



Journées d'Informatique Musicale

15^{ème} édition

Rennes, 18 – 20 mai 2010

Organisées par

l'équipe Arts : Pratiques et Poétiques,
laboratoire Musique et Image : Analyse et Création (MIAC),
Université Rennes 2 Haute Bretagne

les équipes METISS et IMADOC,
Institut de Recherche en Informatique et en Systèmes Aléatoires (IRISA),
CNRS/INRIA/INSA de Rennes

la Société Française d'Analyse Musicale (SFAM)

Sous l'égide de

l'Association Française d'Informatique Musicale (AFIM)

Partenaires

Université Rennes 2 Haute Bretagne

Ville de Rennes

Rennes Métropole

Conseil Général d'Ille et Vilaine

Inria Rennes - Bretagne Atlantique

Programme émergent, Equipe APP,

Rennes 2

Association Rhizome

Remerciements

Services recherche, culturel, logistique, reprographie de l'Université
Rennes 2 Haute Bretagne



INRIA



IRISA



Comité d'organisation

Eric Anquetil

Frédéric Bimbot

Antoine Bonnet

Bruno Bossis

Frédéric Dufeu

Julien Rabin

Gabriel Sargent

Emmanuel Vincent

Comité scientifique

Eric Anquetil

Daniel Arfib

Gérard Assayag

Frédéric Bimbot

Antoine Bonnet

Bruno Bossis

Marc Chemillier

Myriam Desainte Catherine

Dominique Fober

François Giraudon

Mikhail Malt

Yann Orlarey

François Pachet

Laurent Pottier

Jean Michel Raczinski

Jean-Claude Risset

Anne Sèdes

Emmanuel Vincent

Comité de programmation

Eric Anquetil

G rard Assayag

Charles Bascou

Fr d ric Bimbot

Antoine Bonnet

Bruno Bossis

Guilherme Carvalho

Marc Chemillier

Bertrand Co asnon

Myriam Desainte-Catherine

Emmanuel Favreau

Hugues Genevois

Martin Lalibert 

Mikhail Malt

Bertrand Merlier

Fran ois Pachat

Laurent Pottier

Norbert Schnell

Anne S des

Ant nio de Sousa Dias

Emmanuel Vincent

Présentation

Bruno Bossis

SFAM, MINT/OMF Paris Sorbonne, MIAC/APP Rennes 2

Les technologies continuent d'exercer une très forte influence sur de nombreux aspects de la musique et de notre environnement sonore. L'informatique musicale est au centre d'une intense activité qui touche aussi bien le domaine de la recherche que celui de la pratique musicale sous toutes ses formes. Sous l'égide de l'Afim, Association Française pour l'Informatique Musicale, les Journées d'Informatique Musicale ont lieu tous les ans, rassemblent les personnes et les institutions et favorisent des échanges fructueux.

L'université de Rennes 2 est heureuse d'accueillir la quinzième édition des JIM. Si le lieu d'accueil est l'université Rennes 2, les institutions organisatrices sont au nombre de trois. Elles représentent la diversité des domaines concernés par les JIM. Les équipes METISS et IMADOC, l'Institut de Recherche en Informatique et en Systèmes Aléatoires (IRISA) et le CNRS/INRIA/INSA de Rennes sont des lieux de recherche en informatique. La SFAM, Société Française d'Analyse Musicale, est une société savante qui étudie et promeut toutes les formes d'analyse musicale, dans tous les répertoires. Enfin, le laboratoire MIAC, Musique et Image : Analyse et Création, est une composante de l'université Rennes 2 qui s'inscrit dans la recherche en musique et en musicologie. Concernant les arts numériques, le MIAC est lié à un Programme émergent de recherche et à un master recherche (ATN) dont les objectifs sont de développer cet axe dans le domaine des Sciences Humaines.

S'appuyant sur le caractère interdisciplinaire de l'organisation particulièrement visible cette année, deux thèmes ont été mis en évidence lors de l'appel à communications. « La musicologie et les STIC » d'une part, et « L'œuvre musicale face à l'hétérogène : problématique de la mixité » d'autre part. Le premier thème concerne l'analyse à travers la numérisation de partitions musicales et d'enregistrements qui rend possible l'analyse musicologique comparée d'un grand ensemble d'œuvres ou d'interprétations par l'intermédiaire d'outils semi-automatisés issus des STIC. Inversement, l'incorporation de connaissances issues de la musicologie peut bénéficier à des applications des STIC telles que l'indexation de contenus musicaux ou la saisie de partitions. La fertilisation croisée entre ces domaines apparaît donc comme un enjeu essentiel de recherches multidisciplinaires impliquant les communautés concernées. Le second thème concerne l'apparition de nouvelles pratiques sonores à vocation artistique qui accompagne la crise de l'écriture et incite l'œuvre musicale à se confronter, via notamment les nouvelles technologies, à des domaines et modes de pensée qui lui sont hétérogènes : univers sonores, textes, images, cinéma, danse, installations, etc. Telle peut se présenter, dans sa généralité, l'idée de mixité. Complétant utilement ces deux thèmes, les problématiques issues des précédentes éditions et demandant à être approfondies ont été conservées.

Les propositions ont été de grande qualité. En dehors des organisateurs, de nombreuses équipes et institutions sont représentées (Ircam, McGill university, Labri, Scime, Gmea, Grame, universidade Catolica Portuguesa, université Paris 8, conservatoire supérieur de Paris, université de Montréal, Ina-Grm, université Paris Sorbonne, Fondation Destellos en Argentine, université de Saint-Etienne...)

Huit sessions différentes parcourent quelques grands axes très actifs dans l'informatique musicale aujourd'hui : Plateformes de création et de composition, Musique mixte, Temps réel et performance, Notation et visualisation, Image / geste / mouvement, Ecoute et pédagogie, Analyse, Temps réel et instrumentalité. L'ensemble de ces sessions reflète à la fois la diversité des recherches menées autour de l'informatique musicale, et la cohérence d'un domaine qui touche aussi bien l'informatique, l'analyse musicale et la création.

Ces conférences sont enrichies par quatre ateliers pratiques, trois concerts, et deux concours. Les ateliers permettront une approche pratique de logiciels comme iAnalyse, Jamona, Omax et Virage avec leurs concepteurs. Le premier concert est dédié aux musiques improvisées avec ordinateurs, le second aux musiques écrites avec dispositif, et le dernier à la découverte de quelques œuvres audio/vidéo sur support. Enfin, un Prix du jeune chercheur sera attribué et le prix LoMus sera annoncé en fin de conférence.

Les organisateurs tiennent à remercier l'Afim, les équipes de l'université Rennes 2 (Communication, Crea, Recherche), et les différentes institutions et organisations ayant soutenu le projet (Conseil général d'Ille et Vilaine, Conseil régional de Bretagne, Equipe de recherche APP, Inria, Rennes Métropole, Rhizome, SFAM, université Rennes 2, ville de Rennes).

Sommaire

Session 1 : Plateformes de création et de composition

<i>Composer les mélanges sonores avec les maquettes spectrales</i> Philippe Esling, Carlos Agon	15
<i>Représentation des données de contrôle pour la spatialisation dans OpenMusic</i> Jean Bresson, Carlos Agon, Marlon Schumacher	25
<i>A Model For Interactive Scores With Temporal Constraints And Conditional Branching</i> Mauricio Toro-Bermúdez, Myriam Desainte-Catherine, Pascal Baltazar	31

Session 2 : Musique mixte

<i>Penser la musique mixte à la lumière de la mathématique des extensions</i> François Nicolas	41
<i>Aura et perception, Réflexions sur la musique mixte à partir de Benjamin</i> Antoine Bonnet	61

Session 3 : Temps réel et performance

<i>L'instrument augmenté pour l'improvisation : le cas de l'alto avec Rose Amère</i> Karim Barkati, Stéphanie Réthoré	69
<i>La remise en cause de la légitimité de l'œuvre-événement par les nouvelles technologies</i> Anthony Papavassiliou	81
<i>Vertiges de l'espace : un instrument pour la performance électroacoustique improvisée</i> António de Sousa Dias	89

Session 4 : Notation et visualisation

<i>Partitions musicales augmentées</i> Dominique Fober, Christophe Daudin, Stéphane Letz, Yann Orlarey	97
<i>L'acousmographe 3</i> Emmanuel Favreau, Yann Geslin, Adrien Lefèvre	105
<i>Utilisations avancées du logiciel iAnalyse pour l'analyse musicale</i> Pierre Couprie	113

Session 5 : Image, geste, mouvement

<i>Vidéo musique : illusions perceptives et discours esthétique</i> Elsa Justel.....	121
<i>Une approche polyvalente. Direction musicale dansée, captation gestuelle causale</i> Jean-Marie Adrien.....	129
<i>Simultanéité et coïncidence : une proposition de réflexion autour de pratiques artistiques transversales</i> Luc Larmor	135

Session 6 : Écoute et pédagogie

<i>Musique et musicologie à l'Université de Saint-Étienne. Recherches en informatique musicale</i> Laurent Pottier	147
<i>Acousmatic Reasoning: An Organised Listening With Imagination</i> Suk-Jun Kim	153
<i>Instrumenter les opérations d'écoute analytique ? Un bilan du projet 'Écoutes signées' (2003-2006)</i> Nicolas Donin, Samuel Goldszmidt, Jacques Theureau	165

Session 7 : Analyse

<i>Un système de détection pour la description de la structure des morceaux de musique</i> Gabriel Sargent, Frédéric Bimbot, Emmanuel Vincent	177
<i>Décomposition en blocs autonomes comparables</i> Frédéric Bimbot, Olivier Le Blouch, Gabriel Sargent, Emmanuel Vincent	187
<i>La notion de morphologie sonore et le développement des technologies en musique électroacoustique : deux éléments complémentaires d'une unique esthétique ?</i> Gaël Tissot	197

Session 8 : Temps réel et instrumentalité

<i>Tapemovie : un environnement logiciel pour la création temps réel intermédia</i> Tom Mays, Renaud Rubiano.....	207
<i>Multicurve : un nouvel objet graphique de gestion de données de contrôle dans Max/MSP</i> Jean-François Baud.....	215
<i>L'instrument numérique comme objet d'analyse des musiques mixtes</i> Frédéric Dufeu	221

Textes des communications

Session 1

Plateformes de création et de composition

COMPOSER LES MÉLANGES SONORES AVEC LES MAQUETTES SPECTRALES

Philippe ESLING *Carlos AGON*

Institut de Recherche et Coordination Acoustique / Musique
1, place Igor Stravinsky - 75004 Paris
{esling, agonc}@ircam.fr

RÉSUMÉ

Nous présentons dans cet article un système novateur permettant d'explorer l'univers de mélanges sonores à des fins compositionnelles. Ce système intègre plusieurs paradigmes de composition et permet un lien entre la micro-structure du timbre et la macro-forme du discours musical, en imposant des relations temporelles sur des descripteurs spectraux. Il a également été conçu dans le but d'intégrer des unités de signal et de symbolique dans un cadre commun. Le système repose sur une architecture SQL qui agit comme un lexique organisé. Cette base de données contient à la fois des informations symboliques ainsi que des descripteurs spectraux calculés sur des sons provenant de plusieurs collections. Chaque descripteur est modélisé temporellement en utilisant un modèle de mélange gaussien à variance fixe (GMMFV). Nous introduisons également un système de requêtes temporelles permettant de trouver intuitivement des sons correspondants à des courbes d'évolution temporelle des descripteurs spectraux. Finalement, nous présentons les maquettes spectrales pour OpenMusic (OM), qui permettent des interactions temporelles et structurelles complexes pouvant être imposées entre des unités musicales de natures diverses.

1. INTRODUCTION

Au cours du siècle dernier, s'est opéré un revirement dans la musique instrumentale, tendant à dépasser de plus en plus les catégories traditionnelles de l'écriture (hauteur, durée et intensité) et s'attachant à l'inharmonique, la composition de sons bruités et à forte variation de timbre dans le temps. Les marques les plus évidentes de cette évolution sont la *musique spectrale*, la *musique concrète*, ainsi que les écoles de la complexité (Ferneyhough, Dillon) qui cherchent à saturer l'écriture musicale. Parmi les compositeurs contemporains, beaucoup sont résolument tournés vers les qualités expressives du son, grâce au potentiel des sons complexes, bruités ou issus de sources électroniques. L'avènement de la technologie a repoussé la frontière entre le bruit et la musique, comme le montre la démarche de Pierre Schaeffer. En effet, il estime [17] la nécessité de « remplacer la variété limitée de timbres instrumentaux possédée par un orchestre avec l'infinité de timbres bruités obtenus grâce à des mécanismes spéciaux ».

Depuis le XIX^e siècle, le timbre auparavant considéré comme une propriété intrinsèque des instruments s'en est progressivement détaché afin de devenir d'abord un élément nouveau dans le langage musical, évoluant jusqu'au point d'être considéré comme l'épine dorsale de l'écriture musicale. Arriva alors le temps de créer des timbres nouveaux, « composer le timbre » en puisant dans les multiples richesses du vocabulaire instrumental et électronique. Le timbre est devenu le *matériau* principal de la composition, organisant un ordre complexe qui s'élève au-dessus du système instrumental. Le langage musical s'est ainsi progressivement enrichi et ces innovations font maintenant partie intégrante du discours musical contemporain. Ainsi, considéré en termes d'écriture musicale, le timbre joue un double rôle suivant les deux pôles de l'univers instrumental que sont l'articulation et la fusion [5]. L'articulation du discours musical est basée sur la fonction identitaire du timbre, la fusion (ou illusion) qui assure l'impossibilité d'identification est utilisée pour orienter l'auditeur vers de nouveaux éléments du discours musical. Nous pouvons voir ici que chaque pôle s'appuie sur la notion de mélange sonore. Ainsi l'aventure des sons nouveaux, de la musique expérimentale, suscitée par les possibilités de synthèse et de traitement du son ont véritablement fait émerger le problème du timbre des mélanges sonores comme un axe central de recherche en informatique musicale. Dans cette ère nouvelle, l'écoute semble plus orientée sur le *son* de la *musique*. Il s'agit alors d'élever le signal au rang de musique. Ces nouvelles démarches appellent à un approfondissement de nos connaissances sur les mélanges sonores pour arriver à suivre et accompagner les nouvelles orientations musicales. Dans celles-ci, le timbre peut être considéré comme un vocabulaire d'expressivité. Pourtant, il a toujours semblé échapper à l'abstraction pure. Bien sûr, différents modèles de synthèse permettent aux compositeurs de produire des timbres. Avec la synthèse additive, par exemple, les compositeurs peuvent produire des sons par la superposition de plusieurs sinus qui sont considérés comme les unités primitives du système. Cependant, il semble impératif que ces unités primitives soient plus complexes et musicalement pertinentes, comme des sons instrumentaux ou synthétisés. Il est donc nécessaire d'imaginer un système aussi bien qualitatif que quantitatif afin d'explorer les possibilités d'expressivité du timbre. Une approche exploratoire permettrait d'établir

de nouvelles stratégies compositionnelles, profilant ainsi une nouvelle dimension dans l'imaginaire du son.

Cependant, nous pouvons déplorer l'absence de système de composition qui pourrait faire face à ces nouvelles pratiques. Le matériau musical a souvent été pensé et écrit pour un instrument. Aujourd'hui, les systèmes de composition suivent généralement ce paradigme traditionnel et sont inscrits dans un moment ponctuel de l'écriture. Cependant, de nombreux contextes musicaux (figures, textures, gestes) semblent relever conjointement de l'écriture et des mélanges sonores. La vision traditionnelle sur les mélanges sonores, à savoir l'orchestration, peut être vue comme l'ensemble des techniques d'écriture dans lequel le timbre agit comme paramètre principal. Aussi, au-delà du point de vue conventionnel, nous pouvons envisager une généralisation de cette question en comment les sons se combinent pour produire une unité musicale agréable. Dans cette perspective, les compositeurs superposent les unités élémentaires pour en former de plus grandes. Toutefois, dans leur pratique quotidienne de l'écriture, les compositeurs ne travaillent pas seulement par la superposition de sons stationnaires, mais plutôt par la sculpture, simultanément verticale et horizontale du son. Ainsi le processus du mélange des sons est directement lié à l'écriture, voire indissociable de la composition. Toutefois, le problème de l'écriture musicale est un domaine sensible pour la recherche scientifique étant donné que son évaluation technique est soumise à l'appréciation esthétique, et dont les mécanismes semblent difficiles à formaliser. Les fonctions d'analyse acoustique et de traitement symbolique des objets musicaux devraient donc coexister et interagir dans un environnement commun, subordonné à la pensée créatrice du compositeur.

Il semble que le processus compositionnel ne puisse pas contourner la notion de mélanges sonores. Toutefois, celle-ci représente un problème difficile qui soulève de nombreuses questions complexes. En effet, outre les possibilités infinies offertes par un seul signal, les mélanges sonores y ajoutent une nouvelle dimension combinatoire. Le premier problème est donc de trouver des unités sonores qui correspondent à des critères complexes et de natures variées. Il est possible de pallier cette question en mettant en place une base de données organisée. Cependant, celle-ci introduit les questions de taille de stockage ainsi que de recherche rapide basée à la fois sur des informations symboliques et spectrales. Le deuxième problème est la question fondamentale du mélange des sons et donc de trouver des opérations significatives à appliquer aux unités.

Afin de maîtriser la complexité des mélanges sonores, nous introduisons un système de composition qui offre des capacités d'exploration de cet univers combinatoire. L'objectif de ce système est d'abord de permettre un lien entre la micro-structure du timbre et la macro-forme du discours musical, en permettant d'imposer des relations temporelles à partir de descripteurs spectraux. Toutefois, il a également été conçu dans le but de permettre l'interaction entre des unités issues du signal et de la symbo-

lique dans un cadre commun. Afin de permettre des requêtes efficaces fondées sur des critères à la fois symboliques et spectraux, nous avons développé une architecture SQL spécifique. Cette base de données contient des informations symboliques sur les sons provenant de multiples collections ainsi qu'un ensemble complet de descripteurs spectraux pré-calculés sur ces sons. Comme nous l'avons dit précédemment, nous voulons utiliser la micro-structure du timbre comme un élément clé dans le processus compositionnel. Afin de préserver toute la structure spectro-temporelle, chaque descripteur est stocké dans la base de données en utilisant une modélisation temporelle. Les avantages de cette approche sont doubles. Premièrement, la modélisation permet une forte réduction des données en lissant la forme des descripteurs. Deuxièmement, en gardant la structure temporelle du timbre, il est également possible de faire des requêtes basées sur l'évolution des descripteurs plutôt que simplement sur des critères statiques à partir de descripteurs moyennés. Ainsi l'utilisateur peut trouver de manière intuitive des sons suivant une évolution déterminée de volume ou de fréquence fondamentale et ce tout en imposant des contraintes symboliques. Cette architecture permet donc d'établir un lexique organisé pour notre système. Enfin, après avoir permis la possibilité de recherches complexes sur les sons, nous présentons un système permettant d'utiliser pleinement ces nouvelles possibilités. Afin de composer avec des sons jusqu'au niveau du discours musical, nous utilisons le système de maquettes d'OpenMusic [3] comme point de départ. Une maquette peut être considérée comme une « métapartition » englobant dans un seul document, la notation musicale et des programmes audiovisuels [1]. Nous présentons dans cet article les maquettes spectrales, une extension des maquettes classiques permettant l'organisation dans le temps d'événements allant du macro-niveau du discours musical jusqu'aux propriétés micro-temporelles et spectrales.

Le reste de cet article est organisé comme suit. Comme la composition des mélanges sonores se situe à la croisée des mondes du signal et de la notation symbolique, nous discutons dans la section 2 des interactions entre ces deux domaines. Nous présentons ensuite dans la section 3 la modélisation temporelle qui permet d'utiliser les propriétés micro-structurelles du timbre comme un élément compositionnel central. Dans la section 4, nous introduisons la base de données SQL qui permet une organisation hiérarchique ainsi que des requêtes rapides basées sur des critères à la fois spectraux et symboliques. Enfin, dans la section 5 nous présentons le système des maquettes spectrales dans lequel tous ces éléments interagissent et comment des interactions temporelles et structurelles complexes peuvent être imposées entre des unités musicales de diverses natures.

2. DU SIGNAL À LA SYMBOLIQUE

La composition musicale peut être considérée comme un acte symbolique de projection d'un signal hypothétique

qui n'existe que dans l'esprit d'un compositeur vers une représentation symbolique efficace. Cette approche a été largement utilisée depuis des siècles, car elle permet le partage et la distribution de cette image mentale aux interprètes qui essayeront de le reproduire de la manière la plus exacte. La notation symbolique doit donc être la plus représentative possible afin de maintenir l'idée génératrice aussi proche que possible de l'intention du compositeur. La musique occidentale a ainsi été largement fondée sur ce « paradigme harmonique », c'est à dire qu'elle a construit ses concepts musicaux à partir de sons harmoniques et considérés comme stationnaires. La notation musicale occidentale est également sans coïncidence, cohérente avec cette vision (noter la hauteur d'une note comme un point consiste à réduire un son à sa fondamentale, et considérer qu'à l'intérieur de ce son, les harmoniques suivent une évolution parallèle à celle de la fondamentale), tel que le sont les autres systèmes de musique (polyphonie, tonalité).

Cependant, cette tradition symbolique doit coexister avec une préoccupation des compositeurs de plus en plus marquée pour le timbre et une conception musicale qui consiste à former un substrat et une évolution musicale à partir d'éléments caractérisés, de motifs courts, qui se combinent ou se chevauchent de manière à créer une forme plus holistique. L'évolution des pratiques musicales nous a amenés à reconsidérer les sons harmoniques et bruités comme partie intégrante de notre perception musicale. Ainsi, un des intérêts théoriques majeurs provenant de la composition de mélanges sonores est l'utilisation de représentations multi-échelles pour relier des données hétérogènes. De là, émerge une question contemporaine de la recherche en informatique musicale : l'interaction signal / symbolique. Ces deux courants de recherche sont longtemps restés imperméables l'un à l'autre, en partie en raison de l'apparente hétérogénéité de leurs objets d'étude. D'une part, l'analyse et la synthèse de sons numériques ont contribué à la compréhension du son et ont permis la production de sons inouïs. D'autre part, l'approche algorithmique a mis l'accent sur l'analyse des structures symboliques de la notation musicale. Mais la composition des mélanges sonores se situe précisément à l'intersection de ces deux axes de recherche. Si elle prétend créer des timbres, c'est à nouveau à travers un processus d'écriture. C'est donc un point de rencontre essentiel entre les domaines symboliques et spectraux. Ainsi, cette question ne peut être saisie par la musique assistée par ordinateur seulement à la condition de la convergence de l'approche signal et de l'approche algorithmique.

La question des mélanges sonores pourrait encore être manipulée dans le domaine de l'écriture du timbre instrumental. Mais même si nous pouvons la considérer comme une forme de composition, de par son utilisation extensive du timbre comme élément esthétique central, la portée des mélanges sonores en est beaucoup plus profonde. Le son comme matériau coexiste et interagit avec des structures formelles. De nos jours, les ordinateurs offrent la possibilité de manipuler un objet sonore dans un processus de

composition, tout en étudiant simultanément ses propriétés acoustiques au moyen d'outils analytiques. On peut ainsi relier le domaine symbolique de l'écriture avec le domaine spectral des possibilités du timbre. Il devrait donc être possible pour le compositeur de relier l'exploration du son à l'organisation des données à caractère symbolique.

Cependant, même avec des systèmes de composition modernes, les unités spectrales sont généralement réduites à de simples « blocs » symboliques qui seront placés sur une ligne temporelle. Dans ce schéma, les relations temporelles sur la variabilité du signal et la multidimensionalité du timbre résultant sont complètement perdus, afin de faire face à une facilité ergonomique d'utilisation pour l'interaction homme-machine. Nous présentons ici un nouveau schéma compositionnel dans lequel des relations temporelles peuvent être imposées sur les descripteurs du signal tout autant qu'entre des unités spectrales et symboliques. Comme nous considérons une nouvelle approche de la composition musicale, où les propriétés du timbre sont un élément clé du sens musical, il est nécessaire de modéliser précisément l'évolution temporelle des descripteurs spectraux.

3. MODÉLISATION TEMPORELLE

La musique est un processus articulé, ce qui comprend aujourd'hui non seulement la définition des structures macro-temporelles, mais aussi une compréhension plus fine des propriétés micro-temporelles du timbre. Par conséquent, si nous intégrons le timbre comme élément central du processus de composition, nous devons nécessairement étendre sa compréhension afin de composer à des échelles temporelles supérieures. Le territoire de timbre n'est pas limité à une structure statique des proportions. Il est plutôt composé de « lois de variation » qui régissent l'interaction entre les fonctions de fréquence et d'amplitude dans un contexte évolutif dans le temps. Il est donc indispensable de passer à un niveau supérieur de modélisation, par la compréhension des qualités micro-temporelles du timbre afin de capturer le son comme une structure spectro-temporelle complète. Nous proposons donc une modélisation temporelle aussi bien des descripteurs spectraux que des harmoniques par l'utilisation de Gaussian Mixture Models (GMM) [18]. La spécificité de notre modèle est que les moyennes des fonctions gaussiennes sont espacées par une distance proportionnelle à un paramètre de diffusion. Ce paramètre représente également l'écart-type commun à toutes les distributions, conduisant ainsi à un GMM à variance fixe (GMMFV). Ce modèle permet une grande réduction de la quantité d'informations pour le stockage dans la base de données SQL tout en conservant une bonne précision globale. Ce modèle permet également une inférence automatique des propriétés d'un son identique de longueur modifiée.

3.1. Notre modèle

Lors de l'établissement d'un modèle, il est toujours nécessaire d'effectuer un compromis entre la précision et la réduction de données. Pourtant, face à ce problème d'incertitude, nous pouvons utiliser des études perceptives afin de guider notre discrimination des propriétés inutiles. Une étude particulièrement pertinente dans le contexte des modèles instrumentaux a été menée par McAdams et al. [14] et traite de la discrimination des sons instrumentaux aux paramètres spectro-temporels simplifiés. Cette étude donne un aperçu de l'importance perceptuelle de diverses caractéristiques temporelles et spectrales, et lesquelles peuvent donc être simplifiées. D'abord, deux simplifications semblent extrêmement discriminantes. La cohérence des enveloppes d'amplitude et le lissage des enveloppes spectrales montrent que l'évolution indépendante de l'amplitude des partiels est une caractéristique fondamentale. Ces résultats mettent en évidence le fait que le calcul d'amplitudes moyennes pour chaque partiel est largement insuffisant. Il est donc crucial de représenter l'enveloppe de façon indépendante pour chaque partiel. En revanche, la simplification des microvariations, tant en amplitude que en fréquence semble beaucoup moins discriminante perceptivement. Il est donc possible de lisser (jusqu'à un certain niveau) les diverses enveloppes, tout en conservant l'information essentielle.

Pour représenter l'évolution de chaque partiel ou d'un descripteur spectral, nous utilisons les modèles de mélanges gaussiens (GMM). L'évolution d'un partiel est donc définie comme la somme pondérée de distributions gaussiennes

$$B_{x;\theta} = \sum_{c=1}^C \alpha_c \mathcal{N}(x; \mu_c, \sigma_c) \quad (1)$$

avec C le nombre de composants, μ_c la moyenne, σ_c la variance et α_c le poids du composant c , $0 < \alpha_c < 1$ pour tous les composants et $\sum_{c=1}^C \alpha_c = 1$. La liste des paramètres

$$\theta = \{\alpha_1, \mu_1, \sigma_1, \dots, \alpha_C, \mu_C, \sigma_C\} \quad (2)$$

défini donc un modèle particulier de mélange gaussien. Le problème est donc de trouver l'ensemble des paramètres $\hat{\theta}$ qui maximise la probabilité $\mathcal{L}(X; \theta)$ d'observer les données

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}} (\ln \mathcal{L}(X; \theta)) \quad (3)$$

Afin d'estimer les paramètres du modèle, nous utilisons l'algorithme Expectation Maximization (EM). La manière dont cet algorithme est appliqué aux mélanges gaussiens [9] est la suivante. Les données connues X sont interprétées comme des données incomplètes. La partie manquante Y consiste à savoir quel composant produit chaque échantillon x_n . Pour chaque x_n existe un vecteur binaire $y_n = \{y_{n,1}, \dots, y_{n,C}\}$, avec $y_{n,c} = 1$ si l'échantillon est produit par le composant c , ou zéro autrement. L'étape

d'expectation consiste à calculer l'espérance conditionnelle de la log-vraisemblance de toutes les données. La probabilité peut être calculée en utilisant la loi de Bayes :

$$w_{n,c} = \frac{\alpha_c^i p(x_n | c; \theta^i)}{\sum_{j=1}^C \alpha_j^i p(x_n | j; \theta^i)} \quad (4)$$

où α_c^i est la probabilité a priori (de l'estimation θ^i) et $w_{n,c}$ est la probabilité *a posteriori* que x_n soit produit par le composant c . La particularité du modèle GMMFV est que les moyennes des fonctions gaussiennes sont espacées d'une distance proportionnelle au paramètre de diffusion σ_k qui est également l'écart-type commun à toutes les distributions. On obtient donc le mélange suivant :

$$B_{x;\theta} = \sum_{c=1}^C \alpha_c \mathcal{N}(x; \mu_k + c \cdot \sigma_k, \sigma_k) \quad (5)$$

Il suffit donc de conserver uniquement la moyenne μ_k de la première gaussienne et l'écart-type commun σ_k . En appliquant l'étape de maximisation pour estimer la distribution des paramètres pour un mélange de C composants, on obtient les formules de mise à jour

$$\alpha_c^i = \frac{1}{N} \sum_{n=1}^N w_{n,c} \quad (6)$$

$$\mu_k^i = \frac{\sum_{c=1}^C \sum_{n=1}^N (x_n - c \cdot \sigma_k^{i-1}) w_{n,c}}{\sum_{c=1}^C \sum_{n=1}^N w_{n,c}} \quad (7)$$

$$\sigma_k^i = \frac{1}{2} \left(\left((a_k^i)^2 + 4b_k^i \right)^{1/2} - a_k^i \right) \quad (8)$$

$$\begin{cases} a_k^i = \sum_{c=1}^C \sum_{n=1}^N c (x_n - \mu_k^i) w_{n,c} \\ b_k^i = \sum_{c=1}^C \sum_{n=1}^N (x_n - \mu_k^i)^2 w_{n,c} \end{cases} \quad (9)$$

Les nouvelles estimations sont réunies dans θ^i . Les équations sont évaluées à nouveau avec ces nouvelles estimations jusqu'à ce que le critère de convergence soit satisfait.

3.2. Modification de durée

La plupart des instruments de l'orchestre produisant des sons entretenus, il semble obligatoire qu'une modification de durée puisse être faite sur les descripteurs des sons instrumentaux. En utilisant les paramètres du modèle, il devient automatique d'en déduire la structure d'un son instrumental identique, mais de durée différente. Nous considérons la modification par expansion ou compression de l'enveloppe temporelle sur la partie entretenue des sons uniquement. Ainsi les segments d'attaque et de relâchement sont conservés identiques par le calcul de leurs emplacements respectifs et en appliquant ensuite les changements uniquement aux paramètres appartenant à la partie entretenue. Ce choix provient du fait que l'excitation durant l'attaque et le relâchement est généralement identique quelle que soit la durée de la note. Les gaussiennes modifiées sont donc celles dont la moyenne satisfait

$$pos_{att} + \epsilon < \mu_k < pos_{rel} - \epsilon \quad (10)$$

avec pos_{att} la fin du segment d'attaque, pos_{rel} le début du segment de relâchement et ϵ un facteur de tolérance temporelle. On obtient ainsi l'ensemble de gaussiennes à modifier :

$$k \in K = \{\mathcal{N}(\mu_k, \sigma) / pos_{att} + \epsilon < \mu_k < pos_{rel} - \epsilon\} \quad (11)$$

Deux méthodes différentes peuvent alors être appliquées afin de modifier avec précision la durée de la structure.

3.2.1. Modification par dilatation

Dans notre modèle, toutes les gaussiennes sont reliées entre elles par un paramètre de diffusion commun σ_k . Ce paramètre peut être utilisé d'une manière simple pour réaliser l'expansion ou la compression sur le mélange de gaussiennes. Si nous voulons changer la longueur du son par un facteur γ , il nous suffit de multiplier la variance de chaque gaussienne par ce facteur, soit $\sigma'_k = \gamma\sigma_k$. Nous devons ensuite mettre à jour les poids des gaussiennes en conséquence afin de compenser la modification de variance $\alpha'_k = 2 \cdot \gamma \cdot \alpha_k$ puis normaliser tous les poids. Les moyennes des gaussiennes étant reliées par un facteur de diffusion commun, il suffit de calculer les moyennes pour $n \geq \min(k)$

$$\mu_n = \mu_{n-1} + (\sigma'_{n-1} + \sigma'_n) \quad (12)$$

3.2.2. Modification par répétition

Il est possible pour des cas particuliers comme le vibrato et le trémolo d'effectuer une modification plus pertinente par répétition. Cela est fait en conservant les paramètres d'origine mais en changeant le nombre de gaussiennes, puis en mettant à jour les poids des composants correspondants. Nous devons d'abord calculer le nouveau nombre de gaussiennes

$$N_{new} = \lceil \gamma \cdot N_{k \in K} \rceil + N_{g \notin K} \quad (13)$$

Les paramètres des composants pré-existants restent inchangés, et nous considérons que le poids des composants nouvellement ajoutés suit un comportement cyclique

$$\alpha_n = \alpha_{n-card(K)} \quad (14)$$

Il est nécessaire de noter que les deux approches sont applicables à des cas spécifiques. En effet, s'il s'agit d'une enveloppe linéairement décroissante, la modification par répétition apparaît irréaliste. Inversement, pour les trémolos ou vibratos, même si la modification par dilatation semble suggérer une structure valide, il est évident que l'approche par répétition est beaucoup plus pertinente. Ainsi, il serait optimal de choisir la méthode de modification de durée en suivant une taxonomie des échantillons en cours d'analyse.

3.3. Résultats

Nous discutons ici des résultats de notre modèle. Pour cela, nous le comparons avec les polynômes de Legendre et les GMM à variance libre. Il est possible d'observer leur performance sous des angles différents. Nous soulignons l'importance de la précision du modèle en calculant l'écart entre le spectre d'origine et un spectre resynthétisé à partir de données réduites. Cette précision est confrontée à la réduction de données induites par le modèle. Cela nous permet d'évaluer la performance du modèle à saisir les caractéristiques essentielles d'un son.

3.3.1. Précision

La précision est comparée en utilisant deux mesures d'erreur différentes, la norme L^2 et la divergence de *Kullback-Leibler* généralisée. On calcule celles-ci entre le spectre d'origine et le spectre resynthétisé. Toutefois, la durée et le contenu des échantillons pouvant être très variables, on normalise ces mesures avec la divergence entre chaque échantillon et une période équivalente du silence. De cette façon nous pouvons comparer les résultats avec des sons de natures diverses. Enfin nous ajoutons à ces résultats le temps de resynthèse ainsi que la réduction de données induites par les modèles. Le tableau 1 résume les résultats de l'analyse sur les 21.381 échantillons de la base de données.

	Polynômes de Legendre	Gaussian Mixture Models	
		Variance libre	Variance fixe
Synthèse	0.003 +/- 0.004	0.021 +/- 0.003	0.025 +/- 0.013
L^2	39.7 +/- 836.9	19.43 +/- 218.3	8.307 +/- 37.34
$L^2/Sil.$	0.57 +/- 0.152	0.214 +/- 0.027	0.089 +/- 0.022
KL	335.7 +/- 6987	272.8 +/- 458.9	201.7 +/- 362.7
$KL/Sil.$	0.97 +/- 0.504	0.618 +/- 0.004	0.381 +/- 0.015
Réduction	77.2 +/- 498.7	52.20 +/- 339.3	48.68 +/- 246.6

Table 1. Valeur moyenne et variance pour le temps de resynthèse, la réduction de donnée et la précision des différents modèles.

Comme nous l'espérons, nous obtenons les meilleurs résultats avec le modèle à variance fixe qui permet d'obtenir la meilleure précision et ce pour une réduction de données équivalente. La précision du modèle à variance libre pourrait être accrue en augmentant le nombre de composants mais cela augmenterait proportionnellement les chances d'obtenir des estimations singulières.

4. BASE DE DONNÉES SPECTRALE

Durant les dernières décennies, un grand nombre de bases de données ont été établies afin de faire face à l'incroyable diversité et la variabilité des possibilités instrumentales. Les bases de données les plus importantes étant *Studio On Line* (SOL) [4], *Real World Computing* (RWC) [10], et *Vienna Symphonic Library* (www.vsl.co.at).

Cependant, celles-ci manquent clairement d'une organisation architecturale car elles sont uniquement composées de fichiers sonores et de répertoires sans schéma organisationnel complexe. Nous présentons ici une première base de données structurée qui intègre des informations symboliques, spectrales et textuelles sur les fichiers sonores contenus dans toutes les bases précédemment citées. Notre base de données repose sur une architecture SQL et contient pour chaque son des informations sur les descripteurs symboliques et spectraux calculés avec *IRCAM-Descriptor* [16] puis modélisés temporellement grâce au modèle GMMFV.

Cette base de données permet d'effectuer des recherches rapides basées aussi bien sur des descripteurs spectraux que sur de l'information symbolique simultanément. En conséquence, le système est capable de proposer des solutions basées sur des critères à la fois spectraux et symboliques. De plus le système de requêtes temporelles permet de trouver intuitivement des sons correspondant à des courbes d'évolution des descripteurs tels que l'amplitude ou la fréquence fondamentale.

4.1. Architecture SQL

Afin d'organiser et d'accélérer les requêtes sur une large base de données sonores, nous avons conçu une architecture SQL simple et efficace. Comme nous l'avons mentionné précédemment, chaque son de la base de données est représenté par des informations textuelles, symboliques ainsi que des descripteurs spectraux. Pour cela, chaque son nécessite une description qui peut contenir pour les instruments les mode de jeu, hauteur, intensité et autres descriptions informatives. Nous analysons ensuite chaque son en utilisant l'ensemble des descripteurs disponibles dans *IRCAMDescriptor* afin d'obtenir autant d'informations spectrales que possible. Chaque descripteur est ensuite modélisé temporellement en utilisant le modèle présenté dans la section précédente. Les descripteurs représentant la moyenne, l'écart-type et la modélisation temporelle sont stockés dans la base de données. Cette approche permet de faire des recherches conjointes sur des informations issues de la symbolique ou du signal.

Les informations contenues dans la base de données peuvent être de natures diverses :

- Les valeurs atomiques qui représentent une seule information. Les informations symboliques sont stockés avec des chaînes de caractères de longueur variable. La majorité des descripteurs moyens et d'écart-type étant stockés sous forme de valeurs à virgule flottante.
- Les tableaux sont utilisés pour certains des descripteurs moyens qui sont constitués de valeurs multiples (par exemple fréquence moyenne des partiels, platitude spectrale par sous-bande)
- Les modélisations GMMFV sont stockées avec les informations sur la moyenne de la première gaussienne, le facteur de diffusion commun et une liste des amplitudes.

- Les tableaux de gaussiennes sont utilisés pour stocker la modélisation temporelle de descripteurs dont la valeur moyenne est représentée par un tableau.

Pour chaque type d'information, un ensemble spécifique de relations est disponible. Il est donc possible de faire des requêtes simultanément sur des ensemble de données hétérogènes.

4.2. Requêtes temporelles

Pour permettre aux utilisateurs d'effectuer des recherches intuitives dans le lexique, nous introduisons un système de requêtes temporelles. Celui-ci peut être couplé avec n'importe quel type de requête symbolique ou statique et permet ainsi de trouver rapidement et efficacement des sons correspondant à des contraintes complexes. Il permet ainsi de définir des critères d'évolution temporelle sur tout sous-ensemble des descripteurs disponibles. L'utilisateur peut ainsi chercher des sons correspondant à des idées complexes d'évolution. Le principe de requêtes temporelles sur la base de données peut être placé dans le contexte plus global de la *similarité des séries temporelles*. Ce problème peut être ainsi découpé en trois composantes principales :

- La *représentation* des données qui va déterminer la manière dont la recherche et le stockage seront effectués.
- La mesure de *similarité* qui permettra de définir une hiérarchie de ressemblance entre les séries temporelles.
- La technique d'*indexation* des séries qui permet d'organiser les données pour accélérer la recherche dans de grandes bases.

Dans un premier temps, la représentation des données doit fournir un type compact et permettant une recherche rapide. Cette représentation doit également permettre d'évaluer une relation de similarité entre les séries. Un grand nombre de représentations pour la similarité de séries temporelles ont été introduites, notamment la transformée de Fourier discrète [8], la transformée en ondelettes [7], les modèles d'approximation par morceaux [15] et la décomposition en valeurs singulières [11]. Des travaux récents suggèrent que ces modèles ne présentent que peu de différences en termes de possibilités d'indexation [12]. Cependant de nombreuses différences existent quant à la qualité des résultats ainsi que la rapidité des requêtes effectuées sur la base de données. Une propriété principale de toutes les représentations précédentes est qu'elle sont à valeurs réelles. Celle-ci limite donc les algorithmes, structures de données et définitions disponibles pour ces représentations. Ces limitations ont amené à considérer l'utilisation de représentations symboliques pour les séries temporelles. Nous utilisons donc comme représentation un dérivé du système SAX (Symbolic Aggregate approximation) [13]. Celui-ci permet de réduire une série temporelle de longueur arbitraire n à une chaîne de caractères de longueur w (typiquement $w \ll n$) suivant un alphabet de taille a . Cette représentation permet donc d'obtenir une version symbolique de taille très réduite des différentes sé-

ries temporelles. La série à représenter est découpée suivant des points uniformément espacés puis chaque unité est transformée en un caractère en fonction de la valeur moyenne de la série à cet instant. Notre décomposition prends également en compte la valeur moyenne de la dérivée pour chaque unité.

Cette représentation symbolique permet de définir une fonction de similarité sur des chaînes de caractères, prenant en compte la distance entre les différents symboles. Ainsi la similarité entre requêtes temporelles peut dépasser le schéma classique de la distance euclidienne qui ne prend pas en compte des critères perceptuels. En effet, les humains sont capables de détecter des similarités à différentes échelles sur une courbe. En observant de manière simultanée les caractéristiques d'une série, nous sommes capables de nous abstraire des problèmes d'amplitude, de mise à l'échelle temporelle et de bruit. Il est évident qu'une distance point-à-point telle que la distance euclidienne est incapable d'un tel niveau d'abstraction. Une mesure de similarité entre séries temporelles devrait ainsi être invariante suivant une famille de transformations. Soit une série temporelle représentée par la fonction $f(t)$, nous considérons les transformations suivantes :

- **Décalage** : La série $g(t)$ obtenue par décalage linéaire de la série originelle $g(t) = f(t) + k$ avec k une constante.
- **Amplification uniforme** : La série $g(t)$ obtenue par multiplication de la série $g(t) = k.f(t)$ avec k une constante.
- **Modification temporelle uniforme** : La série $g(t)$ obtenue par modification constante de l'échelle temporelle $g(t) = f(kt)$ avec k une constante.
- **Amplification dynamique** : La série $g(t)$ obtenue par multiplication de la série par une fonction d'amplitude $g(t) = h(t).f(t)$ avec $h(t)$ une fonction telle que pour chaque t , $h'(t) = 0$ si et seulement si $f'(t) = 0$.
- **Modification temporelle dynamique** : La série $g(t)$ obtenue par modification dynamique de l'échelle temporelle $g(t) = f(h(t))$ avec $h(t)$ une fonction positive et strictement croissante.
- **Bruit additif** : La série $g(t)$ obtenue par ajout d'une composante bruitée $g(t) = f(t) + \epsilon(t)$ avec $\epsilon(t)$ un bruit blanc indépendant identiquement distribué.

Ces transformations pouvant être combinées simultanément pour former de nouvelles transformations. Pour pallier au problème de décalage et d'amplification uniforme, toutes les séries sont d'abord décalées et amplifiées de manière à ce que leur moyenne soit nulle et que leur déviation standard soit unitaire. Le bruit additif est alors implicitement géré par le passage à la représentation symbolique. Enfin nous utilisons une mesure de similarité [13] sur les chaînes de caractères résultantes permettant de prendre en compte les transformations présentées précédemment. Soit deux chaînes Q et C de longueur w , la chaîne C représentant une série de taille originelle n , la similarité

entre celles-ci est définie par

$$D(Q, C) = \sqrt{\frac{n}{w}} \sqrt{\sum_{i=1}^w (dist(q_i, c_i))^2}$$

avec q_i et c_i le i^{eme} caractère des chaînes correspondantes et $dist(a, b)$ une fonction permettant d'obtenir la distance entre les symboles a et b . Grâce à cette mesure, il est alors possible d'obtenir la similarité entre une série donnée en requête et les séries présentes dans la base de données. L'utilisateur peut alors chercher intuitivement des sons dans la base dont les descripteurs suivent une évolution particulière. Ainsi il suffit de dessiner une courbe pour trouver les sons dont les descripteurs correspondent à celle-ci. Il est alors facilement possible de transposer une idée sonore complexe en recherche intuitive.

4.3. Serveur OSC

Pour que cette base de données puisse être utilisable dans tout type de situations, nous avons développé un serveur OSC qui peut être utilisé comme une interface entre un client et la base de données SQL. Ce serveur a été développé en Matlab et permet à plusieurs clients de se connecter simultanément. Le serveur peut être utilisé pour se connecter à partir d'un client distant sur une base de données distante ou sur une architecture locale. Un ensemble de fonctions encapsule les principales fonctionnalités de la base de données. De cette façon, les clients peuvent ajouter des sons ou effectuer des requêtes sur des descripteurs spécifiques d'une manière totalement transparente. Pour des descripteurs complexes telle que la modélisation GMMFV, le serveur récupère les valeurs du modèle et synthétise la fonction temporelle complète. De cette façon, le client ne voit que la forme définitive des descripteurs. Il est également possible pour le client de demander un aperçu du son. En utilisant la modélisation appliquée à l'évolution des partiels, le serveur est en mesure de calculer une version approximative de l'onde sonore correspondante. Il est enfin possible de sérialiser la base de données entière dans un seul fichier afin de l'utiliser hors connexion.

5. MAQUETTES SPECTRALES

Nous allons voir dans cette section, l'interface pour la base de données spectrale décrite à la section 4. Celle-ci est implémentée comme une bibliothèque dans le logiciel OpenMusic. Tout d'abord nous avons défini une nouvelle classe de sons appartenant à la base de données (`db-sound`). L'utilisateur peut faire des instances de cette classe en envoyant des requêtes SQL à la base de données spectrale (voir Figure 1). Plus qu'un son, un `db-sound` est une collection de sons dont la cardinalité est spécifiée par le paramètre `max-solutions` de la requête. Comme les fichiers sonores classiques, les instances `db-sound` sont des citoyens de première classe (c'est à dire qu'ils peuvent être inclus dans les programmes visuels, notes ou

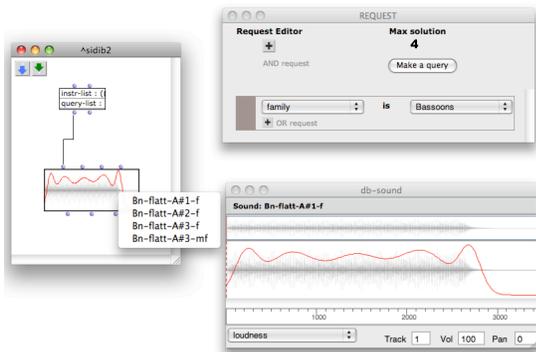


Figure 1. Création d'une instance de la classe `db-sound` par une requête SQL. Dans cet exemple, nous cherchons des sons appartenant à la famille des bassons. La nouvelle instance `db-sound` contient quatre sons. Un éditeur spécial permet d'afficher différents descripteurs de ceux-ci.

d'autres éditeurs OpenMusic). En utilisant des instances `db-sound` ainsi que les outils généraux de programmation visuelle de l'environnement OpenMusic, les compositeurs sont ainsi pourvus de moyens pour développer des processus complets en ce qui concerne la question des mélanges sonores. Dans ce contexte, les structures du programme fonctionnel les laissent gérer la complexité en maintenant un contrôle hiérarchique d'abstractions musicalement pertinentes jusqu'aux processus de synthèse.

En particulier, nous avons mis en place une nouvelle classe de maquettes appelées `Spectral Maquettes`. La maquette est une extension de la notion de programme visuel avec ajout d'autres dimensions spatiales et temporelles, permettant de mettre les éléments du cadre de composition (structures de données et processus) en relation étroite avec ces deux dimensions. Dans un éditeur de maquette, les cases (appelées boîtes temporelles) représentent des unités fonctionnelles (programmes) qui produisent des réalisations musicales (voir Figure 2). La position et les propriétés graphiques de ces boîtes sont associées à un sens temporel et structurel, surtout, l'axe horizontal de l'éditeur représente le temps, afin que la position et l'extension horizontale puissent être liées à des terminaisons et des durées.

Les boîtes temporelles peuvent également être liées par des liaisons fonctionnelles, afin que la maquette entière puisse finalement être considérée comme un programme, comprenant une sémantique fonctionnelle et temporelle. Les relations temporelles et les contraintes (synchronisation, terminaisons relatives, etc) peuvent donc être imposées entre les unités en définissant les paramètres temporels dans leurs programmes correspondants. Par conséquent, le calcul permet de déterminer la structure temporelle. En outre, la possibilité d'intégrer des maquettes dans une maquette permet de construire des structures hiérarchiques temporelles.

La maquette spectrale est une extension de la maquette classique visant à fonder la relation du matériau musical et des processus avec les propriétés micro-temporelles du timbre. En effet, les mélanges sonores exigent d'ac-

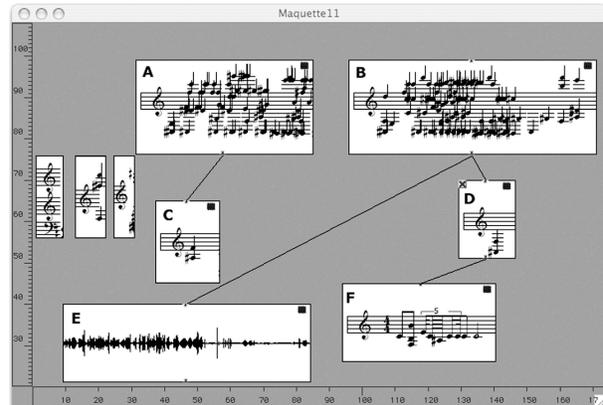


Figure 2. Un éditeur de maquette. Chaque unité représente un programme qui permet le calcul d'objets musicaux, intégré dans une référence temporelle suivant l'axe horizontal. (L'axe vertical est constitué d'unités arbitraires et sans dimension pour permettre une meilleure organisation spatiale.) Les connexions fonctionnelles sont visibles entre les entrées / sorties de certains de ces programmes.

corder une attention particulière à l'évolution temporelle des composantes individuelles ainsi que l'organisation, le séquençage, et les articulations d'une forme musicale globale. Nous montrons dans la Figure 3 un exemple de maquette spectrale. Quand un `db-sound` est mis dans la maquette, la boîte sous-jacente montre la forme d'onde du son augmentée avec plusieurs informations comme des points clef (par exemple *attaque*, *sustain* et *release*) ou l'évolution de certains descripteurs (par exemple, *loudness*, *noisiness*, centroïde spectral, etc.). Les unités à l'intérieur d'une maquette spectrale peuvent être reliées entre elles de différentes manières. Dans la Figure 3 (a) nous pouvons voir comment le démarrage d'un son est lié à la première note d'une mélodie ; déplacer une de ces unités force le déplacement de l'autre, afin de respecter cette contrainte. Dans la partie (b) trois sons se superposent, on peut voir dans cet exemple plusieurs types de relations temporelles : le plus bas est attaché à un marqueur dans la *timeline* globale, ce son démarrera toujours à partir de ce moment, la boîte du milieu quant à elle, se termine en même temps que la plus basse (d'autres relations de Allen peuvent aisément être imposées entre les boîtes), et enfin les relations entre la boîte la plus haute et celle du milieu sont définies au moyen de marqueurs internes aux sons (ces points peuvent être définis à la main ou provenir de requêtes sur les bases de données). Dans la partie (d) le `db-sound` est considéré comme l'entrée d'un programme qui génère une séquence de notes dont le profil mélodique suit l'évolution de l'amplitude du son. Enfin, dans la partie (e) un visualisateur permet de sélectionner un descripteur particulier et calculer le résultat de la superposition de tous les `db-sounds` dans la maquette pour celui-ci. Dans un premier temps ce visualisateur permet de voir l'évolution globale de la superposition, mais il n'est pas difficile d'imaginer qu'en changeant la courbe, l'uti-

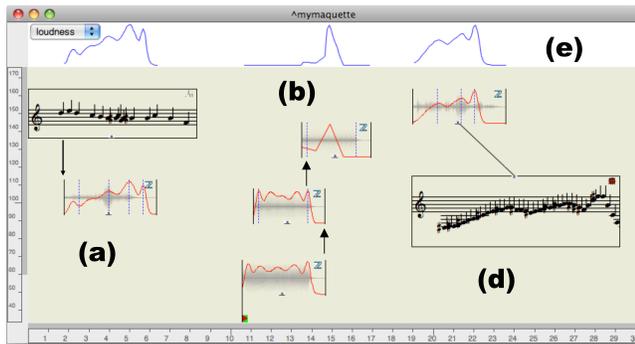


Figure 3. Exemple concret de l'utilisation des maquettes spectrales. De multiples relations fonctionnelles, macro et micro-temporelles sont définies entre les unités.

lisateur puisse lancer une requête qui recherche de nouvelles unités permettant d'approcher cette nouvelle courbe. Dans l'exemple précédent, nous avons essayé de montrer comment les maquettes spectrales permettent de structurer les informations musicales à différents niveaux :

- Le niveau statique de la forme, à savoir l'organisation des unités dans le temps.
- Le niveau dynamique et paradigmatique de la forme (c'est à dire les contraintes temporelles entre les unités).
- Le niveau syntaxique, c'est à dire le calcul permettant la construction du discours musical entre les unités temporelles.
- Le niveau du matériau, c'est à dire les db-sounds qui participent à la maquette.

Ces quatre niveaux d'information sont évidemment liés. L'avantage le plus important dans le concept de maquette spectrale est d'offrir une visualisation de cette interaction et dans le même temps, un contrôle interactif de celle-ci. Cela produit une source d'exploration et d'expérimentation :

- La recombinaison au niveau de la forme globale. Les unités temporelles sont déplacées et étendues dans le temps sans changer les trois autres niveaux.
- Modification des relations fonctionnelles. Nous ne modifions pas la position des blocs, mais leur relation de causalité.
- Modification de la syntaxe. Les algorithmes construisant le matériau sonore peuvent être changés en fonction des objectifs de composition.
- Modification du matériau en lançant de nouvelles requêtes à la base de données.

En combinant ces différentes procédures, des expérimentations musicales sophistiquées peuvent être effectuées. Il est important de noter que dans les maquettes spectrales, l'utilisateur ne travaille pas directement avec des sons. Le son résultant d'une maquette spectrale est synthétisé par le serveur afin de créer un nouveau son de la base ou simplement d'être écouté.

6. CONCLUSION

Nous avons présenté dans cet article un nouveau système permettant d'explorer l'univers combinatoire des mélanges sonores à des fins compositionnelles. En s'appuyant sur une architecture SQL, il est possible de faire des requêtes rapides basées simultanément sur des critères symboliques et spectraux. Un système de requêtes basé sur la forme temporelle de descripteurs a été mis en œuvre afin d'ajouter la possibilité de rechercher des évolutions de descripteurs spécifiques à l'intérieur de la base. Cette approche permet de trouver des unités qui correspondent à des idées sonores complexes. La modélisation temporelle appliquée à chaque descripteur spectral ainsi qu'aux différents partiels permet de préserver l'ensemble de la structure spectro-temporelle. Ce lexique organisé a été utilisé comme une base pour le système de maquettes spectrales. Les maquettes permettent des interactions temporelles et structurelles complexes qui peuvent être imposées entre des unités musicales de diverses natures. Ce système devrait permettre des possibilités d'exploration en profondeur du matériau sonore grâce à l'accès à la multidimensionalité de timbre et à la complexité des mélanges sonores.

Plusieurs directions de recherche peuvent être entreprises sur les maquettes spectrales. Premièrement, s'agissant du modèle temporel, le type de distributions peut être modifié afin d'augmenter la précision globale et la réduction des données du modèle. Le modèle de mélange gaussien asymétrique [2] pourrait être appliqué et renforcerait la réduction des données. Pour l'architecture SQL, un système d'indexation efficace permettrait d'imaginer la mise à l'échelle de la base de données à l'intégralité des collections sonores disponibles. Enfin, l'intégration du logiciel d'orchestration *Orchidée* [6] développé à l'Ircam permettrait des possibilités d'exploration supplémentaires. Ce système est basé sur l'idée de son *cible* qui doit être reproduit en faisant une recherche de combinaisons sonores au sein de grandes banques d'échantillons instrumentaux. La recherche est effectuée par un algorithme génétique qui permet une recherche multi-objectif vers un timbre cible, dans laquelle plusieurs dimensions perceptives sont optimisées conjointement. Il deviendrait ainsi possible pour le compositeur de dessiner des formes de descripteurs et le système ne proposerait pas seulement une solution unique mais plusieurs unités résultant en un mélange sonore correspondant aux différents critères.

7. REFERENCES

- [1] C. Agon, G. Assayag, O. Delerue, and C. Rueda, "Objects, time and constraints in OpenMusic," *Proceedings of the International Computer Music Conference*, Ann Arbor, USA, 1998, pp. 406 – 415.
- [2] R. B. Arellano-Valle, H. Bolfarine, and V. H. Lachos, "Skew-normal linear mixed models," *Journal of Data Science*, vol. 3, pp. 415 – 438, 2005.
- [3] G. Assayag, C. Rueda, M. Laurson, C. Agon, and O. Delerue, "Computer-assisted composition at Ir-

- cam: From Patchwork to OpenMusic,” *Computer Music Journal*, vol. 23, no. 3, pp. 59 – 72, 1999.
- [4] G. Ballet, R. Borghesi, P. Hoffmann, and F. Levy, “Studio online 3.0 : An internet “killer application” for remote access to ircam sounds and processing tools,” in *Actes des Journées Informatique Musicale*, Paris, France, 1999.
- [5] P. Boulez, “Le timbre et l’écriture, le timbre et le langage,” in *Le timbre, métaphore pour la composition*. J.B. Barrière, 1991, pp. 541 – 549.
- [6] G. Carpentier, “Approche computationnelle en orchestration musicale - optimisation multicritère sous contraintes de combinaisons instrumentales dans de grandes banques de sons,” Ph.D. dissertation, University UPMC Paris 6, 2008.
- [7] K. Chan and A. Fu, “Efficient time series matching by wavelets,” in *Proceedings of the 15th IEEE International conference on data engineering*, Sydney, Australia, 1999, pp. 126 – 133.
- [8] C. Faloutsos, M. Ranganathan, and Y. Manolopoulos, “Fast subsequence matching in time-series databases,” *SIGMOD Record*, vol. 23, pp. 419 – 429, 1994.
- [9] M. A. T. Figueiredo and A. K. Jain, “Unsupervised learning of finite mixture models,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 381 – 396, 2002.
- [10] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka, “Rwc music database : music genre database and musical instrument sound database,” in *Proceedings of the 4th International Conference on Music Information Retrieval*, Washington, USA, 2003, pp. 229 – 230.
- [11] E. Keogh, K. Chakrabarti, and M. Pazzani, “Locally adaptive dimensionality reduction for indexing large time series databases,” in *Proceedings of ACM SIGMOD conference on management of data*, Santa Barbara, USA, 2001, pp. 151 – 162.
- [12] E. Keogh and S. Kasetty, “On the need for time series data mining benchmarks : a survey and empirical demonstration,” in *Proceedings of the 8th ACM SIGKDD International conference on knowledge discovery and data mining*, Edmonton, Alberta, Canada, 2002, pp. 102 – 111.
- [13] J. Lin, E. Keogh, L. Wei, and S. Lonardi, “Experiencing SAX: a novel symbolic representation of time series,” *Data Mining and Knowledge Discovery*, vol. 15, no. 2, pp. 107–144, 2007.
- [14] S. McAdams, J. W. Beauchamp, and S. Meneguzzi, “Discrimination of musical instrument sounds resynthesized with simplified spectrotemporal parameters,” *Journal of the Acoustical Society of America*, vol. 105, no. 2, pp. 882 – 897, 1999.
- [15] G. P., “Pattern extraction for time series classification,” in *Proceedings of the 5th European conference on principles of data mining and knowledge discovery*, Freiburg, Germany, 2001, pp. 115 – 127.
- [16] G. Peeters, “A large set of audio features for sound description (similarity and classification) in the cuidado project,” IRCAM, Paris, Tech. Rep., 2004.
- [17] P. Schaeffer, *Traité des objets musicaux*, Seuil, 1966.
- [18] J. J. Verbeek, N. Vlassis, and B. Krose, “Efficient greedy learning of gaussian mixture model,” *Neural Computation*, vol. 15, no. 2, pp. 469 – 485, 2003.

REPRÉSENTATION DES DONNÉES DE CONTRÔLE POUR LA SPATIALISATION DANS OPENMUSIC

Jean Bresson, Carlos Agon
IRCAM - CNRS UMR STMS
{jean.bresson,carlos.agon}@ircam.fr

Marlon Schumacher
IDMIL - DCS - CIRMMT
Schulich School of Music of McGill University
marlon.schumacher@music.mcgill.ca

RÉSUMÉ

La prise en compte de l'espace dans les démarches musicales contemporaines entraîne une nouvelle réflexion sur les méthodes et environnements informatiques de composition. Nous présentons des travaux actuels réalisés dans l'environnement de composition assistée par ordinateur OpenMusic. Plusieurs directions sont explorées, concernant la génération de trajectoires dans l'espace, les structures de contrôle pour la localisation et la spatialisation des sources sonores, ou encore la création de sons spatialisés au sein de cet environnement. L'export et la communication de données avec des systèmes de spatialisation externes sont également envisagés, pour lesquels nous proposons une solution utilisant le format SDIF.

1. INTRODUCTION

Malgré les progrès réalisés dans les techniques de rendu sonore spatialisé, la prise en compte de l'espace dans les processus de création musicale reste un aspect problématique et relativement peu exploré en composition assistée par ordinateur.

La majorité des outils existant actuellement se situe dans le domaine du traitement audio temps-réel ou dans des environnements et stations audio-numériques, et propose des interfaces de contrôle permettant de déterminer de façon plus ou moins directe un *panning* d'amplitude des sources sonores sur différentes configurations de haut-parleurs.¹ La détermination des paramètres spatiaux (trajectoires, effets de salles, etc.) reste cependant relativement limitée par un contrôle généralement décorrélié de la perception, mais également des autres paramètres musicaux. Ces limitations, par ailleurs, s'affirment d'autant plus lorsque le nombre de sources sonores est élevé et/ou lorsque les configurations de restitution ne sont pas fixées dès le départ.

Le Spatialisateur de l'IRCAM [13] est un exemple d'application intégrée dans l'environnement Max/MSP, permettant de contrôler les paramètres de spatialisation à l'aide de données telles que les positions et mouvements de sources dans l'espace, les caractéristiques acoustiques des salles ou autres paramètres perceptifs. Cependant, ces

paramètres restent généralement déconnectés de l'univers compositionnel et des représentations symboliques qui permettraient de les insérer dans une démarche musicale.

La recherche de contrôle musical de plus haut niveau pour la spatialisation a été abordée dans différents projets tels que le logiciel MusicSpace [10] qui permet de contrôler les positions de différentes sources sonores à l'aide d'un système de contraintes, établies et propagées dynamiquement entre celles-ci. L'interface Holo-Edit, créée dans le cadre du projet Holophon [8] propose une autre approche et permet de créer et éditer des trajectoires tridimensionnelles graphiquement ou à l'aide de fonctions de génération et de transformation [19]. Cet aspect fonctionnel, ou algorithmique, également présent dans des projets plus anciens tels que MoveInSpace [23], est un pas significatif dans l'idée d'un contrôle compositionnel des paramètres spatiaux. En effet il est important pour les compositeurs voulant intégrer l'espace dans leurs démarches musicales de pouvoir définir leurs propres processus génératifs et de les lier à un univers ou un projet compositionnel donné [2].

Le contrôle de la spatialisation dans le contexte de la composition assistée par ordinateur (CAO) a également fait l'objet de précédentes expériences dans l'environnement OpenMusic [3]. Avec OpenSpace [11], il a été envisagé de contrôler le système de contraintes de MusicSpace, ajoutant et modifiant ces contraintes dynamiquement dans un déroulement temporel déterminé depuis l'environnement de CAO. Un autre projet important, à la base de nos travaux actuels, est la bibliothèque OMSpat [17] qui permettait d'associer trajectoires et paramètres spatiaux générés dans OpenMusic à des sources sonores, par l'intermédiaire d'une structure matricielle formatée et transférée à une interface de contrôle pour le Spatialisateur.

Les travaux que nous présentons étendent le système OMSpat sous différents points de vue, reprenant et généralisant l'idée des objets et interfaces primitives créées alors pour la représentation des courbes et trajectoires dans l'espace, ainsi que celle d'une représentation matricielle pour la description des scènes sonores. Rattachée au système de contrôle de la synthèse OMChroma [1], cette représentation matricielle peut générer un rendu sonore spatialisé directement dans OpenMusic. Nous insisterons ici sur une solution complémentaire consistant à transférer

1. Une revue et description des principaux outils et logiciels de spatialisation est disponible sur le site
http://multiphonie.free.fr/log_spat.htm

ces données vers des systèmes externes de spatialisation, en utilisant le format SDIF.

2. GÉNÉRATION DE TRAJECTOIRES

Une spécificité du contrôle de spatialisation dans le contexte de la CAO est de travailler sur des représentations abstraites des objets musicaux, et non sur des données de paramétrage comme dans le cas des contrôleurs audio-numériques. Nous nous intéresserons principalement ici aux trajectoires tridimensionnelles, regroupant des points localisés dans l'espace et leurs évolutions temporelles. Le même type de représentations pourra cependant être utilisé pour des ensembles de positions statiques, ou même pour d'autres données telles que la directivité des sources sonores.

De nouveaux objets viennent compléter les représentations traditionnelles des courbes dans OpenMusic. L'objet *3DC* est une suite de points en 3 dimensions, et l'objet *3D-Trajectory* étend ce dernier à la dimension temporelle, en permettant d'associer une date aux différents points (tous ou certains points d'intérêt). Les outils de traitement et de génération de courbes disponibles dans OpenMusic (par exemple le rééchantillonnage ou la génération de courbes B-splines) permettent de créer ces données parfois complexes à partir de nombres réduits de points entre lesquels une trajectoire peut être calculée suivant une résolution donnée (spatiale ou temporelle).

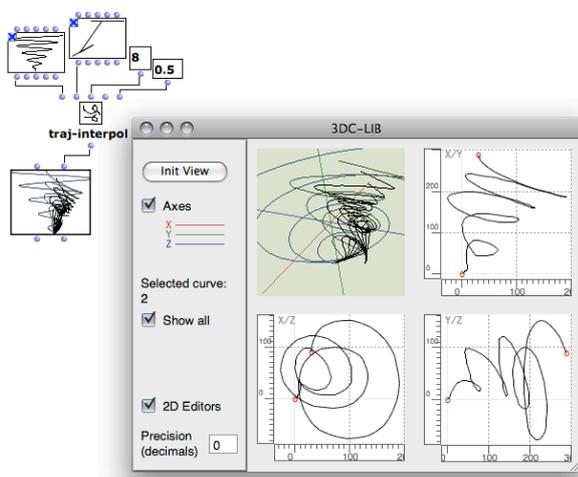


Figure 1. Création et édition de courbes et trajectoires 3D.

Suivant un principe de dualité entre processus et objets musicaux, courbes et trajectoires peuvent être générées par des processus fonctionnels et algorithmiques au sein de l'environnement de programmation visuelle, mais également directement à l'aide d'éditeurs graphiques permettant de dessiner les courbes, placer / déplacer les points, modifier les points d'inflexion ou les repères temporels. La figure 1 montre un objet *3DC-Lib* (collection d'objets *3DC*) créé à l'aide d'une fonction d'interpolation

dans un programme visuel, ainsi que l'éditeur graphique associé à cet objet.

3. MATRICES ET REPRÉSENTATION DES SCÈNES SONORES SPATIALISÉES

Comme nous l'avons évoqué en introduction, les structures matricielles, utilisées notamment pour la synthèse sonore dans OpenMusic [1], présentent un intérêt pour le contrôle de la spatialisation et la représentation des scènes sonores spatialisées sous forme d'objets compacts, symboliques et opérationnels.

Dans une de ses dimensions, une matrice permet en effet de mettre en relation une source sonore et ses différents paramètres de spatialisation (positions, trajectoires, mais éventuellement aussi d'autres paramètres acoustiques comme la réverbération, la directivité, etc.) Dans sa deuxième dimension, elle met en relation les différents éléments constituant un son (dans le cas de la synthèse sonore) ou une scène sonore spatialisée, associant entre elles les différentes valeurs de chaque paramètre du processus. A titre d'exemple, il est ainsi possible d'associer aux différents composants d'une scène spatialisée une même trajectoire, d'établir des rapports entre leurs trajectoires respectives, ou encore de définir un processus qui générera dynamiquement l'ensemble de ces trajectoires.

La figure 2 montre un exemple de construction d'une "matrice de spatialisation", associant sources sonores et paramètres de spatialisation. Cette structure constitue une représentation abstraite susceptible d'être intégrée dans un processus musical de plus haut niveau, et éventuellement transférée vers un système de spatialisation sonore.

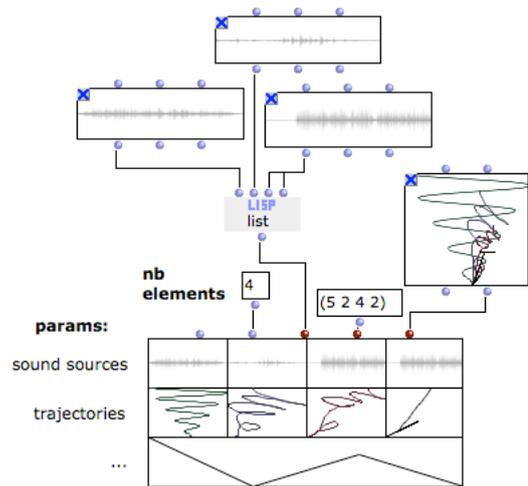


Figure 2. Objets et sources sonores spatialisées dans OpenMusic : représentation matricielle.

Ce transfert pourra avoir lieu dans OpenMusic, comme illustré dans la section 4, ou par le formatage et la communication avec des environnements externes (section 5).

4. SYNTHÈSE ET RENDU SONORES SPATIALISÉS DANS OPENMUSIC : OMPRISMA

OMPrisma [21] est une bibliothèque OpenMusic étendant à la spatialisation le système de contrôle de la synthèse OMChroma. Ce système propose un ensemble de classes correspondant à différents processus de synthèse (ou de spatialisation, avec OMPrisma). Une classe dans OMChroma (et OMPrisma) est une matrice associée à une unité de génération ou traitement de signal (implémentée sous forme de code ou *instrument* Csound [5]). Ainsi les différents paramètres de cette unité de traitement (correspondant au *score* dans Csound) sont générés à partir des valeurs de la matrice suivant l'une de ses dimensions, et instanciés un nombre de fois déterminé par sa deuxième dimension. Un appel au compilateur Csound permet ensuite de générer un fichier audio à partir de l'instrument associé et le *score* généré par la matrice.

Suivant ce principe, OMPrisma propose des modules de spatialisation de sources sonores basés sur des techniques telles que le *panning* sur différents canaux, le VBAP (*vector-based amplitude panning* [20]), ou encore la synthèse *ambisonics* [9]. Celles-ci permettent de générer des scènes sonores complexes en relation aux processus compositionnels développés dans OM. L'interchangeabilité des différentes classes permet de concevoir ces processus et scènes sonores de façon relativement indépendante du système de rendu et de reproduction spatialisée.

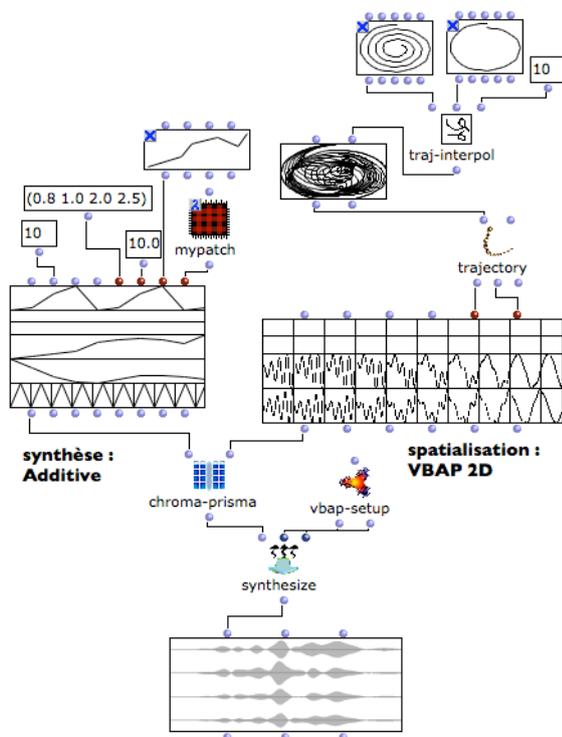


Figure 3. Synthèse sonore spatialisée avec OMChroma et OMPrisma : chaque composant du processus de synthèse (partiels, ici dans le cas d'une synthèse additive) est spatialisé suivant une trajectoire calculée par interpolation.

Par ailleurs, il est possible aussi de fusionner les classes de spatialisation de OMPrisma avec les différentes classes proposées pour la synthèse sonore dans OMChroma (par exemple, synthèse additive, FM, FOF, granulaire, etc.) Chaque composant d'un processus de synthèse (respectivement, chaque partiel, instanciation d'un générateur FM, fonction d'un banc de FOF, grain sonore, etc.) peut alors être associé à un ensemble de paramètres et à un module de traitement intégré pour sa spatialisation (voir figure 3). Couplé à l'environnement de contrôle, ce système implémente une approche générique de la notion de synthèse sonore spatialisée (également abordée dans des travaux tels que [24] ou [15]) adaptée au contexte compositionnel.

5. EXPORT DES DONNÉES : UNE PROPOSITION AVEC LE FORMAT SDIF

La standardisation des formats d'échange pour la spatialisation est une question centrale, dans la mesure où un nombre croissant de technologies se développent, et où les activités de contrôle, de rendu et de reproduction sont très souvent réalisées dans des environnements et configurations variables. Au delà des possibilités offertes par la spatialisation réalisée dans OpenMusic (via OMPrisma et Csound), il serait avantageux, en effet, de générer des données également exploitables dans d'autres environnements, notamment dans les environnements de traitement temps-réel comme le Spatialisateur ou avec des systèmes plus complexes comme la WFS [4].

Différentes tentatives de standardisation ont été proposées pour la description de scènes sonores spatialisées, notamment autour des formats MPEG ou du standard XML (voir le panel dédié à ce sujet lors de la conférence ICMC 2008 [14]). Parmi celles-ci, les formats de stockage (par exemple ASDF [12], pour reprendre l'une des propositions du précédent panel, mais également des formats plus généraux comme VRML/X3D) sont généralement axés sur des descriptions et spécifications de haut niveau (par exemple avec des déclarations d'objets et de transformations telles que des rotations ou translations).

Dans le cadre qui nous concerne ici, cependant, nous souhaitons décrire des scènes sonores d'un point de vue « dénotationnel », c'est-à-dire écrire des fichiers décrivant des trajectoires et courbes de paramètres échantillonnés et dans leur résolution finale (cet échantillonnage pouvant d'ailleurs être pris en charge par l'environnement de contrôle en tant que spécification "compositionnelle"). Ces descriptions pourront alors être lues et interprétées directement, sans déduction ni calcul, par différents environnements de rendu spatialisé.

A l'inverse des précédents, les formats ou protocoles dédiés à la communication temps-réel (par exemple SpatDIF [18]) proposent des descriptions plus directes et précises mais n'offrent actuellement pas de solution de stockage satisfaisante.²

². Dans l'idéal, un standard devrait à terme inclure ces deux modalités et permettre à la fois le transfert (*streaming*) et le stockage des données, à l'instar du protocole MIDI et du Standard MIDI File [16].

5.1. Le format : SDIF

Suivant les précédentes évolutions de l'environnement de CAO en direction du domaine de l'analyse et la synthèse sonore, nous avons choisi d'utiliser le format SDIF (*Sound Description Interchange Format* [22]) comme support de nos descriptions spatiales dans OpenMusic.³ Plusieurs raisons peuvent être évoquées à cela, parmi lesquelles :

- L'existence d'outils permettant la manipulation et le traitement des données SDIF dans OpenMusic, et l'utilisation préalable et établie de ce format pour le stockage et transfert des descriptions sonores [6] ;
- L'extensibilité et l'ouverture du format, qui propose un protocole aisé pour l'ajout de nouveaux types de données ;
- La structure du format, qui permet d'envisager les données sous forme de flux temporels multiples.

Le format SDIF se base sur une description matricielle (qui convient donc a priori à notre représentation interne, voir section 3) décrivant les valeurs d'un certain nombre de paramètres (ou champs de description – *fields*) pour un certain nombre de composants. Ces matrices ont un "type" (identifié par une *signature* de 4 caractères) et sont incluses dans des *frames* déterminant un ordonnancement temporel. Les *frames* ont également une signature correspondant à un type, et un identifiant de flux (*stream ID*) permettant de décrire et différencier un ensemble de flux entrelacés dans un même fichier.

Il existe un certain nombre types SDIF standards et pré-définis, correspondant aux descriptions sonores les plus courantes, mais il est possible d'étendre ceux-ci par de nouveaux paramètres, ou d'en définir de nouveaux. La déclaration de types peut avoir lieu au niveau global dans la librairie SDIF, pour les types standardisés, ou localement à l'intérieur d'un fichier dans les autres cas. Une fois ces types déclarés, une application lisant les données dans un fichier est ainsi capable, d'après le type des matrices, de les identifier et éventuellement de les interpréter de façon adéquate.

5.2. Déclaration des types

Notre description, dans un premier temps, comprend uniquement les trajectoires tridimensionnelles des sources sonores dans l'espace. Voici la déclaration du type de matrice proposé selon le protocole de spécification SDIF :

```
1MTD XPOS {x, y, z}
```

Nous définissons ici le type de matrice "XPOS" (MTD = *Matrix Type Definition*) dont les trois champs sont les coordonnées cartésiennes des points dans l'espace.⁴

3. SDIF a également été envisagé comme support de stockage pour des données de contrôle SpatDIF [18].

4. Selon les conventions, les types SDIF standardisés commencent généralement par un numéro – "1" la plupart du temps, et les types plus expérimentaux par un "X".

D'autres types peuvent être déclarés sur ce modèle, correspondant par exemple aux coordonnées sphériques ou cylindriques :

```
1MTD XSPH {azimuth, zenith, distance}
1MTD XCYL {azimuth, distance, elevation}
```

Comme nous l'avons dit, les matrices sont intégrées dans un flux temporel par l'intermédiaire de *frames*, pour lesquels il est nécessaire de définir également un type. Le type "XSPA" défini ci-dessous (FTD = *Frame Type Definition*) correspond à un *frame* SDIF pouvant contenir des matrices de type XPOS, XSPH, ou XCYL.

```
1FTD XSPA {
  XPOS cartesian_coord;
  XSPH spherical_coord;
  XCYL cylindrical_coord;
}
```

La figure 4 montre un exemple de stream SDIF (séquence de *frames*) décrivant l'évolution de la position d'une source sonore dans l'espace.

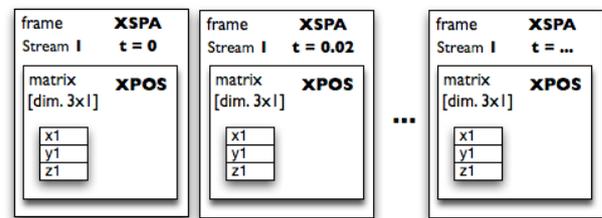


Figure 4. Séquence de *frames* SDIF de type "XSPA" décrivant l'évolution dans l'espace d'une source sonore.

De la même façon peut être envisagée la définition de types de matrices correspondant aux autres paramètres entrant en jeu dans la spatialisation, comme les effets de salle ou la directivité des sources. Concernant les effets de salles, nous pouvons par exemple imaginer un nouveau type de matrice (par exemple "XRFX" pour "room effects") qui serait inclus dans un flux de *frames* d'un nouveau type (par exemple lui aussi de signature "XRFX") :

```
1MTD XRFX {param1, param2, ...}
1FTD XRFX { XRFX room_effects; }
```

Si l'on souhaite attacher ces paramètres à des sources particulières, il est également possible d'étendre le type de *frame* "XSPA" pour y autoriser les matrices de type "XRFX" :

```
1FTD XSPA {
  XPOS cartesian_coord;
  XSPH spherical_coord;
  XCYL cylindrical_coord;
  XRFX room_effects;
}
```

De cette manière, les paramètres de salle, soit au niveau général, soit au niveau de chaque source individuelle, peuvent être inclus sans incidence dans les flux de contrôle SDIF au(x) moment(s) où ces paramètres auront à être communiqués ou modifiés.

5.3. Description de scènes sonores "polyphoniques"

Les aspects polyphoniques, soit la possible multiplicité des sources sonores présentes dans une scène spatialisée, peuvent être considérés de différentes manières.

L'une d'elles serait une description mettant en jeu un certain nombre de sources sonores identifiées (typiquement, des fichiers audio) : dans ce cas il est préférable de générer un flux séparé pour chaque source, ce qui permettra de leur assigner des granularités ou précisions de contrôle indépendantes (par exemple certaines peuvent être immobiles et décrites par un ou peu de *frames*, quand d'autres décriront des mouvements nécessitant une plus grande précision temporelle). La figure 5 illustre ce type de description où chaque source est liée à un *stream* SDIF et donc encapsulée dans un *frame* indépendant.

Le système de spécification de relations entre les *streams* proposé dans [7] nous invite à ce type d'organisation dans laquelle les flux peuvent être liés les uns aux autres ou à des ressources externes par l'intermédiaire de tables déclarées en tête de fichier.

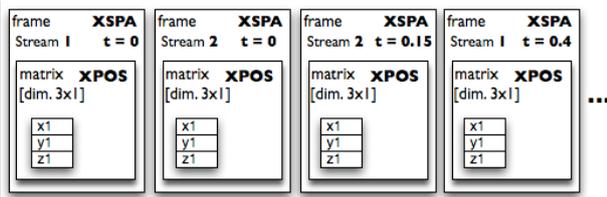


Figure 5. Séquence de *frames* SDIF "XSPA" décrivant l'évolution dans l'espace de deux sources sonores, correspondant respectivement aux *streams* 1 et 2.

Lorsque les sources ne sont pas des ressources identifiées mais par exemple les multiples éléments (grains, partiels, etc.) d'un processus de synthèse sonore (comme dans le cas de la synthèse sonore spatialisée, voir section 4), il peut être préférable d'envisager un flux unique de *frames* dont les matrices n'auront pas un, comme dans le cas précédent, mais un ensemble d'éléments synchronisés. Chacun de ces éléments aura alors des caractéristiques spatiales individuelles mais spécifiées dans une même structure localisée dans le temps. La figure 6 illustre ce deuxième cas.

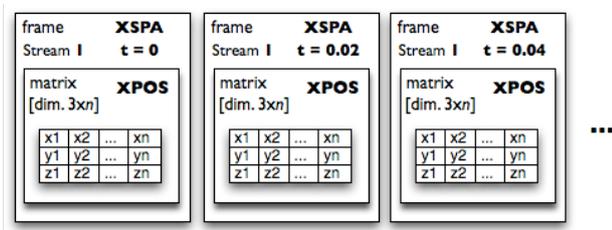


Figure 6. Séquence de *frames* SDIF "XSPA" décrivant l'évolution dans l'espace de n sources sonores. Chaque *frame* décrit les n positions dans une même matrice.

Enfin les deux cas envisagés précédemment peuvent être mis en œuvre simultanément, permettant ainsi de grouper certains éléments dans un flux commun et en parallèle à un certain nombre d'autres flux indépendants.

La figure 7 montre un exemple de description de trajectoires associées à des sources sonores, générée comme « matrice de spatialisation » dans OM et écrite sous forme de fichier SDIF suivant les précédentes spécifications. L'éditeur SDIF-Edit [6] est utilisé pour la visualisation de ces trajectoires.

6. CONCLUSION

L'accessibilité des techniques de spatialisation sonore et l'interpénétration de plus en plus courante des pratiques musicales "symboliques" et électroacoustiques font de la spatialisation un enjeu important des recherches en composition assistée par ordinateur.

Nous avons essayé de donner un aperçu de nos travaux actuels sur ce sujet, et des orientations choisies dans l'environnement OpenMusic. Les outils que nous proposons visent avant tout une ouverture et une généralité qui doivent leur permettre de s'adapter à la multiplicité des approches et aux diversités esthétique et opératoire des compositeurs. En particulier, nous avons mis en avant des choix de représentations et de formats dans l'objectif d'insérer notre environnement dans une interaction possible, et à différents niveaux des processus compositionnels, avec les différents outils de création musicale existant à l'heure actuelle.

7. REFERENCES

- [1] C. Agon, M. Stroppa, and G. Assayag. High Level Control of Sound Synthesis in OpenMusic. *Proc. International Computer Music Conference*, Berlin, Germany, 2000.
- [2] G. Assayag. Computer assisted composition today. *Ist symposium on music and computers*, Corfu, 1998.
- [3] G. Assayag, C. Rueda, M. Laurson, C. Agon and O. Delerue. Computer Assisted Composition at IR-CAM : From PatchWork to OpenMusic. *Computer Music Journal*, 23(3), 1999.
- [4] A. J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *Journal of the Acoustical Society of America*, 93, 1993.
- [5] R. Boulanger, editor. *The Csound Book. Perspectives in Software Synthesis, Sound Design, Signal Processing and Programming*. MIT Press, 2000.
- [6] J. Bresson and C. Agon. SDIF Sound Description Data Representation and Manipulation in Computer-Assisted Composition. *Proc. International Computer Music Conference*, Miami, USA, 2004.
- [7] J. J. Burred, C. E. Cella, G. Peeters, A. Röbel, and D. Schwartz. Using the SDIF Sound Description Interchange Format for Audio Features. *Proc. Inter-*

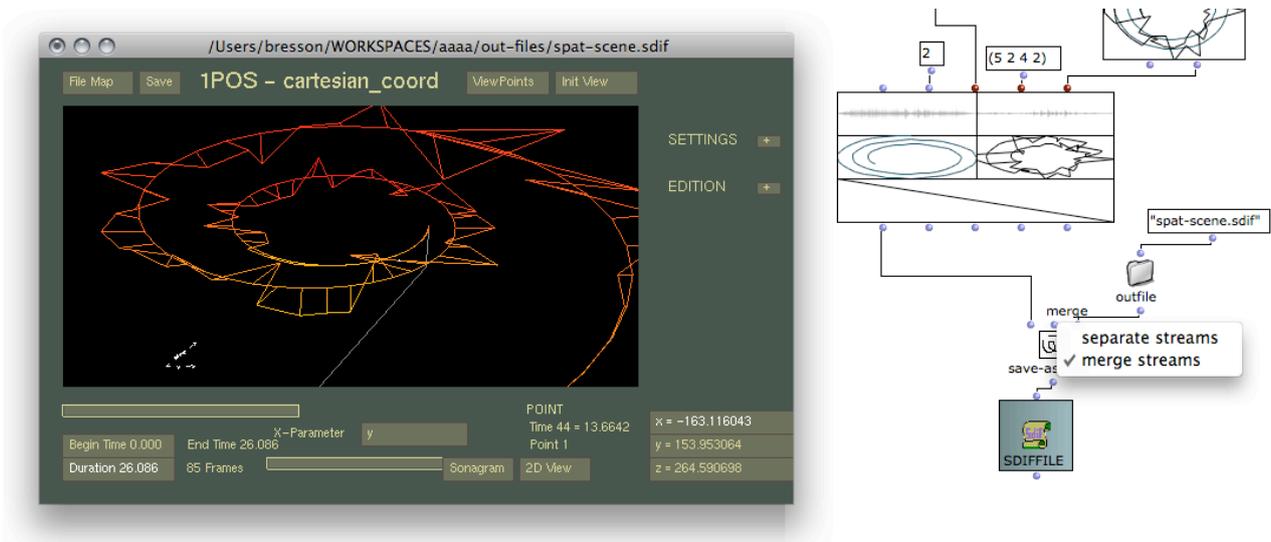


Figure 7. Sauvegarde de trajectoires sous forme de fichier SDIF. Dans cet exemple les trajectoires sont contenues dans un même *stream* SDIF (option “merge streams”) et donc visualisées simultanément. La visualisation courante sur la fenêtre SDIF-Edit (à gauche), montre l’évolution des valeurs x par rapport à y en fonction du temps, pour les deux trajectoires.

- national Conference on Music Information Retrieval*, Philadelphia, USA, 2008.
- [8] B. Cabaud and L. Pottier. Le contrôle de la spatialisation multi-sources – Nouvelles fonctionnalités dans Holophon version 2.2. *Actes des Journées d’Informatique Musicale*, Marseille, France, 2002.
- [9] J. Daniel. *Représentation de champs acoustiques, applications à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimedia*. PhD thesis, Université Pierre et Marie Curie - Paris 6, France, 2001.
- [10] O. Delerue. *Spatialisation du son et programmation par contraintes : le système MusicSpace*. PhD thesis, Université Pierre et Marie Curie - Paris 6, France, 2004.
- [11] O. Delerue and C. Agon. OpenMusic + MusicSpace = OpenSpace. *Actes des Journées d’Informatique Musicale*, Issy-les-moulineaux, France, 1999.
- [12] M. Geier and S. Spors. ASDF : Audio Scene Description Format. *Proc. International Computer Music Conference*, Belfast, Ireland, 2008.
- [13] J.-M. Jot and O. Warusfel. A Real-Time Spatial Sound Processor for Music and Virtual Reality Applications. *Proc. International Computer Music Conference*, Banff, Canada, 1995.
- [14] G. S. Kendall, N. Peters, and M. Geier. Towards an Interchange Format for Spatial Audio Scenes. *Proc. International Computer Music Conference*, Belfast, Ireland, 2008.
- [15] D. Kim-Boyle. Spectral Spatialization - an Overview. *Proc. International Computer Music Conference*, Belfast, Ireland, 2008.
- [16] MIDI Manufacturers Association, *Complete MIDI 1.0 Detailed Specification*, 1996.
- [17] G. Nouno and C. Agon. Contrôle de la spatialisation comme paramètre musical. *Actes des Journées d’Informatique Musicale*, Marseille, France, 2002.
- [18] N. Peters, S. Ferguson, and S. McAdams. Towards a Spatial Sound Description Interchange Format (Spat-DIF). *Canadian Acoustics*, 35(3), 2007.
- [19] L. Pottier. Dynamical Spatialisation of sound. HOLOPHON : a graphical and algorithmical editor for $\Sigma 1$. *Proc. International Conference on Digital Audio Effects (DAFx-98)*, Barcelona, Spain, 1998.
- [20] V. Pulkki. Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society*, 45(6), 1997.
- [21] M. Schumacher and J. Bresson. Compositional Control of Periphonic Sound Spatialization. *Second International Symposium on Ambisonics and Spherical Acoustics*, Paris, France, 2010.
- [22] D. Schwartz and M. Wright. Extensions and Applications of the SDIF Sound Description Interchange Format. *Proc. International Computer Music Conference*, Berlin, Germany, 2000.
- [23] T. Todoroff, C. Traube, and J.-M. Ledent. NeXTSTEP Graphical Interfaces to Control Sound Processing and Spatialization Instruments. *Proc. International Computer Music Conference*, Thessaloniki, Greece, 1997.
- [24] D. Topper, M. Burtner, S. Serafin. Spatio-Operational Spectral (S.O.S.) Synthesis. *Proc. International Conference on Digital Audio Effects (DAFx-02)*, Hamburg, Germany, 2002.

A MODEL FOR INTERACTIVE SCORES WITH TEMPORAL CONSTRAINTS AND CONDITIONAL BRANCHING

Mauricio Toro-Bermúdez
LABRI
mtoro@labri.fr

Myriam Desainte-Catherine
LABRI / SCRIME
myriam@labri.fr

Pascal Baltazar
GMEA
pb@gmea.net

ABSTRACT

Interactive Scores (IS) are a formalism for the design and performance of interactive multimedia scenarios. IS provide temporal relations (TR), but they cannot represent conditional branching and TRs simultaneously. We propose an extension to Allombert et al.'s IS model by including a condition on the TRs. We found out that in order to have a coherent model in all possible scenarios, durations must be flexible; however, sometimes it is possible to have fixed durations. To show the relevance of our model, we modeled an existing multimedia installation called Mariona. In Mariona there is choice, random durations and loops. Whether we can represent all the TRs available in Allombert et al.'s model into ours, or we have to choose between a timed conditional branching model and a pure temporal model before writing a scenario, still remains as an open question.

1. INTRODUCTION

Interactive Scores (IS) are a formalism for the design and performance of scenarios represented by *temporal objects (TO)*, *temporal relations (TR)* and discrete interactive events. Examples of TOs are videos, sounds or audio processors. TOs can be triggered by interactive events (usually launched by the user) and several TOs can be active simultaneously. TRs are used, for instance, to express precedence between objects, relations on their durations and explicit values for their durations.

IS have been subject of study since the beginning of the century [6], [9]. IS were originally developed for interactive music scores. Recently, the model was extended by Allombert, Desainte-Catherine, Laralde and Assayag in [4]. Hence IS can describe any kind of TOs, Allombert *et al.*'s model has inspired two applications: *iScore* [2] to compose and perform Electroacoustic music and *Virage* [5] to control image, video, audio and lights on live spectacles and interactive museums.

Allombert *et al.* showed on *iScore* and *Virage* that IS are successful to describe TRs, but IS have not been used to represent scenarios that require *conditional branching*. Conditional branching is commonly used in programming to describe control structures (e.g., *if/else* and *switch/case*). It provides a mechanism to choose the state of a program based on a condition and its current state.

As in programming, using conditional branching, a designer can create scenarios with loops, concurrent execution of multiple instances of the same TO and choice. Using conditional branching, the user or the system can take decisions on the performance of the scenario with the degree of freedom that the designer described –while the system maintains the TRs of the scenario. For instance, the designer can specify a condition to end a loop: “when the user change the value of the variable *end* to *true*, the loop stops”. The designer can also specify that such choice is made by the system: “the system non-deterministically chooses to stop the loop or to continue”.

Allombert *et al.* represent conditional branching and TRs separately, but there is not an unified way to represent conditional branching together with quantitative and qualitative TRs in the same scenario.

Quantitative TRs are those involving a proportional or explicit duration; for instance, “the duration of *A* is one third of the duration of *B*” or “the duration of *A* is 3 seconds”. On the other hand, *qualitative TRs* represent precedence between the start and end points of two TOs; for instance, “*A* must be played during *B*” or “*C* must be played after *D*”.

In this paper we propose a new model for IS. It extends Allombert *et al.*'s model by including a condition on the qualitative TRs. We do not include quantitative TRs because we found out that durations must be *flexible* (i.e., they can have any duration) to be coherent in all scenarios. However, we show that in some scenarios it is possible to respect *rigid* durations (i.e., durations with values in a finite interval).

The remainder of this paper is structured as follows. Section 2 shows related work on formalisms and applications for interactive multimedia. Section 3 presents our model for Interactive Scores. Section 4 explains the relation between our model and Allombert *et al.*'s model. Section 5 shows how to represent fragments of *Mariona*¹ in our model. Finally, section 6 gives concluding remarks and proposes future works.

