.

*Carlos:* A challenge of the instrument-as-controller approach is that it seemed to me to have some compositional and technical limitations. Incorporating a physical interface with different multi-slider controllers that would also let *me* manipulate the behavior of the balls in a more flexible way seemed to be a good solution. Using the multi-slider allowed me to improvise with Nicole and create a wider range of musical ideas than if I just let the system react to her input. It also gave me the possibility of an interpretive role in the piece, not only a compositional one.



**Figure 4.** System design for the performance of *Points, lines, motion*

## 9. ANALYSIS

Our framework for analysis of these descriptive accounts draws upon the work of John-Steiner, Weber and Minnis in two ways [14]. Firstly, we draw upon their definition of collaboration:

The principals in a true collaboration represent complementary domains of expertise. As collaborators, they not only plan, decide, and act jointly, they also *think together*, combining independent conceptual schemes to create original frameworks. Also, in a true collaboration, there is a commitment to shared resources, power, and talent: no individual's point of view dominates, authority for decisions and actions resides in the group, and work products reflect a blending of all participants' contributions [14].

We also adopt John-Steiner, Weber and Minnis' approach of “looking for commonalities and differences across settings, tasks, working methods, goals, and

values” [14] as a tool in understanding and interpreting our collaborative process.

## 10. DISCUSSION

In reflecting upon the ways in which changes to our collaborative approach facilitated or constrained the development of an equal, interactive, creative working process we identified some key influential factors. These resonate throughout the Case Study descriptions offered here. We were able to observe similarities and differences from case to case in three main areas:

1. Challenging traditional roles
2. Changes to practice
3. Reflection: individual reflection and joint discussion

### 10.1. Challenging traditional roles

Changes of approach to our roles within the collaborative environment highlight advantages and problems with the notion of true collaboration as outlined by John-Steiner, Weber and Minnis. Case Study 1 provided an example of a fairly standard commissioning arrangement. Within that framework, roles of composer and performer were essentially quite separate and the collaborative aspect of the work is best described as a ‘layering’ of these elements to produce the finished piece. Advantages of this approach within the context of *Not Alone* included differing concerns and levels of expertise that needed to be taken into account for both practical and artistic reasons. The performer, for example, was more than happy to defer to the views of the composer in relation to the presentation of the score and the live electronics environment because that was a new area in which she had limited experience. This was also the first time the artists had worked together. Limitations of this approach impacted the artistic outcome in the sense that the relationship with performer, electronics and video did not always appear to be connected in the performance of the piece.

In Case Study 2, our roles had begun to shift. One advantage of this shift was that discussion and early work sessions reflected a desire to let performer and composer into the domain of each other's practice. Interestingly, and perhaps unsurprisingly, the direction and timing of these conversations followed the development of the piece in relation to composer and performer roles – that is, early conversations had a lot to do with Carlos' compositional ideas and approach, and later discussions just prior to, and after the premiere were more concerned with the technical challenges faced by Nicole and Carlos in performing the work. A limitation faced in Case Study 2 was the incompatibility of the timelines of composer and performer as they are traditionally understood: often a work can be almost complete before a performer becomes involved in a practical sense. Additionally, planning, development

and rehearsal environments present different challenges to a live performance context which usually takes place later in the collaborative process, rather than at the beginning, for example.

Case Study 3 reflects a different approach again in that both performer and composer perspectives were included in the planning and discussion from the beginning, with systems design, compositional ideas and approaches to interpretation being given more consistent consideration. Advantages of this level of integration compared with the other two case studies are that both composer and performer felt more able to interact across the three key areas of the collaborative partnership. No one is excluded from a particular artistic relationship combination in this format. A limitation of this approach is that the amount of video and audio processing that can take place in real-time is determined by the system's computing capabilities [15]. This creates a situation in which the video and audio materials have to be designed within the system's constraints for the sake of a higher degree of interactivity. From the composer's perspective, this is not necessarily a drawback, given that composing interactive visual music is not only about combining sounds and images, but also about composing the relationships that will be established between them and the performer(s). However, reflecting on how to minimize the compromise between what a system can do in real-time and what the artist would want it to do is an important factor to consider in contexts like this.

## 10.2. Changes to practice

There were four notable changes to practice across the three case studies. Face to face interaction, joint decision making and regular and repeated collaborative work facilitated a number of changes to individual performer, composer and collaborative working approaches. Choices in the way in which computer systems were designed and employed also played a significant role in the evolution of a sense of joint practice.