2. RELATED WORK

Mariona is a multimedia installation, created by Pol Perez in 2007, that includes temporal relations, conditional branching, random durations, choice and hierarchy (i.e., an object can contain other objects). *Mariona* has a vision-based hand and body tracking system, speakers and a video display. A similar installation that uses motion sensors, instead of a tracking system, is described in [16]. Both installations are controlled by a Max/MSP [13] program –like most interactive multimedia applications.

2.1. Applications for Interactive Multimedia

In the domain of interactive music, there are applications such as *Ableton Live*². Using *Live*, the composer can write loops and the musician can control different parameters of the piece during performance.

An application to define a hierarchy and temporal relations among temporal objects is OpenMusic Maquettes [7]. Unfortunately, OpenMusic is designed for composition and not real-time interaction.

Another model related to Interactive Scores (IS) is *score following* [8]. Such systems “follow” the performance a real instrument and may play multimedia associated to certain notes in the score of the piece. However, to use these systems it is necessary to play a real instrument. On the other hand, using IS the user only has to control some parameters of the piece such as the date of the events, and the system plays the temporal objects described on the score.

2.2. Formalisms for Interactive Multimedia

To handle more complex synchronization patterns and to predict the behavior of interactive scenarios, formalisms such as the *Hierarchical Time Stream*

Petri Networks (HTSPN) [14] and the *Non-deterministic Timed Concurrent Constraint Programming (ntcc)* calculus [12] have been used to model IS in [4] and [3], respectively.

In HTSPN we can to express a variety of temporal relations, but it is not easy to represent global constraints (e.g., the number of temporal objects playing simultaneously). On the other hand, *ntcc* makes it possible to synchronize processes through a common constraint *store*, thus global constraints are explicitly represented in such store. We plan to write our model on *ntcc* because we can easily represent time, constraints, choice, and we can verify the model.

An advantage of using formal methods to model interactive multimedia is that they usually have automatic verification techniques. There are numerous studies to verify liveness, fairness, reachability and boundness on Petri Networks [11]. On the other hand, in the last years, calculi similar to *ntcc* have been subject of study for automatic model checking procedures [10].

For instance, using formal methods, it is possible to verify that will not be deadlocks, and also that certain temporal objects will be reached (played) during performance. This kind of properties are not possible to verify on applications for interactive multimedia with no formal semantics.

3. INTERACTIVE SCORES WITH CONDITIONAL BRANCHING

In this section we show how we can extend Allombert *et al.*'s model with conditional branching. We also show the new possibilities that our model offers.

Our model is based on the concept of *points*, and we provide relations for the points. The *before* relation is the only type of relation in our model. A relation *p before q* means that the execution of *q* is preceded by the execution of *p* if the condition in the relation holds. Although relations also have a nominal duration, it may change during the performance. The nominal duration is computed during the edition of the scenario using constraint programming.

Relations and temporal objects build up Interactive Scores (IS), thus a *score*³ (i.e., the specification of a scenario) is defined by a tuple $s = \langle T, R \rangle$, where *T* is a set of temporal objects and *R* is a set of relations. Relations and temporal objects are described using points.

¹http://www.gmea.net/activite/creation/2007_2008/pPerez.htm

²<http://www.ableton.com/live/>

³ We still use the term *score* for historical reasons.

3.1. Points

A *Point* is defined by $p = \langle D, b_p, b_s \rangle$, where $D \subseteq \mathbb{N}$ is the set of its possible dates of execution. Intuitively, we say that point p is a *predecessor* of q , if there is a relation p before q . On the other hand, we say that a point p is a *successor* of r , if there is a relation r before p .

There are two behaviors for a point. b_p defines whether the point waits until all its predecessors transfer the control to it –*Wait for All (WA)*– or it only waits for the first of them –*Wait for the First (WF)*–. b_s defines whether the point transfers the control to all its successors which conditions hold –*No Choice (NCH)*– or it chooses one of them –*Choice (CH)*–. In this model, a point can only belong to a single temporal object to avoid ambiguities.

3.2. Temporal Objects

A *temporal object (TO)* is defined by $t = \langle p_s, p_e, c, d, proc, param, N, vars \rangle$, where p_s is a point that starts a new instance of t and p_e ends such instance; c is a constraint attached to t (i.e., a constraint with local information for t and its children); d is the duration; $proc$ is a process which executes along with t ; $param$ are the parameters for the process; and N is the set of TOs embedded in t , which are called children of t . Finally, $vars$ represents the local variables defined for the TO. Local variables can be used by t 's childrens, process and local constraint.

The reader may notice that a TO does not provide a relation between its start and end points. For that reason, we must define a Timed Conditional Relation (TCR) between the start and the end points of the TO. We must also define a TCR between the start point of a father and the start point of at least one of its children.

3.3. Timed Conditional Relations

A *Timed Conditional Relation (TCR)* is defined by $r = \langle p_1, p_2, c, d, b, e \rangle$, where p_1 and p_2 are the points involved in the relation, c is the condition to determine whether the control *jumps* from p_1 to p_2 (i.e., the control is transferred from p_1 to p_2), d is the duration of the relation, b is the interpretation for c , e describes whether the condition is evaluated as soon as it holds, or at the end of the duration of the relation.

In what follows we explain some of the parameters of a TCR. We recall from [4] that a duration is *flexible* if it can take any value, *rigid* if it takes values between two fixed integers and *semi-rigid* if it takes values greater than a fixed integer. In our

model, we only guarantee the coherence of flexible durations. When the durations are flexible, there are two possible values for e : *now* or *wait*; otherwise, only *wait* is possible. Finally, there are two possible values for b : *when* means that if c holds, the control jumps; *unless* means that if c does not hold or its value cannot be deduced from the environment (e.g., $c = a > 0$ and $-\infty < a < \infty$), the control jumps.

3.4. Example: A Score with a Conditional Loop

The following example describes a score with a loop. During the execution, the system plays the sound B , a silence of one second, and the video C . If the signal *finish* becomes true, it ends the scenario after playing the video C ; otherwise, it comes back to the beginning of the sound B after playing the video C . To define the score of this scenario, we define a local boolean variable *finish* inside the structure A and we use it as the condition for the relations.

Figure 1 is a representation of the scenario. The duration for B is three seconds and the duration for C is four seconds (these values are calculated during the edition phase); however, they can change during the performance of the scenario depending on the behavior of the points and the relations.

The points have the following behavior. Point e_c (the end of the TO C) is enabled for choice and the other points transfer the control to all their successors. All the points wait for the first predecessor that transfer the control to them. Finally, the relations wait until their durations finish before evaluating their conditions, preserving rigid durations. Formally,

$$\begin{aligned}
 s_A &= s_B = \langle \{d, d \geq 0 \wedge d \% 8 = 0\}, WF, NCH \rangle \\
 s_C &= \langle \{d, d \geq 4 \wedge (d - 4) \% 8 = 0\}, WF, NCH \rangle \\
 e_A &= \langle \{d, d \geq 8 \wedge (d - 8) \% 8 = 0\}, WF, NCH \rangle \\
 e_B &= \langle \{d, d \geq 3 \wedge (d - 3) \% 8 = 0\}, WF, NCH \rangle \\
 e_C &= \langle \{d, d \geq 8 \wedge (d - 8) \% 8 = 0\}, WF, CH \rangle \\
 B &= \langle s_B, e_B, \text{true}, 3, \text{playSound}B, \phi, \phi, \phi \rangle \\
 C &= \langle s_C, e_C, \text{true}, 4, \text{playVideo}C, \phi, \phi, \phi \rangle \\
 A &= \langle s_A, e_A, \text{true}, 8, \text{silence}^4, \phi, \{B, C\}, \\
 &\quad \{finish\} \rangle \\
 T &= \{A\} \\
 R &= \{ \langle s_A, s_B, \text{true}, 0, \text{when}, \text{wait} \rangle, \\
 &\quad \langle e_B, s_C, \text{true}, 1, \text{when}, \text{wait} \rangle, \\
 &\quad \langle e_C, e_A, \text{finish}, 0, \text{when}, \text{wait} \rangle, \\
 &\quad \langle e_C, s_B, \text{finish}, 0, \text{unless}, \text{wait} \rangle, \\
 &\quad \langle s_B, e_B, \text{true}, 3, \text{when}, \text{wait} \rangle, \\
 &\quad \langle s_C, e_C, \text{true}, 4, \text{when}, \text{wait} \rangle \} \\
 s &= \{T, R\}
 \end{aligned}$$

⁴ *silence* is a process that does not perform any action.

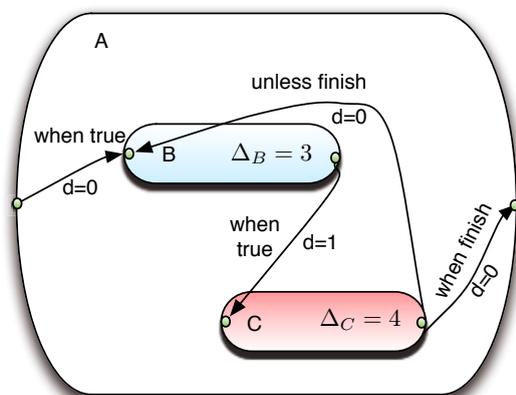


Figure 1. A score with conditional branching.

3.5. Multiple Instances Concurrently

Some scenarios require multiple instances of the same TO executing concurrently. This must be treated in a special way because not all the processes associated to a TO accept that sort of *polyphony*. Consider, for instance, a TO that displays a video. There may not be a defined behavior for more than one instance of it at the same time.

We propose four behaviors (fig. 2) to manage concurrent instances of the same TO: splitting them, delaying them, cancelling them or allowing them. This behavior must be selected by the designer of the scenario according to the nature of the TO.

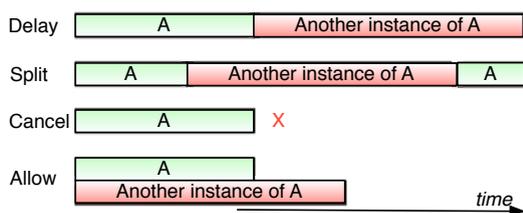


Figure 2. Multiple instances concurrently.

3.6. Limitations

In some cases (e.g., fig. 1), we can guarantee that rigid and semi-rigid durations will be respected during the performance of the scenario. Unfortunately, there is not a generic way to interrupt a rigid TO in a score that contains conditional branching. For instance, figure 3 shows a scenario where we cannot preserve the rigid durations of the TOs. T_1 , T_4 and T_5 have fixed durations, but T_1 can take different values between Δ_{min} and Δ_{max} . There is not a way to predict whether T_2 or T_5 will be chosen af-

ter the execution of T_1 , thus we cannot compute the duration of T_1 before the choice.

The problem is that choices do not allow us to predict the duration of the TO's successor; therefore, it is not possible to determinate *a priori* the duration of the TO.

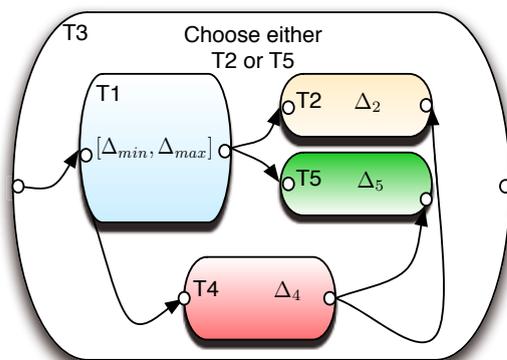


Figure 3. Limitation of rigid durations.

4. RELATION TO ALLOMBERT ET AL.'S MODEL

In this section we show the relation of Allombert *et al.*'s model and ours. First, we briefly present their model; then, we show the encoding of their model into ours.

4.1. Allombert et al.'s Model

Allombert *et al.*'s model is hierarchic: a temporal object (TO) can contain other TOs. A TO that has children is called a structure. There are two types of structures: *linear structures* have an unidirectional timeline with temporal relations inspired on Allen's relations [1]; and *logical structures* based on a state-transition model, equivalent a to Finite State Machine.

The problem of such classification arises when there is a linear structure nested on a logical structure or vice versa: there is not a semantic to handle children of a different type. In what follows, we will focus on representing *linear structures* on our model. Representing *logical structures* is trivial.

We recall some important notations from Allombert *et al.*'s model. A *score* is a tuple $s = \langle T, R \rangle$. A *temporal object* is defined by $t = \langle s, d, p, c, N \rangle$, where s is its start date, d is its duration, p is its attached process, c is its local constraint and N are its children. A *temporal relation* is defined by $r = \langle a, t_1, t_2 \rangle$, where a is an Allen's relation. We recall that we do not consider linear temporal

relations (e.g., the duration of A is k times the duration of B) because they are quantitative temporal relations.

4.2. Encoding Allombert *et al.*'s Model

The function $[[\cdot]]$ takes a score $s' = \langle T', R' \rangle$ on Allombert *et al.*'s model and returns a score $s = \langle T, R \rangle$ on our model. To encode their model, the behavior of each point is to transfer the control to each of its successors and to wait until all its predecessors transfer the control to it.

A TO $t' = \langle s, d, p, c, N \rangle$ is codified into $t = \langle p_s, p_e, c, d, p, \phi, \{x | x \in [[y]] \wedge y \in N\}, \phi \rangle$. This definition basically codifies a TO and all its children recursively. We also need to include a Timed Conditional Relation (TCR) $\langle p_s, p_e, \text{true}, d, \text{when}, \text{wait} \rangle$ for each TO.

A TR $r' = \langle a, t_1, t_2 \rangle$ is codified into a TCR $r = \langle p_1, p_2, \text{true}, \text{duration}, \text{when}, \text{wait} \rangle$. Note that the condition is always `true` and its behavior is `when`, thus the condition is always valid. In addition to the value of the condition, the end behavior `wait` is crucial to define Allen's relations.

4.2.1. Representing Allen's Relations

Allen's relations represent qualitative TRs between two objects (fig. 4). On Allombert *et al.*'s model, they post constraints during the edition phase to maintain Allen's relations during the execution. For instance, for a relation B *meets* A , they post the constraint $start(A) = start(B) + duration(B)$.

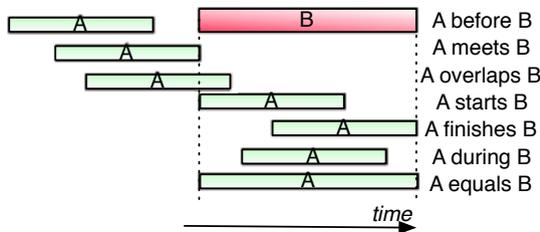


Figure 4. Allen's relations.

We propose a representation of Allen's relations in terms of our *before* relation. For instance, the relation A *meets* B is represented by a relation $\langle e(A), s(B), \text{true}, 0, \text{when}, \text{wait} \rangle$, where e and s are functions that return the end dates and the start dates of a TO, respectively. Figure 5 shows the representation for Allen's relations on our model.

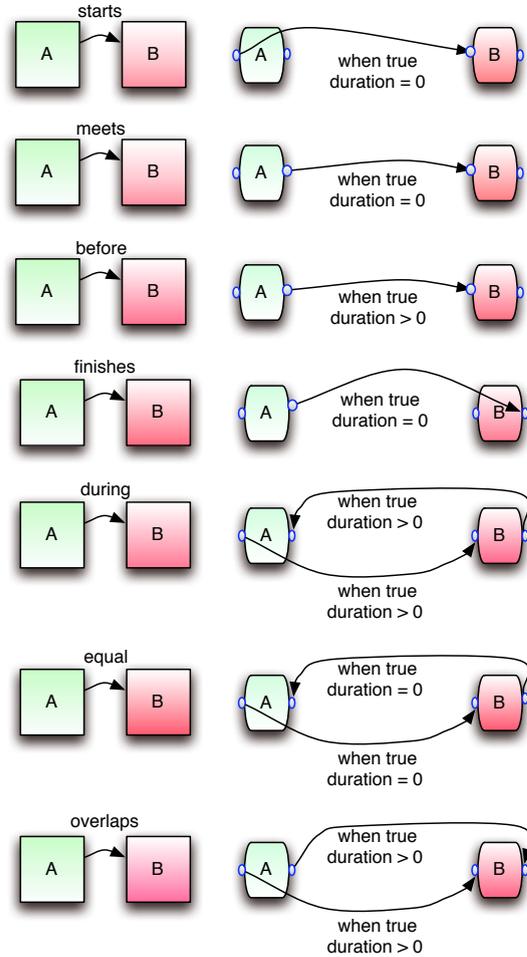


Figure 5. Representing Allen's relations with Timed Conditional Relations.

5. STUDY CASE: MARIONA

Mariona (french acronym for automatic machine with memory, iconography, orinic, narrative and acoustic) is a multimedia installation capable to generate images, analyze the movements of the users and produce sounds. Its control is described by three temporal objects (TO) that interact concurrently: *Global*, *Speed*, *Aléatoire* (random). In its current implementation, these TOs are written on Max/MSP.

Figure 6 describes the main TOs and *Bug*. Figure 7 describes some TOs contained in *Global*. The notation used in the figures is the one used by Pol Perez. In *Mariona* –as we may see in figures 6 and 7–, there is choice, trans-hierarchical relations, random durations and loops (finite and infinite). In what follows, we give some examples about these phenomena and how we can model them.

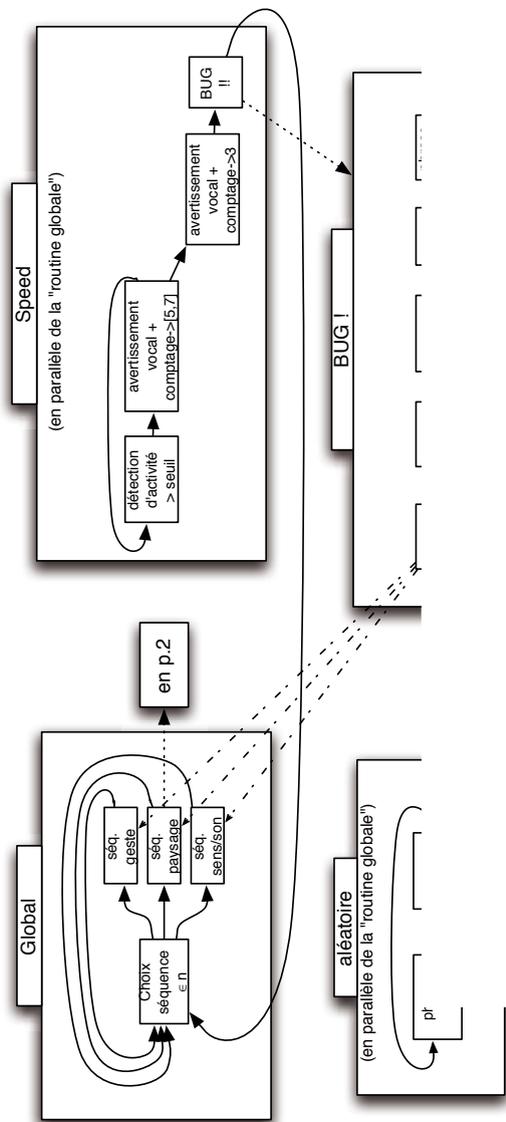


Figure 6. Marion: *Global*, *Speed*, *Aléatoire* and *Bug* TOs

5.1. Choice

Choice is the main application of conditional branching. It makes it possible to post different relations among TOs during the edition of the scenario and then choose one during performance. For instance, in the *Global* TO, the system makes a choice among three different TOs: *seq. geste*, *seq. paysage*, *seq. sens/son*.

5.2. Loops and Random Duration

In the *aléatoire* (random) TO, we may observe loops and a random delay from one to five seconds (*at-*

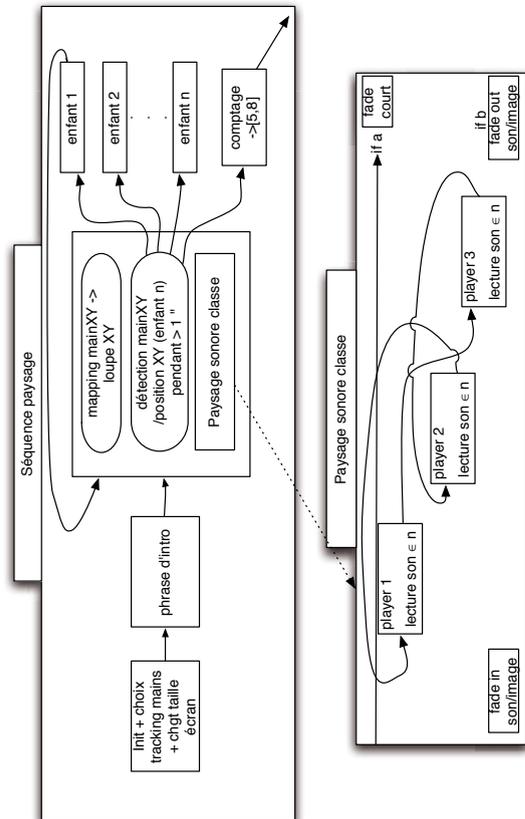


Figure 7. Marion: *Séquence paysage* and *Paysage sonore classe* TOs.

tente [1'-5']). There is also a random duration in the *flash* and *bip* (beep) TOs, which are inside the *Bug* TO.

5.3. Trans-hierarchical Relations

Timed Conditional Relations are usually defined between brothers (i.e., two objects that are children of the same TO). For some scenarios, it is necessary to transfer the control between two TOs with different parents. The conditions for these relations are evaluated in the environment (i.e., the local variables and the local constraint) where it starts. For instance, in Marion, the *Bug* TO (contained in *Speed*) finishes either *Séq. geste*, *Séq. paysage* or *Seq. sens/son* (contained in *Global*). In our model based on points this does not pose any problem.

6. CONCLUDING REMARKS

We presented a new model for Interactive Scores (IS) capable to represent conditional branching together with temporal relations. The model is based

on points. Temporal objects (TO) and Timed Conditional Relations (TCR) are built upon such points. As far as we know, there are not related work on models for interactive multimedia that support conditional branching besides Allombert *et al.*'s model.

We encoded Allombert *et al.*'s model into our model. We found out that without rigid and semi-rigid durations, nor linear temporal relations, it is always possible to transform a temporal relation into a TCR; although sometimes it is also possible to preserve rigid and semi-rigid durations. Unfortunately, rigid, semi-rigid and random durations cannot be always preserved in scenarios with conditional branching.

Whether we can express all the temporal relations presented Allombert *et al.*'s model (e.g., linear temporal relations and rigid durations) into ours, or have to choose between a timed conditional branching model and a pure temporal model before defining a scenario, still remains as an open question.

6.1. Future Work

We want to explore how we can have rigid, semi-rigid durations, and conditional branching simultaneously. We already determined that it is not possible in the general case, but we want to know in which cases we can have such type of durations. We also want to model random durations as the ones presented on Mariona.

Another research line is the *waiting behavior* of TOs. For instance, when a TO is supposed to last five seconds, but it lasts ten seconds during performance, what should be the behavior for the additional five seconds?

Finally, we plan to extend the `ntcc` model for IS [3] to support conditional branching and execute it on *Ntcrt* (a real-time capable interpreter for `ntcc`) [15]. We also plan to develop automatic verification tools for *Ntcrt* in the lines of [10].

7. ACKNOWLEDGEMENTS

We want to thank Antoine Allombert, Matthias Robine and Raphael Marczak for their valuable comments on our model.

8. REFERENCES

- [1] J. F. Allen. Maintaining knowledge about temporal intervals. *Communication of ACM*, 26, 1983.
- [2] A. Allombert, G. Assayag, and M. Desainte-Catherine. Iscore: a system for writing inter-
action. In *Proc. of DIMEA '08*, pages 360–367, New York, NY, USA, 2008. ACM.
- [3] A. Allombert, G. Assayag, M. Desainte-Catherine, and C. Rueda. Concurrent constraint models for interactive scores. In *Proc. of SMC '06*, May 2006.
- [4] A. Allombert, M. Desainte-Catherine, J. Larralde, and G. Assayag. A system of interactive scores based on qualitative and quantitative temporal constraints. In *Proc. of the 4th International Conference on Digital Arts (Artech 2008)*, Novembre 2008.
- [5] P. Baltazar, A. Allombert, R. Marczak, J. Couturier, M. Roy, A. Sèdes, and M. Desainte-Catherine. Virage : Une réflexion pluridisciplinaire autour du temps dans la création numérique. In *Proc. of JIM*, 2009.
- [6] A. Beurivé and M. Desainte-Catherine. Representing musical hierarchies with constraints. In *Proc. of the 7th International Conference on Principles and Practice of Constraint Programming, Musical Constraints Workshop, Paphos*, 2001.
- [7] J. Bresson, C. Agon, and G. Assayag. Openmusic 5: A cross-platform release of the computer-assisted composition environment. In *10th Brazilian Symposium on Computer Music*, Belo Horizonte, MG, Brésil, Octobre 2005.
- [8] A. Cont. Antescofo: Anticipatory synchronization and control of interactive parameters in computer music. In *Proceedings of International Computer Music Conference (ICMC)*. Belfast, August 2008.
- [9] M. Desainte-Catherine and N. Brousse. Towards a specification of musical interactive pieces. In *Proc. of CIM XIX, Firenze, Italy.*, May 2003.
- [10] M. Falaschi and A. Villanueva. Automatic verification of timed concurrent constraint programs. *Theory Pract. Log. Program*, 6(4):265–300, 2006.
- [11] T. Murata. Petri nets: Properties, analysis and applications. In *Proc. of the IEEE*, volume 77, pages 541–580, April 1989.
- [12] M. Nielsen, C. Palamidessi, and F. Valencia. Temporal concurrent constraint programming: Denotation, logic and applications. *Nordic Journal of Computing*, 1, 2002.

- [13] M. Puckette, T. Apel, and D. Zicarelli. Real-time audio analysis tools for Pd and MSP. In *Proc. of ICMC 98*, 1998.
- [14] P. Sénac, P. d. Saqui-Sannes, and R. Willrich. Hierarchical time stream petri net: A model for hypermedia systems. In *Proc. of the 16th International Conference on Application and Theory of Petri Nets*, pages 451–470, London, UK, 1995. Springer-Verlag.
- [15] M. Toro-Bermúdez, C. Agón, G. Assayag, and C. Rueda. Ntcrt: A concurrent constraint framework for real-time interaction. In *Proc. of ICMC 09*, August 2009.
- [16] T. Yamauchi and T. Iwatake. An interactive installation through spatial sensing. In *Proc. of the 4th International Mobile Music Workshop*, May 2007.

Session 2

Musique mixte

PENSER L'ŒUVRE MUSICALE MIXTE À LA LUMIÈRE DE LA MATHÉMATIQUE DES EXTENSIONS

François Nicolas

Compositeur, chercheur associé Ens-Ircam
fnicolas@ens.fr / fnicolas@ircam

RÉSUMÉ

Appelons *œuvre musicale mixte* l'œuvre musicale qui accueille un flux temporel non musical (un poème récité/chanté, un ballet, une vidéo...) pour lui entrelacer son propre flux musical – le modèle le plus simple en sera celui du lied (à la condition – nécessaire – que le texte chanté reste intelligible).

On se propose d'analyser une puissance musicale particulière de ce type d'œuvre, puissance qui tient non seulement au fait que la musique s'y trouve « éclairée » par le poème-ballet-vidéo mais, plus encore, que la musique, qui pourtant ne parle pas, s'y trouve en retour apte à « éclairer » le sens littéraire du texte, ou le sens chorégraphique du ballet, ou le sens plastique de la vidéo.

Appelons « aura » de l'œuvre cet effet général d'éclairage dont elle est le siège.

On analysera, à la lumière des mathématiques, comment une telle irradiation est possible et de quelle manière concrète la musique contrôle, pas à pas et de l'intérieur de l'œuvre, l'engendrement d'une telle « aura ».

Pour ce faire, on s'appuiera

- sur les extensions algébriques de corps (singulièrement celle du corps \mathbb{Q} des rationnels) ;
- sur les extensions génériques d'ensembles dénombrables (dans la théorie du *forcing* de Paul Cohen).

1. INTRODUCTION

Rehaussons sur un point précis l'éclairage dont les mathématiques sont capables pour le musicien soucieux de mieux comprendre théoriquement son art.

Le point musical spécifique qu'il s'agit d'examiner ici est le suivant : d'où procède cette capacité propre de l'œuvre musicale *mixte* d'engendrer « une aura poétique » qui l'irradie ?

Qu'en un tel point, corrélant compréhension musicale et appréhension imaginaire d'une œuvre, *a priori* très éloigné de la mathématique, celle-ci puisse guider le musicien devrait valoriser, s'il en était besoin, la variété des *raisonances* (ou *résonances entre raisons hétérogènes*) dont la mathématique est porteuse pour peu qu'on la désenclave de sa seule figure calculatrice et qu'on lui reconnaisse la puissance de pensée dont elle

fait patiemment et obstinément preuve depuis l'Antiquité.

2. PROBLÉMATIQUE

Partons de ce constat : l'œuvre musicale mixte – entendue, en première approche, comme cette œuvre musicale accueillant aux côtés de son propre développement un flux temporel composé selon une tout autre logique que musicale (un texte, un discours fait d'images visuelles, une représentation scénique, une chorégraphie...) – met en jeu quelque rayonnement singulier (que ne génère pas l'œuvre purement musicale) qui donne forme à une extension de l'œuvre qu'on propose d'appeler son « aura poétique ».

Comme on va l'examiner attentivement, et pour des raisons précises qu'on s'emploiera à clarifier à la lumière des mathématiques, cette extension de l'œuvre musicale mixte – son « aura poétique » – va se présenter comme sensation à la fois certaine et floue (il est certain qu'il y a bien une telle extension, mais sa manière propre d'être *une* reste floue), comme irradiation à la fois patente et difficilement objectivable comme partie propre.

En première approche, « aura poétique » nommera ici une capacité que l'œuvre musicale mixte a en propre grâce à ses interactions internes entre flux musical et flot non-musical : une capacité de s'entourer d'un halo « poétique » de « sens », de s'augmenter selon un ordre « poétique » et plus seulement musical en sorte d'acquérir une aptitude à cette « signifiante » qu'on entendra comme cette capacité générique à la signification qui n'a nul besoin d'une signification précise pour s'attester. L'œuvre musicale mixte apparaît ainsi comme œuvre musicale élargie ou étendue : non pas seulement œuvre musicale accompagnante ou accompagnée par un projet non musical, capable donc de coexister avec d'autres logiques que musicales, mais plus encore œuvre où la musique atteste de sa capacité à se dialectiser intimement avec autre chose qu'elle.

2.1. Une expérience sensible ordinaire...

Tout ceci, somme toute, entreprend de formuler dans la langue commune cette expérience musicale tout à fait ordinaire et originaire : un lied, un vrai lied (non pas un texte accompagné d'une mélodie mais une musique autoconstituée s'enlaçant à un texte lui-même

littérairement autoconstitué en sorte que flux musical et flot textuel restent appréhendés à la fois en leur autonomie et en leur enlacement), c'est à la fois une musique qui sert un texte et un texte qui nourrit une musique, mais c'est également une musique qui éclaire un texte de manière finalement assez mystérieuse (s'il est vrai qu'il ne s'agit pas là d'illustration, d'imagerie ou de commentaires) : ainsi, qui n'a pas, à écouter un tel lied, le sentiment qu'il matérialise une musique agrandie à mesure de sa capacité à nourrir musicalement un texte littéraire ?

Cette capacité musicale (plutôt que littéraire) à nourrir un texte littéraire (plutôt que musical) se donne dans le sensible, non par des mots ; elle procède du travail de l'oreille musicale tout en restant difficilement délimitable sous la forme d'opérations compositionnelles précises (qu'on pourrait auditivement objectiver). D'où que l'auditeur confronté à cette expérience sensible recourra spontanément aux termes d'« atmosphère » musicale ou de « climat » pour désigner cette capacité proprement musicale d'irradier le sens du texte plutôt que de le commenter.

2.2. Une extension moins commune

Le point qui nous intéresse plus particulièrement est que cette aptitude musicale à irradier le texte poétique puisse, dans certaines conditions, se trouver appréhendée comme extension (« aura ») poétique de l'œuvre musicale : non seulement comme geste extraverti d'éclairage mais comme nouvel état de l'œuvre – état « étendu » –, cette nouveauté d'état relevant alors d'un ordre non exclusivement musical (il y s'agit moins d'une musique extérieurement « enrichie » que d'une musique enrichie de sa capacité à enrichir autre chose qu'elle) qu'on dira donc, pour cette raison, extension « poétique ».

On insistera donc ici sur une dissymétrie inhérente à cette réciprocité musique/texte : si le texte, à sa manière propre, éclaire bien la musique, l'œuvre musicale mixte est avant tout éclairage musical du texte, et cette capacité musicale immanente de « nourrir » le texte en auréolant musicalement son sens agrandit l'œuvre musicale d'une puissance poétique matériellement inscriptible non comme auréole poétique (comme ajout exogène entourant extérieurement l'œuvre et par là détachable et effaçable) mais comme irradiation endogène transfigurant son *inspect* sensible (c'est-à-dire son apparence telle qu'elle se donne de l'intérieur d'elle-même, de manière endogène, plutôt que son *aspect* exogène).

2.3. Et une question encore moins ordinaire...

Notre problème est donc le suivant : s'il est vrai qu'une telle aura poétique de l'œuvre musicale mixte existe bien, comment rendre compte de son effet sensible, de son engendrement musical (dans la composition puis dans son interprétation), comment

somme toute l'œuvre musicale mixte contrôle-t-elle effectivement la production d'une telle aura poétique ?

On devine que ce type de questionnement n'appartient guère aux soucis ordinaires du musicien en train de faire de la musique. Celui-ci, qu'il soit compositeur ou interprète de lieds, de ballets, ou d'opéras, s'attachera à la résolution pratique de ce type de questions en mobilisant son métier, sa connaissance intuitive de la langue mobilisée (ou de la vidéo ou de la danse ou du théâtre...) : il engendrera l'aura poétique dont il est ici question non pas sans le savoir, pas davantage sans le réfléchir, mais plus exactement sans le verbaliser, sans expliciter et discriminer les opérations effectuées à cette fin propre.

Si le musicien, réfléchissant sur son travail, s'attache un temps aux questions précédentes, ce n'est pas tant en vue d'une fin pratique – tous les musiciens composent de la même manière (intuitivement, en faisant confiance à leur sensibilité musicale propre, à leur expérience et à leur métier), et le musicien (réflexif ou pensif) ne cherche pas ici des « recettes », ou des « procédures ». Il en va plutôt en ces questions d'un problème d'orientation de pensée, singulièrement d'orientation réfléchie et explicitée, thématisée et argumentée. Sans plus insister sur ce point, indiquons malgré tout qu'il en va ici – en matière d'aura poétique propre à l'œuvre musicale mixte – d'une question sans doute décisive pour l'œuvre musicale contemporaine s'il est vrai que la musique est dans un moment singulier où elle doit remettre sur le chantier cette directive générale d'Adorno : « *L'art a besoin de quelque chose qui lui est hétérogène pour devenir art.* » [1]

*

Pour nous attaquer à notre question, commençons par formaliser plus précisément ce que nous entendons ici par *œuvre musicale mixte*.

3. « ŒUVRE MUSICALE MIXTE » ?

3.1. Cinq conditions

On appellera œuvre musicale mixte (OMM) une œuvre répondant simultanément à cinq conditions.

1) Il y a, au principe de cette œuvre, un flux musical composé selon une logique musicale propre. La musique s'y expose en son autonomie de déploiement ; elle n'est pas là pour illustrer un développement exogène (celui d'un texte narratif par exemple), pour le commenter ou l'ornementer. La musique dispose de sa pleine autonomie logique, garantie sous sa forme traditionnelle de partition.

2) Le flux musical y est associé à un autre flux temporel, composé par ailleurs et selon une tout autre logique que musicale : littéraire, théâtrale, cinématographique.... Il y a donc au départ de l'œuvre une mise en parallèle de deux flux hétérogènes qui vont partager le même temps chronométrique.

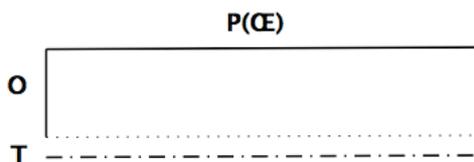
3) Ce flux hétérogène est lui-même présenté sous sa forme sensible naturelle : par exemple, si le texte y a, ce texte sera prononcé et ce de manière qu'il reste auditivement compréhensible (plutôt que simplement consigné dans des notes d'exécution ou exposé aux seuls regards par quelque prompteur). On exclut donc, ce faisant, que ce texte ne constitue qu'un programme littéraire sous-jacent, ou que sa présentation au cœur de la musique conduise à effacer son hétérogénéité logique en raturant sa compréhension auditive.

4) C'est la musique qui assure la coordination étroite et précise de ces deux flux. Le rapport musique/texte (ou musique/images ou musique/danse) se trouve ainsi dissymétrisé au profit de la musique : la musique *dirige* le rapport, quand la poésie ou la danse sont dirigées. La musique constitue donc ici l'aspect principal, celui qui ultimement commande la logique même de l'œuvre mixte. C'est bien à ce titre qu'on parle ici d'œuvre *musicale*, et non pas d'œuvre littéraire (un poème), chorégraphique (un ballet) ou cinématographique (un film).

5) La musique exerce sa direction générale au moyen de ses opérateurs traditionnels d'écriture et de notations, donc de sa partition propre. À ce titre, le flux temporel non musical se trouve, d'une manière ou d'une autre, inscrit dans la partition et scripturalement contrôlé, au moins dans ses grandes lignes : par exemple, le déroulement du texte sera noté dans la partition (comme on le fait usuellement pour un texte à chanter ou à réciter), comme le seront – au choix – les principales articulations d'une chorégraphie, d'une action théâtrale ou du discours plastique d'une vidéo.

3.2. Un schéma

Au total, l'OMM peut être ainsi schématisée par sa partition P(Æ) :



où O désigne la partie proprement musicale de la partition (son système de portées orchestrales) et T ce qui y note le flux hétérogène horizontalement synchronisé avec la musique (songeons tout simplement au texte chanté inscrit sous la portée consignante la mélodie du chant).

4. QUELQUES REMARQUES PRÉLIMINAIRES

4.1. S'il s'agit d'un texte...

Si le flux hétérogène est composé d'un texte, la présentation de ce texte pour l'auditeur de l'œuvre (qui n'est pas *a priori* un lecteur de la partition) doit être compréhensible dans l'ordre propre du sensible : une œuvre musicale qui disloquerait son texte au point de

rendre son flux signifiant insaisissable d'oreille¹ ne saurait constituer une œuvre « mixte » au sens où nous l'entendons ici. Une caractéristique décisive de l'OMM est donc que la logique de consistance propre au flux hétérogène reste exposée dans son ordre sensible propre – comme on va y revenir, ceci constitue une condition nécessaire pour que la dialectique propre à l'OMM puisse opérer.

4.2. S'il s'agit d'un flot visuel...

Si le flux hétérogène est composé d'un film, d'une représentation scénique, d'une chorégraphie, il faut que l'organisation globale de ce flux – ses articulations, périodisations, segmentations – soit représentée dans la partition P(Æ) en sorte que celle-ci contrôle effectivement la synchronisation des deux flux et par là les principales arêtes de leur dialectisation.

Précisons : on n'attendra pas nécessairement que la partition contrôle la logique immanente du discours non-musical (par exemple, il ne sera pas nécessaire que la partie T de la partition P inscrive en détail la logique minutieuse de l'éventuelle chorégraphie, en recourant par exemple à un des types existants de notation, Laban ou autres...) mais il sera nécessaire que la partition contrôle la synchronisation précise du flux hétérogène avec le flux proprement musical (dont par ailleurs elle contrôle la logique immanente propre).

On exclura donc ici de l'OMM les duos du type Cunningham/Cage où chorégraphe et musicien travaillaient indépendamment l'un de l'autre pour ne superposer leur flux, en un même temps chronologique, qu'au moment ultime de la représentation. Sans préjuger ici de l'intérêt artistique d'une telle méthode de travail, on posera simplement que, dans de telles improvisations des rencontres, il ne saurait y avoir composition d'une véritable dialectique (entendue à minima comme interactions soutenues et développées, disciplinées et conséquentes) entre flux relevant de logiques hétérogènes.

4.3. Dissymétrie au profit de la musique

Le point précédent relève la dissymétrie au principe de l'OMM : la musique constitue la partie dirigeante de l'OMM, ce qui s'assigne matériellement au fait qu'elle s'y trouve intégralement écrite, que sa logique propre y est donc explicite et que la synchronisation dialectique entre différents flux se trouve contrôlée et mise en œuvre *du point de cette musique composée*.

Dit autrement, la musique de l'OMM se trouve dotée de l'intégralité de ses atouts logiques et se présente à l'auditeur en tirant parti maximal de son autonomie de pensée et d'exposition. Tel n'est pas, le plus souvent, le cas du flux hétérogène.

Pour prendre un exemple, si l'OMM articule musique et flot visuel d'images (vidéo ou « cinéma »), il ne

¹ Comme Boulez par exemple aime à le faire, en particulier pour les poèmes de Mallarmé qu'il insère dans son chant...

saurait être ici question d'une simple musique de film entendue comme contribution musicale subordonnée à une logique globalement cinématographique. Une œuvre cinématographique, sa musique comprise – disons un film - ne constitue donc pas en tant que telle une OMM au sens où nous l'entendons ici ; tout au plus constituerait-elle éventuellement ce qu'on pourrait appeler une « œuvre cinématographique mixte » (si ce terme ne constituait pas déjà par lui-même un pléonasme en raison du caractère essentiellement « impur » de l'art cinématographique).

4.4. La musique de l'OMM est écrite

Si la synchronisation et la dialectique entre flux est musicalement contrôlée de manière très précise, ceci interdit que l'OMM soit musicalement improvisée.

S'il est vrai qu'improvisation et composition ne misent pas sur les mêmes disciplines – l'improvisation mise sur le charme éphémère des conjonctions imprévues, la composition sur la puissance du faisceau interprétatif d'une partition intégralement rédigée -, alors l'OMM relève musicalement de la composition, non de l'improvisation.

4.5. La musique, qui dirige, n'hésite pas à violenter le flux hétérogène qu'elle accueille.

Corrélat de sa disposition fermement dirigeante : dans l'OMM, la musique s'autorise de transformer le flux hétérogène qu'elle accueille, par exemple de ne pas le restituer dans son intégralité (on pourra ainsi chanter un poème sans pour autant chanter tous ses vers ou toutes ses strophes), de couper dans ce flux, d'y insérer des blancs, ou de lui imposer des répétitions tenues pour musicalement bienvenues.

La règle est alors la suivante : la musique, accueillant en son intimité propre un flux hétérogène, préserve la logique spécifique de ce flux (donc en particulier son ordre chronologique) en sorte que la compréhension de son sens reste assurée. Mais il n'est pas pour autant requis que ce respect par la musique de la singularité logique de son invité conduise à effacer ses prérogatives d'hôte et l'oblige par exemple à restituer *tout* le texte accueilli selon le principe non musical d'une sorte de dictée continue. Respecter musicalement l'autonomie de sens et de forme d'un poème ainsi « mis en musique » n'interdit nullement de le déformer, de le replier, de le contraindre localement ou même régionalement (ce qui n'est pas dire le disloquer, l'écarteler, ou l'éparpiller).

4.6. Synthétisons...

Synthétisons ces contraintes en posant que l'OMM capte l'*aspect* du flux hétérogène, épouse son *intension* (c'est-à-dire son projet constituant) mais assume de déposer son *inspect* propre (soit sa forme saisie de l'intérieur, et non plus en vis-à-vis extérieur).

C'est en ce sens précis que la musique assume, dans l'OMM, de faire violence au flux hétérogène en le

remodelant selon un *inspect* proprement musical : par exemple, le texte chanté pourra être fragmenté, localement distendu, régionalement précipité selon des exigences proprement musicales.

4.7. Violence n'est pas brutalité

L'orientation ici en jeu se condense alors sur ce point traditionnel (dont Jean Genet nous a particulièrement instruit) : violence n'est pas brutalité, et la violence nécessaire dans les rapports entre libertés hétérogènes n'implique nullement leur brutalisation.

La violence impose par la force à une liberté quelque chose qu'elle ne veut pas (ici la musique impose au poème une répétition, une suspension, l'effacement de quelques vers dont le poème n'a nulle envie propre²). La brutalité consisterait à surajouter à cette violence, qui génère une résistance légitime de la liberté violentée, le projet de casser cette résistance et de briser cette liberté rétive *comme telle*. La brutalité ne tolère pas que la liberté violentée regimbe ; elle lui impose alors soit de donner son accord à cette violence qu'on exerce sur elle (c'est le vieux thème de la paix des vainqueurs), soit d'être purement et simplement annihilée comme liberté propre.

Dans l'OMM, la musique violente l'*inspect* du flux hétérogène mais respecte localement ou régionalement son aspect propre (son allure littéraire par exemple) et globalement son *intension* (l'affaire, ici, est plus subtile à décrire, et moins aisément objectivable : disons que la musique s'attache à capter les *raisonances* musicales de l'*intension* littéraire ou cinématographique ou chorégraphique ou théâtrale à l'œuvre et ne lui est donc nullement indifférente³).

*

Cette acception de l'OMM posée, comment comprendre cette puissance génératrice d'aura poétique qui y trouve ses conditions spécifiques d'émergence ? Pour cela, précisons ce que nous entendons par « aura poétique ».

5. « AURA POÉTIQUE » ?

La notion d'*aura* ici convoquée ne recouvre pas exactement celle, très célèbre, que Walter Benjamin a proposée en 1936 dans *L'œuvre d'art à l'époque de sa reproductibilité mécanisée* :

« Qu'est-ce en somme que l'aura ? Une singulière trame de temps et d'espace : apparition unique d'un lointain, si proche soit-il. »⁴

Si notre « aura poétique » pourra bien être conçue comme « une singulière trame de temps et d'espace » (énoncé qui, à dire vrai, n'engage pas à grand-chose), par contre elle ne sera pas tant « l'unique apparition d'un lointain, si proche soit-il » que l'effet sensible

² Rappelons-nous Victor Hugo se désolant que l'on couche de la musique le long de ses vers...

³ comme elle l'était nécessairement dans le duo Cunningham-Cage...

⁴ [2] p. 183

d'une transfiguration interne de l'œuvre, d'une irradiation globale à l'œuvre.

À ce titre, notre « aura poétique » se rapprochera plutôt de ce qu'Adorno entendait pour sa part sous le terme *aura* et qui touchait chez lui à « l'atmosphère » des œuvres d'art entendue comme « le mélange de leur effet et de leur constitution interne »⁵ :

« Ce qu'on appelle aura est familier à l'expérience artistique sous le nom d'atmosphère de l'œuvre d'art dans la mesure où le rapport entre ses éléments renvoie au-delà de ceux-ci et permet à chacun des éléments particuliers de se dépasser. »⁶

« La musique est l'art auratique par excellence : une cohérence musicale se réalise à la seule condition que ce qui apparaît isolément soit davantage que lui-même, qu'il soit transcendé vers une non-présence, un éloignement, et cette relation significative, médium de toute logique musicale, produit sans doute inmanquablement quelque chose comme une atmosphère qui enveloppe la musique. »⁷

L'« aura atmosphérique »⁸ tient ainsi au point que « l'œuvre d'art renvoie au-delà d'elle-même. »⁹

Notre « aura poétique » s'attachera en effet à une atmosphère créée spécifiquement par l'OMM autour d'elle, une atmosphère dont on pourra dire qu'elle l'enveloppe au sens où elle renverra à un au-delà de l'œuvre qu'on dira « poétique ».

Si, au sens strict, l'aura que génère l'OMM ne lui appartient pas – c'est en ce sens que cette aura sera dite *poétique* plutôt que strictement *musicale* –, cette aura procède cependant bien d'un travail intérieur à l'œuvre musicale quand elle est irriguée et fécondée par un flux hétérogène qui la pénètre. Il faut alors la concevoir moins comme une bordure extérieure, moins comme une sorte d'aurole séparable la contournant que comme constituant un nouvel éclairage interne servant de base à une adresse rayonnante de l'œuvre.

Et, bien sûr, cette aura va relever du sensible, du travail de l'oreille. C'est donc bien d'une sensation d'aura qu'il faut tenter d'objectiver l'engendrement.

6. DYNAMIQUE GÉNÉRALE

6.1. Deux temps

La dynamique générale de notre compréhension théorique de cette opération d'engendrement va s'articuler en deux temps :

- un premier temps de fécondation (ou d'éclairage) de la musique propre à l'OMM par le flux hétérogène qu'elle accueille – on nommera « modulation » ce premier temps ;

- un temps d'effet en retour de la musique ainsi modulée vers le flux hétérogène (ou irradiation), qu'on appellera « extension ».

⁵ [3] p. 349

⁶ [3] p. 347-348

⁷ [4] p. 98

⁸ [4] p. 99

⁹ [3] p. 348

6.2. Première formalisation du morceau mixte

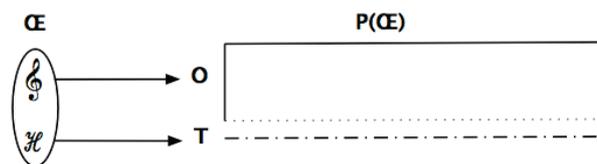
Pour formaliser tout cela, partons d'un morceau mixte.

La dimension spécifique d'œuvre d'art ne sera déterminante que pour le second temps : celui de l'extension ou irradiation. L'idée, en effet, est que tout morceau de musique – même une simple pièce – est susceptible de tirer profit pour lui-même du texte accueilli (chanté le plus souvent) mais que l'œuvre seule est en capacité d'irradier en retour le flux hétérogène qui l'a éclairée-modulée.

Nommons \mathcal{E} ce morceau mixte (dont la dimension d'œuvre nous sera bientôt nécessaire) qui associe un flux musical \mathcal{M} et un flux hétérogène \mathcal{H} .

Soit P sa partition. Comme on l'a vu, P(\mathcal{E}) se disjoint en

- une partition strictement musicale qu'on note O pour « orchestre »,
- une inscription du flux \mathcal{H} qu'on note T comme « texte » puisqu'on prend ici pour modèle d'OMM l'œuvre chantant un texte littéraire préexistant.



6.3. Deux segmentations...

Pour engager une formalisation des interactions entre \mathcal{M} et \mathcal{H} , nous allons partir de leur segmentation respective : nous allons supposer pour ce faire – ce qui constitue une hypothèse tout à fait banale – que chaque flux est doté d'une logique autonome et endogène apte à le périodiser c'est-à-dire à le segmenter hiérarchiquement selon l'ordre chronologique : le flux \mathcal{H} (qu'on suppose ici littéraire par commodité d'exposition) se découpe ainsi en syllabes, mots, groupes de mots, phrases, vers ou paragraphes... qui ossaturent son phrasé envisageable. Tout de même le flux musical \mathcal{M} se découpe en phrases et motifs, cellules et « thèmes », « objets » et périodes, entités musicales significatives (clusters, agrégats, blocs, gestes...).

Inscrivons ce principe de la segmentation en notant $S(\mathcal{H})$ et $S(\mathcal{M})$ nos deux segmentations respectives, la différence de lettres inscrivant l'hétérogénéité des deux logiques de segmentation : $S(\mathcal{H})$ se fait selon une logique langagière, quand $S(\mathcal{M})$ se fait selon une logique musicale.

Précisons : quand nous parlons ici de discours musical (à l'œuvre), nous ne supposons nullement pour autant que la musique serait un langage ; nous ne prenons nullement le discours langagier comme paradigme (moins encore comme transcendantal) de tout discours. Nous prenons ici la notion de discours dans sa plus grande généralité : celle, par exemple, qui préside aux travaux de François Wahl sur

Le discours du tableau. Disons qu'on entend ici par *discours* l'exposition, linéairement ordonnée (dont le temps chronologique constitue alors l'index le plus commode) et hiérarchiquement articulé d'une consistance logique (quel que soit par ailleurs le type de cette logique).

Rappelons en ce point un « résultat » essentiel de *Logiques des mondes* (Alain Badiou) : toute logique langagière (« petite » logique) est subsumable par une logique qui ne l'est pas (« grande » logique), l'inverse n'étant pas vrai si bien que la « petite logique » langagière s'avère n'être qu'un cas particulier d'une « grande logique » bien plus vaste et non restreinte aux seuls effets spécifiques de langue.

Dit autrement : si *logique* nomme le régime de consistance des relations propres à l'apparaître, parler de flux musical comme discours rehausse la consistance logique de ses relations chronologiques sans aucunement nécessiter de prendre le langage comme paradigme d'un tel type de relations.

6.4. ... synchrones

Le point est que deux segmentations procédant de deux logiques différentes aboutissent ici à des segmentations du même temps chronologique : celui, précisément, que partagent nos deux flux ♩ et ♩ . C'est cette homo-chronologie des deux segmentations qui va autoriser leurs interférences.

7. PREMIÈRE ACTION : UNE MODULATION

Notre hypothèse de travail va être la suivante : la première interaction va circuler de $S(\mathcal{K})$ vers $\int \text{♩}$ et c'est cette interaction qu'on va nommer *modulation* en allant chercher notre inspiration du côté de Richard Wagner.

7.1. Fécondation de la musique

Le premier temps de la dialectique entre ♩ et \mathcal{K} est donc celui où \mathcal{K} agit sur ♩ .

Pour utiliser le vocabulaire d'*Opéra et Drame* de Richard Wagner, en ce premier temps, la musique est fécondée par le poème (i.e. par ce texte T qu'on prend, dans notre formalisation, comme paradigme de l'inscription, dans la partition musicale, du flux hétérogène \mathcal{K}). La matérialité de cette fécondation va se donner dans la production d'une nouvelle entité musicale : dans le cas où \mathcal{K} est un texte, par la production d'un chant de ce texte, autant dire d'une mélodie.

8. L'EXEMPLE DE LA MÉLODIE INFINIE WAGNÉRIENNE

De ce point de vue, la manière dont la mélodie infinie de Wagner est engendrée est tout à fait éclairante pour la dynamique de notre OMM. On sait que le compositeur

rédigait, très en amont de toute composition musicale¹⁰, ses livrets donc son texte (ou flux verbal), puis qu'il composait une *particella* dans laquelle le flux orchestral était esquissé en une réduction pour piano, la mélodie du chant apparaissant alors « entre » l'orchestre et le texte en sorte que cette mélodie (notons-la M) puisse être vue comme engendrée par l'interaction du flux orchestral O et du texte T :

$$O \otimes T \Rightarrow M$$

Une fois ceci posé, Wagner composait la partition intégrale, avec ses détails instrumentaux et dramaturgiques.

8.1. Une description

Comme l'on sait, la mélodie que cette dialectique poème/orchestre engendrait, Richard Wagner l'a appelée *mélodie infinie* [unendliche Melodie] et la meilleure présentation qu'il en ait faite est celle qu'on trouve dans « *Musique de l'avenir* » [5] :

« La grandeur du poète se mesure surtout par ce qu'il s'abstient de dire afin de nous laisser dire à nous-mêmes, en silence, ce qui est inexprimable ; mais c'est le musicien qui fait entendre clairement ce qui n'est pas dit, et la forme infaillible de son silence retentissant est la *mélodie infinie* [unendliche Melodie]. Évidemment, le symphoniste ne pourrait former cette mélodie, s'il n'avait son organe propre : cet organe est l'orchestre. Mais, pour cela, il doit en faire un tout autre emploi que le compositeur d'opéra italien, entre les mains duquel l'orchestre n'était qu'une monstrueuse guitare pour accompagner les airs. [...] L'orchestre du symphoniste moderne, au contraire, est mêlé aux motifs de l'action par une participation intime ; car si, d'une part, comme corps d'harmonie, il rend seule possible l'expression précise de la mélodie, d'autre part, il entretient le cours interrompu de la mélodie elle-même. [...] La grande mélodie, telle que je la conçois, qui embrasse l'œuvre dramatique tout entière, [...] doit produire dans l'âme une disposition pareille à celle qu'une belle forêt produit, au soleil couchant, sur le promeneur qui vient de s'échapper aux bruits de la ville. Cette impression consiste [...] dans la perception d'un silence de plus en plus éloquent. [...] Celui qui se promène dans la forêt [...] distingue avec une netteté croissante les voix d'une variété infinie [unendliche], qui s'éveillent pour lui dans la forêt ; elles vont se diversifiant sans cesse ; il en entend qu'il croit n'avoir jamais entendue ; avec leur nombre s'accroît aussi d'une façon étrange leur intensité ; les sons deviennent toujours plus retentissants ; à mesure qu'il entend un plus grand nombre de voix distinctes, de modes divers, il reconnaît pourtant, dans ces sons qui s'éclaircissent, s'enflent et le dominant, la grande, l'unique mélodie de la forêt. »¹¹

¹⁰ Le cas extrême de cette précérence du poème est celui du *Crépuscule des dieux* dont le livret fut achevé en novembre 1848, la partition de l'opéra ne l'étant qu'en novembre 1874, soit 27 ans plus tard !

La composition de *Parsifal* se situe par contre dans la moyenne : la fin de la partition (janvier 1882) intervient un peu plus de quatre ans après la fin de la rédaction du poème (avril 1877).

¹¹ [6] VI.239-241 (en all. : VII.130-131)

Il faudrait prendre le temps de commenter en détail cette présentation passionnante. Retenons simplement ici l'idée que cette mélodie est un produit singulier de l'organe-orchestre, engendré par le bruissement propre à la variété infinie des voix composant la forêt orchestrale. Ce produit compose la grande et unique mélodie embrassant l'œuvre toute entière, sorte de ligne sismique enregistrant de part en part l'intégralité des dynamiques tant littéraires qu'orchestrales.

8.2. Une infinité... en tout point

Notons au passage que l'infinité de la « mélodie infinie » désigne moins une infinité horizontale, son absence d'extrémité (ses extrémités sont simplement celles de l'œuvre) qu'une infinité intrinsèque et en tout point : la mélodie mixe la variété infinie des voix qu'elle capte en sorte de se trouver « infinie » à tout moment, tout de même que la grande œuvre musicale est grande non parce qu'elle durerait grandement mais bien parce qu'à tout moment elle met en œuvre une grande musique.

8.3. Une modulation de fréquence

Captons l'opération de synthèse dont la mélodie infinie est spécifiquement porteuse comme une opération de modulation de fréquence c'est-à-dire comme modulation de la fréquence d'une *porteuse* (ici l'orchestre) par la fréquence d'une *modulante* (ici le texte) engendrant une fréquence *modulée* (ici la mélodie).

On va ainsi voir combien « l'informe »¹² de la mélodie infinie (par comparaison à la « forme » traditionnelle des airs d'opéra) relève bien d'une nouvelle logique – « modulée » - de la forme mélodique : la forme traditionnelle de l'air renvoyait à la mise en arabesque d'une structure tonale et métrique préétablie, quand la nouvelle forme de la mélodie infinie renvoie à la sensibilité sismographique d'une oblique.

8.4. Un exemple

Donnons-en un exemple, extrait du très long monologue (près d'un quart d'heure) de Gurnemanz au début de *Parsifal*. Voici successivement (de haut en bas) la modulante (le texte), la porteuse (la *particella* orchestrale) et la modulée (la mélodie du texte chanté) :

La mélodie résultante porte la trace synchronisée tant du périple tonal et motivique de l'orchestre que des signifiants et signifiés charriés par le poème.

D'un côté, le texte impose sa propre logique littéraire : significations et phrasé propre. De l'autre la masse orchestrale déploie ses courants et contre-courants tant harmoniques que leitmotiviques (sans compter les dynamiques instrumentales, déjà implicites dans la tête du compositeur). La mélodie germe ainsi d'un produit de convolution des deux, telle une ligne de flottaison dont les infimes oscillations reflètent aussi bien les grandes dynamiques souterraines que les fluctuations, plus en surface, du souffle poétique.

8.5. Petite formalisation

Formalisons cela ainsi :

$$O \otimes T \Rightarrow M$$

La modulation se matérialise en l'ajout de M (mélodie) à O et T. M, ce faisant, va se trouver en position de frontière entre \mathcal{G} et \mathcal{H} ce qui, dans le cas de la mélodie, correspond au fait que c'est bien la même voix qui chante la musique et qui parle le texte.

Ce point – c'est la même voix qui chante et qui parle – est aussi capital que mystérieux : il n'y a pas deux voix – l'une qui parle, l'autre qui chante – qui seraient conjointes mais bien une seule et même voix qui à la fois parle et chante. Ainsi la voix mixe étroitement les sons du chant et les sons de la parole. Le mystère et la puissance propre de cette unique voix est que cette mélodie (« l'unique mélodie de la forêt ») se donne comme infime ligne où coexistent deux versants incompatibles (discursivités musicale et textuelle), comme instabilité d'une ligne de crête plutôt que comme stabilité du fond d'un vallon, soit la voix comme constitution d'un lieu mobile, éphémère et infiniment fragile où convergent et rayonnent les forces de toutes natures au travail dans l'OMM.

8.6. Une clôture

M, frontière commune interne à P entre O et T, vient en un certain sens clôturer O (au sens topologique du terme) puisque ce nouvel objet M est bien musical et appartient bien à la musique. Ainsi le nouvel ensemble $\{O+M\}$ devient-il un fermé (un espace topologique

¹² On sait que tel était le grand reproche adressé à la mélodie wagnérienne par les partisans de l'opéra traditionnel.

auquel sa frontière appartient) et la modulation peut-elle donc être vue comme opération de clôture musicale.

Comme on va le voir, cette clôture dote la musique d'une puissance supplémentaire, apte à aimer en retour le poème qui l'a nourrie, ce qui incidemment pointe qu'en certaines circonstances, l'opération de fermeture d'un espace de pensée est bien plus puissante, productive et dynamisante que celle d'ouverture - plus exactement, la capacité endogène de fermer un espace de travail selon une logique autonome inaugure ici une capacité de rayonner et résonner quand une ouverture à tous vents tendrait plutôt à diluer toute consistance endogène...

8.7. L'engendrement d'une *modulée*

Revenons à notre OMM. L'idée que nous allons retenir de l'exemple wagnérien d'une mélodie infinie est la suivante : la modulation de la segmentation musicale de O par la segmentation littéraire de T produit une nouvelle entité musicale M (M pour *modulée* mais aussi, dans notre exemple, pour *mélodie*) venant fermer O.



L'interaction entre \mathcal{H} et \mathcal{H} s'oriente vers le flux \mathcal{H} pour déposer - ici composer - une nouvelle partie musicale M complétant et fermant O.

Comment opère, dans le cas général, une telle « modulation » ? Par le biais des segmentations respectives dont nous avons précédemment parlé : la segmentation (d'ordre littéraire) $S(\mathcal{H})$ de T vient moduler la segmentation (d'ordre musical) $\int(\mathcal{H})$ de O ce qui est possible puisque \mathcal{H} et \mathcal{H} partagent le même temps chronométrique (ce qui n'est pas exactement dire la même temporalité subjective, moins encore la même logique temporelle) et que T et O partagent un même espace d'inscription dans la partition $P(\mathcal{E})$ ¹³.

Dans le cas général - œuvre musicale s'alliant à un flux d'images, à un texte récité, à un flot chorégraphique -, la partie modulée M n'est pas objectivée c'est-à-dire présentée comme objet musical spécifique (comme mélodie par exemple). Il s'agira plutôt de la suggestion d'une ligne de crête, d'une sorte d'enveloppe intérieure ou d'oblique transversale faite de pointillés plutôt que tracée selon une inscription littérale propre. Cette ligne modulée ressemble, par bien des points, à une ligne d'écoute, soit le filigrane d'une texture, ou le fil rouge d'un discours.

Pour l'auditeur, cette modulée se donne subjectivement comme de nouveaux rapprochements entre parties musicales induits par l'entrelacement de la musique et du flux hétérogène, les nouvelles connexions musicales ainsi suggérées par \mathcal{H} tendant à dessiner un

nouveau cours possible du discours musical, une nouvelle manière de le « comprendre », de l'épouser, de s'y incorporer d'oreille.

8.8. Une nouvelle composante musicale donc

Au total, le discours musical se trouve ainsi enrichi d'une nouvelle composante, de nouvelles relations endogènes, d'un nouveau fil conducteur en position cette fois frontière puisqu'à la fois d'ordre musical et en capacité *immédiate* d'agir à l'extérieur, comme si la musique s'était ainsi dotée¹⁴ d'une peau apte à réguler ses interactions avec l'hétérogène et surtout - cela va constituer le second temps du travail propre de l'OMM - apte à servir d'appui pour une rétroaction de la musique sur le flux hétérogène \mathcal{H} : en un sens cette nouvelle peau musicale autorise que la musique puisse désormais « toucher » en retour le flux \mathcal{H} et, par là, l'affecter de frottements, d'appuis ou de caresses.

Formalisons plus avant notre premier temps de modulation. On écrira :

$$S_{\mathcal{H}}(T) \otimes \int_{\mathcal{H}}(O) \Rightarrow M$$

D'où un flux musical \mathcal{H} augmenté selon cette complémentation de la partition $P(\mathcal{E})$:

$$O \Rightarrow \{O+M\}$$

Voyons maintenant comment opère notre second temps, celui que nous proposons d'appeler *extension* et pour lequel nous allons nous appuyer cette fois sur une problématique d'ordre mathématique.

9. SECONDE ACTION : UNE EXTENSION

Notre problème est le suivant : comment penser la manière dont la musique peut agir (ici rétroagir), à sa façon musicale propre, sur un flux hétérogène \mathcal{H} ?

Comprenons-nous bien : il ne s'agit pas ici à proprement parler de « musicaliser » ce flux \mathcal{H} , de le rendre musical, de l'incorporer comme tel à la musique comme la musique incorpore les sonorités de la parole préférant le texte T. Il ne s'agit pas d'homogénéiser en musique les bruits inhérents à l'exposition sensible de \mathcal{H} (comme pourrait le faire par exemple un musicien composant une OMM avec film sonore obligé, incorporant alors à son discours musical les bruits de la ville qui serait présentée dans le film). Il s'agit de la capacité de la musique d'agir sur une tout autre logique discursive que musicale, en préservant à la fois l'autonomie musicale (la musique, ici pas plus qu'ailleurs, ne capitule - n'envisage même de capituler) et l'indépendance propre du flux \mathcal{H} accueilli dans l'OMM.

Reprenons pour cela notre exemple où \mathcal{H} est un texte chanté et M la mélodie de ce chant : il s'agit d'examiner comment la musique (matérialisée désormais par le nouveau ensemble O+M) va pouvoir faire résonner ce

¹³ Rappelons que nous supposons qu'il existe ici une inscription minimale du flux \mathcal{H} dans la partition P sous forme d'une notation T, précisément synchronisée.

¹⁴ Précisément parce qu'elle n'a pas renoncé, dans cette interaction exogène, à son autonomie logique de développement...

qui de T échappe à M c'est-à-dire sa signification langagière propre. En effet, ce qui reste ici essentiellement hétérogène à la logique musicale de M, c'est bien ce que dit le texte chanté (rappelons que nous avons installé, au seuil de cet examen, comme principe le fait que l'OMM préserve la logique propre de \mathcal{H} qu'elle accueille, ce qui veut ici dire qu'elle préserve la compréhension auditive du langage que le texte mobilise).

Comment alors la musique – qui, au sens propre, ne dit et ne signifie rien – peut-elle agir (ici rétroagir) sur la part proprement langagière (et non plus seulement acoustique) de \mathcal{H} donc de T ? Telle est notre seconde question.

9.1. Un type nouveau d'interaction et d'enrichissement

On se doute que ceci va passer à nouveau par une résonance entre segmentations logiquement hétérogènes mais chronologiquement homogènes. Le point essentiel auquel nous allons aboutir sera alors le suivant : cette *raisonance* de la musique vers le texte (ou mise en *raisonance* du texte par la musique) va certes produire un enrichissement du flux hétérogène de départ, donc ici du texte, une « compréhension » renouvelée de ce qu'il dit et porte comme signification, un élargissement de son « sens » littéraire, cinématographique, chorégraphique ou pictural, mais surtout – surtout pour nous musiciens – un élargissement proprement musical de l'OMM elle-même. Très précisément – et c'est ce que l'expression « aura poétique de l'OMM » va venir nommer – la capacité de la musique d'enrichir le flux \mathcal{H} (et donc le texte T) d'un nouveau « sens » (de cette logique non musicale qu'on nomme ici génériquement « poétique ») va s'avérer constituer simultanément un nouvel objet possible dans l'espace (inaccessible à la musique) des significations hétérogènes (langagières dans notre exemple) et une nouvelle capacité (elle proprement musicale) de l'OMM. Soit le point délicat suivant : la musique peut agir de manière musicalement contrôlée (par des opérations de part en part musicales) sur autre chose qu'elle (non seulement sur la forme mais sur le sens même – non musical – de \mathcal{H}) sans pour autant s'incorporer cette chose sur laquelle elle agit, sans pour autant rendre intégralement musical ce flux hétérogène, sans pour autant « musicaliser » le discours poétique qu'elle accueille et qui déborde essentiellement le monde-*Musique*.

9.2. Une question

Comment penser, en musicien, une telle composition musicale d'une « chose » extramusicale, un tel contrôle par la musique de l'engendrement de quelque chose qui essentiellement lui échappe ? Comment est-il possible de l'intérieur d'un monde d'agir sur la logique propre – hétérogène – d'un visiteur qu'on accueille sans aucunement viser à l'« intégrer », d'influer sur la

pensée propre de l'ambassadeur temporaire d'un autre monde (ici celui de la poésie) qu'on invite, le temps d'une OMM, à participer au concert propre au monde-*Musique* ?

9.3. Une méthode

C'est en ce point que nous allons nous tourner vers la mathématique, en l'occurrence la mathématique des extensions, en un geste dont nous assumons la nature essentiellement métaphorique (logique du « comme ») et/ou fictionnelle (logique du « comme si »).

9.4. Deux appuis du côté mathématique

Nous allons ainsi prendre appui sur deux problématiques mathématiques (au demeurant mathématiquement corrélées) des extensions :

- les extensions algébriques de corps ;
- les extensions génériques d'ensembles.

10. EXTENSION ALGÈBRIQUE DE CORPS

10.1. Un exemple

Donnons un exemple très simple du premier type d'extension.

L'ensemble \mathbb{Q} des nombres dits rationnels q (ceux qui se présentent comme fractions de deux nombres entiers $q=n_1/n_2$ avec $n_1, n_2 \in \mathbb{Z}$, ensemble des nombres relatifs) est clos sur lui-même et forme une sorte de petit monde numérique dont on ne sort pas par les opérations usuelles d'addition et multiplication, petit monde que la mathématique appelle « corps » (comme il y a, mathématiquement, des « groupes » et des « anneaux »).

On sait, depuis les Grecs, que $\sqrt{2}$ et une infinité d'autres nombres de ce type ne sont pas rationnels car il n'existe pas deux nombres entiers n_1 et n_2 tels que $\sqrt{2}=n_1/n_2$. Pour cette raison, ces nouveaux nombres¹⁵ ont été appelés « irrationnels ». L'ensemble des nombres rationnels et irrationnels constitue l'ensemble des nombres dits « algébriques » car chacun d'eux est solution d'une équation dite *algébrique*¹⁶ c'est-à-dire racine d'un polynôme : $\sqrt{2}$ par exemple est solution de l'équation algébrique $(x^2-2=0)$ et racine du polynôme $\{x^2-2\}$.

Si l'on appelle A ce nouvel ensemble des nombres algébriques, A est clos sur lui-même pour les mêmes opérations d'addition et de multiplication que \mathbb{Q} et forme ainsi un nouveau corps englobant le corps des seuls nombres rationnels. On dit mathématiquement que A constitue une extension de \mathbb{Q} (plus précisément que le corps A constitue une extension algébrique du corps \mathbb{Q} car \mathbb{Q} est inclus dans A ($\mathbb{Q} \subset A$) et l'engendre (par solution des équations algébriques qu'il institue via les

¹⁵ Rappelons : pour les Grecs, $\sqrt{2}$ n'était pas vraiment un nombre car il ne nombrait nulle grandeur.

¹⁶ Par opposition, par exemple, à équation trigonométrique ou exponentielle/logarithmique...

polynômes qu'il génère). On dit que \mathbb{Q} clôture \mathbb{Q} pour les équations algébriques.

10.2. Un contrôle interne de l'extérieur engendré

Le point qui va nous intéresser en cette construction mathématique est le suivant : il est possible de l'intérieur du corps \mathbb{Q} (de l'intérieur donc du monde des nombres rationnels) de contrôler certaines opérations sur les nombres irrationnels c'est-à-dire sur des nombres extérieurs à \mathbb{Q} et que \mathbb{Q} ignore (ne peut connaître). Prenons un exemple très simple.

\mathbb{Q} connaît l'équation algébrique $x^2-2=0$ (puisque cette équation s'écrit dans la langue de \mathbb{Q} : avec des nombres rationnels et des opérations internes à \mathbb{Q}) mais il ne connaît pas de solution numérique à cette équation (puisque ces solutions sont ici irrationnelles donc essentiellement exogènes à \mathbb{Q}). On sait de l'intérieur de \mathbb{Q} que cette équation a normalement deux solutions $\pm\sqrt{2}$; on peut même nommer, de l'intérieur de \mathbb{Q} , ces deux solutions au moyen de l'écriture purement formelle $\pm\sqrt{2}$ mais on ne saura, de l'intérieur de \mathbb{Q} , de quoi ce symbole $\sqrt{2}$ est-il exactement le nom : un fantôme numérique, une chimère quantitative, le spectre d'une grandeur inaccessible ?

Présenté autrement : si l'on pose qu'un polynôme est le signifiant dont ses racines sont les signifiés, on dira que le polynôme $\{x^2-2\}$ signifie les racines $\pm\sqrt{2}$ sans qu'on puisse, de l'intérieur de \mathbb{Q} , savoir le référent exact de ce $\pm\sqrt{2}$ dans l'espace des grandeurs numériques.

Pour les mêmes raisons, \mathbb{Q} connaît le polynôme (x^2-8) et l'équation algébrique $\{x^2-8=0\}$. Il peut tout de même nommer $\pm\sqrt{8}$ ses deux solutions sans plus savoir ce que nomme l'écriture $\sqrt{8}$.

Le point qui pour nous devient intéressant est alors le suivant : si l'on note $P_1=x^2-2$ et $P_2=x^2-8$, on peut remarquer, de l'intérieur de \mathbb{Q} , que $P_2=P_1-6$ puisque cette opération est parfaitement « rationnelle » au sens de \mathbb{Q} c'est-à-dire au sens où elle se fait par addition entre nombres rationnels.

On peut également remarquer, de l'intérieur de \mathbb{Q} , que les racines respectives r_1 et r_2 de P_1 et P_2 entretiennent la relation suivante $r_2=2r_1$ puisqu'on peut écrire, de l'intérieur de \mathbb{Q} : $r_1^2=2$, $r_2^2=8$, donc $r_2^2=4r_1^2$, donc $r_2=\pm 2r_1$.

Certes on ne sait pas dans \mathbb{Q} ce que nomment ici les symboles respectifs r_1 et r_2 mais on peut cependant calculer dans \mathbb{Q} sur les référents numériques inconnus de ces deux noms et poser donc :

$$P_2=P_1-6 \Rightarrow r_2=2r_1 \quad ^{17}$$

Bien sûr, pour qui se meut dans le monde \mathbb{A} des nombres algébriques et, plus généralement dans le monde \mathbb{R} des nombres dits « réels », ceci est parfaitement clair car les quantités que nomment ces deux nombres r_1 et r_2 sont clairement établies : $\sqrt{2}=1,414\dots$ et $\sqrt{8}=2\sqrt{2}=2,828\dots$

¹⁷ Plus généralement, on peut tout de même établir de l'intérieur de \mathbb{Q} la loi suivante :

$$P_n=P_1-2(n^2-1) \Rightarrow r_n=nr_1$$

10.3. Telle une angéologie...

Pour qui habite \mathbb{Q} , même si r_1 et r_2 désignent des chimères numériques, la relation entre ces deux chimères reste cependant accessible : il y a sens à dire, dans \mathbb{Q} , que retrancher 6 au polynôme P_1 ($P_2=P_1-6$) engendre un nouveau nom pour des chimères deux fois plus grandes ($r_2=2r_1$).

On va ainsi pouvoir contrôler, de l'intérieur de \mathbb{Q} , par des opérations internes et compréhensibles dans \mathbb{Q} (addition, soustraction, multiplication et division) son extension algébrique, tout de même somme toute, que l'humanité religieuse s'est doté d'une extension angélique dont elle contrôle nominalement la hiérarchie interne (des Séraphins aux simples Anges en passant par les Chérubins ou les Archanges...¹⁸) sans pour autant savoir grand-chose de ce à quoi réfèrent réellement ces noms : disons qu'il n'est pas nécessaire de croire à l'existence des anges pour reconnaître l'exercice d'une certaine forme de rationalité interne aux théories angéologiques. Autrement dit : même si on ne sait pas trop ce qu'elles disent, elles ne disent pas pour autant n'importe quoi et leur discours répond à une logique interne qui relève d'une certaine forme de rationalité.

De même, dire dans \mathbb{Q} que $r_2=2r_1$ car $P_2=P_1-6$ n'est pas dire n'importe quoi lors même qu'on ne sait trop de l'intérieur de \mathbb{Q} ce que r_1 et r_2 nomment ici exactement.

11. UN PREMIER « MODÈLE THÉORIQUE »

Nous tenons ici un premier « modèle théorique » – une première maquette ou modèle réduit¹⁹ – pour notre extension poétique de l'OMM : tout de même que le corps \mathbb{Q} génère l'extension algébrique \mathbb{A} dont il ne connaît pas les nouveaux termes (« irrationnels ») mais qu'il peut cependant doter de noms « rationnels » au moyen desquels il peut alors établir entre ces noms des relations aptes à calculer sur leurs références inconnues en sorte, finalement, d'exercer ainsi un certain contrôle immanent sur une réalité extérieure inaccessible, tout de même l'OMM génère, de l'intérieur du monde-*Musique*, une « extension poétique » dont elle ne connaît pas les termes exacts (il y s'agit de significations langagières qui lui échappent par essence) mais qu'elle peut cependant nommer musicalement (au moyen des opérations musicales de périodisation qui engendrent ces chimères poétiques) en sorte de pouvoir contrôler musicalement certaines des relations que ces chimères poétiques ou langagières entretiennent entre elles dans leur monde propre.

¹⁸ L'angéologie chrétienne traditionnelle distingue ainsi neuf classes ; de la plus élevée à la plus proche des hommes : les *Séraphins*, les *Chérubins*, les *Trônes*, les *Dominations*, les *Puissances*, les *Vertus*, les *Principautés*, les *Archanges* pour finir par les simple *Anges*...

¹⁹ Le mot *modèle* est pris ici à rebours de son sens logico-mathématique canonique (voir la théorie des modèles dans la logique mathématique) ; il est pris, par commodité d'exposition, dans son sens empirique, aujourd'hui hégémonique dans la langue commune...

11.1. Une fiction

Notre fiction mathématico-musicale se bâtit donc sur l'équivalence suivante :

MATHEMATIQUE	MUSIQUE
Le corps \mathbb{Q} des nombres « rationnels »	monde- <i>Musique</i>
Les nombres rationnels q_n	Les objets musicaux (ou morceaux de musique) de ce monde
Des polynômes donnés P_i	Des opérations musicales propres à une OMM donnée
Des noms « rationnels » pour les nouvelles entités « irrationnelles », racines r_i des polynômes P_i précédents	De nouvelles relations musicales (entre objets musicaux) dont la signification musicale est indéceise
Un calcul « rationnel » au moyen de ces noms « rationnels » sur les nouvelles quantités numériques inaccessibles	Une paralogie musicale régionale qu'on dira <i>extime</i> (musicalement intime par ses opérations, exogène par ses résultats)
L'extension de \mathbb{Q} en un nouveau sous-corps $\mathbb{Q}[\sqrt{2}]$ de type $\{\mathbb{Q}+\sqrt{2}\mathbb{Q}\}$	Une aura poétique de l'OMM
Le nouveau corps A qui englobe tous les sous-corps de ce type	Soit la question suivante : y aurait-il une extension poétique du monde- <i>Musique</i> qui récapitulerait les extensions poétiques de toutes les OMM ?

11.2. Un sous-corps

Nous introduisons, au passage, l'idée d'un sous-corps intermédiaire entre \mathbb{Q} et A que nous notons $\mathbb{Q}[\sqrt{2}] = \{\mathbb{Q}+\sqrt{2}\mathbb{Q}\}$ et qui désigne l'extension de \mathbb{Q} obtenue quand on lui ajoute non seulement $\sqrt{2}$ mais l'ensemble des nombres algébriques irrationnels de la forme $a+b\sqrt{2}$ (avec a et b rationnels) : par exemple les nombres $2\sqrt{2}$ ou $10/3-\sqrt{2}/5$ appartiennent à ce nouvel ensemble dont on montre très facilement qu'il constitue un corps car il est intérieurement fermé pour l'addition et la multiplication²⁰. Dans notre fiction, ce nouveau corps (sous-corps de A) va désigner l'extension poétique spécifique dont une OMM donnée est musicalement capable (son « aura poétique » donc) laquelle n'est qu'une des infinités d'extensions auratiques musicalement concevables pour les morceaux de musique.

Tout de même que de l'intérieur de \mathbb{Q} , on ne peut « imaginer » – en un imaginaire cependant intérieurement contrôlé et discipliné – que des extensions particulières, des sous-corps limités du type $\mathbb{Q}[\sqrt{2}]$, tout de même de l'intérieur du monde-*Musique*, on ne peut contrôler que des auras poétiques

²⁰ Il est très facile de vérifier que si q_1 et q_2 appartiennent à $\mathbb{Q}[\sqrt{2}]$ (c'est-à-dire si $q_1=a_1+b_1\sqrt{2}$ et $q_2=a_2+b_2\sqrt{2}$), alors q_1+q_2 comme q_1q_2 appartiennent aussi à $\mathbb{Q}[\sqrt{2}]$ c'est-à-dire peuvent s'écrire sous la forme $n+p\sqrt{2}$.

particulières, propres à chaque fois à telle OMM donnée ; on ne saurait pas plus imaginer une extension générale du monde-*Musique* qu'on ne saurait, de l'intérieur de \mathbb{Q} , prétendre contrôler l'engendrement intégral de A c'est-à-dire des nombres irrationnels à partir des polynômes rationnels (il est déjà clair que, de l'intérieur de \mathbb{Q} , il n'y a aucun accès à \mathbb{Q} comme totalité, c'est-à-dire comme ensemble et comme corps : de l'intérieur de \mathbb{Q} on ne peut que constater, comme le faisaient les Grecs, que certaines opérations « marchent » c'est-à-dire conduisent bien à des grandeurs « rationnelles » quand d'autres aboutissent à des chimères numériques). Certes, de l'extérieur de \mathbb{Q} et du point d'une mathématique non grecque, on peut plonger \mathbb{Q} dans A (et plus généralement dans \mathbb{R} , corps des nombres réels) en s'autorisant ainsi d'un concept plus général de nombre (réel) qui englobe hiérarchiquement les nombres entiers (\mathbb{N}), relatifs (\mathbb{Z}), rationnels (\mathbb{Q}), algébriques (A), sans oublier les nombres transcendants qui, au demeurant, s'avèrent incommensurablement les plus nombreux²¹ ! Mais une telle compréhension est radicalement exogène à \mathbb{Q} et ne saurait donc nous orienter sur la manière dont l'OMM comprend et contrôle de manière musicalement endogène son aura poétique propre.

11.3. Pas d'aura globale du monde-*Musique* et pas de sens général...

Répondons ainsi à la question qui conclut notre petit tableau précédent : dans notre fiction, l'équivalent musical de l'extension algébrique A du corps \mathbb{Q} devrait désigner une sorte de nouveau monde, englobant notre monde-*Musique*, monde où la notion mathématique généralisée de nombre désignerait quelque chose comme une catégorie généralisée de *sens*, l'idée étant alors que l'OMM viendrait plonger le « sens » musical restreint dans un « sens » poético-musical élargi (l'étape suivante, équivalant au passage mathématique de A à \mathbb{R} , voir aux nombres surréels, correspondrait alors à une théorie généralisée du « sens » dans laquelle le sens proprement musical apparaîtrait rétroactivement comme une simple région d'un sens global).

Disons-le clairement : nous nous interdisons cette voie et posons, avec la même assurance que Parménide pouvait poser, dans son registre propre « *Le néant n'est pas ; voilà ce que je t'enjoins de considérer.* »²² ceci : le sens de la musique, s'il existe, n'est pas musical (il procède d'un regard extérieur sur la musique), et le musicien ne doit pas *comme musicien* le prendre en considération pour s'orienter musicalement (il peut, par contre, toujours le faire comme individu).

²¹ Où l'on vérifie, au passage, que l'invisibilité dans un monde est la propriété caractéristique non de l'exception rare mais bien de la multitude peuplant en vérité ce monde et débordant de toute part la représentation endogène de lui-même dont ce monde est ordinairement capable.

²² « *Nécessaire est ceci : dire et penser de l'étant l'être ; il est en effet être, le néant au contraire n'est pas : voilà ce que je t'enjoins de considérer.* » [7] Fragment VI (p. 81)

12. UNE DIFFICULTÉ

Comment penser plus avant la manière propre dont l'OMM va musicalement contrôler cette aura poétique dont nous avons pris un premier « modèle » théorique dans la théorie mathématique des extensions algébriques de corps ?

12.1. Le danger d'une homogénéisation

Rappelons : la dimension poétique de l'aura d'une OMM est à la musique comme le sous-corps $\mathbb{Q}[\sqrt{2}]$ de \mathbb{A} l'est à \mathbb{Q} .

Notre modèle théorique bute alors sur la difficulté inhérente suivante : si l'OMM est augmentée de son aura comme \mathbb{Q} l'est par $\sqrt{2}$ dans $\mathbb{Q}[\sqrt{2}]$ ²³, pour autant il n'y a pas d'équivalent musical à la nouvelle notion mathématique de nombre (nombre « algébrique ») venant subsumer nombres « rationnels » et nouveaux nombres « irrationnels ». Autrement dit, là où la mathématique homogénéise corps de base et extension au moyen d'une notion généralisée de nombre autorisant une construction hiérarchique de corps et subsumant ainsi la différence locale entre rationnels et $\sqrt{2}$, la musique ne saurait le faire sauf à s'annexer, en une conception alors inconsistante d'une musique « généralisée »²⁴, la nouvelle signification poétique dont l'aura de l'OMM est tramée.

Il nous faut donc affiner notre modèle théorique pour serrer de plus près la manière dont l'OMM travaille, de manière musicalement endogène, l'aura poétique qu'elle génère.

C'est ce que nous allons maintenant faire en nous tournant vers une seconde théorie mathématique des extensions : cette théorie des extensions génériques d'ensembles qui s'est précisément déployée en analogie avec la théorie des extensions algébriques de corps.

Paul J. Cohen, le mathématicien à l'origine de cette nouvelle théorie, a en effet explicitement bâtie son entreprise en « analogie à la construction de l'extension d'un corps K obtenue par adjonction de la racine α d'une équation irréductible $f(x)=0$ »²⁵.

²³ Attention : rappelons-nous que $\mathbb{Q}[\sqrt{2}] \neq \mathbb{Q} + \sqrt{2}$. $\mathbb{Q}[\sqrt{2}] = \{a + b\sqrt{2}\}$. Il comporte donc tous les nouveaux nombres faisant intervenir $\sqrt{2}$ par addition, soustraction, multiplication et division (et il n'est donc pas augmenté du seul nombre $\sqrt{2}$). $\sqrt{2}$ n'intervient donc pas ici comme un objet-nombre isolé qu'on ajouterait-accolerait aux anciens objets-nombres mais comme une sorte de navette avec laquelle tisser tout un réseau de nouveaux nombres à partir des anciens...

²⁴ Celle-là même dont il est question, métaphoriquement mais aucunement réellement, quand on parle par exemple de « musique de la langue », de « musique de l'architecture » ou de « musique du geste chorégraphique »...

²⁵ "The situation is analogous to the construction of the extension of a field k formed by adjoining the root α of an irreducible equation $f(x)=0$. The elements of the extension field are all of the form $p(\alpha)$ where p is a polynomial and α is taken as a formal symbol, but we identify $p(\alpha)$ and $q(\alpha)$ if $p(x)-q(x)$ is divisible by $f(x)$." [8] p. 113

13. EXTENSION GÉNÉRIQUE D'ENSEMBLE

« *What made it so exciting to me was how ideas which at first seemed merely philosophical could actually be made into precise mathematics.* » P. Cohen

La théorie mathématique des extensions génériques, relativement récente²⁶, est d'une difficulté intrinsèque tout autre que la précédente.

On en trouvera un exposé synthétique dans les conférences faites en 1965 à Harvard [8] et une présentation mathématique plus didactique dans les textes indiqués en note.²⁷ Par ailleurs, cette théorie mathématique est au principe du concept philosophique de forçage d'une vérité dans *L'être et l'événement* d'Alain Badiou²⁸.

Nous ne pouvons ici que survoler cette théorie mathématique en indiquant ses grandes articulations théoriques susceptibles de nous guider pour nos propres objectifs de musicien.

13.1. Une question mathématique

La question de Cohen est la suivante : construire de l'intérieur d'un ensemble Θ ²⁹ (doté bien sûr de propriétés adéquates)³⁰ un ensemble Γ d'un nouveau type (dit *générique*), ensemble fait de parties de Θ mais n'appartenant pas lui-même à Θ , tel que l'extension $\Theta[\Gamma]$ obtenue par interaction de Θ avec ce nouvel ensemble Γ ³¹ forme alors un nouvel espace doté de propriétés inconnues dans l'espace de départ.

Notre expérimentation (ou *pratique théorique*) va se déployer au fil de la fiction suivante : faisons *comme si* la partition musicale donnée d'une OMM donnée \mathbb{E} tenait lieu de Θ et voyons si la construction mathématique de Γ puis de $\Theta[\Gamma]$ nous éclaire quant à une composition musicale de l'aura poétique \mathbb{A} et de l'extension poético-musicale $\mathbb{E}[\mathbb{A}]$ de cette OMM.

Si le schéma général d'une telle expérimentation musicienne est le suivant :

²⁶ 1963. Elle a valu la médaille Fields à son inventeur, Paul J. Cohen

²⁷ Voici par exemple : Thomas Jech : *What is forcing ?* ; Timothy Y. Chow : *Forcing for dummies, A beginner's guide to forcing* ; Patrick Dehornoy : *La méthode du forcing*. On trouvera aisément tous ces textes sur Internet.

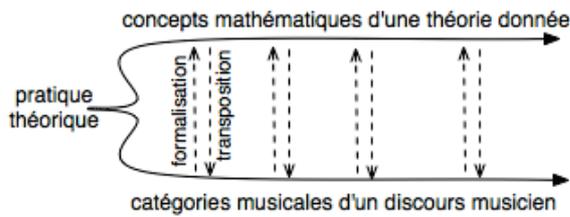
²⁸ On y trouvera une présentation détaillée et minutieusement commentée philosophiquement dans ses chapitres 33, 34 et 36 soit, avec les annexes qui s'y rapportent, près d'une centaine de pages, denses et conceptuellement tendues.

²⁹ On notera ici systématiquement de lettres grecques les notions mathématiques convoquées, réservant ainsi les lettres latines aux catégories musicales. Ce faisant, on réécrit ici Θ ce que Cohen note M de même qu'on va, un peu plus loin, réécrire Π et Γ ce qu'il note respectivement P et G .

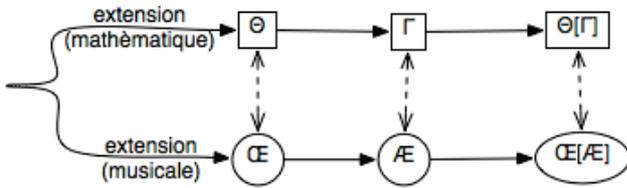
Pour ses besoins philosophiques propres, Badiou renomme Θ en S (« *situation fondamentale quasi-complète* »), Γ en φ et $\Theta[\Gamma]$ en $S[\varphi]$.

³⁰ Θ doit être un modèle transitif dénombrable de la théorie des ensembles (au sens de Zermelo-Frankel).

³¹ L'équivalent de notre précédent $\mathbb{Q}[\sqrt{2}]$



alors celui de notre expérimentation va, en première approche, se présenter ainsi :



Pour indiquer qu'il s'agit bien ici de fiction, nullement d'application, il suffit de relever que l'ensemble mathématique de départ est nécessairement infini quand notre partition musicale est nécessairement finie. La différence est de taille !

Il est vrai que cette même partition sert de support à un faisceau d'interprétations musicales lui-même infini et il est également vrai qu'il nous faut prendre en compte cet espace des interprétations dont une même partition musicale est capable pour rendre compte de notre rétroaction du flux musical \mathcal{L} sur le flux hétérogène \mathcal{H} .

Cependant nous allons être amenés à nous situer de l'intérieur même d'une de ces interprétations ce qui va nous ramener, ce faisant, à la finitude intrinsèque de cette interprétation.³²

Où l'on mesure que la vérité dont une fiction est capable ne tient pas à la véracité de ses hypothèses, à la véridicité de ses procédures, et à la vérifiabilité de ses enchaînements mais à la seule discipline endogène de ses effets dynamiques.

Remarquons d'ores et déjà que l'extension auratique dont il va être ici question ne se cantonnera pas à une aura délimitée (ici à \mathcal{A}), à un « objet » poétique surajouté mais bien à la manière dont cette aura, générée par \mathcal{E} , rétroagit sur toute l'OMM (ce que désigne l'écriture $\mathcal{E}[\mathcal{A}]$) en sorte que l'extension auratique de l'OMM devienne légitimement concevable comme une irradiation globale de l'œuvre, non pas comme une sorte de verrue locale, comme une boursoufflure ou une distension régionale.

Suivons les grandes étapes de la construction par Cohen de son ensemble générique ajouté Γ puis de l'extension de Θ en $\Theta[\Gamma]$ qu'il autorise en examinant, pas par pas, ses effets d'éclairage sur notre problème musical.

14. DES « CONDITIONS »

Cohen va construire le nouvel ensemble Γ à partir d'un type singulier de parties de Θ qu'il appelle des

« conditions » π ³³. Pour cela, il compose dans Θ une partie Π (dont les éléments sont les conditions π) dotée d'une structure spécifique d'ordre : l'idée va être, en effet, que pour composer Γ à partir de conditions π , il faut que celles-ci soient reliées entre elles par une relation d'ordre³⁴, qu'elles soient ordonnables, c'est-à-dire qu'on puisse dire par exemple que l'une affine ou raffine une autre (en étant plus précise), que l'autre est compatible avec une troisième, etc.

Dans notre propre situation musicale, cet ensemble ordonné de conditions va avoir pour équivalent immédiat notre segmentation musicale de l'OMM : un segment s de cette œuvre \mathcal{E} ³⁵ – un motif, un accord, une phrase, un détail, un développement... – sera l'équivalent musical d'une « condition » π pour l'espace mathématique Θ .

Appelons \int une telle segmentation musicale de l'OMM dont la partition $P(\mathcal{E})$ est désormais complète : $P(\mathcal{E})=(P+M)+T$

14.1. Remarque capitale

À ce stade, nous allons travailler sur une segmentation de la partition telle qu'une interprétation musicale donnée la met en œuvre : il ne s'agit plus seulement d'analyser visuellement le texte de la partition et d'y inscrire des continuités et des ruptures par quelque notation (comme nous l'avons fait plus haut dans notre présentation de la mélodie infinie wagnérienne) ; il s'agit désormais de ce type d'analyse de la partition dont une interprétation musicale donnée est spécifiquement capable par sa manière de phraser le texte en y arrimant le jeu d'un corps-à-corps musical. Certes, le résultat de ce phrasé sonore produit des phrases identifiables dans la partition (et donc inscriptibles : les élèves de Chopin ainsi s'attachaient à noter sur leur exemplaire des *Préludes* ou des *Mazurkas* le phrasé de son jeu pianistique) mais l'opération de segmentation dont il est ici question (dans ce second temps de la dialectique entre flux musical et flot hétérogène) diffère de la première en ce qu'elle relève cette fois du jeu instrumental ou vocal, de l'interprétation musicale concrète donc.

La contrepartie décisive de ce point est que cette « seconde » segmentation (rétroaction de la première qui s'est inscrite dans la partition comme ajout de la partie M) est directement et essentiellement audible (quand la première était en premier lieu lisible) : c'est désormais d'oreille que cette segmentation consistant à phraser la partie proprement musicale de l'OMM va être accessible, et « l'aura poétique » qui va résulter de telle ou telle manière de phraser sera elle-même abordée d'oreille : la puissance auratique de l'OMM va ainsi s'avérer relever de l'écoute musicale d'une interprétation donnée.

³³ Cohen les note p .

³⁴ Ordre partiel – on parle ici d'ensemble partiellement ordonné...

³⁵ Rappelons qu'à ce stade, cette OMM est désormais augmentée d'une partie musicale M produite par la modulation $T \otimes P$.

³² Ce qui correspond au fait que le morceau de musique est un ensemble infini d'interprétations finies.

14.2. Une construction mathématique par étapes

Reprenons le fil de notre expérimentation théorique, donc celui de la construction mathématique d'une extension générique $\Theta[\Gamma]$.

L'enjeu est de construire, à partir de Π (inclus dans Θ) et de ses éléments π (« conditions »), un nouvel ensemble Γ qui n'appartienne ni à Π ni à Θ , qui soit doté intérieurement de propriétés spécifiques (on va voir lesquelles) et qui soit tel qu'ajouté et composé à Θ , il engendre une extension $\Theta[\Gamma]$ qui assure les propriétés suivantes.

D'une part

– L'ensemble de départ Θ sera bien inclus dans son extension $\Theta[\Gamma]$.

– Le nouvel espace d'arrivée $\Theta[\Gamma]$ conservera bien la même structure que l'espace de départ Θ .³⁶

D'autre part

– L'ensemble générique ajouté Γ sera bien élément de l'extension $\Theta[\Gamma]$.

– Cet ensemble Γ conservera bien dans la situation étendue $\Theta[\Gamma]$ sa propriété caractéristique d'être générique.

Tout le point pivote donc autour de cette notion de généralité singulièrement attachée à l'ensemble Γ . Détaillons-la quelque peu.

15. UNE « GÉNÉRICITÉ »

« *One should try to adjoin to M an element which enjoys no "specific" property to M i.e. something akin to a variable adjunction to a field. I called such an element a "generic" element. Now the problem is to make precise the notion of a generic element.* » P. Cohen

La généralité de la partie Γ désigne la propriété d'être quelconque en un sens très précis : techniquement dit, Γ intersecte toute partie dense de Π ce qui revient à dire qu'elle n'est pas mathématiquement constructible c'est-à-dire délimitable par une propriété caractérisant tous ses éléments. En quelque sorte, la partie générique Γ diagonalise l'ensemble des propriétés susceptibles de partitionner l'ensemble de départ (ici Π) en deux classes disjointes : grossièrement dit, on y retrouve un peu de tout, c'est-à-dire de tout ce qui est caractérisable dans la langue de la situation, en sorte que si une propriété définissable sépare les éléments de Π en deux parties disjointes, on retrouvera alors dans la partie générique Γ à la fois des éléments relevant de l'une et de l'autre partie complémentaire.

15.1. Un paradoxe

Le point délicat, qui fait toute la difficulté du propos, tient alors au projet *a priori* insensé de construire un tel ensemble quelconque, ce qui de prime abord s'énonce

sous la forme d'un oxymore puisque construire un ensemble, c'est *ipso facto* l'identifier par sa particularité et ce faisant lui faire perdre cet anonymat caractéristique du quelconque.

Voyons comment Cohen va résoudre ce paradoxe.

15.2. Remarque

Remarquons qu'un même type de paradoxe se loge au cœur de notre théorie musicale de l'aura poétique propre à l'OMM : cette aura doit simultanément être *une* aura (spécifiée par une interprétation musicale donnée d'une partition donnée qui compte musicalement pour *un* cette aura) et non pas rester une dimension ou une puissance auratique indéfinie ou fantasmée, mais aussi être aura « poétique » c'est-à-dire rester musicalement quelconque ou générique pour l'OMM, donc musicalement incomptable pour un.

L'enjeu musical d'une interprétation d'une OMM donnée est ainsi homologue à l'enjeu mathématique de Cohen : donner accès pour l'oreille à une œuvre mixte *musicalement* étendue au gré d'une aura *poétique*.

16. DES « NOMS »

« *Trois fois heureux, qui dans le chant glisse un nom ! / Le chant qu'embellit le mot qui nomme / la longue survie parmi les autres.* » Mandelstam

Pour lever ce paradoxe, Cohen va travailler – et c'est là l'idée fondamentale – non directement sur ses éléments constitutifs mais indirectement sur des noms (v) possibles dans Π de ces éléments en sorte qu'on pourra dire – et c'est là sa manière de résoudre le paradoxe relevé précédemment, très exactement de *forcer*³⁷ l'oxymore – que le nom contrôle l'appartenance à Γ des conditions π que ce nom mobilise, tout de même que nos polynômes pouvaient être vus comme noms contrôlant dans \mathbb{Q} les relations dans \mathbb{A} ou \mathbb{R} entre les racines qu'ils nommaient.

16.1. Le nom et sa double face

« *Ce sont les noms qui créent la chose.* » Badiou³⁸

Toute la puissance et la difficulté de l'opération tient ici à la torsion intérieure qu'elle engage ; un *nom*, en effet, a une double face : d'un côté celle du *mot* fait de la matérialité propre du signifiant (ce qui n'est pas tout à fait dire « matérialité signifiante ») : le mot « arbre » est fait de phonèmes et de lettres, lesquels par eux-mêmes ne signifient pas ; d'un autre côté celle de la référence qui signifie une *chose* laquelle relève, dans son ordre propre, d'un tout autre type de matérialité (la chose nommée « arbre » est faite de bois et végétaux, non de lettres et de phonèmes).

³⁶ Techniquement dit, celle d'être un modèle transitif dénombrable de la théorie des ensembles (version ZF).

³⁷ Sa méthode, en effet, s'appelle *forcing*, qu'Alain Badiou traduit judicieusement par *forçage* là où l'habitude des mathématiciens est de conserver le terme anglais.

³⁸ [9] p. 543

Ainsi la nomination polynomiale dans \mathbb{Q} de nombres irrationnels articule la matérialité algébrique propre d'un polynôme dans le corps des rationnels \mathbb{Q} (un polynôme est fait... de polynômes plus petits³⁹ et/ou de lettres et chiffres) quand sa référence signifiée est affaire de nombres algébriques.

Par exemple x^3-x^2-2x+2 s'écrira $(x-1)(x+\sqrt{2})(x-\sqrt{2})$ et nommera ainsi les racines 1 et $\pm\sqrt{2}$.

Cette écriture, dans \mathbb{Q} , sera marquée par le fait que « $\sqrt{2}$ » y restera sans référent numérique et ne sera qu'une simple lettre, commodité d'écriture décomposant le polynôme rationnel x^2-2 selon la même logique qui décompose le polynôme x^2-4 en $(x+2)(x-2)$.

16.2. D'où une torsion singulière

La torsion, au principe des opérations de Cohen, va alors consister en ceci.

Il va d'abord s'agir de travailler la *chose* indirectement en travaillant directement les mots nommant la chose.

Il va ensuite s'agir de constituer la chose à partir de ces mêmes *mots*, de composer la chose en travaillant sur les mots susceptibles de la nommer et de travailler ces mots en sorte de la doter d'un nouveau pouvoir de nomination.

Il va enfin s'agir (et c'est ici que la torsion se matérialise abruptement) que ce nouveau pouvoir de nomination des mots ainsi travaillés tienne au fait que ces *noms* particuliers vont nommer... d'autres mots du même type !

Au total, chaque mot de ce nouveau lexique particulier va être ainsi simultanément... nom et chose : il va être *nom* d'autres mots (du même lexique), et *chose* nommée par d'autres mots (du même lexique).

16.3. Et sa portée musicale

Pour ne pas trop nous perdre dans ce labyrinthe, indiquons ce que cette manière de travailler éclaire quant à notre problème musical d'aura poétique.

L'équivalent musical de la nomination inventée par Cohen va tenir à un type particulier de segmentation auditive du flux musical : l'idée est que l'interprétation musicale du flux écrit $\{P+M\}$ met en œuvre une périodisation singulière (qu'on attachera à la manière interprétative propre de *phraser* le flux sonore) qui se trouve alors intérieurement marquée, dans le cas de l'OMM, par le jeu d'une périodisation non musicale (celle du texte entendu et compris qui opère au lieu même de la musique puisque c'est la même voix qui chante et qui parle).

D'où que le phrasé de l'interprétation imbrique segmentations musicale et non musicale (littéraire, par exemple) *en un même corps* (l'interprétation musicale est bien ce qui ajoute à la partition le jeu d'un corps

musical projetant/rayonnant une matérialité sonore audible).

La torsion propre (à notre rétroaction $\mathcal{L} \rightarrow \mathcal{H}$) va alors tenir au fait que toute segmentation musicale ainsi phrasée par l'interprétation va être saisissable d'oreille comme nomination possible de segments prélevés cette fois dans \mathcal{H} , donc dans T, autant dire comme signifiant de nouveaux rapports cette fois littéraires (ou chorégraphiques...) dans le texte T : pour user d'une métaphore cinématographique, cette segmentation musicale synthétique, phrasée par l'interprète, devient audible comme opération parallèle de *montage* (du texte, ou des images, ou des gestes dansés) : ainsi phraser musicalement la partition enrichie de M se projette extérieurement comme nouvelle opération de montage sur le flux hétérogène \mathcal{H} ; bien sûr ce flux n'est pas ici entièrement démonté puis remonté par la musique - disons que son montage originaire (d'ordre littéraire, chorégraphique ou cinématographique) se trouve désormais *modulé* en retour par le phrasé qu'une interprétation musicale effectuée corporellement et projetée acoustiquement. *L'aura poétique désignera ici la chose non musicale (« poétique ») dont ce phrasé musical est devenu le nom*, ce qui revient à dire - et ceci nous ramène à la mathématique de Cohen - que le phrasé interprétatif va simultanément s'affirmer comme phrasé musicalement ordinaire de l'OMM et comme phrasé surnuméraire de son aura poétique.

16.4. Trois temps, pour la musique comme pour les mathématiques...

Si l'on reprend en effet les trois temps de l'opération mathématique de torsion précédemment mise à plat, on dira :

1) Il va d'abord s'agir de travailler l'aura (*i.e. la chose*) indirectement en travaillant directement les segmentations qui phrasent l'aura (*i.e. les mots qui nomment la chose*).

2) Il va ensuite s'agir de constituer l'aura (*i.e. la chose*) à partir de ces segmentations musicale - de composer l'aura à partir de ces segmentations - en sorte de les doter d'un nouveau pouvoir de phraser (*i.e. d'un nouveau pouvoir de nomination*).

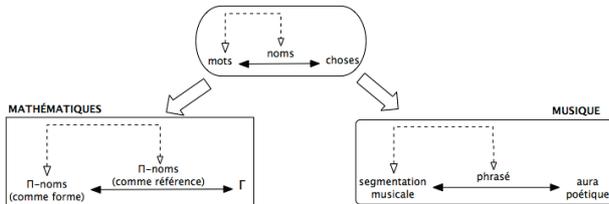
3) Il va s'agir enfin que ce nouveau pouvoir interprétatif (consistant à phraser des segmentations enrichies par le jeu de la modulation $T \otimes P \Rightarrow M$) tienne au fait que cette nouvelle manière de segmenter le texte musical enrichi phrase bien... ce texte musical et non pas phrase un programme littéraire sous-jacent en convoquant pour ce faire un principe narratif de nature extra-musical : le péril, pour l'OMM, serait ici que le phrasé musical se dégrade en narration non musicale. C'est, en effet, en cette torsion singulière où le nouveau pouvoir musical de segmenter - phraser - doit se rabattre sur la musique (et non pas se laisser happé par une pseudo-problématique du sens, de la narration ou du récit) que réside tout le sel et la puissance propre de ce jeu de part en part musical, l'enjeu de cette puissance

³⁹ C'est-à-dire d'ordre inférieur

propre étant précisément la capacité de l'OMM de générer – composer et contrôler musicalement – une aura poétique et non pas de prétendre raconter une histoire (fut-ce un pattern générique d'histoire), faire récit musical d'une aventure anthropologique ou donner sens musical à un programme littéraire.

16.5. Une analogie globale

Somme toute, nous travaillons ainsi sur l'analogie suivante :



17. UNE « ÉVALUATION » DES « NOMS » ET UN « FORCING »

Revenons à notre construction mathématique d'une partie générique Γ au moyen de Π -noms sans trop nous attarder sur la technicité proprement mathématique de la chose.

N'oublions pas, cependant, le danger spécifique de ce type de raccourci : celui de glisser sans vergogne d'une rigoureuse discipline discursive au jeu périlleux des images...

17.1. Une nouvelle étape capitale

L'étape décisive de la torsion effectuée par Cohen sur ses mots-noms (étape rendant raison mathématique de cette construction si singulière où l'auto-référence, ancrée sur une première opération d'apparence anodine sur le vide⁴⁰, s'avère n'être pas une tautologie mais bien le socle d'un déploiement hiérarchique de très grande portée conceptuelle) va tenir à ce qu'il appelle l'évaluation de ses noms c'est-à-dire précisément leur capacité de nommer quelque chose qui concerne l'ensemble générique Γ visé.

17.2. Une évaluation autorisant la nouvelle opération de forcing

Cette évaluation va en effet déboucher sur le grand résultat de cette théorie : l'opération dite de *forcing* qui va permettre de composer et contrôler de l'intérieur de l'ensemble de départ Θ l'ensemble générique Γ qu'on lui ajoute.

L'idée est la suivante : le contrôle de G par les Π -noms internes à la situation de départ Θ va se faire via la construction de formules ϕ (internes à Θ) sur ces noms en sorte qu'une formule ϕ soit, comme les Π -noms, dotée d'une double face : l'une qui énonce sur les

noms dans Θ , l'autre qui énonce sur les référents de ces noms dans l'extension $\Theta[\Gamma]$.

Rappelons : l'extension n'est pas l'ajout d'un simple appendice Γ à Θ (pas plus que l'extension de \mathbb{Q} ne se fait pas simple ajout du seul nombre « irrationnel » $\sqrt{2}$) mais bien par adjonction de tout ce que cet appendice autorise comme nouvelle composition avec Θ .

Pour nous, l'aura poétique de l'OMM ne consiste pas seulement en un nouveau « sens » donné par la musique au texte de départ quand elle le chante mais bien dans l'ensemble des effets *sur la musique* que produit cette nouvelle capacité musicale à phraser un texte non musical. Ainsi l'aura est ici dite *poétique* pour indiquer sa généricité (certes musicalement composée mais cependant musicalement indiscernable de l'intérieur du monde-*Musique*) et cette aura poétique ne se réduit nullement à une auréole de sens poétique circonscrivant l'OMM : elle s'attache à une irradiation générale et en tout point de l'OMM.

L'opération de *forcing* va se matérialiser dans des énoncés de ce type : « la condition π force la formule ϕ » où « forcer » prend le sens précis suivant : si la formule ϕ (sur des noms internes à la situation de départ Θ) est vérifiable dans cette situation de départ Θ , alors l'énoncé sur les référents de ces noms dans l'extension $\Theta[\Gamma]$ sera véridique ; soit : une relation vérifiable dans notre situation de départ Θ « force » une véridicité dans la situation étendue $\Theta[\Gamma]$.

17.3. Portée musicale de tout cela

Pour nous, la transposition de cette dernière étape va être immédiate (au prix, il est vrai, de nombreuses approximations, à commencer par celle – qui n'est pas mince – consistant à transposer une situation infinie sur la structure finie d'une œuvre musicale...): l'appropriation (musicale) d'un phrasé (musical) aux segmentations (musicales) dont une partition (musicale) est (musicale) capable *force* la pertinence (poétique, c'est-à-dire non exclusivement musicale) d'un énoncé (poétique) sur l'OMM (musicale et poétiquement) irradiée par son aura (poétique).

Ainsi, pour donner deux exemples de ce « forçage » en matière d'OMM, il y aura pertinence *musicale* (tout le point, bien sûr, est dans cette qualification) à ce qu'un musicien déclare que l'interprétation par Thomas Hampson des *Dichterliebe* de Schumann force à admettre que ce cycle de lieder met bien musicalement à l'œuvre une position sexuée d'homme (lors même, comme je le soutiens opiniâtrement, que le monde-*Musique* ignore la différence des sexes qui pourtant marque tout musicien).

Tout de même, il y a pertinence *musicale* à soutenir en musicien que l'interprétation de mon *Duelle* phrase les langues anglaise, allemande et russe (telles que musicalement segmentées dans l'œuvre⁴¹) et que leur mise en œuvre force la musique à énoncer quelque vérité sur certaines particularités de ces trois langues.

On pressent ici l'immense pas franchi par ce type d'extension poétiquement auratique.

⁴⁰ Ici le premier Π -nom est \emptyset , ce qui autorise que le second soit $\langle \emptyset, \pi_1 \rangle$, tout de même que pour la théorie des ensembles le premier ensemble est \emptyset ce qui autorise que le second soit $\{\emptyset\}$.

⁴¹ Dans ses parties III, V et VII.

18. RAPPELS

Rappelons ici deux points cruciaux.

18.1. Auto-limitation musicienne

Il faut autolimiter l'exaltation du musicien ; l'extension dont il est ici question est celle d'une OMM, puis d'une autre, non pas celle du monde-*Musique* comme tel.

Si tout morceau de musique est bien susceptible de servir de point de départ pour une OMM, encore faut-il pour cela (pour transformer donc ce morceau en OMM) d'abord compléter ce morceau d'un flux hétérogène \mathcal{H} ; ensuite transformer ce morceau non seulement pour qu'il soit bien modulé par ce nouveau flux \mathcal{H} (et ainsi augmenté, sans déchirure ou boursoufflure induite, d'une partie M : d'une mélodie, par exemple) mais, plus encore, pour que la segmentation écrite du nouveau morceau (augmenté de M) soit effectivement phrasable par un interprète en *raisonance* musicale avec T.

18.2. Œuvre, et pas simple pièce

Autant dire : il faut que ce morceau assure désormais d'être une œuvre (une œuvre « musicale » mixte) et pas simplement une pièce de musique hybridée d'un texte (mélodrame plaqué), d'une chorégraphie (Cunningham sauvant chorégraphiquement l'insignité musicale de Cage) ou d'un film (un cinéaste venant sauver une musique comme on peut sauver telle symphonie de Chostakovitch grâce à de judicieuses mises en scène cinématographiques).

Bref, le monde-*Musique* ne saurait être irradié de part en part en tout point selon de telles auras et il n'y a pas lieu pour le musicien de concevoir une aura poétique du monde-*Musique* comme tel (qui viendrait alors tenir fantasmatiquement lieu de foyer musical apte à réenchanter le *chaosmos*...).

Notre aura poétique de l'OMM est donc à la fois plus précise, plus matériellement inscriptible dans les opérations proprement musicales du musicien (écrire, composer, lire, interpréter, jouer, phraser, entendre, écouter...) et finalement plus décisive qu'une supposée irradiation générale du monde-*Musique* car cette capacité musicale à l'aura poétique repose ultimement sur l'existence d'une *intension* proprement musicale (celle précisément qui distingue l'authentique œuvre de la simple pièce) c'est-à-dire d'une véritable ambition musicale mise en œuvre par tel morceau de musique (et, bien sûr, qui dit ambition dira alors discipline, en particulier discipline des conséquences et donc des gestes musicaux de pensée posés à tel ou tel moment du discours musical⁴²).

⁴² Il est clair que c'est souvent à ce point précis que l'oreille reconnaît l'absence d'ambition ou de projet musical d'une simple pièce : quand l'oreille comprend que le développement musical de la pièce, savant ou maladroit, n'est en tous les cas pas autonomé par une discipline immanente de décisions singulières puis de leurs conséquences...

19. RÉSUMÉ

Résumons notre propos.

L'OMM accueille un flux hétérogène \mathcal{H} , préservant l'hétérogénéité propre de son *aspect* (le poème mobilisé reste présenté à l'oreille comme texte signifiant, et pas décomposé en phonèmes insignifiants) mais violentant son *inspect* (le poème n'est pas nécessairement restitué en sa continuité native, en son tempo primitif, en son phrasé poétiquement naturel, en sa forme globale constitutive). Cet aspect hétérogène module l'aspect du flux musical et le complète d'un nouvel objet musical M qui vient dans l'OMM faire office de frontière musicale entre logique musicale et logique hétérogène.

Si l'*intension* hétérogène propre à \mathcal{H} est alors bien mise en *raisonance* avec la musique (c'est ici que la composition et l'interprétation propres à une œuvre – qui n'est pas une simple pièce – entrent spécifiquement en scène), alors la musique module en retour le flux \mathcal{H} , le modèle selon un *inspect* proprement musical qui s'avère apte à faire retour sur l'*intension* « poétique » (c'est-à-dire hétérogène) du flux \mathcal{H} et à irradier toute la musique d'une nouvelle puissance auratiquement poétique.

En bref, la musique, violentant l'*inspect* du poème mais tirant parti de son aspect pour moduler le flux musical, pour peu qu'elle sache alors ne pas brutaliser le flux hétérogène et entrer en *raisonance* avec son *intension* proprement poétique, peut en retour moduler son *inspect* selon une logique cette fois musicale et par là moduler musicalement son *intension* poétique.

On schématisera ainsi la dynamique ici esquissée :

	\mathcal{H}	\mathcal{T}	
Aspect	• →→→→	•	composition
Inspect	• ←←←←	•	interprétation
Intension	• ←.....→	•	écoute
	<i>raisonances</i>		

20. SCHÉMA GÉNÉRAL DE NOTRE « PRATIQUE THÉORIQUE »

Ramassons notre expérimentation théorique sur la théorie des extensions génériques.

À l'enchaînement des notions mathématiques suivantes – espace de travail (Θ), ensemble ordonné des conditions (Π), ensemble générique ajouté (Γ), Π – noms dotés de leur Γ – évaluation, extension $\Theta[\Gamma]$ et relation de *forcing* – nous avons fait correspondre l'enchaînement des catégories musicales suivantes : OMM, espace des périodes produit par une segmentation musicale de \mathcal{T} , rétroaction auratique ajoutée, phrases musicales de \mathcal{T} produites par le phrasé interprétatif et modelant le flux \mathcal{H} , extension auratique

21.3. Aujourd'hui...

Cette fécondation possible de la musique (de son discours, de sa logique, de ses *intensions* propres) par de tout autres logiques artistiques constitue une ressource mobilisable par les musiciens dont l'importance aujourd'hui n'a sans doute rien à rendre à celle que Richard Wagner relevait pour son propre compte (au lendemain des Révolutions écrasées de 1848-1849) sous le signe d'un Drame où le poème devenait susceptible de féconder la musique.

Posons-le tout net : il nous semble aujourd'hui indispensable de mobiliser ces ressources *musicales* de l'œuvre musicale mixte si l'on veut continuer la musique comme art de l'écoute par-delà un XX^e siècle qui fut certes musicalement flamboyant mais s'avère désormais saturé en sorte que le XXI^e siècle musical ne saurait en être la continuation comme le XIX^e musical a pu l'être du XVIII^e.

21.4. Le musicien, moins encore que la musique, ne pense seul !

Qu'en ce point (où il s'agit de continuer la musique en relançant, à nouveaux frais, le vieux geste par lequel la musique se régénère par accueil et captation prédatrice d'autres *intensions* artistiques), la mathématique soit la pensée qui éclaire le musicien et l'aide à s'orienter n'est pas pour déplaire à l'auteur de ce texte.

22. RÉFÉRENCES

- [1] Adorno, T, *Die Kunst und die Künste*. Conférence du 23 juillet 1966 à Berlin.
- [2] Benjamin, W. *Écrits français*. Folio, essais.
- [3] Adorno, T, *Théorie esthétique*. Klincksieck, 1989.
- [4] Adorno, T, *Moments musicaux*. Contrechamps, 2003.
- [5] Wagner, R, *Lettre sur la musique*, à Fr. Villot, 1860.
- [6] Wagner, R, *Œuvres complètes*.
- [7] Parménide, *Le Poème* (trad. Jean Beaufret). PUF ; 1984.
- [8] Cohen, Paul J. *Set Theory and the Continuum Hypothesis*. Mathematics Lecture Note Series; The Benjamin/Cummings Publishing Company; 1966.
- [9] Badiou, A., *L'être et l'événement*. Seuil, 1988.

AURA ET PERCEPTION RÉFLEXIONS SUR LA MUSIQUE MIXTE À PARTIR DE BENJAMIN

Antoine Bonnet
Compositeur, Professeur à
l'Université Rennes 2
antoine-bonnet@wanadoo.fr

RÉSUMÉ

Pour Benjamin, la reproduction mécanisée détruit « l'aura » de l'œuvre, ce qui en émane du fait de l'unicité de son existence au lieu où elle se trouve, la salle de concert par exemple. A la « perception auratique », fondée sur la « contemplation » de l'œuvre propre à l'époque de la « fonction artistique de l'art », se substituerait alors la « perception tactile », fondée sur la « distraction et l'accoutumance » propre à notre époque de la « fonction politique de l'art ».

La technologie mise en œuvre dans le cadre de la musique mixte ne permet-elle pas en réalité une appropriation matérielle de cette aura ? La question est ici réfléchie à partir de l'exploitation musicale de la notion de résonance et des « effets de durée » générés par le temps réel.

*

On connaît la thèse de Benjamin dans son célèbre texte des années 1930 *L'Œuvre d'art à l'époque de sa reproduction mécanisée* [1]. En substance, l'invention de la photographie dans les années 1820-1830 a inauguré une époque absolument nouvelle au cours de laquelle la notion d'œuvre d'art, telle que constituée *grosso modo* depuis la Renaissance, ne pouvait être que radicalement bouleversée.

Ce n'est pas la reproduction en elle-même qui est ici en cause. Benjamin souligne d'ailleurs qu'il est du principe de l'œuvre d'art d'avoir toujours été reproductible, de la copie des œuvres dans les ateliers, à des fins d'apprentissage, aux techniques artisanales toujours plus sophistiquées de la gravure jusqu'à la lithographie qui, de façon plus significative, hissa la reproduction des dessins au niveau de celui des textes qu'ils pouvaient alors quotidiennement illustrer.

Mais le pas effectué par l'invention de la photographie, quelques années seulement après celle de la lithographie, est pour Benjamin sans commune mesure : au profit de l'œil elle décharge la main des tâches artistiques les plus importantes. Et parce que l'œil saisit plus vite que la main ne dessine, la reproduction des images se fait à un rythme assez accéléré pour suivre la cadence de la parole. « Si la lithographie contenait virtuellement le

journal illustré, dit Benjamin, la photographie contenait virtuellement le cinéma parlant ».

Avec l'invention de la photographie puis du cinéma et entre temps du phonographe et du téléphone, c'est en effet toute une conception de l'espace et du temps qui vacille : la fin de l'unicité de l'existence d'une chose au lieu où elle se trouve. Pour l'œuvre d'art, cette fin est celle de son ici et maintenant constitutif et la perte de son authenticité, le détachement de sa tradition et la destitution de son autorité – bref, la destruction de ce que Benjamin appelle son *aura*.

A l'exception d'Adorno, cette notion d'aura, à l'instar d'ailleurs de l'œuvre entière de Benjamin, n'a guère plus interpellé les musiciens et penseurs soucieux de la musique que Benjamin lui-même ne s'est soucié de la musique ; autant dire quasiment pas. Près de cinquante ans plus tard le terme apparaît toutefois très discrètement dans un article de Boulez, *Le système et l'idée* [2]. Il n'y fait nullement référence à la notion de Benjamin ni d'ailleurs à Benjamin tout court. Du reste, à première vue très loin de la signification générale que revêt l'aura chez le penseur allemand, le terme désigne chez le compositeur une notion très spécifiquement musicale et d'apparence plutôt anodine. A l'instar de l'appoggiature dans le système tonal, elle correspond à une nécessité d'ordre expressive et technique. Dans le prolongement implicite de ce que Messiaen appelle l'agrandissement des notes étrangères, il s'agit de greffer une structure adjacente sur une structure principale, sans départir cette dernière de ses fonctions constitutives, en sorte de créer localement un halo sonore qu'il appelle une *aura*.

J'attache pour ma part une importance décisive à cette notion d'aura ainsi définie par Boulez pour la compréhension de son œuvre en général et à partir du milieu des années 70 en particulier, c'est-à-dire de l'époque où il élabore le projet de l'Ircam avec déjà en tête cette idée du temps réel dont on sait qu'il l'imposera très vite comme programme de recherche de l'institut. Pour aller vite je dirai simplement ceci : la technologie du temps réel est ce qui permet à Boulez de générer à profusion des structures adjacentes - des auras donc - et de les greffer sur les structures principales du discours que conduisent les gestes instrumentaux. Pour reprendre le parallèle avec le système tonal, les événements

généralisés par la technologie du temps réel, les auras, sont donc au langage boulezien ce que les appoggiatures sont au langage tonal. Le médium de l'expression se déplace ainsi de l'appoggiature vers l'aura, autant dire de la note vers le son.

Il y a là un modèle général très simple, comme toujours chez Boulez, mais qui est susceptible d'infinies ramifications et transpositions, y compris dans le domaine de la musique purement instrumentale. Ainsi par exemple dans *Messagesquisse* [3] les violoncelles secondaires ont la même fonction par rapport au violoncelle principal que les événements sonores en temps réel par rapport aux solistes dans *Répons* : une fonction expressive de halo sonore, une fonction d'aura. Concrètement, de quoi sont faites ces auras ? De ce dont est capable la technologie en temps réel, soit d'opérations se ramenant à des répétitions variées dans l'espace et dans le temps, essentiellement des transpositions et des délais. C'est à la fois une limitation - on le sait à l'écoute des effets produits, plutôt redondants - mais également un avantage pour un compositeur aussi soucieux de cohérence structurale que Boulez : cela lui permet de réfléchir, dans tous les sens du terme, le temps réel dans le cadre général des couples dialectiques figure/structure ou prolifération/intégration.

A la même époque Boulez parle volontiers d'hétérophonie, d'écriture virtuelle, d'illusion acoustique, de contrastes de plan, d'accumulations de points de vue ou encore de « surdimension [qui] ne crée pas l'objet proprement dit [mais en] multiplie l'image ». Ce qui est ici frappant, c'est qu'il commente son travail de composition en des termes qui non seulement conviendraient au cinéma mais, plus frappant encore pour ce qui nous occupe ici, qui sont similaires à ceux qu'utilise Benjamin pour décrire les opérations du cinéma, opérations qui sont justement pour lui les opérations techniques de la destruction de l'aura. Il y a là un chassé-croisé des plus amusants qui, dans un premier temps, ne peut que conduire à chercher ailleurs que dans le sens donné à l'aura par l'un et l'autre un lien quelconque entre leurs préoccupations.

Comme je l'ai montré ailleurs [4], ce qui rapproche la musique de Boulez de ce que dit Benjamin du cinéma et plus généralement de l'œuvre d'art à l'époque de sa reproduction mécanisée, c'est un mode de perception dont Benjamin voit le modèle canonique dans le plus constant des arts en dépit de l'évolution considérable de ses matériaux et de ses techniques : l'architecture.

Mais avant de poursuivre une précision concernant Benjamin. Du fait des termes mêmes de sa célèbre définition de l'aura - « Une singulière trame d'espace et de temps : l'unique apparition d'un lointain, si proche soit-il » -, certains - notamment des penseurs d'obédience marxiste - taxèrent Benjamin de penseur obscurément mystique et messianiste, condamnant l'art du XX^e siècle au nom d'un art auratique qui tirerait son autorité de son articulation consubstantielle à quelque principe transcendant. Il est vrai que les textes de

Benjamin sont foisonnants et pas toujours d'une grande clarté d'exposition. Il faut cependant être d'assez mauvaise composition pour y trouver - en particulier à partir de la fin des années 20, qui sont les années d'expansion du fascisme et concomitamment de l'inflexion matérialiste de sa pensée - une trace quelconque de nostalgie à l'endroit d'un certain régime de l'art dont il sait, bien avant tant d'autres, qu'il est irrémédiablement condamné pour toutes sortes de raisons. Certes, comme l'a relevé Anne Boissière [5], Benjamin, à la différence de Marx, ne part pas de l'idéologie pour considérer l'art comme un reflet de la vie matérielle ; il l'appréhende au contraire sans préjuger de ses liens au régime de la production et des échanges. Mais ce n'est nullement pour le délier du processus historique. Plus empirique qu'idéologue, ce qui lui importe avant tout c'est de comprendre, en liaison avec l'invention des nouvelles techniques, l'incidence des mutations de l'art sur l'évolution des modes de perception, elle-même concomitante avec celle des phénomènes de masse.

En substance, pour ce qui nous occupe ici, l'analyse de Benjamin est la suivante : à l'époque des grands tournants de l'histoire - et les nouvelles techniques créent précisément pour lui les conditions d'un tel tournant -, le seul mode de perception sur lequel l'art soit en mesure de pouvoir compter est celui qui depuis toujours s'attache à l'architecture : celui, collectif, fondé sur l'attention distraite et l'accoutumance, qu'il appelle la perception tactile par opposition à la perception auratique reposant sur la contemplation.

Revenons maintenant à Boulez. Ce qui frappe dans sa musique depuis le milieu des années 70, c'est ce que j'appelle leurs effets de durée. Je désigne par là l'effet que produit un passage voire une œuvre entière dont les caractéristiques globales sont telles que son expression est une fonction de la durée. L'influence de Wagner, qu'il a dirigé à cette époque six années de suite à Bayreuth, est ici manifeste ; que l'on songe au début de *l'Or du Rhin* ou encore à la *Mort d'Iseult*. Une conséquence de ces effets de durée sur la perception, effets largement favorisés par l'ample déploiement sonore du temps réel, est précisément de ne demander qu'une attention distraite à l'auditeur. Je sais que le propos a de quoi surprendre tant Boulez est dans l'opinion le paragon du compositeur exigeant et plus soucieux de la pureté de son art que du confort de son auditeur. Pourtant qu'est-ce d'autre, de prime abord, qu'une attention distraite que demandent à l'auditeur les halos sonores l'enveloppant de toute part dans *Répons* ou *Dialogue de l'ombre double*, les répétitions incantatoires d'un simple profil infiniment orné dans *Rituel* ou les résonances inlassablement explorées des gestes virtuoses dans *Sur Incises* ?

Certes, la musique de Boulez ne se réduit pas aux effets sonores qu'elle produit, loin s'en faut. En référence à Kafka j'ai du reste appelé *terrier* le centre nerveux de sa musique, lequel constitue l'autre face de sa dialectique compositionnelle. En réalité, l'aura et le terrier sont les

deux termes d'une stratégie. D'une part le terrier - l'écriture pour aller vite - préserve les exigences de l'art telles qu'héritées d'une tradition savante occidentale à laquelle Boulez ne saurait renoncer mais dont la logique combinatoire en elle-même n'est plus en mesure de produire un discours musical convaincant ; d'autre part, et conséquemment, l'aura offre une réponse possible au problème de l'évolution des modes de perception en relation avec celle des techniques. En somme, Boulez prend acte de ce que - aux grands tournants de l'histoire, pour parler comme Benjamin - la musique - sans s'y réduire - doit proposer un régime de perception offrant de prime abord un niveau tactile, c'est-à-dire reposant sur la distraction et l'accoutumance. Ce qui se traduit par une recherche expressive fondée sur le déploiement local de structures adjacentes foisonnantes, de halos sonores incessamment dupliqués et variés à la faveur du temps réel qu'il appelle des auras.

Ce déploiement local d'auras, au sens restreint que leur donne Boulez, participe bien sûr de manière décisive à façonner l'œuvre au niveau global, à lui attacher un univers sonore spécifique que l'on peut comprendre à son tour comme une aura, mais en un sens général cette fois : non plus l'aura locale, transposition de l'appoggiature, mais l'aura globale de l'œuvre tout entière, l'univers sonore qu'elle instaure et qui coïncide avec son apparition. Autrement dit, l'aura, au sens restreint, est l'opérateur de l'aura au sens général de l'œuvre, l'opérateur logique de son apparaître.

Voilà qui cette fois nous rapproche de l'aura telle que l'a définie Benjamin : « l'unique apparition d'un lointain, si proche soit-il ». Le terme *apparition* désigne le résultat de l'invention/production même de l'œuvre, la manifestation sensible de sa réalité matérielle. Celle-ci ne s'épuise cependant pas comme telle, d'où les termes de *proche* et de *lointain* ; ils traduisent que cette apparition ne saurait être intégralement reproductible et qu'elle est donc unique. L'évocation benjaminienne d'un horizon associé à la contemplation d'un paysage est à cet égard significative. Il s'agit de souligner la dimension d'infini qui se dégage de toute expérience vécue, de ce qui est unique au sens strict de non reproduit.

Cette définition de l'aura est évidemment très générale et vaut en droit pour toute œuvre non mécaniquement reproduite. Elle est toutefois singulièrement parlante concernant la musique en raison du caractère évanescant du son. Tout morceau joué en un lieu et un moment donnés est bien en effet l'unique apparition d'une réalité à la fois tangible et insaisissable. Soulignons en outre que depuis le XIX^e siècle la tendance de la musique à déplacer le vecteur de l'expression de la note vers le son accentue la pertinence de la notion d'aura. Le son, en effet, en tant qu'il ne se réduit pas aux paramètres de hauteur et de durée à partir desquels il était autrefois contrôlé et qu'il est de plus en plus mis en avant dans sa complexité constitutive, s'accorde bien à la notion d'aura, tant au sens propre d'émanation d'une substance que figuré d'atmosphère dégagée par un être. Dans le cas de la musique mixte en général et du temps réel en

particulier - surtout lorsque s'y associe un dispositif de spatialisation, ce qui est généralement le cas -, la notion d'aura paraît alors d'autant plus appropriée que l'œuvre amplifie la sensation paradoxale de proche et de lointain à proportion de ce que le travail sur le son crée à la fois une présence enveloppante, donnant l'impression d'être au cœur du phénomène sonore, et des illusions acoustiques, donnant au contraire l'impression d'un univers insaisissable.

Autrement dit, la technologie ouvre à la possibilité d'une appropriation matérielle de l'aura. Non bien sûr que cette appropriation soit totale et sans reste, mais que ce qui était autrefois hors de toute possibilité de contrôle, et d'ailleurs hors de toute visée musicale proprement dite, se trouve désormais investi et porté au centre du processus compositionnel. Là encore Wagner et sa fosse d'orchestre de Bayreuth pourraient nous guider dans la réflexion. Je prendrai toutefois ici le cas de la résonance acoustique. Non que la résonance ait évidemment le monopole de l'aura - le terme, qui désigne étymologiquement le souffle, conviendrait tout aussi bien à l'orgue par exemple - mais que la résonance d'un instrument, disons celle du piano, correspond bien intuitivement à la traduction dans le domaine sonore de ce que peut être une aura au sens matériel du terme : l'émanation d'un corps physique, en l'occurrence le rayonnement acoustique d'un instrument.

On peut faire un parallèle entre la résolution des problèmes techniques de facture instrumentale liés à la production de la résonance du piano et l'exploitation musicale de cette résonance. Il est en effet manifeste que la progressive constitution du piano moderne a accompagné la non moins progressive évolution du langage harmonique. Mais à partir de Debussy - dont précisément le langage harmonique cesse d'être fonctionnel au sens de la tonalité pour mieux s'ouvrir à l'univers des sons - se fait jour un souci nouveau, celui de la résonance pour elle-même, souci que je serai tenté de nommer ici celui de l'aura pianistique elle-même. Bien des titres de ses œuvres pour piano semblent d'ailleurs le suggérer ; pensons par exemple aux préludes *La terrasse des audiences du clair de lune* ou *Les sons et les parfums tournent dans l'air du soir*.

En liaison avec le raffinement des pédales de piano - *forte* et *una corda* d'abord puis *sostenuto* et plus récemment *harmonique* - l'intérêt pour la résonance sera celui de bien d'autres compositeurs à la suite de Debussy, et en particulier de Boulez qui sera un des premiers à organiser systématiquement ses compositions à partir de la catégorie spécifique d'instruments résonants. A côté d'œuvres pour piano, la *Troisième sonate* et le *Deuxième Livre des structures* notamment, *Improvisation II* sur Mallarmé, *Eclats* et *Sur Incises* en sont les manifestations les plus évidentes dans le domaine acoustique ; et bien sûr *Répons*, dans celui de la musique mixte, avec le dialogue constitutif de l'œuvre entre la technologie en temps réel et les six solistes résonants.

Il est clair qu'il y a dans l'intérêt initial de Boulez pour la résonance le germe de ce qu'il appellera l'aura à l'époque de l'Ircam. Mais une aura qu'il ne peut alors se déployer au-delà des capacités acoustiques naturelles des instruments en liaison avec les paramètres élémentaires des attaques qui les mettent en vibration. Jusque dans les années 60 Boulez explore ainsi systématiquement la tripartition attaque/résonance/extinction en fonction du registre et de l'inertie, des modes d'attaque et de l'intensité, du poids acoustique et du mélange des instruments mis en jeu. Mais à partir des années 70 Boulez prolonge et entretient la résonance par la démultiplication des figures autrefois simplement dévolues au rôle de son déclenchement. Cette démultiplication des figures en fonction de l'impulsion initialement donnée par les instruments - que l'on pourrait thématiser sous le chef de la notion d'écho et dont l'arpège d'arpège est un exemple-type - est le second temps du travail de Boulez sur la résonance et donne le programme de ce que devra concrètement réaliser la technologie du temps réel.

Du coup la résonance cesse de devenir l'apanage des instruments résonnants naturellement pour devenir une idée compositionnelle à part entière et indépendante des instruments mis en jeu. Les *Notations* pour orchestre, par exemple, ne sont ainsi pas de simples orchestrations des pièces de jeunesse pour piano du même nom mais leur transmutation sonore, leur mise en résonance orchestrale, l'écriture contrôlée de leur aura [6]. Le temps réel et la transposition de son principe de démultiplication dans le domaine purement acoustique témoignent donc bien d'une appropriation matérielle de l'aura.

Toutefois cette appropriation matérielle de l'aura ne règle pas le problème de sa reproduction, c'est-à-dire de son enregistrement. On peut même dire qu'elle l'aggrave, en particulier dans le cas des œuvres associées à des dispositifs de diffusion et de spatialisation. Les techniques de démultiplication de l'image sonore amplifient le phénomène auratique de l'œuvre dans le registre de sa production à proportion de ce qu'elles le réduisent dans celui de sa reproduction. Ce n'est en effet plus alors seulement l'aura de l'œuvre au sens de son ici et maintenant qui est perdu, mais également son aura matérielle, constitutive de son projet même.

En réalité le problème touche plus généralement à celui de la diffusion des sons par les haut-parleurs dans les musiques mixtes, y compris en situation de concert. Celles-ci tendent en effet à créer deux mondes séparés : d'un côté celui des sons acoustiques rayonnant à partir des instruments, de l'autre celui des images sonores projetées par les haut-parleurs, lesquelles ne paraissent pas homogènes alors même qu'elles sont objectivement l'émanation auratique des instruments acoustiques par le truchement du temps réel. Autrement dit, via la diffusion des haut-parleurs, le problème de la destruction de l'aura au sens de Benjamin resurgit au

lieu même du concert où par définition il ne devrait pas se poser.

Des tentatives de résolution de ce problème sont en cours d'élaboration ; elles consistent à créer des haut-parleurs se modelant sur le rayonnement acoustique des instruments. C'est par exemple le cas de la *Timée* développée à l'Ircam. Traduit dans les termes de la présente communication, il s'agit de doter les haut-parleurs d'une aura au sens de Benjamin, c'est-à-dire comparable à celle des instruments acoustiques entendus en situation de concert, mais une aura augmentée d'une autre au sens de Boulez, c'est-à-dire constitutive du projet compositionnel. En somme la *Timée* dans le domaine de la diffusion et l'aura dans le domaine de la composition s'emploient à résoudre un problème d'homogénéisation : l'une par transposition du rayonnement instrumental naturel par le médium électroacoustique, l'autre par transposition de la démultiplication sonore artificielle dans le monde acoustique.

La tendance est donc ici à domestiquer les nouvelles technologies en sorte d'en transposer les possibilités sur le modèle du monde instrumental. La raison en est simple : les machines ne sont pas suffisamment expressives. Voilà, du point de vue strictement musical, qui devrait donner le *la* des priorités dans tous les secteurs de la recherche en informatique musicale et en acoustique.

Il serait pourtant vain d'escompter des machines les modes d'expression auxquels la musique instrumentale nous a accoutumés. Ceux-ci n'ont cessé d'évoluer, certes, mais en se cantonnant dans le registre que circonscrivait l'artisanat, le *fait main* comme dit le mot allemand. Depuis l'invention de l'enregistrement, la musique est confrontée au franchissement du cap auquel les arts visuels ont été confrontés avec le passage de la lithographie à la photographie et ce que dit Benjamin de la vitesse de l'œil vaut désormais pour l'oreille. En 1957, après avoir composé *Gruppen* qui lui avait demandé deux ans de travail manuel, Stockhausen déclarait sous forme de boutade vouloir composer à la vitesse de la musique. Cinquante ans plus tard le temps réel est devenu une seconde nature pour les musiciens.

Mais plus encore que la production, la reproduction et la diffusion des œuvres, le temps réel affecte le mode de perception de la musique. Ainsi ce qui se présentait comme difficile et problématique il y a encore peu de temps semble s'être déplacé : non plus la musique contemporaine ou même moderne par opposition à la musique tonale réputée plus accessible, mais l'écoute elle-même. A l'heure de l'internet, de l'instantanéité, du zapping généralisé et de l'immersion quasi permanente dans quelque atmosphère musicale, être confronté à une musique quelle qu'elle soit et devoir ne rien faire d'autre que l'écouter pour elle-même, y faire face, en faire l'expérience intégrale dans sa durée propre, paraît devenir une exigence éminemment singulière, nullement spontanée voire même angoissante. A moins, et c'est là

encore quelque chose de symptomatique, d'être en mesure de se l'approprier en la produisant ou la jouant soi-même, activités tendant d'ailleurs à devenir étroitement liées avec les nouvelles technologies.

On rejoint là une autre caractéristique de la perception tactile de Benjamin, caractéristique qu'il dit politique en ce qu'elle est susceptible de transformer en profondeur la cité elle-même : celle d'être productrice, c'est-à-dire de conduire tendanciellement tout être qui perçoit à pratiquer à son tour. Cette fonction productrice de la perception, en tant qu'elle en deviendrait une condition, déplace en effet de façon radicale le problème de la perception elle-même puisqu'à terme elle ne produirait plus que des acteurs de la musique. La chose paraît curieuse. Mais n'est-ce pas ce à quoi semble déjà conduire d'un côté l'enfermement individuel dans les baladeurs, dont on peut prévoir qu'ils finiront par se transformer en studios miniatures personnalisés, et de l'autre la participation corporelle à la célébration rituelle d'un son démesurément amplifié lors de cérémonies collectives de type *rave party*.

Voilà qui nous amène à une autre partie de l'analyse de Benjamin. Selon lui, la perception tactile aurait commencé à s'imposer avec les passages parisiens du XIX^e siècle [7]. Le passant, distrait et s'appropriant par accoutumance tant la nouvelle architecture de fer et de verre que les panoramas y préfigurant le cinéma, s'opposait alors déjà au flâneur d'antan et redoublait la figure naissante du travailleur. En réalité, pour Benjamin, ce qui commence à s'épuiser à partir de l'ère industrielle, c'est la fonction artistique de l'art, celle qui opère *grosso modo* de la Renaissance à la Révolution française. Il lui oppose l'antique et préhistorique fonction religieuse de l'art d'une part, et son actuelle fonction politique d'autre part. Il souligne que la nécessité pour l'art d'être vu ou entendu a été son propre à l'époque de sa fonction artistique, mais nullement à celle, antérieure, de sa fonction religieuse où il suffisait à l'objet d'être présent. Sa valeur était alors culturelle et sa fonction magique dans la mesure il s'agissait de dominer par un geste mimétique ce qu'il appelle la *première nature*. L'art à l'époque de sa fonction artistique aurait ensuite acquis une valeur d'exposition qui logiquement allait développer l'acuité visuelle et auditive de qui l'appréhendait et faire émerger la problématique de la perception. Le rapport de domination à la nature devenait alors non plus magique mais rationnel. Enfin, avec l'émancipation de la technique, apparaissait une seconde nature requérant de l'homme un apprentissage analogue à celui dont il avait eu besoin face à la première nature. « Une fois de plus, dit-il, la fonction [de l'art] est de soumettre l'homme à un entraînement ; il s'agit de lui apprendre les aperceptions et les réactions que requiert l'usage d'un appareillage technique dont le rôle s'accroît presque tous les jours. Faire de l'immense appareillage technique de notre époque l'objet de l'innervation humaine, telle est la tâche historique au service de laquelle le cinéma [l'art] trouve son véritable sens ».

Il serait donc vain d'attendre des machines les modes d'expression conjointement inventés par l'artisanat rationnel et la perception auratique à l'époque de la fonction artistique de l'art. Elles ne peuvent à terme que modifier notre rapport au sensible, aiguïser nos sens par l'audition répétée de ce dont elles sont capables et nous imprégner tactilement d'une réalité nouvelle mais appelée à devenir notre nature à l'instar de ce que sont déjà les studios électroacoustiques.

Pour autant, en dépit de leurs promesses, les machines sont-elles en mesure de nous offrir des modes d'expression dont nous pourrions nous satisfaire ? Quand bien même l'époque de la fonction artistique de l'art serait définitivement révolue, peut-on faire abstraction du degré d'exigence expressive qu'elle nous a léguée ? Si l'époque préhistorique de la domestication de la *première nature* a bien été celle qui demandait à l'art non d'être perçu mais seulement d'être présent, faut-il, à notre époque de domestication de la *seconde nature*, se contenter de la seule perception tactile et s'abandonner sans plus d'exigence à quelque rituel célébrant la magie du son ?

Formuler ces questions, c'est dire toute la difficulté de notre temps pour qui s'obstine à penser que le destin de la musique n'est ni de ressasser son glorieux passé ni de s'abolir dans l'exhibition rituelle de son matériau par l'entremise de machines chargées d'exorciser notre compulsion à nous y soumettre faute de pouvoir les maîtriser. Ces questions ne sauraient bien sûr trouver de réponse ailleurs que dans la musique elle-même, ce pourquoi il n'y a d'autre solution que de continuer à œuvrer, quoiqu'il en soit, en sachant qu'il en va également de la transmission d'exigences ne reposant désormais plus que sur l'opiniâtreté de qui décide de les soutenir.

A titre donc simplement indicatif et en guise de conclusion provisoire, je me contenterai de trois remarques.

Tout d'abord, je crois que nous sommes engagés, et sans doute pour longtemps encore, dans une période de transition obscure, à la fois hantée par les spectres du passé et aspirée par les chimères d'un avenir qui tarde à tenir ses promesses. Je ne vois personnellement guère comment prendre son parti de ce que propose la technique dans son état actuel. Mais je ne vois guère non plus comment ne pas s'en soucier de quelque manière tant les pratiques de la composition relevant de l'écriture telle qu'héritée du passé ne sont plus guère aptes à produire par elles-mêmes un discours musical susceptible de mobiliser nos oreilles.

Dans le cadre de la musique pure, c'est-à-dire ne faisant appel à aucune autre logique que la sienne propre, je crois donc envisageables les stratégies compositionnelles qui donnent le change à notre époque en laissant prise à une perception tactile, mais en même temps qui perpétuent les exigences d'écriture inhérentes à la constitution d'une pensée musicale et laissent ainsi

ouverte la possibilité d'une écoute ne s'épuisant pas dans la simple audition des effets sonores.

Ensuite, quelque vaste que soit le champ de la musique mixte au sens de l'association des mondes acoustiques et électroacoustiques, je crois que c'est à une mixité généralisée que doit songer la musique. Il ne s'agit pas là d'une fuite en avant dans l'association de la musique avec tout ce qu'elle peut croiser sur son chemin comme le multimédia joint à l'électisme en donne souvent l'impression, mais de la prise au sérieux de la possibilité pour elle de s'enrichir de la confrontation avec d'autres arts, avec d'autres logiques que la sienne, avec ce qui lui est hétérogène.

Enfin, le problème essentiel me paraît être celui de l'expression. C'est un problème dont on ne sait finalement trop comment parler tant il paraît consubstantiel aux moyens techniques dont se dote la musique. Il me semble qu'il devrait pourtant être posé à nouveaux frais, en liaison étroite avec la réalité concrète dont se nourrissent les œuvres mais à distance de ce que l'on a coutume d'appeler les langages musicaux, dont on tient un peu vite pour acquis qu'ils se confondent avec l'expression musicale. Le problème n'est d'ailleurs pas nouveau. Si Schönberg se plaignait déjà de ce que sa musique n'était pas moderne mais mal jouée, c'est bien parce qu'il insiste un problème de lecture, d'interprétation, de liaison des événements sonores, mais plus profondément aussi parce que ce problème de liaison se pose en amont, au moment même de la composition dès lors que les enchaînements ne sont plus surdéterminés par quelque nécessité objective.

Ces questions pourraient être appréhendées sous le chef de celle de phrasé, de façon non restreinte à la musique. Benjamin a caractérisé notre temps comme celui de l'expérience du choc et, en écho à Freud, de la perception traumatisante parce que ne bénéficiant plus du temps nécessaire à l'appréhension d'un réel faisant effraction. Cette effraction du réel ne me paraît pas étrangère aux conditions musicales et de mixité dans lesquelles se présente la question du phrasé en tant qu'elle est intimement liée aux modes d'expression que notre temps nous requiert d'inventer.

1. REFERENCES

- [1] Benjamin, W., « L'Œuvre d'art à l'époque de sa reproduction mécanisée », Œuvres III, Gallimard, Paris, 2000.
- [2] Boulez, P., « Le Système et l'idée », Leçons de musique, Bourgois, Paris, 2005.
- [3] Bonnet, A., « *Écriture* and perception : on *Messagesquisse* by Pierre Boulez », Contemporary music review, London, 1987.
- [4] Bonnet, A., « L'Aura et le terrier », La Pensée de Pierre Boulez à travers ses écrits, Delatour, 2010.
- [5] Boissière, A., « La reproductibilité technique chez Benjamin », revue électronique du Centre d'Etude des Arts Contemporains de l'Université Lille-3, <http://demeter.revue.univ-lille3.fr>,
- [6] Bonnet, A., « De l'idée à l'œuvre. Figures, fonctions, formes, langage dans la *Notation I* pour orchestre de Pierre Boulez », revue Circuit, Montréal, 2007.
- [7] Benjamin, W., « Paris, capitale du XIX^e siècle », Œuvres III, Gallimard, Paris, 2000.

Session 3

Temps réel et performance

L'INSTRUMENT AUGMENTÉ POUR L'IMPROVISATION : LE CAS DE L'ALTO AVEC ROSE AMÈRE

Karim Barkati *
Université Paris VIII – CICM
karim.barkati@mines-paristech.fr

Stéphanie Réthoré
CNSMDP – Alto, Improvisation générative
stephanierethore@gmail.com

RÉSUMÉ

La singularité de l'augmentation instrumentale tient à une propriété remarquable : c'est l'instrumentiste lui-même qui pilote les contrôleurs en plus de jouer de son instrument. Rose amère est un logiciel musical développé en collaboration avec l'altiste et improvisatrice Stéphanie Réthoré, sur un projet d'alto augmenté pour de l'improvisation, avec un dispositif podophonique adapté : un pédalier MIDI multiple, un microphone et un ordinateur. Il répond à plusieurs contraintes musicales adressées par l'improvisation – comme l'équilibre sonore, la variété musicale et la promptitude réactionnelle – grâce à plusieurs astuces de programmation avec les interfaces graphiques et pédestres. Ainsi, Rose amère pourrait être décrit comme un auto-échantillonneur interactif, doté de quatre voix paramétrables en temps réel, permettant à l'instrumentiste de jouer avec des boucles, des impacts, du délai et des effets (transposition et distorsion). Ce logiciel, programmé dans l'environnement Max/MSP, a déjà fait ses preuves lors de plusieurs occasions, a déjà été joué avec plusieurs instruments de musique et se trouve en téléchargement libre sur mon site.

1. Introduction

Rose amère ¹ est un logiciel musical développé en collaboration avec l'altiste et improvisatrice Stéphanie Réthoré, sur un projet d'alto augmenté pour de l'improvisation, avec un pédalier MIDI « multiple » ², un microphone et un ordinateur. Le logiciel « Rose amère », joué par Stéphanie sous la forme d'une extension logicielle de l'alto au sein de notre dispositif podophonique interactif, a fait ses preuves lors de plusieurs occasions ³ :

- lors du Prix d'improvisation générative du CNSMDP, en trio ⁴ alto augmenté / contrebasse / piano, le 13 juin 2008 ;
- lors de l'émission « À l'improviste » d'Anne Montaron au studio 106 de la Maison de la Radio, en trio

* Adresse actuelle : Centre de recherches en informatique, MINES ParisTech.

1. Rose amère est téléchargeable à l'adresse http://karim.barkati.online.fr/Logiciels/Rose_amere/

2. Nous avons utilisé un pédalier Berhinger FCB1010, constitué de dix pédales interruptrices et de deux pédales progressives.

3. Des extraits audio sont écoutables à l'adresse <http://rose.amere.online.fr/>

4. Avec Charlotte Testu à la contrebasse et Laurent Durupt au piano.

- puis en tutti ⁵, enregistrée le 21 juin 2008 ;
- pour un CD « Rose amère » constitué d'improvisations solo enregistrées en mai 2008, réalisé en juillet 2008 ;
- aux Journées de la profession organisées au CNSMDP, en trio, le 11 septembre 2008 ;
- au vernissage de l'exposition de la photographe Geneviève Hofman au Scriptorial d'Avranches « Le Serpent, le Dragon et les Ailes », en solo, le 19 septembre 2008.
- une pièce pour alto et Rose amère est en cours d'écriture, composée par Iván Solano.

Cet article donne un retour sur la réalisation en informatique musicale de ce projet, avec Max/MSP ⁶, essentiellement du point de vue de la programmation.

Un premier résultat, musical, de ce développement logiciel tient à l'équilibre entre la *vocalité* et la *percussivité* du dispositif (au sens de Martin Laliberté). Un second résultat, informatique, tient aux stratégies élaborées pour réduire la complexité du jeu de contrôle pédestre tout en prélevant une certaine complexité du résultat sonore, par l'implémentation de comportements différents selon les pédales (clics simples ou doubles et permutations circulaires sur des ensembles de deux ou trois fonctions).

La première partie, succincte, présente d'abord une forme de cahier des charges à travers trois contraintes qui nous ont paru pertinentes à retenir dans le cadre de l'improvisation générative : l'équilibre sonore, la variété musicale et la promptitude réactionnelle. La seconde partie, plus importante, présente le logiciel lui-même : l'organisation de l'interface graphique, l'enregistrement, les boucles, les « impacts », le délai, les effets, le mixage, l'accrochage et les fichiers de préreglages.

2. Des contraintes pour l'improvisation générative

Assez tôt dans la collaboration, nous avons entrepris d'inventorier ensemble les contraintes majeures qui nous paraissaient liées à l'improvisation générative ⁷, dans son

5. Alto augmenté, clarinette, violoncelle, batterie, deux contrebasses, piano, saxophone.

6. « MSP gives you over 170 Max objects with which to build your own synthesizers, samplers, and effects processors as software instruments that perform audio signal processing. [...] Max enables you to design your own programs for controlling MIDI synthesizers, samplers, and effects processors. » [8, p. 9]

7. « Ceci dit, "générative", pour Alain [Savouret], c'est l'idée de "l'entendre génère le faire", et puis ça lui a été aussi inspiré par le mot

exercice collectif ou individuel, puis de les hiérarchiser ; ce recensement nous a guidé durant le développement de Rose amère⁸. En outre, la spécificité de l'improvisation recentre radicalement la conception d'un tel dispositif musical à partir de l'interprète et non d'une œuvre, d'une idée, ou de « l'utilisateur » imaginaire du génie logiciel⁹.

2.1. L'équilibre sonore

Dans un contexte d'improvisation à plusieurs musiciens, les besoins musicaux peuvent être de plusieurs ordres, mais celui de l'équilibre sonore est primordial. L'amplitude sonore naturelle de l'alto acoustique souffre d'un déficit important par rapport à celle d'autres instruments comme la batterie, le saxophone, ou encore le piano. Il s'agit donc dans un premier temps d'« augmenter » l'instrument au sens de le renforcer, éventuellement de le multiplier, en tout cas de lui donner une plus grande portée sonore, davantage de place dans le collectif.

Par contre, il faut éviter, lors de cette opération, de tomber dans un excès inverse où cette extension prendrait trop d'importance au point de soumettre et de contraindre en permanence les autres improvisateurs. Autrement dit : comment augmenter l'instrument tout en évitant que l'improvisation ne s'organise seulement autour de cette augmentation ?

Ainsi, paradoxalement, cette question de l'équilibre sonore au sein d'un groupe fût une motivation historique importante pour notre augmentation de l'alto mais reste aussi la réserve majeure quant à l'augmentation d'un instrument dans un groupe, car toute amplification a tôt fait de bouleverser l'équilibre entre les instruments, à plus forte raison si certains instruments ne sont pas amplifiés, ou bien si le dispositif emploie de l'informatique en temps réel.

2.2. La variété musicale

La deuxième contrainte relève du domaine directement musical : l'extension logicielle doit pouvoir s'adapter à ce qui se passe *musicalement*, par exemple pouvoir alternativement accompagner, se taire, répondre, proposer, etc. La variété expressive devient ici une nécessité absolue, sous peine de se voir rapidement exclu du jeu musical collectif, ou pire, d'appauvrir les possibilités expressives du groupe lui-même.

On peut légitimement envisager de déployer deux catégories musicales dans ces situations : d'une part les déroulements étales, continus et plutôt arythmiques, et d'autre part les éléments rythmiques, accentués et éventuellement périodiques. Ces deux catégories rejoignent d'ailleurs la formalisation duale entre *voix* et *percussion* proposée par

generative arts en anglais, qui désignait les performances de *dripping* sur des tableaux [...] C'est surtout cette idée que le son, la musique, est générée(e) en temps réel, à l'instant même. » Alexandros Markeas [6].

8. Ce nom a été choisi par Stéphanie, de même que les codes couleur de l'interface graphique, respectant un esprit qui m'est cher : celui du « sur-mesure » (cf. [1, p. 92-93]).

9. « Personne n'a jamais rencontré l'utilisateur des informaticiens, parce que les utilisateurs de la plupart des logiciels sont légions. » [5]

Martin Laliberté dans son analyse *Archétypes et paradoxes des nouveaux instruments* :

Elle [la *percussion*] recherche plutôt les objets sonores ambigus, tels les composites timbre / harmonie, que les objets sonores mélodiques simples. Il s'en suit une nature verticale, harmonique ou massive, contrastant avec la nature horizontale de la *voix*. [4, p.125]

[...] la richesse de telle nouvelle forme instrumentale se mesure dans sa réalisation convaincante des grands courants de fond du développement organologique, par un succès dans la réalisation des tendances *vocales* secrètes d'un instrument de percussion ou par la floraison de l'aspect *percussif* d'un instrument en apparence *vocal*. [4, p.137]

Il y a bien sûr d'innombrables gradations entre ces deux archétypes musicaux et, précisément, la réussite ou l'échec du logiciel dépend en grande partie de sa capacité à proposer un éventail le plus large possible entre ces deux extrêmes, idéaux.

2.3. La promptitude réactionnelle

La troisième contrainte concerne surtout la rapidité, voire l'instantanéité (perceptive) des transitions entre les différents modes de réponse musicaux de l'extension logicielle. C'est en effet souvent sur cette promptitude réactionnelle que repose la dramatisation du jeu collectif et la construction des évolutions du discours musical ensemble et au cours du jeu. Martin Laliberté rappelle la nécessité de prendre en compte les aspects pratiques dans la réalisation et pas seulement les aspects musicaux :

La grande majorité des nouveaux instruments furent des échecs à cause de lacunes fondamentales dans un des aspects essentiels propres aux instruments de musique. Non seulement les nouveaux instruments de musique ont-ils besoin d'une ouverture sonore, d'une richesse des approches et des solutions proposées aux différents problèmes musicaux mais aussi doivent-ils s'incarner de façon probante. [4, p.129]

En particulier, dans le contexte musical de l'improvisation, qui se caractérise par une versatilité potentiellement importante, une extension logicielle insuffisamment réactive constituerait de toute évidence un frein inacceptable.

3. Présentation générale du logiciel

L'ensemble des contraintes musicales vues précédemment – l'équilibre sonore, la variété expressive et la promptitude réactionnelle – laisse encore une grande marge de créativité logicielle, à partir du dispositif informatique et électroacoustique choisi.

3.1. Un auto-échantillonneur interactif

Dans son principe, Rose amère est un auto-échantillonneur ¹⁰ interactif, pour pédalier MIDI, et orienté vers l'improvisation. Ainsi, Rose amère comporte quatre voies paramétrables à gérer par l'instrumentiste lui-même en temps réel : les boucles, les impacts (enveloppes percussives), le délai et les effets (transposition et distorsion).

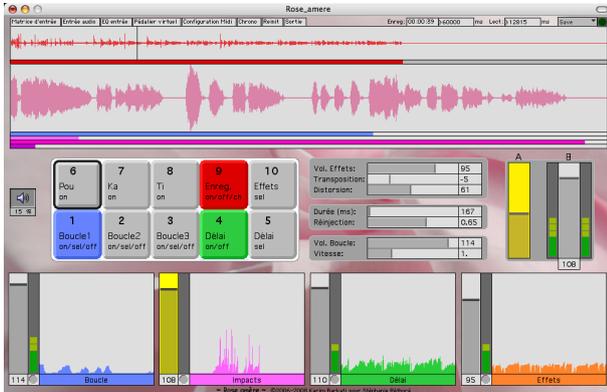


Figure 1. Capture de l'interface visuelle de Rose amère

La conception et le développement se sont déroulés en collaboration avec Stéphanie Réthoré, dans un aller-retour fréquent entre la programmation et les tests musicaux.

Globalement, le résultat sonore relève à la fois de l'augmentation et de l'anamorphose. L'augmentation provient de l'amplification mais aussi de la polyphonie qui démultiplie véritablement l'instrument ; par exemple une boucle s'entend clairement comme un discours musical parallèle à celui de l'instrument acoustique, donnant nettement la sensation d'un dédoublement. L'anamorphose provient quant à elle des opérations numériques programmées dans chacun des modules sonores, par exemple les modifications de vitesse de relecture ou d'enveloppe dynamique ; certaines déformations sont trop importantes pour reconnaître la source, ce qui accentue l'effet de démultiplication.

3.2. Interaction pédestre

La fluidité et l'ergonomie de l'interface homme-machine est cruciale en musique, aussi plusieurs stratégies ont été déployées pour Rose amère, dont la rigidification fonctionnelle de la pédale B, la A-permutation, la A-sélection et l'accrochage (cf. section 11).

3.2.1. Rigidification fonctionnelle de la pédale B

Nous avons décidé, en consultation avec plusieurs instrumentistes, que la pédale progressive de droite du FCB1010,

10 . « [...] le préfixe auto indique que l'instrumentiste a non seulement accès aux fonctions de relecture mais aussi à la fonction d'enregistrement : c'est-à-dire qu'il peut enregistrer lui-même (avec le dispositif) le son qu'il aura produit lui-même (avec l'instrument). » [1]



Figure 2. Le pédalier Midi FCB1010 de Behringer

la pédale B, située à l'extrémité du pédalier donc facilement repérable, servirait systématiquement à contrôler le volume ¹¹ de sortie générale du logiciel.

De fait, l'abstraction d'une interface ou d'un système MIDI comporte des limites pratiques : si tout est toujours permutable, alors aucune association réflexe n'est possible. Or, du point de vue de l'instrumentiste en situation de répétition ou de concert, le contrôle du volume général du dispositif doit être le plus instinctif possible. Conséquemment, cette association volontairement « rigidifiée » de la pédale de droite au même paramètre du volume général, salutaire, a été adoptée et bien accueillie par l'ensemble des collaborateurs – compositeurs comme interprètes.

3.2.2. A-permutation

Musicalement, les pédales progressives possèdent un avantage majeur sur les pédales interruptrices : une *course*, qui permet de varier, de doser, de moduler, de se positionner progressivement à l'intérieur d'un intervalle (de 0 à 127) et pas seulement de déclencher un événement ponctuel. En cela, les pédales progressives sont tout à fait précieuses à l'instrumentiste, à l'improvisateur et à la composition, comme accès à un geste instrumental perceptivement continu. Après qu'on a décidé de réserver systématiquement la pédale B pour le volume général du logiciel, il ne reste donc que la pédale A de disponible pour le contrôle d'un second paramètre, mais on devine que le contrôle progressif de deux paramètres seulement – le volume et un seul autre – peut rapidement se révéler insuffisant à renouveler l'intérêt musical. Comment donner la possibilité de contrôler plus de deux paramètres avec seulement deux pédales progressives, dont l'une est rigidifiée ? Une première réponse consiste à permuter les fonctions qui sont associées à la pédale A : la « A-permutation », qui permet de faire une « A-sélection » de fonction.

3.3. Quatre modules sonores

Quatre modules sonores structurent Rose amère, constituant une polyphonie de quatre voies pour le logiciel. Les deux premiers modules sont fondés sur des relectures déformantes à partir de l'auto-enregistrement, alors que les

11 . Incidemment, il se trouve que cette pédale B envoie par défaut un message MIDI *volume*...

deux derniers relèvent du traitement du signal audio numérique entrant, en temps réel :

1. le module « Boucle » (en bleu) permet de déclencher trois processus macro-temporels au choix, variés sur la vitesse de lecture ;
2. le module « Impacts » (trois impacts en rose, fuchsia et violet) génère trois enveloppes dynamiques de type percussif, à partir du son enregistré en direct, avec des paramètres semi-aléatoires bornés ;
3. le module « Délai » (en vert) retarde le son entrant selon deux paramètres contrôlables : la durée du délai et le taux de réinjection ;
4. le module « Effets » (en orange) applique deux effets en série : la transposition du son et la distorsion numérique.

Chacun de ces quatre modules est relié à la pédale d'expression gauche (A), en respect du principe de permutation fonctionnelle, ce qui permet au musicien de réaliser lui-même le mixage des modules, de paramétrer chaque fonction et de doser l'équilibre avec son instrument. Quant à la pédale d'expression droite (B), elle permet simplement de gérer le volume global sortant du logiciel, en respect de sa « rigidification » fonctionnelle, ce qui revient à gérer la dynamique générale de l'instrument augmenté.

4. L'organisation de l'interface graphique

L'interface graphique apporte ici à l'instrumentiste un retour précieux sur ce qu'enregistre ou ce que joue Rose amère, ce que Max/MSP permet intrinsèquement ¹².

4.1. Présentation synthétique : quatre étages

L'interface graphique de Rose amère s'organise en trois étages bien visibles, pour le jeu en direct, surmontés d'un étage plus discret, pour la configuration des entrées. Ces quatre étages sont synthétisés ¹³ dans le 1 et détaillés dans la suite de cette section.

4.2. Premier étage : configuration des entrées



Figure 3. Le premier étage visuel de Rose amère

Le premier étage reste le plus discret visuellement, parce qu'il est dédié essentiellement à la configuration et non à la performance, avec 8 petits boutons textuels qui ouvrent chacun une fenêtre de configuration :

12. « Un patch est en effet à la fois la description d'un algorithme de traitement, mais également la réalisation d'une interface graphique de contrôle de l'algorithme. » [3, p. 88]

13. De haut en bas.

Étages	Représentations	Fonctions
1	8 boutons textuels et 3 boîtes numériques	Configuration générale des entrées et informations numériques mineures
2	2 formes d'ondes et 5 barres de progression	Visualisation de l'enregistrement et de la lecture
3	12 boutons et 9 curseurs	Visualisation des fonctions de performance à partir du pédalier
4	4 curseurs, 4 VU-mètres, 4 historiques déroulants et 4 boutons textuels	Visualisation de l'activité sonore de chaque voie

Table 1. Organisation de l'interface graphique de Rose amère

- « Matrice d'entrée », qui permet de configurer le nombre d'entrées externes et de les prémixer ;
- « Entrée audio », qui permet de sélectionner l'entrée effective parmi Silence, Carte son et Fichier son, et de régler son niveau ;
- « EQ entrée », qui ouvre un éditeur graphique de 5 filtres fréquentiels ;
- « Pédalier virtuel », qui affiche le pédalier virtuel, soit pour la visualisation, soit pour l'interaction directe (à la souris ou au clavier alphanumérique ¹⁴, cf. 2) ;
- « Configuration Midi », qui permet de rappeler et d'enregistrer ses propres réglages de correspondance MIDI avec les trois entrées logiques Déclencheurs, Pédale A et Pédale B, ainsi que de visualiser les différents messages MIDI entrants.
- « Chrono », qui permet de visualiser l'écoulement du temps depuis le premier clic ¹⁵ dans sa fenêtre ;
- « Reinit », qui réinitialise l'ensemble du patch [Rose_amere] ;
- « Sorties », qui ouvre le sous-patch de routage des sorties audio.

a	z	e	r	t	y	→	1	2	3	4	5	UP
q	s	d	f	g	h		6	7	8	9	10	DN

Table 2. Les raccourcis-clavier du pédalier virtuel

4.3. Deuxième étage : visualisation de l'enregistrement et de la lecture

Le deuxième étage montre clairement deux formes d'ondes occupant toute la largeur :

- la forme d'onde supérieure, plus petite et en rouge, représente le tampon d'enregistrement ;
- la forme d'onde inférieure, plus grande et en rose, représente le tampon de lecture.

14. Les raccourcis-clavier ont été programmés par défaut pour un clavier azerty et pour la main gauche, la main droite pouvant ainsi s'occuper des curseurs avec la souris.

15. Un double clic réinitialise le chronomètre.

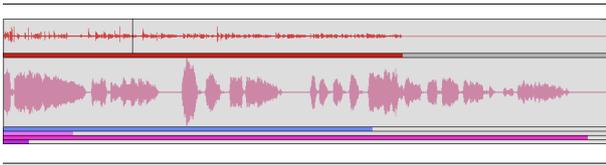


Figure 4. Le deuxième étage visual de Rose amère

En cours d'enregistrement, la barre de progression correspondante, en rouge, avance de gauche à droite. Le principe reste le même pour la lecture, sauf qu'il y a quatre barres de progression différentes, soit de haut en bas :

- le pointeur de lecture du module « Boucles », en bleu ;
- le pointeur de lecture de l'impact « Ti » (le plus aigu), en rose ;
- le pointeur de lecture de l'impact « Ka » (médium), en fuchsia ;
- le pointeur de lecture de l'impact « Pou » (le plus grave), en violet.

4.4. Troisième étage : visualisation des fonctions de performance



Figure 5. Le troisième étage visual de Rose amère

Le troisième étage s'organise horizontalement en quatre parties, soit de gauche à droite :

- 2 boutons pour les statuts audio (marche/arrêt et DSP Status) ;
- 10 boutons correspondant aux 10 pédales interruptrices ;
- 7 curseurs horizontaux pour le paramétrage interactif ;
- 2 curseurs verticaux correspondant aux 2 pédales progressives A et B.

Cette disposition horizontale reprend donc la métaphore visuelle du pédalier, en intercalant les 7 curseurs de paramétrage entre les 10 boutons¹⁶ et les 2 curseurs verticaux.

La restriction à une unique banque implique en contrepartie une certaine complexité pour l'accès aux différentes fonctions : une diversité de mécanismes comme des doubles clics et des permutations circulaires, ces comportements différant d'un bouton à l'autre (cf. sections suivantes).

16 . Les 2 boutons UP et DOWN n'apparaissent pas ici, car Rose amère n'utilise que la première banque, c'est-à-dire les messages MIDI de 0 à 9, afin d'optimiser la rapidité de l'utilisation du dispositif, en supprimant la navigation dans les 10 banques du pédalier FCB1010 (10 à 19, 20 à 29, etc).

4.5. Quatrième étage : visualisation de l'activité polyphonique

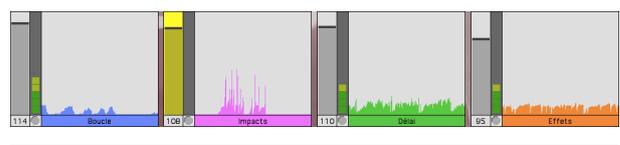


Figure 6. Le quatrième étage visual de Rose amère

Le quatrième étage s'organise horizontalement en quatre modules identiques, regroupant chacun :

- un curseur vertical pour le volume (de 0 à 127), assorti d'un VU-mètre ;
- un historique déroulant du signal, pour voir une trace de 10 secondes¹⁷ ;
- un bouton textuel qui ouvre la fenêtre de pré-réglage du module idoine¹⁸, 15.

Ainsi, cet étage se révèle tout à fait précieux pour le mixage entre les quatre voies en situation de direct.

5. L'enregistrement

L'enregistrement demande de maîtriser différents types de déclenchement avec la pédale 9, selon les transitions d'états représentées 7 :

- le simple clic (« 1c ») qui déclenche l'enregistrement (états « enregistrement » e_0 et e_1) ;
- le simple clic (« 1c ») qui arrête l'enregistrement (états « prêt » p_0 et p_1), le tampon d'enregistrement devant prêt à charger ou à réenregistrer ;
- le double clic (« 2c ») qui stoppe d'abord les éventuelles lectures en cours (boucle et impacts) et qui charge ensuite l'enregistrement dans le tampon de lecture (état « chargé » c_1).

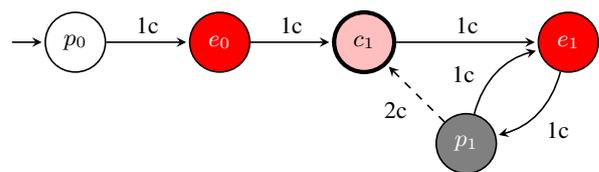


Figure 7. Diagramme de performance de l'enregistrement dans Rose amère (bouton 9)

Ce mécanisme qui utilise deux tampons différents – un pour l'enregistrement et un pour la lecture – permet de ne pas interrompre les boucles en cours de lecture pendant les enregistrements, de façon à rendre l'opération d'enregistrement plus discrète à l'écoute, sans rupture. Ainsi, l'arrêt de l'enregistrement (un simple clic) est dissocié du chargement effectif dans le tampon de lecture (un double clic) ; sauf au départ, où l'arrêt de l'enregistrement (un

17 . La trace dure exactement 195 pixels × 50 ms = 9 750 ms.

18 . Ces fenêtres ne sont pas montrées ici, par souci de concision, à part la fenêtre du module « Effets ».

simple clic) suffit à charger aussitôt le tampon de lecture (pas de double clic), puisque ce dernier est nécessairement vide. Cette dissociation permet en outre de réenregistrer sans avoir chargé (retour p_1 vers e_1); ceci permet donc de charger uniquement ce que l'on souhaite vraiment et seulement au moment voulu (chargement p_1 vers c_1).

Par ailleurs, la fenêtre temporelle de détection du type de clic (simple ou double) a été fixée à 500 ms pour tous les boutons qui font appel au double clic, les autres boutons ne subissant pas ce retard. Cette durée nous a semblé raisonnable pour réaliser un double clic avec les pieds; en revanche, les simples clics d'un tel bouton subissent aussi un retard équivalent de 500 ms, à cause de la fenêtre temporelle de détection.

Enfin, la durée maximum d'un enregistrement a été fixée à 1 mn : si cette durée est atteinte, alors l'enregistrement s'arrête automatiquement, sinon, si l'enregistrement est arrêté avant cette durée, alors le tampon est redimensionné automatiquement jusqu'à ce point d'arrêt (mais le prochain enregistrement se verra à nouveau allouer une minute entière).

6. Les trois boucles

6.1. Présentation générale des trois boucles

Une seule des trois boucles peut être lue à la fois, car elles partagent le même moteur de lecture – techniquement le même objet `Max/MSPgroove~`. Cependant, elles procèdent chacune d'un fonctionnement différent :

- la boucle 1 déclenche une lecture oscillante lente et grave ;
- la boucle 2 déclenche une lecture aléatoire « agitée » dans les médiums/aigus ;
- la boucle 3 déclenche une lecture dont la vitesse est contrôlable avec la pédale A.

name	#1	#2	#3	#4	#5	#6
loop-bp-VitesseLecture						
Vitesse_lecture_min	0.15	1.	5.	-3.	-7.	-0.98
Vitesse_lecture_max	0.16	4.	8.	-0.91	-2.	-1.02
Vitesse_lecture_init	0.156214	2.978718	5.383405	-0.927099	-2.379416	-1.02
loop-bp-DureeOscillations						
Duree_oscillation_min	10.	0.36	1.	0.36	0.36	0.8
Duree_oscillation_max	40.	2.	2.	1.	1.47	1.5
Duree_oscillation_init	29.266001	1.086	1.336	0.36	1.165	0.8

Figure 8. Préréglages des boucles de Rose amère

La 8 présente le fichier `loop-ps` qui contient les pré-réglages retenus pour la boucle 1 (le pré-réglage n° 1) et la boucle 2 (tous les autres pré-réglages, ici n°s 2 à 6); la boucle 3 n'utilise pas ces pré-réglages puisque la vitesse de lecture est confiée à la pédale A (dans un intervalle compris entre 0,25 et 3).

Un mécanisme supplémentaire intervient pour l'interaction avec les boucles : la « A-sélection », qui correspond à la permutation fonctionnelle de la pédale A. Ce mécanisme, qui sera repris pour le délai et les effets, permet d'affecter à la pédale A le premier paramètre du module *sélectionné* (le volume le plus souvent). Afin de visualiser cette sélection pedestre à l'écran, d'une part, un épais

contour noir est dessiné sur le dernier bouton sélectionné (ainsi que sur les diagrammes de performance des figures 9 et 10) et d'autre part, le curseur concerné et celui de la pédale A s'allument en jaune à l'écran.

6.2. La boucle 1

La boucle 1 utilise de l'aléatoire borné dans les graves, en appliquant le pré-réglage n° 1 (cf. 8). Ce dernier correspond à une vitesse de lecture comprise entre 0,15 et 0,16, soit une lecture ralentie plus de six fois¹⁹ qui sonne plus grave de deux octaves et une sixte mineure ou majeure²⁰, modulée par un oscillateur à très basse fréquence, dont la période varie aléatoirement entre 10 s et 40 s²¹.

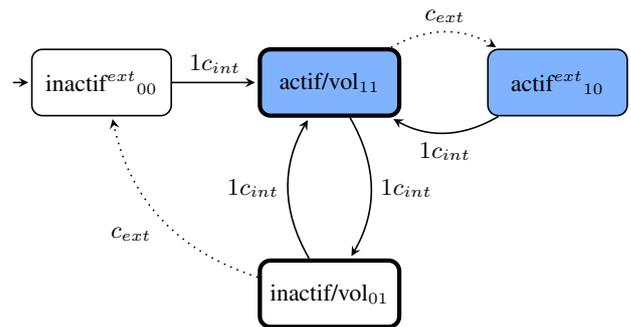


Figure 9. Diagramme de performance des boucles 1 et 2 dans Rose amère (boutons 1 et 2)

À cause du mécanisme de A-sélection, il faut distinguer 4 états possibles²² et 2 types de clics pour la boucle 1 (cf. 9) :

- le clic interne « $1C_{int}$ », correspondant à un appui sur la pédale 1, qui permet alternativement d'activer et de désactiver la lecture ;
- le clic externe « C_{ext} », correspondant à un appui sur toute autre pédale que la pédale 1, qui n'interrompt pas nécessairement la lecture.

Ainsi, la lecture de la boucle 1 n'est pas non plus interrompue quand on « arrive dessus » (*i. e.* dans le cas où la boucle est active sans être sélectionnée et qu'on effectue un clic interne) : dans ce cas, ce clic se contente de A-sélectionner la boucle 1, de façon à pouvoir ajuster son volume sans rupture même en venant d'ailleurs. Autrement dit, un *premier* clic interne (*i. e.* depuis une A-sélection externe) n'interrompt jamais la fonction ; c'est un clic « d'entrée ». Il faut donc généralement appuyer deux fois de suite pour pouvoir désactiver une fonction quand on vient d'une autre pédale. En outre, un premier clic interne implique toujours une A-sélection interne, que

19. $1 \div 0,15 = 6,25$ et $1 \div 0,16 \approx 6,25$.

20. Une vitesse de lecture de 0,15 correspond environ à 33 demi-tons ($2^{\frac{33}{12}} \approx 0,15$) et 0,16 à 32 demi-tons ($2^{\frac{32}{12}} \approx 0,16$); soit à une transposition comprise entre deux octaves inférieures et une sixte mineure (32) ou une sixte majeure (33).

21. Soit une fréquence comprise entre 0,1 Hz et 0,025 Hz.

22.

	processus inactif	processus actif
non A-sélection	00	10
A-sélection	01	11

la fonction soit active ou non (sauf pour l'enregistrement, soit le bouton 9).

6.3. La boucle 2

La boucle 2 applique un préréglage tiré au hasard à partir du deuxième préréglage dans le fichier idoine (donc ici du n° 2 au n° 6, cf. 8). Chaque tirage aléatoire est renouvelé lorsqu'un cycle de l'oscillateur basse fréquence se termine, soit, avec ce fichier de réglage, dans un laps de temps compris entre 0,36 s et 2 s, donc assez fréquemment, ce qui donne son côté « agité ». De plus, les vitesses de lectures sont beaucoup plus variées, éventuellement à l'envers et beaucoup plus aiguës : jusqu'à 8 fois plus vite (préréglage n° 3) et à peine moins que la vitesse normale pour le minimum (-0, 91 pour le préréglage n° 4).

En dehors de ce tirage aléatoire récurrent, le fonctionnement de la boucle 2 reste identique à celui de la boucle 1 (cf. 9).

6.4. La boucle 3

La boucle 3 permet à l'instrumentiste de piloter lui-même la vitesse de lecture à partir de la pédale A, en plus du volume, sur un intervalle [0, 25; 3, 0]. Elle se distingue des deux premières boucles par son absence de préréglage dans le fichier et par la permutation fonctionnelle de la pédale A sur deux paramètres (une « 2A-permutation ») : le volume et la vitesse de lecture.

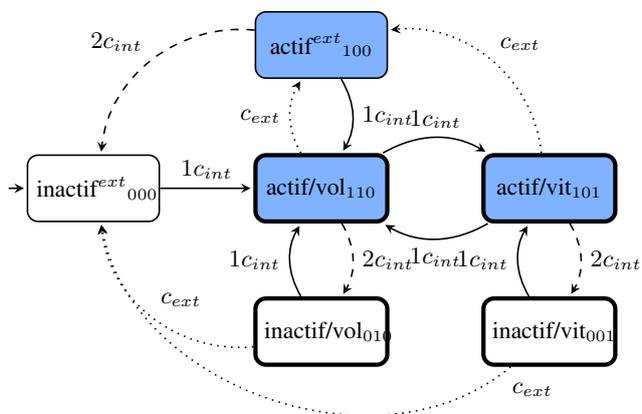


Figure 10. Diagramme de performance de la boucle 3 dans Rose amère (bouton 3)

Le diagramme de performance de la boucle 3 comporte alors 6 six états différents ²³ et 3 types de clics :

- le clic interne « $1C_{int}$ », correspondant à un appui sur la pédale 3, qui permet cette fois-ci de permuter alternativement la A-sélection entre le volume et la vi-

23. Deux cas ne sont pas autorisés parmi les 8 cas théoriques (²³) formulables à partir des trois conditions binaires – la lecture, la A-sélection du volume et la A-sélection de la vitesse. De fait, puisque la sélection du volume et celle de la vitesse se font par permutation circulaire, ces deux paramètres s'excluent mutuellement, éliminant les deux cas « 011 » et « 111 » du tableau suivant :

- le clic interne « $2C_{int}$ », correspondant à un double appui sur la pédale 3 (sous 500 ms), qui permet de désactiver la lecture ;
- le clic externe « C_{ext} », correspondant à un appui sur toute autre pédale que la pédale 3, qui permet de passer sur une autre fonction (externe par rapport à celle-ci) mais qui n'interrompt pas la lecture si elle est en cours.

Par rapport aux boucles 1 et 2, il faut donc apprendre un nouveau comportement du simple clic interne – la A-permutation –, anticiper le retard de détection de 500 ms, et intégrer le double clic pour stopper la lecture. De plus, visuellement, il y a deux curseurs à guetter lors des paramétrages en direct : « Vol. boucle » (de 0 à 127) et « Vitesse » (de 0,25 à 3), en bas de la zone des sept curseurs horizontaux.

7. Les trois impacts

7.1. Présentation des impacts

La création des « impacts » répond à une certaine frustration due à l'absence de sons percussifs et potentiellement puissants à l'alto – le volume des *pizzicati* naturels restant très modéré –, et réalise en quelque sorte ce que Martin Laliberté appelle « la floraison de l'aspect *percussif* d'un instrument en apparence *vocal* » (cf. 2.2).

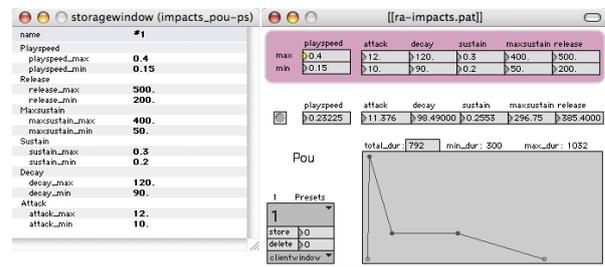


Figure 11. Édition des impacts dans Rose amère

Techniquement, les impacts appliquent une enveloppe dynamique de type ADSR ²⁴. Même si Curtis Roads notait déjà en 1998 l'aspect anachronique de ces enveloppes ²⁵, elles peuvent trouver une application pertinente dans un processus aléatoire sur de l'enregistrement en direct. En

lecture	Asel-vol	Asel-vit	
0	0	0	inactif
0	0	1	inactif/vit
0	1	0	inactif/vol
0	1	1	(n'existe pas)
1	0	0	actif
1	0	1	actif/vit
1	1	0	actif/vol
1	1	1	(n'existe pas)

24. Attack, decay, sustain, release.

25. « [...] pour la spécification d'une enveloppe musicale, une limite de quatre étapes est anachronique. Le façonnage d'une amplitude est une opération délicate, et c'est ainsi que des éditeurs d'enveloppe plus précis permettent au musicien de tracer des courbes arbitraires. » [7, p. 64]

effet, il ne s'agit pas de synthèse : le son provient du tampon de lecture, donc préenregistré en direct, et possède déjà sa propre richesse sonore. D'abord, ce son est renormalisé au maximum pour chaque impact ; ensuite, une vitesse de lecture est tirée aléatoirement (entre les bornes minimum et maximum spécifiées dans le fichier de pré-réglage) et déclenche la lecture ; enfin, une enveloppe ADSR tirée aléatoirement de la même façon module l'amplitude au cours du temps.

7.2. Fonctionnement des impacts

Ainsi, chaque type d'impact – Pou, Ka et Ti – possède son propre fichier de pré-réglage (dont seul le premier pré-réglage est pris en compte par le logiciel, cf. 11 pour Pou et 12 pour Ka et Ti), pour un résultat globalement différent entre les types d'impacts qui deviennent souvent identifiables :

- Pou sonne plutôt grave et un peu mou ;
- Ka sonne plutôt médium et plus franc, avec une lecture inverse ;
- Ti sonne plutôt aigu et parfois fluet.

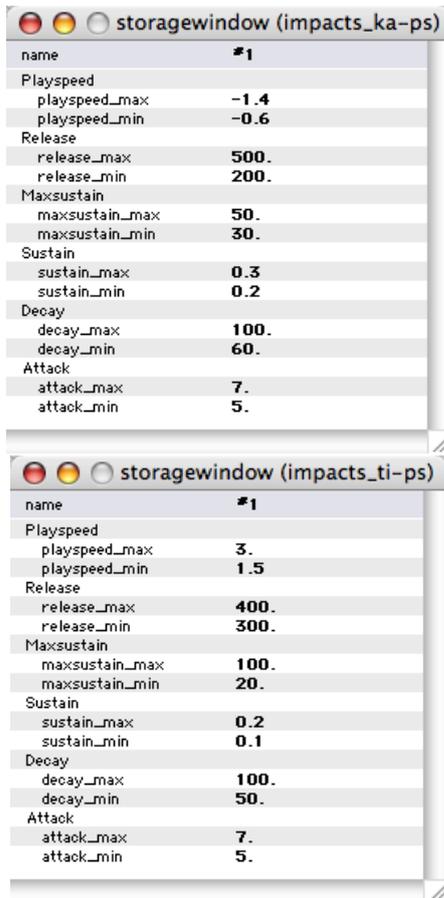


Figure 12. Préréglages des impacts Ka et Ti dans Rose amère

Au niveau du jeu musical, deux déclenchements successifs d'un même impact donnent un résultat différent pour le deuxième impact afin de contrer un problème de monotonie : si le deuxième déclenchement a lieu après la

fin de l'enveloppe, alors le deuxième impact ressemble au premier impact, car le pointeur de lecture revient à zéro ; sinon, lorsque le deuxième déclenchement a lieu avant la fin de l'enveloppe, le deuxième impact sonne différemment du premier impact car le pointeur de lecture continue sa progression dans le tampon de lecture (de façon circulaire si la fin du tampon est atteinte). Ainsi, les déclenchements rapprochés donnent une plus grande variété spectrale.

8. Le délai

8.1. Présentation du délai

À la différence des boucles et des impacts, le délai n'utilise pas le tampon de lecture, car il s'applique directement sur l'entrée audio, par exemple sur la captation de l'alto en direct. En revanche, le délai comporte ici trois paramètres – le volume, la durée et la réinjection –, au lieu de deux pour la boucle 3 et un seul pour les boucles 1 et 2.

Paramètre	Minimum	Maximum	Évolution
Volume	0	127	exponentielle
Durée (ms)	40,0	1000,0	exponentielle
Réinjection	0,0	1,0	linéaire

Table 3. Les trois paramètres du délai de Rose amère

Les pré-réglages sont enregistrés dans le fichier *fib-ps* présenté 13, où seul le premier pré-réglage est pris en compte par Rose amère. Le 3 synthétise l'ensemble des informations pratiques sur les trois paramètres du délai, y compris sur ceux qui n'appartiennent pas au fichier de pré-réglage.

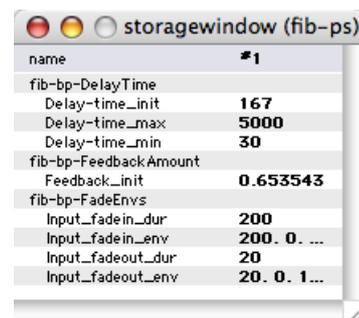


Figure 13. Préréglages du délai dans Rose amère

8.2. Fonctionnement du délai

On a vu avec la boucle 3 que l'augmentation du nombre de paramètres combinée à la gestion de l'activation entraîne une complexité croissante, pour l'instrumentiste comme pour le logiciel. Ainsi, pour pouvoir gérer le délai correctement, Rose amère utilise deux boutons au lieu d'un seul, en reprenant toutefois les principes de fonctionnement des boutons des boucles :

- le bouton 4 reprend exactement le double mécanisme d'activation/désactivation et de A-sélection du volume présenté dans le diagramme de performance des boucles 1 et 2 (cf. 9);
- le bouton 5 reprend le principe de A-permutation (entre deux paramètres) exposé pour la boucle 3, mais de façon bien plus simple (cf. 14), puisque la gestion de l'activation est déportée sur un autre bouton (le bouton 4 en l'occurrence).

Ainsi, il existe une véritable indépendance entre le bouton 4 (la commutation du délai et la A-sélection du volume) et le bouton 5 (la A-permutation entre la durée et la réinjection); en particulier, si le délai est inactif, le bouton 5 fonctionne quand même et permet de préparamétrer silencieusement la durée ou la réinjection.