### 10.2.1. Face to face interaction

Face to face interaction became an increasingly important factor in facilitating the development of new working processes between performer and composer. This has not only been the case for joint experiments, but also for determining the nature and structure of our independent work. Increased awareness of each other's individual skills and expertise brought about through our joint creative work has played a major part in shifting priorities reflected in the decision making and planning processes, allowing for the development of both more nuanced roles for composer and performer, and a reflective approach to systems design.

### 10.2.2. Joint decision making

Composer and performer jointly agreed what they would independently work on in Case Study 3. Resulting changes to practice have included more deliberate division of tasks and responsibilities, defining together pre-compositional material and determining parameters for performance. A fluid approach to systems design has greatly facilitated the conceptual shifts evident through these three cases of collaboration: the system has become an interface between composer and performer in which new roles, approaches and artistic outcomes have been defined and facilitated. In contrast, the approach taken in Case Study 1 was much less concerned with details of the role and tasks of the other with relation to the use of digital technology: it was designed by one artist for a specific purpose, and presented to the other without a great deal of prior discussion or consultation.

### 10.2.3. Regular and repeated collaborative work

The opportunity to undertake three different collaborative projects in a relatively short space of time (less than two years) is viewed by both participants as a significant factor in the development of their joint collaborative practice. In that sense, the case studies form a single example of an ongoing, evolving collaborative practice. The possibility of the collaborative 'next step' in the form of ongoing performance opportunities has been a critical factor – from both research and artistic perspectives – in providing an outlet and incentive in which to explore and implement changes to practice.

### 10.2.4. Changes to systems design/tailorability

Another observable feature of the evolving collaborative process is demonstrated in the systems design choices and/or use of digital technology in each case. In addition to discussions about the artistic possibilities and uses of different programs, an important consideration in the development of the system for Case Study 3 was the style of interaction between composer and performer that this system could afford. Notably different in this approach, as compared with case studies 1 and 2 was the objective to create a system that would facilitate jointly agreed artistic and collaborative objectives. This is in contrast to the idea of composer and performer supplying complementary, but separate components of the collaboration. In this sense, digital technology has a dual function, and in a way that mirrors changes to performer/composer interactions, a more nuanced role. It plays both a critical artistic part, and also functions as a tool for achieving new collaborative objectives.

### 10.3. Reflection : individual reflection and joint discussion

Reflection upon a range of factors over the course of the collaboration was highly illuminating for both participants. Artistically, discussing and evaluating the strengths and weaknesses of each piece was an important factor in understanding quite different traditional roles and creative perspectives. Conducting a review of theories of systems design, creativity and collaboration for the purposes of this study provided us with ideas for the theoretical framework in which we could place our actions and decisions, or interrogate other understandings of collaboration.

In more general terms, reflection and discussion seemed to facilitate a blurring of boundaries between performer and composer roles as our conversations generated new levels of interest and ownership in a wider range of the creative tasks. As the collaboration unfolded, the consequences of these discussions and reflection is evident in the revised approaches to creative decision making, including the different approaches taken to systems design. Case Study 2 provides a good example of the composer increasing his frame of activity to music and visuals, whilst we see the performer increasing the scope of her concerns to include her relationship to the visual elements of the piece alongside her music making/interpretive role.

### 11. CONCLUSIONS

In this study we have aimed to document and describe key factors contributing to an evolving collaborative relationship between a performer and a composer. Three areas which most facilitated change in the collaborative contexts outlined were 1) challenging traditional roles, 2) changes to practice and 3) reflection (both separately and together).

As a contribution to understandings of collaboration beyond the scope of these three cases, both from the perspective as researchers drawing upon established theories of collaboration and our own professional experiences as artists, we suggest that our study is useful for:

Highlighting ways in which the challenging or blurring of traditional roles can facilitate new insights and approaches to practice;

Appreciating the capabilities of systems design as both an essential artistic tool and facilitator of new levels of interaction between performer and composer in the composition and performance of visual music;

Recognising the value of collaborative work environments for artists as a place of ongoing development and learning.

To that end, discussing, reflecting upon and employing in practice a range of theoretical perspectives on collaboration may be highly useful for composers and performers. Within the visual music making framework, working with tailor-made computer technology in a collaborative environment offers great possibility as both creative tool and interface in new performer/composer relationships and interactions.