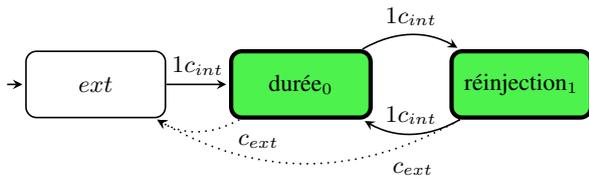


Figure 14. Diagramme de performance de la A-permutation des paramètres du délai dans Rose amère (bouton 5)

Il y a d'autres avantages à cette répartition sur deux boutons. D'une part, le double clic n'étant pas nécessaire, les 500 ms de retard dues à l'analyse n'ont pas lieu. D'autre part, afin de faire économiser une mémorisation supplémentaire à l'instrumentiste, un premier clic interne A-sélectionne systématiquement le paramètre de la durée (jamais celui de la réinjection), quelle que soit la dernière A-sélection quittée (la durée ou la réinjection).

Le résultat sonore du délai dépend d'abord de la durée et celle-ci peut varier dans des proportions considérables : depuis 30 ms – pour des effets de filtre en peigne –, jusqu'à 5 s – pour une répétition retardée au-delà de l'écho traditionnel. Rose amère ne propose que 126 valeurs intermédiaires, un nombre finalement restreint en regard de la durée importante à parcourir, mais cette quantification garde ici une pertinence musicale grâce à l'échelle exponentielle²⁶ utilisée pour mettre en correspondance les valeurs MIDI de la pédale A avec le contrôle de la durée.

De fait, une échelle linéaire aurait progressé régulièrement par pas de 40 ms, alors qu'il est nettement plus intéressant musicalement de progresser finement pour les durées courtes que pour les durées longues : une palette large dans les durées courtes donne une grande variété d'effets psychoacoustiques jusqu'à 100 ms environ – *phasing*, *flanging*, réverbération, écho – pour devenir ensuite un effet essentiellement rythmique où la précision devient moins cruciale.

Le paramètre de réinjection se contrôle quant à lui de façon linéaire, entre 0 et 1.; le fait que le maximum ne

26. Il s'agit de l'objet `scale` paramétré comme suit : `scale 0 126 50. 5000. 1.06`, dont le cinquième argument indique un mode exponentiel lorsqu'il est supérieur à 1.

dépasse pas 1. permet de poser une première limite aux risques de larsen, même si ce risque n'est jamais nul pour du délai, surtout avec des durées courtes.

9. Les effets

9.1. Présentation des effets

Le quatrième et dernier module sonore, nommé « effets », offre trois paramètres : le volume du module, une transposition et une distorsion, en série sur l'entrée audio.

La transposition progresse par demi-ton entre -12 et $+12$, en colorant le son à cause de l'algorithme de transposition²⁷; le son subit un léger retard et les attaques sont un peu gommées; le résultat sonne donc assez doux, comme « poli » par le traitement.

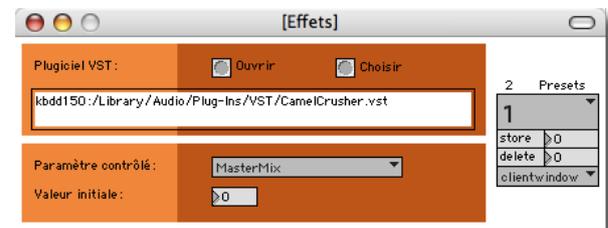


Figure 15. Choix du plugiciel à contrôler avec Rose amère

Pour la distorsion, Rose amère utilise par défaut le plugiciel *CamelCrusher*²⁸ et contrôle le mixage de l'effet par rapport au signal entrant (le paramètre *MasterMix*). Cependant, Rose amère accepte n'importe quel plugiciel VST²⁹ et permet de sauvegarder son chemin, le paramètre à contrôler, ainsi que la valeur initiale de ce paramètre (cf. 15).

Lorsque le volume est supérieur à 0 et que les autres paramètres sont à 0, le circuit des effets devient simplement un amplificateur de l'entrée audio.

9.2. Fonctionnement des effets

D'abord, le bouton des effets n'a pas de statut pour l'activation, car il est activé en permanence; il suffit que le volume soit à zéro pour ne pas l'entendre, et le traitement de la transposition se désactive automatiquement si elle est réglée sur zéro. Ainsi, grâce à cette économie du statut d'activation, un seul bouton suffit.

Ensuite, ce bouton unique fonctionne sur le principe de la A-permutation circulaire, comme les boutons 3 et 5 mais avec trois paramètres : le volume, la transposition et la distorsion. La A-sélection d'entrée se fait toujours par le volume et dans cet ordre.

Enfin, il y a une astuce pour remettre la transposition et la distorsion à zéro en même temps : un double clic.

27. Cet algorithme utilise une combinaison des objets `pfft~` et `gizmo~`, sur une fenêtre d'analyse de 4096 points avec un facteur de recouvrement égal à 8.

28. Par défaut, les plugiciels se trouvent dans le répertoire `/Library/Audio/Plug-Ins/VST/` sous Mac OS.

29. *Virtual Studio Technology*.

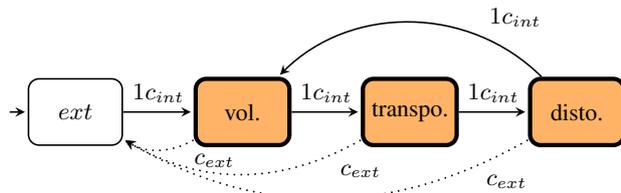


Figure 16. Diagramme de performance de la A-permutation des paramètres des effets dans Rose amère (bouton 10)

Cela évite de devoir répéter l'opération de remise à zéro avec la pédale A pour chacun des deux paramètres, d'autant que cette opération reste délicate pour la transposition (le zéro se trouvant en milieu de course, entre -12 et $+12$). Comme la distorsion est susceptible d'entraîner des larsens, la priorité a été donnée à la vitesse de réaction : j'ai supprimé le fenêtrage temporel pour l'implémentation de ce double clic rapide, afin d'éviter le retard de 500 ms de l'analyse des clics ; ce double clic rapide continue la A-permutation en parallèle des deux simples clics équivalents qui déplacent la A-sélection, mais c'est un compromis que nous avons jugé acceptable.

10. Le mixage

Dans la polyphonie à quatre voies que propose Rose amère, le mixage apparaît rapidement comme une fonctionnalité indispensable : pour que le résultat sonore global ait un sens, il faut que les différents plans puissent être ajustés en direct et de la façon la plus commode possible.

Cependant, sachant que la pédale B est strictement réservée au volume global, il ne reste qu'une seule pédale pour mixer les quatre voies. Le mixage se fait donc par A-sélection :

Module	Boutons	A-permutation	Norm.	Sortie
Boucle	1, 2, 3	3 (2 params)	oui	dac 3
Impacts	6, 7, 8	–	oui	dac 4
Délai	4	5 (2 params)	non	dac 5
Effets	10	10 (3 params)	non	dac 6

Table 4. Récapitulatif des accès au mixage de Rose amère

Le niveau (pré-facteur) des boucles et des impacts est approximativement prévisible, grâce à la normalisation automatique du tampon de lecture lors de son chargement (objet `normalize~`). En revanche, le niveau du délai et des effets n'est pas prévisible, car il dépend de l'entrée audio. Tous les niveaux sont à 0 au démarrage, sauf les impacts qui sont à 70, déjà disponibles.

L'accès au volume d'un module (*i. e.* l'accès à la A-sélection du volume) se fait toujours dès le *premier* clic entrant sur le bouton correspondant (cf. 4), ce qui répond à l'aspect prioritaire du mixage en situation de direct. Excepté donc pour les boutons qui utilisent la A-permutation entre plusieurs paramètres (les boutons 3, 5 et 10), cet ac-

cès direct au volume est maintenu le plus souvent possible. De cette façon, le volume peut être ajusté même si le module n'est pas actif, par exemple pour préparer un fondu entrant sur une boucle à venir.

Pour pouvoir s'adapter aux différents systèmes de diffusions et aux différentes cartes audios, Rose amère possède une sortie stéréophonique prémixée, complétée par six sorties séparées. Ainsi, la sortie audio générale, mono-phonique pour l'instant, est dupliquée sur les sorties 1 et 2, tandis que les quatre modules Boucle, Impacts, Délai et Effets sortent respectivement en 3, 4, 5 et 6 ; enfin, l'entrée pré-traitement sort en 7 et l'entrée post-traitement (après égalisation et volume) sort en 8.

11. L'accrochage

Nous avons contourné le manque de pédales progressives de notre dispositif par la A-permutation, mais permuter le paramètre à contrôler pose un nouveau problème de rupture des valeurs, que nous avons surmonté grâce à un mécanisme d'accrochage de la valeur du nouveau paramètre. En effet, lorsqu'on change le paramètre associé à la pédale A, il y a de fortes chances pour que la valeur actuelle du nouveau paramètre soit en décalage avec la position effective de la pédale progressive. Ce décalage peut se révéler rédhibitoire à l'usage, par exemple sur le volume qui prendrait une valeur tout à coup très éloignée, au moindre mouvement de la pédale si elle se trouve à l'opposé...

L'accrochage empêche précisément ces sauts brutaux de se produire, en bloquant la transmission tant qu'une valeur envoyée par la pédale A n'est pas suffisamment proche de la valeur actuelle du paramètre nouvellement sélectionné. La zone d'accrochage a été fixée par défaut entre -10 et $+10$ autour de la valeur actuelle du paramètre, sur l'intervalle des 128 valeurs possibles, à travers l'abstraction `kb-accrocheur127` développée pour Rose amère.

La notification visuelle du bon accrochage à l'utilisateur se fait par un basculement de la couleur de deux curseurs : le curseur concerné et le curseur de la pédale A, qui passent tous deux du jaune – la couleur de première A-sélection, donc en instance d'accrochage – à la couleur du module concerné – soit bleu, rose, vert ou orange. Tant que l'accrochage n'a pas eu lieu, le curseur du module concerné reste immobile et jaune, indiquant passivement la valeur que la pédale A doit approcher pour le débloquent.

12. Les fichiers de pré-réglages

Rose amère compte en tout 8 fichiers de pré-réglages³⁰, rassemblés et commentés dans le tableau 5.

L'édition de ces fichiers de pré-réglages est unifiée par le `bpatcher [kb-Storage-mng]` qui permet d'accéder aux principaux messages adressables à l'objet `pattrstorage` : `store`, `delete`, `clientwindow`, `storagewin-`

³⁰. Le suffixe « `-ps` » fait référence à l'objet `pattrstorage` qui gère ces fichiers.

Entrée	entree_matrice-ps.xml filtergraph5-ps.xml	détermine le nombre d'entrées et leur routage stocke les coefficients des 5 filtres d'égalisation
Boucles	loop-ps.xml	paramètre les boucles 1 et 2 pour la vitesse de lecture et le LFO
Impacts	impacts_pou-ps.xml impacts_ka-ps.xml impacts_ti-ps.xml	contiennent les bornes du tirage aléatoire pour la vitesse de lecture et les segments de l'enveloppe ADSR
Délai	fib-ps.xml	spécifie les courses de la durée et du taux de réinjection
Effets	plugiciel-ps.xml	mémorise le plugiciel à utiliser et le paramètre à contrôler

Table 5. Les fichiers des préréglages de Rose amère

dow, read, readagain, write, writeagain, getslotlist, renumber et clear.



Figure 17. Capture visuelle du bpatcher kb-Storage-mng

Remarquons que la prolifération des préréglages dans des logiciels comme Rose amère, manifestement destinés à la performance en direct – soit au temps réel musical³¹ –, tend à confirmer l'idée que pour fonctionner efficacement en temps réel (musical ou pratique), un logiciel tire avantage de l'utilisation d'éléments temps différé (pratiques ou techniques).

13. Conclusions et perspectives

Après avoir relevé l'épreuve du concert à plusieurs reprises, Rose amère nous semble répondre de façon satisfaisante aux contraintes et aux souhaits formulés par rapport à la situation du direct, à la pratique de l'improvisation et à l'augmentation de l'alto. En effet, ce logiciel s'insère dans un dispositif podophonique interactif en proposant une certaine variété sonore – à travers les quatre modules et leurs potentiels *vocal* et *percussif* – et en opérant les compromis nécessaires à la fluidité de l'interaction pédestre, avec la restriction à une banque unique (10 boutons seulement). Cette restriction à dix pédales interruptrices seulement engendre des conséquences pratiques : le recours à plusieurs mécanismes locaux, comme la A-sélection et la A-permutation, selon différents schémas de transition d'états, parfois complexes.

³¹. Cf [2] concernant les champs sémantiques *musical*, *pratique* et *technique* appliqués aux notions de temps réel et de temps différé.

Rose amère devient ainsi une extension logicielle qui permet à l'instrumentiste de jouer de l'ordinateur en même temps qu'il joue de son instrument et qu'il s'auto-enregistre. Néanmoins, ce double jeu – acoustique et électronique – demande un temps d'apprentissage conséquent et surtout une virtuosité supplémentaire importante : au niveau du jeu pédestre bien sûr, mais aussi pour l'effort d'écoute multiplié par les quatre modules, ainsi que pour la visualisation des retours graphiques. Quant à l'improvisation collective, elle exige une disponibilité aux autres qui rend évidemment l'exercice encore un peu plus virtuose...

Enfin, une pièce est actuellement en cours d'écriture, composée par Iván Solano pour alto et dispositif podophonique dérivé de Rose amère, à l'attention de Stéphanie Réthoré. Cette pièce sera créée au CNSMDP en juin 2010.

14. Remerciements

Nous remercions chaleureusement Pierre Jouvelot (MINES ParisTech) pour ses relectures éclairées.

15. Références

- [1] Karim Barkati. *Entre temps réel et temps différé – Pratiques, techniques, et enjeux de l'informatique dans la musique contemporaine*. PhD thesis, Université Paris 8, 2009.
- [2] Karim Barkati. La polysémie du temps réel et du temps différé. In *Actes des JIM*, 2009.
- [3] François Déchelle. jMax : un environnement de programmation pour l'interactivité et le temps réel. In *Interfaces homme-machine et création musicale*, chapitre 4, pages 85–94. Hermes, Paris, 1999.
- [4] Martin Laliberté. Archétypes et paradoxes des nouveaux instruments. In *Les nouveaux gestes de la musique*, pages 121–138. Parenthèses, Marseille, 1999.
- [5] Pierre Lévy. *De la programmation considérée comme un des beaux-arts*. La Découverte, Paris, 1992.
- [6] Philippe Michel and Makis Solomos. Entretien avec Vincent Lê Quang et Alexandros Markeas. In *Fili-grane n°8 : Jazz, musiques improvisées et écritures contemporaines : convergences et antinomies*. Philippe Michel, 2008.
- [7] Curtis Roads. *L'audionumérique*. Dunod, 1998. Version française : Jean de Reydellet.
- [8] David Zicarelli, Gregory Taylor, Joshua Kit Clayton, jhno, Richard Dudas, R. Luke DuBois, Andrew Pask, and Darwin Grosse. *MSP 4.6 Tutorial and Topics*. Cycling'74, jun 2006.

LA REMISE EN CAUSE DE LA LEGITIMITE DE L'ŒUVRE-ÉVÈNEMENT PAR LES NOUVELLES TECHNOLOGIES

Anthony Papavassiliou

Étudiant en Maîtrise en musicologie à l'Université de Montréal

anthony.papavassiliou@gmail.com

RÉSUMÉ

Les musiques électroniques et électroacoustiques, dont la naissance remonte à un peu plus de cinquante années, ont introduit un concept jusqu'alors inconnu dans l'histoire de la musique : celui de la production musicale décorrélée de la performance. Ce phénomène peut être principalement expliqué par deux facteurs majeurs qui sont **la possibilité d'enregistrer** une partie ou la totalité de l'œuvre, remettant ainsi en question les formes de l'interprétation ainsi que le principe de l'« évènement » créé par la représentation, et **l'infinie possibilité de manipulation** sonore offerte par les nouvelles technologies qui permettent de dépasser les limites autrefois imposées par l'interprétation.

Le but de cet exposé sera de dresser un constat de la situation actuelle, de s'interroger sur les conséquences de l'utilisation des nouvelles technologies dans la création, la diffusion et la représentation de la musique et d'illustrer par des exemples précis les premières manifestations concrètes de la remise en question suscitée par cette évolution récente.

Nous essayerons également d'établir des pistes, laissant la voix à diverses méthodes de réconciliation entre la technologie, l'artiste et son public.

1. INTRODUCTION

Il y a quelques jours s'est tenue à l'Université de Montréal une conférence sur le geste, un thème très en vogue ces derniers temps, par un étudiant qui tentait de mettre au point une interface de contrôle du son qui permette à la fois de produire des sons riches et complexes par une manipulation intuitive, mais également d'établir un lien avec le public par le jeu instrumental.

Ce qui était intéressant, ce n'était pas tant que la facture de cet objet comprenait une dimension qui n'était pas du tout du ressort de la musique, mais que, contrairement à ce que pourrait être la couleur, la forme ou tout ajout décoratif, que cette dimension était primordiale ; lors de la présentation des différentes interfaces déjà connues et utilisées par les musiciens du monde entier, certaines avaient été écartées par l'artiste lors de l'élaboration de son prototype d'interface pour le simple fait de leur très pauvre communicabilité gestuelle.

L'étudiant, Martin Marier, n'était pas le premier à se préoccuper des fonctions du geste de l'interprète dans cette crise de la représentation que la musique est en train de vivre actuellement. Cette préoccupation est en réalité en chacun de nous depuis toujours et il convient d'aborder notre sujet en explorant les motivations qui nous poussent à vouloir ressentir un certain degré d'humanité dans l'élaboration d'une œuvre musicale.

2. PSYCHOLOGIE DE L'ÉVÈNEMENT

Qui irait voir un concert donné par des enceintes diffusant les œuvres d'un quatuor enregistré en *live* 3h auparavant ? Probablement personne. Pourtant il s'agirait d'une réelle diffusion du quatuor dont l'œuvre aurait réellement été jouée en *live*. Il y a donc lieu de se demander les raisons qui nous auraient poussés à faire un tel choix, si on prenait soin d'écartier l'influence que la qualité sonore d'une telle diffusion peut avoir sur notre décision.

L'intérêt du *live* ne réside pas moins dans l'unicité de l'évènement que dans l'interprétation qu'il peut nous offrir et c'est en assistant à la représentation que nous pouvons témoigner ou non de sa légitimité ; une représentation est une œuvre à part entière, qu'elle diffère ou non de l'œuvre originale, son caractère fugace lui confère toute sa valeur et c'est probablement la raison pour laquelle nous ressentons le besoin primordial de **voir** les auteurs, autrement dit les interprètes, nous apporter **la preuve** que cette œuvre, telle que nous la concevons lorsque nous faisons la démarche de venir l'apprécier, existe bel et bien.

C'est cette conception « évènementielle » de l'interprétation qui serait anéantie si nous ne pouvions pas constater son évidence, comme ce serait le cas par exemple avec une représentation où les interprètes seraient absents ou cachés.

Que se passerait-il alors, si les interprètes étaient bien présents et visibles, mais qu'ils donneraient l'impression de ne pas jouer la musique que nous serions en train d'entendre, de n'avoir aucune maîtrise sur elle ?

C'est précisément le problème soulevé dans la conception d'une interface musicale par Martin Marier et un problème qui touche particulièrement tous les amateurs de musiques créées à partir des nouvelles

technologies¹ : si nous venons à douter de la légitimité de la présence de l'interprète, nous remettons alors en cause la légitimité de l'œuvre-événement.

3. REPRÉSENTATION ET MUSIQUES ÉLECTRONIQUES

3.1. Introduction

Dans le monde des musiques électroniques, le problème de l'interprétation ne s'est pas posé avant l'avènement très récent des *laptops*² et des platines numériques. Jusqu'alors, les DJ enchaînaient les morceaux de musique gravés sur des disques vinyles ou bien triturait les sons, préparés à l'avance, sur des appareils munis d'une multitude de boutons et potentiomètres qu'il fallait ajuster en permanence. Ces techniques de production sonore dédiées aux événements *live* avaient comme avantage de mettre l'interprète en action quasi-permanente, le geste physique accompagnait les modifications sonores et garantissaient la légitimité de l'œuvre produite en temps réel.

Depuis peu, les platines numériques ont envahi le marché pour remplacer les platines vinyles et les *laptops* remplacent les séquenceurs et sampleurs par leurs homologues virtuels. Dans le cas des platines, le support, le disque, qui constituait à lui seul la raison d'être de l'instrument a littéralement disparu, tandis que ce qui demandait autrefois une certaine expression et précision du geste, dans le cas des appareils matériels de séquençage, s'est réduit aux simples clics d'une souris.

Bien sûr, les DJ continuent à utiliser des appareils qu'ils relient à leur *laptop* et continuent à effectuer quasiment les mêmes gestes sur leurs nouvelles platines, mais les liens forts qui unissaient le son au geste s'en sont trouvés très grandement affectés : **la vue** ne suffit plus car l'instrument s'est dématérialisé et est devenu virtuel. **La preuve** a laissé place à la confiance et à la réputation.

Malgré la confiance et la réputation, le doute peut subsister et mettre en branle un certain nombre d'interrogations : dans la mesure où le DJ ou l'interprète crée l'œuvre-événement en accordant volontairement une part plus ou moins grande à la manipulation du son lors de la représentation, où se situe la limite de la légitimité de l'œuvre ? De ce fait, puisque nous perdons peu à peu notre capacité de juger par nous-même cette légitimité, comment savoir si cette limite a été franchie ou pas ? Et aussi, bien que nous savons pertinemment qu'un DJ ou interprète gagne son statut par la qualité de son travail mais qu'il peut malgré tout nous tromper, chose qui était impossible avant, pourquoi ne le ferait-il pas ? Cela a-t-il une importance si nous conservons l'illusion que l'œuvre est légitime ?

¹ Plutôt qu'avec des instruments classiques qui demandent une participation active et continue de l'interprète sans quoi l'instrument n'émet tout simplement plus de son !

² Ordinateurs portables.

3.2. Le « mystère » Daft Punk

Un exemple assez concret de confrontation avec la légitimité du *live* est celui des représentations données par le groupe de *house music* Daft Punk, composé de Thomas Bangalter et Guy-Manuel de Homem-Christo. Ces artistes originaires de Versailles ont pris, avec le temps et le succès, l'habitude d'apparaître masqués non seulement à chacune de leurs représentations mais également sur chacune des photos officielles que l'on peut trouver, que ce soit dans la presse ou bien dans leurs albums. Lors de leur tournée mondiale de 2006 (9 dates au total dans cinq pays différents), les Daft Punk mettaient en scène leur représentation d'une manière à la fois remarquable et troublante : nichés en haut d'une pyramide, ils œuvraient déguisés pendant toute la durée du concert. Cette tournée amena plusieurs constatations et interrogations :

- Déguisés, personne ne peut attester qu'il s'agit bien d'eux, en conséquence, seule la virtuosité de certains gestes (aussi bien musicaux que corporels) de l'exécution pourrait éventuellement attester de la légitimité de l'œuvre à laquelle on s'attend.

- Le *set*, incroyablement millimétré, ne connaît aucune imperfection. Puisqu'on ne voit pas les artistes, ce show aurait-il pu être enregistré et présenté par n'importe qui ?

- Le jeu de lumières scéniques, à la fois complexe et varié, concorde parfaitement avec les inflexions de la musique. Dépassant l'algorithme traditionnel de réponse de la lumière aux variations du son, le jeu de lumières est ici clairement programmé pour être en phase avec la musique, ce qui renforce l'idée d'une possible rediffusion d'une œuvre enregistrée déjà envisagée par l'aspect sonore de la représentation.



Figure 1. Les Daft Punk mixant au centre de leur pyramide.

La question ici n'est évidemment pas de mettre en doute la qualité de l'événement produit, qui était à chaque représentation d'un haut niveau selon les participants, mais de se poser la question de son intérêt à partir du moment même où il est possible pour les participants de douter de la légitimité de l'événement qu'ils sont en train de vivre. Notamment pour ceux qui s'attendraient à une œuvre-événement : comment pourraient-ils être certains que l'événement auquel ils

assistent n'est pas simplement une version « en plein air » d'un produit qu'on sera susceptible de leur proposer un peu plus tard dans le commerce. Une version sans performance, sans imprévu, sans aucune réelle interaction interprète/public. Il serait d'ailleurs intéressant d'enquêter pour savoir si la version du spectacle offert a pu différer d'une scène à l'autre.

3.3. L'œuvre figée

Bien sûr, nous pourrions parler du cinéma puisque la version d'une œuvre offerte dans les salles est supposée être identique à celle proposée un temps plus ou moins variable après dans le commerce.

Comme pour le concert de Daft Punk supposé enregistré, nous pouvons aller au cinéma pour, non pas avoir l'exclusivité au sens de l'unicité de l'œuvre, mais pour découvrir l'œuvre au plus proche de sa création. Dans ce cas, il est évident que le fait qu'il n'y ait pas de performance ou d'interaction n'a aucune importance. Nous allons également au cinéma pour l'immersion, pour avoir la sensation d'être « plongés » au cœur du récit ou de l'action, ce qui correspond plutôt assez bien à l'idée que l'on se fait du concert musical. Nous allons également au cinéma pour profiter d'une qualité et d'un confort de son et d'image inaccessibles chez soi, ce qui peut être plus discutable dans le cas d'un concert.

Pour résumer, si nos attentes concernent moins la totale unicité de l'œuvre dans l'interprétation que les qualités de sa reproduction, l'œuvre musicale, qu'elle soit enregistrée ou performée, conserve sa légitimité. Tant que nous dissociions les paramètres de représentation de l'œuvre de l'œuvre elle-même, nous n'assistons pas à l'**œuvre-événement**, éphémère et unique, mais à une **œuvre figée**, reproductible, dans un contexte plus ou moins favorable à sa diffusion. En d'autres termes, dans l'œuvre-événement, l'unicité de l'évènement, qui est objective, confère à l'œuvre son identité (et donc son unicité), tandis que dans l'œuvre figée, l'unicité de l'évènement, subjective, n'agit pas sur l'identité de l'œuvre.

C'est là tout l'enjeu de l'interprétation et du doute semé par son absence : suis-je en train d'assister à une œuvre figée ou une œuvre-événement³ ?

3.4. L'exemple du « scandale Justice »

Un autre cas de dissonance amené par le *live* est la fameuse photo de Gaspard Augé, membre du groupe Justice, en train de manipuler un contrôleur midi qui n'est manifestement pas branché.

Cette fois-ci il n'est plus question de la présence ou non des artistes ou des interprètes, qui semblent bien engagés dans l'élaboration d'une représentation *live*,

³ Notons que la réponse à une telle question n'a probablement de l'importance que si nous avons envisagé une œuvre-événement.

mais de la duperie qui peut exister dans la relation établie entre le geste physique et la musique ; encore une fois, certes, on peut très bien imaginer que si les artistes feignent l'action il s'agit d'une œuvre préenregistrée sur laquelle il n'effectuent que peu ou pas de changement, mais surtout, il s'agit dans ce cas d'une volonté manifeste de simuler la réalisation d'une œuvre-événement alors qu'il pourrait tout à fait s'agir d'une œuvre figée (en tout cas, plus « figée » qu'on pourrait le penser).

Ce n'était pas vraiment le cas quand nous avons abordé les représentations de Daft Punk qui, volontairement, s'efforcent d'effacer la présence de l'interprète pour privilégier le show. Ce n'est d'ailleurs peut-être pas pour rien si leurs déguisements sont ceux de robots, représentant sans aucun doute le reniement du caractère humain et sensible de l'œuvre, mettant ainsi l'emphase sur son caractère mécanique, mais aussi et surtout, déterminé et autonome.



Figure 2. La photographie qui a fait scandale.

Un tel scandale prouve que la présence de l'artiste ainsi que le geste ne constituent pas en totalité l'interprétation : le rôle de la technologie n'est plus simplement de faciliter la tâche de l'artiste dans sa quête de production sonore, que les fins visées soient la richesse, la simplicité, l'expérimentation ou l'efficacité, mais également d'assurer son intégrité et celle de l'artiste vis-à-vis de l'œuvre.

4. LES SOLUTIONS POSSIBLES

4.1. Introduction

Si la conception de l'œuvre-événement est mise à l'épreuve par les artistes travaillant les musiques électroniques c'est parce que les nouvelles technologies ont ouvert la voie à de nouvelles façons d'envisager l'œuvre, tant au niveau de la composition, de la diffusion que de l'écoute.

Il est clair, à la lumière de ces faits, qu'il nous faut reconsidérer notre approche de l'œuvre musicale comme

les mutations du support audio nous l'ont déjà suggéré ces quelques dernières années : l'œuvre, en tant qu'évènement, n'est qu'une autre de nos conceptions ébranlées par l'utilisation des nouvelles technologies dans la musique. Nous pouvons citer le home studio et la capacité pour chacun de produire de la musique sans maîtriser d'instrument sous sa forme classique, évacuant ainsi les préoccupations liées à la performance et la technicité humaine. Citons également la musique numérique, entraînant le piratage et le problème de la rémunération des artistes.

Les mutations technologiques de la musique ne sont pourtant pas obligatoirement bénéfiques pour l'artiste et néfastes pour le public ou inversement. L'enregistrement est une des premières révolutions qui a permis non seulement aux artistes d'être mieux diffusés, mais également au public de pouvoir écouter de la musique de manière totalement illimitée. Nous pouvons donc envisager un destin pour l'œuvre-évènement plus glorieux que celui qui semble se dessiner, de la même façon qu'il suffirait de penser un système adéquat pour que les nouvelles technologies soient à la fois bénéfiques aux artistes et au public dans le domaine de la diffusion de la musique.

4.2. Renforcer l'évènement

L'idée est ici de concevoir l'œuvre-évènement de manière explicite, d'envisager l'utilisation des nouvelles technologies comme des outils servant cet objectif et non comme simples manières de produire du son, à la manière des concerts traditionnels mettant en jeu des instruments qui, sans l'action de l'homme, n'émettent aucun son. Cela ne veut pas forcément dire qu'il faut créer une réplique électronique d'un instrument classique pour rendre honneur à l'interprétation, mais qu'il faudrait à la fois envisager une technologie impliquant la participation de l'artiste par le geste, permettant au public de faire le lien entre le geste corporel et le geste musical, et également persuader à raison le public de l'action déterminante de l'artiste sur son instrument, quelle que soit sa forme.

Bien que l'interaction geste/instrument soit primordiale dans la représentation interprétée, il ne faut pas négliger le fait qu'ils servent la relation artiste/public si particulière à l'œuvre-évènement, puisqu'il ne s'agit pas seulement d'apprécier l'œuvre mais de la voir se créer devant nos yeux.

L'œuvre n'a pas nécessairement besoin d'être improvisée pour être « évènement », mais il paraît clair que ce qui nous amène à considérer qu'elle puisse revêtir un caractère unique et fugace est la participation primordiale de la dimension humaine, conférant à l'œuvre à la fois sa part d'inexactitude (ou d'erreur) non reproductible ainsi que sa capacité de réaction et d'adaptation au public. Pour nous apporter la preuve de la réelle participation humaine, nous devons avoir recours à une technologie sans laquelle l'œuvre ne pourrait pas fonctionner et qui ne fonctionnerait pas non

plus sans la manipulation de l'homme. En effet, une technologie qui n'a pas nécessairement besoin d'être manipulée par l'homme n'a pas d'utilité dans la mesure où l'œuvre pourrait alors très bien se dérouler sans l'artiste, une telle technologie n'apporterait ni la preuve, ni le degré d'humanité nécessaire à la naissance de l'évènement tel que nous le concevons.

La technologie qui nous intéresse ne doit pas non plus permettre des manipulations qui n'ont pas de signification musicale, ce qui tombe sous le sens, mais doit également amener un geste dont la signification s'accorde au maximum avec l'effet désiré : le public doit pouvoir comprendre le sens de l'interaction entre l'artiste et la technologie, entre ses gestes et le rendu sonore, de sorte à ce qu'il puisse apprécier la qualité de maîtrise de l'artiste sur son œuvre.

De telles exigences nous permettent d'établir trois conditions auxquelles la technologie utilisée pour l'œuvre-évènement devrait satisfaire. Les détails de ces conditions seront discutés dans le chapitre suivant. Il nous faudrait donc une technologie :

- Capable de produire une gamme de sons et samples la plus étendue possible ;
- Dépendante à chaque instant de l'action de l'artiste sur elle ;
- Dont la modification des paramètres nécessitent une implication gestuelle explicite et efficace;

La supposition qu'il puisse exister une technologie dédiée à l'œuvre-évènement regroupant ces trois conditions et qui puisse susciter une approche des musiques électroniques propice à la mise en œuvre d'une représentation à la fois technique et vivante nous amène à nous demander si l'idéal que nous souhaiterions atteindre ne serait donc pas la création d'un instrument dit « total », un outil qui permette un contrôle absolu sur la musique et demandant une grande agilité pour être maîtrisé, ou encore, un instrument dont un panel de paramètres seraient influencés par certaines manifestations du public.

Dans ce dernier chapitre nous allons développer notre propos en explorant les différentes pistes qui s'offrent à nous.

Bien sûr l'idée ici n'est pas d'imaginer imposer un instrument ou type d'instrument pour l'ensemble des pratiques artistiques impliquant de la musique électronique. Il s'agit plutôt de faire le constat de ce qui manque aux technologies utilisées actuellement et, par la même occasion, de proposer des pistes d'élaboration basées sur la prise en compte de considérations jugées nécessaires pour le maintien et le développement de l'œuvre-évènement au sein des arts impliquant des musiques électroniques.

4.3. L'importance du geste dans l'œuvre-événement : quels types de gestes ?

Selon Leticia Cuen⁴, « la matière expressive de la musique est gestuelle », la musique « nous renvoie toujours à la reconnaissance et/ou au souvenir des états moteurs, cinétiques, physiologiques qui sont à l'origine des sensations et des émotions humaines ». Comme nous l'avons vu, lors d'une représentation musicale, le public voit en la présence d'un jeu scénique de l'artiste une preuve de la légitimité de l'œuvre en tant qu'évènement : celui-ci se meut pour insuffler à l'œuvre son caractère unique.

Cette action déterminante sur l'œuvre doit être favorisée par l'instrument dont la facture doit être réalisée dans ce sens tout en respectant la cohérence du geste.

En effet, le public ne doit pas avoir l'impression que le musicien surjoue son œuvre en y ajoutant une composante gestuelle inutile : pourquoi courir le long de la scène pour augmenter le volume alors qu'il suffirait à un artiste muni d'une console des plus simples de tourner un potentiomètre pour obtenir le même rendu ? Cependant, comme le souligne Martin Marier⁵, il faut tenir compte de l'effet des *ancillary gestures* (Cadoz, 1988; Cadoz et Wanderley, 2000), ou gestes auxiliaires, en expliquant qu'il « s'agit des gestes qui ne servent pas directement à la production du son, mais qui font tout de même partie de la prestation. Dans le cadre d'une étude sur les mouvements des clarinettes, Wanderley et al. (2005) concluent que les mouvements de l'interprète contribuent à mieux communiquer les intentions musicales au public. »

La manipulation de l'instrument doit donc être à la fois élémentaire et expressive : elle doit rendre compte des processus cognitifs et émotionnels mis en avant par Leticia Cuen inspirés par les travaux d'Imberty et Cuen tout en évitant l'exagération.

Martin Marier, en concevant son éponge, poursuivait un objectif similaire : il voulait produire du son mais également que la manipulation sur son instrument rende compte de manière visuelle le geste musical induit. Ainsi, secouer l'éponge engendrait un vibrato, lui faire effectuer des rotations faisait varier le pitch. Bien sûr il s'agissait là d'une technologie assez rudimentaire, mais le concept avait le mérite de viser juste en donnant au geste une importance capitale; un type de communication élémentaire entre l'artiste et son public résumable en un mot : l'**interaction**.

4.4. Rétablir la confiance : l'action déterminante de l'artiste sur l'instrument

Il n'est pas aisé d'imaginer une technologie à la fois avancée, efficace, pratique, simple et démocratique qui nous permette de concevoir en live la musique que l'on

veut tout en conservant la présence indispensable de l'artiste, en pratique mais aussi aux yeux du public. Nous savons que le choix vers du matériel toujours plus réduit par les DJ et autres artistes de live se fait pour nombre de raisons avantageuses et il paraît que pour rétablir certains des aspects essentiels à l'œuvre-événement il faudrait planifier l'amointrissement voire la mise à l'écart de certains de ces avantages.

Comment bénéficier alors des avantages procurés par l'informatique dans le domaine de la production sonore tout en garantissant l'intégrité du live ?

Des appareils ont déjà été créés dans l'optique de réaliser un « instrument total ». C'est le cas du Méta-Instrument développé par Puce Muse. Le Méta-Instrument est décrit par ses concepteurs comme suit : « Le MI3⁶ s'apparente à un double Joystick portable qui peut manipuler, simultanément et indépendamment les unes des autres, 54 variables continues. La position des avant-bras se contrôle sans les mains, l'orientation de la poignée se contrôle par la paume, et chaque doigt comporte cinq capteurs indépendants. La mesure des rotations est au 1/20 de degré et la pression au 1/10 de gramme. »

C'est donc un instrument complexe, permettant de contrôler un nombre très important de paramètres musicaux et d'agir sur eux par le geste. Cependant le Meta-Instrument n'est pas parfait : c'est un objet très onéreux, encombrant, lourd, qu'il faut porter sur soi. De plus, il est conçu comme une amélioration ou une extension du corps humain et non comme un outil générateur de son.

Quand, chez les Daft Punk, l'identité de l'artiste disparaît au profit de l'instrument et, d'une manière plus prononcée, de l'œuvre, ici l'instrument s'efface et laisse place à un « homme-musique ».

Il est important de différencier ce qui pourrait faire la différence entre l'homme-orchestre, dont parle Puce Muse au sujet des musiciens utilisant le Méta-Instrument, et l'homme-musique : dans le cas du Méta-Instrument, la frontière entre l'homme et l'instrument disparaît, or, sur quelle base pouvons-nous qualifier la maîtrise de l'artiste sur son instrument et donc sur la musique si nous n'avons pas d'élément objectif et concret nous permettant de nous en rendre compte ?

Le Méta-Instrument est un formidable outil d'expérimentation et de création musicale, cependant il ne permet pas de communiquer convenablement avec le public dans le contexte d'une œuvre-événement : ce dernier doit pouvoir apprécier ce que l'artiste insuffle à l'œuvre dans le moment par la transcendance des contraintes inhérentes à la production musicale et l'interprétation, effacer les contraintes d'interprétation défie en quelque sorte l'artiste, annihilant la part d'humanité nécessaire à l'œuvre-événement.

⁴ *Temps, geste et musicalité*, 2007, p. 151.

⁵ *Le rapport geste-son en musique électroacoustique*, 2010, p. 9.

⁶ Acronyme pour désigner la troisième génération de Méta-Instruments.

4.5. Le contrôle des paramètres

Des expériences ont montré que la qualité de l'intégrité de la manipulation sonore par un interprète était fortement influencée par notre longue tradition de manipulation d'instruments classiques. Certaines des observations faites vont dans le sens d'une relation geste/son capitale dans l'interprétation. Ainsi, on sait qu'il existe un lien fort entre effort et expressivité (Ryan, 1991). Marier décrit cette relation ainsi : « L'effort physique est essentiel pour jouer d'un instrument acoustique. Pour jouer plus fort, on doit fournir plus d'énergie ; c'est en quelque sorte un combat. Il s'agit là, en soi, d'une image qui reflète la condition humaine et qui peut ouvrir la route vers l'expression de sentiments forts. ».

Les expériences de Hunt et Kirk (2000) montrent par exemple que la manipulation du volume doit être sujette à un effort fourni pour donner sa pleine expressivité, malheureusement, le volume est beaucoup moins un paramètre dynamique dans les musiques électroniques que dans les musiques instrumentales puisque la production sonore est davantage axée sur la composition du spectre dans l'espace et le temps que sur les variations d'intensité.

Une méthode pourrait être d'envisager des moyens inhabituels ou inédits d'agir sur la production sonore tout en conservant un certain degré d'expressivité et d'interaction avec le public. Le cas de Tim Exile⁷ est très intéressant à ce sujet.

Tim Exile est un producteur et interprète de musiques électroniques doté d'une solide formation au violon et titulaire d'une maîtrise en composition électroacoustique. Cet artiste est connu mondialement pour ses *live* particulièrement inventifs; en effet, muni d'un *setup*⁸ quasiment toujours identique, il enregistre sa voix ou encore les bruits de la scène ou du public pour l'intégrer en direct à ses compositions. La part d'interaction est donc importante, l'intégrité également puisqu'il n'y a rien de plus « concret » en musique que l'utilisation de la voix humaine. Concernant le geste, son expressivité est garantie par le *setup* composé d'un clavier midi, de pads et de divers potentiomètres. Quand Tim Exile est équipé d'un *laptop*, ce n'est pas pour garder les yeux rivés dessus et y effectuer des manipulations mais simplement d'un moyen rapide de visualiser certains paramètres de sa configuration.

4.6. Conclusion

Il est certain que l'avènement de l'électronique et de l'informatique amène avec son lot infini de nouvelles possibilités une reconsidération de notre manière de penser l'art musical. Les nouvelles technologies ont bouleversé nos conceptions et remis en question notre approche du monde esthétique. En abolissant les

limites qui autrefois garantissaient l'authentique unicité de l'évènement musical, elles ont semé le doute parmi les attentes de l'auditeur. Pourtant celui-ci n'a pas changé ; en quête de cette relation privilégiée accordée par l'œuvre-évènement entre lui, l'interprète et la musique, il aime se reconnaître dans une temporalité synonyme de vie et avoir la sensation d'être un acteur, sinon un témoin, de cette forme d'art.

Que faut-il alors ? Un instrument total ? Une interface contrôlant un maximum de paramètres ou plusieurs interfaces manipulées par plusieurs personnes ? Faut-il un orchestre ? Peut-on encore envisager la création musicale et l'interprétation par l'instrumentation, comme nous le faisons depuis des siècles ? Il est sûr cependant que notre perception musicale restera encore longtemps liée à un certain nombre de paramètres intemporels que la technologie non seulement, n'abolira pas, mais qu'elle devra plutôt satisfaire pour entretenir les meilleures relations possibles entre l'homme et l'œuvre-évènement.

L'exploration de ces paramètres tout au long de cet exposé a permis de comprendre qu'il existe des issues et que le domaine est fertile à de nombreuses innovations. Peut être que le *setup* de Tim Exile est une bonne voie d'exploration ou que le Méta-Instrument entraînera des améliorations qui octroieront à ce type de technologie une place de choix dans le cœur du public.

De nombreux efforts restent à fournir dans ce domaine, ce n'est pas moins une quête de l'instrument parfait que de l'intégrité parfaite, une préoccupation presque aussi cruciale que de vouloir comprendre ce qu'est l'art et l'esthétique : vouloir faire perdurer l'œuvre-évènement en la développant, c'est continuer à considérer plus que tout que la musique n'est pas un art figé mais un art du mouvement, un art du temps que l'homme crée à sa mesure en lui insufflant ses caractéristiques de momentanéité.

5. RÉFÉRENCES

- [1] Arfib, D., Couturier, J. M., Kessous, L., Verfaillie, V., « Strategies of mapping between gesture control parameters and synthesis models parameters using perceptual spaces », *Organised Sound*, vol. 7, N°2, 2002, pp. 127-144.
- [2] Battier, M. « L'approche gestuelle dans l'histoire de la lutherie électronique. Étude de cas: le theremin », *Les nouveaux gestes de la musique*, H. Genevois et R. de Vivo (eds). Marseille: Éditions Parenthèses, 1999, pp. 139-156.
- [3] Beaudouin-Lafon, M. « Moins d'interface pour plus d'interaction », *Interfaces Homme-Machine et Création Musicale*, H. Vinet et F. Delalande (eds). Hermès, 1999, pp. 123-141.
- [4] Cadoz, C. et Wanderley, M. « Gesture and Music », *Trends in Gestural Control of Music*, M. Wanderley & M. Battier (eds). Paris: IRCAM, 2000.

⁷ Tim Shaw de son vrai nom.

⁸ Mot couramment utilisé pour désigner une configuration quelconque permettant d'exécuter la représentation prévue.

- [5] Cadoz, C. « Musique, geste, technologie », Les nouveaux gestes de la musique, H. Genevois et R. de Vivo (eds). Marseille: Editions Parenthèses, 1999, pp. 47-92.
- [6] Choi, I. « Gestural Primitives and the context for computational processing in an interactive performance system », CD-room Trends in Gestural Control of Music, Wanderley M. et Battier B. (eds), publication Ircam, 2000.
- [7] Couturier, J. M. « Espaces Interactifs Visuels Pour le Contrôle des Sons Musicaux », Journées d'études Espaces Sonores 2002, Marseille : 2002.
- [8] Cuen, L., « Le geste sonore et le geste cinétique : Une évocation des sensations et des émotions humaines », Temps, geste et musicalité, Imbert, M. et Gratier, M. (eds), Paris: L'Harmattan, 2007.
- [9] de Laubier, S. et Goudard, V. « Meta-Instrument 3: a look over 17 years of practice », Proceedings of International Conference on New Interfaces for Musical Expression (NIME'06), Paris: 2006.
- [10] Gritten, A. et King, E. « Music and gesture », Gritten, A. et King, E (eds). Aldershot ; Burlington: Ashgate, 2006.
- [11] Hunt, A., Wanderley, M., Kirk, R., « Towards a Model for Instrumental Mapping in Expert Musical Interaction », Actes de la International Computer Music Conference ICMC'2000, Berlin, Allemagne, ICMA, pp. 209-212.
- [12] Hunt, A. et Kirk, R., « Mapping strategies for musical performance », Trends in Gestural Control of Music, 2000, pp. 231-258.
- [13] Marier, M., « Le rapport geste-son en musique électroacoustique », Mémoire de Maîtrise, Montréal : 2010.
- [14] Ryan, J., « Some remarks on musical instruments design at STEIM », Contemporary music review, 6(1) :3-17, 1991.
- [15] Stowell, D., Plumbley, M.D. et Bryan-Kinns, N. « Discourse analysis evaluation method for expressive musical interfaces », Proceedings of New Interfaces for Musical Expression (NIME'08), Genova: 2008.
- [16] Trueman, D. « Why a laptop orchestra ? », Organised Sound, vol. 12, n°2, 2007, pp. 171–179.
- [17] Wanderley, M. et Depalle, P. « Contrôle Gestuel de la Synthèse Sonore », Interfaces homme-machine et création musicale, Vinet H. et Delalande F. (eds), pp. 145-163. Hermès, 1999.

VERTIGES DE L'ESPACE : UN INSTRUMENT POUR LA PERFORMANCE ÉLECTROACOUSTIQUE IMPROVISÉE

António de SOUSA DIAS

CITAR - Centro de Investigação em Ciências e Tecnologias
das Artes (Universidade Católica Portuguesa)
CICM - université Paris 8, MSH Paris Nord
sousa.dias@wanadoo.fr

RÉSUMÉ

Nous présentons ici un instrument conçu pour le projet *Vertiges d'espace* et destiné à la diffusion sonore dans le contexte d'une performance de musique électroacoustique improvisée. Cet instrument a été expérimenté en concert et les premiers aboutissements nous permettent d'entrevoir des solutions concernant l'équilibre entre le nombre de contrôles nécessaires, le caractère et la gestion de la mise en espace envisagée.

Mots clé : Interfaces logicielles ; interprétation musicale ; spatialisation du son.

1. PRÉSENTATION DU PROJET ET DU DISPOSITIF

1.1. Le projet *Vertiges de l'espace*

Vertiges de l'espace est une performance de musique électroacoustique improvisée. Elle réunit quatre musiciens : *Les Phonogénistes* (Laurence Bouckaert, Pierre Couprie, Francis Larvor) et António de Sousa Dias.



Figure 1. *Vertiges de l'espace* : performance au *New Morning*, Paris, le 25 janvier 2010, Soirée SOS-Art.

La particularité de cette performance tient au fait que les trois premiers musiciens improvisent le matériau musical (cf. §2.1 *Les modes de jeu des musiciens* ci-après) et le quatrième musicien se charge de

l'improvisation de la spatialisation de ce matériau sur des enceintes disposées autour du public (cf. Figure 2).

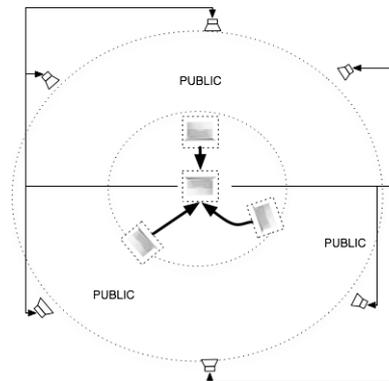


Figure 2. *Vertiges de l'espace* : disposition des musiciens, public et haut-parleurs.

1.2. Le dispositif de mise en espace

Pour ce projet, après avoir choisit la disposition des musiciens, il a fallu déterminer la façon d'improviser la mise en espace. Ainsi, nous avons conçu un instrument constitué d'une application programmée dans *Max 5* [3], piloté par une surface de contrôle et une carte-son.

Cet instrument récupère l'audio produit par les trois musiciens, chacun délivrant deux canaux. Il gère un nombre arbitraire d'enceintes disposées en cercle (cf. Figure 3).

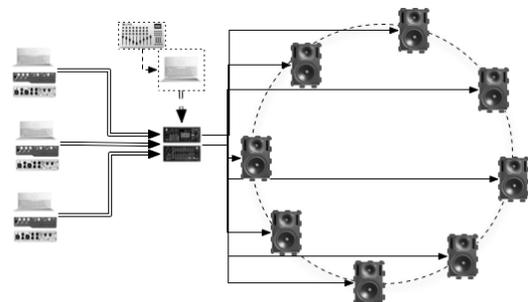


Figure 3. *Vertiges de l'espace* : flux de signal audio et de contrôle.

D'une façon générale, il respecte les consignes suivantes :

- surveiller le volume générale de chaque musicien ;
- avoir un mode de jeu souple concernant la gestion de figures spatiales ;
- s'adapter à différents dispositifs, notamment le nombre de haut-parleurs disponibles et leur emplacement en fonction de la salle.

Nous verrons le mode de fonctionnement de cet instrument par rapport aux choix qui ont été faits. Puis, nous présenterons brièvement le système conçu pour garder et restituer la trace d'une performance et finalement nous exposerons des perspectives issues des résultats obtenus.

2. LES MODES DE JEU ET LES MODES DE DIFFUSION

2.1. Les modes de jeu des musiciens

Les modes de jeux des musiciens produisant du son ont été déterminant pour la conception de la diffusion et des possibilités de contrôle choisies.

Dans un premier temps, nous avons repéré de façon fort « brutale », voire réductrice, leurs modes de jeux, très riches, variés et différents :

- Musicien 1 – le jeu avec une flûte conduit soit à des granulations et des accumulations, soit à des objets sonores de courte durée (impulsions) et aussi à des notes et tenues longues à caractère granuleux ;
- Musicien 2 – l'emploi d'un clavier produit soit des trames, soit des ponctuations à travers des impulsions et des notes ;
- Musicien 3 – un mode de jeu avec une tendance à la construction des longs objets à caractère stable ou évolutif, articulant trames et accumulation, avec un caractère parfois itératif.

Ensuite, nous avons envisagé une mise en correspondance entre les tendances détectées et une façon de les souligner au travers de certaines figures spatiales :

- Musicien 1, nous proposons des mouvements diversifiés, rapides, déployés partout dans l'espace, correspondant à ses granulations. Vitesse rapide de rotation et d'oscillation du rayon. Des « figures-fleur », presque mouvement brownien ;
- Musicien 2, nous avons imaginé deux types de mouvement : soit des rotations très lentes, soit des mouvements rapides de déplacement comme un moyen de souligner les ponctuations où chaque note se déplace rapidement pendant toute sa durée"=
- Musicien 3, les mouvements plus évolutifs, par vagues, trames, font appel soit à des rotations lentes, envoûtantes, soit à des déplacements plutôt du type oscillation/rebondissement.

Acceptant le concert comme événement audiovisuel, nous avons ajouté deux autres propositions :

- la possibilité d'aligner le son avec le musicien qui le produit pour un repérage causal ;
- la possibilité de ramener l'audio de chaque musicien au centre du dispositif pour un « solo ».

2.2. Les modes de diffusion

L'application doit aussi être souple, reprogrammable et adaptable à différentes situations. Ainsi, nous sommes partis des nécessités suivantes :

- garder un caractère proche du concert, du point de vue de la complémentarité *solo-tutti* remplacée par le couple figure-fond ;
- placer et déplacer des sons rapidement dans l'espace défini par le cercle de haut-parleurs ;
- contrôler les six entrées avec agilité.

Pour implanter l'instrument, nous avons d'abord décidé de grouper les entrées de chaque musicien et prendre le cercle de haut-parleurs comme contrainte privilégiée. Ensuite, nous avons pris les décisions suivantes :

- programmer des contrôles permettant de placer le son de chaque musicien dans son entourage, pour permettre un repère audiovisuel ;
- programmer la possibilité d'avoir des trajectoires automatiques ;
- concevoir un instrument modulaire où le nœud central sauvegarde tous les préréglages des modules subordonnés ;
- contrôler l'instrument à partir d'une surface de contrôle et par conséquent permettre la reconfiguration MIDI pour interfacer avec n'importe quelle surface de contrôle ;
- de même pour la configuration audio et pour le nombre de haut-parleurs qui peut varier à chaque concert ;
- utiliser l'objet « ambipan~ » développé par Benoit Courribet et Rémi Mignot¹ au CICM, pour un faible usage des ressources au niveau du processeur.

Finalement, nous avons envisagé trois modes d'exécution :

- le mode *trajectoire circulaire*, spécialement envisagée pour ce projet ;
- le mode *orbite de Lissajous* – possibilité de réaliser de figures harmoniques basées sur des figures de Lissajous ;
- le mode *figures spatiales* – ce mode, encore en développement, est basée dans les figures définies par Vande Gorne [2]0

¹. Téléchargement sur le site <http://cicm.mshparisnord.org>, section *Téléchargements*, Outils de spatialisation (ambipan~, vbapan~, ambicube~) (03.03.2010).

Ces modes, présentés sur la figure suivante, sont gérés par deux pistes d'une surface de contrôle comme la *Evolution UC-33*. Le choix de cette surface de contrôle est dû à l'accessibilité de plusieurs contrôleurs (*faders* et potentiomètres) programmables et assignables selon nos besoins.

Ainsi, chaque musicien (deux pistes audio d'entrée) est assigné à huit paramètres, dont la signification varie selon le mode de diffusion choisi.

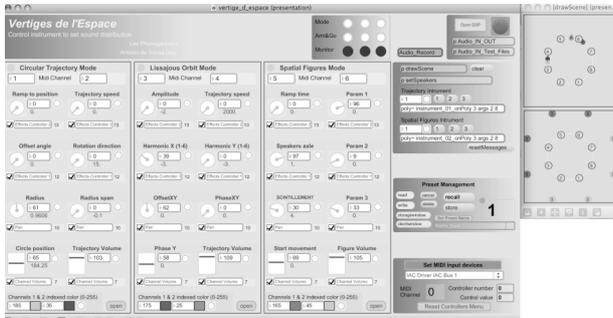


Figure 4. Application *Vertiges de l'espace* : fenêtre principale et repérage de position des sources.

Ces modes sont des interfaces vers des *patches* qui contrôlent le routage et la distribution des signaux audio vers les sorties. Ils sont totalement configurables en ce qui concerne le nombre d'entrées audio par musicien, le nombre de sorties/haut-parleurs disponibles, leur emplacement et la méthode de diffusion.

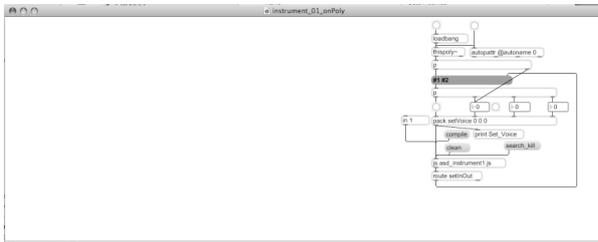


Figure 5. Patch pour le routage de la distribution du signal. Pour des exemples d'instanciation voir la Figure 6 et la Figure 7 plus bas.

Les deux pistes audio d'entrée de chaque musicien sont toujours coordonnées et commandées par les mêmes paramètres, d'où l'existence, pour les deux premiers modes, de paramètres agissant sur la différence de « phase » de l'emplacement des sources.

Nous présentons ensuite un résumé de chaque mode de diffusion.

2.2.1. Trajectoire circulaire

Ce mode a été conçu spécialement pour ce projet. Partant du postulat que les haut-parleurs sont disposés sur un cercle de rayon 1 (sans unité de mesure) pour la diffusion, il propose un fonctionnement assez souple, même s'il est limité. Les contrôles sont repartis selon les fonctions présentées dans le Tableau 1.

		Trajectoire circulaire		
		(Positionnement manuel)	(Contrôle orbite)	
1-5	Ramp to position (Temps pour atteindre la position suivante)	to	Trajectory speed (Vitesse maximale de rotation et de variation d'amplitude d'oscillation)	(Temps)
2-6	Offset angle (Angle de départ)		Rotation direction (Direction et vitesse relative de la rotation)	(Angle)
3-7	Radius (Distance des sources par rapport au centre)		Radius span (Amplitude et direction de l'oscillation autour du rayon)	(Rayon)
4-8	Circle position (« Phase » entre les deux entrées audio : superposées – opposées)		Trajectory Volume	

Tableau 1. Mode *Trajectoire circulaire* : mise en correspondance des paramètres de contrôle et leur fonction.

2.2.2. Orbite de Lissajous

Ce mode engendre des courbes de Lissajous. Les paramètres choisis proposent une articulation entre deux courbes de Lissajous (une courbe par entrée audio) et les paramètres disponibles (cf. Tableau 2).

Nous remarquons ici un des problèmes de mise en correspondance entre domaines d'application. Par exemple, le contrôle *Offset XY* réalise une application d'un domaine avec une dimension sur un domaine à deux dimensions. Selon la valeur du contrôle *V* (entre 0 et 127), le centre de l'orbite est placé sur le plan *XY* selon la transformation :

$$x = \left[\frac{3 * v}{127} - \text{int} \left(\frac{3 * v}{127} \right) \right] * 3 - 1.5$$

$$y = \begin{cases} 1 & 0 < v < 43 \\ 0 & 44 < v < 86 \\ -1 & 87 < v < 127 \end{cases}$$

Ces deux modes de diffusion sont plus adaptés à une mise en place basée sur des coordonnées polaires. Ainsi, le *patch* de routage du signal pour ces deux modes est basé sur l'utilisation d'un objet « *ambipan~* » distribuant spatialement chaque entrée.

3. L'ENREGISTREMENT DES DONNÉES

Dans un premier temps, la conception et la construction de l'instrument pour l'improvisation de la diffusion a été la priorité. Ensuite, nous avons commencé à enregistrer les répétitions et les concerts en enregistrant, à partir de l'instrument de diffusion, les 6 canaux audio des musiciens et les canaux audio de diffusion, en général 8, obtenant ainsi, en général, un fichier audio de 14 pistes.

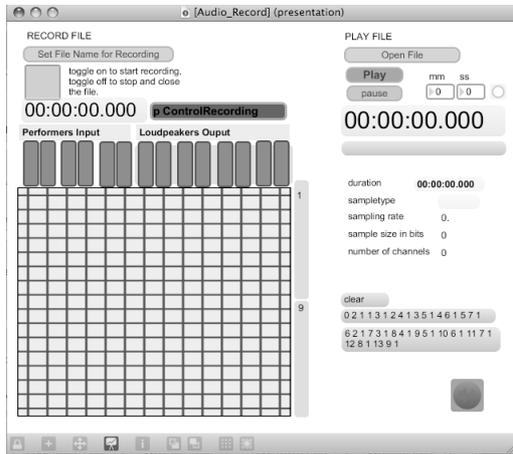


Figure 8. Fenêtre d'enregistrement.

En complément de cet enregistrement audio, deux fichiers au format MIDI sont également sauvegardés. Le

premier fichier sauvegarde les commandes effectuées lors de la performance et, en raison du caractère dynamique de cet instrument, le deuxième fichier contient les mouvements de déplacement de chaque entrée audio.

Suite à des essais, pour le choix de la méthode de transcription, d'enregistrement et de restitution des données de diffusion nous nous sommes fixés sur la méthode suivante :

- enregistrer les données sous forme de fichier MIDI ;
- faire correspondre chaque piste MIDI à un canal audio d'entrée, contenant les données attribuées à ce canal ;
- enregistrer les valeurs de volume, assigné au contrôleur 7 et de position (x, y), assignée aux contrôleurs 14 pour la valeur de x et 15 pour la valeur y ;
- convertir linéairement les valeurs de volume entre 0 et 158 en valeurs entre 0 et 127 ;
- convertir linéairement les valeurs de coordonnées de la position x et y, entre -2.0 et 2.0, de chaque entrée audio en valeurs entre 0 et 127.

Pour la restitution de ces données, nous avons conçu un *patch* Max qui fait le rendu des mouvements sonores (cf. Figure 9).

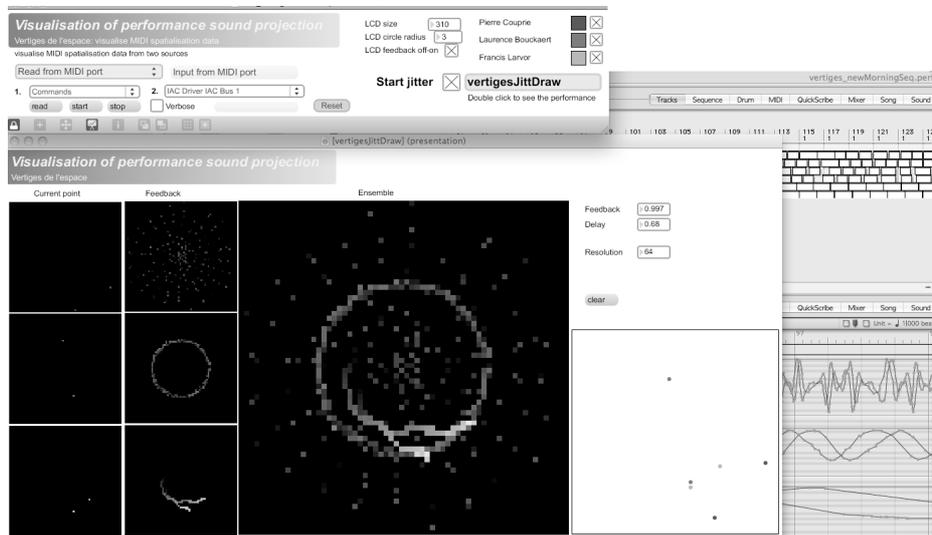


Figure 9. Patch de lecture des données de diffusion. La diffusion des trois musiciens est vue sous trois formes : (1) position instantanée des deux entrées audio [*Current point*] ; (2) position avec *feedback* pour suivi de trajectoire [*Feedback*] et (3) deux vues d'ensemble [*Ensemble*]. En arrière plan on aperçoit la séquence MIDI qui pilote ce *patch* et les données MIDI des pistes 1, 3 et 5.

4. CONCLUSIONS ET PERSPECTIVES

Ce projet est encore en cours de développement et nous ne présentons ici que les premiers résultats. Toutefois, les solutions trouvées par rapport à la mise en espace répondent à des besoins ressentis par les musiciens.

L'instrument est assez robuste et stable en situation de concert. Pour l'instant, notre priorité a été le mode de diffusion *trajectoire circulaire* que nous trouvons le plus adapté au projet *Vertiges de l'espace*. Des développements et améliorations sont en projet car celui-ci nous a permis de poser des questions lors de la performance électroacoustique en situation d'improvisation [1].

Par exemple, l'essai d'adaptation des figures d'espace de Vande Gorne nous a permis de réfléchir sur les exigences d'une performance de diffusion « interprétée » (dans le sens où il y a une musique déjà fixée) et d'une performance de diffusion « improvisée » (dans le sens où il faut réagir à « ce qui arrive ») en prenant compte des différences de propos et de position. La solution entamée ici, résultant de cette réflexion, a été de construire un agencement direct sur les niveaux de sortie au contraire des autres modes. De même pour une question pratique : comment transcrire vers huit paramètres de contrôle les gestes implicites par ces figures et comment préparer, enchaîner et déclencher ces figures dans un temps suffisamment rapide pour qu'elles puissent avoir une pertinence musicale.

Un autre aspect a été le besoin d'intégrer dans l'instrument des outils qui puissent ouvrir la performance à l'analyse dans un sens de transparence. Au-delà de la sauvegarde du fichier audio résultant, la sauvegarde des entrées audio ainsi que les données de commande de la diffusion et de la distribution spatiale représente un premier pas. Les modes de restitution de ces mêmes données représentent un aspect qui n'est pas négligeable car il a eu comme conséquence d'ouvrir notre travail à des questions et des possibilités à explorer prochainement. À savoir : comment améliorer l'interaction entre les musiciens produisant du son et celui qui le diffuse : quel meilleur rendu visuel des signaux d'entrée pour identification ? Est-ce que l'information visuelle de la mise en espace de chaque musicien lui apporte vraiment des informations pertinentes ?

Nous envisageons d'obtenir des réponses à ces questions dans le cadre de la poursuite de ce travail.

5. REMERCIEMENTS

Ce projet est développé dans le cadre d'une bourse de postdoctorat attribué par Fundação para a Ciência e a Tecnologia / Ministério da Ciência da Tecnologia e do Ensino Superior (Portugal).

6. RÉFÉRENCES

- [1] Couprie, P. ; Sousa Dias, A. « Vertiges de l'espace : analyse d'une performance électroacoustique improvisée », *Comment analyser l'improvisation - Colloque international*, Ircam, Paris, 12-13 février 2010.
- [2] Vande Gorne, Annette « L'interprétation spatiale. Essai de formalisation méthodologique », *DÉMéter - Revue électronique du Centre d'Etude des Arts Contemporains de l'Université de Lille-3*, 2002. <http://www.univ-lille3.fr/revues/demeter/interpretation/vandegorne.pdf> (10/02/2010)
- [3] Zicarelli, D. *et al.*, *MAX 5 : Documentation*. <http://www.cycling74.com/docs/max5/vignettes/intro/docintro.html> (15/02/2010)

Session 4

Notation et visualisation

PARTITIONS MUSICALES AUGMENTÉES

D. Fober, C. Daudin, S. Letz, Y. Orlarey

{fober,daudin,letz,orlarey}@grame.fr

Grame - Centre national de création musicale

RÉSUMÉ

Une *partition musicale augmentée* est une partition mettant en relation un objet musical symbolique avec différentes représentations de son interprétation. La partition musicale est à considérer au sens large, comme un objet graphique permettant de représenter un objet temporel. L'interprétation représente une instance sonore ou gestuelle particulière de la partition. Nous présenterons les fondements théoriques qui sous-tendent la *partition augmentée*, ainsi qu'une application sous forme d'*afficheur* mettant en oeuvre les solutions proposées.

1. INTRODUCTION

La notation musicale s'inscrit dans une longue histoire et a beaucoup évolué à travers les âges. Des neumes aux notations contemporaines, notre culture est riche de toutes les voies explorées pour représenter la musique. Des formes symboliques ou prescriptives de la notation aux représentations purement graphiques [5], la partition musicale reste en constante interaction avec le processus de création artistique.

Cependant, bien que l'on ait assisté à une explosion des représentations avec l'avènement de l'informatique musicale [7, 15, 11], la partition, destinée à l'interprète, n'a pas évolué en proportion des nouvelles formes musicales. En particulier, il y a un fossé important entre les musiques interactives et la manière statique de les représenter. L'interprète ne dispose souvent que d'une partition traditionnelle à laquelle s'ajoute une représentation électronique minimale de l'état du système, sous forme rudimentaire de compteur ou de lettres.

La notation musicale par ordinateur a toujours été et demeure une tâche complexe [2]. En dehors du monde commercial de l'édition de partition, les systèmes les plus aboutis, i.e. ceux qui vont au delà de la notation traditionnelle, proposent des approches de type gravure musicale [10, 14]. Bien que n'ayant pas l'ambition de produire des partitions mises en page, l'approche proposée par ENP [12] est certainement la plus ouverte à l'extension de la notation, notamment grâce au langage Lisp sur lequel s'appuie ENP. Toutefois aucune de ces approches ne permettent la prise en compte d'aspects dynamiques dans la partition.

A partir du constat que le temps est une propriété commune à tous les objets musicaux, nous proposons d'étendre la partition musicale à tout objet graphique possédant des

propriétés temporelles. Nous appellerons *partition musicale augmentée*, l'espace permettant de représenter, composer et manipuler des objets musicaux hétérogènes, aussi bien dans le domaine graphique que temporel. La définition des propriétés graphiques et temporelles requises définit dans le même temps toute une classe de partitions musicales.

Nous allons considérer des objets arbitraires (partitions musicales, images, texte, représentation du signal, graphiques vectoriels) comme candidats possibles d'une partition augmentée en leur donnant une position et une dimension temporelles nécessaires à leur statut de partition musicale, ce qui implique également de rendre les relations temporelles visibles.

L'organisation consistante de l'espace graphique par rapport à l'espace temporel est à la base de la notation musicale traditionnelle. Elle a toutefois été bousculée dans les cinquante dernières années [3] et particulièrement avec les langages informatiques pour la composition musicale, en raison notamment de la nature particulière de la composition algorithmique. Cependant, la question de cette consistance reste ouverte, y compris dans les recherches récentes [4] aussi bien que pour satisfaire les besoins de synchronisation de différents médias [1, 13].

La synchronisation dans le domaine graphique de médias hétérogènes soulève des questions de non-linéarité, discontinuité, non-bijectivité. Nous avons abordé ces questions par le biais de la *segmentation* et de la description de relations entre *segments*.

Enfin, la représentation de l'interprétation musicale, soit gestuelle ou sonore, est abordée avec un renversement de perspective : la représentation graphique d'un signal y est vue comme un *signal graphique*, c'est à dire comme un signal composite comportant toute l'information nécessaire à son rendu graphique. Cette approche, qui abstrait le calcul de la représentation, constitue un système ouvert et dynamiquement extensible.

Nous nous attacherons tout d'abord à décrire les principes de base des *mappings* ainsi que le contexte de leur utilisation dans le cadre de la partition augmentée. Nous expliquerons ensuite comment le système traite les signaux de l'interprétation pour en faire des *signaux graphiques*. Enfin nous présenterons l'*afficheur* de partitions augmentées en nous attachant plus particulièrement à son API de contrôle qui est une API de messages OSC[16].

2. RELATIONS TEMPORELLES ET GRAPHIQUES

Nous parlerons de *synchronisation temporelle dans le domaine graphique* pour faire référence à la représentation graphique des relations temporelles entre composants d'une partition. Notre expérience dans ce domaine [6, 9] nous a conduit à aborder la question de ces relations par le biais de la *segmentation* et de la description de relations entre *segments*. Nous utiliserons par la suite le terme *mappings* pour faire référence à ces relations.

Le rôle d'un *mapping* est d'établir des correspondances entre les *segments* de ressources différentes. Un *segment* est une zone contiguë d'une ressource. Ainsi que mentionné en introduction, ces ressources peuvent être de nature arbitraire (images, texte, notation musicale, signaux...). Ces mappings permettent typiquement de lier des positions graphiques, du temps musical et des positions dans des ressources audio.

Un mapping entre un enregistrement audio et une partition musicale permet par exemple d'exprimer une correspondance entre une position audio (exprimée en frames) et un temps musical (exprimé en subdivision de la noire). Un mapping entre une partition musicale et sa représentation graphique permet d'exprimer une correspondance entre une position exprimée en temps musical en une position graphique. La combinaison de ces mappings suffit alors à exprimer des relations entre toutes les ressources ayant une base temporelle.

Les sections qui suivent définissent un cadre général pour les notions de *segment*, *segmentation* et *mapping*. Ces définitions sont indépendantes de toute implémentation et de toute information spécifique aux différentes ressources. Elles sont suivies de cas concrets d'utilisation, implémentés dans le cadre de l'afficheur de partition augmentée.

2.1. Définitions

Nous allons tout d'abord introduire les notions de segments graphiques et temporels. Puis nous généraliserons ces définitions concrètes en une version abstraite et générale.

2.1.1. Segment temporel

Un *segment temporel* est défini comme un intervalle $i = [t_0, t_1[$ tel que $t_0 \leq t_1$.

Un intervalle $i = [t_0, t_1[$ est dit vide quand $t_0 = t_1$. Il sera noté \emptyset .

L'intersection de segments temporels est telle que communément définie : c'est le plus grand intervalle tel que

$$\forall i_m, \forall i_n, i_m \cap i_n := \{j \mid j \in i_m \wedge j \in i_n\}$$

2.1.2. Segment graphique

Un *segment graphique* g est défini comme un rectangle donné par deux intervalles $g = (i_x, i_y)$ où x est un intervalle sur l'axe des abscisses et y , sur l'axe des ordonnées.

Un segment graphique $g = (i_x, i_y)$ est dit vide quand $i_x = \emptyset$ ou $i_y = \emptyset$

L'opération d'intersection \cap entre segments graphiques est définie telle que :

$$\forall g = (i_x, i_y), \forall g' = (i'_x, i'_y), g \cap g' = (i_x \cap i'_x, i_y \cap i'_y)$$

2.2. Generalisation de la notion de segment

Nous allons étendre les notions de segment temporel et graphique à une définition plus générale de segment à n dimensions. Un segment de dimension n , noté s^n , est défini comme une liste de n intervalles $s^n = (i_1, \dots, i_n)$ où i_j est un intervalle sur la dimension j .

Un segment s^n de dimension n est dit vide quand $\exists i \in s^n \mid i = \emptyset$

L'intersection de segments de dimension n est définie comme la liste des intersections de leurs intervalles :

$$s_1^n \cap s_2^n = (i_1 \cap j_1, \dots, i_n \cap j_n) \quad (1)$$

où $s_1^n = (i_1, \dots, i_n)$ et $s_2^n = (j_1, \dots, j_n)$

2.3. Segmentation d'une ressource

Une ressource R de dimension n est *segmentable* quand elle peut être vue comme un segment S^n de dimension n . La segmentation d'une ressource R est l'ensemble des segments $Seg(R) = \{s_1^n, \dots, s_i^n\}$ tels que :

$$\begin{array}{ll} \forall i, j \in Seg(R) & i \cap j = \emptyset & \text{les segments sont disjoints} \\ \forall i \in Seg(R) & i \cap S^n = i & \text{tous les segments sont} \\ & & \text{inclus dans la ressource} \end{array}$$

2.4. Mapping

Un *mapping* est une relation entre *segmentations*.

Pour un mapping $M \subseteq Seg(R_1) \times Seg(R_2)$ nous définissons deux fonctions :

$$M^+(i) = \{i' \in Seg(R_2) \mid (i, i') \in M\} \quad (2)$$

qui donne l'ensemble des segments de R_2 associés au segment i de R_1 .

et la fonction inverse :

$$M^-(i') = \{i \in Seg(R_1) \mid (i, i') \in M\} \quad (3)$$

qui donne l'ensemble des segments de R_1 associés au segment i' de R_2 .

Ces fonctions sont définies sur un ensemble de segments comme l'union des mappings de chaque segment :

$$M^+(\{i_1, \dots, i_n\}) = \{M^+(i_1) \cup M^+(i_2) \dots \cup M^+(i_n)\} \quad (4)$$

et

$$M^-(\{i_1, \dots, i_n\}) = \{M^-(i_1) \cup M^-(i_2) \dots \cup M^-(i_n)\} \quad (5)$$

2.5. Composition de mappings

La composition de mappings se fait de manière naturelle. Pour les mappings $M_1 \subseteq \text{Seg}(R_1) \times \text{Seg}(R_2)$ et $M_2 \subseteq \text{Seg}(R_2) \times \text{Seg}(R_3)$:

$$(M_1 \circ M_2)^+(i) = M_2^+(M_1^+(i)) \quad (6)$$

De manière similaire :

$$(M_1 \circ M_2)^-(i') = M_1^-(M_2^-(i')) \quad (7)$$

2.6. Segmentations et mappings d'une partition augmentée

Toutes les ressources qui composent la partition augmentée ont en commun une dimension graphique et une dimension temporelle exprimée en temps musical. Elles sont donc toutes *segmentables* dans l'espace graphique et dans l'espace temps musical. Par ailleurs, chaque type de ressource est *segmentable* dans un espace qui lui est propre : espace linéaire des frames audio pour un signal audio, espace à 2 dimensions organisé en lignes/colonnes pour du texte, etc.

La table 2.6 décrit les segmentations et mappings utilisés par les différents types de composants. Les mappings sont indiqués par des flèches (\leftrightarrow). Les segmentations et mappings en *italique* sont automatiquement calculés par le système, ceux en **gras** sont fournis de manière externe.

C'est la composition de ces mappings qui permet d'adresser et de synchroniser les composants d'une partition augmentée aussi bien dans l'espace graphique que temporel. La segmentation de l'espace temporel constitue la colonne vertébrale du système : l'espace temporel est commun à tous les composants et permet, par le biais de la composition, de mettre en relation toutes les autres segmentations.

type	segmentations et mappings requis
texte	<i>graphic</i> \leftrightarrow text \leftrightarrow temps
partition	<i>graphic</i> \leftrightarrow <i>temps enroulé</i> \leftrightarrow <i>temps</i>
image	<i>graphic</i> \leftrightarrow pixel \leftrightarrow temps
graph. vectoriel	graphic \leftrightarrow temps
signal	<i>graphic</i> \leftrightarrow échantillon \leftrightarrow temps

Table 1. Segmentations et mappings pour chaque type de composant.

Pour un élément de type *texte* par exemple, cela revient à établir la composition suivante :

$$(\text{graphic} \leftrightarrow \text{text}) \circ (\text{text} \leftrightarrow \text{temps})$$

Le temps fait référence au temps musical métronomique (i.e. relatif à un tempo). La relation "*temps enroulé* \leftrightarrow *temps*" est nécessaire pour traiter les sections d'une partition avec reprises ou avec des sauts (à la coda, au signe...).

2.7. Exemples de synchronisation

Soit deux composants A et B d'une partition, ainsi que leurs segmentations graphiques et temporelles :

$$\text{Seg}(A_g), \text{Seg}(A_t), \text{Seg}(B_g), \text{Seg}(B_t).$$

De plus, B possède une segmentation intermédiaire $\text{Seg}(B_t)$ exprimée dans les coordonnées de son espace local (par exemple : frames pour un signal audio). Les mappings

$$M_A \subseteq \text{Seg}(A_g) \times \text{Seg}(A_t)$$

$$\text{et } M_B \subseteq \text{Seg}(B_t) \times \text{Seg}(B_t)$$

donnent la correspondance entre l'espace graphique et temporel pour A et entre l'espace temporel et local pour B .

Lors de la synchronisation et afin de spécifier quelle position graphique doit être utilisée comme base, nous avons introduit une relation maître/esclave entre composants : un esclave est contraint à la position de son maître.

2.7.1. Alignement graphique de positions temporelles

Considérons que B est l'esclave de A et que nous voulons aligner graphiquement B et A à un temps t . Soit s , le segment de A qui contient la date t , alors le segment graphique correspondant est :

$$M_A^-(s) = \{g_i \in \text{Seg}(A_g) \mid (g, s) \in M_A\}$$

Quand $M_A^-(s)$ est constitué d'un seul segment, la position de B peut être calculée par simple interpolation linéaire i.e. :

$$(x_B, y_B) = (g_{x0} + (g_{x1} - g_{x0}) \cdot \delta, g_{y0})$$

où g_{x0} et g_{x1} sont les coordonnées x de début et de fin du segment graphique et $\delta = (t - s_0)/(s_1 - s_0)$.

y_B est fixé de manière arbitraire à g_{y0} . Dans la pratique, il est contrôlé par le mode de synchronisation (*over*, *above*, *below*).

Quand $M_A^-(s)$ contient plusieurs segments, l'opération peut-être répétée pour chacun de ces segments.

2.7.2. Alignement de segments graphiques

Le principe de base de l'alignement par segment consiste, pour chaque segment du composant maître, à prendre le segment correspondant de l'esclave, exprimé dans son espace local, et à faire le rendu de ce segment dans l'espace graphique correspondant du maître.

Si $\text{Seg}(A_t) = \text{Seg}(B_t)$, l'opération peut être vue comme une composition de mappings :

$$M_A \circ M_B \subseteq \text{Seg}(A_g) \times \text{Seg}(B_t)$$

La figure 1 donne l'exemple d'une même image alignée sur une partition à des dates différentes.

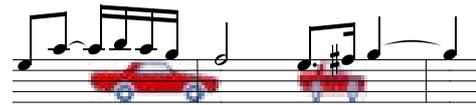


Figure 1. La même image d'une voiture synchronisée à des dates différentes. L'image a la durée d'une noire, elle est étirée sur la longueur du segment graphique de la partition qui lui correspond.

3. REPRÉSENTATION DE L'INTERPRÉTATION

Le travail sur la représentation de l'interprétation s'appuie sur des expériences de visualisation du jeu instrumental réalisées dans un contexte pédagogique [8]. Dans ce cadre, nous avons développé un moteur de rendu graphique prenant des signaux ainsi qu'un type de représentation en entrée pour calculer l'image correspondant au type désiré et aux valeurs des signaux. L'inscription statique des types de représentation supportés dans le moteur de rendu constitue une des limitations importante de cette approche : elle implique une modification du moteur de rendu pour chaque nouveau type de représentation.

Dans le cadre du travail sur la partition augmentée, nous avons souhaité lever cette limitation et définir un système extensible dynamiquement. A cet effet, la représentation graphique d'un signal est vue comme un *signal graphique*, c'est à dire comme un signal composite comportant toute l'information nécessaire à son rendu graphique.

3.1. Des signaux graphiques

Nous définissons un signal graphique comme un signal composite constitué :

- d'un signal y : les coordonnées en y du graphique
- d'un signal h : qui décrit l'épaisseur du graphique à la position y
- d'un signal c : qui décrit la couleur du graphique à la position y

Pour simplifier, nous supposons que l'espace des couleurs décrit par c n'a qu'une dimension. La figure 2 donne un exemple de ces paramètres dans l'espace graphique.

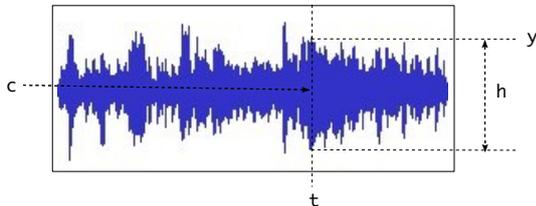


Figure 2. Paramètres graphiques d'un signal.

Considérons maintenant que nous disposons d'un signal S défini comme une fonction du temps :

$$f(t) : \mathbb{R} \rightarrow \mathbb{R}^3 = (y, h, c) \mid y, h, c \in \mathbb{R}$$

alors ce signal contient toute l'information pour être dessiné directement i.e. sans calcul supplémentaire.

Une autre manière de voir revient à considérer le système comme un *oscilloscope* qui prendrait les composantes d'un signal graphique en entrée.

3.2. Composition de signaux

Afin de construire des signaux composites pouvant servir de signaux graphiques, nous avons introduit une opération de parallélisation de signaux.

Soit \mathbb{S} , l'ensemble des signaux $s : \mathbb{N} \rightarrow \mathbb{R}$. Nous définissons l'opération *parallèle* '/' comme :

$$s_1/s_2/\dots/s_n : \mathbb{S} \rightarrow \mathbb{S}^n \mid s_i \in \mathbb{S} \quad (8)$$

La fonction du temps d'un signal parallèle $s^n \in \mathbb{S}^n$ est la mise en parallèle de la fonction du temps de chaque signal :

$$f(t) = (f_0(t), f_1(t), \dots, f_n(t)) \mid f_i(t) : \mathbb{N} \rightarrow \mathbb{R} \quad (9)$$

3.3. Types de signaux parallèles

Pour les besoins de l'afficheur de partition augmentée, nous avons défini plusieurs types de signaux parallèles :

- le type *signal de couleur* qui utilise le modèle HSBA [hue, saturation, brightness, transparency] de représentation des couleurs :

$$c ::= \overrightarrow{(h, s, b, a)} \mid h, s, b, a \in \mathbb{R}$$

- le type *signal graphique* qui comporte un signal y , un signal d'épaisseur th suivi des 4 composantes du signal de couleur :

$$g ::= \overrightarrow{(y, th, h, s, b, a)} \mid y, th, h, s, b, a \in \mathbb{R}$$

- en enfin le type *signaux graphiques parallèles* qui permet de composer plusieurs *signaux graphiques* en parallèle :

$$g^n ::= \overrightarrow{g} \mid g \in \mathbb{R}^6$$

3.4. Exemples de signaux graphiques

A des fins de validation du modèle, nous allons décrire plusieurs types de représentations qui étaient implémentées de manière statique dans notre approche précédente.

3.4.1. Représentation de la hauteur de note

Consiste à représenter les hauteurs de note à partir de leur fréquence fondamentale, sur l'axe des y (figure 3).

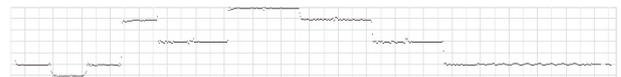


Figure 3. Graphique des hauteurs de notes.

Le signal graphique correspondant s'exprime de la manière suivante :

$$g = S_{f_0} / k_t / k_c$$

où S_{f_0} : fréquence fondamentale exprimée en 1/2 tons
 k_t : signal d'épaisseur constant
 k_c : signal de couleur constant

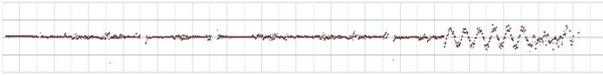


Figure 4. Graphique de justesse.

3.4.2. Représentation de la justesse

Consiste à représenter la différence entre une fréquence fondamentale et une fréquence de référence (figure 4).

Le signal graphique correspondant s'exprime de la manière suivante :

$$g = S_{f_0} - S_{f_r} / k_t / k_c$$

où S_{f_0} : fréquence fondamentale exprimée en 1/2 tons

S_{f_r} : fréquence de référence exprimée en 1/2 tons

k_t : signal d'épaisseur constant

k_c : signal de couleur constant

3.4.3. Représentation sous forme d'enveloppe

Consiste à utiliser les valeurs RMS d'un signal pour contrôler l'épaisseur du graphique. (figure 5).



Figure 5. Graphique d'articulations.

Le signal graphique correspondant s'exprime de la manière suivante :

$$g = k_y / S_{rms} / k_c$$

où k_y : signal y constant

S_{rms} : signal RMS

k_c : signal de couleur constant

3.4.4. Combinaison d'enveloppe et de hauteur

Consiste à utiliser la fréquence fondamentale et les valeurs RMS d'un signal dessiner un signal d'articulation modulé en fonction de la hauteur de note. (figure 6).



Figure 6. Combinaison de hauteur et d'enveloppe.

Le signal graphique correspondant s'exprime de la manière suivante :

$$g = S_{f_0} / S_{rms} / k_c$$

où S_{f_0} : fréquence fondamentale exprimée en 1/2 tons

S_{rms} : signal RMS

k_c : signal de couleur constant

3.4.5. Combinaison d'harmoniques et de hauteur

Consiste à combiner la fréquence fondamentale et l'intensité des premiers harmoniques. (figure 7). Chaque harmonique est représenté avec une couleur différente.



Figure 7. Combinaison de la hauteur et de l'enveloppe des harmoniques.

Le signal graphique correspondant s'exprime en plusieurs étapes. On construit d'abord le graphique de la fréquence fondamentale comme précédemment (voir section 3.4.4) :

$$g_0 = S_{f_0} / S_{rms0} / k_{c0}$$

où S_{f_0} : fréquence fondamentale exprimée en 1/2 tons

S_{rms0} : valeurs RMS du signal f_0

k_{c0} : signal de couleur constant

Puis on construit le graphique de l'harmonique 1 :

$$g_1 = S_{f_0} / S_{rms1} + S_{rms0} / k_{c1}$$

S_{rms1} : valeurs RMS de l'harmonique 1

k_{c1} : signal de couleur constant

Puis de l'harmonique 2 :

$$g_2 = S_{f_0} / S_{rms2} + S_{rms1} + S_{rms0} / k_{c2}$$

S_{rms2} : valeurs RMS de l'harmonique 2

k_{c2} : signal de couleur constant

etc.

Pour finalement les combiner en un signal graphique parallèle :

$$g = g_2 / g_1 / g_0$$

4. L'AFFICHEUR DE PARTITION AUGMENTÉE

Les travaux sur la partition augmentée ont été réalisés dans le cadre du projet Interlude¹. Ils ont fait l'objet d'une implémentation sous forme de librairie C++ - la librairie Interlude - ainsi que sous forme d'un afficheur de partitions, construit au dessus de cette librairie. Cet afficheur n'a pas d'interface utilisateur car il a été conçu pour être contrôlé par des messages OSC i.e. par des applications externes (typiquement Max/MSP ou Pure Data).

4.1. Format général des messages

Un message OSC/Interlude est constitué d'une adresse OSC, suivie d'un nom de message, suivi de 0 à n paramètres. Le nom du message peut être vu comme le nom d'une méthode de l'objet identifié par l'adresse OSC. Cette adresse peut comporter des expressions régulières désignant un ensemble de destinataires.

1. ANR-08-CORD-010

L'espace d'adressage OSC inclut des nodes statiques prédéfinies :

- /ITL correspond à l'afficheur Interlude.
- /ITL/scene correspond à la scène de rendu graphique, dans les faits, l'adresse de la partition augmentée.

La section qui suit présente un exemple de messages OSC décrivant une partition comportant des éléments synchronisés. La liste des messages correspond strictement au format de fichier d'une partition augmentée. Notez que l'exemple ci-dessous est statique mais qu'une interaction dynamique avec la partition est toujours possible, par exemple en déplaçant des objets dans le temps en leur envoyant des messages `date` ou `clock` (d'une sémantique similaire aux message `clock` MIDI).

4.2. Exemple de synchronisation imbriquée

Cet exemple met 3 composants en oeuvre : le premier est maître du second, qui est maître du troisième. Les lignes qui commencent par '#' sont des commentaires imbriqués avec les messages.

```
# crée une partition basée sur une image
/ITL/scene/score set img "score.jpg"
# décrit son mapping entre les espaces
# temps et graphique
/ITL/scene/score mapf "score.map"

# crée un texte à partir d'un fichier
/ITL/scene/text set txtf "comment.txt"
# change l'échelle du texte
/ITL/scene/text scale 3.0
# ainsi que sa couleur
/ITL/scene/text color 0 0 240 255
# met le texte en 'avant' (z order)
/ITL/scene/text z 0.5
# et décrit le mapping du texte au temps
/ITL/scene/text mapf "comment.map"

# crée un cercle (graphique vectoriel)
/ITL/scene/ball set ellipse 0.2 0.2
# le mets en 'avant'
/ITL/scene/ball z 0.4
# change la couleur du cercle
/ITL/scene/ball color 250 50 0 255

# change la date de tous les objets
# la date est exprimée sous forme
# de rationnel (1 valant une ronde)
/ITL/scene/* date 4 1

# rend le texte esclave de la partition
/ITL/scene/sync text score
# rend le cercle esclave du texte
/ITL/scene/sync ball text
```

Le résultat correspondant est donnée par la figure 8.

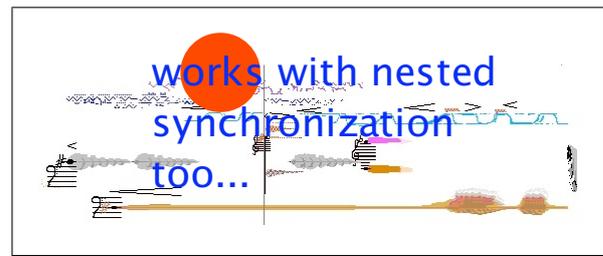


Figure 8. Une partition avec synchronisation imbriquée.

5. CONCLUSION

La méthode proposée pour la synchronisation graphique d'objets arbitraires en fonction de leurs relations temporelles combine les avantages de la simplicité et de la flexibilité : une grande diversité de comportements peut être décrite par la simple définition de segmentations et de mappings. Cette méthode est décrite indépendamment de toute implémentation.

La simplicité et la flexibilité caractérisent également l'approche mise en oeuvre pour intégrer la représentation de l'interprétation au sein de la partition. Elle consiste à abstraire le calcul de la représentation du moteur de rendu, ce qui en fait un système ouvert et dynamiquement extensible.

Il y a de nombreux domaines d'application potentiels au concept de partition augmentée - domaine pédagogique, jeux,... - mais nous souhaitons également que ce travail puisse servir les besoins de la création contemporaine et notamment de ses nouvelles formes telles que les musiques interactives.

Remerciements

Cette recherche a été menée dans le cadre du projet Interlude qui est soutenu par l'Agence Nationale pour la Recherche [ANR-08-CORD-010].

6. REFERENCES

- [1] D. Baggi and G. Haus, "IEEE 1599 : Music encoding and interaction," *COMPUTER*, vol. 42, no. 3, pp. 84–87, March 2009.
- [2] D. A. Bird, "Music notation by computer," Ph.D. dissertation, Indiana University, 1984.
- [3] A. Boucourechliev, *Archipel 1*. Universal, 1967.
- [4] J. Bresson and C. Agon, "Scores, programs, and time representation : The sheet object in openmusic," *Computer Music Journal*, vol. 32, no. 4, pp. 31–47, 2008.
- [5] E. Brown, *December 1952*. AMP/G.Schirmer, 1952.
- [6] Y. Chapuis, D. Fober, S. Letz, Y. Orlarey, and C. Daudin, "Annotation de partitions musicales dynamiques," in *Actes des Journées d'Informatique Musicale JIM'07 - Lyon*, Grame, Ed., 2007, pp. 18–27.

- [7] R. B. Dannenberg, "Music representation issues, techniques and systems," *Computer Music Journal*, vol. 17, no. 3, pp. 20–30, 1993.
- [8] C. Daudin, D. Fober, S. Letz, Y. Orlarey, and Y. Chappuis, "Visualisation du jeu instrumental," in *Actes des Journées d'Informatique Musicale JIM'07 - Lyon*, Grame, Ed., 2007, pp. 64–72.
- [9] D. Fober, S. Letz, and Y. Orlarey, "Vemus - une école de musique européenne virtuelle," in *Actes des Journées d'Informatique Musicale JIM'07 - Lyon*, Grame, Ed., 2007, pp. 57–63.
- [10] K. Hamel, "NoteAbility, a comprehensive music notation system." in *Proceedings of the International Computer Music Conference.*, 1998, pp. 506–509.
- [11] W. B. Hewlett and E. Selfridge-Field, Eds., *The Virtual Score ; representation, retrieval and restoration*, ser. Computing in Musicology. MIT Press, 2001.
- [12] M. Kuuskankare and M. Laurson, "Expressive notation package," *Computer Music Journal*, vol. 30, no. 4, pp. 67–79, 2006.
- [13] L. A. Ludovico, "An XML multi-layer framework for music information description," Ph.D. dissertation, Università degli Studi di Milano, 2006.
- [14] H.-W. Nienhuys and J. Nieuwenhuizen, "LilyPond, a system for automated music engraving," in *Proceedings of the XIV Colloquium on Musical Informatics*, 2003.
- [15] E. Selfridge-Field, Ed., *Beyond MIDI : the handbook of musical codes*. MIT Press, 1997.
- [16] M. Wright, *Open Sound Control 1.0 Specification*, 2002.

L'ACOUSMOGRAPHE 3

Emmanuel Favreau
Ina-GRM
efavreau@ina.fr

Yann Geslin
Ina-GRM
ygeslin@ina.fr

Adrien Lefèvre
Ina-GRM
alefevre@ina.fr

RÉSUMÉ

L'Acousmographe est un outil d'analyse et représentation des musiques électroacoustiques, et par extension, de tout phénomène sonore enregistré. Son développement est un projet au long terme né du besoin pour les compositeurs et musicologues de disposer d'outils de transcription des musiques non écrites, par des représentations graphiques et annotations textuelles synchronisées à l'écoute et aux représentations techniques usuelles du signal (amplitude - analyse spectrale).

Dans sa troisième génération, initiée en 2003 et largement distribué depuis 2005, l'Acousmographe est multi plateforme, multi canal et comprend notamment des capacités techniques avancées telles l'analyse quasi illimitée, la représentation synchrone ou asynchrone de multiples plans graphiques, ou l'analyse par ondelettes du signal [6]. De nombreux développements ultérieurs sont envisagés, grandement facilités par la structure en plugin du logiciel. L'Acousmographe 3 est réalisé principalement par Adrien Lefèvre, avec les contributions d'Emmanuel Favreau, Yann Geslin, et les stagiaires accueillis au GRM.

1. INTRODUCTION

La nécessité de représenter le sonore, l'entendu, a accompagné dès ses débuts la production de musique électroacoustique, et ceci tout particulièrement au Groupe de Recherches Musicales : les premières inventions de Pierre Schaeffer sont élaborées à partir de sons préenregistrés. Dans cette configuration créative, le compositeur échappe par nature au dogme classique de la musique occidentale : écrire puis interpréter. Néanmoins, l'absence de support écrit se fait cruellement ressentir dans le travail. Pierre Schaeffer et l'équipe qui forma progressivement le GRM durent dès leurs débuts faire face à plusieurs nécessités :

- La première est d'obtenir une trace du sonore pour confronter une analyse objective à la perception subjective de l'objet entendu. Ceci est réalisé à l'époque par l'emploi assidu du bathygraphe [3], outil permettant d'obtenir un relevé thermo-imprimé de l'amplitude du signal (avec une constante temporelle assez petite) ; et du sonographe, cet appareil donnant une analyse spectrale relativement grossière, et de quelques secondes seulement. Ces outils technologiques qui nous font sourire aujourd'hui de par leurs limitations frustrantes, fondaient la référence et la preuve

scientifique du discours phénoménologique, de Schaeffer, mais d'autres également, quand on considère la page de représentation sonographique de l'œuvre de Karlheinz Stockhausen, produite pour fournir une référence indiscutable de contrôle en vue de réalisations ultérieures de l'œuvre.

- La deuxième nécessité vient de la difficulté à mémoriser le son dans sa temporalité ; la fugacité du phénomène sonore rend particulièrement subjective son inscription dans la mémoire. De l'ébauche d'une réalisation à la projection finale en concert, le compositeur a besoin de fixer son travail, ses intentions, et il en vient tout naturellement à utiliser des formes d'annotations personnelles. Ces annotations, parfois de nature textuelle, mais le plus souvent de nature graphique, se placent à mi-chemin entre une forme réductrice de la notation musicale classique et un système intuitif de représentation hauteur - temps - timbre. Cette représentation est extrêmement primitive, schématique, non codée, mais tout le monde s'accorde à l'accepter à défaut de système et code préexistant et commun. Notons également que la nécessité pendant des années de répondre à l'exigence de produire une « partition » de l'œuvre pour son dépôt à la Sacem, afin qu'elle puisse s'élever au rang d'œuvre musicale protégeable dans ses droits, a encouragé cette pratique tout d'abord intime et lui a donné un statut très usurpé de représentation autorisé de l'œuvre.

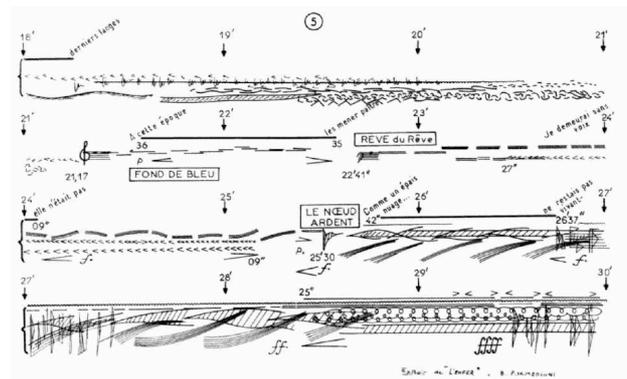


Figure 1. Bernard Parmegiani *L'Enfer* transcription manuelle par l'auteur

- La troisième nécessité vient de la difficulté à produire un discours sur une œuvre temporelle sans matérialisation de cette temporalité. Les musicologues, fondant habituellement leur analyse sur le repérage des découpages de l'œuvre, se sont dès le départ appuyés sur

les graphes sommaires des compositeurs, ou en ont produits personnellement. Ainsi par exemple, dans les années 70-80, l'admission définitive en composition électroacoustique au Conservatoire de Paris était assujettie à la production en fin de premier trimestre d'une « partition » (on dira aujourd'hui transcription) d'une œuvre imposée du répertoire, que les étudiants réalisaient laborieusement sur papier à l'aide d'un chronomètre, d'un magnétophone partagé quelques heures durant au casque, et de l'usage vaillant de la lecture en demi vitesse pour améliorer la précision temporelle du relevé.

De ces usages vient le désir d'une meilleure efficacité de la représentation, de la facilitation du travail par les moyens informatiques, ou même d'une normalisation utopique.

2. L'ACOUSMOGRAPHE, UN PROJET DE LONGUE HALEINE

2.1. Précursions et prémisses

Qu'ils aient réalisé leurs travaux soit à l'aide des appareils cités précédemment ou par l'usage intensif du crayon à dessin, les chercheurs, musicologues et compositeurs ont bien souvent ressenti le caractère arbitraire et superficiel du dispositif, à opposer aux représentations savantes, par courbes, graphiques orthogonaux et tableau de réalisation produits à Cologne et dans les pays de l'Est. Quelques tentatives pour élever au rang de pratique recommandable ces graphiques ont été explorées. La plus notable est la « partition animée » réalisée par Jacques Vidal, François Delalande, etc. dans les années 70, produite par l'animation (prise en caméra de banc-titre image par image) d'une représentation succincte manuelle... Ce film, peu exploité, probablement parce que le report du négatif en positif a fait perdre la partie gauche de chaque page, soit 2 à 3 secondes de représentation, quasi oublié, montre les attentes en un système dynamique de la représentation ; laissant penser que les insuffisances des représentations tenaient plus à leur mode de fixité sur papier, qu'aux modalités du système même.

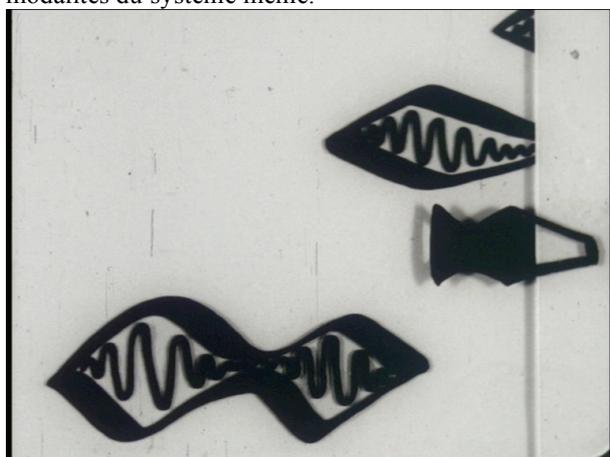


Figure 2. Pierre Schaeffer - Partition animée, par F. Delalande et J. Vidal - film 16mm 1972 ©Ina

Vers 1975, François Delalande et François Bayle chargent deux jeunes polytechniciens en stage de modifier le fonctionnement du Sonographe pour permettre l'analyse de durées sonores d'environ trente secondes¹, supérieures aux 2,4 secondes imposées par l'appareil [5].

Si le premier outil de représentation graphique informatique au GRM est la GT40 monochrome vectorielle, celui-ci sert en 1978 à représenter l'onde du signal avec une définition microscopique, en vue de la réalisation d'un éditeur audio ultra précis. Le second outil informatique graphique d'importance est l'écran bitmap 16 couleurs du système Syter (1984). Jean-François Allouis réalisera une version éphémère de représentation sonographique du signal, montrée dans un dispositif truqué faisant croire à la compréhensibilité par le système d'un signal polyphonique. La puissance relativement limitée de Syter ne permit pas d'intégrer ce sonographe tel quel.

Citons enfin le travail de Denis Dufour pour les besoins de sa classe de composition à Lyon, patient, tant par l'analyse que par l'exploitation extrême des capacités graphiques limitées de son premier ordinateur personnel, un MSX de Sony. Dufour a réalisé l'analyse et la représentation de plusieurs œuvres du répertoire, dont nous avons aujourd'hui la trace imprimée, ainsi que les fichiers originaux - hélas désormais inexploitable. L'utilisation de l'informatique, sous une forme encore assez frustrante, laissait espérer une rationalisation des codes de représentation. On y voit un des premiers exemples de la frise synoptique, désormais si prisée des transcripteurs, par un emploi astucieux de vignettes réduites de chaque page recollées en un conducteur.

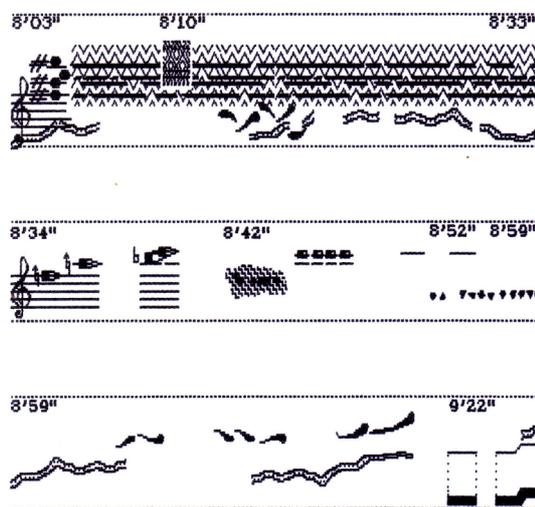


Figure 3. Ivo Malec : Luminétudes par Denis Dufour, transcription sous MSX (1988)

¹ Le Sonographe du GRM, probablement un Kay Electric, était constitué d'un cylindre accouplé à un disque magnétique enregistreur d'une durée de 2,4 secondes. Un stylet gravait sur un papier enroulé sur le cylindre l'amplitude détectée par un filtre glissant progressivement du grave à l'aigu. Le processus de mesure durait approximativement 20 minutes en de très nombreuses rotations (500 à 600).

2.2. L'Acousmographe 1 (1989-1998)

En 1988, le passage bref d'un chercheur rompu au travail graphique et à la microinformatique, Olivier Koechlin (alors que nos chercheurs patentés étaient experts en traitement du signal), offre l'opportunité à François Bayle de demander la réalisation d'une ébauche d'outil d'aide à la représentation. Ce sera l'Acousmographe 1 (version attribuée beaucoup plus tard), réalisé sur Macintosh et sous Hypercard puis Supercard, et avec les premières cartes audionumériques d'alors, Dyaxis (Studer) puis Audiomedia (Digidesign) en 1989 [10]. Ce premier projet suscite un engouement sans précédent des compositeurs, mais aussi des musicologues, François Delalande en première ligne, engouement à la hauteur de la déception qui suivra, car le logiciel ne sera jamais totalement stabilisé : il avait notamment la fâcheuse propension à planter et perdre le travail réalisé au bout de quelques minutes de musique représentée, c'est-à-dire malheureusement après de nombreuses journées de travail. Le logiciel offrait pour la première fois un outil adapté au son, notamment l'écoute quasi synchrone à l'image, la couleur, une gamme d'objets graphiques variée, une représentation de l'amplitude et de la FFT [1]. Enfin, dernière qualité – repérée tardivement, l'emploi de la « toolbox graphique » Apple standard, qui permettait l'export dans d'autres logiciels en vue d'une réutilisation des objets graphiques.

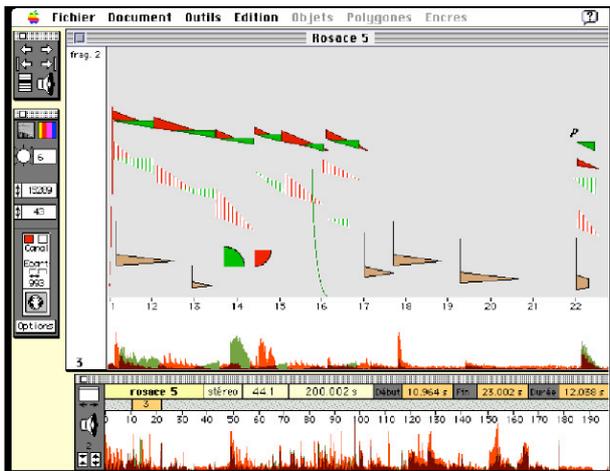


Figure 4. l'Acousmographe, première version

L'outil a servi de base à quelques-unes des représentations du cd-rom expérimental « *Les Musicographies* » (1993-95) [2]. Ce produit exploratoire proposait plusieurs modalités de représentations temporelles, tant de musiques électroacoustiques que de musiques non écrites. La mise en forme très avancée, sur le plan de la représentation et sur la programmation interactive, laissait penser que l'Acousmographe permettait aisément ces représentations, ce qui était un peu abusif.

2.3. L'Acousmographe 2 (1998-2005)

Le succès et la déception du premier Acousmographe donnèrent lieu à la production du deuxième Acousmographe (version 1.2 nommée désormais v2). La demande émanait du ministère de l'Éducation Nationale, notamment de son inspecteur général Vincent Maestracci, qui voyait dans les potentialités de représentations hors système de notation musicale traditionnelle un intérêt évident pour les professeurs de musique, leur permettant d'asseoir un discours sur des représentations du sonore adaptées à toute musique enregistrée. La demande fut par ailleurs celle d'un outil sur ordinateur PC plutôt que Macintosh. Cette version fut développée par Didier Bultiauw, puis Emmanuel Favreau avec l'aide de Mathieu Rainaud et Jean-Baptiste Thiébaud. La version sur Macintosh en fut donc moins finalisée que celle pour PC.

Les propriétés principales en sont le multicouche, l'export et l'impression du résultat, l'export textuel des occurrences d'objets, l'insertion d'un dispositif d'analyse automatique – placement de marqueurs, etc.

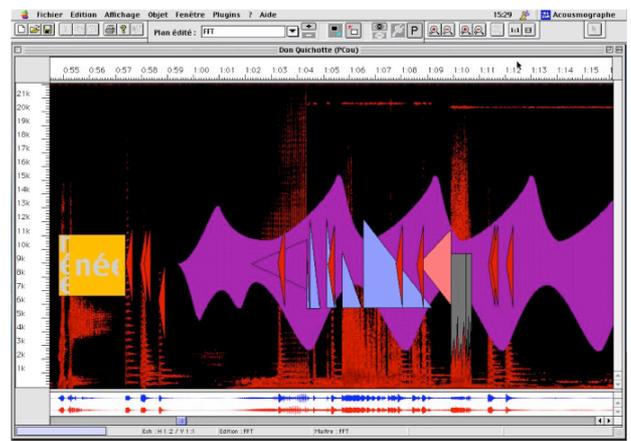


Figure 5. l'Acousmographe deuxième version

La majeure partie des représentations proposées dans le cd-rom « *La Musique Electroacoustique* » (2000) donne une bonne palette des possibilités [8]. Toutefois, là encore, ce cd-rom fait croire qu'il est facile de réaliser une acousmographie aboutie et de grande qualité graphique et mise en page, ce qui n'est pas exact. Par ailleurs, certains « acousmologues » préférèrent opter pour l'emploi de logiciels graphiques commerciaux majeurs tels ceux d'Adobe, soulignant que les possibilités graphiques de l'Acousmographe étaient différentes et restreintes.

Après de bons et loyaux services de cette version, la complexité de la maintenance « cross-plateform » par des intervenants successifs pèse dans la viabilité du logiciel, tandis que la demande d'une version plus moderne et plus évoluée devient pressante. Le Ministère de l'Éducation Nationale encourage une fois encore le projet, ce qui permet notamment la sous-traitance du travail à l'extérieur, en faisant appel aux compétences d'Adrien Lefèvre, alors développeur indépendant.

3. L'ACOUSMOGRAPHE 3

3.1. Le projet, ses contraintes

Ce projet de l'Acousmographe 3 débute en juillet 2003, lorsque l'Ina-GRM en passe la commande à Adrien Lefèvre, pour une refonte radicale de l'environnement. Dès le départ, la demande porte sur un projet ambitieux, un logiciel doté d'une interface utilisateur et d'une ergonomie proches de celles des logiciels du commerce. Les éléments clés caractérisant cette application peuvent se résumer en quelques points: dans le plan, la visualisation de la forme d'onde et du spectrogramme d'un son (guide impartial pour l'analyse musicale) donnant les repères temps-fréquences pour placer des symboles divers et variés regroupés sur des calques. L'utilisateur doit pouvoir visualiser ces calques dans différentes configurations: superposés ou séparément, synchronisés temporellement ou non, à différents niveaux de zoom pour chacun ou à un même niveau pour l'ensemble. Pour couronner le tout, le logiciel doit supporter l'analyse et la visualisation de sons multipistes, avec forme d'onde et spectrogramme pour chaque piste.

3.2. Propriétés et singularités

La synthèse de l'ensemble de ces contraintes en une réponse simple ne fut pas une mince affaire. Premièrement, étant donnée la diversité des symboles potentiellement existants pour représenter un son, un timbre; n'étant de limite que celle de l'imaginaire humain, on ne pouvait imposer dès le départ aucune forme graphique de quelque nature que ce soit. L'Acousmographe fut donc conçu comme un logiciel ouvert, ne comportant à sa base qu'un système de commande central - un contenant - dont le contenu (les symboles) serait pris en charge par des plugins. Deuxièmement, les différents modes de visualisation des calques, formes d'onde et spectrogrammes étaient difficilement conciliables en une seule fonctionnalité simple. C'est pourtant une solution compacte qui fut apportée; après un an de réflexions et de dialogues, l'Acousmographe laissait entrevoir son allure finale: une acousmographie était représentée au sein d'une fenêtre standard divisée horizontalement en cadres. Cette solution répondait à l'ensemble des contraintes de manière simple et efficace, c'est certainement celle qui offre le maximum de compacité fonctionnelle.

3.3. Quelques mots du code

Dans sa réalisation actuelle, l'Acousmographe gère une structure de donnée que l'on pourrait joliment nommer « un arbre à objets ». D'une part, la structure de donnée principale est un arbre binaire doublement chaîné, sur chaque noeud de cet arbre viennent « s'accrocher » des objets, d'autre part ces objets sont instanciés et gérés par les plugins ou les classes internes du logiciel. Ainsi, du plus haut niveau de la hiérarchie

(le calque maître), en passant par les groupes de calques, contenant des calques, contenant eux-mêmes des groupes de symboles, jusqu'aux symboles; les traitements sur cette structure récursive sont homogènes et unifiés (comme pour le copier-coller ou le *wf q* illimité par exemple).

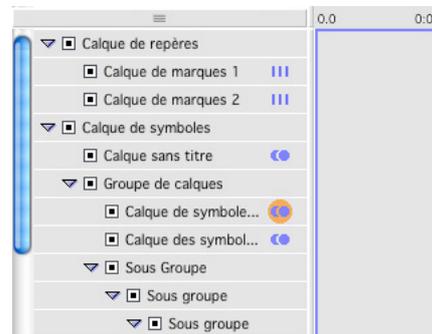


Figure 6. Calques hiérarchiques

Un autre aspect essentiel du logiciel est sa gestion des spectrogrammes de longue durée (théoriquement jusqu'à trois heures); grâce à un gestion bas niveau des pages mémoire (*o go qt{ o crrkpi*), et une répartition astucieuse des données brutes du spectrogramme sur les mémoires de masse, l'Acousmographe permet de visualiser plus d'une heure de signal sur un écran standard. Cette fonctionnalité s'est avérée précieuse - entre autres - pour détecter les mouvements d'ensemble d'une pièce électroacoustique.

Pour conclure, signalons encore d'autres propriétés remarquables du logiciel, telles l'ouverture simultanée de plusieurs acousmographies, la possibilité d'organiser les objets graphiques en bibliothèques et de transférer les propriétés graphiques (couleur, styles) d'un ensemble à un autre, la quasi parfaite compatibilité de plateforme Macintosh et PC, et enfin les capacités d'écoutes ralenties ou filtrées du signal.

3.4. Développements en cours et futurs

L'Acousmographe est devenu le cadre idéal pour expérimenter de nouveaux algorithmes d'extraction de descripteurs, sous forme de plugins. C'est déjà le cas avec deux plugins spécialisés, l'un pour le repérage automatique des similitudes du timbre au sein d'une même pièce (stage de Sébastien Gulluni), l'autre pour la détection automatique d'événements (stage de Sébastien Roger). Nous prévoyons de nombreuses nouvelles fonctions dans les prochaines versions, comme l'import-export XML d'une acousmographie, l'exportation sous forme de film avec plus d'options, la recherche de symboles par type et par nom, et bien d'autres encore. Soulignons le fait que les nouveautés peuvent aussi venir de développeurs tiers qui créeront des plugins (à cet effet nous préparons la documentation de l'API à plugins de l'Acousmographe), ou encore d'utilisateurs chevronnés qui sauront profiter de la formidable opportunité qu'offre ce logiciel, à savoir, l'expérimentation libre de nouveaux systèmes de notations symboliques.

4. REPRÉSENTER LE SON ET LA MUSIQUE

4.1. Quelques considérations

La représentation (graphique) du sonore pose un problème inédit au musicien. D'une part, il est vain d'imaginer représenter au mieux graphiquement les caractéristiques du son. La représentation la plus exacte est tout simplement la forme d'onde ! La transcription vise à extraire des informations pertinentes du perçu, à souligner des traits importants, essentiels, ou rares du son. Du côté de l'analyse sonore, la représentation est toujours la matérialisation d'un point de vue de l'analyste [4], c'est son but, son rôle, et il est donc normal que plusieurs analyses, plusieurs représentations d'une œuvre puissent coexister. Un exemple explicite est la double représentation de l'œuvre de Bernard Parmegiani *Fg" Pcwmc" Uqqtwo* dans le cd-rom « *La Musique Electroacoustique* ». La première représentation est un synoptique extrême du mouvement, permettant une compréhension réelle et la mémorisation de la forme unique du mouvement (l'apparition décalée de trois couches sonores, leur progression en densité et intensité, puis leur disparition). La deuxième représentation détaille chaque unité sonore, classe chaque objet dans sa famille d'appartenance par la couleur, etc. Le prix de cette exhaustivité remarquable dans le détail est la mise en retrait du geste global (relativement simpliste) de l'œuvre. Les deux représentations sont donc utiles, nécessaires même, pour la compréhension de l'œuvre.

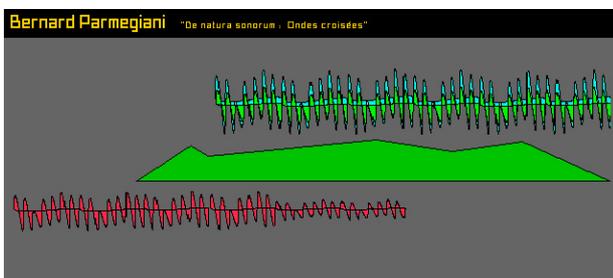


Figure 7. Bernard Parmegiani - *De Natura Sonorum* – *Ondes croisées* par Daniel Teruggi : représentation synoptique des trois voix (2000)

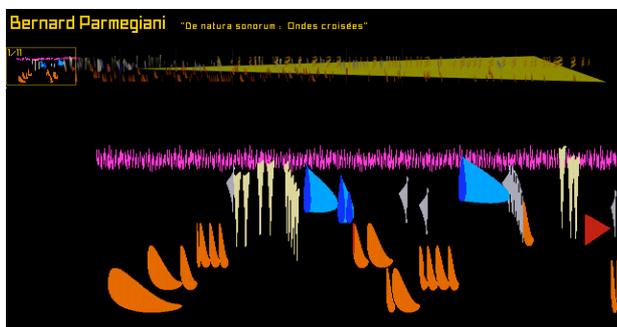


Figure 8. Bernard Parmegiani *Ondes croisées* par Pierre Couprie : représentation synoptique (partie haute) et zoom détaillé des 20 premières secondes du mouvement (2000)

4.2. Faut-il un Acousmographe ?

La problématique posée par la question de la représentation des musiques électroacoustiques et autres sons enregistrés ne saurait se résumer à la question des capacités du logiciel proposé. D'une part, la réalisation d'une acousmographie prend du temps ; ce qui est préjudiciable à la généralisation de l'usage du logiciel. L'Acousmographe a tout d'abord été conçu pour répondre à une demande relativement professionnelle, dans le but implicite de favoriser la production de représentations abouties, à montrer, donc d'une indéniable qualité graphique. Une part non négligeable de ces utilisateurs professionnels demande que soient offertes désormais de très nombreuses capacités graphiques, une sorte de synthèse des logiciels commerciaux disponibles, ce qui est hors d'atteinte d'une petite équipe de travail, et peut-être même hors de propos. Dans ces circonstances, la possibilité d'intégrer des objets réalisés à l'aide d'autres logiciels semble une bonne solution.

Un aspect notable dans la qualité imitant le trait de crayon « artisanal » et naturel est que le graphique laisse entrevoir l'aspect dessiné, donc humain, permettant probablement de souligner qu'il s'agit là de graphismes plus artistiques que codés symboliquement. Même si l'on s'attache parfois à la qualité graphique de la notation musicale de tel ou tel compositeur, il est clair pour le lecteur qu'il s'agit d'un attribut accessoire du code que l'écriture est censée produire. Ainsi donc la demande de capacités graphiques très abouties pour l'Acousmographe serait éventuellement liée à la nécessité de signaler par des indices de factures qu'il s'agit d'un dessin libre, même informatisé, et non d'un code (cf. figure 9)



Figure 9. Michel Chion *Danse de l'ombre* par Muriel Adamo : Élégance de tracés d'encres et gouaches sur papier (2005)

Or, dans de nombreuses circonstances, le besoin se fait plus souvent ressentir d'un ensemble de repérage explicitant quelques aspects communs du signal, ceci dans un emploi rapide et efficace. De premiers pas ont été faits dans la version 2 pour proposer des outils de détection automatique. Nous fondons quelques espoirs dans les travaux de recherches sur les détections automatiques de similarités, pouvant aboutir éventuellement à des propositions automatisées de découpage formel. D'autre part, une recherche sur des

modalités de représentations simplifiées nous semble nécessaire.

C'est le sens du travail entrepris en 2006 avec Noémie Springer-Ohana, proposant un panorama de

représentations simplifiées à l'extrême, celles-ci coïncidant toutes dans leur cadre de représentation, et rendant toutefois compte d'aspects importants du signal analysé [7].

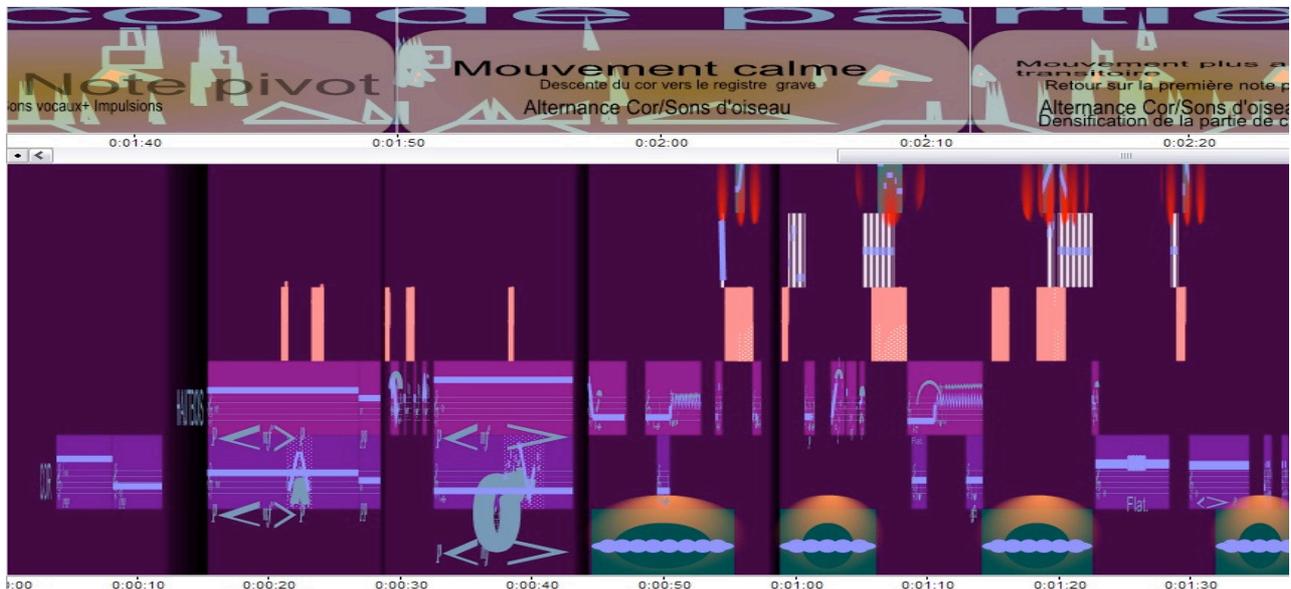


Figure 10. François Bayle *l'Oiseau moqueur*
Représentations multiples & isomorphes par Noémie Springer-Ohana (2008)

4.3. Quelques réalisations, au GRM et ailleurs

On connaît à l'étranger la célèbre partition de Rainer Wehinger établie à posteriori sur Artikulation de Ligeti, et avec l'aide du compositeur, notamment pour les aspects analytiques. Citons, en Angleterre, des représentations réalisées par Andrew Lewis ou Michael Clarke qui font également appel aux outils logiciels, ou en Norvège, les représentations symboliques de Lasse Thoresen [11], qui développe depuis plusieurs années déjà un ensemble particulièrement développé de codes graphiques symboliques, en s'appuyant, tant sur les analyses schaefferiennes que les Unités Sémantiques Temporelles (UST - MIM - Marseille) ou les propositions d'analyse fonctionnelle de Stéphane Roy.

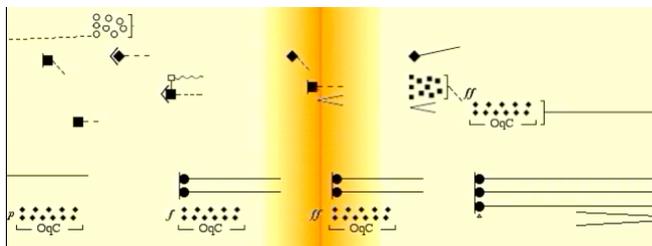


Figure 11. Ake Parmerud *Les Objets Obscurs*
Transcription codée par Lasse Thoresen (2004)

Au Groupe de Recherches Musicales, une intense activité d'exploitation des potentialités du logiciel se manifeste dans le cadre de la collection d'ouvrages monographiques « *Portraits Polychromes* » publiée sous la direction d'Évelyne Gayou [6]. Riche de maintenant

15 numéros, la collection offre l'opportunité de nouvelles analyses et transcriptions des auteurs étudiés, pour lesquelles une association remarquable – édition papier – site interactif permet d'offrir différentes modalités de transcriptions selon le support choisi.

Du côté de l'Éducation Nationale, les enseignants de musique se sont approprié le logiciel dans des usages parfois originaux, et ont notamment mis l'accent sur la simplification nécessaire des modalités de représentations du son, pour des raisons pédagogiques évidentes : un élève qui peine déjà à assimiler la notation musicale traditionnelle, n'a que faire d'une complexité de codes nouveaux et non-pérennes.

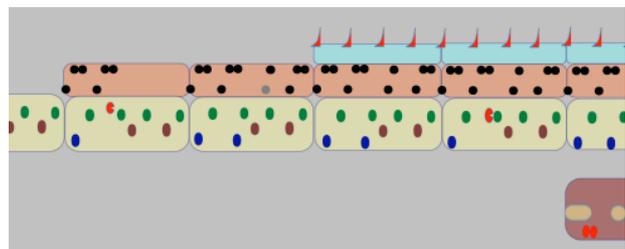


Figure 12. Herbie Hancock *Water Melon Man*
Transcription par Guillaume Caplain Saint-André pour ses élèves (2005)

À côté de ces travaux très rigoureux, on peut trouver chez certains transpositeurs issus plus particulièrement de milieux du graphisme des adaptations très singulières de la notion de transcription. Le renvoi à un code d'écriture ou à une description spectro-morphologie est délaissé pour faire place à des éléments strictement

graphiques d'évocation du cadre général de la pièce, ou faisant sans vergogne appel à l'entendu causal. La beauté et la puissance évocatrice de ces réalisations troublent le lecteur. S'agit-il vraiment de transcription, ou bien d'interprétations d'autres natures de l'entendu, artistiques, esthétisantes dans leur graphisme, scénarisées dans leur choix de point de vue de représentations ?

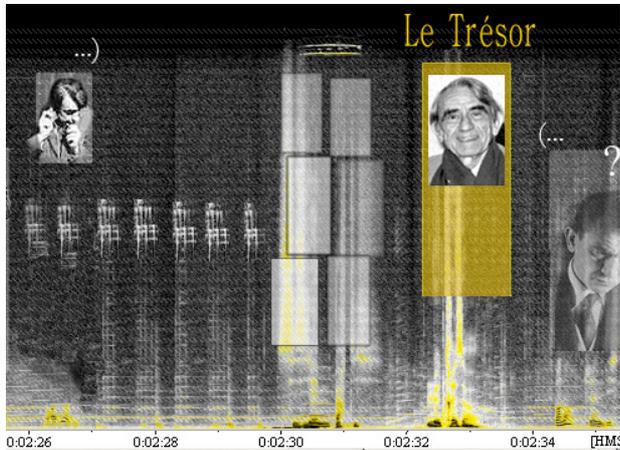


Figure 13. Michel Chion *La Tentation de St Antoine* Transcription figurative (collages) par Yoann Samson (2005)

5. CONCLUSION

L'analyse et la représentation des musiques électroacoustiques est une pratique de plus en plus répandue, la disponibilité d'outils informatiques ayant libéré le transcritteur des contraintes du relevé temporel. Toutefois, cette pratique étendue n'a pas encore permis de dégager une méthodologie acceptée par tous et pouvant servir de socle commun à toutes les modalités de représentations existantes. De grandes divergences se font jour dans les demandes d'évolutions qui nous sont soumises : tandis que certains souhaitent ardemment un enrichissement des possibilités graphiques, dans le sens d'une qualité plus manuelle, artistique, d'autres proposent une animation de la lecture, notamment l'apparition-disparition des objets pendant l'écoute; d'autres enfin espèrent la 3D et la représentation de l'espace, mais selon quel usage ? Maintenant que les outils existent, il importe que la pratique permette de dégager les enjeux de la transcription.

L'Acousmographe a atteint désormais une maturité enviable, tant au niveau de son code informatique, que dans le nombre important de transcriptions produites avec son emploi. Il est souhaitable maintenant d'ouvrir ce code aux autres développeurs, et rendre transparent le format de représentation, afin que d'autres logiciels et d'autres utilisateurs puissent s'en emparer et en généraliser l'usage. À ce prix, les utilisateurs actuels et potentiels pourront se sentir rassurés quant à l'avenir et la pérennité de leurs travaux de transcription.

6. RÉFÉRENCES

- [1] Bayle, F. « *Musique Acousmatique, propositions... positions* » pp. 224 – planches XXV à XXXII INA-Buchet/Chastel. Paris. ISBN 2-7020-1584, 1993.
- [2] Besson, D., Koechlin, O., Bonfils, M. « *Les Musicographies* » CD-rom produit par l'Ina-Grm – non édité. 1995.
- [3] Delalande, F. « *Pratiques et objectifs des transcriptions des musiques électroacoustiques* », pp. 130-153 in Rémy Campos, Nicolas Donin, *L'analyse musicale, une pratique et son histoire*. Droz, Genève, 2009.
- [4] Delalande, F., Besson, D. « *Problèmes théoriques et pratiques de la transcription des musiques électroacoustiques* » in Actes du second congrès européen d'analyse musicale, Université de Trente, Italie, 1991 2-1.
- [5] Gayou, E. « *Le GRM Groupe de Recherches Musicales, Cinquante ans d'histoire* ». pp. 182-183 Fayard, Paris, 2007.
- [6] Geslin Y., Lefevre A. « *Sound and musical representation: the Acousmographe software* ». International Computer Music Conference, Miami, USA, 2004.
- [7] Geslin, Y., Spinger-Ohana, N. « *Abstraction et symbolisation dans la représentation des musiques électroacoustiques : un exemple de transcription à niveaux multiples de l'Oiseau moqueur réalisé à l'aide de l'Acousmographe* », in Electronic Music Studies - EMS08 proceedings, Paris, 2008-2010. <http://www.ems-network.org/ems08/paper.html>
- [8] Groupe de Recherches Musicales. « *Les Musiques Électroacoustiques* » CD-Rom interactif, dirigé par François Delalande © Ina 2001.
- [9] Groupe de Recherches Musicales. « *Portraits Polychromes* », collection monographique et site web (2001-2010), sous la direction d'Évelyne Gayou, réalisation Dominique Saint-Martin. © Ina <http://www.ina-grm.com/portraits-polychromes.html>
- [10] Koechlin, O. et Vinet, H. « *The Acousmographe, a Macintosh software for the graphical representation of sounds*'i in Proceedings of the International Computer Music Conference, pp. 586–588. Montreal: Faculty of Music, McGill University 1991.
- [11] Thoresen, L. « *Spectromorphologic Analysis of Sound Objects. An adaption of Pierre Schaeffer's Typomorphology*'i. Inédit, présenté au GRM, Paris Juin 2002 ; également à la Library of the Norwegian Academy of Music, Oslo, Norvège.

UTILISATIONS AVANCÉES DU LOGICIEL IANALYSE POUR L'ANALYSE MUSICALE

Pierre Couprie

MINT-IUFM de l'université de Paris-Sorbonne Paris 4

pierre.couprie@paris.iufm.fr

RÉSUMÉ

iAnalyse est un logiciel d'aide à l'analyse musicale permettant de synchroniser une partition numérisée sur un fichier audio ou vidéo et d'annoter cette partition à l'aide de graphiques ou de textes. L'article se concentre sur les outils d'analyse qui complètent ce système avec la création de graphiques tels que la structure en vue synoptique, le diagramme formel, la visualisation des variations de tempo dans l'interprétation ou les associations de cellules musicales sous la forme de cartes. En conclusion, l'auteur termine par une présentation des fonctions qui seront ajoutées dans la prochaine version du logiciel.

1. INTRODUCTION

Cet article fait suite à deux autres articles présentés lors des JIM 2008 [2] et lors d'EMS 2008 [3]. Ces deux présentations consistaient en une introduction au logiciel iAnalyse¹ et une exploration de ses possibilités dans le cadre de l'analyse de la musique électroacoustique. Le logiciel a beaucoup évolué depuis la version de 2008 et commence à être très utilisé par des chercheurs ou des enseignants², il m'a donc semblé important de compléter ces premières publications en présentant ses nouvelles fonctions : les outils d'analyse. Ces outils sont un ensemble d'utilitaires permettant d'aider l'utilisateur dans son travail d'analyse. Après une courte introduction exposant l'historique du logiciel et son fonctionnement, je décrirai en détail comment utiliser 4 de ces outils d'analyse.

2. IANALYSE : UNE COURTE INTRODUCTION

2.1. Présentation

iAnalyse est un logiciel d'aide à l'analyse musicale, il permet de synchroniser une partition musicale numérisée avec un fichier audio ou vidéo. Il est ensuite possible d'ajouter sur la partition des annotations graphiques, elles-mêmes synchronisées sur le déroulement temporel du média, afin d'analyser une œuvre ou de créer un guide d'écoute.

2.2. Historique des versions

Le développement de iAnalyse a commencé en 2006, date à laquelle j'ai appris que l'IRCAM développait un logiciel pour l'Education Nationale qui permettrait d'annoter les partitions³. Ce logiciel ne devait fonctionner que sur la plateforme Windows, j'ai donc décidé de développer un équivalent sur Macintosh. C'est ainsi que la première version de iAnalyse est apparue, le logiciel se nommait alors EDiMu⁴ et était développé dans l'environnement Realbasic. En 2007 est sortie la première version nommée iAnalyse, le logiciel était alors relativement simple : un lecteur audio et vidéo, une synchronisation des pages d'une partition numérisée sur le déroulement temporel du média, un dessin d'une tête de lecture (nommée curseur) sur la partition permettant de suivre le déroulement de la musique et un ensemble d'annotations graphiques à ajouter sur la partition. La version 2, disponible en 2008, contenait les prémices de ce qui deviendrait les outils d'analyse. Poussant l'expérience et l'environnement Realbasic dans ses limites, j'ai décidé de reprendre le développement dans l'environnement Xcode en Objective-c, ce qui m'a permis d'améliorer la réactivité du logiciel et d'augmenter les fonctions disponibles. Cette nouvelle version, iAnalyse 3, est disponible depuis début 2009 en deux versions : une version simple gratuite et une version contenant de nombreuses fonctions supplémentaires, dont les outils d'analyse présentés dans cet article, au prix de 15€.

2.3. Le fonctionnement du logiciel

Comme le montre la figure 1, la structure de iAnalyse est très simple. Il fonctionne comme un logiciel de diaporama avancé. L'utilisateur commence par importer un fichier audio ou vidéo, toutes les fonctions de lecture sont alors disponibles (lecture, sélection, boucle, variation du tempo sans modifier la hauteur, etc.). Il va ensuite importer une partition numérisée (images ou

1 Cet article est basé sur la version 3 du logiciel disponible à partir de mars 2009. Il est disponible sur le site : <http://ianalyse.pierrecouprie.fr>.

2 Il est possible de voir plusieurs exemples d'analyses ou de guides d'écoute réalisés avec iAnalyse sur la chaîne Youtube du logiciel : <http://www.youtube.com/user/pierrecouprie#g/c/90708293B6AACA01>

3 Ce logiciel devait se nommer MusiqueLab 2 Annotations [5]. L'IRCAM semble ne jamais avoir finalisé le projet, il n'est donc pas disponible.

4 Edition de Diaporamas Musicaux.

fichier PDF) et synchroniser cette partition sur le média en positionnant chaque page en fond d'une diapositive⁵. Il peut alors créer un curseur (une tête de lecture) qui se déplacera sur la partition au rythme de la musique. Il peut aussi ajouter sur cette partition un ensemble d'annotations graphiques ou textuelles permettant de souligner certains aspects de la musique (harmonie, thème, structure, indications analytiques, etc.). Enfin, un ensemble d'outils d'analyse sont disponibles afin de visualiser le média (sonagramme), créer des graphiques à partir du curseur (variation de tempo) ou des annotations (structure, diagramme formel, carte) ou encore assister l'utilisateur dans son travail analytique (navigateur d'annotations, visualisation des échelles et des modes, calcul des valeurs de *set-theory*).

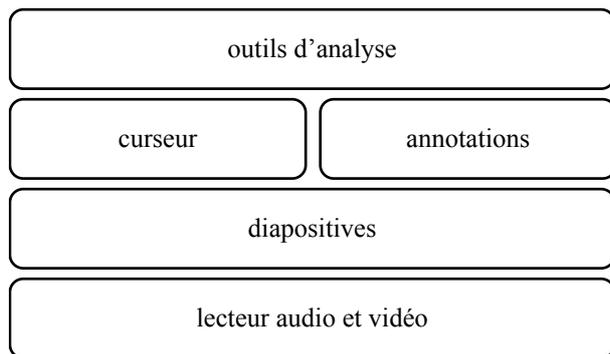


Figure 1. La structure du logiciel iAnalyse.

L'utilisateur peut ensuite présenter son travail à l'aide du logiciel, en mode normal ou en mode plein écran, ou l'exporter sous la forme d'images ou d'un film Quicktime.

3. LES OUTILS D'ANALYSE

3.1. Présentation

Les outils d'analyse du logiciel constituent un ensemble de 10 utilitaires dont l'objectif est d'aider le chercheur dans son travail d'analyse ou lors de ses présentations. Parmi ces différents outils, 4 se distinguent particulièrement par leur intégration au logiciel et/ou l'aboutissement de leur développement.

3.2. L'analyse des variations de tempo

Un premier outil permet de visualiser assez facilement les variations de tempo. Il est ainsi possible de comparer plusieurs interprétations d'une même œuvre.

Le principe de fonctionnement est le suivant. Après avoir créé un curseur⁶ positionné sur chacun des temps de la partition, le logiciel va calculer les intervalles temporels entre chacune des images-clés du curseur (chacun des temps) et les afficher sous la forme d'un graphique (figure 2).

L'ensemble de l'interface est relativement facile à utiliser, mais présente plusieurs réglages complexes.



Figure 2. L'interface permettant de visualiser les variations de tempo par mesure.

⁵ Cette étape peut être omise si l'utilisateur travaille sur des musiques de tradition orale, les annotations seront alors de simples graphiques sans lien avec une partition.

⁶ La fonction curseur créée une tête de lecture se déplaçant sur la partition au rythme de la musique. Pour permettre l'interpolation de position de cette tête de lecture, l'utilisateur doit marquer un certain nombre d'images-clés (les positions de la tête) sur, par exemple, chacun des temps.

Comme le montre la figure 3 — extraite de la figure 2 — il est ensuite possible de choisir comment seront regroupés les intervalles temporels entre chaque curseur afin de calculer les valeurs de tempo minimum, moyen et maximum. Trois types de regroupements sont proposés :

- par curseurs : c'est le moyen le plus simple de regrouper les temps et d'obtenir une valeur par mesure. Toutefois, ce choix ne fonctionne que si les mesures sont régulières ;
- par marqueurs⁷ : cette fonction permet de faire des calculs sur des intervalles temporels irréguliers ;
- par annotations : en cas d'utilisation des annotations pour marquer la structure, il est, par exemple, possible de réaliser le calcul en fonction des cellules thématiques.



Figure 3. Les différents choix pour regrouper les temps lors du calcul des variations de tempo.

Ces différents réglages permettent d'obtenir tous les types de configurations possibles et ainsi d'adapter les calculs à n'importe quel type d'enregistrement.

Dates	Tempo moyen	Tempo minimum	Tempo maximum
00:07:44	46,01141	39,1097	53,96319
00:07:55	43,03938	37,54831	49,42434
00:13:23	39,71404	34,86522	48,36148
00:19:27	39,02486	31,75357	44,53082
00:25:42	42,03379	36,43934	47,55742
00:31:13	41,74122	35,98095	47,34934
00:36:58	38,8563	28,64722	50,69536
00:43:06	44,42086	40,21364	49,75088
00:48:46	42,67185	34,86529	57,70729
00:54:09	40,39173	34,83741	46,1605
01:00:03	37,32462	25,99775	51,55376
01:06:46	38,03484	28,64732	49,44589
01:12:77	37,25452	31,02981	49,70905
01:19:09	38,35921	35,27103	42,08625
01:25:35	48,79004	38,4411	55,27821
01:30:27	55,80867	48,18174	67,77892
01:34:57	51,49538	41,47106	62,27339
01:39:23	35,96088	31,08918	38,26194
01:45:90	37,08855	30,45648	51,47548
01:52:17	39,28453	25,74751	55,28803
01:58:48	40,06103	28,67605	51,49368

Figure 4. Les différentes valeurs calculées à partir du tempo.

Etant donné que le graphique n'intègre pas l'affichage des valeurs exactes, il est aussi possible d'afficher ces valeurs sous la forme d'un tableau (figure 4).

3.3. La structure musicale

Le deuxième outil présenté dans cet article permet de visualiser la structure d'une œuvre de manière synoptique (figure 5).

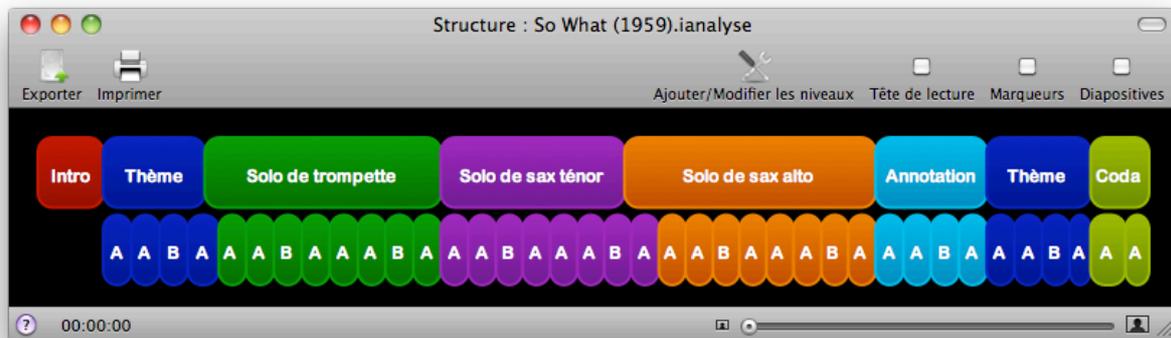


Figure 5. L'interface permettant de visualiser la structure musicale.

L'utilisateur doit d'abord marquer la structure sur la partition à l'aide d'annotations. La figure 7 montre une annotation indiquant l'introduction de *So What*⁸. L'utilisateur a ici dessiné un cadre entourant les deux premiers systèmes de la partition et indiquant le début de l'introduction musicale. Le logiciel va ensuite récupérer la date temporelle du début de cette annotation pour construire le graphique. Comme le montre la figure 5, il est possible de visualiser plusieurs niveaux de structure. Pour ce faire, l'analyste doit inscrire les

annotations par niveau dans des calques⁹ différents. Il va ensuite sélectionner les différents calques qui entreront dans le dessin de la structure (figure 6) ainsi que la manière dont ils apparaîtront (la couleur et le texte qui seront affichés).

7 Les marqueurs sont des traits verticaux apparaissant dans le scénario et permettant de repérer des positions temporelles précises.

8 Cette figure reproduit une analyse réalisée par Laurent Cugny. L'original a été réalisé pour l'exposition Miles Davis à la Cité de la Musique de Paris sur le logiciel metaScore.

9 Le fonctionnement des calques dans iAnalyse est sensiblement identique à l'utilisation des calques dans les logiciels d'image comme Photoshop : chaque annotation graphique peut être associée à un calque qui sera affiché ou masqué en fonction des besoins.

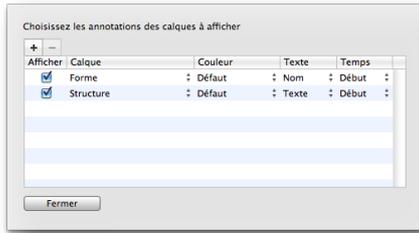


Figure 6. Les paramètres servant à l’affichage de la structure.

Ce graphique figurant la structure musicale est présenté dans une nouvelle fenêtre, mais il peut aussi être intégré au plan graphique de la fenêtre principale (figure 7). Dans ce cas, lors d’une exportation vidéo, la structure sera aussi visible. Cette fonction d’incrustation est entièrement paramétrable (position, largeur, taille des marges, zoom, couleur d’arrière-plan, présence ou non d’une mini-tête de lecture) et est aussi interactive puisque l’utilisateur peut cliquer directement dessus pour repositionner la tête de lecture générale.

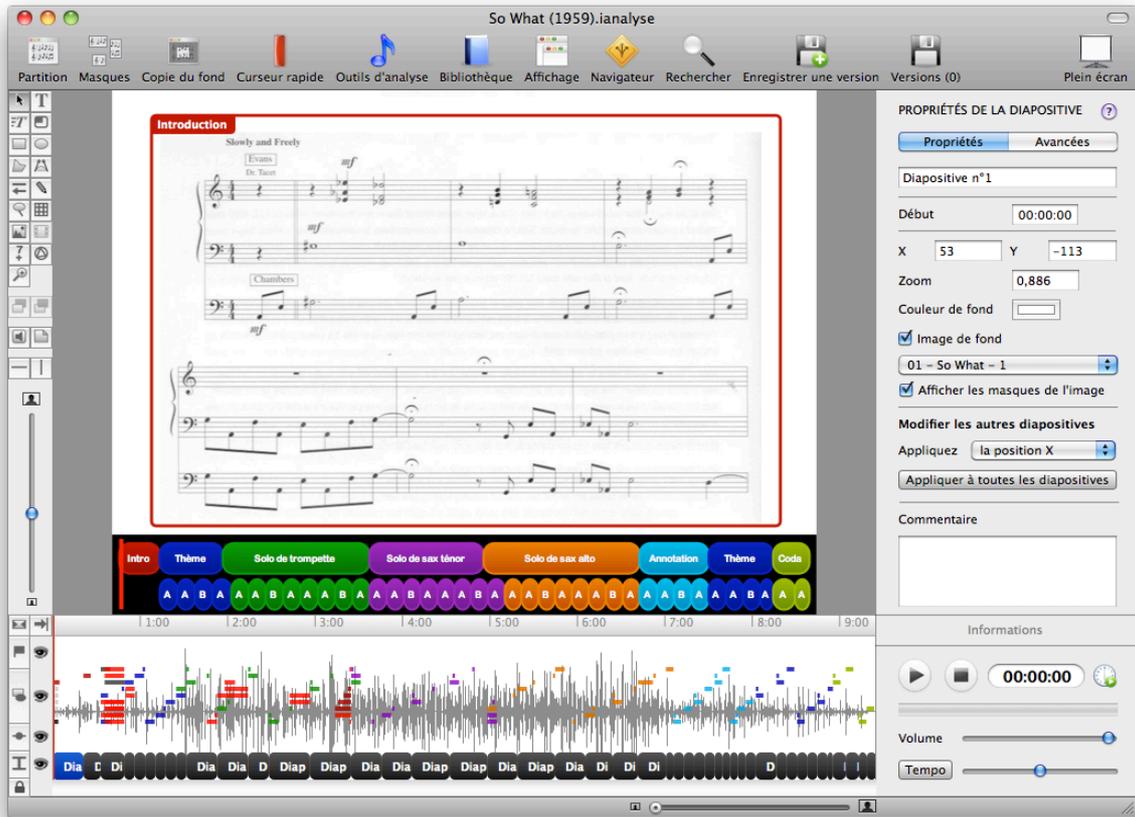


Figure 7. L’incrustation de la structure dans le plan graphique.

3.4. Le diagramme formel

Le diagramme formel [1] est complémentaire de l’outil de visualisation de la structure. En effet, il fonctionne sensiblement de la même manière et permet d’obtenir un graphique dans lequel les cellules musicales sont représentées par ordre d’apparition (axe horizontal) et segmentées selon la méthode d’analyse paradigmatisée (axe vertical). Le graphique ainsi obtenu (figure 8) permet de visualiser facilement la ou les structures sous-jacentes aux cellules segmentées.

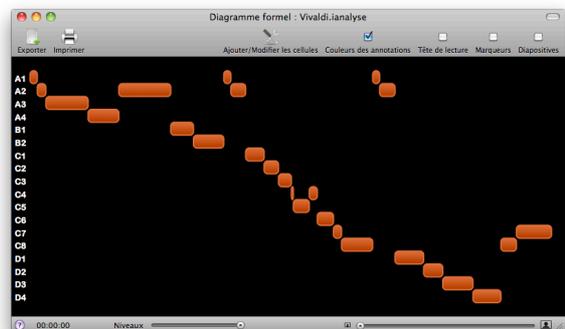


Figure 8. La visualisation du diagramme formel. La réalisation d’un tel graphique nécessite, comme pour celui de la structure, que l’analyste ait auparavant

marqué les différentes cellules à l'aide d'annotations (figure 9).



Figure 9. Le marquage des cellules afin de réaliser un diagramme formel.

Ces annotations peuvent être dessinées sur le même calque ou sur des calques différents. Le logiciel va alors calculer le diagramme en fonction des noms donnés aux annotations (cellule A, cellule B, etc.) ou du texte qu'elles contiennent en fonction des indications de l'utilisateur (figure 10). N'importe quel nom peut convenir puisque le logiciel détectera simplement les répétitions. Toutefois, une fonction permet de réduire le niveau d'analyse (figure 11). Dans ce cas, il est préférable d'utiliser un codage simple alternant lettres et nombres tel que : A1, A2, B1, B2, etc.

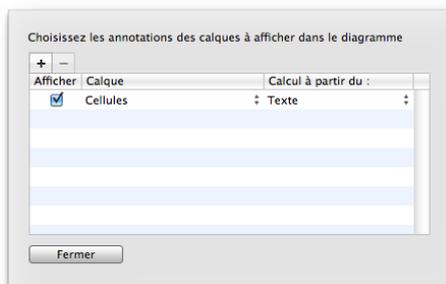


Figure 10. Les paramètres servant au calcul du diagramme formel.

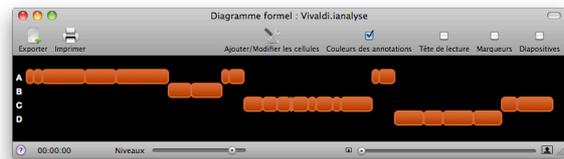


Figure 11. Le diagramme formel de la figure 8 avec une réduction des niveaux.

3.5. Les cartes

Le quatrième et dernier outil présenté dans cet article permet d'afficher les fragments de partitions présents sous les annotations indépendamment du reste de la partition.

Après avoir créé des annotations recouvrant des fragments de la partition, par exemple la cellule 1 de *NoaNoa* de Kaija Saariaho dans la figure 12, l'utilisateur va retrouver l'ensemble des fragments recouverts par les annotations dans l'outil carte (colonne de droite de la figure 13). Il ne restera plus qu'à utiliser ces fragments de partitions en les glissant sur l'espace blanc de la carte. Leur positionnement est totalement libre et la surface de cette carte n'est pas limitée, elle est fixée par défaut à 2048x2048 pixels, mais elle peut être agrandie à tout moment. En outre, l'extrait audio

correspondant à chacun des fragments de la partition ajoutés sur la carte (positions et durées des annotations) peut être lu. Il est ainsi possible de faire des écoutes comparatives de ces différents fragments.



Figure 12. La cellule 1 dans la première page de *NoaNoa* de Kaija Saariaho

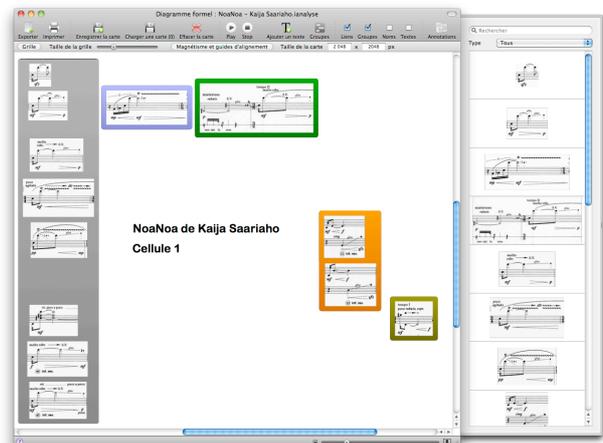


Figure 13. Un exemple de carte représentant une analyse paradigmatique

Les figures 13 et 14 montrent aussi qu'il existe un certain nombre de fonctions graphiques intéressantes pour l'utilisateur. Il peut, par exemple, regrouper différentes cellules, créant ainsi des zones colorées, ajouter du texte (figure 13) ou relier les cellules à l'aide de traits ou flèches avec des styles différents (figure 14).

Une fois la carte réalisée, l'utilisateur peut l'enregistrer dans le projet afin d'en créer d'autres. Il est aussi possible d'exporter les cartes sous la forme d'images.

Les chercheurs commencent à s'approprier cet outil dans des usages très divers allant de la création de tableaux paradigmatiques à la comparaison de cellules musicales dans des musiques de tradition orale, en passant par des graphiques heuristiques ou des cartes résumant le parcours musical d'une œuvre.

Notons que ce type d'outils est aussi en cours d'expérimentation à l'IRCAM avec le logiciel Score Charter [4].

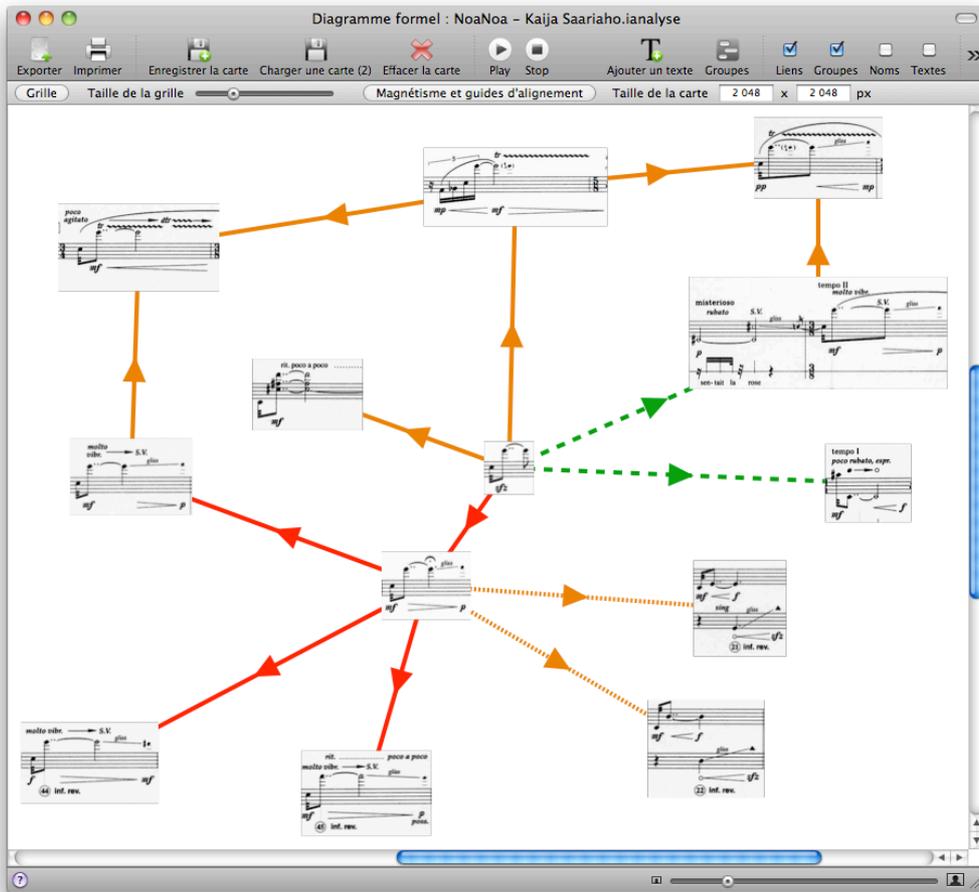


Figure 14. Un exemple de carte représentant les relations entre des cellules thématiques.

4. CONCLUSION

Ce court article m'a permis de montrer quelques-unes des dernières évolutions de mon logiciel iAnalyze. Le développement de la version actuelle (version 3) est terminé, les modifications futures ne concerneront que quelques résolutions de bogues. La version 4 est en cours d'élaboration et sera disponible au quatrième trimestre 2010. Elle reprendra bien évidemment les caractéristiques de la version 3 ainsi l'interface ne sera probablement que peu modifiée. Par contre, de nombreuses nouvelles fonctions verront leur apparition comme, par exemple, l'importation et l'exportation vers d'autres formats (MIDI, SDIF, CSV, etc.), de nouveaux outils permettant d'analyser et d'annoter les films ou la possibilité pour les développeurs de créer des plug-ins.

5. RÉFÉRENCES

[1] Chouvel, JM. *Analyse musicale. Sémiologie et cognition des formes temporelles*, L'harmattan, Paris, 2006.
 [2] Couprie, P. "iAnalyze : un logiciel d'aide à l'analyse musicale", Journées d'Informatique Musicale - JIM08, Albi, 2008, article en ligne sur : <http://www.gmea.net/jim08/index.php/Articles>.

[3] Couprie, P. "Analyser la musique électroacoustique avec le logiciel iAnalyze", Electronic Music Studies Network, Paris, 2008.
 [4] Goldman, J. "Un outils de mise en tableau au service de l'analyse paradigmatique, et quelques divergences interprétatives", L'analyse musicale, une pratique et son histoire, ed. Nicolas Donin et Rémy Campos, Droz/HEM, Genève, 2009.
 [5] Guédy, F. Bresson, J. Assayag, G. "Musique Lab 2 un environnement d'aide à la pédagogie musicale", JIM07, Lyon, 2007, article en ligne sur : <http://www.grame.fr/Recherche/JIM/jim07/download/02-Guedy.pdf>.

Session 5

Image, geste, mouvement

VIDEO MUSIQUE : ILLUSIONS PERCEPTIVES ET DISCOURS ESTHETIQUE

Elsa Justel
Fondation Destellos
justelsa2@gmail.com

RÉSUMÉ

La problématique de la mixité a été un des sujets de préoccupation des compositeurs de musique électroacoustique (ME), depuis les dernières décennies. Dans cet article, nous allons décrire certains des résultats de notre recherche par rapport aux possibles interrelations entre la musique et l'image vidéo. L'analyse du phénomène de la mixité nécessite une observation à double sens : les aspects perceptifs et paramétriques des deux langages. De la bonne correspondance entre ces éléments dépendra le résultat esthétique.

1. INTRODUCTION

La plupart des signaux envoyés au cerveau par l'œil et l'oreille ne sont que des illusions d'une réalité physique. En effet, certains phénomènes comme le son musical et l'image en mouvement sont dotés d'une structure interne très complexe, dont les paramètres sont mesurables dans le laboratoire. Cependant, la perception de ces effets est parfois contradictoire par rapport au phénomène physique qui les produit. Les sens agiront comme des filtres dont la fonction sera de resynthétiser ces paramètres en nous permettant de capter globalement et rapidement le phénomène audiovisuel.

Notre recherche dans le champ audiovisuel [1] tend à découvrir, analyser et comparer certains de ces paradoxes et illusions de la perception. En observant les travaux graphiques de l'architecte et graveur néerlandais Maurits Cornelius Escher [2], nous pouvons affirmer qu'il est possible d'établir des parallèles et des comparaisons entre les effets de l'illusion visuelle et ceux de la perception auditive.

Nous avons pris comme point de départ certaines des prémisses énoncées par l'artiste mentionné et nous avons développé un plan orienté vers la recherche des possibles éléments en commun entre les deux phénomènes perceptifs. Bien que les deux modes d'expression s'articulent sur des paramètres très différents, nous croyons qu'il est possible d'établir des relations pour aboutir à une bonne correspondance du discours audiovisuel. Il est de la fonction de l'artiste de mettre en évidence ces liens subtils entre les deux langages pour parvenir à une structure audiovisuelle organique.

2. FORMES DE VIDEO

Nous devons signaler que le terme vidéo est utilisé d'une façon indistincte pour des expressions visuelles qui ont évolué historiquement dans des directions esthétiques et conceptuelles diverses. Depuis le traditionnel dessin animé, réalisé à la main cadre par cadre, jusqu'aux plus complexes formes d'image de synthèse, la vidéo a été utilisée pour des fins artistiques, publicitaires, scientifiques. Deux des techniques les plus utilisées actuellement sont l'animation 3D¹ et la forme représentative qui utilise un caméscope et décrit une situation ou raconte une histoire avec ou sans des personnages. Ce dernier genre, qui a ses origines dans l'apparition du cinéma expérimental, est très répandu.

L'avènement de l'ordinateur et la prolifération des logiciels de montage et de traitement du son et de l'image, de plus en plus accessibles, ont donné lieu à une vaste production d'objets audiovisuels. Ces audiovisuels constituent un pas intermédiaire entre l'expression abstraite et le cinéma.

Moins répandue est la Forme Abstraite, qui apparaît à la fin des années 40. Entre les représentants les plus intéressants : le Canadien Norman MacLaren, qui dessinait les images sur la pellicule ; l'Américain James Whitney, qui travaillait des figures géométriques élaborées sur le modèle de la musique sérielle; le Français A. Raymond qui photographiait ses propres œuvres plastiques et plus tard, dans les années 70, quelques membres du groupe GRM de Paris, entre autres, Bernard Parmeggiani, qui introduisent la musique électroacoustique dans le genre. Aujourd'hui, ce type de vidéo, réalisé avec la technologie numérique, offre des plus larges possibilités aussi bien techniques qu'expressives. C'est dans cette dernière forme que nous trouvons les exemples d'interrelation entre l'image et le son, les plus intéressants.

« (...) de réelles possibilités créatrices émergent quand la structure de l'image dicte ou "inspire" la structure sonore et vice-versa, ou quand leur conception est simultanée. Ceci évidemment, est mieux ré-

¹ 3D, nomenclature utilisée pour se référer aux procédés numériques de simulation d'image en trois dimensions.

alisé quand les deux parties ont des origines créatrices communes ». [3]

Nous allons employer le terme vidéo musique pour définir notre proposition, puisqu'il nous paraît pertinent par rapport à l'étude du problème de correspondance entre les deux langages. D'ailleurs, et au risque de paraître trop partielle, cette forme est l'option que nous avons choisie pour nos propres réalisations audiovisuelles.

Notre recherche est orientée vers les possibilités expressives de l'image en mouvement et ses affinités avec le discours musical, à travers l'étude de deux aspects principaux :

- **L'aspect perceptif** : qui tend à établir les relations entre les phénomènes physiques et les effets perceptifs qu'ils provoquent. Cette étude commence par l'exploration des matériaux sonores et visuels et la recherche des corrélations possibles.

- **L'aspect formel** : qui cherche à découvrir les paramètres porteurs de forme des deux langages. Toute forme d'expression qui implique du temps et de la dynamique, possède une certaine narrativité. Une structure audiovisuelle équilibrée, même dans sa forme la plus abstraite, se développe à travers un discours articulé par des éléments d'organisation divers.

Pour élaborer la base théorique de notre projet, nous avons commencé par l'analyse d'une série d'éléments perceptifs, qui nous conduisent à l'organisation et articulation du discours audiovisuel. Même si les deux langages fonctionnent dans des champs paramétriques différents, nous pouvons constater qu'il existe des particularités en commun. Par exemple, l'effet tridimensionnel, la réverbération et le phénomène cinétique, sont quelques-uns des aspects intéressants.

Dans la plupart des essais pour établir des parallèles entre musique et image, nous sommes fréquemment confrontés avec des conflits de terminologie. Il est facile de supposer que l'analogie entre certains mots implique une similitude de concepts. Nous trouvons les mêmes mots pour définir différents aspects des arts visuels et de la musique, qui résultent néanmoins, très éloignés à niveau de la perception. Cependant, après avoir expérimenté avec les matériaux, nous avons réussi à établir des associations significatives.

3. LA NOTION D'ESPACE

Les nouvelles technologies de traitement du son et de l'image nous permettent d'élaborer une vaste gamme d'effets spatiaux. Des procédés comme la réverbération, l'équalisation, le filtrage et le délai, pour ne mentionner que les plus conventionnels, sont actuellement d'une extraordinaire ductilité pour la manipulation du son et la création des espaces virtuels. En ce qui concerne l'image, mises à part les techniques en 3D, il existe aussi des traitements numériques d'une grande précision pour l'application à des images en

deux dimensions. Moyennant ces traitements, nous pouvons par exemple faire devenir translucide une figure, un fond, une surface, pour produire des reliefs, ou bien effectuer des procédés de rotation, des étirements, des plissements, et des mouvements multidirectionnels. Ces différentes techniques nous offrent la possibilité de créer des sensations diverses, qui se manifestent souvent par des oppositions et des conflits entre deux paramètres.

Ex. 1 (animation) Déformation d'une figure et déplacement spatial.

3.1. Conflit entre plan et espace

La conception d'Escher, face au phénomène visuel est basée sur la notion d'illusion perceptuelle. Une de ses préoccupations fondamentales était le conflit entre le plan et l'espace. À travers ses travaux d'apparence emmêlée, l'artiste nous surprend par sa capacité pour nous promener depuis le monde plan de la gravure, à la perception d'objets quasi-sculpturaux, tridimensionnels.

« L'espace tridimensionnel est la seule réalité que nous connaissons. Celui de deux dimensions est tant fictif comme celui de quatre, parce que rien n'est plan, même pas le verre le plus finement poli. Nous admettons la convention qu'un mur ou une feuille de papier sont plans, cependant, nous continuons à créer l'illusion d'espace sur des telles surfaces. N'est pas absurde que, après avoir tracé quelques lignes sur un papier, nous disions qu'est une maison? ». [4]

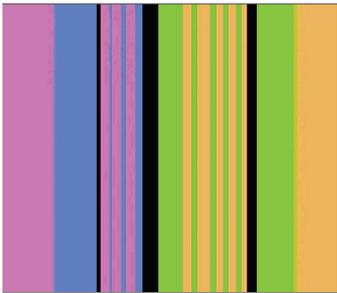
Dans la nature donc, l'espace est toujours tridimensionnel. Cependant dans la représentation graphique des objets, ainsi que dans la vidéo et le cinéma, nous sommes en présence d'un espace fictif, bidimensionnel (même dans les images réalisées par des logiciels appelés en 3D). Pourtant il est possible de créer l'illusion de tridimensionnalité à travers une figure projetée sur l'écran.

Ex.2 (animation) Plissement d'une figure (effet tridimensionnel)

La notion d'espace dans la musique, en revanche, est réelle parce que le son se propage à travers l'air dans un espace physique. Pourtant, la ME nous permet de créer des espaces virtuels. Par exemple, moyennant certaines techniques nous pouvons créer une sensation d'éloignement ou d'approximation du son, tels les procédés de réverbération ou de filtrage spectral. La réverbération sonore est résultat d'une répétition consécutive et très rapide d'un son. L'oreille est incapable de percevoir des répétitions à moins de 20 millisecondes, par conséquent l'effet de réverbération est perçu comme un « étirement » du son, ce qui donne en même temps une sensation d'éloignement dans l'espace. C'est-à-dire : une sensation de profondeur.

Nous trouvons ici, un exemple d'analogie entre la dialectique des arts visuels et celle de la musique, étant donné que le terme réverbération s'applique aussi aux arts plastiques pour exprimer la sensation de

diffusion chromatique ou d'une sorte de vibration entre deux couleurs proches. Par exemple, si nous approchons deux couleurs complémentaires comme le vert et l'orange, ou le fuchsia et le bleu (tous les deux à un même niveau de saturation), nous pouvons observer que la ligne d'union entre les deux couleurs, apparaît fusionné comme un dégradé, entre les deux tonalités. (Ex.4)



Ex.4 Effets de réverbération par proximité de couleurs complémentaires

Autre similitude entre la perception spatiale du son et de la couleur est celle qui se produit par approximation d'éléments contrastants. La perception de deux sons ayant la même intensité, peut varier selon son registre ce qui donne une sensation de proximité, de présence, aux sons aigus, et de profondeur aux sons graves. De la même façon, la profondeur spatiale d'une figure sera différente selon sa couleur (plus ou moins saturée). Existence certaines couleurs qui, combinées entre elles, provoquent un phénomène de contraste spatial. En général les couleurs chaudes (rouge, orange, jaune), tendent à s'approcher et les froids (bleu, vert) à s'éloigner.

Dans la ME nous parlons souvent d'espace réel et espace virtuel et cette dualité de concept convient parfaitement à la notion d'espace visuel. Nous savons que dans l'espace réel les conditions de propagation de la vibration sonore dépendent de la forme et la largeur de la salle et que les vibrations se produisent en dispersion sphérique. Cependant les traitements effectués sur les paramètres physiques du son et imprimés sur le support sonore, nous permettent de modifier cette réalité pour créer des sensations d'espace plus larges, dévier des trajectoires ou fixer des points dans un espace imaginaire.

Cette notion d'espace virtuel dans la ME a cependant ses inconvénients, puisque l'illusion spatiale est une construction mentale qui « peut être discutable par les indices contradictoires provenant de l'espace réel, en particulier ceux qui proviennent d'autres modalités sensorielles ». [5]

3.2. Profondeur et Relief

Ces deux termes qui paraissent se définir par opposition sont, pourtant, complémentaires dans la perception. Si nous essayons d'imiter la réalité tridimensionnelle, nous ne pouvons pas percevoir un relief dans un champ plan. En utilisant une terminologie photogra-

phique, nous dirions que : la profondeur de champs déterminera l'amplitude des reliefs.

Les célèbres gravures Noeud de Moebius ou le Cube des bandes magiques [6] de M.Escher sont des exemples classiques de ce type d'illusion spatiale. (Fig. 1) Une des stratégies pour simuler l'effet de volume est le procédé d'inversion. Le résultat perceptif est l'équivoque entre les notions d'intérieur et extérieur, de creux ou de protubérance, de concavité ou de convexité. Ce jeu d'inversions est une simple transposition intellectuelle.



Figure 1. Noeud de Moebius

En ce qui concerne le son, en plus de la relation entre amplitude et fréquence (que nous avons mentionnée en référence à la proximité d'éléments contrastants), les différences spectrales et de complexité des masses sonores, nous offre un effet similaire qui peut provoquer des sensations de relief plus ou moins profond.

Il est aussi possible d'établir un parallèle entre les notions de grandeur (de l'image) et d'intensité (du son). Un procédé classique pour donner l'illusion d'espace virtuel est la reproduction à l'infini. Si nous voulons suggérer une figure qui s'éloigne, nous devons procéder à la réduction graduelle de sa grandeur jusqu'à arriver à la limite infinitésimale. De cette façon, nous provoquons une illusion de profondeur. Quant au son, puisqu'il se propage dans l'espace et que la perception auditive est déterminée par les limites physiologiques de l'oreille, ce procédé aura comme équivalent la diminution graduelle de l'intensité, ou amplitude. De cette manière, nous aurons la sensation d'éloignement du son dans un espace virtuellement infini.

Autres effets d'association perceptive entre fréquence et couleur peuvent s'obtenir moyennant des traitements d'accumulation et de contraste. Il est très commun d'associer les sons aigus aux tonalités claires et vice-versa. Cependant l'élément psychologique sous-jacent peut provoquer des divergences d'appréciation qui nous éloignent du champ perceptif.

D'ailleurs, dans les structures complexes ce type de correspondance devient plus difficile à établir. Dans certains cas, il est possible d'obtenir de bons effets en associant des sons complexes aux tonalités empâtées, mais, dans ce cas, il sera plus approprié d'établir l'association avec des textures et d'autres qualités du

matériau visuel. Même si l'infographie limite considérablement la qualité de texture des objets représentés, il est possible de réussir des effets intéressants en expérimentant avec les différentes formes de traitement.

Les techniques d'harmonisation du son sont très variées et elles ont été employées depuis le début de la ME à l'époque des moyens analogiques. Ces techniques offrent la possibilité de « colorer » les objets sonores. L'accumulation de composantes permet d'obtenir des objets denses et plombés ou brillants et de timbre contrastant. Ce procédé est très ductile, parce qu'il permet de créer des masses dotées d'une grande mobilité (en raison de sa richesse spectrale), ce qui supporte le passage par des états différents [7] (fusion, ségrégation, condensation), au long du discours musical. De pareils résultats peuvent s'obtenir par des mélanges colorimétriques à l'intérieur des gammes et des graduations. Ainsi, par le mixage de couleurs primaires entre elles s'obtiennent des tonalités secondaires et complémentaires et celles-ci encore mixées donnent des nouvelles tonalités plus ou moins aléatoires (marrons, sépias, ocre, fuchsia etc.). La brillance ou le degré de condensation des tonalités ainsi obtenues dépendra du degré de saturation de chacune de ses composantes. Nous pouvons encore ajouter des valeurs spatiales en fonction du degré de transparence ou opacité appliquée à l'image.

3.3. Relation entre figure et fond

Le principe de figure et fond, propre à toute manifestation visuelle, trouvera son parallèle dans la musique. La vue et l'ouïe sont capables de percevoir et distinguer une quantité déterminée de phénomènes simultanément. Les deux sens ont une tendance à se fixer sur un objet précis, reléguant à l'arrière-plan tout ce qui est autour.

Si nous voulons configurer un champ morphologique capable de contenir dans un même niveau, des éléments visuels et sonores, nous nous trouvons face à deux paramètres de base : la surface et la figure.

Dans la perception auditive, nous entendrons avec plus de relief par exemple : les sons brefs, séparés, qui se détacheront nettement sur les surfaces en continu. La surface représentera ici le fond, par son caractère statique.

Dans l'œuvre graphique d'Escher nous trouvons des dessins symétriques qui montrent comment une surface peut être divisée et remplie avec des figures de formes similaires, sans pour autant laisser des espaces vides. Ces dessins sont d'excellents exemples de surface. (Fig. 2). À partir de ces modèles et avec l'aide de l'informatique, nous pouvons maintenant animer quelques-unes des figures qui constituent la surface, les séparer et les détacher du fond.

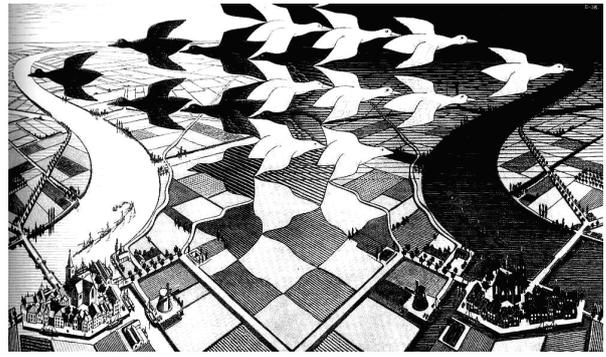


Figure 2. Nuit et Jour.

Dans cette gravure d'Escher, nous observons comme les rectangles gris vont se transformer en silhouettes d'oiseaux blancs et noirs, étant les uns l'inversion des autres. Les oiseaux noirs volent vers la gauche et les blancs vers la droite en deux formations opposées. À gauche de l'image, les oiseaux blancs fusionnent pour former la lumière du jour et un paysage et à droite les oiseaux noirs deviennent la nuit. Les paysages clairs et obscurs constituent aussi des images en miroir.

Un autre procédé que nous avons utilisé dans une image fixe représentant des objets à surface brillante est de recouper les reflets pour les mettre en mouvement, l'image fixe maintenue comme fond. De ce fait les reflets, en dépit de son caractère abstrait, deviennent des figures, tandis que l'objet concret défini par l'image fixe, va constituer le fond.¹

4. RELATION SPECTRE ET COLORIMÉTRIE

L'association timbre - couleur a été un concept traditionnel dans la terminologie musicale. Nous savons que ce paramètre musical est lié au spectre harmonique, par conséquent les procédés de filtrage nous permettront de modifier ses caractéristiques. Plus un son est complexe, plus importantes seront les possibilités de transformations spectrales.

Si nous associons donc la notion de spectre à la gamme chromatique en colorimétrie, nous pouvons aussi appliquer des procédés de filtrage qui nous conduisent à des nuances plus ou moins brillantes ou mates.

Comme nous l'avons mentionné plus haut, la notion de tridimensionnalité est en rapport avec les notions d'espace et de volume. La couleur, dans toutes ses possibilités de mélange, constitue aussi un facteur spatial de relevance dans l'animation vidéo.

En des termes picturaux souvent nous parlons de la gamme chromatique, qui contient la totalité des couleurs dans toutes ses graduations d'intensité. Nous avons ici un conflit de terminologie entre l'image et le son : la gamme chromatique en musique est tout sim-

¹ « Destellos » 2001

plement une série limitée de notes, mais ne contient pas la totalité des fréquences audibles.

Cependant, le concept d'énergie spectrale nous approche plus au concept d'intensité de la couleur. C'est pour cela que nous risquons de proposer le même type de traitement pour modifier le corps d'une couleur. Nous croyons que s'il est possible d'analyser les composants chromatiques d'une couleur déterminée, il sera possible aussi de réaliser des filtrages ou d'ajouter des tonalités (des fréquences en musique) à l'intérieur de son contenu chromatique, pour modifier son degré de transparence ou opacité. De ce fait, avec la suppression de bandes de fréquences d'une région de la gamme chromatique nous obtiendrions des tonalités intermédiaires.

5. STÉRÉOPHONIE ET DÉCALAGE CHROMATIQUE

Nous avons observé la possibilité d'établir des parallèles entre la structure stéréophonique et certains mouvements à l'intérieur de la gamme chromatique.

Pour produire un effet de glissement spatial sur un son, nous pouvons effectuer des procédés divers. Une stratégie simple pour créer cet effet spatial est de décaler les deux voix dans un enregistrement stéréophonique avec une différence minime de temps entre les deux.

Le procédé de délai nous permet d'obtenir aussi bien un rebondissement net qu'une légère sensation d'ampliation de l'espace sonore (quand le délai est très court et avec une proximité de moins de 20 millisecondes, l'oreille ne perçoit qu'un son.).

Quant à la perception de la couleur, cette même ampliation de l'espace sera donnée par l'absence de contours nets. Ainsi, par exemple, le décalage graduel (par fusion) d'une tonalité vers une autre produira des contours en dégradé. Ces effets sont possibles aussi bien entre des tonalités d'une même gamme ou de gammes complémentaires.

6. LA NOTION DE TEMPS

6.1. Mutations et éléments cinétiques

Un autre des problèmes que nous avons abordés dans notre recherche est celui des transformations graduelles. Par exemple, une structure déterminée (intégrée par des éléments de forme, couleur et taille), est transformée graduellement dans le passage d'un état à l'autre. C'est-à-dire : une surface se transforme en figure et vice-versa. Ce procédé implique la notion de temps. Il est bien connu que la perception auditive est plus rapide que la visuelle. L'effet cinétique n'a donc pas une correspondance exacte avec le son et par conséquent doit être considéré avec attention. Pour démontrer cet énoncé nous avons expérimenté avec le son d'un objet qui rebondit. L'ouïe est capable d'entendre les attaques de chaque rebondissement jusqu'à 20 millisecondes de séparation. L'œil, en re-

vanche, perçoit seulement les premiers sauts de l'objet et ensuite voit un mouvement unique, une illusion optique semblable à une ellipse. (Fig. 3)

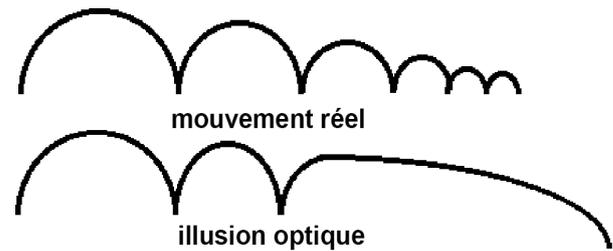


Figure 3. Rebondissement

Le phénomène de stroboscopie dans l'image et l'effet Doppler dans la musique sont deux excellents exemples d'illusion perceptives de ce genre. Dans les deux cas, il existe un décalage entre la vitesse du mouvement de la source et la rapidité de perception (auditive ou visuelle). Dans le cas du son, nous percevons une différence de fréquence selon la vitesse et la distance de la source, dans le cas de l'image nous avons une sensation d'inversion du mouvement.

Nous allons revenir sur cet aspect dans le paragraphe « forme et mouvement ».

7. LES MATÉRIAUX

Pour terminer avec les aspects relatifs aux éléments de la perception nous voulons faire référence brièvement aux matériaux. Personnellement nous travaillons la musique électroacoustique avec des matériaux concrets (c'est-à-dire, des sons enregistrés, non synthétiques). Nous pensons que c'est un bon élément pour créer une symbiose avec les matériaux visuels. Comme une forme de renforcer l'esprit musical avec le discours visuel et vice-versa.

L'association des matériaux sonores et visuels nous semble une stratégie très effective dans la vidéo-musique. Certains réflexes du cerveau tendent à associer un son avec la source qui le produit. Dans la plupart de nos travaux audiovisuels nous avons enregistré le son des objets photographiés pour les utiliser en synchronisation avec l'image en mouvement. Parfois, l'analyse des caractéristiques du matériau peut aider à trouver des bonnes corrélations pour créer et développer des idées artistiques intéressantes. Par exemple, l'analyse spectrale du son et l'analyse colorimétrique de l'image nous permettront d'établir des liens qui contribueront à assembler les matériaux de base pour un projet bien structuré.

8. FORME ET MOUVEMENT : ORGANISATION TEMPORELLE

Notre conception de la vidéo-musique ne se limite pas à l'association des matériaux, des couleurs, des formes et des sons, mais elle vise à la création d'un discours poétique, issu de l'animation des images en combinaison avec les sons. Ainsi pour obtenir la for-

malisation et le développement d'une structure audiovisuelle organique, nous devons tenir compte de plusieurs aspects.

En ce qui concerne la forme temporelle, la musique nous offre une grande flexibilité. Nous pourrions affirmer que c'est son domaine, comme l'espace a été considéré traditionnellement le domaine du visuel. Cependant, grâce aux procédés d'animation que nous propose l'informatique, les morphologies visuelles peuvent acquérir une mobilité et une flexibilité non négligeables. Par exemple, la notion de durée du son (court – long), peut s'appliquer à la longueur ou largeur d'une figure. Même si dans la perception visuelle d'une figure plane le temps n'existe pas, les traitements numériques d'animation nous permettent d'incorporer cet aspect, de façon que la permanence de la figure dans notre champ de vision, peut varier de l'infiniment fugitif à l'extrêmement long. En même temps, cette sensation peut se produire graduellement, par la définition progressive de ses contours, ce qu'impliquera un temps plus long pour l'élaboration intellectuelle qui nécessite la perception de l'image.

La plupart des manipulations appliquées à l'animation de l'image, sont basées sur des transformations et mutations de largeur et perspective. Il est possible, par exemple, de produire des effets d'extension d'une spirale, déformation d'une sphère, étirement d'un cube, etc., simultanément, c'est-à-dire par transformation graduelle.

8.1. La chaîne du développement

Dans notre conception esthétique musicale, la notion de forme implique l'organisation de tous les paramètres du discours musical, conduisant vers un contenu expressif. Cette conception signifie un développement temporel, à travers lequel les idées musicales trouveront sa cohérence et son unité. Ce développement des idées musicales suivra un parcours, plus ou moins accidenté, en passant par des transitions, respirations, ponts, déclinaisons, etc. Ces accidents constitueront les éléments porteurs de forme du discours. (Fig. 4)

Nous avons donc quatre facteurs de structuration fondamentaux : temps, directionnalité, dynamique et transitions. À travers ces éléments, nous circulerons par des moments de Tension et de Repos, nécessaires à toute narrativité temporelle et audiovisuelle, même dans les langages les plus abstraits.

Les sensations de tension et repos peuvent s'exprimer par des constructions opposées en vitesse, intensité, tessiture (rapide-lent, forte-piano, grave-aigu), ou par des combinaisons spectrales plus ou moins denses ou complexes ; par des mixages plus ou moins ouverts ou fermés, etc.

Par rapport à l'image nous pouvons trouver des corrélations possibles avec les paramètres de la forme et largeur des objets, pour les transformations graduelles, et dans les champs de la colorimétrie et la luminosité.

Le discours audiovisuel prendra forme graduellement à travers les différents modes d'articulation et des transitions entre les séquences. Ces transitions peuvent s'obtenir à travers des changements des éléments de texture, de dynamique, de registre, par fusion ou ségrégation des matériaux, par accumulation d'événements etc.

Ainsi nous pouvons associer les effets de vibrato ou glissando avec des filtres d'animation comme le *ripple* ou la dissolution de l'image. Dans de tels cas, les effets cinétiques sont en parfaite correspondance et en équilibre avec la musique.

Vers la fin de notre vidéo « Destellos », nous pouvons observer des cristaux de neige qui se transforment en glace et ensuite se décomposent jusqu'à devenir un tissu presque transparent. La musique s'assimile à ce paysage avec des sons de craquements qui deviennent progressivement des grains infinitésimaux.

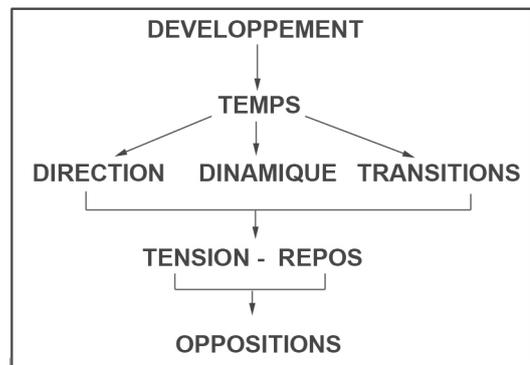


Figure 4. La chaîne du développement

Si nous suivons ce schéma de développement formel, la structure sera équilibrée.

Nous pouvons affirmer aussi, que les associations des effets perceptifs jouent un rôle important dans la structure temporelle de la forme audiovisuelle. L'unité de sens, la cohérence esthétique, la sensibilité expressive, dépendent des multiples interrelations entre les paramètres matériels, temporels, physiques et perceptifs qui participent dans le développement de l'œuvre.

L'histoire du cinéma nous offre de nombreux exemples de cette conception esthétique. Si nous dégageons l'élément narratif du scénario et nous fixons notre attention sur le traitement du temps, les changements de lumière, le cadrage et l'utilisation des différents plans spatiaux, nous pouvons trouver des magnifiques exemples de vidéo dans beaucoup des films classiques de réalisateurs comme Eisenstein, Hitchcock, Fassbinder et tant d'autres.

9. LA TECHNOLOGIE

Pour réussir cette unité formelle, l'artiste doit être capable de maîtriser les outils pour créer des nouvelles perspectives et formes d'illusion, selon sa conception esthétique. L'aspect technologique n'est pas

l'objectif de ce travail, cependant nous ferons une brève référence pour signaler quelques réflexions sur le sujet.

Nous croyons qu'il est important de pouvoir se soustraire aux charmes des moyens technologiques. La séduction qu'exercent les outils numériques provoque souvent des réalisations pléthoriques de fugue et d'enthousiasme, mais exempts de réflexion esthétique.

L'artiste doit posséder le pouvoir de transformer la nature pour la convertir en œuvre d'art. Si les outils se manifestent en forme trop évidente, c'est n'est plus l'artiste mais la machine qui fait l'œuvre. Pour cela il est important de maîtriser les ressources informatiques afin de les diriger vers notre propos esthétique. Cette maîtrise s'apprend avec une bonne dose de curiosité et exploration personnelle. Il n'y a pas de règles établies. Chaque artiste sera responsable de ses propres trouvailles créatives, qui constitueront son style personnel.

La plupart des ressources technologiques nous offrent la possibilité de modifier les *presets* du fabricant pour créer nos propres modèles et les sauvegarder, pour pouvoir les appliquer plus tard à d'autres images ou sons. Cependant il est important d'observer le comportement intrinsèque de chaque type d'image et de son pour appliquer le traitement le plus convenable. Toutes les matières ne supportent pas les mêmes effets. Dans les traitements en 3D, par exemple, la qualité des textures de l'image est un facteur fondamental. Certaines textures opposent résistance à la forme en trois dimensions et peuvent devenir brutes. Les mouvements ondulatoires peuvent résulter de grands effets pour créer des images abstraites de contours imprécis. Le même procédé, appliqué à une image avec des bords nets, aura comme résultat une figure trop géométrique et banale.

Des pareilles considérations pour le traitement du son peuvent être prises en compte. L'artillerie de *plugins* dont nous disposons actuellement nous permet de réussir des transformations inusitées. Cependant, nous devons observer la qualité et le comportement des matériaux, si nous voulons réussir des résultats perceptifs intéressants. L'analyse morphologique et spectrale des matériaux est donc importante pour nous permettre de « doser » les traitements selon les limitations du matériau.

10. CONCLUSION

Se dégagent donc de toutes ces descriptions et réflexions que : d'une association intelligente entre les éléments techniques, la poésie, l'imagination et la créativité, émanera une bonne structuration.

Avec cette présentation, j'ai essayé de décrire certains processus de mon propre travail de création, dans le but de trouver des réponses à la signification de ce nouveau langage. La recherche scientifique et technologique pourra fournir des réponses plus approfondies. En effet, la recherche représente le cadre sur lequel s'appuie la pensée artistique puisque c'est à travers des modèles que s'explicite le résultat ultime de la créati-

tivité. Parfois ces modèles signifient une aide et une stimulation pour la compréhension et pour l'appréciation esthétique, parfois ils posent une entrave à la sensibilité. Ce problème de la synergie entre différentes disciplines a été défini pertinemment par Hugues Dufourt.[8]

« Le chercheur scientifique vise des valeurs de connaissance, de preuve, de vérité. Le chercheur technologique vise une efficacité maximale et recherche un optimum rationnel dans la mise au point des procédés. L'artiste s'approprie souvent des modèles, mais il les détourne de leur finalité théorique, les transpose, les gauchit, inverse les normes logiques et pose parfois des problèmes inédits à la science ou à la technologie. Ce peut être dans l'irrégularité ou le déficit des modèles scientifiques que certains musiciens découvrent des perspectives artistiques fécondes. L'artiste fait souvent interférer deux échelles de grandeur ou de complexité pour produire des effets esthétiques ou des objets artistiques insolites. »

11. RÉFÉRENCES

- [1] Justel, Elsa, « La ilusión sensible », Institut de l'Audiovisual-Université Pompeu Fabra- Barcelone-Espagne, (1995), « Destellos » (2001), « Un regard sur la ville » (2004).
- [2] Escher, Maurits Cornelis, *L'œuvre graphique*, Benedikt Taschen Verlag, Berlin GmbH. 1990.
- [3] Whitney John, *Musique Film*, Yann Beauvais, 1980, p. 29.
- [4] Escher, *op. cit.*
- [5] Risset, Jean-Claude, « Musique, recherche, théorie, espace, chaos », in *revue Inharmoniques*, N°8/9, IRCAM/Centre Pompidou, Paris, pp. 273-313 ; (1991), in Justel, Elsa : « Les structures formelles dans la musique de production électronique », Editions du Septentrion, 2002, p.62.
- [6] Escher, *op.cit.*
- [7] McAdams y Bergman, « Hearing musical streams », in *Computer music journal* 3(4), 1979, p.26-43.
- [8] www.education.gouv.fr/rapport/risset

UNE APPROCHE POLYVALENTE DIRECTION MUSICALE DANSEE, CAPTATION GESTUELLE CAUSALE

Jean-Marie Adrien
jma@jeanmarie-adrien.net

RÉSUMÉ

Cet article traite de la relation entre le mouvement dansé et la musique, en lien avec notre tradition occidentale tout d'abord, par le biais de la direction musicale dansée, puis, sous l'angle des nouvelles technologies, par le biais de la captation gestuelle abordée sous l'angle causal.