### 12. REFERENCES

- [1] Allen, J. A. (1994). The Constructivist Paradigm. *Journal of Teaching in Social Work*, 8(1-2), 31-54. doi: 10.1300/J067v08n01\_03
- [2] Blacking, J. (1973). *How musical is man?* Seattle: University of Washington Press.
- [3] Cockburn, A. & Jones, S. "Four principles of groupware design" *Interacting with Computers* vol.7 no.2 (1995) 195-210.
- [4] Crispin, D. (2009). "Artistic Research and the Musician's Vulnerability." *Musical Opinion*, 132, 25-27.
- [5] Dannenberg, R. B. (2005). "Interactive Visual Music: A Personal Perspective". *Computer Music Journal*, 29(4), 25-35. doi: 10.2307/3681479
- [6] Edmonds, E. A., Weakley, A., Candy, L., Fell, M., Knott, R., & Pauletto, S. (2005). "The studio as laboratory: Combining creative practice and digital technology research." *International Journal of Human-Computer Studies*, 63(4-5), 452-481. doi: 10.1016/j.ijhcs.2005.04.012
- [7] Fleischmann, K. and C. Hutchison (2012). "Creative Exchange: An Evolving Model of Multidisciplinary Collaboration." *Journal of Learning Design* 5(1): 23-31
- [8] Goehr, L. (1989). "Being True to the Work." *The Journal of Aesthetics and Art Criticism*, 47(1), 55-67. doi: 10.2307/431993
- [9] Goehr, L. (1992). *The imaginary museum of musical works an essay in the philosophy of music*. Oxford [England]: Clarendon Press ; Oxford University Press.
- [10] Hatch, J. A. (2002). *Doing qualitative research in education settings*. Albany, N.Y.: State University of New York Press.
- [11] Hayden, S., & Windsor, L. (2007). "Collaboration and the Composer: Case Studies from the end of the 20<sup>th</sup> century." *Tempo*, 61(240), 28-39. doi: doi:10.1017/S0040298207000113

- [12] Hooper, M. (2012). "The start of performance, or, does collaboration matter?" *Tempo*, 66(261), 26-36. doi: doi:10.1017/S0040298212000241
- [13] John-Steiner, V. (2006). *Creative collaboration*. Oxford, Eng.: Oxford University Press.
- [14] John-Steiner, V., Weber, R. J., & Minnis, M. (1998). "The Challenge of Studying Collaboration." *American Educational Research Journal*, 35(4), 773-783. doi: 10.3102/00028312035004773
- [15] Lopez Charles, C. (2012) "La transduction dans la création de la musique visuelle en temps différé" *Revue Francophone d'Informatique Musicale* no. Online versión: <http://revues.mshparisnord.org/rfim/index.php?id=2022>
- [16] Partti, H. (2012). "Cosmopolitan musicianship under construction: Digital musicians illuminating emerging values in music education." *International Journal of Music Education*. doi: 10.1177/0255761411433727
- [17] Partti, H., & Karlsen, S. (2010). "Reconceptualising musical learning: new media, identity and community in music education." *Music Education Research*, 12(4), 369-382. doi: 10.1080/14613808.2010.519381
- [18] Sawyer, K. (2007). *Group Genius*. New York: Basic Books.
- [19] Small, C. (1998). *Musicking : the meanings of performing and listening*. Hanover: University Press of New England.
- [20] Stake, R. E. (2005). *Qualitative Case Studies*. *The SAGE handbook of qualitative research*. Thousand Oaks, Sage Publications. 3rd ed.: 443 - 466.
- [21] Steimerling, O., Hinken, R., & Cremers, A.B. "Distributed Component-Based Tailorability for CSCW Applications" in Proceedings of International Symposium on Autonomous Decentralized Systems, ISADS 1999, 20-23 March 1999, Tokyo, Japan; pp.345-352.
- [22] Taruskin, R. (2003). "Sacred Entertainments." *Cambridge Opera Journal*, 15(2), 109-126. doi: 10.2307/3878330
- [23] Thiétart, R., Wauchope, S. (2001). *Introduction. Doing Management Research*. SAGE Publications Ltd. London, UK: SAGE Publications Ltd.
- [24] Willis, J. (2007a). *Chapter 1. World Views, Paradigms, and the Practice of Social Science Research. Foundations of Qualitative Research: Interpretive and Critical Approaches*. SAGE Publications, Inc. Thousand Oaks, USA: SAGE Publications, Inc.
- [25] Willis, J. (2007b). *Chapter 7. Methods of Qualitative Research. Foundations of Qualitative Research: Interpretive and Critical Approaches*. SAGE Publications, Inc. Thousand Oaks, USA: SAGE Publications, Inc.