La direction musicale dansée est une relation nouvelle entre musique et danse par laquelle le mouvement dansé devient la musique vivante : un danseur-musicien dirige, par sa danse, un ensemble instrumental placé sur scène. La captation gestuelle, envisagée sous l'angle causal, est la modélisation informatique de la direction musicale dansée, le danseur musicien actionnant un ensemble instrumental virtuel par le biais d'un dispositif de captation de son mouvement.

La formalisation de ce travail a été réalisée *a posteriori*, après écriture, production et présentation sur scène dans le cadre du spectacle vivant. Il ne s'agit donc pas ici de la présentation d'un travail de recherche à proprement parler, mais plutôt d'une tentative d'objectivation, une théorisation ex-post, illustrée par des exemples qui sont du ressort de l'auteur et donc subjectifs par nature.

1. INTRODUCTION

La relation musique danse, ou, plus généralement, la relation du mouvement au son, est un sujet très ancien, elle constitue un archétype qui trouve probablement son origine dans le fait que le son physique, pour exister, s'appuie le plus souvent sur un mouvement : une énergie cinétique, par exemple, transformée par le biais d'un instrument en énergie acoustique, ou encore un souffle, mouvement de l'air, transformé en son dans un tuyau etc.

Son et mouvement sont très profondément et très étroitement liés dans l'esprit humain, ils le sont de façon causale, instrumentale. Cette relation causale se matérialise, au cours des époques, et selon l'évolution de notre civilisation, par un grand nombre de dispositifs, du plus archaïque au plus sophistiqué. Il va sans dire que cette relation est source de plaisir, de jouissance, de jubilation, de transe etc.

Pour le sujet qui nous intéresse, à savoir le spectacle vivant, il faut ajouter à cette personne - ce musicien danseur qui actionne donc par son mouvement un dispositif causal produisant du son - un tiers qui observe son semblable aux prises avec cet instrument.

Ce tiers, placé dans la proximité du premier, peut engager un processus d'identification empathique, lui aussi très ancien, reptilien en quelque sorte, qui le fait participer à ce qu'il voit de façon active et partager ainsi le plaisir de l'exécutant [5]. La relation empathique s'appuie sur une bonne compréhension par

l'observant des mécanismes mis en jeu par l'exécutant. Cette compréhension est indispensable, et notamment, ajouté aux facteurs culturels, le caractère causal du dispositif dont dispose l'exécutant est capital en ceci qu'il assure cette compréhension, la capacité de l'observant à anticiper, etc.

Il s'agit de relations archaïques qui sont au fondement du spectacle vivant, et qui sont, bien entendu, encore totalement opérantes de nos jours.

Pour illustrer ce qui figure plus haut dans notre contexte actuel, on pensera à une situation ordinaire de concert, dans laquelle le public, en effet, respire avec la pianiste, anticipe le son et le reçoit donc d'une façon différente à la manière avec laquelle il le recevrait si l'instrumentiste était placée derrière un rideau, ou, pire, diffusée à partir d'un enregistrement.

De la même façon, placé à proximité du danseur, le public danse avec lui parce qu'il a, lui aussi deux jambes, deux bras, et qu'il est, lui aussi en prise avec la gravité [4].

La causalité est au fondement de l'empathie : si dans l'exemple précédent le danseur s'envole subitement dans les airs, contredisant ainsi le principe causal et contredisant donc toute attente, il s'agit là encore d'un jeu spectaculaire qui s'appuie sur la relation causale, en opposition à celle-ci.

La relation empathique fonctionne en permanence, de façon sous-jacente. Bien entendu, elle n'interdit pas la présentation d'un discours d'un autre niveau qui s'adresse non pas au cerveau reptilien de l'observant, mais à son cerveau analytique : le public observe la violoniste, engage une relation empathique avec elle, il profite également de l'architecture de la pièce, ou encore des idées qui lui sont présentées, du scénario etc. Il actionne parallèlement ces différents niveaux de perception : schématiquement, ce va-et-vient, cette articulation entre propos emphatique et propos analytique est un fondement des arts de la scène.

On pourrait objecter que ce préambule relève de l'évidence, pourtant il ne va pas de soi : je n'ai jamais entendu parler de causalité, d'observant et d'empathie, au cours de mes études musicales ou chorégraphiques par exemple, ni au cours de la carrière qui a suivi - auprès d'institutions de premier plan dévolues au spectacle vivant.

Ensuite, pour ce qui concerne l'utilisation de technologies de captation sur scène, ce qui semble aller de soi dans ce que l'on peut lire, voir et entendre, c'est l'inverse : le principe causal, la causalité est au contraire une sorte de repoussoir. Très souvent la relation causale est présentée comme mauvaise par nature, faible, trop simple, gratuite, etc. En bref, la captation ne devrait surtout pas être causale : nous allons examiner plus bas les raisons probables de cet état de fait.

Par ailleurs, on peut tenter, sur scène, de s'extraire de la relation empathique, de la refouler vers une sorte de neutralité, au profit d'un message, d'idées, d'un concept ou d'une pensée¹. Néanmoins, pour un observant normal, l'absence d'empathie est souvent synonyme de désintérêt et d'ennui : l'alternative serait probablement de présenter ce message, cette pensée et ces idées dans un livre, par exemple, plutôt que sur scène.

Enfin, dans cette ligne et de façon plus générale, il semble que l'on ait souvent tendance à oublier l'observant dans le domaine de la médiation artistique des arts vivants dans notre pays depuis une quarantaine d'années, et, plus récemment, dans celui de la médiation arts-technologie, avec pour conséquence directe que l'affluence dans nos salles n'est pas ce qu'elle pourrait être. Cet observant qui, pourtant, est un constituant indispensable du spectacle, et que l'on ne devrait pas évacuer pour rester entre soi.

Autant de facteurs qui justifient, à mon avis cette introduction qui n'a pas dans la pratique, le caractère évident que l'on pourrait attendre et qui fonde le travail présenté plus bas.

2. DIRECTION MUSICALE DANSÉE

La relation du mouvement au son est immémoriale, notre travail s'inscrit naturellement dans une filiation extrêmement riche, dont nous pouvons extraire quelques exemples à des fins d'illustration : après Cro-Magnon avec sa flûte d'os et ses grelots, on peut citer - l'os se transformant en canne dans une ellipse cinématographique - Jean-Baptiste Lully, chef d'orchestre danseur, Fred Astaire avec ses claquettes, et, plus récemment, Franck II Louise, compagnie de danse hip-hop équipée de capteurs [7].

Il faut noter que, dans cette filiation et dans notre culture, en dehors du geste instrumental, le rapport du mouvement au son tend à se distendre conjointement à l'évolution de la civilisation, comme conséquence de la sophistication grandissante de l'expression artistique et de la spécialisation des métiers : les chefs d'orchestre danseurs après Lully ne sont pas très nombreux.

C'est néanmoins à partir de cette perspective du chef d'orchestre que nous nous intéressons à la relation du mouvement au son : nous ne considérons pas ici la fonction d'organisateur du chef d'orchestre ; mais nous nous intéressons à sa fonction d'intervenant sur scène, et au mécanisme de communication qu'il instaure vis à vis des instrumentistes et du public dans une situation de musique vivante.

Dans cette situation tout à fait particulière, le médium du chef d'orchestre est, par excellence, l'empathie : excellence parce que nous sommes ici dans la sophistication ultime de la relation du mouvement au son - dans le cadre de la musique écrite et de notre culture occidentale ; empathie parce que le chef ne produit directement aucun son, et que le son résulte pourtant de son mouvement, de son état de corps.

Cette relation empathique - absolument causale - joue envers les instrumentistes, dans un aller-retour de chaque instant, mais également envers l'observant, le public, et c'est cela qui, en dernière instance, nous intéresse.

De façon sommaire et d'un point de vue technique, le chef impulse autant qu'il reçoit. Son mouvement peut prendre une multitude de formes dès l'instant où il est conduit par une conscience, une écoute et une anticipation musicales. Ce mouvement s'appuie notamment sur le temps : le chef façonne, en fonction de ce qu'il reçoit, le temps optimal pour que les instrumentistes jouent le mieux possible.

La relation entre son mouvement et la matière sonore est à ce point fine et causale, que ce mouvement *est* la musique vivante en train de se développer : il est d'une nature particulière, il est doté d'attributs très spécifiques, comme résultat de la situation dans laquelle il prend corps, on peut écrire qu'il s'agit d'un « mouvement qui est la musique ».

Or, ce mouvement peut prendre une multitude de formes. Dans notre culture, on dirige habituellement avec la main, parce que la main est très agile et qu'elle peut refléter la conscience musicale assez directement, sans une préparation considérable.

Mais, si le corps est agile, rien ne s'oppose à ce que l'on dirige avec le corps : c'est ce que nous appelons la direction musicale dansée (DMD). Un musicien-danseur dirige par ses mouvements un ensemble instrumental placé sur le plateau, à la manière d'un chef d'orchestre. L'information de direction est répartie sur le corps entier, dans une écriture que l'on ose qualifier de chorégraphique [2].

Sur le plan pratique, l'information de direction musicale est, dans un premier temps, débarrassée de tout ce qu'elle contient de redondant et de superflu. L'information restante, essentielle, est alors ventilée ici sur le genou, là sur la hanche et ici sur le talon.

Cette écriture chorégraphique est convenue en répétition avec l'ensemble instrumental, et, ce dernier reçoit l'information - qu'il attend donc sur le genou, sur la hanche et sur le talon - comme il le ferait s'il suivait la main ou une baguette ; sachant donc que, comme la main, le genou, la hanche et le talon sont conduits par une conscience, une écoute et une anticipation musicales.

Il va sans dire que le danseur doit être un excellent musicien et qu'il doit avoir mémorisé la partition pour diriger sans elle.

Cette technique est compatible avec le mode opératoire habituel des instrumentistes, qui, semble-t-il, prennent plus de plaisir à suivre la hanche d'une jolie danseuse dirigeant de mémoire, que le bras d'un chevelu en frac lisant sa partition. Les répétitions ne sont pas plus nombreuses, simplement, elles se font sur un plateau de danse.

Cette technique est compatible également avec l'écriture instrumentale contemporaine usuelle, et, comme dans le cas de la direction musicale ordinaire, la relation au mouvement est d'autant plus riche et

¹ le « mon spectacle ne fonctionne pas sur l'empathie » venant remplacer le génial « ma musique n'est pas faite pour être entendue »

jubilatoire que l'écriture musicale est riche du point de vue rythmique.

Le mouvement de DMD, lu en direct par les instrumentistes et s'appuyant sur cette lecture, *est* la musique vivante en train de se développer. Cette technique permet de produire et de travailler un matériau chorégraphique spécifique, ce « mouvement qui est la musique » que nous évoquons plus haut.

D'une façon prosaïque, nous pouvons écrire que la DMD renouvelle la technique de direction musicale par l'émancipation complète du mouvement, lui conférant la complexité du mouvement dansé, et lui donnant ainsi une forme plus riche, plus moderne, plus spectaculaire, plus engagée.

Réciproquement, du point de vue de la danse, la DMD permet de créer et de travailler un matériau nouveau, « le mouvement qui est la musique ». On peut écrire, d'une manière plus simple, qu'elle permet de doter la danse d'une signification musicale pertinente et opérante sur scène.

Quel est le sens d'une telle démarche dans le cadre du spectacle vivant ? En référence à notre propos introductif, il apparaît que l'observant, situé dans la proximité du chef danseur, peut mettre en œuvre un double processus empathique d'identification, avec un interprète placé dans la jubilation créer du son par sa danse.

L'observant, le public, n'a donc pas besoin de connaître la technique de direction musicale, ni la teneur de cet article : il lui suffit de jubiler en laissant aller son cerveau reptilien.

Notons que l'observant ne peut pas mettre en œuvre un tel processus lorsqu'on lui présente les dispositifs habituels de mise en relation de la danse avec la musique : orchestre en fosse, ensemble sur le plateau mais sans relation véritable avec les danseurs, ou encore musique enregistrée. La DMD prend donc tout son sens dans le cadre du spectacle vivant.

Concrètement, elle a été élaborée sur le terrain, au fil des projets, en réponse à une frustration quant à la véritable relation entre musiciens et danseurs sur scène en situation de production.

Il faut noter enfin que la DMD est une conséquence directe d'une approche polyvalente musique-danse : pour la pratiquer, il suffit d'être un excellent musicien et un excellent danseur, comme au temps de Lully.

Les limites de l'approche, dans l'état, sont les suivantes :

Tout d'abord, la direction musicale dansée est lourde de mise en œuvre : il est difficile de produire, en dehors des institutions, un plateau comportant cinq à six interprètes tous les six mois.

Par ailleurs, en supposant que l'on dispose du financement, il est impossible de chercher ou d'écrire en présence des musiciens. Le planning de répétitions impose d'inventer le mouvement de direction musicale en amont, en l'absence des musiciens, et d'en constater le résultat en répétition sans pouvoir fondamentalement le modifier.

D'autre part, la DMD ne permet pas, dans l'état, d'improviser, or la phase d'improvisation est sans doute primordiale dans la découverte et l'exploration du « mouvement qui est la musique ».

Enfin, en situation de direction musicale dansée, il persiste une séparation ultime sur le plateau entre instrumentistes et musiciens-danseurs en ce sens où le musicien-danseur, comme le chef, ne produit pas directement de son. En conséquence, il existe, sur le plateau, un plan B vers lequel les musiciens peuvent se replier en cas de défaillance de la direction, comme au concert : en cas d'accident, les instrumentistes peuvent choisir de se synchroniser entre eux à la façon d'un ensemble de chambre, ce qui constitue, du point de vue de l'observant, une perte de qualité.

Ces raisons pratiques nous ont conduit à réaliser la modélisation informatique de la direction musicale dansée - la captation gestuelle causale - le recours à technologie permettant de contourner la plupart de ces limites.

3. CAPTATION GESTUELLE CAUSALE

S'il n'y a pas eu, à notre connaissance, de nombreuses tentatives d'émancipation de la direction musicale vers la danse, ou réciproquement de la danse vers la direction musicale ; en revanche, les dispositifs de captation gestuelle ont été, dans un passé récent, mis en œuvre à maintes reprises dans le spectacle vivant, les références étant trop nombreuses pour être citées.

Dans notre travail toutefois, la captation gestuelle est envisagée comme le modèle informatique de la direction musicale dansée, et la technologie est convoquée, au départ, afin d'en contourner les limites.

C'est la raison pour laquelle notre approche, ses résultats et ses perspectives, sont spécifiques, et c'est la raison pour laquelle nous avons présenté en détail la DMD dans la première partie de cet article.

En particulier, nous abordons cette problématique de la captation dans un angle causal, le modèle de référence, la direction musicale dansée étant elle-même fortement causale. Il s'agit donc de captation gestuelle causale (CGC).

Dans notre travail, le danseur musicien conduit par sa danse un ensemble instrumental virtuel, comme il le faisait en situation de direction musicale dansée : il a mémorisé la partition, et il pilote, en fonction du contexte, des fragments musicaux correspondant à des sections (un pupitre de contrebasses) des incises (un fragment de clarinette) des notes isolées (une attaque de grosse caisse) ou tout autre son et matière sonore [3].

Le dispositif embarqué a été développé à l'Ircam par Emmanuel Fléty [6] : il s'agit, pour chaque chef danseur, d'une Wise Box connectée à quatre accéléromètres bi-dimensionnels. Les données de captation sont traitées par un logiciel temps réel type Pure Data, alimentant lui-même un moteur de synthèse sonore. Les instruments virtuels sont, dans l'état, des fichiers texte - des partitions instrumentales - synthétisés en temps réel à partir d'échantillons à un tempo flexible.

Il s'agit d'un modèle de la direction musicale : l'interprète polyvalent délègue à la machine la réalisation d'un certain nombre de tâches, comme le fait le chef avec les instrumentistes.

La grandeur physique mesurée est l'accélération, ce qui caractérise bien le mouvement, mais ce qui réduit considérablement, dans le modèle, le phénomène de communication empathique qui s'établit entre le chef et les instrumentistes : on sent bien que, lorsqu'on modélise un ensemble instrumental avec un ordinateur, il va en résulter une perte, en ce sens qu'une machine va difficilement remplacer un ensemble d'êtres humains.

En revanche, la communication empathique entre le chef-danseur et le public est conservée en grande partie, dans le sens où le chef-danseur est dans une situation instrumentale, aux prises avec un dispositif de captation causal dont le fonctionnement est bien compris, et dans le sens où l'accélération est un bon support de l'empathie, du point de vue de l'observant.

La captation gestuelle causale permet notamment, comme la DMD, de produire et de travailler ce matériau que nous avons qualifié de « mouvement qui est la musique », et, du point de vue de l'observant, d'établir le même double lien empathique.

Dans le travail de ce matériau, un certain nombre de choses sont perdues par rapport à la situation où l'on dispose d'un véritable ensemble instrumental, mais les limites de la DMD que nous avons évoquées plus haut sont contournées :

Il est possible d'explorer par l'improvisation, les coûts de production sont réduits de façon significative, et, enfin, le chef-danseur produisant lui-même du son, la fusion entre ce dernier et l'ensemble instrumental serait complète, dans le cas où CGC et DMD seraient associées.

Les limites techniques du modèle, dans l'état, sont liées notamment aux limites des outils de synthèse en temps réel - ici l'échantillonnage. Il serait intéressant d'utiliser dans l'avenir des méthodes de synthèse sonore plus sophistiquées [1].

D'autres limites, plus importantes, sont culturelles : elles sont liées à la médiation du spectacle vivant en général. La majorité des médiateurs de l'offre culturelle, quand elle n'est pas accaparée par sa propre cooptation, son esprit grégaire ou sa myopie, n'évalue pas l'offre culturelle dans une perspective polyvalente, intégrée, comme le fait le public, mais dans une perspective de spécialiste, danse, musique ou théâtre.

Ainsi par exemple, cette majorité n'a pas vu ni entendu, au cours de l'année passée un concert, ni un chef d'orchestre ; ou, autre exemple, cette majorité admet à l'évidence, que dans un contexte de spectacle vivant, il est possible de présenter des danseurs sur un disque, ce qui, dans notre perspective, constitue une dégradation insoutenable : en fait, la musique vivante, qui est le référent de la captation gestuelle causale, tend peut-être elle-même à disparaître.

À la vue d'une situation de CGC donc, une oreille non avertie peut nous objecter que, finalement, c'est comme un disque, tout simplement parce qu'il n'y a pas d'ensemble instrumental apparent sur le plateau.

Nous savons bien qu'il n'en est rien mais nous sommes obligés de considérer une telle observation, en particulier si elle provient d'un inspecteur spécialisé dans le théâtre ou la danse.

4. CONCLUSIONS

Ce travail conduit à poser quelques remarques de caractère général sur l'emploi des outils de captation et

plus généralement, des outils technologiques, dans le cadre du spectacle vivant, ainsi que sur le sujet de la pluri-disciplinarité :

Une première remarque est liée à la technologie :

Sur le plan pratique, le caractère causal d'une application de captation gestuelle est ajustable au stade de la programmation informatique. Tout est possible entre une application fortement causale - une application instrumentale type grelots associant une note à un mouvement - et une application non causale - le déclenchement de la lecture d'un disque, ou le contrôle de textures à évolution lente. La véritable question, on l'aura bien compris, est l'intérêt de l'une ou de l'autre du point de vue de l'observant.

Le cas limite est celui, habituel, où l'on équipe un excellent danseur avec des capteurs, et où, bien évidemment, on lui confie des processus non causaux, parce que, pour prendre en charge un dispositif causal, ce danseur devrait également être un musicien virtuose, ce qu'il n'est pas.

La conséquence directe est que le dispositif technologique perd immédiatement son utilité, avant même que l'on mette un pied sur le plateau ou que l'on écrive une note : pourquoi ce danseur est-il bardé de capteurs si l'observant n'en perçoit aucune conséquence ?

Si l'on se hasarde toutefois à lui confier des processus causaux, le résultat est alors très fortement limité par son incapacité musicale.

De la même manière, si l'on équipe un instrumentiste avec des accéléromètres on va inévitablement se heurter à des limites en matière de mouvement. Dans ce cas néanmoins l'approche est le plus souvent, causale ; les limites, du point de vue de l'observant, proviennent des contraintes pesant sur le geste instrumental.

Il semble donc que, dans le cadre du spectacle vivant, la captation gestuelle gagne à s'adresser à des interprètes polyvalents afin d'être pertinente du point de vue du public.

Il ne s'agit pas, bien entendu, d'être, quant à la causalité, perpétuellement dans une position démonstrative. Il s'agit plutôt, dans l'écriture et dans l'interprétation, d'en jouer librement, sans contrainte et sans impuissance.

Plus généralement, on peut formuler l'hypothèse que le bon usage de la technologie sur scène s'appuie sur un effet du premier ordre, du point de vue du public, l'utilisation au premier ordre permettant de porter le lien empathique.

Par exemple, si nous envoyons des décharges électriques à un danseur sur scène afin de modifier son mouvement - avec un faible voltage - l'utilisation que nous faisons de la technologie est du second ordre. Les décharges électriques ne concernent pas directement le public, et on pourrait probablement, pour le même résultat, en faire l'économie. À l'inverse, si le danseur est capable d'interpréter de façon instrumentale une partition musicale, alors l'utilisation de la technologie est du premier ordre, directe, et elle permet en effet l'établissement du lien empathique.

Une autre façon de déterminer l'ordre d'investissement de la technologie est de se demander si la représentation pourrait avoir lieu avec un dispositif en panne ; et, dans l'affirmative, si le résultat serait

significativement différent du point de vue de l'observant.

En retour, une pratique du premier ordre de la technologie induit souvent une redéfinition profonde de l'interprétation sur scène et de l'écriture, à l'opposé d'utilisations anecdotiques et autres collages ou juxtapositions. Dans le cas de la captation gestuelle causale, par exemple, la technologie requiert rien de moins que l'acquisition de deux métiers d'interprète, - musicien *et* danseur - et permet la mise en évidence et l'exploration d'un matériau chorégraphique nouveau.

Tout ceci n'interdit pas de convoquer la technologie au deuxième ou troisième ordre, bien entendu : dans ce cas néanmoins, le propos n'est pas la technologie elle-même, le propos - et l'empathie - est ailleurs ; et il est important, alors, de ne pas évaluer la technologie et son intérêt potentiel en fonction de cette mise en œuvre particulière.

C'est peut-être ce dont a souffert la captation gestuelle appliquée à la scène, pour laquelle nous avons vu au cours des dernières années et en premier lieu des réalisations très visibles et très médiatisées, basées sur des approches non polyvalentes, ce qui a profondément handicapé aux yeux de tous l'idée même de la captation, avec pour effet les opinions que nous décrivions plus haut, selon lesquelles par exemple la captation ne doit surtout pas être causale.

Une deuxième observation générale est liée à la polyvalence :

La direction musicale dansée et la captation gestuelle causale sont les résultats directs d'une approche polyvalente, une approche intégrée, qui permet d'aller au delà de limites infranchissables par les méthodes habituelles pluri- ou multi-disciplinaires, qui procèdent, invariablement, par juxtaposition d'experts, dans un schéma ancien de division du travail.

Pour faire évoluer la relation musique-danse, il est peut-être nécessaire d'être musicien et danseur à la fois, et de le revendiquer, envers et contre tout ce qui s'oppose, aujourd'hui encore, à la polyvalence.

La même remarque vaut aussi probablement pour le croisement art-technologie, pour lequel la pensée habituelle, les cadres de recherche et de financement, les institutions fonctionnent souvent en conformité avec un modèle qui date d'un demi siècle : on tend encore invariablement, comme on le faisait il y a cinquante ans, à associer un chercheur et un artiste, le chercheur refoulant l'artiste qui est en lui et inversement.

Si, à l'époque, le modèle était pertinent, dans la mesure où, notamment, les équipements informatiques n'étaient accessibles qu'en laboratoire, il semble qu'il ait aujourd'hui perdu une partie de son sens et qu'au contraire, dans un certain nombre de domaines au moins, les approches intégrées soient nécessaires. Elles devraient, à ce titre, être suscitées, encouragées et reconnues. En effet, paradoxalement, la complexité accrue des problèmes ne peut pas se résoudre, dans certains cas, par l'accumulation d'experts de plus en plus spécialisés.

A cet égard, la captation gestuelle est un exemple pour lequel l'écriture artistique et le développement technologique sont très étroitement imbriqués : dans une approche polyvalente - en contraste avec ce que l'on peut attendre d'une équipe d'experts regroupant un ingénieur, un compositeur, un chorégraphe et un interprète - l'intuition artistique et l'intuition

technologique se fertilisent l'une l'autre dans un aller-retour très productif et très ... jubilatoire.

5. RÉFÉRENCES

- [1] Adrien, JM. *The Missing Link : modal synthesis*. in Representations of Musical Signals, MIT Press, Cambridge 1991, pp 269-298.
- [2] Adrien, JM, *L'intrus*. Pour trois chefs danseurs, ensemble instrumental et disp. électroacoustique, Commande de l'Etat 2004, Site internet <http://jeanmarie-adrien.net>
- [3] Adrien, JM, *Vue sur Jardin* Pour deux chefs danseurs et dispositif de captation gestuelle causale, Commande de l'Etat 2006, Site internet <http://jeanmarie-adrien.net>
- [4] Akinleye, A. « *Geography of the Body* », Dance UK News, Issue 70, Autumn 2008, Paris, 2005.
- [5] Decety, J; & Meyer, M. « *From emotion resonance to empathic understanding: A social developmental neuroscience account.* » in Development and Psychopathology, 20, 1053-1080
- [6] Flety, E. *Wise Box : a Multi-performer Wireless Sensor Interface using WiFi and OSC*, NIME 05.
- [7] Franck II Louise, « *Konnecting souls* ». [En ligne] Disponible <http://www.culture.fr/creation-numerique/pdf/konnect.pdf?11689733152>

SIMULTANÉITÉ ET COINCIDENCE, UNE PROPOSITION DE RÉFLEXION LIÉE AUX PRATIQUES ARTISTIQUES TRANSVERSALES

Luc Larmor

Ecole des Beaux-Arts de Rennes

Université Rennes 2

Groupe d'Expérimentation Plastique du Sonore

contact@artplug.net

RÉSUMÉ

En situation de création croisée ou de propositions interdisciplinaires, il est parfois difficile, même si chacun sait que toute nouvelle hypothèse ou proposition, sans venir questionner de fond en comble sa propre discipline, peut entraîner pratiquement de profondes et durables remises en question, d'appréhender pleinement l'impact d'un choix ou d'un dispositif sur un projet de création transversal.

Porosité du lieu, comportement du public, modalité de diffusion, complexité des médias, nombreux sont les éléments qui (inter)agissent dans une proposition artistique interdisciplinaire. Cet article se propose d'aborder la complexité propre aux créations ouvertes, contextualisées ou déployant des flux, par le biais d'une notion de co-présence. Co-présence d'éléments qui peuvent se voir, se lire ou s'écouter en eux-mêmes, composantes à part entière, ou qui peuvent se comprendre en regard de ce qu'ils sont susceptibles de s'apporter mutuellement : décalages significatifs, réunions improbables, trajectoires et volumes, vecteurs de destination, d'emplacement, en dehors de toute référence directe à leurs attributs de base.

1. INTRODUCTION

La question est de savoir comment, face à des formes assurément fragmentaires, des constructions volontairement ouvertes, aux composantes parfois non-locales ou au sein de processus plus ou moins aléatoires, mettre en évidence des modes de connexion ou de contact peut-être plus spatiaux ou temporels, que discursifs et symboliques, plus cartographiques que syntaxiques.

Comment, en présence d'une réalité artistique complexe et hybride, environnementale et contextuelle, parcourue de protocoles de mise en présence tout azimut, envisager développer un champ artistique propre, qui ne soit pas (simplement) la somme d'éléments empruntés aux différentes disciplines concernées ?

Comment contribuer à l'émergence de zones d'influences mutuelles, permettant d'aborder ce qui se passe de façon spécifique ? Complexes, mixtes voire hétérogènes, les dispositifs proposés à la monstration ne réclament-ils pas une approche intégrée, un positionnement s'ouvrant sur une appréhension multisensorielle et corporelle ?



Figure 1. Shining field, *Cécile Babiolo* (Rennes 2007)

Il est important de noter que le discours artistique se met en place, dans le domaine qui nous intéresse ici, de manière indirecte et contextuelle. Qu'il a tendance à s'imbriquer dans les interstices de l'espace, rebondir sur un objet, mêler les médias, convoquer une pluralité de savoirs et de pratiques, chez un même artiste, diversement en fonction des propositions ou des projets. « ... Alors que dans *Circulez y'a rien à voir* Cécile Babiolo a cherché à inscrire la notion d'interaction au sein même de l'espace social (lieu de passage ouvert), alors qu'elle a cherché dans *Incessamment* à jouer avec la plasticité de notre environnement (urbain) en manipulant notre perception du temps qui passe, ou dans *Sensors Sonics Sights* (avec Laurent Dailleau et Atau Tanaka) à évacuer de la performance tout recourt à l'improvisation "sans fin", elle a voulu, dans *Shining field*, sortir l'image du cadre de

l'écran pour l'installer en plein dans l'espace de l'installation, matérialisation et détournement du pixel même » [1] . Ce que nous venons d'esquisser au niveau de la démarche de l'artiste ou des artistes, et consistant à déployer un espace-temps capable d'accueillir, comme une niche écologique, la cohabitation de phénomènes indépendants, peut également s'appliquer au processus perceptif mis en place par le public. Jean-Yves Bosseur dans son ouvrage *La musique du XX^e siècle à la croisée des arts* nous dit à ce sujet :

« Par exemple, l'intrusion d'un phénomène imposant un mode de perception unidirectionnel ou dramatique nuirait inévitablement au caractère ouvert et partiellement imprévisible d'un événement basé sur des principes de discontinuité. L'auditeur-spectateur n'a plus à vouloir retrouver à tout prix la connexion éventuellement prévue par les auteurs, mais se crée des réseaux de relation à travers sa propre perception, qui devient, de ce fait, active et créative » [2] .

La notion de co-présence, dans sa dynamique structurelle, peut contribuer à donner la clé des relations qui s'établissent entre des contenus ou des réalités plastiques, sonores et musicales, chorégraphiques, environnementales, sociales, apparemment sans commune mesure, univers en soi chacun dans son domaine, communiquant en surface, selon les canaux d'une porosité (analogique) parfois trompeuse, qui nécessitent une appréhension active pour être perçus au travers des réseaux de relation que leur discontinuité notamment, contribue à nourrir.

2. SIMULTANÉITÉ ET DISPERSION, UN CHAMP RELATIONNEL « CARTOGRAPHIQUE »

Accident ou proposition sonore, le « larsen » réalise la co-présence du couple micro / haut-parleur, se traduisant par la mise en place d'une relation de type feedback à l'origine du son (à partir de rien d'autre, en l'occurrence, que cette proximité, et la capacité des circuits à entrer de manière plus ou moins stable et entretenue en oscillation). Il est remarquable de noter, que le larsen n'est pas particulier au sonore. Il peut apparaître lorsque l'on met en présence une caméra vidéo et une télévision, et s'exprimer alors dans la production d'une masse (monochromatique) plus ou moins instable.

De tels phénomènes, par delà leurs particularités, ont participé au départ, à la formulation d'une hypothèse plus générale consistant à imaginer comment, à l'intérieur de dispositifs de création mêlant des médiums en apparence « non-communicants » (du fait de nos habitudes de perception et de la disjonction physique qui caractérise leur domaine respectif de manifestation), peut se tisser un réseau de relations susceptible de porter, au coeur de son maillage, la continuité d'un point de vue (quelque soit

par ailleurs l'angle d'attaque exercé) ?

Cet article tentera d'étayer le propos par des exemples directement issus des propositions de recherche, de formation et de création que j'ai pu mettre en place ces dix dernières années dans différentes structures¹, et qui m'ont permis d'envisager et tester différents fragments d'outils d'ordre pédagogique, méthodologique, expérimental, nourris, souvent avec passion, des réflexions artistiques des étudiants et des artistes invités.

Une approche élargie de la notion de larsen, nous a été proposée par Tom Mays, avec sa création *Feedback*, créée au Centre Culturel Colombier à Rennes en 2003 (une résidence proposée par le GEPS en partenariat avec l'école des Beaux-Arts de Rennes et France Télécom R&D - Studio Créatif). L'idée artistique globale reprend la fabrication d'un dispositif susceptible de générer des éléments visuels à partir de sons, et inversement, puis de ré-injecter ces productions dans le contexte de départ. Une manière de questionner l'environnement dans lequel on évolue, de questionner également les relations que l'on entretient avec l'espace public, et qui consiste à lier de façon originale et en temps réel les informations visuelles et sonores.



Figure 2. *Feedback*, Tom Mays (Rennes 2003)

¹ Il s'agit en particulier des Ateliers de Recherche et Création (ARC), développés à l'école des Beaux-Arts de Rennes. Il s'agit également des séminaires thématiques approches disciplinaires réalisés à l'Université de Rennes 2 dans le Département Arts Plastiques, Master Arts et Technologies Numériques (ATN), et des résidences de création mise en place dans le cadre du GEPS (Groupe d'Expérimentation Plastique du Sonore). Celui-ci peut rapidement se définir comme la volonté de développer une synergie entre différentes structures, lieux ou institutions, de manière variable et diversement en fonction des besoins et des projets de création accueillis et développés.

Plus précisément, au sujet de l'articulation de sa proposition, Tom Mays nous dit : « Imaginons un espace public dans lequel sont captés des sons et des images. Ces bribes de contenus sont analysées en temps réel, traitées et diffusées *in situ* de manière fragmentaire et décalée. Les sons issus des micros sont décomposés en hauteur, intensité, brillance, quantité de bruit, rythme, densité, et vont servir à générer ou traiter les images. Dans l'autre sens, les vidéos captées par les caméras, analysées à leur tour en valeurs d'intensité, couleurs, zones formelles, vont interagir sur les sons comme pourraient le faire un grand nombre d'interrupteurs, de potentiomètres, de filtres revisitant en continu le sonore. Le son va en quelque sorte peindre par touches l'images qui, en retour, va musicaliser le son, offrant du lieu une représentation sensiblement décalée, une ré-interprétation en temps réel dans laquelle tous les éléments qui composent cet espace-temps anonyme et fonctionnel révéleront leurs relations complexes et insoupçonnées » [3]. Ainsi dans l'installation de Tom Mays, le rapport du son et de l'image n'est pas métaphorique. Il se fonde sur une série de filtres logiciels dont le rôle est de dégager au niveau de la captation (sonore et vidéo), des éléments saillants, pour les ré-injecter d'une manière croisée dans la production *in situ* d'images et de sons, attestant ainsi de la possibilité d'un dialogue audio-visuel au sein de l'environnement immédiat. Par delà les effets produits, et puisque le propos ne semble manifestement pas développer une vision « esthétique » du monde qui nous entoure, c'est bien dans l'égalité, mais en tout point diverse ré-interprétation de l'image et du son, que se trouve le lieu de l'émergence d'une distance critique.

Le travail de création dont nous allons parler maintenant, concerne toujours le rapport image/son abordé sous l'angle d'un principe de feedback beaucoup plus direct celui-ci. Il s'agit de l'installation *100 Hz*, une création collective d'un groupe d'étudiants de quatrième et cinquième année de l'école des Beaux-Arts de Rennes, réalisée dans le cadre d'un Atelier de Recherche et de Création qui a eu lieu au printemps 2004. La proposition s'appuie sur l'utilisation d'un principe élémentaire de manipulation du son, basé sur un échange d'énergie et une gestuelle appliquée à la production d'un larsen lumineux provoqué entre une caméra vidéo et une télévision. La masse lumineuse apparaissant à l'écran peut être déplacée et déformée en intercalant et en bougeant les mains dans le faisceau engendré entre les deux appareils. Le dispositif étant physiquement placé au centre d'un groupe de six enceintes acoustiques disposées de manière circulaire, la caméra fixée au plafond et l'écran de télévision lui faisant face à hauteur de table. Une captation temps réel de la taille et de la déformation de la masse lumineuse contrôle l'intensité et le timbre des sons diffusés, tandis que la captation de l'emplacement de cette même masse à la surface de l'écran, pilote la spatialisation et le mouvement des sons dans l'espace de l'installation.

Ce travail de recherche et de création est l'occasion

pour nous de (re)dire, en marge de nos préoccupations présentes, combien, au fil du temps, les manipulations électroacoustiques du sonore ont permis (d'une manière générale) aux sons de se détacher plus ou moins radicalement de toute référence à la réalité de leur source énergétique, et ainsi accéder à la « plasticité » potentiellement nécessaire à une rencontre transversale et inter-médiatique. Nous pouvons d'ailleurs appliquer cette remarque au dispositif sonore et lumineux rapporté ici, dont la manipulation, au travers de l'inertie (analogique) relative de la masse lumineuse (par rapport au sonore), et les phénomènes d'amortissement qui peuvent accompagner ses déformations, contribue à produire un champ de variation sonore bien particulier, d'autant plus particulier que le geste semble non pas commander le son, comme dans l'utilisation d'un potentiomètre par exemple, mais lui offrir une matérialité inédite.

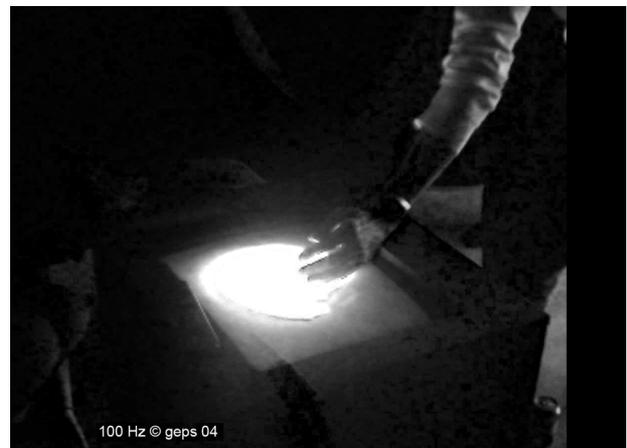


Figure 3. *100 Hz, ERBA (Rennes 2004)*

L'utilisation d'éléments aussi différents que la lumière, l'espace, le corps et le son, à l'intérieur d'un système de relations relativement élémentaires, contribuant à l'émergence d'une proposition « d'art hybride » dans laquelle le son peut être considéré comme élément à part entière d'une vision élargie de l'objet plastique ou de la sculpture produisant et habitant l'espace de l'installation.

Pour clore provisoirement cet exposé sur l'appréhension des phénomènes de feedback et de réinjection, sous l'angle d'une notion de co-présence, je voudrais aborder un dispositif consistant, diversement, à prélever, transformer et réintroduire dans l'environnement sonore des bribes de sons, par un double mouvement temporel de bouclage et de réactualisation du flux sonore ambiant. Comme s'il était question de diriger lentement l'écoute vers ce flux, et donc, en l'occurrence, vers elle-même, dans le sens d'une exploration des sensations auditives, de la perception, de notre attention d'écoute et notre situation dans l'espace.

Nous venons ici de faire une description rapide de l'installation *Do not disturb*, création collective, là encore, d'un groupe d'étudiants de l'école des Beaux-Arts de

Rennes (Atelier de recherche et de Création « Pratiques sonores » 2005), en lien étroit avec un workshop dirigé par Yannick Dauby, autour de la fabrication de petits dispositifs sonores simples et multiples (« patches-modules »), en interaction sonore les uns avec les autres. Yannick Dauby est en effet venu à mon invitation, proposer aux étudiants un atelier basé sur la conception, dans l'environnement Max/MSP, de petits dispositifs autonomes aux comportements simples (« insectes numériques ») reliés chacun à un micro et un haut-parleur, et pensés selon un principe d'auto-propositions sonores, de mutuelles influences via une écoute automatique, une captation et des prélèvements plus ou moins aléatoires, et un dialogue robotique. Réagissant de manière sonore les uns avec les autres et réagissant également au bruit ambiant (condition initiale de l'interaction sonore), ces petits dispositifs ont permis d'engendrer finalement un phénomène de *feedback* complexe. Rétroaction logique, inertielle et statistique des éléments sonores les uns sur les autres, en fonction de l'état informatique des automates (enregistrement, traitement et/ou diffusion), entraînant une transformation lente de l'espace sonore, et une mise en abîme fragmentaire, incorporant dans son processus une importante et fondamentale dimension environnementale. La capacité à créer un espace sonore remarquable et à part entière, constitué de résonances, de dispersions, de colorations et d'atténuations, d'accumulations et de déformations, ouvert dans l'instant (et du fait de son inertie) à toutes les rencontres sonores, conscience « augmentée » de ce qui était déjà là.



Figure 4. Do not disturb, ERBA (Rennes 2005)

A l'occasion d'une résidence à Taipei Artist Village, Taiwan, été 2004, autour de la thématique *écoute, collecte, intervention*, Yannick Dauby définit son projet de création en des termes qui peuvent également expliciter son intervention à l'école des Beaux-Arts de Rennes : « L'essentiel du projet est d'inviter le visiteur à créer un parcours personnel dans les paysages au travers des formes proposées. La remise en situation d'éléments captés n'est pas d'obliger un point de vue subjectif dans une relation d'artiste-émetteur à un public-récepteur, mais de laisser à disposition quelques indices, provoquer quelques

croisements et inciter à l'attention à nos environnements » [4].

Le phénomène du *feedback*, une fois amorcé, peut également s'appréhender comme quelque chose de l'ordre d'un flux. J'ai pour ma part abordé cet aspect des choses en 2006 par le biais d'une expérience très simple dans l'environnement de programmation Max/MSP/Jitter, et consistant à produire un larsen directement à l'intérieur d'un objet matrice. Au départ, la matrice est vide de toutes données numériques. Ses entrée et sortie sont alors reliées entre elles, via un dispositif permettant d'éviter que la boucle ne déclenche trop rapidement un phénomène de saturation préjudiciable à l'expérience. En créant de la perturbation (introduction d'un décalage, d'une « rotation » dans les données contenues dans la matrice), on obtient des formes remarquables, caractéristiques d'un flux (à ne pas confondre avec des « formes fractales »).

Cette expérience montre à quel point une situation de *feedback*, et plus généralement un flux, peut être sensible aux conditions initiales et à la perturbation. Il en va ainsi des flux de données et d'informations que l'on peut déclencher dans les dispositifs multimédias, qui circulent ou s'échangent, et se transforment dans les machines, et qui peuvent finalement être contraints par une configuration, un contexte, un lieu, un espace ou un mode d'interaction. En se proposant d'exercer méthodiquement sa sensibilité à l'encontre de ces phénomènes de flux, on peut très vite revisiter sa manière de percevoir le sonore. Engager une réflexion à l'endroit des réseaux de forces qui l'anime, en y décelant des principes d'auto-régulation, de résonance, de perturbation, de boucles locales, de symétries, de mouvements rétrogrades et plus encore, se situer différemment par rapport globalement à l'environnement, non plus pour s'étonner des sons inouis qui peuvent le parcourir, mais pour véritablement en réaliser l'étendue et le potentiel formel.

En optique, la vitesse de la lumière dans le vide ne dépend pas de sa longueur d'onde, mais dans certains milieux, elle peut en dépendre : c'est le phénomène de la dispersion (à ne pas confondre avec un éclatement suivi d'une dilution, mais à prendre dans le sens de l'apparition de traits distinctifs, source de rapports sensibles et articulés). L'association de deux dioptries plans non parallèles (un prisme), mis en co-présence avec un faisceau lumineux, produit une dispersion colorée de la lumière en ses constituants, champ relationnel et formel potentiel. La co-présence d'une séquence sonore et d'un espace réverbérant peut, dans certaines conditions acoustiques, engendrer une dispersion sonore, avec l'apparition de sons nouveaux, et/ou supplémentaires (spatialisation sonore, déambulation, stratégie d'écoute, centre perceptif). Cette réflexion autour de la notion de simultanéité nous renvoie d'emblée vers des caractérisations « hors temps », sortes de qualités inhérentes à la co-présence des choses (images, sons, objets) susceptible de représenter un répertoire

de formes.

Plastiquement ou musicalement, c'est l'apparition d'une surface, d'un plan, un jeu de contrastes, une variation de milieu, une variation de fréquence, un grain de matière, un accident de forme, de matérialité, une disjonction sonore (...), qui peut, et sans préjuger ni de l'effet, ni de l'ampleur de celui-ci, ni de l'aspect local ou non-local de ses conséquences, ni du médium affecté, ou de la durée et la régularité de la perturbation, contribuer à former un regard, une écoute, une approche. Offrir une continuité à un déplacement, déclencher aussi le processus complexe d'intégration par lequel les impressions visuelles, auditives, tactiles s'associent pour former un point de vue, une présence dans l'espace.

Un phénomène d'incidence perceptive peut sembler relativement évident à ressentir. On peut ainsi éprouver l'impact que peut avoir un son sur les proportions, l'échelle et l'atmosphère d'un lieu ou d'un espace. Mais il est peut-être moins évident d'envisager éprouver avec la même acuité, une qualité perceptive transversale, issue d'un jeu de mutuelles déterminations entre les éléments co-présents. C'est en ce sens que la notion de simultanéité et la réflexion qui lui est associée, peut nous servir, nous permettre, de manière sans doute fragmentaire et souvent éphémère, de dépasser, sans les effacer, la matérialité et le contenu des choses, au profit de leur association. Nous permet de nous affranchir du rapport (même élargi et interchangeable) du fond et de la forme, de considérer différemment la question de la narrativité et celle de la représentation, ainsi que la distance qu'entretient le public avec l'oeuvre. Nous montrant enfin l'existence de champs de tensions, de frictions, de significances ouverts et contextualisés, territoires artistiques habités de propositions et de dispositifs parfois difficiles à conjuguer.

3. LA QUESTION DU TEMPS, COINCIDENCE ET DIFFÉRENCE

La simultanéité recouvre des phénomènes « presque physiques » (qui nous renvoie à des propriétés liées aux éléments dans leur co-présence spatiale et peut-être locale). Mais une création multimédia et trans-disciplinaire intègre une dimension temporelle importante. Il y a la temporalité propre aux éléments articulés entre eux, celle des sons et des images, mais aussi celle des phases de transition, de mutation ou encore des phénomènes d'accumulation ou de dégradations. Il y a par ailleurs la temporalité de l'expérience perceptive proprement dite, liée à la durée de l'appréhension de l'espace et des déplacements effectués, allers et retours et manipulations éventuelles. Une autre dimension temporelle, plus proche en l'occurrence de nos préoccupations, peut s'entrevoir en avançant une notion de coïncidence, laquelle se réfère directement à une idée « d'exception temporelle », troublante, qui nous donne à

voir ou à entendre quelque chose différemment, du seul fait de la présence fortuite (pour l'observateur) d'une autre chose. On est toujours quelque part, dans une situation de co-présence, mais celle-ci peut être qualifiée de spontanée, et peut, contrairement à ce qui a été dit au sujet de la simultanéité, s'exercer de manière locale ou non-locale, dans le cas de pratiques artistiques ouvertes aux télescopes temporels (le fait de ne plus l'entendre ne détruit pas le son. celui-ci est toujours présent et peut à tout moment ressurgir. Un savoir sonore intime, finalement contemporain de nos expériences auditives quotidiennes nous a été sans doute transmis de génération en génération, nous mettant fréquemment en résonance, pour peu que l'on prête l'oreille, avec des sensations que l'on pourrait croire à jamais perdues. Il est important, pour retrouver cette chaîne auditive, de prêter l'oreille, car les éléments indicels qui déclencheront cet écho et ce retentissement se trouvent bien, ne nous trompons pas, ici et maintenant dans notre environnement sonore actuel, pour peu que celui-ci ne soit pas simplement là pour masquer le silence).

Il est possible d'exemplifier ce rapport au temps (instantanéité/rencontre), source de différence, avec *Step to step*¹, une proposition de Thierry Fournier, lequel décrit son installation de la manière suivante : « Dans une salle sont disposés, face à face, un socle blanc et une projection vidéo grandeur nature représentant un coach de fitness qui donne un cours de "step" de vingt minutes, en boucle... . Dès qu'un visiteur pose un pied sur le socle, la vitesse de la vidéo se ralentit proportionnellement à l'emprise de son corps sur le piédestal, jusqu'à se geler presque entièrement. La musique est également ralentie, ainsi que la voix du coach qui demeure toutefois compréhensible, même dans l'extrême lenteur. ... Entre les visiteurs et le coach, à la fois modèle et victime de leurs gestes, s'instaure un bouclage généralisé, double contrainte d'une imitation impossible. L'action se transfère du côté du spectateur, "mise en oeuvre" sur le socle, sculpture vivante en perpétuelle instabilité » [5].

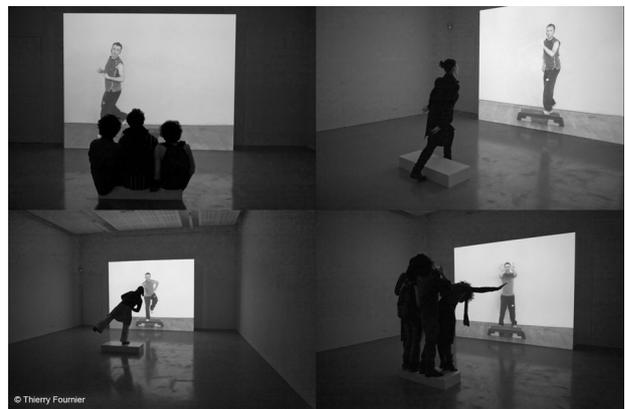


Figure 5. Step to step, Thierry Fournier (ERBA 2008)

¹ *Step to step* est une installation réalisée dans le cadre des résidences de création du GEPS et présentée dans la galerie du Cloître à l'école des Beaux-Arts de Rennes en 2008.

Tout au long de la création de *Step to step*, Thierry Fournier a veillé scrupuleusement à ce qu'une conception ouverte et élargie de l'articulation puisse se mettre en place entre tous les éléments constituant l'installation. J'ai écrit à ce sujet dans la publication consacrée à son travail de création : « Méthodologiquement, Thierry Fournier a abordé son propre projet de création d'une manière véritablement matricielle, nous conviant pratiquement à participer d'une démarche que je qualifierais globalement de "transversale et multiforme". Et ceci dans le sens d'une profonde et solide approche pluridisciplinaire qui nous a finalement plongés, en complète résonance avec l'avancement du projet (et l'apparition des difficultés), au cœur de problématiques aussi diverses et soudaines que la programmation informatique, la diffusion sonore, la diffusion vidéo, la scénographie, la technologie de la captation, la fabrication des différents éléments inhérents à l'installation, et, surtout, moment précieux, l'articulation de tous ces vecteurs expressifs et perceptifs, au sein d'un processus artistique d'une grande rigueur (vision précise de l'interdépendance des éléments constitutifs de l'œuvre, conception dynamique et ouverte de l'articulation globale du dispositif, approche tantôt historique, tantôt analytique des fonctionnalités, des couches structurelles propres à la création et de leurs liens, soucis de la lisibilité et de l'intégration des éléments non essentiels, mise en perspective continue des choix pratiques et de leurs conséquences), afin d'engendrer un dispositif "d'une seule pièce", débarrassé de tout ce qui pourrait créer des tensions superflues voire perturbantes dans le dispositif proposé au public.

In situ, *Step to step* a vu s'exercer une pratique, physique, identifiée comme telle dans la mise en œuvre de gestes se déclinant sur des modes variés et une énergie toujours renouvelée, dépassant et de loin une simple (et commune) phase exploratoire. Installation vidéo et sonore catalysant rencontres, échanges, dialogues, envies de bouger et de participer, collectivement le plus souvent, *Step to step* a été décrite presque unanimement comme une expérience "chorale" dans l'instant vécu, mêlant l'altérité du corps, une communauté de mouvements (plus ou moins imitatifs) et le temps qui passe, alors que seul le geste en qualifie la perception » [6].

Pour des raisons de transversalité et d'échanges interdisciplinaires, ainsi que pour des raisons de pédagogie, de méthodologie de l'expérimentation et finalement de conception technique (et technologique), il nous est apparu utile de mettre en avant et de réfléchir à ces notions de simultanéité et de coïncidence. Elles se sont révélées très vite capables, dans un premier temps, de rendre compte de l'existence de « répertoires » de formes et de relations, non-réductibles à des questions de simple détournement ou de pure combinatoire. Parler de simultanéité source de dispersion (réseaux de relations), nous permet d'entrevoir l'espace des relations artistiques transversales sous le modèle d'une logique modale du « ou », par laquelle la co-présence des éléments fondateurs du dispositif

inter-disciplinaire se comprend en regard d'une subordination tout azimut, multiple, liée aux « conditions initiales » et globalement au point de vue exercé.

Parler dans le même temps de coïncidence source de différence, c'est convoquer une logique sérielle du « et », qui, sans début ni fin, nous permet, quelque soit l'échelle retenue, d'opposer des intérieurs et des extérieurs, d'introduire et de mesurer des présences, de dévoiler des localisations, des étendues géographiques, de coordonner des instants au sein de temporalités utiles à une navigation intégrée au sein de l'œuvre transversale.

Je voudrais, pour montrer un autre aspect de la coïncidence et de son pouvoir de distribution temporel, évoquer rapidement *HK_sound_mapping*, une création réalisée en février 2010, dans laquelle on assiste à un processus d'éclatement spatial d'une séquence vidéo sous l'effet du sonore. Chaque image entre dans un court cycle de répétitions, jeux d'échelle et de dispersion à la surface de l'écran. Les densités d'images, les trajectoires et la

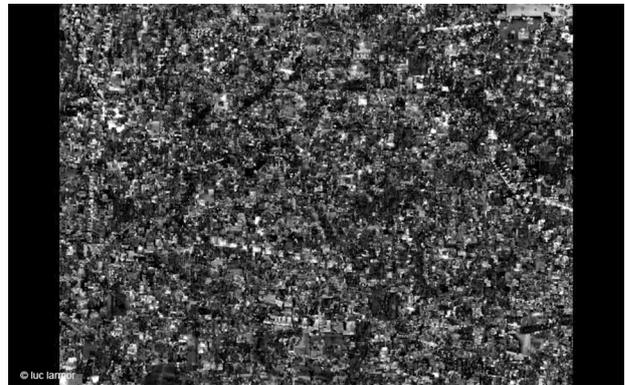


Figure 6. *HK_sound_mapping*, Luc Larmor (2010)

répartition des images à la surface de l'écran, sont elles aussi pilotées, par seuils, par le son. Pour, globalement, une occupation de l'espace qui n'est pas sans rappeler une cartographie urbaine. Au niveau de la structuration de la pièce, nous avons affaire à deux processus (l'un visuel et l'autre sonore), autonomes. Le flux sonore venant, de manière discontinue, perturber et actualiser l'affichage du flux vidéo. Rencontres fortuites de flux, génératrices de différence entre les images, à l'origine aussi de rapports d'échelle en apparence improbables, responsables de la production d'une cartographie sonore qui, dans l'épaisseur de l'accumulation des images et par le jeu de positionnements topographiques « vectoriels », évoque une urbanité en train de se construire [7]. Dans cette création, le dialogue entre les médias n'est pas direct. La relation entre le son et l'image se fait par flux interposés. Une manière de créer un système de relations qui, même s'il finit logiquement (accumulation) par entraîner un phénomène d'inertie, interdit, dans l'expressivité des formes visuelles, et même si l'évocation urbaine est évidente, l'articulation de figures (vites) interchangeables,

sans logique interne propre, entrant avec le sonore dans un rapport de dépendance symbolique, d'objectivation directe, à mi-chemin entre la représentation indicielle (haut/aigu, bas/grave, défilement, fausses couleurs ...), et l'illustration. La mise en rapport « frictionnel » et discontinu de deux flux indépendants, permet de créer des passerelles, des interfaces non tributaires directement du matériau, de son histoire ou de ce qui se raconte au niveau de son contenu. L'éclatement du « vocabulaire » propre à chaque média n'est pas synonyme d'effacement ou de masquage. Bien au contraire, la référence au matériau reste bien présente, sous forme d'ellipses venant localement et obstinément s'opposer au mouvement d'ensemble très certainement quant à lui, générateur de dilution et d'indifférenciation. Une manière décisive de donner à ce qui est en train de se passer une présence, et d'amener la lecture que l'on peut en avoir vers plus de profondeur et de « relativité ».

Pour conclure, du moins le référencement de notre problématique par rapport aux créations et aux oeuvres qui lui ont donnée corps et substance, je voudrais parler de *When durations collide*, un projet de performance que Laurent Dailleau est venu développer et diriger à l'école des Beaux-Arts de Rennes (annexe du Bas Village, un hangar industriel d'environ 350 m²) en 2009 à mon invitation.



Figure 7. *When durations collide*, Laurent Dailleau (annexe du Bas Village - ERBA 2009)

La proposition initiale de Laurent Dailleau portait sur une notion de flux et d'interpénétration, de transitions floues, de mélanges, d'ambiguïtés perceptives et sensorielles. Afin de guider et d'orienter le groupe d'étudiants qui devaient participer à cette performance, un premier workshop a été organisé autour de propositions de masses plus ou moins diffuses, plus ou moins denses et toujours en mouvement (nuages), même imperceptible. Les phénomènes d'écoulement, d'étirement, d'accumulation, d'élasticité, de fluidité, de contact « aqueux », furent longuement abordés, discutés, expérimentés visuellement, auditivement et matériellement. Les choses devaient s'articuler autour de propriétés telles que la résistance, la transformation, l'échange. Et cela dans le but de trouver un grain, un contact, une matière, une chaleur, une

émotion, une idée ou la rémanence d'un souvenir enfoui, quelque soit le médium sollicité ou interagissant. Exercer ses yeux et ses oreilles, ses doigts et tout son corps à la découverte de phénomènes sans début ni fin, en perpétuelle mutation (interrogation), attentif à l'environnement proche ou lointain et soucieux de conserver une certaine idée de la forme, une cohérence, rémanence, « comme un air de ... » (l'expression est de Laurent Dailleau).



Figure 8. *When durations collide*, Laurent Dailleau (annexe du Bas Village - ERBA 2009)

Cette proposition de performance a été en tout point riche d'enseignements au niveau justement de l'articulation d'une création interdisciplinaire¹. *When durations collide* a permis la confrontation, le contact, entre un performeur et les « manettistes » (quatre musiciens équipés de manettes « Nintendo Wii »), donnant aux gestes de ces derniers une ambiguïté très intéressante, superposant à leurs gestes musiciens un questionnement corporel et spatial ouvert. Différents objets vibrants installés, sensibles diversement au sonore, matérialisaient pour leur part, ponctuellement dans l'espace, des vents de vibration en fonction du déplacement des fréquences, répondant également, au détour de la déambulation, aux vibrations qui parcouraient certaines images vidéos, en incitant le public à se déplacer et varier les angles d'écoute. Notons par ailleurs la présence d'une vidéaste dont le rôle semblait être à première vue de filmer l'événement, mais qui participait en fait dynamiquement à la « chorégraphie » générale en matérialisant physiquement des points de vue, en témoignant par ces déplacements incessants de leur diversité. Sa présence peut exprimer à un autre niveau, un rapport particulier au temps, celui de la mémoire de l'instant, d'un instant saisi parmi d'autres. Un rapport particulier voire subjectif à la conservation d'une trace, en regard de situations artistiques éphémères, impossibles à saisir dans leur globalité, que l'on ne peut finalement que vivre, et dont on se souvient, corporellement.

¹ *When durations collide* fût l'occasion d'une collaboration technique (programmation informatique) avec un groupe d'étudiants de l'IFSIC (Institut de Formation Supérieure en Informatique et Communication) de l'Université de Rennes 1.

Ce détour par les aspects formels et relationnels, intentionnels et imbriqués des pratiques artistiques transversales, et la difficulté d'en rendre compte dans un document d'archive, pose à l'évidence la question de la trace sonore et de sa lecture. Une forme sonore qui prend son « sens » au moment où se produit un son, au moment où s'isole du continuum sonore un élément qui va entrer en relation avec « quelque chose » de son environnement, lorsque mon attention, au détour d'une image, d'un point de vue, sera attirée par ou retiendra une séquence, lorsque celle-ci, pour une raison ou pour une autre deviendra à son tour remarquable, structurante et parviendra à modifier mon rapport à l'espace, au mouvement d'ensemble ou à un objet particulier (dans sa globalité ou pour un détail jusque là ignoré). Des conditions d'apparition et de présence des sons qui limitent grandement la portée d'une description, toujours arbitraire. Attestant en cela de la nécessité d'en parler autrement. Soulignant les contraintes et les limites des « typo-morphologies » partant d'un postulat de différenciation et d'imbrication posé une fois pour toutes, comme si les choses n'avaient qu'une seule manière d'être. D'où la nécessité, à notre avis, d'aborder les questions de contiguïté et de juxtaposition, à la lumière de problématiques faisant appel, par exemple, aux notions de co-présence, de simultanéité et de coïncidence, à l'image de celles que nous venons d'avancer dans cet article.



Figure 9. When durations collide, *Laurent Dailleau* (annexe du *Bas Village* - ERBA 2009)

En situation de jeu (performance), il est possible d'ajuster tel ou tel paramètre et influencer plus ou moins radicalement ou profondément le déploiement des processus. Un contrôle qui s'exerce à un niveau d'ensemble, celui des modifications de direction, de densité, de registre, de variabilité ou de dégradation. Un contrôle qui peut conduire à un changement de contexte, d'environnement, de « niche écologique », sans que rien, sinon la volonté de franchir le sérail n'altère dans l'instant la prolifération des formes sensibles. Le sonore, engendré par l'utilisation de flux numériques, est envisagé en compréhension. Il est pensé par rapport à une loi d'ensemble qui ne préjuge en rien de l'extension des éléments effectivement produits et articulés. C'est une qualité que l'on prête aux sons, un comportement. Une

manière de circuler, de s'ouvrir, de se transformer, de se fragmenter, de se fondre, de se disperser, d'inventer (composer spontanément avec le contexte). Des formes sonores qui sont virtuellement maîtrisables, mais qui n'ont pour seul substrat nerveux et réflexif, pour seule certitude au moment de leur actualisation, que la sensation d'être en devenir, purs questionnements.

Comme si la pièce musicale jouée en temps réel, en relation avec un dispositif visuel ou scénographique, déroulement temporel non linéaire, incompressible, théâtre de rencontres, d'interactions, développements, transformations, et disparitions de bribes d'univers sonores, n'existait véritablement que dans la mémoire de chaque auditeur, visiteur, interacteur, affirmant alors comme une expérience personnelle ou une histoire individuelle, directions, affinités, drames et significations...

Expérience individuelle ne veut pour autant pas dire perception forcément isolée dans une altérité profonde, incommunicable. A travers l'exercice d'écoute envisagé ici, il nous faut « naturellement » chercher à sortir de soi-même, et à s'ouvrir à la diversité du sonore et de ses occurrences (d'où l'importance du positionnement spatial du son dans le dispositif de diffusion). Le partage, l'expérience collective, l'échange, se situent sans doute là, dans le croisement des trajectoires de perception. Nous devons accéder à une qualité d'écoute qui en passe certainement par une salutaire vacuité de l'esprit, un loisir. Ne pas être réticent aux formes de dispersion, de délocalisation perceptive. S'ouvrir à des temporalités parallèles, à des mises en perceptives entre des événements actuels et passés, réels, présents ou imaginaires.

Nous évoluons au sein de décalages perpétuels sans même y songer. Nous accueillons le temps dans la succession des instants, c'est-à-dire, régulièrement, en dehors de toute chronologie. Des instants, tour à tour inconséquents, ou lourds de rencontres infiniment présentes, un voyage au coeur du mécanisme de la mémoire, des résonances perceptives, et d'une écoute spatiale qui a besoin d'un corps pour exister, pour se rassembler et tisser des liens et du réel. Cette manière d'être du sonore (en temps réel), place la question de son éventuelle signification non pas dans quelque chose qui se raconte, mais dans notre propre rapport au monde. Jeux très « plastiques » et discrets (au sens des unités discrètes de la linguistique), entre des matières éclatées, déformées, recomposées dans l'instant, hybridées, formant une musique qui n'a intrinsèquement pas de sens, ne délivre aucun message, et se situe de prime abord au coeur du sonore, dans sa spécificité phénoménologique. Et dans la mesure ou le processus trans-disciplinaire s'attache à des comportements, manières d'être au monde, co-présence, « mode d'être » dirait Antoine Schmitt, se pose également, avec une acuité particulière, la question de la « volonté » qui anime ainsi le sonore, le pousse à dialoguer avec son environnement, et pour le moins, nous invite à méditer

sur sa vitale obstination.

4. CONCLUSION

Les notions avancées ici et la problématique qui leurs est associée, peuvent, en situation de recherche et de « confrontation disciplinaire », servir à mettre en place des protocoles d'échanges et de communication favorisant l'appréhension et la compréhension des phénomènes de co-présence artistique. Notre préoccupation est d'entrevoir des zones d'échange à l'intersection des domaines de référence (son, image, objet, espace, numérique ...). Une manière (et sans chercher à établir de comparaisons), de faire également apparaître des distinctions importantes entre disciplines, et de tisser des liens nouveaux ou dynamiques entre elles, en tenant compte des différences qui existent, dans l'approche de la spatialité et celle du temps notamment.

Notre recherche nous a permis en particulier de constater que, malgré (et peut-être en raison) de la transversalité des moyens, les pratiques tendent souvent à se singulariser, en s'attachant de plus en plus finement aux « aspérités médiatiques » susceptibles d'établir des connexions significatives, par exemple. Des tentatives visant à dépasser les particularités des matériaux (considérés dans leur discontinuité) pour autoriser l'émergence de liens et de lieux relationnels qui ne soient plus des lieux de transition mais des lieux d'échange stables et largement partagés.

Les notions mises en avant, peuvent encore contribuer, d'une certaine manière, à « dématérialiser » l'approche que l'on peut avoir de certaines formes artistiques actuelles. Et ceci en contribuant à poser différemment la question de la relation entre l'art et le réel, entre l'art et le domaine social (réseaux sociaux, pratiques nomades, contextuelles, utilisation et détournement d'outils non-dédiés ...). Une manière aussi d'éclairer la démarche consistant parfois à varier les formes d'apparition et de réalisation des propositions artistiques, plastiques et/ou sonores, en fonction des lieux, des occasions (et des budgets). Des projets artistiques et des créations qui peuvent désormais changer rapidement de statut, passer de celui de performance, à celui d'installation, ou de traces vidéo, sonore, on-line (versions capables d'être échangées, connectées, transférées, re-mixées, hybridées). Outils partiels mais utiles également pour parler des oeuvres avec d'autres mots que ceux liés à la représentation (ou à l'abstraction). Pour rendre compte de projets à plusieurs dimensions, intégrant concept et démarche, attitude et positionnement, vécu et éléments environnementaux, co-présence. La possibilité pour nous offerte de tenir compte dans nos propos de la relation étroite qui relie proposition artistique, modélisation (simulation informatique) et processus d'élaboration (lequel, parfois, se donne à voir en lui-même comme œuvre à part entière).

D'une manière plus générale, les conditions de la « monstration » des œuvres plastiques et picturales ont elles aussi, dans leur spécificité, grandement changées. Les artistes sont à l'évidence désormais sensibles à l'environnement dans lequel leurs créations se donnent à voir. Préoccupés par le temps (qui passe), la dégradation (programmée) des œuvres ou du contexte de monstration, ils interrogent aujourd'hui le support même de l'accrochage des œuvres, la fabrication ou l'investissement de l'environnement immédiat de la galerie, sans parler des modalités de la communication autour de leur travail de création et de l'utilisation contextuelle du son. Toutes propositions utiles aujourd'hui pour provoquer une multiplicité de contacts avec la création et les œuvres, qui ne sont semble-t-il pas étrangères à notre champ d'étude.

5. RÉFÉRENCES

- [1] Larmor, L., « Cécile Babiole artiste sonore et visuel », in Cécile Babiole *Shining field*, responsable de la publication Jacques Sauvageot, conception Jérôme Saint-Loubert Bié, Ecole des Beaux-Arts de Rennes, France, février 2008, ISBN : 2-908373-65-3.
- [2] Bosseur, J.-Y., *La musique du XX^e siècle à la croisée des arts*. Editions Minerve, Paris, 2008.
- [3] Mays, T. *Feedback*, document GEPS, Rennes, France, septembre 2003.
- [4] Dauby, Y. et Tsai W.-S., *Shejingjen*, document en ligne, http://www.kalene.net/shejingren/shejingren_presentation.pdf, novembre 2004.
- [5] Fournier, T., « Step to step », in Thierry Fournier *Step to step*, responsable de la publication Jacques Sauvageot, conception Jérôme Saint-Loubier Bié, Ecole des Beaux-Arts de Rennes, France, juin 2009, ISBN : 2-908373-77-7.
- [6] Larmor, L., « L'intelligence du geste », in Thierry Fournier *Step to step*, responsable de la publication Jacques Sauvageot, conception Jérôme Saint-Loubier Bié, Ecole des Beaux-Arts de Rennes, France, juin 2009, ISBN : 2-908373-77-7.
- [7] Larmor, L., *HK_sound_mapping*, création sonore et visuelle, France, février 2010, extrait visible à l'adresse : http://www.dailymotion.com/video/xcdg9l_hk-sound-mapping_creation

Session 6

Écoute et pédagogie

MUSIQUE ET MUSICOLOGIE À L'UNIVERSITÉ DE SAINT-ETIENNE RECHERCHES EN INFORMATIQUE MUSICALE

Laurent Pottier
Université Jean Monnet
Laurent.Pottier@univ-st-etienne.fr

RÉSUMÉ

Présentation des formations à l'université Jean Monnet de Saint-Etienne en musique et musicologie offrant un parcours complet, orienté à la fois recherche et production, en informatique musicale.

1. INTRODUCTION

Le département de musique est relativement récent (1992) et appartient à une université pluri-disciplinaire qui a fêté ses 40 ans en 2009. L'équipe enseignante comporte deux professeurs : Alban Ramaut et Béatrice Ramaut-Chevassus ; quatre maîtres de conférences : Nathalie Fernando, (en détachement au Canada), Marc Desmet, Jean-Christophe Branger et Laurent Pottier ; un PRAG : Julien Garde et plusieurs chargés de cours.

Une cinquantaine d'étudiants entrent chaque année en licence, pour un effectif total en licence d'une centaine d'étudiants. Nous avons actuellement une cinquantaine d'étudiants en master et quinze étudiants inscrits en thèse (le master pour sa partie recherche est co-habilité avec l'ENS de Lyon).

Le programme des cours en licence est la fois musical (technique et pratique) et musicologique. Si toutes les périodes de la musique occidentale savante sont étudiées, trois disciplines sont également proposées sur les trois ans de licence, pouvant déboucher sur des masters recherche et des thèses de doctorat : l'ethnomusicologie, le jazz et les musiques électroniques.

2. LA FORMATION EN INFORMATIQUE MUSICALE

Le programme de formation en informatique musicale porte sur les outils de l'informatique musicale, la musique électronique et l'ingénierie du son. Il s'étend sur cinq ans.

2.1. Le programme de licence

Il comporte six semestres. Les enseignements des semestres S1, S4, S5, S6 et l'option *Musique et Image* ont lieu en salle d'informatique musicale sur machines et les élèves sont évalués principalement sur des réalisations pratiques.

- S1 : le MIDI

Ce cours présente dans le détail la norme MIDI et permet une initiation aux programmes Finale, Cubase et Max/MSP. Les étudiants apprennent ainsi l'édition de partitions avec Finale, ils réalisent des séquences MIDI avec Cubase et ils abordent la construction de dispositifs MIDI interactifs avec Max/MSP.

- S2 : histoire des musiques électroniques

Dans ce cours (théorique) sont présentées l'histoire des instruments électroniques et les principales écoles de la musique électroacoustique : musique concrète, musique électronique, musique par ordinateur. Les pionniers et le répertoire de la musique électroacoustique sont étudiés et nous effectuons un panorama des principaux studios de création en France et à l'étranger.

- S3 : acoustique

Ce cours (théorique), axé sur l'acoustique musicale, présente la théorie du son et les caractéristiques du son numérique. Les différentes techniques d'analyse et de représentation du son sont abordées, pour déboucher sur les principales techniques de synthèse et de traitement du son et leurs applications. Les principales théories de l'acoustique architecturale et de l'acoustique des instruments sont également étudiées. Enfin, un panorama des outils de l'informatique musicale est également proposé.

- S4 : l'audio numérique

Ce cours est un atelier de composition électroacoustique. Les étudiants y apprennent les techniques de montage, l'utilisation des outils de traitements audio numériques et le mixage. Une formation à la synthèse du son est proposée avec des outils réalisés avec Max/MSP. Les pièces réalisées au cours de cet enseignement sont présentées en concert en fin d'année sur un acousmonium lors d'un stage de diffusion multipoints.

- S5 : les outils d'aide à la composition (Open-Music) et leurs applications

Ce cours présente les principales techniques de composition utilisées par les compositeurs contemporains (modèles sériels, modèles spectraux, modèles stochastiques, graphisme et interpolation, contrôle de la synthèse, applications à l'analyse d'œuvres) à travers l'étude de l'environnement d'aide à la composition Open-Music de l'Ircam. Les étudiants apprennent à utiliser cet outil et doivent produire des partitions utilisant ces technologies. Ces pièces peuvent ensuite être présentées en concert lors des festivals

organisés chaque semestre par l'association des étudiants en musicologie.

- S6 : le temps réel

Ce cours est un atelier de composition basé autour du logiciel Max/MSP et de la synthèse et du traitement du son en temps réel. Pendant ce semestre, le dernier de la licence, les étudiants doivent concevoir une pièce mixte, pour instruments acoustiques traités par ordinateur, en écrivant la partition et en réalisant la partie électronique dans Max/MSP. Ces pièces sont ensuite jouées en concert lors du festival de fin d'année. D'autres concerts sont également organisés pour présenter ces travaux, notamment avec la classe de musique électroacoustique du conservatoire à rayonnement régional de Saint-Etienne.

Les enseignements optionnels

- Musique et image

Cet enseignement optionnel (en semestre 6 de licence) est destiné à introduire une réflexion sur l'introduction de l'image vidéo dans la production musicale contemporaine. De plus en plus de compositeurs, en particulier ceux qui utilisent l'électronique dans leurs créations (musiques électroacoustiques, musiques mixtes...), font appel à des projections d'images dans leurs spectacles ou créent des dispositifs multimédias. Par ailleurs, chez de nombreux artistes, plasticiens, vidéastes, le rôle de la musique devient de plus en plus important et partie prenante de l'acte créatif. Nous étudions les relations « Musique / Images » chez plusieurs compositeurs et artistes (Dominique Barbier, Jean-Baptiste Barrière, Thierry Demey, Catherine Ikam, Serge de Laubier, Tristan Murail, Patrick Portella).

Par ailleurs, ces questionnements sont mis en application à travers des travaux réalisés par les étudiants pendant cette formation, en utilisant les outils qui permettent d'associer les deux médias (Cubase, Max/Jitter, Flash).

- Esthétique du cinéma : Musique et cinéma.

Cette option est une option transversale ouverte à tous les étudiants de la faculté. Elle permet dans un premier temps d'étudier les fonctions de la musique de film, puis de présenter les différents styles de musiques utilisés dans l'histoire du cinéma. Enfin, le cas des clips et des films musicaux est également abordé.

2.2. Les séminaires de Master

- M1 : l'analyse et la synthèse du son (Open-Music et Csound)

Le programme Csound¹ est un de programmes de synthèse les plus généraux qui soient et également un des plus puissants puisqu'il n'a pas les limitations des programmes temps réel. C'est aussi un des programmes de synthèse du son les plus utilisés². Combiné à un outil

¹ Barry Vercoe – MIT - 1985.

² Richard Boulanger, *The Csound Book*, avec 2 CD-Rom, Cambridge: Massachussets, 2000, 740 p.

comme Open-Music, les possibilités d'exploration du son sont immenses. Dans ce cours les étudiants apprennent à combiner analyse du son (Audiosculpt, Diphone)³, traitement et transformation des données, composition algorithmique (Open-Music) et synthèse du son.

- M2 : (a) les dispositifs gestuels pour le temps réel

Dans ce séminaire de Master, nous présentons aux étudiants divers capteurs gestuels que l'on peut utiliser pour contrôler des dispositifs électroniques temps réel (claviers, pédales, capteurs sur dispositif Eobody, tablettes graphiques, système Lemur de Jazz Mutant, caméras vidéo avec softVNS⁴, ...) en décrivant des exemples de réalisations utilisant ces dispositifs.

- M2 : (b) analyse et transcriptions des musiques électroniques

Ce séminaire forme à l'utilisation de l'Acousmographe⁵, outil qui permet de réaliser des représentations graphiques, des descriptions des objets sonores pour l'analyse musicologique de pièces électroacoustiques. Ce séminaire s'appuie notamment sur les analyses disponibles sur le site du GRM dans la collection des portraits polychromes⁶.

Avec le développement de l'informatique musicale, d'autres options vont être mises en place en partenariat avec la faculté des sciences (traitement du signal, acoustique et langages informatiques).

- M2Pro « administration et gestion de la musique » : communication et multimédia.

Ce cours (pratique sur ordinateur) enseigne les outils pour l'image et le son utilisés dans le domaine de la communication. L'accent est mis sur la réalisation de sites Internet interactifs en utilisant les logiciels Dreamweaver et Flash.

3. LES LOCAUX

Une salle de cours est équipée de seize ordinateurs PC avec Max/MSP, Cubase, Finale, Open-Music, l'Acousmographe, Csound, des claviers Midi et des cartes audio.

³ Audiosculpt et Diphone sont des programmes Ircam diffusés par le forum (<http://forumnet.ircam.fr/>).

⁴ Créé par David Rokeby.

⁵ L'acousmographe est un logiciel du GRM (<http://www.ina-entreprise.com/entreprise/activites/recherches-musicales/acousmographe.html>).

⁶ Portraits polychromes du GRM : <http://www.ina-entreprise.com/entreprise/activites/recherches-musicales/portraits-polychromes.html>.



Figure 1. Salle de cours pour l'informatique musicale.

Une salle de répétition est équipée d'un dispositif de sonorisation APG, d'un piano à queue, d'une batterie et de divers amplificateurs.



Figure 2. Salle de répétition.



Figure 3. Studio recherche.

Un studio a été construit en 2009 pour la recherche. Il est équipé d'un Protools HD et dispose d'une quadriphonie Focale (extensible en hexaphonie).

Une salle de spectacle de 180 places, une salle de répétition et un studio d'enregistrement pour la pédagogie sont actuellement en construction et seront utilisables dès la rentrée de septembre 2010.

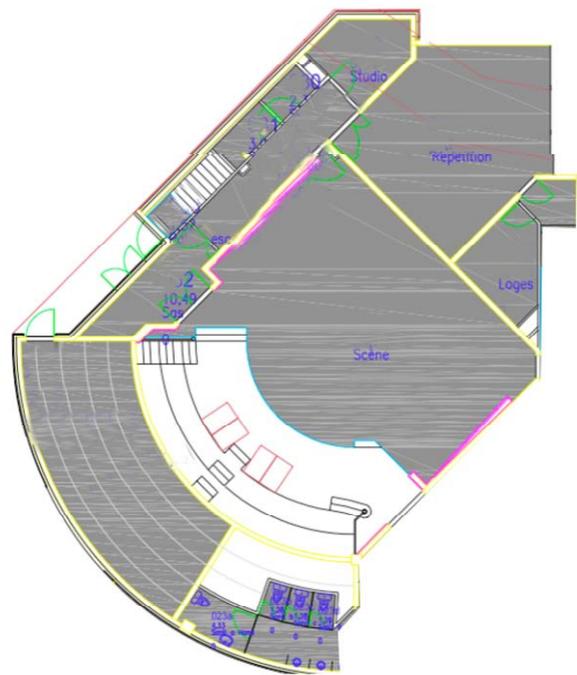


Figure 4. Plan des futurs bâtiments abritant une salle de spectacle, une salle de répétition et un studio.

3.1. Les productions

Les étudiants en musicologie participent à de nombreuses réalisations musicales (chorale, musique d'ensemble, stages, rencontres) qui débouchent sur diverses productions (2 festivals par an et une saison de concerts).



Figure 5. Concert de musique électronique temps réel au Fil (mai 2008), présentant les résultats des cours de musique d'ensemble. Ici, un groupe ayant travaillé sur l'improvisation et la composition collective sur instruments électroniques ou acoustiques traités en temps réels.



Figure 6. Concert de musique électronique temps-réel au FIL présentant les compositions de musique mixte interactive réalisées par les étudiants de 3^e année de licence (mai 2009).



Figure 7. Concert de musique électronique temps-réel à Saint-Victor/Loire présentant les compositions de musique mixte interactive réalisées par les étudiants de 3^e année de licence (juin 2009).



Figure 8. Spatialisation au Fil (mai 2009) sur l'acousmonium de J.-F. Minjard des pièces acousmatiques des étudiants en 2^e année de licence.



Figure 9. Stage de musique improvisée au CRR (mars 2009) dirigé par le *Trio Bic*.

Ces productions utilisent les diverses ressources en lieux de spectacles de la ville de Saint-Etienne, du département de la Loire et de la région Rhône-Alpes.

4. L'EQUIPE POUR L'INFORMATIQUE MUSICALE

L'équipe enseignante dans le domaine de l'informatique musicale et des musiques électroacoustiques est formée de Béatrice Ramaut-Chevassus, professeur en musicologie, Laurent Pottier, maître de conférences en musicologie, Max Bruckert, réalisateur en informatique musicale (permanent à Gramme), Vincent Rollin, allocataire moniteur en thèse, Lorène Giron et Romain Michon, étudiants en Master. L'an prochain, l'équipe sera augmentée d'un ingénieur de recherche embauché sur le projet ANR ASTREE.



Figure 10. B. Ramaut-Chevassus, L. Pottier, M. Bruckert, V. Rollin, R. Michon et L. Giron.

5. LA RECHERCHE

La formation en informatique musicale est fortement adossée aux activités de recherche qui ont lieu au Cierec (Centre Interdisciplinaire d'Etude et de Recherche sur l'Expression Contemporaine - EA 3068).

L'équipe transversale E-Formes, regroupant les chercheurs en littérature, arts plastique et musique s'intéressant aux arts numériques permet de confronter les pratiques de ces disciplines complémentaires.

Les étudiants en master et en thèse peuvent réaliser des recherches dans des domaines variés, grâce à une équipe enseignante diversifiée. Dans le domaine de l'informatique musicale, l'accent est mis sur le domaine de la synthèse du son, et son contrôle interactif en concert. Les outils pour l'aide à l'écriture musicale sont également au centre de nos préoccupations pour le contrôle des processus de synthèse du son.



Figure 11. Concert de musique électronique à la Cité du Design, Remi Dury et Karlax (nov. 2009).

Actuellement, nous travaillons sur un projet de recherche financé par l'ANR (ASTREE) dans lequel sont impliqués le CiREC, l'Ircam, Grame et Armines. Ce projet s'intéresse à la préservation des œuvres électroniques utilisant des dispositifs temps réel. Nous cherchons notamment à proposer des versions des programmes de synthèse et de traitements temps réel écrits dans un langage de description générique : le langage Faust développé à GRAME. L'ambition du projet ASTREE est donc de s'appuyer sur la représentation en langage FAUST afin de permettre la transformation de l'existant, quel que soit le langage ou le système originel (Max/MSP, PureData, ...) en une forme concise, expressive, adaptée et d'effectuer des travaux de classification de l'existant, afin de construire les bases d'une organologie des processus temps réel.

6. BIBLIOGRAPHIE

Quelques publications récentes des Presses Universitaires de Saint-Etienne :

Monique Maza et Alessandra Saemmer (dir.), *E-Formes 2, Arts & écritures numériques, au risque du jeu*, Saint-Etienne : PUSE, 2010.

Laurent Pottier (dir.), *Le calcul de la musique - Composition, Modèles & Outils*, Saint-Etienne : PUSE, 2009, 477 p.

ACOUSMATIC REASONING : AN ORGANISED LISTENING WITH IMAGINATION

Suk-Jun Kim

Helmstedter Strasse 27, 10717 Berlin, Germany

E-mail : sukjunkim@reddoorsound.com

ABSTRACT

This paper proposes acousmatic reasoning as a way of organised listening to electroacoustic music. It also examines the inference processes of acousmatic reasoning—overcoded, undercoded, and creative abductions—as well as syntactically relevant properties (SRPs), properties of sound-images identified by listeners and closely related to the processes of acousmatic reasoning. The paper then sets out an analytical listening exercise using *The Flywheel Dream* (1994) by Paul Koonce to show how acousmatic reasoning works in electroacoustic music listening.

1. INTRODUCTION TO ACOUSMATIC REASONING

Considering that examining the possibility of making-sense in electroacoustic music listening means observing how listeners use inference, a process of reasoning based on evidence acquired not only by listening, but also by imagining, it is believed that electroacoustic music listening involves a two-sided process: on the one hand, it affords listeners unlimited freedom in exercising their imagination, a topic I have examined elsewhere [13]; on the other hand, it involves their processes of reason, which delimit, realize or materialize their imagining.

Listeners imagine sound-images while listening to electroacoustic music as it actively employs, as its key feature, everyday sounds whose dis-embodied and displaced presentation inspires listeners to engage their memories of sound and sense to imagine that which is missing. As Casey [3] might say (and I paraphrase), listeners imagine more than is heard.

Yet, there is more to the task of listening, for participating in the play of associations is only part of the game; the greater challenge is to follow the way sounds and sound-images (gradually or suddenly) transform and navigate in a theater of semiotic phenomena where the reading of sounds depends upon 'inferential processes' [6]. Furthermore, the drama of changing sound-images is not the listener's only concern; with less and less quotidian connotations cueing listeners to the nature of scene and sound, listeners soon recognize that they are no longer simply playing the game of signifiers. Consequently, electroacoustic music offers listeners the opportunity to listen in two different ways, one that focuses on issues

of representation and mimesis and another that focuses on the sounds alone and how they reference their own developing aural orders. [6]

However, the duality only explains part of the complexity of electroacoustic music listening, for with the duality comes an underlying process, emerging out of the ability and interest of listeners concerned with how the two modes interact. The idea that listeners can balance 'concrete and abstract attitudes' [18] and travel back and forth between semiotic and spectromorphological listening depends, first and foremost, upon their ability to negate semiotic references while listening. While this may be a challenging skill to acquire (as conceptual as it is perceptual), once mastered, a field of play between the semiotic and the spectromorphological can open up, offering listeners the opportunity to engage in what I call *acousmatic reasoning*.

1.1. Definition and issues of acousmatic reasoning

I define acousmatic reasoning as *a process of listening to and composing with dis-embodied and displaced sound materials based on their spectromorphological and semiotic significations and connotations*. So defined, acousmatic reasoning involves three primary subjects: sounds dis-embodied and dis-placed by technological manipulation, semiotic and spectromorphological listening as opposing modes of listener engagement, and finally, the listener's inference processes.

1.1.1. Dis-embodied and Dis-placed Sounds

Audio recording and technologically manipulated sounds foster the perceptual dis-embodiment and displacement of sounds, which can inspire listeners to imaginatively construct sound-images that ground and complete their perceptual experience. [13] Yet, however natural, the imaginative construction of sound-images may not be a practical listening goal. For much like other products of the imagination, sound-images are as easily destroyed as they are constructed, especially in pieces in which the mix and shift of aural cues challenges listeners to hear sound-images as anything more than ephemera. The question then is: how do listeners successfully turn sound into sound-images? And moreover, why do they base their images on some sounds rather than others? I argue that, as primary

processes, semiotic and spectromorphological listening provide insight into the way listeners choose and turn sounds into images.

1.1.2. *Semiotic listening: listening for sound*

One form of listening used in acousmatic reasoning is semiotic listening. Semiotic listening is *a way of listening to electroacoustic music in which listeners entertain sounds and their potential sound-images based on semiotic significations*.¹

In semiotic listening, listeners focus not only on moment-to-moment sound-images, but also on how they change, sequence, or collect together to create meaning. Depending on what they find, listeners may then imagine a *sound-image framework* capable of encompassing and explaining the function of one or more sounds, which become, through their reinforcement of it, *components* of the framework. For example, a listener, upon hearing the sound of typing and telephone calls, would conceivably progress from an initial identification of the sound with the human action of typing to its possible use as a *sound-image component* in the *sound-image framework* for office scenes.

In the formation of plausible sound-images, source-identification and context-identification are co-dependent processes, and that, being co-dependent, the source and context identification processes continually impact each other. If listeners hear, for example, a traffic sound added, at low volume, to the office example given above, they either ignore the sound, or most likely, recognize it as outside traffic, heard through an open office window. Having identified the traffic as such, listeners maintain the sound-image of the office by simply modifying their listening framework to now include an open window near a busy street. But if the traffic sounds become louder and louder, gaining foreground prominence, the framework is challenged and even changed as the listener is seemingly taken through the open window and into the street, although when the shift from inside to outside occurs exactly is hardly clear, since the spaces we hear are never so

¹ Semiotics as used in my definition of *Semiotic Listening* should be distinguished from semiotics as used by others. For example, discussing mainly music in the Classical style in his book *Playing with Signs*, Agawu distinguishes between two forms of semiosis: the semiosis identified through the internal structuring of the music and the semiosis identified by the extramusical 'topics' the music draws upon. By borrowing the concepts of Roman Jakobson, he terms the former as 'introversive semiosis' and the latter as 'extroversive semiosis.' By distinguishing the two different types of semiosis, he proposes to investigate the 'interaction between topical signs and structural signs, a notion that might be described in terms of play' [1]. As such, Agawu's use of semiotics is concerned with the identification of both intrinsic and extrinsic sign systems, whereas my use of semiotics is concerned only with matters of source and context identification, and their signifying processes. Interestingly, however, just as Agawu is interested in the play between topical signs (extroversive semiosis) and structural signs (introversive semiosis), this dissertation is also interested in the interplay between sounds and what they represent or signify; my definition of semiotic listening is a step toward recognizing and investigating this interplay.

clearly demarcated as the ones we see. In fact, the growing traffic volume could suggest that we are merely walking toward the open window, or more fantastically, that the window is coming closer to us or getting bigger and bigger just before moving us through it into the out-of-doors.² As the traffic sound gains more of the listener's attention, listeners fail to assimilate the sound into the image of an office; the added sound can destroy the office framework altogether.

Newly introduced sounds can, no doubt, change a listening framework. However, the identity of an established framework can be challenged not only through new sounds, but also through changes to component sounds already supporting the framework. If, for example, a rhythmic feature were introduced into the typewriter sound of the office scene and became increasingly regular, so much so that it began to sound as if someone were performing music with the typewriter, then the listener would be forced to question the plausibility of the sound as having come from a typewriter, since, conceivably, no one types that way (Although, in theory, one could 'play' a typewriter; however, typists rarely show such ability or concerns short of boasts of speed; more, musical performability is not a common property of typewriters.). At best, it is imaginable that, because of the central role of typewriters in offices in listeners' imagination, listeners might attempt to keep the typewriter in its place by changing it into, perhaps, a musical typewriter; however, doing so would no doubt render the office scene more musical and comical than believable.

The emergence of the office framework in the example above is as much a consequence of the assimilation and inclusion of available sound-image components as it is the downplaying and exclusion of other sounds deemed irrelevant or outside the domain of the office framework. Yet, the task of creating and maintaining plausible frameworks is more daunting (if not entertaining), for added to the difficulties of navigating the inclusion and exclusion of sounds to a framework is the problem of dealing with components that betray their original identities or contextual frames; transformations of sound composed to betray a listener's initial reading are common to the field's more engaging electroacoustic compositions, especially since they actively engage listeners with the veracity of the frameworks they imagine.

1.1.3. *Spectromorphological listening: listening to sound*

Given the above, how then do sounds come to betray the readings made of them? And what circumstances of sound, development, and transformation lead listeners to incorrectly identify sounds? As discussed, changes to

² One can argue that listeners may not necessarily attempt to combine these two sound-images as one; rather, they may entertain the two as they are. This argument may be true. Nevertheless, it should be noted that the previous sound-image of 'a busy office' has effectively been changed due to the introduction of the traffic sound.

critical sound features (such as the volume of traffic heard in an office scene) can, at their extreme, render a listening framework implausible. As aptly supported by Smalley [18], the decision to reevaluate an established sound-image depends upon engagement with the sound's spectromorphological properties. In fact, the source-identification process relies mostly on spectromorphological properties, which is why a sound-image perceived as no longer tenable inevitably leads a listener back to not just the source-identification process, but to a deeper examination of a sound's spectromorphology. [17]

There are two issues in spectromorphology that need to be discussed in order to relate spectromorphological listening to acousmatic reasoning. First, spectromorphological listening takes reduced listening as an investigative strategy, which means that spectromorphological listening is subject to the same dialectical conditions as found in reduced listening. Listeners engaged with reduced listening and its removal or denial of the representational and semiotic connotations of the sounds they hear face an equivalent sensitization to the spectromorphological properties of sound that allow sounds to signify; in short, focus on sound can actually *encourage a new spectromorphological awareness of how sounds represent*. The second issue is closely related to the first: despite its goal of investigating the abstract dimension of sound, spectromorphological listening reveals, and is affected by, mimetic structures and the sound-image traces they carry and preserve. It is particularly interesting how the aural abstractions of gesture and texture can be identified, however remotely, with sound-images, specifically those of body and place. [7][9][10][11][13][16][18]

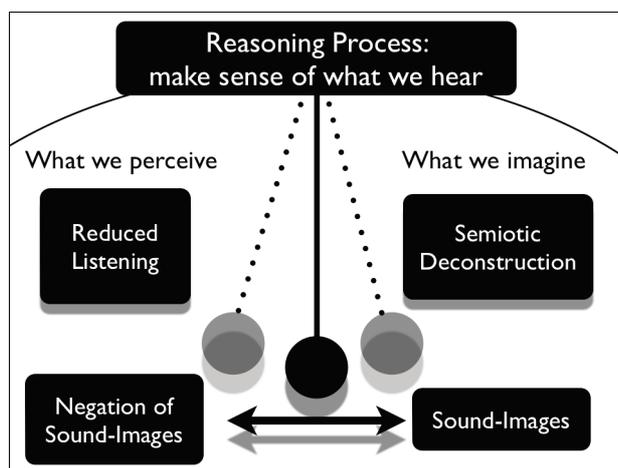


Figure 1. Listening process in an acousmatic situation

As such, by employing acousmatic reasoning in electroacoustic music, listeners listen to dis-embodied and dis-placed sounds by traversing back and forth, like a pendulum, between the semiotic and the spectromorphological (Figure 1). But how can we examine this act of travel? To answer this question, it is

necessary to explore the third subject of acousmatic reasoning, that is, listeners' inference processes.

2. ACOUSMATIC REASONING AT WORK

2.1. Inference processes in acousmatic reasoning

The power of listening modes and frameworks to orient the listening experience may render new sounds or changes in sound as random and trivial—as extraneous to the work and, therefore, *semantic noise*.³ With time, however, what initially is taken for noise or meaningless distortion can, through emerging order, design, and exposition, come to be heard as noteworthy and even essential to how one understands a piece. Bateson gives an intriguingly similar example of such a process in the perception of anatomy: 'To the aesthetic eye, the form of a crab with one claw bigger than the other is not simply asymmetrical. It first proposes a rule of symmetry and then subtly denies the rule by proposing a more complex combination of rules'. [2]

Notably, recognition of the crab's asymmetry comes only after the viewer's presupposition that the crab is, at some fundamental level, symmetrical (even though the two process appear to occur at the same time). By extension to music, Bateson's example suggests that the evaluation of new sounds or changes to the properties of existing sounds proceeds along a three-stage process: first, that, regardless of the listening situation, listeners bring whatever rules or codes they know to listening (no matter how limited or naive they may be); second, once a music succeeds in challenging a presupposed system of codes, listeners make new assumptions and broaden or change existing code systems or propose new ones; and then finally, listeners choose the best, most explanatory code system from amongst the plausible systems they have imagined.

This three-step listening process is based on an inferential process called *abduction*, termed by Pierce and further expanded by Eco. Acousmatic reasoning begins with the process of abductive listening and ends with the choices listeners make, however intuitively, consciously or personally. However, listening to music is an ongoing process that continuously offers up new sounds that can (and will) call into question the explanations listeners have made; consequently, the process of acousmatic reasoning needs to be applied continually, using it to accumulate additional input that either furthers previous explanations or undermines them on the way toward the invention of newer, more plausible ones. Listeners use abduction to imagine 'the

³ With regard to semantic noise compared to engineering (or channel) noise, Hockett notes: 'Ordinary Spanish usage reflects this distinction quite neatly. If A speaks to B and B responds with *no entiendo*, it means "I have not heard your words, because of interfering sound or lack of attention; please transmit the same message again"; if he responds with *no comprendo*, it means "I heard you all right, but what I heard doesn't make sense; please paraphrase or explain." Channel noise, thus, is the responsible factor when that which leaves a transmitter is not that which reaches the receiver; semantic noise is a discrepancy between the codes used by transmitter and receiver.' [12]

Rule' that might explain 'the Case' they are hearing. Once proposed, listeners 'widen (and verify)' the suppositional rule relative to new sounds and the ability of the rule to continue to sufficiently explain what they imagine. [6] Furthermore, Eco differentiates three types of abduction, *overcoded*, *undercoded*, and *creative abduction*, which correspond neatly to the three types of abductive listening found in acousmatic reasoning.

2.1.1. Overcoded abduction

Eco defines an abduction as *overcoded* when 'the law is given automatically or quasi-automatically'. [6] Assuming that there is some kind of process predisposing them to listen according to one or the other mode, we can say that each mode by itself, can, in essence, function as a kind of 'automatic' rule, in the spirit of Eco's overcoded abduction, specifying the way in which sounds will generally be heard.

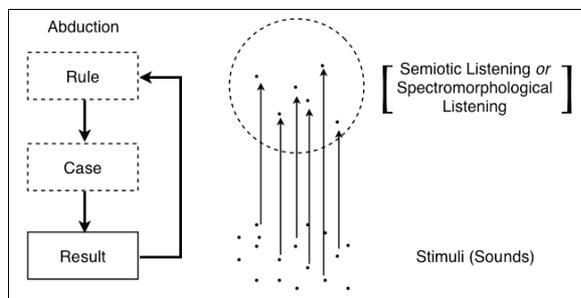


Figure 2. Overcoded abduction

Looking back at the discourse explored here, I have argued that within the practice of acousmatic reasoning, listeners have the choice to listen semiotically or spectromorphologically. However, given that listening according to one or the other mode is a choice, the rule is therefore adopted 'quasi-automatically,' since certain choices or conditions must first be met to incline or direct a listener into one or the other mode. Figure 2 shows the process by which sounds (aural stimuli), perceived according to one or the other mode, are explained by an overcoded abduction; the sufficiency of the abductive reading is determined by the ability of the stimuli to map into the rule's codes.

2.1.2. Undercoded abduction

In acousmatic reasoning, however, listeners suppose much more than any one listening mode might necessarily address. In particular, listeners imagine sound-images that function, however temporarily and tentatively, as rules, explaining the sequence of sounds they hear.

According to Eco, this form of abduction that employs imagination and supposition in the ongoing creation of explanatory rules (in our case, the sound-image as explanatory rule) is *undercoded* because 'the rule must be selected among a series of equiprobable

alternatives'. [6] Sound-images must be seen as *undercoded* due to the uncertainty surrounding them, the possibility of multiple explanations that introduce doubt; in spite of the doubts at play, the rules of sound-images, however undercoded, are standardly, albeit provisionally, embraced by listeners given the tenable explanations they momentarily offer.

Figure 3 illustrates how listeners, searching for explanation, choose the most plausible sound-image. In undercoded abduction, conditioned by the lack of pre-established rules, 'certain macroscopic portions of certain texts are provisionally assumed to be pertinent units of a code in formation, even though the combinational rules governing the more basic compositional items of expressions, along with the corresponding content-units, remain unknown' [5].

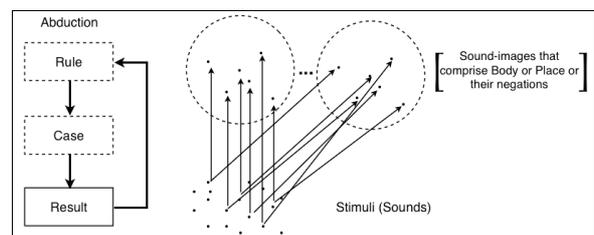


Figure 3. Undercoded abduction

Likewise, compared to the two listening modes that function as a pre-established and quasi-automatic rule in acousmatic reasoning, sound-images in undercoded abduction are based on a listener's provisional assumptions that 'certain macroscopic' properties of sound can be considered as 'pertinent units' of the supposed sound-images; but, as shown in Figure 3, the ability of some sound components to code into multiple sound-images introduces doubt, judgment, and even guessing into the explanatory process. Consequently, listeners, faced with sound-images that are only plausible, *entertain* them until they are seriously challenged, requiring the development of new sound-images—a process easily identified with what Eco calls *creative abduction*.

2.1.3. Creative abduction

In a *creative abduction*, 'the rule acting as an explanation has to be invented *ex novo* [anew or once again]'. [6] Eco further notes: 'In creative abductions one is not sure that the explanation one has selected is a "reasonable" one.... Many cases in which language is used not to confirm but to challenge a given world view or a scientific paradigm, and to decide that certain properties cannot belong any longer to the meaning of a given term require an interpretive cooperation that displays many characteristics of a creative abduction'. [6]

In a similar fashion, listeners who realize that some of the sounds or properties of sound that can no longer be explained by the sound-images they suppose need to

establish new sound-images. As shown in Figure 4, this process can involve not only the introduction of new sounds or new properties of existing sounds, but also the continuation of previous sounds into new contexts.

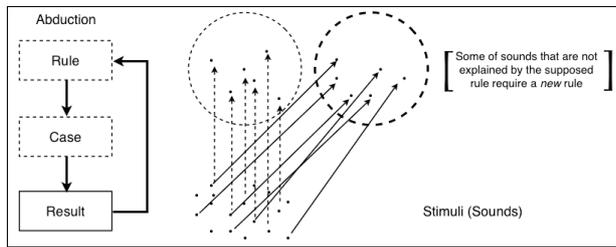


Figure 4. Creative abduction

2.2. Syntactically relevant properties (SRPs) in acousmatic reasoning

The abductive processes in acousmatic reasoning presuppose that sounds in an electroacoustic composition may signify more than one thing, and that, relative to a particular context, listeners choose the most plausible meaning. Using Eco's terms, the possibility of multiple significations in electroacoustic music listening is based on the assumption that listening to an electroacoustic composition is a process that examines 'a polydimensional network⁴ of properties, in which some properties are the *interpretants* of others'. [6] Given this, the question then becomes: what properties in acousmatic reasoning are meaningful to listeners?

The answer to this question depends upon a listener's current listening mode; as shown in Figure 5, semiotic listening focuses on sound-image components and their potential sound-image frameworks, while spectromorphological listening focuses on spectral types and morphological shapes.

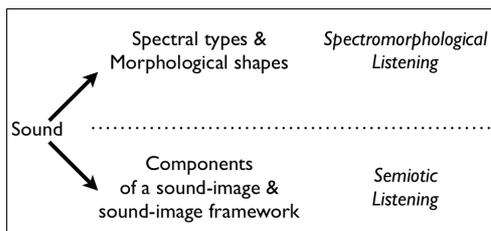


Figure 5. Properties of sound in acousmatic reasoning

Furthermore, not every property that can be categorized in Figure 5 is significant in acousmatic reasoning; some sound properties are more relevant than others, and that these differences affect acousmatic reasoning. The differences among properties of competing relevancy to the perception at hand is fittingly articulated by John V. Kulvicki in his book *On*

Images: Their Structure and Content, in which he proposes the concept of Syntactically Relevant Properties (SRPs): 'Let the *syntactically relevant properties* (SRPs) of tokens in a certain system of representation be those properties on which the syntactic identities of the tokens supervene. Changing the syntactic identity of a token requires changing its SRPs, though there can be changes in a token's SRPs that do not affect its syntactic identity. Shape, for example, is an SRP of most alphabets. One can change the shape of a letter without altering its syntactic identity, but if one wants to change a letter's syntactic identity, one must change its shape'. [15]

To look at this from a sound perspective, let us consider an audio example where we hear a voice whose first five formants are gradually shifted; in particular, formants centered at 609Hz, 1000Hz, 2450Hz, 2700Hz, and 3240Hz are shifted to 650Hz, 1100Hz, 2860Hz, 3300Hz, and 4500Hz, respectively. [4] What we hear is a person singing the vowel 'a' whose gender changes from male to female. In terms of SRPs, the SRPs associated with the gender of the voice can be identified as having changed from those of a male to those of a female; however, the SRPs associated with the voice itself has not changed. It is matters of audio detail like this, seemingly hidden, that can have a significant impact on how we hear electroacoustic music. Therefore, any discussion of the listening process based on acousmatic reasoning requires a close examination of SRPs, acknowledging their dependence upon not only the operative listening mode, but, as well, the image or framework to which listeners are, at any given moment, most attentive.

3. LISTENING TO THE FLYWHEEL DREAM USING ACOUSMATIC REASONING

3.1. Introduction to *The Flywheel Dream*

The Flywheel Dream by Paul Koonce [14] is replete with everyday sounds whose semiotic and spectromorphological features offer ample opportunities for listeners to engage in acousmatic reasoning. In line with the quaternary framework I have proposed elsewhere [13], the work's sounds often conjure up images associated with the human body or familiar places; more important, however, is the way these sounds undergo various processes of transformation that constantly force into question the identity of sound-images. For listeners, the continuous transformation of sound-images places them in a state of unrelenting tension where their uncertainties about what they are listening to keep them working to determine what each sound is (and just as relevantly, what it is not).

Given this, the focus of this (analytical) listening exercise is on the properties of the piece's sounds and sound-images, in particular, their SRPs. Furthermore, this exercise seeks to identify the functional relationship between SRPs and acousmatic reasoning and the imaginative process by which listeners create, question

⁴ Eco terms this network *Model Q*. For a further discussion about Model Q, refer to Eco's *A Theory of Semiotics*. [5]

and ultimately negate sound-images of body and place in their imaginative and acoustically reasoned movement through the quaternary framework for listening.

3.2. Analytical listening to *The Flywheel Dream*

In this analytical listening, we will look at the use of acousmatic reasoning to position various moments from 00:00 to 01:24. Following this, we will make detailed observations of some key sounds from 01:24 to 02:48, with the purpose of examining the interplay between semiotic and spectromorphological listening and the distinctive shifts that occur between the four poles of the framework—Body, Place, Non-body, and Non-place—during the segment.

3.2.1. *Acousmatic reasoning from 00:00 to 01:24*

Figure 6 shows the results of the analytical listening exercise I conducted from 00:00 to 01:24. Moving from top to bottom, we see spectromorphological features (SRPs) collected by a sonogram as well as through our spectromorphological listening followed by semiotic features (SRPs) collected through the semiotic listening. Below the semiotic features are the sound-image frameworks, identified by dotted circles, which I inferred through undercoded and creative abduction; the figure is completed at the bottom with the quaternary framework. As shown on the right side of the figure, both spectromorphological and semiotic listening modes are overcoded abduction processes as they progress according to relatively rigid rules connected with and distinguished by the particular form of listening. Sound-image frameworks, however, as shown, are products of undercoded or creative abduction, representing a listener's explanation of sonic events based not on hard facts or rules but on imagination.

Dotted-line arrows connecting one or more semiotic features (SRPs) to a sound-image framework indicate an undercoded abduction where the semiotic features become components of the sound-image framework. An important feature of sound-image frameworks is the question mark '?' beneath the name, which identifies that the imagined framework is in some way fleeting and difficult to verify as it is likely challenged by *missing, conflicting or distorting information*. [13] Question marks combined with exclamation points and arrows, '!→?' or '?←!', identify sound-image frameworks that shift from being affirmed to negated or vice versa.

Sound-image frameworks can combine or merge to become *super-set* sound-image frameworks, for example, sound-image frameworks like 'Kitchen' and 'Office' becoming sound-image components of another sound-image framework called 'Indoor.' Consider, also, a sound-image framework called 'Musical combo,' as found in Figure 6, which intersects with both 'Indoor' and 'Outdoor,' suggesting that, in this section of the

piece, 'Musical combo' could support both 'Indoor' and 'Outdoor' frameworks.

The Quaternary Framework, included beneath the sound-image frameworks, is divided into two levels: Body/Non-body and Place/Non-place. Gray boxes identify poles on the quaternary framework and solid arrows between them indicate shifts from one pole to another. If the shift is contained within a gray box, then the shift takes place *within* the same context. For example, in Figure 6, **BODY → NON-BODY** shows that the transition from Body to Non-body—in this case, the shift of 'Child!→?'—is due to the changes in the SRPs of 'Child,' which should therefore be identified together. Compared to the shift from **BODY** to **NON-BODY**, the shift from **BODY → NON-BODY** to **PLACE**, which is not in a single box and has an angled arrow shifting focus to a different axis, indicates that **BODY → NON-BODY** does not become, but is rather subsumed by **PLACE**.

Section (a) has a quality of relentless locomotion that is supported by both semiotic and spectromorphological readings of the work. This sense of locomotion is caused not only by the spectromorphological features of the collaged sounds, but also by the diversity of materials, which resist being circumscribed by any one sound-image of body or place. The sound-image frameworks in section (a) are potentially somewhere indoors, possibly an office or a kitchen or a domestic place with added environmental noise. However, none of these sound-image frameworks is definitively distinguished by support through sounds, context, or circumstance. Section (a) may offer listeners the opportunity to imagine a sound-image framework connected with place, but there is just *too much conflicting information*, or better, too many plausible yet unconfirmed frameworks that leave listeners little choice but to take what they hear, for the moment, as **NON-PLACE**, as shown in the Quaternary Framework.

In section (b), listeners get a strong sense of the human body as a result of SRPs associated with the body: the voice of a young child or even baby, and footsteps moving through shallow water. However, these SRPs are *distorted* by expansion in time and the introduction of reverberation; consequently, listeners are led away from 'Child!' toward 'Child?' as the image of the voice becomes less and less natural and the distortions of a reverberant space are added.

Our traversal through the quaternary framework in section (b) begins with Body, due to the strong sense of the human body as we have discussed, but quickly moves to Non-body. To form Non-body in this section, listeners collect together the reverberated and time-expanded voices, joining them together based on their spectromorphological similarities. Place soon subsumes the shift in the quaternary framework from Body to Non-body as a result of the cave framework whose SRPs include not only reverberated voices and footsteps, but also the sound of birds and reverberated water drops. In section (b), thus, the listener's experience, as mapped onto the quaternary framework,

shifts from Body to Non-body, and then finally, to Place.

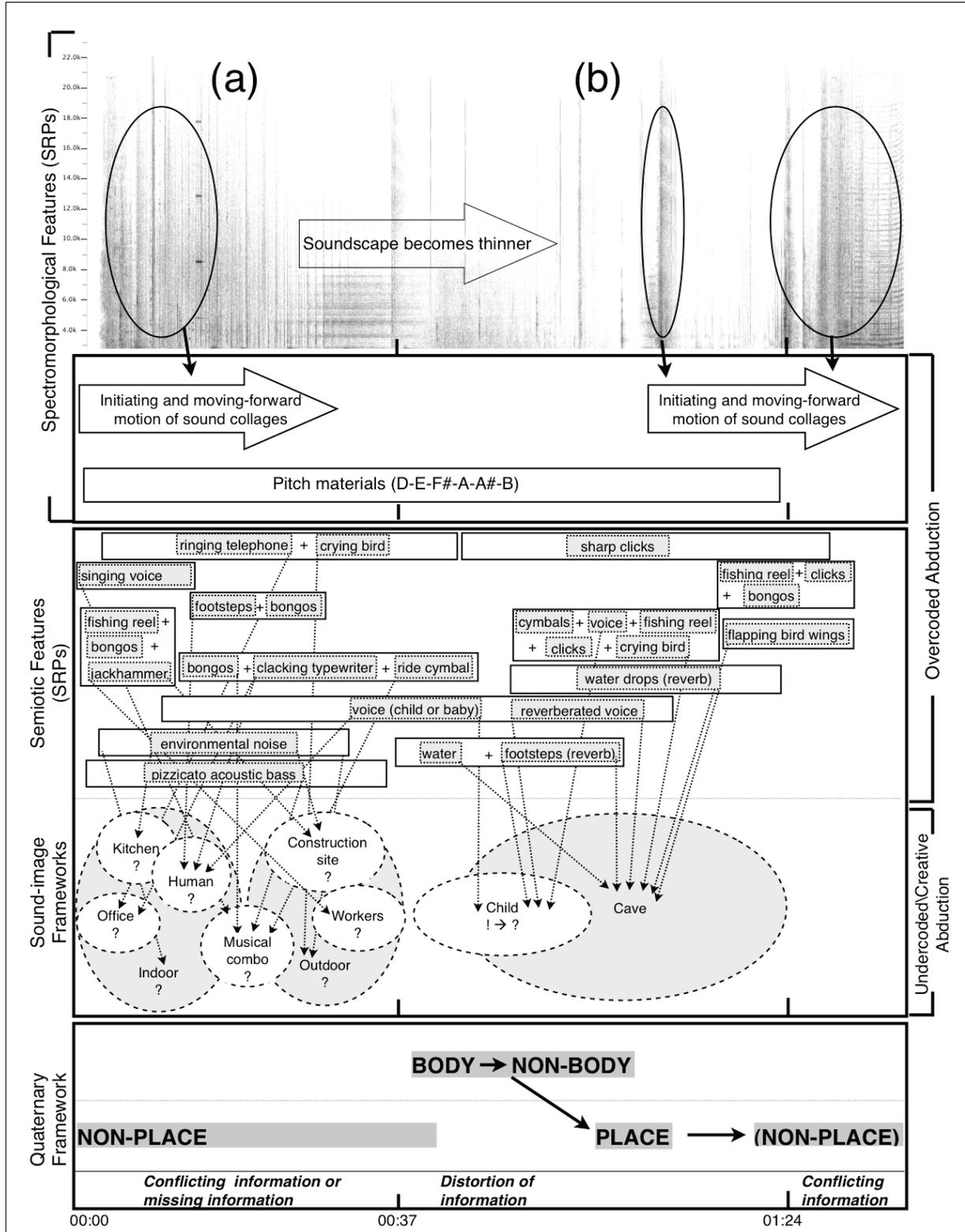


Figure 6. Detailed listening exercise for *The Flywheel Dream* (00 :00 – 01 :24)

Section	Symbols and Codes	Explanation
Spectromorphological Features (SRPs)	(a), (b)	• Subsections
Semiotic Features (SRPs) and Sound-image Frameworks		<ul style="list-style-type: none"> • Sounds in gray, dotted-line rectangles denote preliminarily identified sound-image features and possible sound-image framework components. • Sound-image features enclosed in white boxes and linked by “+” denote spectromorphologically linked sounds, regardless of semiotic implications. • Dotted lines with an arrow identify the use of sound-image features in sound-image frameworks as a result of undercoded or creatively coded abduction. • Sound-images in white, dotted-line circles denote sound-image frameworks formed through undercoded or creatively coded abductions using one or more sound-image features as framework components. • Sound-image frameworks with gray backgrounds indicate ‘super-set’ frameworks; super-set frameworks are formed through the joining together of related frameworks into more generalized frameworks. • Question marks beneath the name of a sound-image framework identify the framework as questioned, challenged, or potentially negated. • An exclamation point followed by a question mark beneath the name of a sound-image framework identifies the framework to be first affirmed and then questioned, challenged, or even negated. • A question mark followed by an exclamation point beneath the name of a sound-image framework identifies the framework to be initially questioned, challenged, or even negated, but then finally affirmed.
• Quaternary Framework		<ul style="list-style-type: none"> • Quaternary framework poles in gray boxes denote the pole currently in effect; boxes extend over the length of time during which the pole is identified. • Gray boxes containing processes of negation or affirmation (denoted by arrows) indicate framework shifts operating within the context of a single sound-image framework. • Solid arrows denote transition between two or more poles of the framework. • Poles enclosed in parentheses indicate preceding or upcoming poles otherwise not indicated by accompanying information.

Table 1. Explanation of symbols and codes used in Figure 6 and Figure 7

One benefit of the detailed listening is how it illustrates the interaction between spectromorphological and semiotic listening modes. In Figure 6, we can see that, in the semiotic listening section, there are various sounds linked by a plus sign (+) and circumscribed by a rectangle. These notations indicate that the sounds are linked together spectromorphologically, despite the possibility that they may be semiotically unrelated. *The Flywheel Dream* offers listeners copious examples of these unlikely semiotic bedfellows cultivated through spectromorphological congruencies; notable examples include the fishing reel + bongos + jackhammer sounds

in the beginning to 00:05, the footsteps + bongos from 00:19 to 00:22, the interplay in rhythm and pitch between the ringing telephone + crying bird sounds from 00:24 to 00:46, and the rhythmic interplay between the bongos + clacking typewriter + ride cymbal from the beginning to 00:40. I have mentioned that, in section (b), the voice loses its identity when it is combined with other sounds. Between 01:07 and 01:12, you can hear a collage of sounds that starts with clicks, followed by (in series) the ride cymbal, fishing reel sounds with a time-expanded voice, water drops, fishing reel sounds with clicks, and a bird call. This is another example of the

juxtaposition of semiotically heterogeneous sounds based on the similarity of their spectromorphological designs. The role of this sound collection is twofold: it functions to not only move away from the cave as a place (a semiotic function), but to also reinitiate the motion observed in section (a) and move into the next section (a spectromorphological function), thus, effectively encouraging listeners to move from Place to Non-place.

3.2.2. *Interplay and metamorphosis of SRPs in acousmatic reasoning from 02:10 to 02:50*

The analytical exercise explored above using *The Flywheel Dream* shows how acousmatic reasoning works in listening. This paper will now examine, in more detail, the interplay of SRPs (semiotic and spectromorphological properties) in acousmatic reasoning. The segment to be examined is from 01:24 to 02:48, which includes the sound sequence fishing reel → clicks → water drops → ticking bicycle derailleur → ping-pong balls → ticking clock → alarm clock from 02:10 to 02:50.

Shown in Figure 7 are the results of an analytical listening exercise for the segment mentioned above from 02:10 to 02:50. This segment offers another example of the interplay between spectromorphological and semiotic features in the piece, displayed primarily through its rhythm (a spectromorphological feature). The continual metamorphosis of the segment's rhythmic features causes it to momentarily conjoin with different semiotic features and sound-images.

The segment begins with the sound of a reverberated and time-expanded trilling flute, along with the sound of a clicking fishing reel, which, by this point in the piece, should be familiar to listeners. One of the spectromorphological features of the fishing reel is its decelerating rhythm. Based on semiotic listening, one can hear the clear combination of the fishing reel with the clicks, which are then followed (through superimposition) by the sound of water drops, the bicycle derailleur, ping-pong balls, a clock, and finally, an alarm. Listening to the metamorphosis of the fishing reel into the water drops and finally the clock requires, among other things, close attention to the regularity and irregularity of sound events. For example, where the water drops are intermittent, the clock, conversely, stands out through its much faster, regular ticking. The fishing reel, which is regular while changing in speed, decelerates and metamorphoses into the sound of the water drops, gradually adopting their irregular rhythmic identity. At the same time, the water drops, the ticking bicycle derailleur, and quasi-regular bouncing ping-pong balls metamorphose into the regularity of the clock. As shown in the spectromorphological listening diagram, the rhythmic features progress from 'decelerating' to 'irregular' to 'regular,' which coincides with the contributing sounds, from clicking fishing reel to ticking clock to finally alarm (the solid arrows indicate the contribution that spectromorphological

rhythmic features make to the sequence and identity of sounds).

In addition, it should be noted that the change of sound-images through changes in their degree of regularity is a feature that combines with other sounds to create new sound-images and shift preexisting ones. The rhythmically irregular sounds identified as water drops combine with the running water to suggest a cave as a place, and the rhythmically regular sounds that eventually become identified as a clock, underscored at the end of the segment by a disrupting alarm, combine with the sound of soft, restful breathing to suggest a bedroom as a place.

The stability of Place is present and evident in this metamorphosing segment. However, the cave place, which is a reprise of the previous cave image, carried over by reverberated water drops and running water, is not as strong a sound-image as when it was first introduced. Compared to the earlier moment, this sound-image lacks the information listeners need to identify it as a cave; even the most important cave SRPs, such as those connected with the water drops, quickly change to suggest a bicycle derailleur or ping-pong game, causing our listening experience to shift from a cave to something that is not a cave ("Cave (Reprise)! →?"), and, by correspondence within the quaternary framework, from Place to Non-place. Furthermore, the bedroom, while identifiable, appears but for a few seconds, as it is quickly eclipsed by the awakening call of the alarm clock and the collage of noisy sounds that follow; the intrusion of construction sounds and honking cars combine with the return of familiar sounds to re-engage the toy box spilling moment. Thus, sound-images and frameworks connected with Place are as soon challenged as they are imagined, as denoted by "Bedroom! →?".

3.3. Macro-experience of *The Flywheel Dream* based on acousmatic reasoning

Despite the various easily recognizable sounds that we hear in *The Flywheel Dream*, listening to it semiotically is never easy; the piece resists the listener's semiotic reading of it. While we are able to detect some place-like images in the piece, they are, however, never verified. Interestingly, spectromorphological features—the primary pitch content and morphological shapes—offer listeners a simpler way to 'understand' *The Flywheel Dream*. Sounds carrying these spectromorphological features are the common thread that runs through the piece, knitting together all of the heterogeneous materials that do not otherwise go together. Listening to the piece according to acousmatic reasoning, one can regard the piece as being similar in structure to the mobiles of Alexander Calder in that the spectromorphological features of the piece act as the armature holding together copious and seemingly mobile objects. Furthermore, like a mobile whose moving parts are powered by wind, a motor, or the observer, *The Flywheel Dream* is powered by the

semiotic whimsicality of the sounds as well as the listener's imagination.

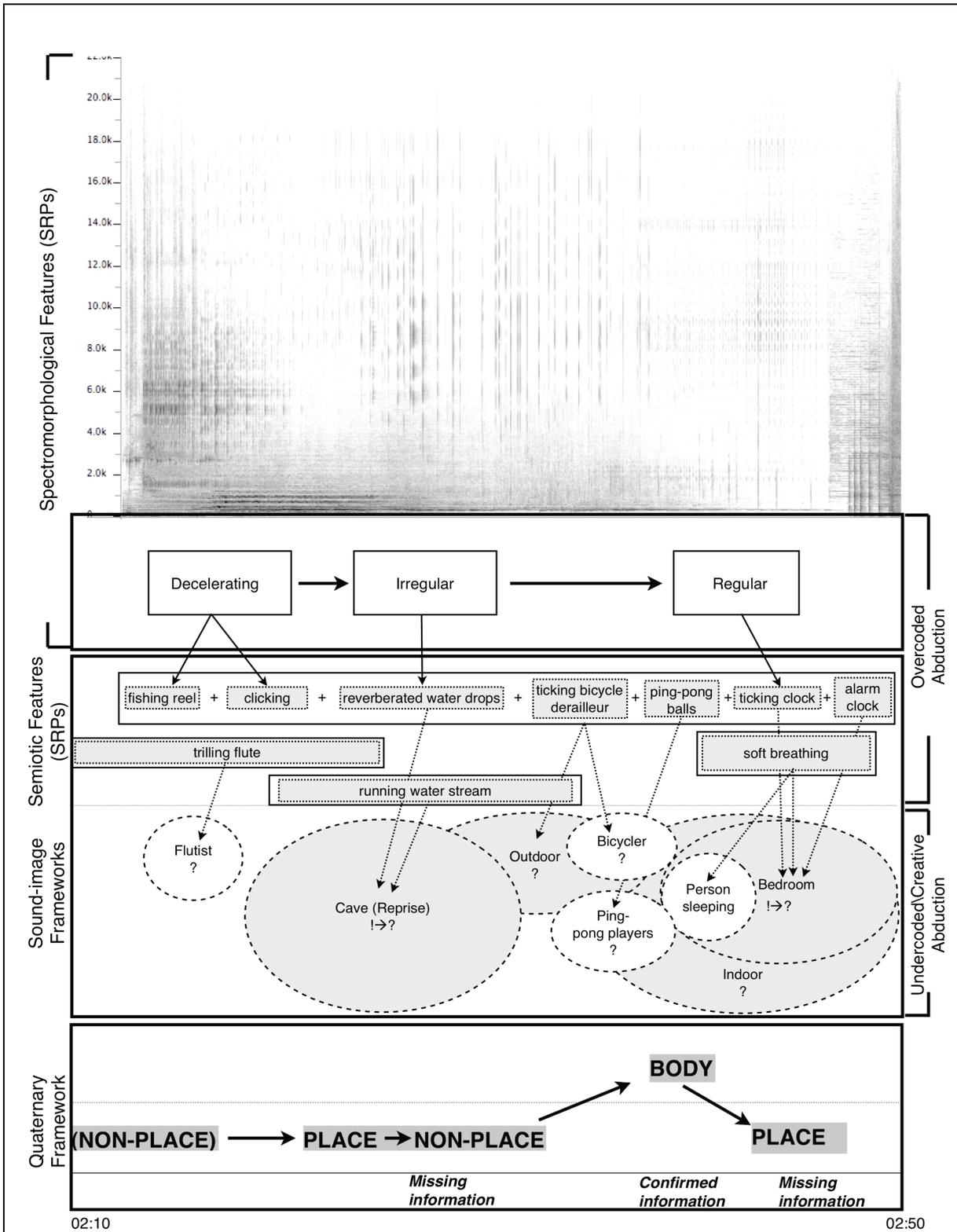


Figure 7. Detailed listening exercise for *The Flywheel Dream* (02 :10 – 02 :50)

4. CONCLUSION

The goal of the proposed acousmatic reasoning, well stated, has been to address how listeners listen to electroacoustic music that is clearly engaged with everyday sounds and their connotations. By definition, other forms of electroacoustic music, not so engaged with everyday sounds, are not addressed. Additionally, acousmatic reasoning is a process not just for listeners, but for composers as well. As composers more and more realize how they are, in fact, their first listener, they need to learn how to be attentive to what they, and others, experience when listening and work to examine how acousmatic reasoning, as a fundamental process behind how they listen, plays a role in their compositional processes.

Listening to electroacoustic music is as much an act of reasoning as of imagining. Good listeners are not only attentive to both modes of listening, but, as well, balance their attention in accord with the pieces they hear. Acousmatic reasoning, as proposed in this paper, shows how listeners proceed to listen.

5. REFERENCES

- [1] Agawu, V. K. *Playing with Signs: A Semiotic Interpretation of Classic Music*. Princeton, N.J: Princeton University Press, 1991.
- [2] Bateson, G. *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. Northvale, N.J: Aronson, 1987.
- [3] Casey, E. S. *Imagining: A Phenomenological Study*. Bloomington: Indiana University Press, 1976.
- [4] Dodge, C., Jerse, T. A. *Computer Music: Synthesis, Composition, and Performance*. New York: Schirmer Books, 1997.
- [5] Eco, U. *A Theory of Semiotics*. Bloomington: Indiana University Press, 1976.
- [6] Eco, U. *Semiotics and the Philosophy of Language*. Bloomington: Indiana University Press, 1984.
- [7] Emmerson, S. « The relation of language to materials », *The Language of Electroacoustic Music*. London: Macmillan, 1986.
- [8] Emmerson, S. *Living Electronic Music*. Aldershot, UK; Burlington, VT: Ashgate Publishing, 2007.
- [9] Godøy, R. I., Haga, E., Jensenius, A. R. « Playing 'air instrument': mimicry of sound-producing gestures by novices and experts », *Gestures in Human-Computer Interaction and Simulation: 6th International Gesture Workshop GW*, 2006.
- [10] Godøy, R. I. « Gestural-sonorous objects: embodied extensions of schaeffer's conceptual apparatus », *Organised Sound* 11(2), 2006.
- [11] Godøy, R. I. « Imagined action, excitation, and resonance », *Musical Imagery*. Exton, PA: Swets & Zeitlinger Publishers, 2001.
- [12] Hockett, C. F. « An approach to the quantification of semantic noise », *Philosophy of Science* 19(4), 1952.
- [13] Kim, S.- J. « Imaginal listening: a quaternary framework for listening to electroacoustic music and phenomena of sound-images », *Organised Sound* 15(1), 2010.
- [14] Koonce, P. *The Flywheel Dream* (1994), *Walkabout & Back: Electroacoustic Works*. New York: Mode, 2000.
- [15] Kulvicki, J. V. *On Images: Their Structure and Content*. Oxford; New York: Clarendon, 2006.
- [16] Leman, M. *Embodied Music Cognition and Mediation Technology*. Cambridge, Mass: MIT Press, 2008.
- [17] Neuhoff, J. G. *Ecological Psychoacoustics*. Amsterdam; Boston: Elsevier Academic Press, 2004.
- [18] Smalley, D. « Spectro-morphology and structuring processes », *The Language of Electroacoustic Music*. New York, NY: Harwood Academic Publishers, 1986.

INSTRUMENTER LES OPERATIONS D'ÉCOUTE ANALYTIQUE ? UN BILAN DU PROJET 'ÉCOUTES SIGNÉES' (2003-2006)

Nicolas Donin

Samuel Goldszmidt
S.T.M.S. (UMR 9912)
IRCAM-CNRS
1, pl. Igor-Stravinsky
F-75004 Paris
prenom.nom (at) ircam.fr

Jacques Theureau

RÉSUMÉ

Le projet "Écoutes signées" portait sur la conception d'outils d'assistance à l'écoute analytique de la musique. Maquettes et prototypes de modules informatiques ont été développés collaborativement par un ingénieur multimédia, un musicologue, un ergonomiste et des musiciens dont la pratique d'écoute était ainsi explicitée et outillée. Nous présentons un bilan de l'ensemble des travaux menés dans ce cadre et les inscrivons dans un programme de recherche technologique plus large.

1. INTRODUCTION

S'il y a abondance d'outils de création musicale, il existe peu d'outils d'assistance à l'écoute musicale. Il existe bien des dispositifs portatifs autonomes du type baladeur. Mais ces derniers ont été conçus et vendus sur la seule base de leur capacité de stockage (qui n'a cessé d'augmenter), de la qualité d'encodage des fichiers musicaux (qui, elle aussi, n'a cessé d'augmenter, passant de 128kbps comme norme en 2003, à 190kbps voir 320kbps en 2006), et de l'ergonomie de leur interface d'accès à ces fichiers musicaux (dont les progrès sont mitigés). Une fois la lecture d'une musique déclenchée, très peu de fonctions d'assistance à l'écoute ont été développées et proposées à l'utilisateur hormis une barre de navigation dans le son. Pourtant, bien des technologies pourraient enrichir cette écoute.

Parti de ce constat, le projet "Écoutes signées" a tenté, durant une période déterminée et avec des moyens déterminés, d'explorer différentes instrumentations possibles de l'écoute musicale, susceptibles d'enrichir des pratiques d'écoute existantes et de s'intégrer à des dispositifs techniques actuels ou futurs. Si l'impulsion initiale du projet est due à un philosophe alors directeur de l'Ircam, Bernard Stiegler, sa réalisation et sa transformation ressortissent au programme de recherche d'un groupe associant musicologie, ergonomie cognitive et ingénierie documentaire multimédia (voir [23]). L'intérêt de faire un bilan du projet est multiple : rassemblement d'informations inégalement publiées ; mise en évidence des idées directrices qui les sous-tendent ; retour critique sur une expérience dont les

principes conservent un potentiel de fécondité à exploiter, par nous ou par d'autres.

2. ABORDER L'ÉCOUTE MUSICALE COMME UNE PRATIQUE HISTORIQUEMENT ET TECHNIQUEMENT SITUÉE

Le projet "Écoutes signées" posait que l'écoute musicale est 1) *active* (ce n'est pas une réception passive d'informations), 2) *médiate* (elle a lieu dans un contexte spécifique et au moyen de médiations techniques), et 3) *singulière* (si certaines caractéristiques d'une écoute sont largement partagées, d'autres sont irréductiblement liées à la situation, la pratique et l'individu considérés). En conséquence, concevoir des outils d'aide à l'écoute impliquait 1) une relation à des pratiques individuelles ou collectives, 2) une insertion "écologique" dans un environnement technique et cognitif pré-existant, 3) une prise au sérieux du caractère idiosyncrasique, difficilement généralisable, des manières d'écouter susceptibles d'être observées et instrumentées.

2.1. Musicologie et instrumentation de l'écoute

Ces postulats se réfèrent à divers travaux pionniers en musicologie de l'écoute : une étude empirique de Delalande a proposé une typologie de « conduites d'écoute » face à une même œuvre entendue dans les mêmes conditions par plusieurs sujets [6] ; Hennion et al. ont enquêté sur les formats et dispositifs de l'écoute des amateurs actuels de musique [18] ; Szendy a mis en évidence l'historicité constitutive de l'écoute ainsi que sa capacité à être « signée » et « adressée » [22] ; Sterne [21] et Maisonneuve [19] ont fait l'archéologie de pratiques d'écoute éloignées dans le temps (avant et après l'invention du phonographe) ; nos propres travaux musicologiques, réalisés en parallèle du projet "Écoutes signées", ont montré les liens entre l'analyse musicale et diverses techniques d'écoute attentive depuis la fin du XIX^e siècle [4]. Enfin, parmi les rares précédents reliant réflexion sur l'écoute musicale et projet informatique, on citera les travaux de Philippe Aigrain et al. [1] et ceux relatifs à l'Acousmographe au GRM [15]. S'ajoutent nos propres recherches empiriques (en relation ou en parallèle avec les "Écoutes signées") qui

ont abordé l'écoute dans les activités de composition et d'interprétation musicales (voir, par ex., [12] et [14]).

2.2. Un programme de recherche technologique...

Plus généralement, le programme de recherche technologique dans lequel s'est inscrit le projet "Écoutes signées" peut être caractérisé par son objet (la conception d'outils d'assistance à l'écoute musicale analytique), sa relation avec la formulation et l'étude d'hypothèses empiriques concernant l'écoute musicale, son principe méthodologique essentiel (la participation à cette conception d'« écouteurs qualifiés »), et sa relation privilégiée avec un domaine de l'ingénierie informatique, les techniques hypermédia.

Cet objet ouvre sur une multiplicité d'outils d'assistance à l'écoute musicale analytique, qui peuvent être juxtaposés ou intégrés. Ils sont destinés à participer à la composition de situations dynamiques (c'est-à-dire comprenant l'activité des utilisateurs concernés) dont certains aspects peuvent aussi être conçus (par exemple, peut s'ajouter une conception de la formation nécessaire, etc.), tandis que d'autres aspects ne peuvent constituer que des contraintes de cette conception. Ils sont aussi destinés à connaître des processus d'appropriation et d'individualisation, donc de détournement, divers de la part de ces utilisateurs.

2.3. ... ses hypothèses empiriques...

Dans ce programme de recherche technologique sont en jeu une variété d'hypothèses empiriques sur l'écoute, d'hypothèses sur le développement des processus de conception informatiques et d'hypothèses sur la sélection des techniques hypermédia et les développements à leur faire subir pour les combiner dans les produits de conception. Ces diverses sortes d'hypothèses ne sont pas triviales, en ce sens qu'elles sont pour une bonne part alternatives d'autres hypothèses largement acceptées.

Les hypothèses empiriques sur l'écoute musicale développent les trois principes du projet "Écoutes signées" présentés en introduction à cette section. Parmi ces développements, signalons particulièrement : non monopole de la modalité sensorielle auditive ; insertion de tout moment d'écoute dans une trame de protentions et rétentions ; structure de renvoi (rappel mnémonique à plus ou moins long terme, mais aussi renvoi vers des perceptions et actions futures) ; intégration dans des pratiques plus larges, au passé et au futur (par exemple, la composition, l'interprétation, l'improvisation, mais aussi des pratiques sociales diverses) ; écoute active allant de l'écoute la moins analytique à la plus analytique, celle qu'on peut caractériser comme à la fois attentive au détail, continue, répétée, donnant lieu à inscription, engagée dans des pratiques actives plus larges et susceptible d'être guidée.

2.4. ... son principe méthodologique...

Ce programme de recherche technologique intègre l'ensemble de ces différentes hypothèses sans bénéficier nécessairement d'études empiriques systématiques de l'écoute, et ce grâce à la participation d'« écouteurs qualifiés », c'est-à-dire des individus ou groupes d'individus développant une pratique d'écoute particulière d'un type de corpus particulier, et dont il semble intéressant de concevoir une instrumentation informatisée qui leur permettrait 1) de reproduire et/ou modifier certains aspects de leur pratique d'écoute ; 2) à cette occasion, d'explicitier cette dernière ; 3) de la rendre intelligible à autrui, de la partager, voire d'éprouver sa généralité.

Le collectif de conception regroupe le(s) « écouteur(s) qualifié(s) », un(des) ingénieur(s) multimédia, un musicologue, un ergonomiste. La co-conception passe par cinq étapes : 1) le ciblage d'aspects pertinents de la pratique d'écoute considérée ; 2) la schématisation de divers scénarios d'instrumentation, d'abord sous forme papier puis sous forme de maquettes multimédia non fonctionnelles (mais d'apparence suffisamment réaliste pour pouvoir être ou non validées au titre d'objectifs souhaitables) ; 3) la réalisation d'un prototype fonctionnel ; 4) son aboutissement sous forme d'animation, de module, de publication web ou sur support ; 5) l'intégration au sein d'un environnement d'écoute assistée par ordinateur (non encore réalisée).

2.5. ... et ses techniques hypermédia

Privilégier la transmission implique, soit d'utiliser des techniques de lecture hypermédia existantes et largement diffusées (pour la période considérée, c'est la suite Macromedia Flash qui dominait le marché, avec le *plug-in* de lecture Macromedia FlashPlayer et l'environnement d'édition Macromedia Flash, les alternatives étaient peu disséminées), soit d'articuler des techniques informatiques musicales, en particulier celles développées à l'Ircam, avec les techniques de lecture hypermédia usuelles (en l'occurrence, le FlashPlayer a été exploité extensivement, notamment en relation avec Max/MSP). Nous voulions aussi pouvoir échanger des données avec d'autres systèmes, garantir l'interopérabilité et la lecture des maquettes, ouvrir aux commentaires et à la critique scientifique, publier sur plusieurs supports et interroger l'interaction de l'utilisateur avec des médias typés (par exemple une partition musicale). Macromedia Flash et Macromedia Director étaient, entre 2003 et 2006, les candidats idéaux pour réaliser ces prototypes. Nous avons opté pour Flash, davantage orienté vers la publication Web.

3. DÉROULEMENT DU PROJET

Le phasage du projet fut le suivant :

- 2003 : identification d'un ensemble de pratiques d'écoute et premières maquettes ponctuelles ;

- 2004-5 : version 2 de plusieurs maquettes et adaptation d'une maquette au projet Semantic Hifi [17] ;
- 2005-6 : finalisation et publication de plusieurs "écoutes signées".

Les attendus et les résultats de la première phase ayant été décrits et discutés dans [7], nous nous limiterons ici à les résumer et compléter. La deuxième phase a permis un gain en abstraction par l'identification de certaines « opérations » (ou gestes techniques) d'écoute, que nous retiendrons ici pour rubriquer les différentes maquettes présentées. La troisième phase apparaîtra enfin à travers des renvois bibliographiques systématiques aux résultats publiés.

Les différents sous-projets aboutis ou mis en chantier dans la période 2003-2006 sont recensés ci-dessous, afin d'en montrer la variété et les manques, donc les autres possibilités à explorer :

« Ecouteur »	Corpus	Dispositif
Andrea Cera* (compositeur)	Sélection d'échantillons d'enregistrements discographiques de musiques populaires	Micro-séquenceur permettant l'interaction avec chaque item d'une transcription
Marc Chemillier* (ethnomusicologue)	Enregistrements de harpistes NZakara	Animation visuelle de structure canonique
Nicholas Cook (musicologue)	Ensemble de versions discographiques du 1 ^{er} Prélude du <i>Clavier Bien Tempéré</i>	Interface de navigation fine (multi-échelles) et de représentations de calculs dans la base de données de sons
Nicolas Donin* (musicologue)	Partition + enregistrement d'une pièce de Webern. Enregistrement + partition d'une pièce de Ravel.	Synchronisation multimédia partition + enregistrement + annotations temporalisées
Jonathan Goldman* (musicologue)	Partition + enregistrement d'une pièce de Boulez Articles musicologiques sur la mise en tableau.	Outil de mise en tableau de segments issus de la partition et synchronisés à un enregistrement
Philippe Leroux* (compositeur)	Partition manuscrite et documents liés à la composition : esquisses, brouillons, fichiers son et entretiens filmés de remise en situation de composition	Outil de navigation dans les traces de composition et maquettes interactives détaillant des aspects componentiels de chaque mouvement de <i>Voi(r)ex</i>
Membres du MIM* (compositeurs et théoriciens)	Enregistrements sonores et enquête sur l'activité d'analyse musicale du MIM	Outil de segmentation audio et de descriptions des segments
Makis Solomos (musicologue)	Retranscriptions pré-analytiques de partitions	Outil de transcription

Table 1. Liste des différentes « écoutes signées »

Les modalités de participation des « écouteurs qualifiés » n'ont pas été uniformes, la plupart étant physiquement présents (signalés ci-dessus par une *), tandis que la pratique d'écoute des autres était élicitée sur la base de leurs écrits.

4. COMPARER

Comparer plusieurs versions d'une même œuvre constitue une expérience multiple qui transforme notre écoute de telle ou telle version singulière. En outre, la

navigation / comparaison entre plusieurs interprétations est en soi une pratique d'écoute repérée (bien que mal connue), caractéristique d'un rapport moderne à la musique enregistrée.

4.1. La « Bachtèque » : principes

Le principe de la « Bachtèque », développée en 2004, est de mettre en regard une série d'enregistrements sonores du premier prélude du *Clavier bien tempéré*, et de les rendre comparables par une représentation simultanée de certaines de leurs caractéristiques (durée entre chaque attaque ; tempo moyen par mesure ; intensité moyenne de chaque note), fournies par un algorithme d'alignement entre fichier midi et fichier audio développé à l'Ircam.

Pour ce faire, une référence importante de l'analyse musicale d'interprétation a été mise à contribution : l'étude de N. Cook intitulé "Structure and Performance Timing in Bach's C major Prelude (WTC 1): An Empirical Study" (*Music Analysis*, 6, 1987). Dans cet article, Cook considère que les interprètes sont auteurs d'« analyses musicales » au même titre que les musicologues professionnels, mais à l'importante différence près que leurs analyses sont exprimées sous la forme de réalisations sonores du texte analysé. Par le contrôle des divers aspects expressifs du son (par ex. au piano : l'intensité, le toucher, la pédalisation, le phrasé, etc.), l'interprète propose à sa façon une segmentation, une mise en relation de différents motifs et passages, ou encore un rythme harmonique [8].

4.2. Descriptif fonctionnel

La Bachtèque regroupe des enregistrements de Glenn Gould, Edwin Fischer, Paul Jacobs, Helmut Walcha, Ralph Kirkpatrick et Wanda Landowska. Deux interfaces principales permettent de naviguer entre ces interprétations.

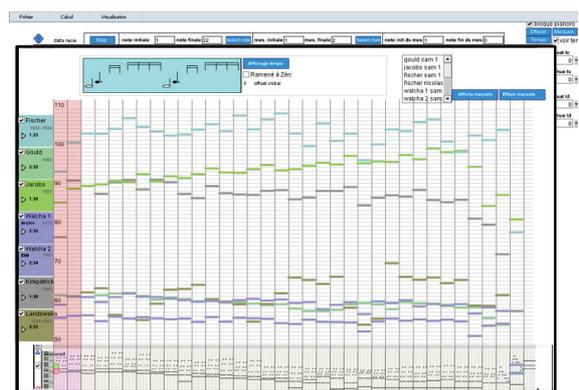


Figure 1. Bachtèque - Comparaison de temps à l'échelle macroscopique

La première interface donne une vision macroscopique du prélude et l'accès à plusieurs types de calculs du tempo moyen par mesure pour chaque interprétation. Un affichage partiel des résultats permet de distinguer des familles d'interprétation sous l'aspect

des variations de tempo. L'écoute d'une section déterminée se fait par une sélection dans une vue synthétique du prélude sous forme de *pianoroll* (en bas sur la Figure 1).

La deuxième interface (Figure 2) propose une visualisation en tableau de chacune des interprétations (forme d'ondes les unes au-dessus des autres) et de la partition (sous deux formes : sur portée linéaire issue du fichier MIDI et sous forme de *pianoroll* – ce dernier étant commun aux deux interfaces). Opérer une sélection depuis l'une de ces représentations sélectionne simultanément les passages correspondants dans les autres lignes, facilitant l'écoute comparée d'un même passage dans chaque version.

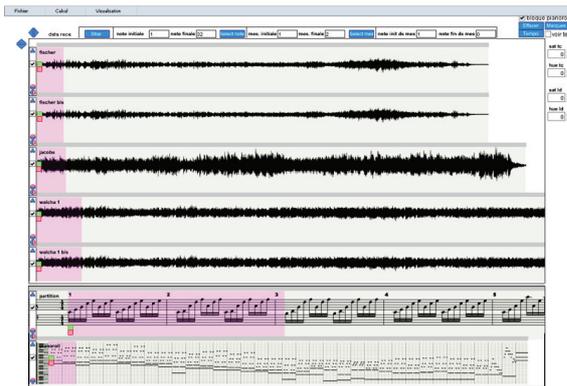


Figure 2. Banothèque – Navigation / sélection à l'échelle microscopique

Cette interface inclut des facilités de zoom et de sélection automatiques. Dans une sélection, l'utilisateur peut appeler une représentation en vis-à-vis de la durée et de l'intensité de chaque note. Ces différentes fonctionnalités permettent de retrouver – ou non – à l'écoute d'infimes variations de durée calculée algorithmiquement. Afin de faciliter la focalisation de l'attention sur telle déviation locale, l'utilisateur peut poser des marques sur chaque note ; afin de garder trace de ce processus, il peut également qualifier par un texte l'état d'une sélection et/ou de son calcul.

4.3. Descriptif technique

L'interface a été développée en ActionScript 2 avec Macromedia Flash MX. Le moteur audio de lecture des fichiers sons est réalisé avec Max/MSP. La communication entre Max/MSP et l'extension FlashPlayer s'effectue à partir de l'objet Max *flashserver*, il permet notamment le positionnement avec précision du curseur sur l'image de la forme d'onde. Un fichier XML, produit par l'algorithme d'alignement, indique pour chaque interprétation, la durée et l'intensité des notes. Ce fichier, invariant, est chargé au lancement de l'application. Les paramètres d'affichage des représentations et les annotations de l'utilisateur sont sauvegardés dans un fichier XML.

Cette maquette ne fonctionnait que hors-ligne. D'abord pour une raison d'économie de temps de développement des processus d'écriture et de lecture des

fichiers d'annotation sur un serveur distant, qui n'avait pas d'intérêt immédiat dans notre projet. Ensuite, du fait que l'objet *Sound* (pour la gestion du son) de la machine virtuelle de Flash ne permettait pas en 2004 une précision à la milliseconde près pour le déclenchement des sons – d'où l'utilisation conjointe de *flashserver* et de Max/MSP.

5. TRANSCRIRE

On détaille dans cette section sur les opérations de transcription des maquettes développées à des différentes périodes du projet "Écoutes signées" avec des objectifs variés.

5.1. Marc Chemillier/Harpistes Nzakara

Plusieurs maquettes ont été réalisées sous la prescription de Marc Chemillier, informaticien et ethnomusicologue (EHESS), à partir de ses recherches sur la musique des Nzakara (cf. [24] [26] pour leur volet multimédia). Une 'tablature animée' lui a permis de mettre en évidence des aspects de la structuration canonique du répertoire de harpe de cette ethnie.

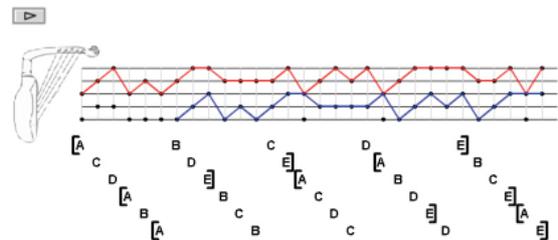


Figure 3. Maquette transcription d'un enregistrement de harpe Nzakara

Sur cette copie d'écran de la maquette (Figure 3), la structure canonique est mise en évidence par les courbes rouge et bleu. Lors de la consultation, l'affichage des lignes est synchronisé avec un enregistrement des harpistes. La maquette a été développée en ActionScript 2. Ce travail a donné lieu à un outil plus générique permettant de réaliser des transcriptions canoniques animées.

5.2. Makis Solomos/Iannis Xenakis

Plusieurs articles de Makis Solomos (musicologue, Université de Montpellier-III) sur les œuvres de Iannis Xenakis proposent des exemples musicaux à mi-chemin entre la notation de la partition commentée et une réécriture analytique : il s'agit plutôt d'un changement de repère mettant en valeur certaines propriétés peu visibles dans l'original, et en amoindrissant d'autres. Souvent, il s'agit d'une réécriture sur papier réglé (évoquant le papier millimétré souvent utilisé par le compositeur avant la rédaction définitive de la partition) dans laquelle les notes isolées sont figurées par des

points et les glissandi par des traits – le temps étant en abscisse et les hauteurs en ordonnée.

Nous avons réalisé une maquette en ActionScript 2 permettant de dessiner des traits et des points dans un espace analogue à celui des exemples analytiques de Solomos. Elle permet de recréer ce type d'exemple de réécriture et de les modifier visuellement. En utilisant des données MIDI riches, on pourrait permettre de partir d'une partition en compréhension, de la représenter à l'intérieur de l'espace pré-analytique défini par l'utilisateur et enfin d'écouter le résultat en le modifiant à volonté, explorant ainsi les limites d'identité musicale de la « sonorité xenakienne » analysée.

5.3. Andrea Cera/Boucles de musiques populaires

Andrea Cera est un compositeur italien de musique contemporaine savante qui a travaillé dans le domaine de la musique commerciale. Quand il se trouve face à une singularité rythmique d'une pièce de *popular music*, Cera cherche à la comprendre analytiquement pour aller au-delà de la pure fascination et, éventuellement, se réapproprier cette singularité rythmique dans son activité créatrice. Une rapide transcription lui permet cette inscription de l'écoute sur un support assurant la mise à distance et rendant possible un rapport analytique. Cette transcription sert de base à des recherches assidues de phénomènes remarquables, par addition de couches : ainsi apparaissent des dactyles, ou des "groove", inaudibles *a priori* car seulement accessibles à travers leur déposition sur le papier.

C'est à partir de cette pratique d'écoute (expliquée par Cera et donnant à lieu à des traces observables) que nous avons travaillé. Il s'agissait d'imaginer des modes de représentation et manipulation de cette transcription, permettant au récepteur de cette "écoute signée" de se caler sur l'écoute de Cera, et de comprendre son propos en se familiarisant avec la maquette.

Nous avons réalisé une maquette affichant la transcription de Cera synchronisée à l'aide d'un curseur avec un extrait de quelques secondes, mis en boucle, du remix de "We have arrived" par Aphex Twin dans 'Classics' – *The Aphex Twin*. La représentation permet de dissocier en partie la transcription – qui n'est qu'une image de l'écoute de Cera, et ne correspond pas nécessairement exactement ni aux superpositions de timbre de l'œuvre, ni aux pistes du séquenceur d'origine, etc. – du cadre temporel, dont l'unité minimale est le plus petit dénominateur commun aux couches rythmiques, soit la double-croche dans la transcription d'origine. Le masquage de certaines lignes de la transcription permet d'isoler des figures remarquables et de diriger l'attention auditive en privilégiant une image partielle de l'écoute de Cera. Plusieurs techniques (ralentissement, emphase sonore sur telle ou telle couche rythmique, etc.) permettent de faire varier l'objet ainsi créé afin de faire coïncider ce que l'on entend avec la transcription. D'autres alternatives de visualisation de cette transcription ainsi

que les dérivés de ce projet dans le cadre du projet Semantic Hifi sont développées dans [17].

Les opérations de transcription abordées dans les trois maquettes précédentes permettent de dégager des invariants : le développement d'un référentiel propre au matériau analysé (une tablature pour les harpes Nzakara étudiées par Chemillier, un repère ressemblant à du papier millimétré pour l'analyse de l'œuvre de Xenakis par Solomos, une grille pour que Cera puisse transcrire des boucles rythmiques), des outils d'écriture sur ce nouveau référentiel, et parallèlement, la synchronisation d'un matériau sonore à cette retranscription.

6. METTRE EN TABLEAU

La mise en tableau est l'une des opérations classiques de la musicologie depuis le XIX^e siècle : tables de *leitmotive*, comparaison entre un thème et ses variations, analyse sémiologique de partition (selon l'approche particulière de Ruwet systématisée par Nattiez), sont autant de façons très différentes les unes les autres de procéder à une telle opération.

Nous avons essayé d'instrumenter ce type de pratique d'écoute et d'analyse sur la base de quelques études de cas (passés – en refaisant certaines des opérations décrites dans les articles analytiques de référence – ou actuels – en réalisant de tels tableaux à partir de recherches en cours). Cette épistémologie pratique s'est développée en relation avec un processus de prototypage d'un outil d'aide à la mise en tableau en analyse musicale, basé sur l'outil d'annotation de partition Musique Lab Annotation [23], et réalisé en 2006 par un élève ingénieur, Thomas Bottini. Dans le cadre de cette conception participative, le musicologue Jonathan Goldman a notamment produit une analyse musicale de *Mémoriale* de Pierre Boulez selon les procédures de la méthode sémiologique de Nattiez. Cette première phase de travail a produit un prototype opérationnel en Flash, une publication analytique, et deux publications sur le projet ([2] et [10]) dont l'une rend accessible en ligne le prototype. Ce dernier est au croisement d'un outil d'annotation manuelle, de découpage automatique et d'accès déterminé à l'ordre de lecture des segments ainsi définis : il permet d'écouter un tableau.

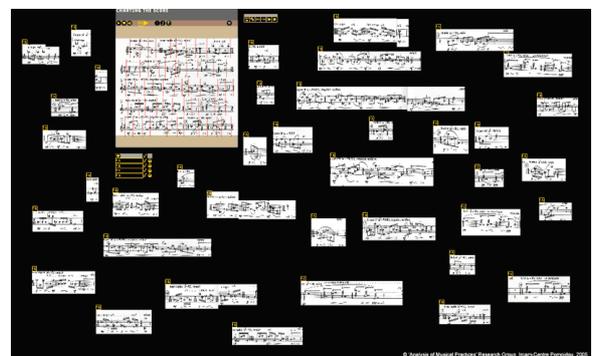


Figure 4. Disposition des segments de partition dans l'espace de travail

La partition est initialement segmentée et synchronisée avec un fichier son. L'outil permet à l'utilisateur de découper des extraits dans la partition et de les déplacer dans l'espace de travail (le fond noir sur la figure 4). Ces segments peuvent être étiquetés à l'aide de mots-clés.

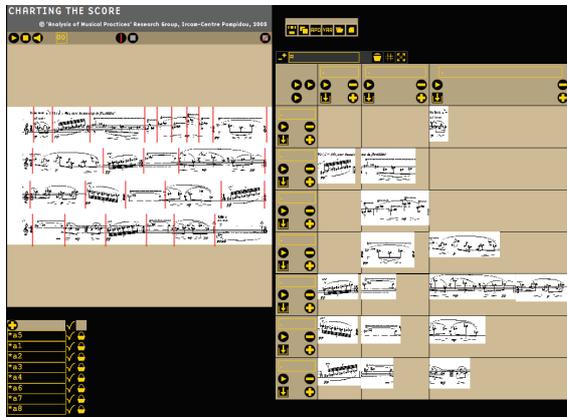


Figure 5. Positionnement des segments dans les cellules d'un tableau

Sur la figure 5, les segments auparavant placés dans l'espace de travail ont été déplacés dans les cases d'un tableau (à droite sur l'image). Le déplacement peut se faire selon différentes procédures :

- algorithmiques, en remplissant automatiquement le tableau à travers une analyse paradigmatique ; tous les segments de la partitions seront mis dans le tableau ;
- automatique, en fonction de l'étiquette que l'on a donné aux segments ; le système informatique les place dans la colonne correspondante à l'étiquette ;
- manuelle, en déplacement manuellement les segments dans les cases du tableau.

Plusieurs tableaux proposant des lectures concurrentes de la partition segmentée peuvent être créés sur le même espace de travail. On peut par ailleurs écouter des lignes, des colonnes ou bien même un parcours dans les segments du tableau.

Cette maquette de mise en tableau développée en ActionScript 2 outille numériquement une pratique existante de la musicologie pour des usages variés. En prenant appui sur l'informatique, la maquette permet des actions impossibles avec le papier :

- l'écoute des segments *via* la synchronisation de la partition avec le fichier son ;
- la simplicité de la manipulation qui encourage les variations de tableaux ;
- la possibilité d'utiliser plusieurs fois un même segment.

7. RECONSTRUIRE UN ASPECT DE L'ŒUVRE

Parmi les pratiques et situations d'écoute étudiées dans le projet "Écoutes signées", un cas particulier est apparu lors de la collaboration de longue durée de Philippe Leroux avec l'équipe Analyse des Pratiques Musicales : peut-on rendre compte de la façon dont un

compositeur écoute sa propre musique ? À défaut de pouvoir répondre à la question dans sa trompeuse généralité, des études de cas pouvaient fournir des éléments de réponse tangibles et situés. En parallèle avec l'étude de la composition de son œuvre *Voi(rex)*, des maquettes d'écoute informée ont été réalisées en suivant deux lignes directrices :

- utiliser lorsque c'était possible du matériel de la genèse de l'œuvre (fichiers sons travaillés, indications sur les brouillons, etc.), situé dans la temporalité de la création de l'œuvre, c'est-à-dire lié à une écoute se projetant dans l'œuvre à venir ;
- explorer pour chaque mouvement une idée directrice, en s'inspirant des notions compositionnelles (par ex. : spatialisation, travail sur les formes d'ondes, rotations harmoniques, etc.) par lesquelles le compositeur caractérise chacun des mouvements de son œuvre, dès l'élaboration de ses brouillons, puis pendant son écriture, et enfin après l'achèvement de l'œuvre et sa création (par exemple lors de conférences, où ces mêmes catégories sont désormais mobilisées comme des guides d'écoute autorisés).

Les maquettes décrites ici sont principalement issues de cette deuxième ligne directrice (la première est traitée dans la partie suivante) et ont été publiées dans le deuxième numéro de *L'Inouï, revue de l'Ircam* [14].

7.1. 4^e mouvement ("Devant tout autour")

La maquette du 4^e mouvement de *Voi(rex)* associe des images animées de certaines esquisses compositionnelles de ce mouvement, avec son enregistrement sonore. Le mouvement ayant été pensé par blocs successifs reliés par des conduits dès les premiers plans, l'animation exploite le décalage partiel entre la structure prévue (matérialisée par la mise en séquence de deux des brouillons disponibles, l'un de l'ensemble du mouvement, l'autre du bloc central) et la structure réalisée (durées effectives des divers blocs et conduits et de leurs parties, élimination de certains d'entre eux).

Une représentation synthétique sous forme de barre de navigation temporelle, reprenant les différents blocs, permet de naviguer directement dans le fichier son (en haut sur la figure 6).

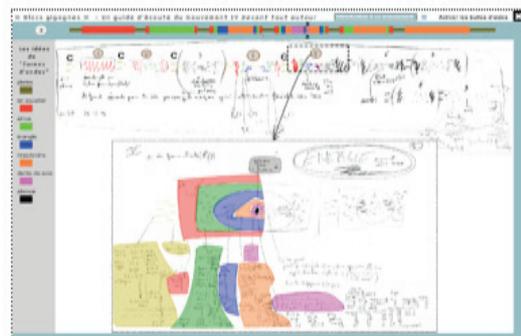


Figure 6. Maquette du 4^e mouvement

La maquette constitue un guide d'écoute généralement jugé efficace pour se repérer dans la forme symétrique « en blocs gigognes » choisie par le compositeur et pour entendre le « silence central » de cette forme – silence qui, par sa brièveté et par le fait qu'il ne se trouve pas, chronométriquement, au milieu du mouvement (tel qu'il a finalement été écrit), est longtemps passé inaperçu à l'oreille des chercheurs.

Cette maquette a été développée dans l'outil Flash MX et aucun élément n'est chargé dynamiquement : les scans des brouillons du compositeur, le fichier son du 4^e mouvement et les événements sur la *timeline* sont intégrés dans le fichier Flash. Dans le cadre du projet Semantic HiFi, une maquette en Flash AS2 de visualisation d'enregistrement sonores segmentés, identique fonctionnellement, avait été générée automatiquement à partir de fichiers XML issus d'algorithmes d'extraction de structure de l'équipe Analyse/Synthèse de l'Ircam [3].

7.2. 2^e mouvement (“Jusque”)

La maquette du 2^e mouvement présente la partition du mouvement (scans du manuscrit) synchronisée avec son enregistrement sonore. Cette partition s'inspire de la calligraphie des lettres de l'alphabet (en l'occurrence les lettres du texte poétique mis en musique). Aussi l'utilisateur peut-il d'une part, lire et écouter ce mouvement en visualisant chacune de ces lettres apparaissant en surimpression sur le passage correspondant, et d'autre part, les effacer pour y substituer sa propre calligraphie et en écouter la version piano MIDI correspondante. Il peut aussi réordonner les lettres du poème, recomposant en ce sens la partition. Des *presets* permettent d'ordonner ces sections en fonction de six catégories musicales déterminées par le compositeur, ou bien de critères liés à l'ordre des lettres.

Le développement informatique en Flash MX est plus conséquent que sur la maquette précédente, l'utilisateur pouvant interagir davantage. Mais l'opération générale est fonctionnellement identique : des métadonnées (extrait de la partition et forme de la lettre dessinée) sont associés à des segments d'un fichier son.

7.3. 3^e mouvement (“De part [...] En part”)

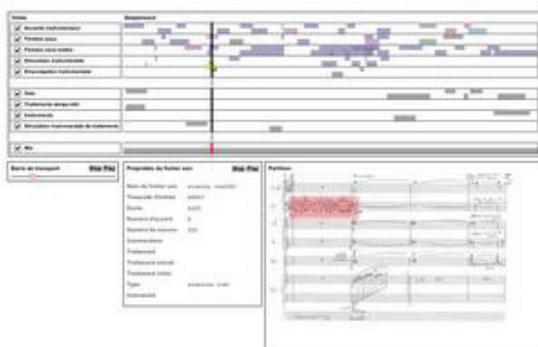


Figure 7. La maquette du 3e mouvement

La maquette du 3^e mouvement (voir Figure 7) reprend plusieurs principes visuels et fonctionnels d'un séquenceur audionumérique. Les différentes pistes du séquenceur correspondent à autant de catégories de sons établies au cours de la composition de l'œuvre : fichier sons, fichiers sons traités, sons joués par les instruments sur scène, simulation instrumentale, voies de mixage, etc. Cette structuration rend possible l'application à chaque fichier son des traitements sonores (*Time Stretching*, *Frequency shifting*, *Doppler*) que Leroux avait lui-même appliqués à cette catégorie de son au cours de la composition.

Cette maquette a nécessité un développement informatique complexe du fait de la communication entre le FlashPlayer et Max/MSP. Du point de vue la généralisation du concept fonctionnel dégagé à partir de la maquette du 4^e mouvement, il s'agit ici du même modèle de segmentation de fichiers sons, mais ici, plusieurs segmentations (les pistes du séquenceur) sont superposées et des actions différentes (les transformations sonores) peuvent être appliquées aux segments de chaque pistes.

Les trois maquettes précédentes (discutées en détail dans [9]) permettent à l'écouteur de sélectionner/écouter (maquette du 4^e mouvement) ou, aussi, de tester des variations dans un champ de possibles restreint aux opérations effectuées par le compositeur (maquettes des 2^e et 3^e mouvement).

8. STRUCTURER UNE DOCUMENTATION

Ce projet est né du besoin et de la possibilité actuelle, en critique génétique de l'œuvre musicale et en analyse de l'activité de composition, d'un outil permettant d'articuler temporellement partition, enregistrement sonore, traces documentant le travail de composition, et commentaires sur ces traces.

Un tel dispositif peut servir à la fois à la réalisation de l'analyse, à la soumission de cette analyse à la contestation scientifique, et finalement à la dissémination (dans un contexte plus large que celui des communautés scientifiques concernées, puisqu'incluant aussi des étudiants en composition, des mélomanes, etc.). Antérieurement à ce projet, une publication sur CD-Rom avait proposé un parcours dans des matériaux variés (esquisses, sons, interviews, etc.) autour d'*Avis de tempête* de George Aperghis [13].

Des maquettes ont été réalisées en vue d'établir l'architecture nécessaire à une telle navigation génétique dans l'ensemble de *Voi(rex)* de Philippe Leroux, articulant la partition et le son de cette dernière avec une base de données constituée de l'ensemble des esquisses, brouillons, fichiers sons, patches, etc., ayant servi à la composition, ainsi que les extraits d'entretien de remise en situation de composition de *Voi(rex)* (cf. <http://apm.ircam.fr/leroux/>). Des tests sur de courts extraits du premier et du deuxième mouvement ont permis de déterminer les opérations de navigation élémentaires et leur représentation à l'écran.

Dans le prototype réalisé par Florine Chamey-Davignon et S. Goldszmidt en 2005, lorsqu'un passage de la partition a fait l'objet d'une reconstitution d'activité pendant le projet de recherche empirique sur l'activité de composition de *Voi(rex)* et qu'ainsi des données d'entretien de Philippe Leroux sont disponibles, des annotations automatiques sur la partition permettent d'accéder à ces données. De plus, lorsque des fragments des esquisses et brouillons ont pu être mis en relation, d'une part avec ces données d'entretien, d'autre part avec des mesures et lignes de la partition, ces fragments sont isolés par un autre mode d'annotation et l'on peut y accéder pareillement à partir de la partition. Enfin, des liens permettent de naviguer entre l'ensemble de ces éléments.

Deux modes de navigation sont proposés :

- un mode d'écoute centré sur la lecture de la partition mais signalant les enrichissements possibles (le lecteur focalise son attention sur la partition) :



Figure 8. Mode d'écoute centré sur la partition

- un mode documentaire permettant l'accès à ces enrichissements à partir d'une vue plus grossière de la partition, ce qui permet de donner accès à des éléments, ici catégorisés (entretiens, esquisses locales, globales...), utilisés par Leroux à ce moment de l'écriture de la partition :

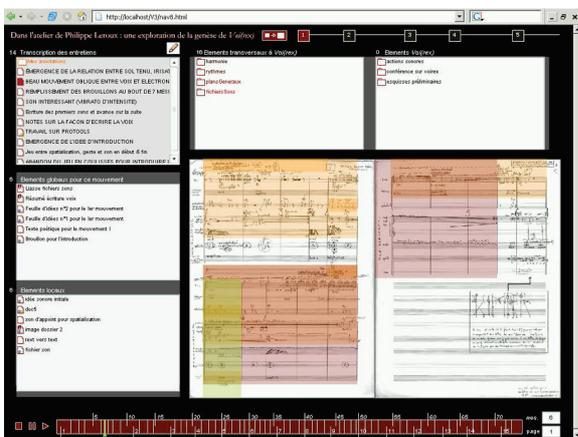


Figure 9. Mode documentaire

La maquette présentée ici a été développée dans une architecture client-serveur. PHP est utilisé côté serveur, et Flash AS2 pour l'interface cliente riche. Les données sont enregistrées dans une base de données MySQL. Flash est utilisé pour faciliter les échanges asynchrones avec le serveur, la solution AJAX n'étant alors pas encore démocratisée, donc difficile à implémenter pour des maquettes informatiques. Une version consolidée de cet outil de navigation est présentée dans [16].

9. L'ANNOTATION : UNE OPÉRATION TRANSVERSALE

Toutes les maquettes précédemment présentées ont en commun une caractéristique générale : elles incluent des outils d'annotation du corpus musical traité. Le geste d'annotation est en effet essentiel à toute « musico-logie » en tant que discours se rapportant à des moments ou morceaux de musique. La notion d'« écoute signée » impliquait, certes, d'éliminer le plus possible le commentaire verbal pour donner à voir / entendre / manipuler directement la musique ; mais l'annotation n'est pas forcément verbale : elle peut être aussi graphique (que ce soit la schématisation visuelle d'informations musicales représentées, ou un commentaire noté de façon conventionnelle, par exemple solfégique). Au fur et à mesure de la conception des diverses maquettes sont apparues plus clairement les spécificités de l'annotation de partition.

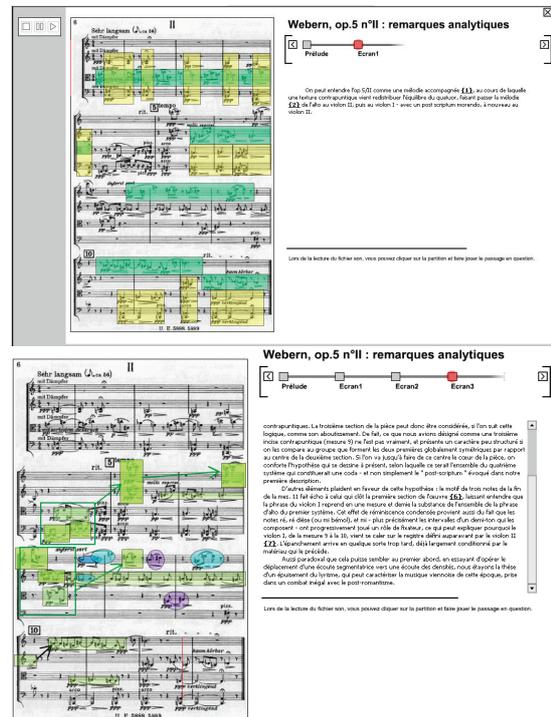


Figure 10. Etapes du parcours analytique dans une œuvre pour quatuor d'Anton Webern

Une première maquette d'outil d'annotation temporelle de partition a été réalisée en 2003 dans le cadre d'un projet européen (MusicWeb, voir <http://musicweb.koncon.nl/>). Il s'agissait d'imaginer un

cours d'analyse musicale en ligne exploitant les possibilités immédiates de la synchronisation manuelle entre un fichier son et la partition correspondante, à partir de laquelle il soit possible de dessiner sur le texte musical des annotations analytiques (graphiques ou verbales) et de les faire apparaître puis disparaître à des moments déterminés au cours de l'écoute-lecture. La maquette proposait un parcours progressif à travers plusieurs guides de lecture et d'écoute de l'un des mouvements pour quatuor op. 5 de Webern (*Fünf Sätze für Streichquartett*, op. 5, n° II, 1909). On part des lacunes d'une écoute un peu superficielle selon laquelle la pièce serait une mélodie accompagnée, donc une mélodie 'horizontale' ponctuée d'accords 'verticaux' (Figure 10a) – mais cela ne permet pas de rendre compte d'une partie des mesures 6-7, 9 et 11-12. On entre ensuite dans une écoute plus attentive au détail, jusqu'à une écoute focalisée sur les hauteurs (relations motiviques et prépondérances intervalliques) introduites par de nouvelles annotations. Enfin, une remise en cause de cette deuxième écoute, passant par la considération du grand nombre d'indications dynamiques dans la page (image au dessus), mène à un guide d'écoute et de lecture moins univoque, plus idiosyncrasique – en fait plus représentative de l'écoute personnelle de l'analyste que les précédentes annotations (Figure 10b).

Au terme du parcours, l'ensemble des annotations est proposé à la consultation et à la manipulation du lecteur, qui accède également aux fonctions d'annotation utilisées par l'auteur pour réaliser son analyse. L'annotation de partition synchronisée au son apparaît ici comme un moyen privilégié non seulement pour guider l'auditeur, mais pour lui transmettre des moyens de s'approprier à son tour l'objet de l'écoute.

Cette maquette a conduit à la réalisation d'un prototype d'Outil d'annotation et de synchronisation image/son (OASIS), réalisé en Flash. Cet outil, redéveloppé en C# et enrichi de nouvelles possibilités, est devenu en 2006 un élément de la suite logicielle MusiqueLab Annotation destinée à l'Education Nationale et aux Conservatoires [20]. Il a aussi inspiré le développement d'iAnalyse par Pierre Couprie [5]. Nous avons par ailleurs réalisé une réduction de cette analyse pour écran de petite taille, plus propice à l'écoute (Figure 11) utilisant un stylet comme moyen de navigation.



Figure 11. Parcours analytique sur Pocket PC

Mais la notion d'annotation ne caractérise pas seulement un geste essentiel de pratiques d'écoute spécialisées telles que l'analyse musicale ; elle est largement transversale. D'une part, elle est présente dans de nombreuses pratiques musicales dans lesquelles l'écoute est orientée vers la production sonore, comme typiquement l'interprétation des musiques écrites [12], ce qui appelle de nouveaux développements. D'autre part, elle a une signification plus générale au sein de l'ingénierie hypermédia mise en œuvre ici : en vue d'une généralisation des maquettes développées, on peut formaliser dans un premier temps l'opération d'annotation comme indépendante des caractéristiques du matériau annoté (partition, extrait sonore, entretien filmé). Informatiquement, annoter signifie alors pointer tout ou partie d'un document numérique et le relier à tout ou partie d'un autre document numérique. L'étape suivante consiste à instancier ces pointeurs en fonction des caractéristiques des matériaux annotés ; pour un fichier son, il s'agit par exemple de segments audio (4^e mouvement de *Voi(rex)*), pour une partition, de zones dessinées sur l'image (guide d'écoute de Webern), etc. Une fois ces annotations, au sens large, effectuées, il s'agit de créer une vue informatique adéquate du matériau annoté (cf. [11]), comme c'est le cas pour la mise en tableau.

10. APPORTS ET PERSPECTIVES

Les apports du projet "Écoutes signées" sont divers.

Chaque maquette a produit des résultats, plus ou moins significatifs selon les cas, d'une part quant à la connaissance actuelle des pratiques d'écoute (récemment [25] à propos de l'écoute en U.S.T.), et d'autre part quant au développement d'interfaces hypermédia adaptées à la représentation et à la transmission de phénomènes musicaux en jeu dans l'écoute.

L'effort d'abstraction consistant à ramener certaines maquettes à des gestes ou opérations analytiques, en relation avec l'identification des formes anciennes de ces pratiques d'écoute, a permis d'esquisser les fonctions d'un environnement, ou d'une boîte à outils, dédié à l'écoute dite 'active'. Les questions afférentes à l'intégration de ces différentes briques et des types de représentations pertinentes des données considérées (par ex. la visualisation en séquenceur à différentes échelles) n'ont cependant pas été abordées.

Au plan informatique, une fois les corpus variés numérisés (partitions, esquisses, sons ...) dans un système d'information intégrant les possibilités d'annotation génériques décrites ci-dessus, il faut développer les interfaces utilisateurs d'annotation et de consultation, pour la plupart issues des maquettes décrites ici. Un tel système permettrait aussi de varier les vues disponibles pour un corpus ; de changer de point de vue, de consulter les matériaux annotés, etc.

Ce projet a enfin constitué un banc d'essai suffisamment varié pour le programme de recherche technologique ambitieux présenté plus haut. En termes

de conception, la variété de nos processus de maquettage constitue un apport aux méthodes de design participatif en IHM et, plus généralement, en ergonomie et en Computer-Supported Cooperative Work (voir [17]). En termes d'hypothèses empiriques sur l'écoute musicale, le projet a été l'occasion d'amorcer une étude scientifique de ces dernières, donc de développer en cette matière une relation organique (c'est à dire dans les deux sens) entre recherche technologique et recherche empirique, par : (1) une ethnologie informelle des pratiques d'écoute (associée à une intervention technologique sur un aspect de ces pratiques); (2) une analyse rigoureuse d'une pratique musicale (la composition) incluant l'écoute ; (3) une provocation technique par des situations d'écoute nouvelles, contribuant à la connaissance des facteurs situationnels-techniques de l'écoute.

11. RÉFÉRENCES

- [1] Aigrain, P. Lepain, P. « Le groupe Ecoute Interactive de la Musique de la Bibliothèque Nationale de France », Journées d'Informatique Musicale (JIM'96), Caen, 1996.
- [2] Bottini, T. *et al.* « Modèle et outils documentaires multimédias pour la mise en tableau de partitions », Document Numérique, 11/3-4, 2008
- [3] Boutard, G. Goldszmidt, S. Peeters, G. « Browsing inside a Music Track, the Experimentation Case Study », SAMT, Athènes, 2006.
- [4] Campos, R. Donin, N. *L'Analyse musicale, une pratique et son histoire.* Droz, Genève, 2009.
- [5] Couprie P. « iAnalyse : un logiciel d'aide à l'analyse musicale », Journées d'Informatique Musicale (JIM'08), Albi, 2008.
- [6] Delalande, F. « Music Analysis and Reception Behaviours: Sommeil by Pierre Henry », Journal of New Music Research, 27, 1-2, 1998.
- [7] Donin, N. « Towards Organised Listening: Some Aspects of the 'Signed Listening' Project », Organised Sound, 9, 1, 2004.
- [8] Donin, N. « Problèmes d'analyse de l'interprétation. Un essai de comparaison assistée par ordinateur d'enregistrements du premier prélude du *Clavier bien tempéré* », Musurgia, XII, 4, 2005.
- [9] Donin, N. « Pour une "écoute informée" de la musique contemporaine : quelques travaux récents », Circuit, 16, 3, 2006
- [10] Donin, N. Goldman, J. « Charting the Score in a Multimedia Context: the Case of Paradigmatic Analysis », Music Theory Online, 14, 4, 2008, <http://mto.societymusictheory.org>
- [11] Donin, N. Goldszmidt, S. « Annoter la musique : de la segmentation de fichiers audio à la publication d'articles multimédia », Annexes des actes d'IHM 2007, Paris, 2007
- [12] Donin, N. Theureau J. « Annotation de la partition par le musicien et (re)distribution de son attention en situation de répétition », Salembier & Zacklad, *Annotation dans les documents pour l'action*, Hermes publishing, Londres, 2007
- [13] Donin, N. Goldszmidt, S. « *Autour d'Avis de Tempête, Un guide d'écoute multimédia de l'Opéra de Georges Aperghis* », supplément DVD-Rom joint à *L'Inoui – Revue de l'Ircam*, 1, 2005.
- [14] Donin, N. Goldszmidt, S. Theureau J. *De Voi(rex) à Apocalypsis, fragments d'une genèse. Exploration multimédia du travail de composition de Philippe Leroux.* DVD-Rom joint à *L'Inoui, revue de l'Ircam*, 2, 2006.
- [15] Geslin, Y. Lefevre, A. « Sound and musical representation : the acousmographe software », Proceedings of the International Computer Music Conference, Miami, USA, 2004.
- [16] Goldszmidt, S. Donin, N. Theureau J. « Navigation génétique dans une œuvre musicale », Proceedings of IHM 2007, Paris
- [17] Goldszmidt, S. Theureau, J. « Conception de situations d'assistance à l'écoute musicale active et analyse de l'activité de composition musicale », dans G. Vallery, M-C Leport et M. Zouinar, *Ergonomie des produits et des services médiatisés*, PUF, Paris, à paraître
- [18] Hennion, A., *et al.* *Figures de l'amateur. Formes, objets, pratiques de l'amour de la musique aujourd'hui.* La Documentation Française, Paris, 2000.
- [19] Maisonneuve, S. *L'Invention du disque, 1877-1949. Genèse de l'usage des médias musicaux contemporains.* Editions des Archives Contemporaines, Paris, 2009.
- [20] Puig, V. *et al.* « Musique Lab 2 : A three level approach for music education at school », Conference Proceedings [of the] International Computer Music Conference (ICMC), 2005.
- [21] Sterne, J. *The Audible Past: Cultural Origins of Sound Reproduction.* Duke University Press, 2003
- [22] Szendy, P. *Ecoute. Une histoire de nos oreilles.* Minuit, Paris, 2001.
- [23] Site internet de l'équipe Analyse des pratiques musicales : <http://apm.ircam.fr/>
- [24] *Musimédiane*, numéro 3 : Musiques non écrites, 2008, <http://musimediane.com/>
- [25] *Musimédiane*, numéro 5 : les U.S.T., 2010, <http://musimediane.com/>
- [26] Chemillier, M. « Pour une écriture multimédia de l'ethnomusicologie », Cahiers de musiques traditionnelles, 16, 2003

Session 7

Analyse

UN SYSTÈME DE DÉTECTION DE RUPTURE DE TIMBRE POUR LA DESCRIPTION DE LA STRUCTURE DES MORCEAUX DE MUSIQUE

Gabriel SARGENT
Université de Rennes 1,
IRISA (- UMR 6074)
gabriel.sargent@irisa.fr

Frédéric BIMBOT
CNRS,
IRISA (- UMR 6074)
frederic.bimbob@irisa.fr

Emmanuel VINCENT
INRIA,
Centre INRIA Rennes
- Bretagne Atlantique
emmanuel.vincent@inria.fr

RÉSUMÉ

Cet article présente une chaîne d'inférence automatique d'éléments de la structure musicale de morceaux de musique sous leur forme sonore. Le système, inspiré de l'état de l'art, est constitué de deux étapes : la segmentation du morceau et l'étiquetage des segments trouvés. Deux contributions sont proposées : un critère de détection des ruptures du timbre des morceaux, initialement proposé pour la séparation entre parole et musique dans un flux audio, est introduit lors de la segmentation. On propose par ailleurs de réaliser l'étiquetage en effectuant un regroupement hiérarchique des segments selon leur timbre modélisé par une gaussienne et en ayant recours à des métriques empruntées au traitement de la parole. Cette étude se termine sur les premiers résultats d'évaluation de cette chaîne en utilisant un corpus de 20 morceaux de musique.

1. INTRODUCTION

1.1. Contexte

Les techniques actuelles de compression sonore facilitent la diffusion de la musique (lecteurs mp3, services Internet d'écoute de musique en ligne...), ce qui conduit à la constitution de bases de morceaux de plus en plus volumineuses. Une indexation efficace est nécessaire afin de conserver un accès facile et rapide à ces « données ». Ce besoin a conduit au développement du domaine de la recherche d'information dans les morceaux de musique, qui a pour but la description et l'organisation automatique des contenus musicaux.

L'analyse de la structure des morceaux en terme de couplets, refrains, ponts... entre dans le cadre de la description haut niveau du contenu musical et possède plusieurs applications intéressantes. L'identification des parties répétées, qui sont en général les plus caractéristiques des chansons, permet la génération de résumés sonores. En production musicale, le travail de l'ingénieur du son est susceptible d'être facilité par un accès précis et rapide aux parties musicales à mixer. Enfin, dans le cadre de la musicologie, il peut être utile de formuler des caractéristiques relatives à un style musical par l'analyse automatique d'un grand nombre de chansons associées à ce style (analyse à grande échelle).

Le travail présenté dans cet article correspond au développement d'un système de référence réalisé dans le cadre d'un mémoire de Master 2 (spécialité Systèmes Intelligents et Communicants).

1.2. Définition de la tâche de structuration

Du point de vue des Sciences de l'Ingénieur, on peut décrire un morceau de musique comme étant l'émission d'un ensemble d'événements sonores agencés de manière cohérente dans le temps. La musique tonale occidentale, qui constitue le matériau expérimental dans ce travail, peut être décrite suivant différentes « strates musicales » : la tonalité, le tempo, le timbre, le rythme, l'harmonie, la mélodie et les paroles. On fait l'hypothèse qu'il existe une description de la structure d'un morceau de musique rendant compte de groupements logiques sur les différentes strates décrivant les morceaux : couplets, refrains, ponts ou tout autre découpage s'appuyant sur les similarités et répétitions à moyen terme dans le morceau.

Le but du système présenté dans cet article est de rechercher ces éléments structurels sur la strate du timbre du morceau. On désignera par « segmentation timbrale » l'opération consistant à déterminer les parties homogènes du point de vue du timbre. Celles-ci permettent la description partielle de la structure d'un morceau à plus ou moins long terme. Une description à court terme considère les notes jouées et l'on peut passer à des groupes hiérarchiques de notes plus conséquents en regroupant ces notes par motifs cohérents (mélodies, séquences harmoniques) ou selon les mesures musicales.

Lorsque l'on s'intéresse à une échelle à moyen terme, la notion de structure musicale devient incertaine, même si certaines parties sont bien identifiées. La définition précise met généralement en jeu les différentes strates. La structure des morceaux n'étant pas explicitement communiquée par les compositeurs (si tant est qu'ils l'aient consciemment fixée), il est possible de rencontrer des descriptions différentes du point de vue de la taille des segments, de leurs bornes ou encore de leur étiquette (voir à ce sujet le travail de Peiszer *et al.* dans [19]). Ceci est dû au choix des strates musicales supposées pertinentes pour un morceau donné, qui peut varier d'un auditeur à l'autre. Il n'y a ainsi pas de réel consensus sur ce sujet et l'on a

choisi ici de s'intéresser prioritairement au timbre en remarquant que ses ruptures coïncidaient souvent avec les frontières des segments structurels (couplets, refrains...).

La partie 2 de cet article décrit brièvement les différentes méthodes de détection de la structure dans l'état de l'art, la partie 3 traite de la chaîne de structuration qui a été mise en place, constituée de deux sous-tâches : la segmentation et l'étiquetage. La partie 4 présente les premiers résultats obtenus sur le corpus de 20 morceaux : une moitié est utilisée pour l'apprentissage de ses paramètres et l'autre moitié pour le test, puis les deux moitiés sont interverties. Enfin, un premier diagnostic est établi suite à l'évaluation de différentes versions du système proposé.

2. PANORAMA DE L'ÉTAT DE L'ART

Les tâches de segmentation et d'étiquetage, réalisées successivement ou conjointement selon les méthodes, s'effectuent par l'intermédiaire de descripteurs extraits du signal musical à intervalles réguliers. Les plus populaires sont les coefficients cepstraux à l'échelle Mel et les vecteurs de chroma [4].

2.1. Descripteurs

Les coefficients cepstraux à l'échelle Mel, que l'on notera ci-après MFCC (pour *Mel Frequency Cepstral Coefficients*), sont des descripteurs de timbre et sont interprétables en terme d'instrumentation. Les 20 premiers coefficients sont en général retenus car ils contiennent l'essentiel de l'énergie du spectre du signal étudié : les vecteurs MFCC sont donc de taille 20 [4]. L'échelle Mel reflète certaines caractéristiques perceptives de l'oreille humaine.

Le vecteur de chroma est un descripteur basé sur le tempérament égal de la musique occidentale. Pour l'obtenir, le spectre du signal obtenu à un instant donné est découpé en intervalles fréquentiels de la largeur d'un demi-ton, puis ramené à une seule octave (constituée de 12 demi-tons). Le vecteur obtenu est ainsi de taille 12, chaque coefficient correspondant à l'amplitude d'un demi-ton, toutes octaves confondues - pour plus de détails, voir [2, 7].

Pour procéder à l'extraction de ces descripteurs, le signal est préalablement découpé en trames régulières de quelques dizaines de millisecondes [6]. Jehan a montré dans [10] qu'étendre les bornes des trames aux pulsations musicales (une échelle de description naturelle des morceaux) pouvait améliorer la précision des frontières détectées.

2.2. Segmentation et étiquetage

Deux approches dominantes se sont développées pour l'extraction automatique de la structure haut niveau : celle ayant recours aux Modèles de Markov à états Cachés (MMC) et celle utilisant les matrices de similarité.

2.2.1. Modèles de Markov à états Cachés

Le Modèle de Markov à états Cachés est un modèle statistique permettant la reconnaissance dynamique et robuste de motifs spécifiques (suite de phonèmes formant des mots et portions de phrases, mélodies, suite d'accords...). Une suite d'observations est supposée être issue d'un modèle constitué de p états inaccessibles, dit cachés. Ce modèle suit un processus markovien tel que la probabilité d'être dans un certain état, à un instant donné, ne dépend que de l'état précédent. Dans notre contexte, les observations sont les descripteurs extraits du signal musical et les p états cachés représentent les étiquettes des segments structurels recherchés. Le problème de structuration revient à estimer le modèle par apprentissage de ses paramètres sur un ensemble de chansons (« étape d'apprentissage »), puis à estimer quelle est la suite d'étiquettes pouvant le mieux représenter les descripteurs (« étape de décodage », réalisée par un algorithme de Viterbi). Pour plus de détails sur les MMC, se référer à [21].

Logan et Chu utilisent une description timbrale (MFCC) du signal musical, découpé en trames régulières de quelques dizaines de millisecondes [14]. Ils proposent une étude comparant les performances d'un algorithme de regroupement hiérarchique de trames voisines de timbre homogène, et un MMC constitué d'un nombre d'états comparable au nombre de segments structurels recherchés. Abdallah *et al.* utilisent les MMC afin d'associer un premier état - une première étiquette - à chacune des trames du signal [1]. Les auteurs décrivent ensuite le signal par des histogrammes en effectuant des groupes de 15 trames voisines. Un *clustering* des histogrammes est effectué via des mesures de distances empiriques (telles la distance cosinus ou la version symétrisée de la divergence de Kullback-Leibler), ou par une approche de type « maximum de vraisemblance ». Rhodes *et al.* introduisent un terme de contrôle de la taille des segments trouvés dans la précédente fonction de coût [22]. Levy *et al.* proposent de leur côté d'effectuer un *clustering* par l'algorithme des K-Moyennes floues [13].

2.2.2. Matrices de similarité

Les matrices de similarité offrent une représentation visuelle de la comparaison des descripteurs de toutes les trames du signal entre elles. Un exemple est donné à la figure 1. C'est Foote [6] qui en proposa initialement l'utilisation pour la tâche de structuration en utilisant une description du timbre des morceaux : les segments structurels à identifier sont caractérisés par des zones rectangulaires d'une texture spécifique, lorsque l'on visualise la matrice. La position des bords de ces zones permet de localiser les instants de séparation entre les parties recherchées. Ceci est facilité par le calcul d'une fonction de « nouveauté timbrale » obtenue par corrélation de la matrice avec un noyau en damier, dont les pics correspondent à des ruptures de timbre. Foote et Cooper proposent aussi d'obtenir les segments structurels par *clustering* spectral sur la

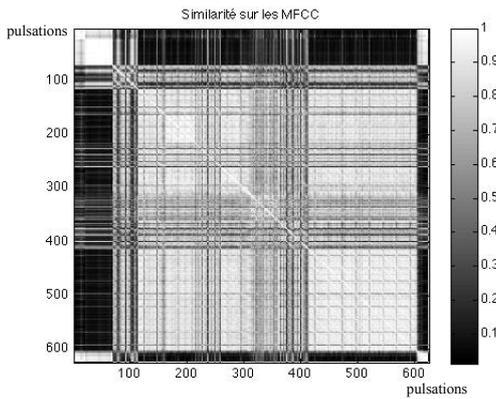


Figure 1. Matrice de similarité calculée sur les MFCC extraits de la chanson « Caribbean Blue » de Enya.

matrice de similarité, puis de modéliser les segments obtenus par des modèles gaussiens comparés par la divergence de Kullback-Leibler [5]. Une autre utilisation possible de la matrice passe par l'utilisation des vecteurs de chroma : il s'agit alors de détecter les séquences de similarité élevées entre trames, visualisables sur la matrice par des bandes sombres, sur ses sous-diagonales. Ces séquences correspondent à la répétition de motifs harmoniques joués lors du morceau analysé. Plusieurs méthodes sont utilisées afin de détecter ces bandes. Lu procède par opérations morphologiques 2D [15]. Goto utilise une matrice temps-décalage dans le cadre précis de la détection de refrains [7]. Shiu propose un noyau de corrélation pour étudier les séquences harmoniques à court terme [24].

2.2.3. Approches mixtes

D'autres méthodes empruntent aux deux approches, comme Peeters [18], qui s'appuie sur la similarité timbrale entre trames voisines pour effectuer une première segmentation du signal, puis utilise les K -Moyennes et l'apprentissage d'un MMC décodé par l'algorithme de Viterbi afin d'obtenir la meilleure séquence de segments étiquetés décrivant les morceaux. Jensen représente l'ensemble des différentes hypothèses de segmentation possibles par un graphe acyclique orienté et ramène le problème à la recherche du chemin de moindre coût le traversant. Il effectue cette étude sur 3 descripteurs différents (rythme, timbre, harmonie) et remarque que la segmentation issue des descripteurs de timbre est la plus proche de sa segmentation de référence [11]. Paulus et Klapuri, dans [16], segmentent le signal musical par l'utilisation de la fonction de « nouveauté timbrale » de Foote et recherchent le meilleur étiquetage des segments par une fonction de coût utilisant leur description harmonique.

2.2.4. Récapitulation

Des diverses méthodes proposées dans la littérature, il n'existe pas à notre connaissance d'étude comparative permettant de les départager.

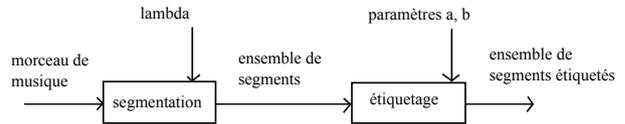


Figure 2. Schéma fonctionnel de la chaîne de structuration. λ est le seuil de détection de rupture timbrale (voir 3.1), a et b sont des paramètres d'ajustement de l'étiquetage (introduits en 3.2.3).

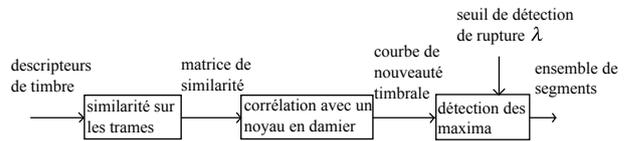


Figure 3. Etape de segmentation

Dans le travail présenté ici, nous nous concentrons sur une approche utilisant des matrices de similarité qui servira de méthode de référence à laquelle nous pourrions comparer par la suite d'autres approches que nous développerons ultérieurement.

3. SYSTÈME DE SEGMENTATION TIMBRALE MIS EN PLACE

L'organisation globale du système de traitement étudié dans ce travail est présenté dans la figure 2. Il est composé d'une étape de segmentation et d'une étape d'étiquetage. Celles-ci font intervenir 3 paramètres à régler, qui peuvent être « appris » sur un corpus d'apprentissage.

Les descripteurs choisis pour l'analyse, les MFCC et vecteurs de chroma, sont exprimés à l'échelle des pulsations des morceaux étudiés.

3.1. Segmentation

La méthode de segmentation utilisée est inspirée de celle de Foote [6] et utilisée sur les 20 premiers coefficients MFCC. La figure 3 rend compte des différentes étapes la composant : similarité sur les trames, puis corrélation avec un noyau en damier et enfin détection des maxima, dont le seuil est déterminé par le paramètre λ . Les trames utilisées sont des fenêtres temporelles centrées sur les pulsations rythmiques. Pour chaque pulsation, la taille de la fenêtre associée est la moitié de la durée entre la pulsation qui la précède et celle qui la suit.

La matrice de similarité, notée A par la suite, est construite en comparant les MFCC des trames du morceau. Les coefficients de A sont obtenus par la formule suivante :

$$[A]_{i,j} = 0.5 + 0.5 \frac{\langle v_i, v_j \rangle}{\|v_i\| \|v_j\|} \quad (1)$$

où v_i et v_j sont les descripteurs associés aux trames i et j , $\langle \cdot, \cdot \rangle$ et $\|\cdot\|$ représentent respectivement le produit scalaire et la norme euclidienne des vecteurs considérés.

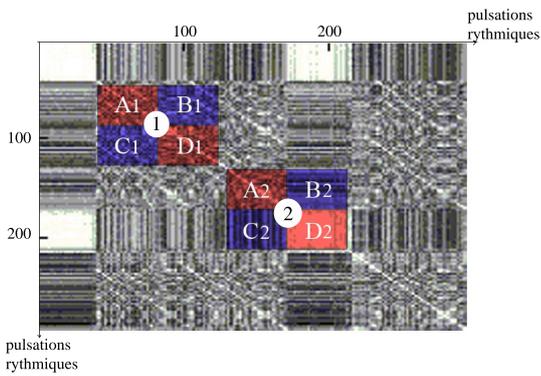


Figure 4. Exemples de position du noyau de corrélation le long de la diagonale de la matrice de similarité (support audio : extrait de Karma Police de Radiohead). La position 1 correspond à une faible valeur de nouveauté (l'instant sur lequel est centré le noyau est contenu dans une zone de texture régulière), la position 2 correspond à un pic de nouveauté, l'instant sur lequel est centré le noyau étant le point de séparation de deux zones de textures différentes.

La courbe de nouveauté timbrale est ensuite calculée par corrélation de la matrice avec un noyau « en damier ». Par exemple un tel noyau, de taille égale à 4, est tel que :

$$K = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix} \quad (2)$$

Ce type de noyau permet de comparer les valeurs des similarités contenues dans les zones passée et future à l'instant considéré (appelées « *similarités propres* », zones A1-2 et D1-2 de la figure 4), avec les zones de similarité comparant les trames passées avec les trames futures (ou « *similarités croisées* », zones B1-2 et C1-2 de la figure 4). L'instant considéré est celui sur lequel le noyau est centré : localisé sur la diagonale, il représente la frontière entre deux trames voisines.

Formellement, si l'on note N la courbe de nouveauté et K le noyau utilisé, on a :

$$N(i) = \sum_{m=-L/2}^{L/2} \sum_{n=-L/2}^{L/2} K(m,n) A(i+m, i+n) \quad (3)$$

Si la zone dans laquelle se trouve un instant t est homogène à l'échelle du noyau, les *similarités propres* et *similarités croisées* sont proches en valeurs. Ainsi la différence des deux types de zones donnera une valeur de nouveauté proche de 0 (exemple à l'instant 1 de la figure), tandis que si l'instant considéré se trouve à la frontière de deux types de textures différentes, on obtient un pic de nouveauté (instant 2 de la figure). On utilise un noyau en damier de forme gaussienne dont l'allure est donnée à la figure 5.

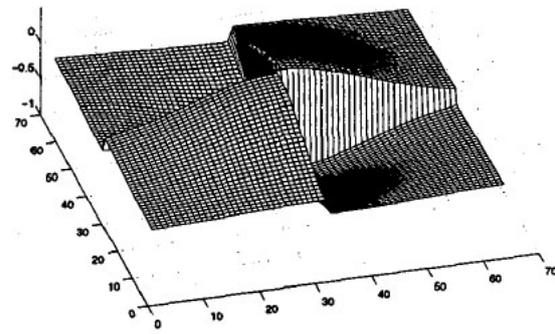


Figure 5. Allure du noyau en damier de forme gaussienne proposé par Foote, extrait de [6].

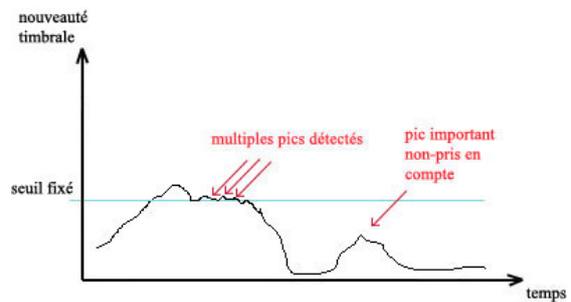


Figure 6. Exemple de détection de maxima par rapport à un seuil fixé.

On est maintenant amené à détecter les maxima de la courbe obtenue, qui représentent les variations importantes du timbre. On les identifie aux frontières des segments structurels recherchés.

Pour détecter les « pics de nouveauté » pertinents, on pourrait se baser sur le seuillage de l'ensemble des maxima locaux détectés [16]. Cette méthode peut cependant conduire à un grand nombre de fausses détections, notamment lorsque la courbe fluctue au voisinage de ce seuil et n'aboutit pas à la détection des maxima les plus pertinents pour la détection de frontières entre segments structurels (voir par exemple la figure 6) : on recherche ici des segments de grande taille. Il est nécessaire d'utiliser un critère qui permet de mettre en valeur les pics de nouveauté qui dominent localement leurs voisins.

Seck *et al.* proposent un tel critère dans [23], basé sur l'analyse d'un indice de présence de rupture statistique. Cet indice, que l'on notera I , mesure la similarité entre les zones passée et future au voisinage de chaque trame du signal (elle est ainsi associable à la courbe de nouveauté timbrale vue précédemment). Le critère de décision C est calculé de manière à prendre la valeur 0 pour tous les instants qui ne sont pas des maxima locaux de l'indice de rupture, et des valeurs plus ou moins élevées suivant l'importance relative du pic par rapport à son contexte.

La méthode de calcul de C est la suivante : on cherche, à partir de chaque instant t , les premiers instants voisins $\tau_1(t) < t$ et $\tau_2(t) > t$ vérifiant $I(\tau_1(t)) > I(t)$ et $I(\tau_2(t)) > I(t)$. On recherche ensuite les quantités :

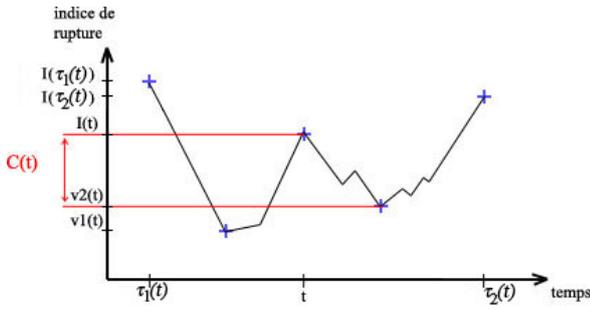


Figure 7. Obtention du critère proposé par Seck *et al.* par l'analyse de l'indice de présence de rupture, associable à une courbe de nouveauté.

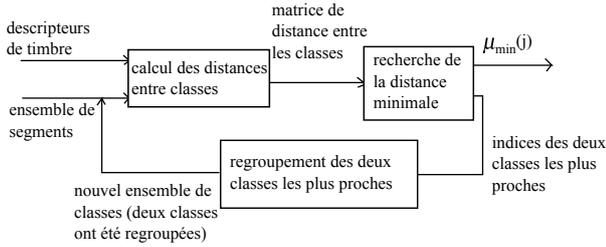


Figure 8. Description du *clustering* hiérarchique ($\mu_{\min}(j)$ est la distance minimale entre deux classes à l'étape j)

$v_1(t) = \min_{\tau_1(t) < i < t} I(i)$ et $v_2(t) = \min_{t < i < \tau_2(t)} I(i)$ afin de calculer :

$$u(t) = \max(v_1(t), v_2(t)) \quad (4)$$

Le critère de présence de rupture est défini par :

$$C(t) = I(t) - u(t) \quad (5)$$

(voir figure 7). Enfin la décision de la présence de rupture à l'instant t est prise par comparaison à un seuil fixe λ : il y a rupture statistique si $C(t) \geq \lambda$. Ce seuil peut être réglé par apprentissage sur un corpus de morceaux de musique annotés. Dans nos expérimentations initiales, l'utilisation de ce critère s'est avéré être un facteur de robustesse.

3.2. Étiquetage

La partie précédente permet d'obtenir un ensemble de segments dont le contenu timbral est supposé homogène. On cherche maintenant à les associer entre eux en comparant leur timbre : deux segments similaires de ce point de vue sont rassemblés dans une même « classe de segments » désignée par un numéro. Ce numéro correspond à leur étiquette. On utilise un *clustering* hiérarchique pour y parvenir ; le processus est décrit à la figure 8.

3.2.1. Clustering hiérarchique

Les regroupements hiérarchiques sont effectués comme suit.

Initialement, chaque segment est caractérisé par une gaussienne et est contenu dans une classe différente. A

chaque itération, on calcule la distance entre toutes les classes deux à deux. Les deux classes de distance minimale sont regroupées : les descripteurs des segments qu'elles contiennent sont regroupés et l'on cherche le modèle gaussien correspondant à cette nouvelle classe. Pour ne pas avoir à calculer les moments à partir des descripteurs regroupés différemment, on utilise les moyennes et matrices de covariances déjà calculées.

Si l'on fusionne deux gaussiennes modélisant les classes S_x et S_y , de nombres d'échantillons de dimension p , de moyennes, et de matrices de covariance respectives (n_x, \bar{x}, X) et (n_y, \bar{y}, Y) , on obtient une gaussienne modélisant une classe S_z ayant les paramètres suivants :

– nombre d'échantillons

$$n_z = n_x + n_y \quad (6)$$

– moyenne :

$$\bar{z} = \frac{n_x \bar{x} + n_y \bar{y}}{n_x + n_y} \quad (7)$$

– matrice de covariance :

$$Z = \frac{n_x (X + \bar{x} \bar{x}^T) + n_y (Y + \bar{y} \bar{y}^T)}{n_x + n_y} - \bar{z} \bar{z}^T \quad (8)$$

Il est ainsi possible de comparer le modèle de cette nouvelle classe avec les classes restantes.

Le procédé est répété jusqu'à satisfaire un certain critère d'arrêt.

3.2.2. Comparaison des segments : modèle gaussien de leur contenu timbral

On fait l'hypothèse que la distribution des MFCC de chaque segment suit une loi gaussienne. Cette approche diffère de celle de [5] au niveau de la métrique de comparaison des modèles correspondants.

La mesure de vraisemblance gaussienne symétrisée utilisée dans cette étude a été introduite dans le cadre de l'identification du locuteur dans des enregistrements numérisés [3]. Elle est définie ci-dessous.

Pour simplifier les prochaines formules, on pose :

$$\delta = \bar{y} - \bar{x}, \Gamma = X^{-\frac{1}{2}} Y X^{-\frac{1}{2}}, \rho = \frac{n_y}{n_x} \quad (9)$$

On considère par la suite que le modèle de S_x est le modèle de référence et celui de S_y est le modèle estimé.

Notons $\{y_t\}_{1 \leq t \leq n_y}$ les descripteurs de S_y . La vraisemblance que y_t ait été généré par le modèle de S_x s'écrit :

$$L(y_t | \bar{x}, X) = \frac{\exp\left(-\frac{1}{2} (y_t - \bar{x})^T X^{-1} (y_t - \bar{x})\right)}{(2\pi)^{\frac{p}{2}} (\det(X))^{\frac{1}{2}}} \quad (10)$$

Sous l'hypothèse d'indépendance *a priori* de ses éléments, on définit la log-vraisemblance moyenne comme suit :

$$\bar{l}(y_{1..n_y} | \bar{x}, X) = \frac{1}{n_y} \log(L(y_{1..n_y} | \bar{x}, X)) \quad (11)$$

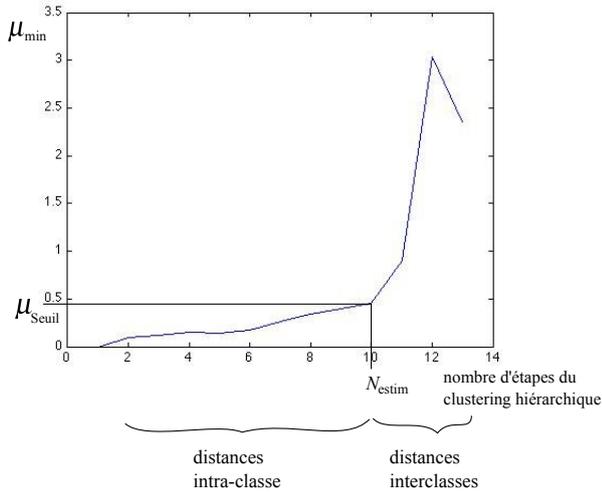


Figure 9. Evolution de μ_{\min} en fonction du nombre d'étapes du *clustering* hiérarchique

La mesure de vraisemblance gaussienne est définie par :

$$\mu(S_y|\bar{x}, X) = \frac{1}{p} [\text{tr}\Gamma - \log(\det(\Gamma)) + \delta^T X^{-1}\delta] - 1 \quad (12)$$

et [3] montre que :

$$\arg \max_{(\bar{x}, X)} (\bar{l}(y_{1\dots n_y}|\bar{x}, X)) = \arg \min_{(\bar{x}, X)} (\mu(S_y|\bar{x}, X)) \quad (13)$$

La recherche du paramètre de gaussienne (\bar{x}, X) maximisant la probabilité que le modèle ait produit les descripteurs de S_y revient à rechercher (\bar{x}, X) qui minimise la mesure de vraisemblance gaussienne. Cette mesure, asymétrique, est positive et telle que $\mu(S_y|\bar{x}, X) = 0$ si les éléments de S_y ont été générés par le modèle gaussien de paramètre (\bar{x}, X) . Cette mesure doit être symétrisée pour pouvoir être interprétable en terme de distance. Pour cela, on effectue une moyenne de la mesure avec sa mesure duale :

$$\mu_{[0.5]}(S_x, S_y) = \frac{1}{2}\mu(S_x|\bar{y}, Y) + \frac{1}{2}\mu(S_y|\bar{x}, X) \quad (14)$$

3.2.3. Estimation du critère d'arrêt

On considère l'ensemble des « distances cumulées » $\mu_{\min} = \{\mu_{\min}(j)\}_j$, avec $\mu_{\min}(j)$ la distance minimale liant deux classes ou groupes de classes à l'étape j du *clustering* hiérarchique. On suppose que cet ensemble est séparable en deux classes : la classe des distances « intra-classe », distances de faible valeur liant deux classes (ou groupe de classes) appartenant à la même étiquette, et la classe des distances « interclasses », de valeurs élevées, liant deux segments (ou groupe de segments) appartenant à deux étiquettes différentes. Ceci est illustré par la figure 9.

On partitionne une première fois ces distances selon ces deux classes via l'algorithme des K -Moyennes (avec $K = 2$).

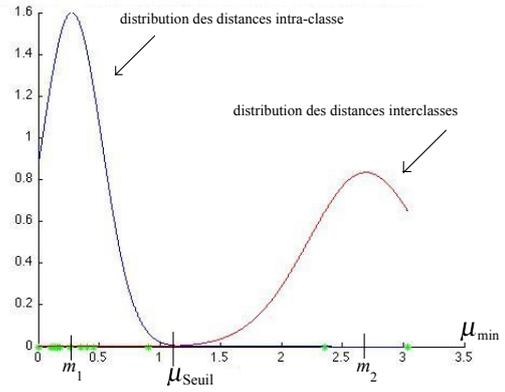


Figure 10. Modélisation gaussienne des classes de distance « intra-classe » et « interclasses », de moyennes respectives m_1 et m_2 .

On modélise ensuite chaque classe obtenue par l'intermédiaire d'une gaussienne, comme l'illustre la figure 10. La distance seuil, notée μ_{seuil} , est estimée comme étant la distance correspondant au « croisement des modèles gaussiens ». Cette distance sépare plus finement les distances « intra-classe » des distances « interclasses ». Le nombre de regroupements hiérarchiques est estimé par le nombre de distances « intra-classe » et noté par la suite N_{estim} . L'ensemble des distances cumulées possède autant d'éléments que le nombre maximum de regroupements hiérarchiques pouvant être fait. Le nombre de segments des chansons étant de l'ordre d'une douzaine, la classe des distances « interclasses » peut ne contenir qu'un seul élément ; dans ce cas on ne considère que le regroupement par les K -Moyennes, sans passer par la détermination de la distance seuil.

On suppose enfin que si N_{estim} et le nombre optimal de regroupements N_{opt} ne sont pas égaux, ils peuvent être linéairement dépendants. On introduit ainsi deux paramètres a et b tels que :

$$N_{\text{opt}} = E(a * N_{\text{estim}} + b) \quad (15)$$

$E(\cdot)$ représente la fonction partie entière. Les valeurs de a et b sont réglées par apprentissage sur un corpus de morceaux.

4. ÉVALUATION

4.1. Corpus d'étude

Un premier corpus d'expérimentation, que l'on note R, a permis la construction de la chaîne de structuration. Il est constitué de cinq morceaux de différents styles, issus de la base *Real World Computing (RWC) Music database*, qui est une base de données musicale libre de droits créée pour la recherche et proposée par le *RWC partnership* au Japon¹ comme base commune de travail. Les références des morceaux sont rassemblées au sein du tableau

1. <http://staff.aist.go.jp/m.goto/RWC-MDB/> et [8, 9]

nom de la chanson	auteur	style
<i>Jinsei konnamono</i>	Fevers	pop
<i>Once in a lifetime</i>	Shinya Iguchi	pop
<i>Everyday lovin'</i>	Takeshi Wada	rock
<i>Woman</i>	Suguru Kawagoshi	heavy-metal
<i>Life gage</i>	Kazuo Nishi	house

Table 1. Référence des morceaux ayant servi de support pour le développement de la chaîne de segmentation.

1 ; les dénominations des morceaux choisis pour la suite de l'étude provient de la section et du numéro qui leur sont attribués dans la base RWC. Les créateurs de la base proposent leurs annotations sur leur site Internet, incluant des annotations structurelles ².

Un deuxième corpus, noté Q, est constitué d'une liste de 20 morceaux de styles variés (pop, rock, techno) et de leurs annotations fournies par l'IRCAM. On nommera par la suite ces annotations « annotations de référence ». La liste est proposée dans le cadre de la tâche « Evaluation of music structuring and summarizing » du projet QUAERO ³. Une première moitié (Q1) de ce corpus est utilisée pour l'apprentissage des paramètres de l'algorithme, la deuxième moitié (Q2) constituant le corpus de test de la chaîne de structuration ; le procédé est repris en échangeant le rôle des deux moitiés.

4.2. Mesures de performance

La segmentation est évaluée par des mesures de précision, de rappel et de F-mesure. La précision est le rapport du nombre de frontières correctement estimées au nombre total de frontières estimées, et le rappel est le rapport du nombre de frontières correctement estimées sur le nombre total de frontières de la vérité terrain ; une frontière estimée est « correcte » avec une tolérance de 3 secondes (non-fixé par l'évaluation, mais fréquemment trouvé dans la littérature [12]), c'est-à-dire que l'écart maximum entre la frontière estimée et celle des annotations de référence qu'on lui associe ne dépasse pas cette durée. La F-mesure est une moyenne harmonique de la précision et du rappel. Cette quantité se détériore lorsque l'un des deux critères diminue ; en particulier elle suit le comportement du plus faible des deux dans le cas où l'un est négligeable devant l'autre.

L'évaluation globale de la structure inférée (segments étiquetés) est effectuée par le calcul de l'erreur de modélisation [17, 20]. Celle-ci effectue une mise en correspondance entre les annotations et les estimations.

Notons T_e^j l'ensemble des segments temporels estimés étiquetés j et T_a^i l'ensemble des segments temporels annotés étiquetés i . Le nombre d'étiquettes utilisées, noté

2. <http://staff.aist.go.jp/m.goto/RWC-MDB/AIST-Annotation/>, voir « chorus sections »

3. Il s'agit d'un programme de recherche et d'innovation franco-allemand visant à développer des technologies de traitement automatique des contenus multimédia et multilingues, afin de proposer de nouveaux produits et services au grand public et aux professionnels. Site internet : www.quaero.org

J pour l'estimation et I pour l'annotation, peut être différent. On cherche la meilleure association entre les segments estimés j et annotés i , c'est à dire les couples (i, j') tel que le recouvrement entre T_a^i et $T_e^{j'}$ soit maximisé. On a ainsi $j'(i) = \arg \max_j (T_e^j \cap T_a^i)$. Une étiquette annotée j ou estimée i ne peut être associée qu'une fois ou ne pas être associée. Après appariement, l'erreur de modélisation est définie comme la partie du signal non-correctement assignée (ou non-assignée) et prend des valeurs comprises entre 0 et 1. Pour l'évaluation de la performance on s'intéresse à la valeur du « score de modélisation », défini par :

$$\text{score} = 1 - \sum_{T_e^{j'}} \sum_{T_a^i} |T_e^{j'} \cap T_a^i| / N \quad (16)$$

où N est la durée totale des annotations et où $|\cdot|$ représente le cardinal de l'ensemble considéré.

Cette quantité tend vers 1 lorsque les annotations et les estimations sont très proches.

4.3. Extraction des descripteurs

Les MFCC sont tout d'abord extraits sur des trames de taille 23 ms qui ne se recouvrent pas, puis ramenés à l'échelle des pulsations rythmiques de la manière décrite en 3.1. Le calcul des descripteurs et l'estimation des pulsations sont faits par le logiciel Sonic Visualiser 1.7.1 muni des plugins VAMP 1.6 développés au *Centre for digital music, Queen Mary University of London* ⁴.

4.4. Évaluation de la segmentation

La taille du noyau de convolution est fixée à $L = 40$ pulsations, soit 16 s en estimant la durée moyenne entre deux pulsations à 400 ms.

Le paramètre λ fixant le seuil de détection de rupture de timbre est fixé de la manière suivante.

La segmentation est effectuée sur le corpus d'apprentissage : on calcule la F-mesure pour chaque morceau et pour chaque valeur de λ qui varie de 5 à 15 avec un pas de 0.1 (une étude préliminaire sur R a permis d'estimer grossièrement le seuil optimal à 10). La valeur de λ utilisée pour le test est celle qui maximise la F-mesure moyenne sur le corpus d'apprentissage.

Le tableau 2 présente les résultats obtenus pour deux expériences : il s'agit des mesures moyennes de précision, rappel, F-mesure sur les corpus de test.

Dans la première expérience, le corpus R est utilisé à la fois en tant que corpus d'apprentissage et de test. La phase d'apprentissage a permis de fixer $\lambda = 9.7$.

Dans la deuxième expérience, le corpus Q1 sert d'ensemble d'apprentissage et Q2 constitue l'ensemble de test. On a ainsi pu fixer $\lambda = 7.5$ sur Q1 et $\lambda = 9.4$ sur Q2.

4. sites Internet associés : <http://www.sonicvisualiser.org/>, <http://www.vamp-plugins.org/>

App.	Test	λ	précision	rappel	F-mesure
R	R	9.7	0.6986	0.6748	0.6798
Q1	Q2	7.5	0.5958	0.7034	0.6273
Q2	Q1	9.4	0.5718	0.6353	0.5840

Table 2. Évaluation de la segmentation : mesures moyennes de précision, rappel et F-mesure obtenues sur les différents corpus de l'étude. R : corpus de développement, Q1 et Q2 : moitiés du corpus servant d'apprentissage (« App. ») et de test.

App.	Test	Segmentation	a	b	ScoreMoy
R	R	connue	0.16	1.67	80.24%
R	R	automatique	0.81	-3.63	64.95%
Q1	Q2	connue	0.49	-0.20	79.07%
Q1	Q2	automatique	0.99	-2.85	59.46%
Q2	Q1	connue	0.23	2.17	73.88%
Q2	Q1	automatique	0.03	4.08	56.27%

Table 3. Évaluation de l'étiquetage : a et b déterminés par apprentissage (« App. ») et scores de modélisation moyens (« ScoreMoy ») sur les corpus d'étude.

4.5. Évaluation de l'étiquetage

Les paramètres a et b sont fixés en déterminant pour chaque morceau du corpus d'apprentissage le nombre de regroupements hiérarchiques optimal N_{opt} , maximisant le score de modélisation, et N_{estim} par la méthode décrite en 3.2.3. L'ensemble des couples (N_{opt}, N_{estim}) obtenus est modélisé par une droite de coefficient directeur a et d'ordonnée à l'origine de la droite b (régression linéaire) qui sont les paramètres utilisés pour la phase de test.

L'étiquetage est évalué en deux phases. Le score de modélisation est tout d'abord calculé sur les segments obtenus avec la segmentation idéale connue : les frontières de ces blocs sont celles des annotations de référence. Ceci permet de s'affranchir des performances de l'étape de segmentation. Ce score est ensuite calculé sur les blocs structurels issus de la phase de segmentation automatique, ce qui permet de mesurer la performance de la chaîne complète. Les valeurs des paramètres fixés par apprentissage et les scores de modélisation associés sont présentés dans le tableau 3.

On observe une différence d'en moyenne 17% entre les scores de modélisation calculés avec la segmentation idéale connue et ceux calculés avec la segmentation obtenue automatiquement.

4.6. Evaluation de systèmes annexes et discussion des résultats

Plusieurs variantes du système proposé ont été évaluées afin de discuter l'importance du critère de rupture utilisé lors de l'étape de segmentation, et de discuter l'influence de la métrique utilisée pour la comparaison des segments obtenus au cours de l'étape d'étiquetage.

s1					
App.	Test	n_p	précision	rappel	F-mesure
R	R	31	0.5677	0.8269	0.6709
Q1	Q2	38	0.4795	0.9207	0.6168
Q2	Q1	23	0.5217	0.6560	0.5726
s2					
App.	Test	λ	précision	rappel	F-mesure
R	R	14.9	0.7667	0.6615	0.7049
Q1	Q2	11.8	0.5611	0.7841	0.6400
Q2	Q1	12.2	0.5471	0.6507	0.5870
s3					
App.	Test	n_p	précision	rappel	F-mesure
R	R	12	0.6603	0.6748	0.6573
Q1	Q2	20	0.5289	0.7660	0.6105
Q2	Q1	17	0.5454	0.6991	0.5919

Table 4. Évaluation de la performance des systèmes de segmentation s1, s2 et s3.

4.6.1. Importance du critère de rupture pour la segmentation

Trois variantes du procédé de segmentation proposé sont testées afin d'évaluer l'intérêt de la détection de frontières par seuillage du critère de Seck.

Les systèmes s1 et s2 se limitent à l'étude de la courbe de nouveauté timbrale. Le système s1 localise les frontières par la détection de tous les pics de nouveauté pour ne garder que les n_p plus grands, tel que n_p maximise la F-mesure moyenne sur le corpus d'apprentissage. Le système s2 ne conserve que les pics dépassant un seuil λ , fixé comme en 4.4. Le système s3 reprend le système s1 mais fixe n_p à partir du critère de Seck.

Les performances obtenues sont référencées dans le tableau 4. Lorsque l'on compare les performances des systèmes s1 et s3, on ne remarque qu'une faible variation des F-mesures moyennes sur les ensembles de test, bien que n_p chute de 14 unités en moyenne d'un système à l'autre. Ceci met en évidence que le critère de Seck parvient à diminuer fortement le nombre de fausses détections et de mettre en valeur les pics de nouveauté pertinents pour notre segmentation. Les performances du système s2 comparées à celles du système évalué en 4.4 ont un comportement similaire, avec néanmoins une légère diminution de la F-mesure moyenne sur ce dernier système. La diminution des fausses détections permet de limiter la présence de segments trop petits par rapport à notre étude à moyen terme de la structure musicale.

Enfin, l'ensemble de ces performances confirme qu'il n'est pas judicieux de fixer le nombre de segments sur un ensemble de morceaux de musique, comparé aux méthodes utilisant le seuillage.

4.6.2. Mesure de vraisemblance gaussienne et rapport de vraisemblance gaussien pour l'étiquetage

Afin de mesurer la pertinence de l'utilisation de la mesure de vraisemblance gaussienne lors du processus d'éti-

App.	Test	Segmentation	a	b	ScoreMoy
R	R	connue	2	-1	87.49%
R	R	automatique	-2	10	72.03%
Q1	Q2	connue	0.58	4.42	75.83%
Q1	Q2	automatique	6.28	-8.11	46.53%
Q2	Q1	connue	0.61	3.77	70.55%
Q2	Q1	automatique	0.22	6.13	53.30%

Table 5. Évaluation de l'étiquetage : cas du rapport de vraisemblance gaussien utilisé en tant que distance entre classes de segments.

quetage, on utilise cette fois un rapport de vraisemblance gaussien (RVG) en tant que distance entre les classes de segments obtenues au cours du *clustering* hiérarchique. Le RVG permet de mettre en relation les vraisemblances de 2 modélisations gaussiennes d'un même couple de classes : une modélisation des deux classes par une seule distribution mono-gaussienne, et une modélisation associant à chaque classe une distribution mono-gaussienne différente.

Formellement, si l'on considère les classes S_x et S_y composées respectivement des descripteurs $\{x_t\}_{1 \leq t \leq n_x}$ et $\{y_t\}_{1 \leq t \leq n_y}$ et modélisées respectivement par les gaussiennes G_x de paramètre θ_x et G_y de paramètre θ_y , on a :

$$\text{RVG}(S_x, S_y) = \frac{L(x_1 \dots x_{n_x} y_1 \dots y_{n_y} | \theta_z)^{n_x + n_y}}{L(x_1 \dots x_{n_x} | \theta_x)^{n_x} L(y_1 \dots y_{n_y} | \theta_y)^{n_y}} \quad (17)$$

θ_z est le paramètre de la gaussienne issue de la fusion de G_1 et G_2 (au sens du paragraphe 3.2.1) et $L(x|\theta)$ est la vraisemblance des échantillons x étant donné le modèle gaussien de paramètre θ .

On regroupe ainsi à chaque étape du *clustering* hiérarchique les deux classes dont le RVG est le plus élevé : le couple qui est le plus susceptible d'être modélisé par une distribution mono-gaussienne devient une seule classe de segments. Les scores de modélisation pour les différents ensembles de morceaux sont référencés dans le tableau 5. On constate que l'utilisation de la mesure de vraisemblance gaussienne permet d'obtenir de meilleures performances pour l'étiquetage des segments. Remarquons enfin que les valeurs idéales $a = 1$ et $b = 0$ (pas d'ajustement de N_{estim}) sont rarement atteintes, ce qui signifie qu'une voie d'amélioration possible de la chaîne serait de modéliser plus finement les classes de segments et les classes de distances « interclasses » et « intra-classe », par exemple en ayant recours à des modèles de mélanges gaussiens.

5. CONCLUSION

Cet article a présenté un système de détection de ruptures timbrales pour l'inférence de structure musicale développé au sein de l'équipe METISS. Ce système, inspiré de l'état de l'art, présente quelques nouveautés : un critère de détection de rupture a été introduit lors de l'étape

de segmentation ; la comparaison des segments timbraux a été faite par des métriques utilisées en traitement de la parole, que l'on a utilisé afin d'estimer le critère d'arrêt du *clustering* hiérarchique. On a montré leur intérêt après avoir évalué plusieurs versions du système proposé. La prochaine étape de notre étude consiste à comparer ces résultats avec ceux d'autres algorithmes existants et de mesurer les progrès accomplis par nos travaux par comparaison à ce système de référence.

6. REFERENCES

- [1] S. Abdallah, K. Noland, C. M. Sandler, M., and C. Rhodes, "Theory and evaluation of a bayesian music structure extractor" *Proceedings of the ISMIR Conference*, London, UK, pp. 420–425, Sep. 2005.
- [2] M. A. Bartsch and G. Wakefield, "To catch a chorus : Chroma-based representations for audio thumbnailing" *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, New York, USA, pp. 15–18, October 2001.
- [3] F. Bimbot, I. Magrin-Chagnolleau, and L. Mathan, Second-Order Statistical Measures for Text-Independent Speaker Identification. In *Speech Communication*, Vol.17, No. 1-2, pp. 177-192, August 1995.
- [4] M. A. Casey, R. Veltkamp, M. Goto, M. Leman, C. Rhodes, and M. Slaney, "Content-based music information retrieval : Current directions and future challenges" *Proceedings of the IEEE*, vol. 96, no. 4, April 2008.
- [5] M. Cooper and J. Foote, "Media segmentation using self-similarity decomposition" *Proceedings of the SPIE Storage and Retrieval for Multimedia Databases*, San Jose, California, USA, pp. 167–175, January 2003.
- [6] J. Foote, "Automatic audio segmentation using a measure of audio novelty" *Proceedings of the IEEE International Conference on Multimedia and Expo*, New York, New York, USA, pp. 452–455, August 2000.
- [7] M. Goto, "A chorus-section detecting method for musical audio signals" *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Hong Kong, China, pp. 437–440, April 2003.
- [8] M. Goto, H. Hashiguchi, T. Nishimura et R. Oka, "RWC Music Database : Music Genre Database and Musical Instrument Sound Database" *Proceedings of the ISMIR Conference*, USA, October 2003.
- [9] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka, "RWC Music Database : RWC music database : Popular, Classical, and Jazz Music Databases" *Proceedings of the ISMIR Conference*, USA, pp. 287–288, October 2002.

- [10] T. Jehan, "Hierarchical multi-class self similarities" *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Mohonk, New York, USA, October 2005.
- [11] K. Jensen, "Multiple scale music segmentation using rhythm, timbre and harmony" *EURASIP Journal on Advances in Signal Processing*, vol. 2007, 2007.
- [12] M. Levy and M. Sandler, "Structural segmentation of musical audio by constrained clustering" *IEEE transactions on audio, speech and language processing*, vol. 16, no. 2, pp. 318–326, February 2008.
- [13] M. Levy, M. Sandler, and M. A. Casey, "Extraction of high-level musical structure from audio data and its application to thumbnail generation," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Toulouse, France, pp. 13–16, May 2006.
- [14] B. Logan and S. Chu, "Music summarization using key phrases" *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Istanbul, Turkey, pp. 749–752, June 2000.
- [15] L. Lu, M. Wang, and H.-J. Zhang, "Repeating pattern discovery and structure analysis from acoustic music data" *Proceedings of the Multimedia Information Retrieval Workshop*, New York, New York, USA, pp. 275–282, October 2004.
- [16] J. Paulus and A. Klapuri, "Music structure analysis by finding repeated parts," *Proceedings of AMCMM*, Santa Barbara, California, USA, p. 59-68, October 2006.
- [17] G. Peeters, "Sequence representation of music structure using higher-order similarity matrix and maximum likelihood approach" *Proceedings of the IS-MIR Conference*, Vienna, Austria, September 2007.
- [18] G. Peeters, A. La Burthe, and X. Rodet, "Toward automatic music audio summary generation from signal analysis" *Proceedings of the International Conference on Music Information Retrieval*, Paris, France, pp. 94–100, October 2002.
- [19] E. Peiszer, T. Lidy, and A. Rauber, "Automatic audio segmentation : Segment boundary and structure detection in popular music" *Proceedings of the Workshop on Learning Semantics of Audio Signals*, Istanbul, Turkey, 2008.
- [20] G. Peeters and E. Vincent, "QUAERO, Task 6.5 : Evaluation of music structuring and summarization, Evaluation Guidelines," Internal Report, April 2009.
- [21] L. R. Rabiner, "A tutorial on hmm and selected applications in speech recognition," *Proceedings of the IEEE Conference*, vol. 77, no. 2, pp. 257– 286, February 1989.
- [22] C. Rhodes, M. A. Casey, S. Abdallah, and M. Sandler, "A markov-chain monte-carlo approach to musical audio segmentation," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Toulouse, France, pp. 797–800, May 2006.
- [23] M. Seck, R. Blouet, and F. Bimbot, "The IRISA/ELISA Speaker Detection and Tracking Systems for the NIST'99 Evaluation Campaign". *Digital Signal Processing*, Volume 10, Issues 1-3, January 2000, Pages 154-171.
- [24] Y. Shiu, H. Jeong, and C. C. Jay-Kuo, "Similarity matrix processing for music structure analysis" *Proceedings of AMCMM*, Santa Barbara, California, USA, pp. 69–76, October 2006.

DÉCOMPOSITION EN BLOCS AUTONOMES COMPARABLES

UNE PROPOSITION DE DESCRIPTION ET D'ANNOTATION DE STRUCTURE POUR LE TRAITEMENT AUTOMATIQUE DE MORCEAUX DE MUSIQUE

Frédéric BIMBOT¹
frederic.bimbot@irisa.fr

Olivier LE BLOUCH²
olivier.le_bloch@inria.fr

Gabriel SARGENT²
gabriel.sargent@inria.fr

Emmanuel VINCENT²
emmanuel.vincent@inria.fr

Equipe-Projet METISS

(1) IRISA, CNRS - UMR 6074 - (2) INRIA, Rennes Bretagne Atlantique
Campus Universitaire de Beaulieu, 35042 RENNES cedex, France

RÉSUMÉ

La structure musicale d'un morceau de musique est un concept auquel il est fréquemment fait référence en musicologie, mais pour lequel il n'existe pas de définition communément admise.

Ceci soulève un problème méthodologique lorsque l'on souhaite concevoir et évaluer un procédé d'inférence automatique de structure musicale mais aussi lorsqu'il s'agit de spécifier une procédure d'annotation manuelle permettant de produire de façon reproductible une description structurelle pour un morceau de musique.

Cet article présente une approche dite de *décomposition en blocs autonomes comparables*, fondée sur des principes inspirés du structuralisme et de la linguistique, et fournissant une méthodologie d'analyse à partir de critères simples et basés sur l'expérience musicale de l'annotateur.

Une évaluation préliminaire de la méthode proposée, appliquée à l'annotation auditive de 20 morceaux, a abouti à une concordance entre annotateurs supérieure à 90 %, puis, ultimement, à une annotation consensuelle.

1. PRÉSENTATION

1.1. Préambule et contexte

Ces derniers mois, des travaux ont été initiés dans notre équipe dans le cadre du projet QUAERO [1], sur l'inférence de structure musicale par des algorithmes automatiques [2].

Les finalités de ces travaux sont multiples, notamment au niveau applicatif : la navigation rapide dans les morceaux de musique, la production automatique de résumés, la détection de couplets et de refrains, le remixage non-supervisé, etc... Mais ces travaux possèdent également un intérêt situé davantage en amont : l'utilisation d'informations structurelles pour la modélisation acoustique et musicale.

Cette thématique de recherche nous a amené à nous interroger préalablement sur la définition même du concept de structure musicale ainsi que sur la façon de spécifier une procédure d'annotation permettant de produire manuellement et de façon reproductible une description structurelle d'un morceau de musique.

Cet article présente le résultat d'une première étape de ces travaux, que nous exposons dans le but de pouvoir échanger des idées et confronter nos points de vue avec les communautés scientifiques des Sciences Pour l'Ingénieur et de la Musicologie

1.2. Cadre du travail

Les notions et la méthodologie présentées dans cet article sont destinées à être appliquées à ce que nous désignerons par la musique « conventionnelle », qui couvre notamment la très grande majorité de la musique populaire actuelle d'inspiration occidentale, mais aussi une fraction significative d'autres familles de musique, notamment une partie de la musique classique. Nous conservons cependant présent à l'esprit que plusieurs types de musique, notamment certains courants de la musique contemporaine, ne se prêtent pas aussi bien, voire pas du tout, à l'approche proposée.

L'entité considérée est le morceau de musique. Toutefois, on peut être amené à s'intéresser à la partie de morceau de musique (section), lorsque le morceau est composé de plusieurs grandes parties sans lien entre elles.

1.3. Objectifs

Le but de ce travail est de proposer une définition opérationnelle d'une certaine acception de la notion de structure musicale et de spécifier une procédure permettant l'annotation quasi-univoque de cette structure par différents annotateurs.

On observe en effet qu'il est fréquemment fait référence au concept de structure musicale dans les ouvrages de théorie musicale, en tant qu'agencement formel régissant la construction d'une œuvre.

Toutefois, plusieurs conceptions de la structure musicale coexistent et il n'existe pas de définition unique communément admise [3]. De plus, les définitions proposées reposent souvent sur le concept musicologique de « phrase musicale » qui est ambigu ou dénué de sens pour de nombreux morceaux.

Ceci soulève donc un problème méthodologique, dès lors que l'on souhaite évaluer un procédé automatique

d'inférence de structure, puisqu'il se pose alors la question de la référence à laquelle comparer le résultat de l'algorithme pour en mesurer les performances. Pour un état des lieux très complet sur la question, voir [4].

L'approche présentée dans cet article vise à définir et à spécifier une méthodologie d'annotation de la structure musicale d'un morceau qui possède les propriétés suivantes :

- *Reproductible* par différents annotateurs
- Basée sur *l'expérience musicale* auditive de l'annotateur (et non sur son expertise musicologique)
- *Déconnectée* de tout a priori algorithmique et de toute application particulière
- *Ne s'appuyant pas* sur des propriétés acoustiques absolues
- *Indépendante* d'un quelconque *rôle musical* attribué aux éléments structurels (notamment « couplet », « refrain », etc...)
- Applicable à un *éventail de genres musicaux* le plus vaste possible

Au stade actuel de notre travail, nous nous sommes focalisés sur la question de la décomposition en éléments structurels et nous reportons à une phase ultérieure la question de l'étiquetage de ces éléments.

Nous présentons d'abord les fondements sur lesquels repose notre approche, puis nous spécifions un procédé visant à obtenir une décomposition structurelle basée sur ces fondements. Nous évaluons la cohérence de la décomposition obtenue par plusieurs annotateurs sur un ensemble de 20 morceaux de musique, l'annotation se faisant par écoute du morceau sous une forme sonore.

En dépit de son côté nécessairement un peu réducteur du fait de ses visées « opérationnelles », nous pensons que l'approche présentée ici est susceptible d'apporter un éclairage intéressant sur des concepts qui sont habituellement abordés sous l'angle de la création artistique ou de la perception, plutôt que sous celui de l'ingénierie.

1.4. Définitions préliminaires

On considère qu'un morceau possède 3 *propriétés de référence* qui peuvent être constantes ou subir des modifications au cours du morceau :

- tonalité/modalité (hauteur et gamme de référence)
- tempo (vitesse d'exécution)
- timbre (instrumentation / texture sonore)

On considère également qu'un morceau de musique possède 4 *niveaux d'organisation temporelle* :

- rythme (durée relative des notes)
- harmonie (progression des accords)
- mélodie (hauteur relative des notes)
- paroles (contenu linguistique)

L'ensemble forme 7 strates musicales dont on fait l'hypothèse qu'elles évoluent de façon indépendante au cours d'un morceau.

2. FONDEMENTS

2.1. Concept de structure musicale

L'approche que nous adoptons se place dans le cadre du *structuralisme*, initié par Ferdinand de Saussure dans le domaine de la Linguistique [5] et étendu ensuite à de nombreuses autres disciplines, notamment à la sémiologie musicale [6]. On peut en résumer ainsi le principe général : ce qui détermine la *structure* d'une entité, ce sont essentiellement les relations que ses constituants entretiennent les uns avec les autres au sein de l'entité, indépendamment de la forme et du sens de ces constituants.

Appliquée à notre objet d'étude, cette approche nous amène à considérer un morceau de musique comme le résultat de l'agencement d'un ensemble d'éléments constitutifs, selon un certain processus d'assemblage (dit *processus syntagmatique*). Les éléments constitutifs entretiennent également entre eux des relations qui permettent de les comparer (dites *relations paradigmatiques*) lesquelles s'expriment sous forme de relations d'équivalence. L'ensemble forme un *système* au sens structuraliste du terme, c'est-à-dire une « entité de dépendances internes », selon la définition de Hjelmslev [7].

Le morceau de musique apparaît alors comme une réalisation particulière (ou *observation*) issue de ce système et le problème d'inférence de structure musicale consiste à déterminer, à partir de cette unique observation, la délimitation des éléments constitutifs du morceau (segmentation ou, plus généralement *décomposition*) et l'attribution d'une classe d'équivalence à chacun d'entre eux (*étiquetage*).

Dans ce cadre général, la spécification d'un type particulier de structure musicale passe donc par la définition :

1. de la nature des éléments constitutifs considérés
2. de leur processus d'agencement
3. de la relation d'équivalence permettant la comparaison des éléments

2.2. Hypothèses de travail

Dans les travaux présentés ici, les éléments constitutifs sont supposés communs aux 4 niveaux d'organisation temporelle.

Ils sont limités dans le temps et ils sont agencés principalement par concaténation, avec toutefois possibilité de chevauchement (tuilage). On désigne ces éléments constitutifs par *blocs*.

Un bloc est *autonome*, c'est-à-dire doté d'une cohérence musicale propre. Il est défini par un *début*, une *durée* et une *taille*. La distinction entre durée et taille est explicitée en section 2.3.1.

Un bloc peut se décomposer en un *radical* (lui-même doté d'une cohérence musicale propre) et un ou plusieurs *affixes* (cf. 2.3).

Différentes relations d'équivalence exprimant la *comparabilité* entre blocs peuvent être considérées, notamment : l'isométrie, l'interchangeabilité, la similarité,...

Ainsi, dans le cadre de cet article, l'inférence de structure musicale consiste à effectuer une décomposition du morceau en *blocs autonomes comparables*.

La notion de *bloc autonome comparable* possède des similitudes avec la notion de « phrase musicale », sans toutefois pouvoir y être assimilée. Elle possède également une parenté avec la notion de *grouping structure*, développée dans [8].

2.3. Définitions spécifiques

On précise dans cette section quelques propriétés et définitions relatives aux segments musicaux d'un morceau, qui sont nécessaires pour spécifier la procédure de décomposition en blocs autonomes comparables.

2.3.1. Cohérence musicale

Pour plusieurs de ces définitions, on considère l'effet de transformations telles que la suppression, l'insertion ou la substitution de segments musicaux dans le morceau considéré et on évalue la *cohérence musicale* du résultat ainsi obtenu vis-à-vis de l'ensemble des strates définies au paragraphe 1.4.

La notion de cohérence musicale est délicate à définir. On peut tenter de la préciser en considérant par exemple que la cohérence musicale du morceau est préservée dès lors que la transformation effectuée ne crée pas de singularité morphologique vis-à-vis du morceau d'origine et est compatible avec son agencement syntagmatique : en d'autres termes, qu'on rencontre (ou qu'on pourrait rencontrer) ailleurs dans le même morceau, un passage semblable à celui résultant de la transformation effectuée.

Cette définition possède une part de subjectivité mais, en l'absence d'une formulation plus rigoureuse (qui reste à trouver), elle fournit un critère qui permet à un auditeur humain de décider si il considère que la transformation effectuée est admissible ou non.

2.3.2. Propriétés intrinsèques

Un segment musical est dit :

- *valide*, si il ne présente pas d'incohérence musicale majeure (sur aucune strate)
- *indépendant*, si son écoute isolée produit une impression de complétude chez l'auditeur.
- *itérable*, si son raccordement avec lui-même donne lieu à un segment valide.
- *autonome, si il est indépendant ou itérable*
- *incomplet*, si il n'est pas autonome
- *homogène*, si ses propriétés de référence sont constantes sur la durée du segment

La *taille* d'un segment musical est exprimée en nombre de *snaps*, celui-ci étant défini comme la période à laquelle on

claque des doigts pour accompagner la musique à une périodicité approchant le plus possible la seconde.

Une portion (éventuellement non-connexe) de segment musical est qualifiée d'*affixe* s'il est possible de le supprimer tout en conservant un segment musical valide. Si l'affixe est situé en début (resp. en fin) de segment et si il est connexe, il est qualifié de *préfixe* (resp. *suffixe*), sinon, on le désigne comme *infixe* (éventuellement non-connexe).

2.3.3. Propriétés syntagmatiques

Un segment musical au sein d'un morceau est dit :

- *suppressible* si sa suppression du morceau n'entraîne pas d'incohérence musicale majeure
- *insérable* à un endroit du morceau si son insertion n'entraîne pas d'incohérence musicale majeure

2.3.4. Propriétés paradigmatiques et relations d'équivalence

Un segment de taille n' est dit *réductible* à un autre segment de taille $n \leq n'$ si le segment de taille n peut s'obtenir à partir du segment de taille n' par suppression de $n'-n$ *snaps*. Si le segment de taille n est valide, alors ce dernier constitue un *radical* de celui de taille n' .

Les principales relations d'équivalences (déjà évoquées plus haut dans la section 2.2) sont :

- *isométrie* : blocs ou segments de taille identique (isométrie absolue) ou réductibles à des radicaux de même taille (isométrie de radicaux).
- *interchangeabilité* : blocs ou segments qu'on peut interchanger dans un morceau sans introduire d'incohérence musicale majeure.
- *similarité* : blocs ou segments partageant des caractéristiques communes sur tout ou partie des strates musicales (cette ressemblance pouvant porter sur tout le bloc, ou seulement sur les radicaux).
- *isomorphisme* : blocs pouvant être obtenus l'un à partir de l'autre par transformation de leurs propriétés de référence.

2.4. Propriétés complémentaires

Par convention, on suppose que les frontières des blocs sont synchronisées avec les débuts de mesures (les anacrouses étant considérées comme faisant partie du bloc auquel elles se rattachent).

Les blocs ne sont pas nécessairement homogènes : leurs propriétés de référence peuvent varier au sein d'un bloc.

Les blocs d'un même morceau sont potentiellement répétés (à des variations près) ou isomorphes.

3. SPÉCIFICATIONS

La définition des blocs est ambiguë pour deux raisons.

D'une part, les notions d'*autonomie* et de *comparabilité* ne sont pas univoques ; elles peuvent s'avérer subjectives et

dépendantes du morceau étudié. D'autre part, à elles seules, ces notions ne mènent pas nécessairement à une structure unique

Pour lever ces ambiguïtés, on convient :

- 1) d'attribuer le même poids à toutes les strates sonores (tonalité, tempo, timbre, mélodie, rythme, harmonie, paroles) quel que soit le morceau étudié
- 2) d'ajouter des contraintes supplémentaires, dits *critères de décomposition*, explicités dans la sous-section suivante

On tente de formuler les critères de décomposition sans faire appel à des propriétés acoustiques absolues, ni au rôle musical des segments, de sorte à être le plus indépendant possible du genre musical auquel se rapporte le morceau.

3.1. Critères de décomposition

3.1.1. Test de suppressibilité

L'ensemble des blocs d'un morceau constitue un ensemble de segments interchangeable (au moins au sein d'une même classe d'équivalence). Il est toutefois difficile de tester cette propriété, à laquelle on préfère une propriété plus faible, mais plus pratique : la suppressibilité (définie en section 2.3.2).

Le test de suppressibilité vise à *identifier les principales hypothèses* plausibles de frontières de blocs. Cependant, tout segment suppressible ne constitue pas nécessairement un bloc.

3.1.2. Régularité de la structure

On fait l'hypothèse que tout morceau est bâti sur un ensemble minimal de *pulsations structurelles* qui régissent la taille des blocs, préférentiellement :

- une valeur (type I)
- deux valeurs (type II)
- une suite régulière de valeurs, dit « patron structurel » (type III)
- un ensemble limité de valeurs mais une absence de régularité structurelle (type IV)
- l'existence de blocs identifiables mais un large éventail de tailles différentes (type V)

On désignera par morceau de type 0 (ou indéterminable) tout morceau pour lequel on constate l'impossibilité de délimiter des frontières bien localisées de blocs autonomes.

On distingue :

- les *blocs réguliers*, majoritaires, dont la taille n est conforme à la pulsation structurelle (ou cohérente avec le patron structurel). Dans ce cas, le bloc sera annoté comme étant de taille n .
- les *blocs dérivés* de blocs réguliers, dont la taille est proche de l'une des pulsations structurelles et qui peuvent se décrire simplement :

- soit par ajout d'un (ou plusieurs) suffixe(s), préfixe(s) ou infixé(s) à un bloc régulier (bloc dit *rallongé*)
- soit par suppression en position finale, initiale ou quelconque d'une (ou plusieurs) portion(s) de bloc par rapport à un bloc régulier (bloc dit *tronqué*).

Le *radical structurel* dont est supposé issu le bloc dérivé est généralement observé (à des variantes près) dans le même morceau (ou pourrait l'être sans sembler incohérent avec le reste du morceau).

- les *blocs irréguliers* dont la taille est différente et généralement éloignée de la (ou des) pulsation(s) structurelle(s) et/ou qui ne peuvent pas se rapporter clairement à un radical structurel. Il s'agit souvent de blocs de taille inférieure au moins de moitié à la pulsation structurelle (*raccord ou pontage*).

En outre, on autorise le *chevauchement de blocs* ou *tuilage*, c'est-à-dire un début de bloc anticipant la fin du précédent pour rendre compte de phénomènes de démarrage de blocs asynchrone avec la fin du précédent, que l'on peut observer dans certaines constructions de morceaux (notamment les canons ou les fugues).

A cet égard, il est important de distinguer le tuilage de l'anacrouse. Dans l'anacrouse, le contenu d'un ou plusieurs niveaux d'organisation temporelle débute avant la frontière de bloc selon un mécanisme d'*anticipation*. Au contraire, dans le tuilage, la frontière d'un nouveau bloc est observée alors que le bloc précédent n'est pas encore terminé, donnant lieu à la *superposition* des deux blocs et à leur coexistence sur tous les niveaux d'organisation temporelle pendant la durée de leur tuilage.

La contrainte de régularité vise à obtenir des *blocs comparables* au sein de chaque morceau.

La figure 1 présente un schéma de décomposition en blocs structurels dans le cas d'une décomposition parfaitement régulière de pulsation structurelle n puis illustre différentes configurations d'irrégularités pouvant être rencontrées.

3.1.3. Durée cible des blocs

En appliquant les deux critères précédents, on peut parfois aboutir à plusieurs solutions possibles composées de blocs suppressibles et présentant une bonne régularité. En général, ces solutions correspondent à des valeurs de pulsation structurelle qui sont multiples ou sous-multiples les unes des autres.

On départage ces situations par un critère de durée cible du bloc, basé sur la minimisation de ce que nous désignons par *contexte informatif prédominant*.

La figure 2 illustre une décomposition structurelle sous forme d'une représentation paradigmatique, mettant en évidence la correspondance existant entre portions homologues dans les différents blocs.

Si cette décomposition est exploitée pour effectuer une prédiction des propriétés musicales du morceau sur un court intervalle de temps, les portions les plus pertinentes pour y parvenir seront d'une part les portions appartenant au

même bloc structurel et d'autre part les portions homologues dans les autres blocs structurels.

Ce contexte informatif prédominant est distinct pour chaque portion du morceau mais il est déterminé uniquement par la structure et il constitue la partie prépondérante de l'entité de dépendances internes évoquée dans la section 2.1.

Si la longueur totale du morceau est égale à N et si la longueur typique des blocs est égale à n , alors le nombre de blocs est de l'ordre de N/n et le contexte informatif prédominant a une longueur cumulée d'environ :

$$C = n + (N/n) - 2 \quad (1)$$

C est alors minimal lorsque $n = \sqrt{N}$.

Pour des morceaux de musique d'une durée typique de 4 minutes (soit $N = 240$, en prenant la seconde comme unité), la valeur de n minimisant C vaut environ 15.49 secondes.

Dans le travail présenté ici, on a retenu cette valeur (arrondie) comme valeur cible pour la pulsation structurelle ce que l'on exprime ainsi : au moins une des valeurs de pulsation structurelle doit être la plus proche possible, en échelle logarithmique, de la *durée cible* de 15 secondes. Avec ce critère, la décomposition structurelle obtenue tend à équilibrer le poids de l'axe syntagmatique et de l'axe paradigmatique dans leur contribution à la description de la structure du morceau.

Plus généralement, on pourra donner un poids relatif λ différent à chacun des deux termes de l'équation (1), conduisant à une fonction du type :

$$C(\lambda) = n + \lambda(N/n) - (\lambda+1) \quad (2)$$

dont la minimisation (en n) induira des décompositions basées sur un équilibre différent.

Le recours à une durée cible permet de résoudre les situations ambiguës telle qu'un bloc composé de plusieurs sous-blocs identiques et favorise l'obtention de blocs comparables entre plusieurs morceaux.

Cette contrainte peut toutefois être relâchée lors de variations importantes du tempo à l'intérieur d'un bloc ou à l'échelle de plusieurs blocs successifs.

3.1.4. Détermination des affixes et des troncatures

Un bloc de taille $n+p$ supérieure à la pulsation structurelle n sera préférentiellement décomposé comme un radical de taille n accompagné d'un affixe p (notations $n+p$, $p+n$ et $n\&p$) si :

- Le radical est autonome et de préférence observé seul dans le morceau tel quel ou sous une forme similaire.
- L'affixe de taille p est suppressible mais pas autonome sans le radical de taille n .
- La valeur de p est plus petite que $n/2$

Un bloc de taille $n-p$ inférieure à la pulsation structurelle n sera préférentiellement décomposé comme un radical de taille n tronqué (ou privé) d'un sous-bloc de taille p (notations $n-p$, $-p+n$ et $n\backslash p$), si :

- Le bloc peut être complété par un sous-bloc de taille p pour former un bloc régulier de taille n , autonome et de préférence observé dans le morceau tel quel ou sous une forme similaire.
- La valeur de p est plus petite que $n/2$

3.1.5. Recours au test d'interchangeabilité

Dans certaines situations ambiguës, notamment lorsque la pulsation structurelle a été fixée et que plusieurs segmentations compatibles avec cette pulsation structurelle mènent à des blocs suppressibles, on pourra avoir recours à un test d'interchangeabilité pour départager les différentes solutions possibles.

3.1.6. Situations particulières

Les tailles fractionnaires seront utilisées si l'annotateur estime que la taille d'un bloc correspond à un nombre non-entier de *snaps*.

Dans certains cas, l'annotateur jugera qu'il est impossible de définir un bloc, ou d'attribuer une taille à un bloc ou à un infixe. Ceci sera annoté par un symbole spécial (tableau 1).

Un symbole particulier sera utilisé dans les cas suivants : incertitude de l'annotateur sur la délimitation du bloc, incertitude sur la taille du bloc, taille indéterminable, segment de silence précédant le début ou suivant la fin du morceau, ou segment ne faisant pas partie du morceau.

3.2. Procédure de référence

A partir des critères énoncés ci-dessus, nous décrivons une procédure destinée à permettre à un annotateur humain de produire une décomposition structurelle conforme aux spécifications proposées.

1) Déterminer une valeur de la pulsation structurelle

Déterminer une ou plusieurs hypothèses de taille n en s'intéressant aux parties structurellement les plus stables, notamment à distance de l'introduction et de la conclusion du morceau ou de la partie de morceau. Privilégier en priorité les hypothèses correspondant à des durées compatibles avec la durée de référence.

2) Décomposer partiellement le morceau par rapport à la pulsation structurelle n

Identifier les hypothèses de début de blocs et déterminer les blocs suppressibles de taille n (réguliers) ou dérivant simplement (par extension ou troncature) de blocs réguliers de taille n (existants ou plausibles). Cette étape est également l'occasion de rechercher des tuilages.

3) Compléter la décomposition

Compléter la décomposition partielle par des blocs irréguliers, notamment des « pontages » de taille $n/2$, $n/4$ ou des blocs complètement irréguliers. S'assurer à cette

étape que les blocs ainsi formés sont suppressibles (sinon, il convient de les rattacher à des blocs déjà existants).

4) *Evaluer la régularité de la décomposition obtenue selon les indices suivants*

- forte majorité de blocs réguliers
- répartition équilibrée entre plusieurs pulsations structurelles / corrélées à la macro-structure (couplet / refrain, sections, etc...)
- patron structurel simple à décrire (répétitif, progressif, etc...)

Pour évaluer la régularité de la structure, on ne tiendra pas compte de façon primordiale de la décomposition de l'introduction ou de la fin du morceau. On n'attachera pas trop d'importance aux « pontages » (surtout si il s'agit de blocs dont la taille est égale à la demi-pulsation structurelle et qui jouent manifestement le rôle de transitions musicales) et on accordera une faible importance aux suffixes, aux préfixes et aux infixes (dans un ordre décroissant de régularité).

5) *Envisager d'autres solutions*

Considérer la/les décomposition(s) qui pourraient être obtenues avec d'autres hypothèses de pulsation structurelle et évaluer si elles donnent un résultat préférable au précédent.

3.3. Conventions d'annotation

Le tableau 1 résume l'ensemble des conventions d'annotation adoptées pour décrire les tailles de blocs.

3.4. Tableau de bord et rapport d'annotation

Au cours (ou à l'issue) de l'étude d'un morceau, l'annotateur peut s'appuyer sur un « tableau de bord » généré automatiquement et qui synthétise les principales propriétés de l'annotation produite pour le morceau considéré, notamment :

- La distribution des tailles des blocs
- Les durées moyennes pour chaque taille (ainsi éventuellement les valeurs médianes, minimales, maximales ...)
- La taille des tuilages
- Une représentation par automate minimal de la décomposition produite et différentes valeurs d'entropie correspondante (propre, croisée, ...)

Par ailleurs, l'annotateur indique sur un compte-rendu d'annotation les informations suivantes :

- Type d'organisation structurelle (0, I, II, III, IV, V)
- Difficulté de l'annotation (sur une échelle de 1 à 5)
- Degré de confiance dans le résultat produit (sur une échelle de 1 à 5)
- Observations (c'est-à-dire toute information qu'il juge pertinente)

Le tableau 2 illustre des exemples de tableaux de bord pour l'annotation de 2 morceaux.

4. ÉVALUATION

4.1. Objectifs

Cette section présente les résultats d'un processus d'évaluation préliminaire visant à valider la pertinence de l'approche proposée. Il s'agit de mesurer la concordance des annotations produites par différents annotateurs sur un même ensemble de morceaux et sa progression en fonction de l'évolution des consignes.

4.2. Protocole d'évaluation

On sélectionne un ensemble de morceaux et on les fait annoter par plusieurs annotateurs, selon les spécifications présentées ci-dessus. On évalue la cohérence des segmentations en prenant les annotations deux à deux et en mesurant la précision P , le rappel R et la F-mesure d'une segmentation par rapport à l'autre.

P , R et F sont des métriques désormais classiques dans le domaine du traitement de l'information : la précision et le rappel entre deux segmentations correspondent au pourcentage de frontières de blocs concordantes, avec une tolérance temporelle $\pm\delta t$, en prenant l'une puis l'autre des annotations comme référence.

On calcule ensuite la F-mesure :
$$F = 2 \frac{PR}{P+R} \quad (3)$$

pour chaque morceau, puis on moyenne ces valeurs sur l'ensemble du corpus.

Dans les expériences présentées dans cet article, le corpus de morceaux est composé de 20 titres dont la liste a été déterminée par l'IRCAM et qui ont servi d'ensemble de développement dans le cadre de la première phase du projet QUAERO pour la tâche 6.5 du projet CTC. La liste est donnée dans le tableau 3.

Quatre annotateurs ont participé à l'expérience. Aucun d'entre eux n'est musicologue, ni musicien professionnel. Toutefois, il convient de préciser qu'il s'agit des quatre auteurs de cet article, ce qui constitue un biais méthodologique qu'il faut prendre en compte dans l'interprétation des résultats.

La précision, le rappel et la F-mesure ont été calculés par le programme utilisé dans le cadre de la campagne d'évaluation QUAERO de Septembre 2009, réalisé par l'IRCAM et validé par l'IRIT [9]. La tolérance δt a été prise égale à 0,75 secondes.

4.3. Résultats

Le tableau 4 résume les résultats obtenus à l'issue d'une première session d'annotation.

Les annotations obtenues à l'issue de la première session ont été passées en revue collectivement pour discuter des discordances observées et notamment les imprécisions de

frontières et les situations où les spécifications n'avaient pas été respectées par tel ou tel annotateur. Ceci a donné lieu, après une période de « décantation » d'un mois environ, à une révision des annotations dont la concordance est résumée dans le tableau 5 ci-après. La figure 3 présente 3 exemples d'annotations multiples après révision.

01	Pink Floyd	Brain Damage
02	Queen	Lazing On A Sunday Afternoon
03	DJ Cam	Mad Blunted Jazz
04	Outkast	Return Of The G
05	ACDC	You Shook Me All Night Long
06	Eric Clapton	Old Love
07	Stan Getz & J. Gilberto	O Pato
08	Enya	Caribbean Blue
09	Mickael Jackson	Off The Wall
10	Bass America Collection	Planet
11	Plastikman	Fuk
12	Shack	Natalies Party
13	Sean Kingston	Take You There
14	Lil Mama	Shawty Get Loose
15	Abba	Waterloo
16	Eiffel 65	Blue (Da Ba Dee)
17	Meat Loaf	I'd Do Anything For You
18	Kaoma	Lambada
19	Vangelis	Conquest Of Paradise
20	Nirvana	Smells Like Teen Spirit

Tableau 3 : Liste des morceaux utilisés pour les expériences décrites dans cet article.

Annotateur	N°1	N°2	N°3	N°4
N°1	100	79.1	90.7	86.6
N°2	79.1	100	77.7	74.5
N°3	90.7	77.7	100	84.5
N°4	86.6	74.5	84.5	100

Tableau 4 : Concordance des annotations à l'issue d'une session initiale, mesurée en terme de F-mesure (en %) sur 20 morceaux. La moyenne globale des concordances inter-annotateurs s'établit à 82.2 %.

Annotateur	N°1	N°2	N°3	N°4
N°1	100	88.9	95.7	92.9
N°2	88.9	100	88.7	88.7
N°3	95.7	88.7	100	92.8
N°4	92.9	88.7	92.8	100

Tableau 5 : Concordance des annotations à l'issue de la phase de révision, mesurée en terme de F-mesure (en %) sur 20 morceaux. La moyenne globale des concordances inter-annotateurs s'établit à 91.3 %.

On observe, qu'après révision des annotations, la concordance inter-annotateur s'établit au-delà de 90 % ce qui valide globalement la cohérence du processus proposé, tout en laissant une petite marge d'amélioration potentielle. La figure 4 détaille la distribution des scores de concordance par morceaux. La médiane est à 95.8 %. Deux morceaux sont responsables de près de 4% des erreurs (sans eux, la concordance moyenne dépasserait 95 %).

Une part importante des discordances résiduelles observées provenant d'ambiguïtés de groupement de segments en nombre impair ayant une taille proche de la demi-pulsation structurelle, il a été décidé de traiter ces cas en adoptant un principe de « groupement par ressemblance ». Il s'agit d'un critère subsidiaire qui consiste, en cas de solutions multiples, à grouper dans un même bloc les segments qui se ressemblent et dissocier ceux qui sont dissemblables.

Ainsi, dans une phase ultime, une session de travail collective a permis d'aboutir à une décomposition consensuelle pour 19 des 20 morceaux, satisfaisant globalement les 4 annotateurs et l'ensemble des critères de décomposition. Le 20^{ème} morceau (n°11) a été classé en catégorie 0 (frontières de blocs indéterminables).

5. CONCLUSIONS

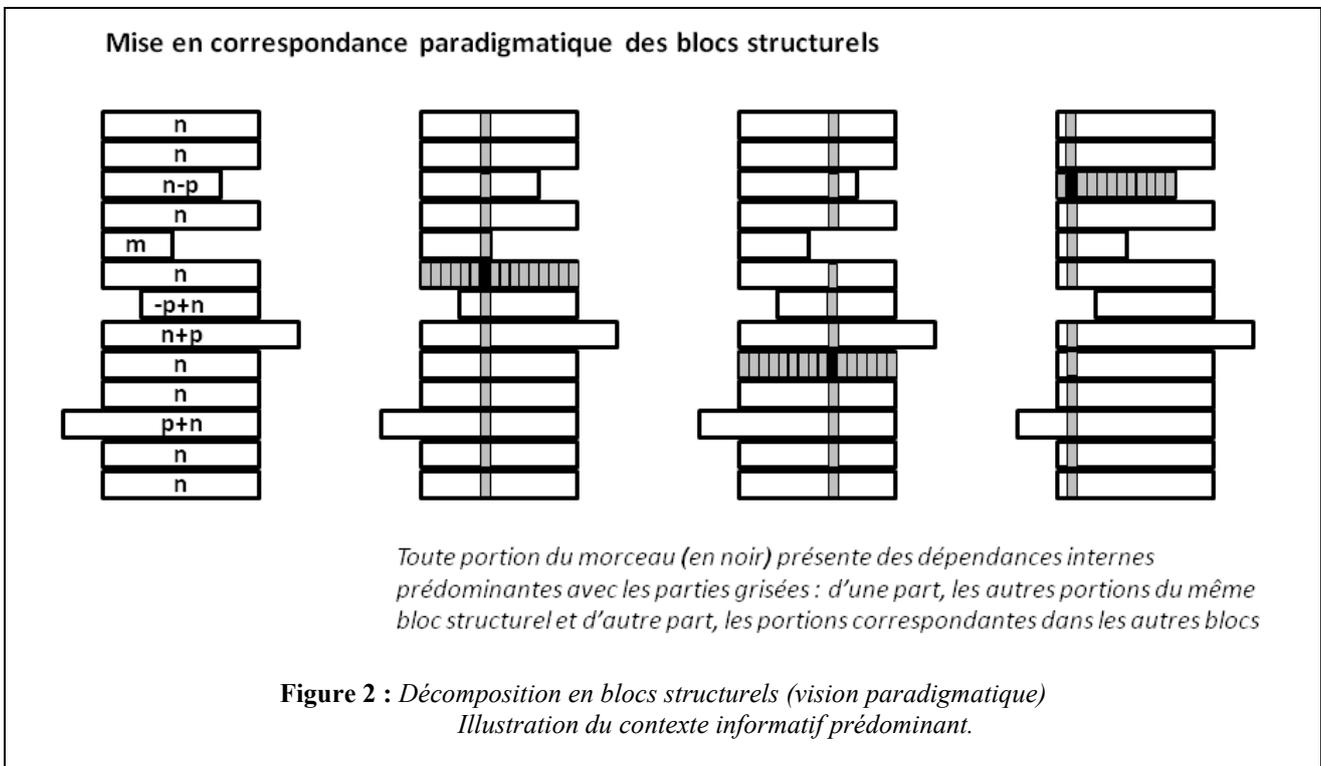
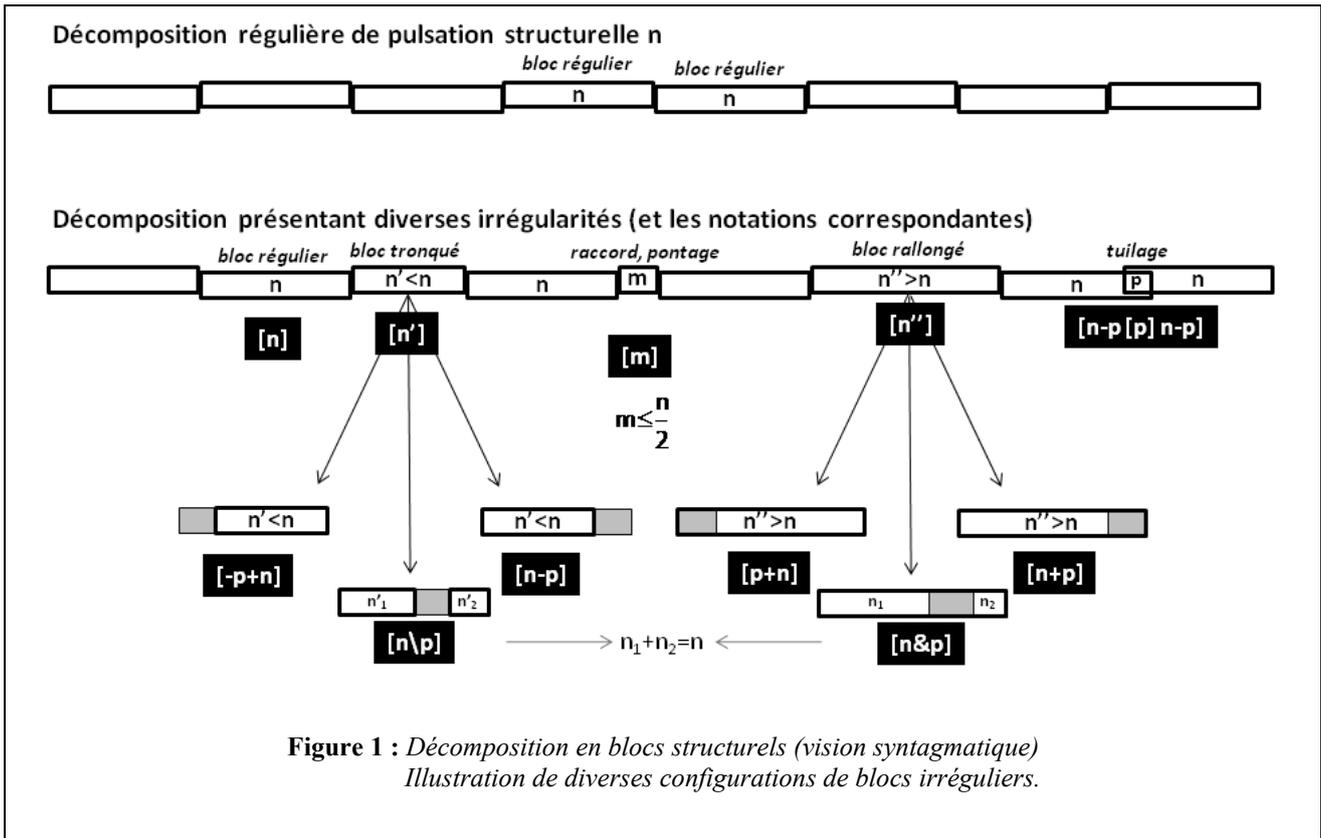
Les travaux présentés dans cet article ont permis de spécifier une procédure d'annotation visant à produire manuellement et de façon raisonnablement reproductible une description structurelle d'un morceau de musique. Conformément aux objectifs visés, la procédure proposée ne s'appuie pas sur une expertise musicologique, ni sur une analyse du rôle musical des éléments constitutifs du morceau.

Une prochaine phase du travail consiste à décomposer selon ces principes une base de données musicales plus large (de l'ordre de 200 morceaux), ce qui permettra dans le même temps de conforter et d'affiner la méthodologie proposée.

Enfin, une procédure sera définie pour attribuer des « étiquettes » aux blocs obtenus, de sorte à enrichir les informations structurelles en tenant compte des similarités entre blocs. Une piste consiste à utiliser un jeu d'étiquettes par strate et aboutir ainsi à une description structurelle multi-dimensionnelle.

6. RÉFÉRENCES

- [1] Projet QUAERO : <http://www.quaero.org>
- [2] Gabriel Sargent, Frédéric Bimbot, Emmanuel Vincent : Un système de détection de rupture de timbre pour la description de la structure des morceaux de musique. JIM 2010.
- [3] MIREX 2009 – Structural Segmentation : <http://www.music-ir.org/mirex/2009>
- [4] Geoffroy Peeters and Emmanuel Deruty : Is Music Structure Annotation Multi-Dimensional ? A Proposal For Robust Local Music Annotation. LSAS, Graz (Austria) 2009.
- [5] F. de Saussure : Cours de Linguistique Générale. 1916.
- [6] http://fr.wikipedia.org/wiki/Sémiologie_de_la_musique
- [7] Louis Hjelmslev : Prolégomènes à une théorie du langage (1943).
- [8] Fred Lerdahl & Ray Jackendoff : A Generative Theory of Tonal Music, MIT Press, 1983.
- [9] Quaero project : ID.CTC.12.R6.5.P2_Mars.pdf



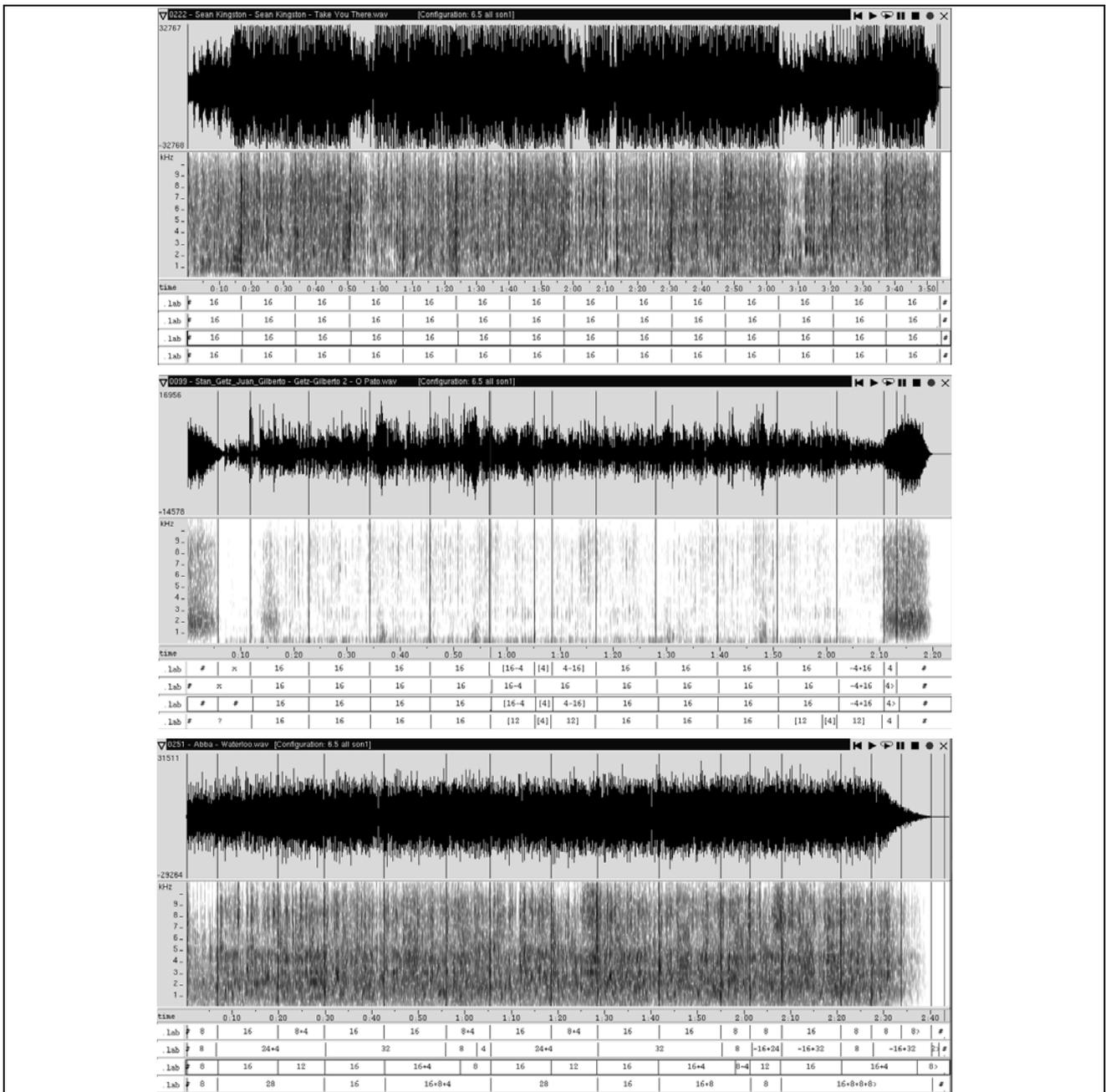


Figure 3 : Exemples d'annotations obtenues pour 3 morceaux.
De haut en bas : concordance totale (100%), moyenne (93,7%) et faible (71,0%).
Visualisation avec le logiciel Wavesurfer.

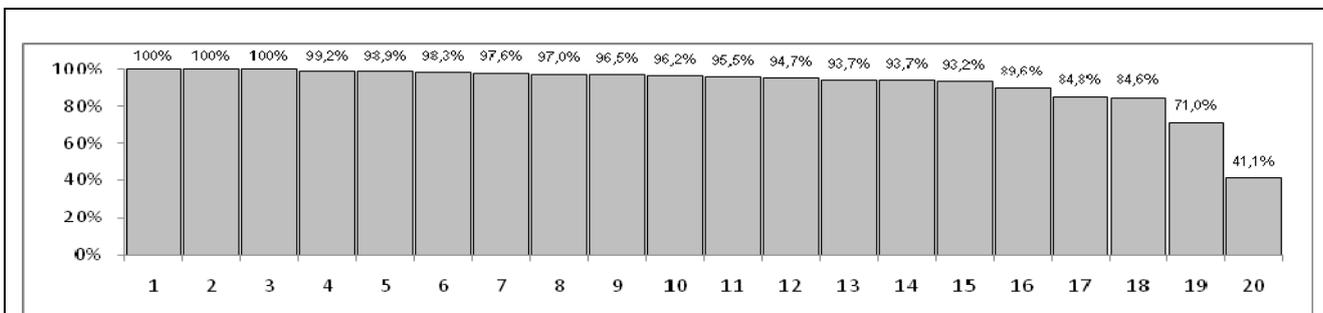


Figure 4 : Taux de concordance entre annotateurs rangé par ordre décroissant sur les 20 morceaux étudiés.

Situation	Notation	Exemple
Bloc de taille n , régulier	n	16
Bloc de taille $n-p$, obtenu par suppression de la fin d'un bloc régulier	$n-p$	16-4
Bloc de taille $n-p$, obtenu par suppression du début d'un bloc régulier	$-p+n$	-4+16
Bloc de taille $n-p$, obtenu par suppression de p snaps d'un bloc régulier	$n \setminus p$	16 \ 4
Bloc de taille $n+p$, obtenu par ajout d'un suffixe de taille p	$n+p$	16+4
Bloc de taille $n+p$, obtenu par ajout d'un préfixe de taille p	$p+n$	4+16
Bloc de taille $n+p$, obtenu par ajout d'un infixe de p snaps	$n \& p$	16 & 4
Tuilage de 2 blocs de taille n avec chevauchement de taille p	$[n-p [p] n-p]$	[12 [4] 12]
Bloc irrégulier de taille m	m	8
Bloc initial du morceau, en fade-in	$<n$	<16
Bloc final du morceau, en fade-out	$n>$	16>
Bloc de taille fractionnaire	valeur décimale	7.5
Délimitation du bloc incertaine pour l'annotateur	??	??
Taille incertaine pour l'annotateur ou qu'il ne parvient pas à déterminer	$n?$ ou $?$	16? ou ?
Taille que l'annotateur considère comme indéterminable	x	x
Segment ne faisant pas partie du morceau	#	#

Tableau 1 : Conventions d'annotation des tailles de blocs structurels.

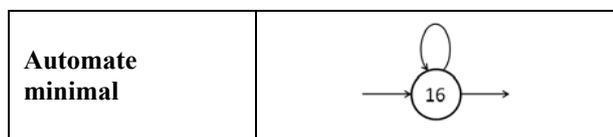
Référence du morceau : 13

Durée totale : 234 s

Taille	Nombre	Taille x Nombre
16	14	224
16	14	224

Tuilages (taille totale)	0
---------------------------------	----------

	Taille	Durée
Pulsation structurelle T1	16	16,7 s



Type d'organisation structurelle : I

Difficulté d'annotation : 1

Degré de confiance : 5

Observations : R.A.S.

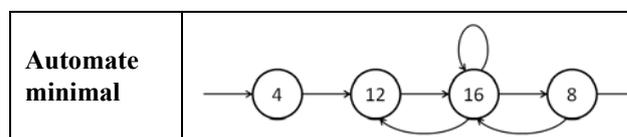
Référence du morceau : 18

Durée totale : 207 s

Taille	Nombre	Taille x Nombre
16	8	128
12	4	48
8	3	24
4	1	4
12,75	16	204

Tuilages (taille totale)	0
---------------------------------	----------

	Taille	Durée
Pulsation structurelle T1	16	16,2 s
Pulsation structurelle T2	12	12,1 s



Type d'organisation structurelle : II/III

Difficulté d'annotation : 2

Degré de confiance : 4

Observations : Les blocs de 12 peuvent éventuellement être vus comme des blocs dérivés de type 16 \ 4.

Tableaux 2a et 2b : Exemples de tableaux de bord résultant de l'annotation de 2 morceaux.

LA NOTION DE MORPHOLOGIE SONORE ET LE DEVELOPPEMENT DES TECHNOLOGIES EN MUSIQUES ELECTROACOUSTIQUES : DEUX ELEMENTS COMPLEMENTAIRES D'UNE UNIQUE ESTHETIQUE ?

Gaël Tissot

Université Toulouse-Le Mirail

gael.tissot@laposte.net

RÉSUMÉ

""La lecture des écrits de compositeurs ou d'analystes de musique électroacoustique fait ressortir l'importance que semble jouer la notion de figure, ou morphologie sonore, d'un point de vue à la fois théorique et pratique lors de la composition d'une œuvre. Le développement de ce vocabulaire correspond à l'utilisation de technologies particulières dans la création du son. Quel rôle ont alors joué ces outils, notamment informatiques, dans une pensée sur la musique en tant qu'agencement de formes ? L'étude des textes et de la pensée de trois compositeurs (Pierre Schaeffer, Iannis Xenakis et Denis Smalley) sur ce point précis permet de dégager trois angles d'approche différents, et cependant reliés à un courant de pensée similaire. La distinction matière/forme opérée par Pierre Schaeffer, la représentation graphique de formes pour Iannis Xenakis, ou les spectromorphologies de Denis Smalley ont en commun de présenter une réflexion sur la continuité/discontinuité du phénomène sonore. Ce questionnement, à la fois induit par la technologie et réalisé grâce à elle, semble caractéristique de la musique électroacoustique, et de manière plus large, de la musique de la seconde moitié du XX^{ème} siècle. Sa compréhension permet de replacer le développement des technologies musicales sous son aspect esthétique et philosophique.

1. INTRODUCTION

Le Groupe de Recherches Musicales (G.R.M.), établi à Paris dès 1958, est un centre de recherches et de composition de musique électroacoustique de renommée internationale. À travers ses activités, qui vont de la recherche scientifique sur le son à l'organisation de concerts, il réunit plusieurs dizaines de compositeurs, invités ou membres, et reflète en grande partie l'esthétique de la musique électroacoustique française [4]. Pierre Schaeffer, fondateur du groupe, est sans aucun doute le premier à avoir utilisé la notion de figure, ou morphologie, en musique. Depuis, le couple fond/forme est devenu une notion théorique centrale pour les compositeurs travaillant au G.R.M. : Iannis Xenakis, Guy Reibel ou François Bayle s'y réfèrent de manière explicite ou implicite. Pourtant, aucune recherche musicologique approfondie n'a encore été menée de ce point de vue. Il semble cependant que l'utilisation de ces termes révèle une pensée sinon nouvelle,

du moins particulière sur la musique : ils renvoient en effet implicitement à des éléments des arts plastiques. La dualité entre figure et fond, par exemple, peut être trouvée dans les œuvres de Vassily Kandinsky. Dans son livre *Point et ligne sur Plan* [5], dont le titre est déjà en rapport avec la problématique, le peintre explique la différence fondamentale entre les figures (points et lignes) et les fonds (plan). Paul Klee suit le même raisonnement dans la *Théorie de l'art moderne* [6] quand il explique le dynamisme d'une ligne en comparaison de la relative immobilité d'une surface. Une telle coïncidence ne peut être fortuite, quand on connaît l'intérêt qu'ont suscité ces artistes chez les compositeurs de la seconde moitié du XX^{ème} siècle.

L'apparition du terme « morphologie sonore » coïncide avec l'apparition de la musique électroacoustique, c'est-à-dire avec l'apparition de nouvelles technologies de création du son : enregistrement sur bande, puis développement de moyens informatiques. Quel rôle ont alors joué ces outils dans une pensée sur la musique en tant que morphologies sonores ? Ces éléments sont-ils perceptibles par l'auditeur (point de vue esthétique) ? Peut-on dire que le développement technologique a influé directement sur un élément du style ?

L'étude des écrits de Pierre Schaeffer, Iannis Xenakis et Denis Smalley devrait permettre d'apporter quelques éléments de réponse. Ces trois compositeurs se distinguent par la place importante qu'ils ont accordé à ce questionnement. En outre, la répartition de leur réflexion dans le temps (de 1948 à 1990 environ) permet d'offrir un aperçu de son évolution "dans"les"cinquante premières années de la musique électroacoustique.

2. PIERRE SCHAEFFER : LE COUPLE MATIÈRE/FORME ET LES PREMIÈRES TECHNIQUES ÉLECTROACOUSTIQUES

Pierre Schaeffer, dès 1948, note dans *Premier journal de la musique concrète : 1948-1949* à la date du 10 mai : « Si j'extrais un élément sonore quelconque et si je le répète sans me soucier de sa forme, mais en faisant varier sa matière, j'annule pratiquement cette forme, il perd sa signification ; seule sa variation de matière émerge, et avec elle le phénomène musical. » [11] Avec le technicien Francis Poullin, il conçoit dans les années 1950 un « modulateur de forme » (le morphophone), qui

permettait des modifications morphologiques par accumulation d'événements, filtrage et réinjection [16]. L'appareil était constitué d'un disque d'environ cinquante centimètres de diamètre, sur la tranche duquel était collée une bande magnétique, face enregistrable vers l'extérieur. Sur le pourtour du disque étaient disposées une tête d'effacement, une d'enregistrement et dix têtes de lecture dont la position était réglable. Chacune de ces têtes de lecture délivraient donc un signal décalé dans le temps, dont le délai était réglable. Les dix signaux étaient amplifiés et filtrés de manière indépendante, avant d'être mixés ensemble et éventuellement réinjectés dans la boucle de lecture. Il était possible d'obtenir, grâce à ce système, un son continu à partir d'un son limité dans le temps, les différents délais se superposant et masquant la reprise de la boucle. Cette « matière », au sens schaefferien du terme, pouvait alors être sujette à un travail morphologique (ajout de contours d'intensité ou de hauteur par exemple).

Dans le *Traité des objets musicaux* [12], le concept de matière/forme est précisé. Selon Pierre Schaeffer, un son pourrait être comparé à une matière possédant une certaine forme, chacun des deux aspects pouvant éventuellement être modifié par des moyens électroacoustiques.

""D'un "point" de "vue" théorique" cependant, "la" notion" de couple matière/forme est contestable d'emblée, car elle se fonde sur une hypothèse impossible à réaliser. Le fait de pouvoir « arrêter » un son est en effet physiquement impossible, étant donné la nature vibratoire de tout phénomène sonore. Les manipulations tendant à faire durer un son – mise en boucle, ou plus récemment gel d'un son –, mettent en jeu en réalité la permanence d'une ou plusieurs caractéristiques sonores, et non du son lui-même. Ainsi, la mise en boucle d'un fragment ne peut être considérée comme l'arrêt du fragment, de la même manière que l'on parlerait d'arrêt sur image au cinéma. Le phénomène de répétition n'est pas un « arrêt » en soi : il produit un nouveau son de durée indéfinie n'entretenant que certains rapports de similitude avec le son d'origine. De même, la technique de gel d'un son consiste à prolonger ses caractéristiques spectrales, caractéristiques qui ne constituent pas un son dans sa globalité : les variations d'intensité ou de grain par exemple ne sont pas prises en compte. Une « matière » sonore, existant hors du temps, est donc physiquement impossible.

En dehors de toute réalité acoustique du couple fond/forme, le concept de morphologie est intéressant en ce qu'il permet de lier une technologie et une pratique musicale. À partir d'une technologie et d'une expérience pratique apparaît une manière de concevoir la musique. Relève de la forme ce qui évolue dans le temps, alors que relèvent de la matière les éléments constants. Dans l'exemple précédent du son « arrêté » donné par Pierre Schaeffer, le timbre est effectivement un élément constant (matière), alors que l'intensité évolue (forme).

La notion de morphologie suppose intrinsèquement l'idée de contour, de délimitation d'une unité. L'idée sous-jacente de Pierre Schaeffer est le fait que ce contour est déterminé par le trajet du silence au silence. On pourrait dire que ce qui donne sa « forme » à un son est sa limite d'avec le silence. Les acousticiens décrivent généralement l'évolution en intensité d'un son en quatre phases : l'attaque, la chute, la tenue, le relâchement. On retrouve ces quatre phases sur les synthétiseurs analogiques traditionnels, généralement désignées sous le terme A.D.S.R. (initiales du nom des quatre phases en anglais : attack, decay, sustain, release). Le contour d'intensité semble donc donner une forme claire, quantifiable au son. Tout le *Traité des objets musicaux* est placé sous cette hypothèse. Cette conception place donc l'intensité au centre des préoccupations de Pierre Schaeffer. Des choix technologiques induisent une certaine esthétique.

À travers la notion de morphologie, c'est un point essentiel que soulève donc Pierre Schaeffer dès les débuts de la musique concrète : celui de variabilité continue des éléments constitutifs d'un son. Cet aspect de continuité, qui se démarque alors des instruments de musique occidentaux pensés avant tout pour utiliser des variables scalaires (hauteurs et durées principalement), fonde l'originalité de la musique électroacoustique. Ce premier aspect de la notion de morphologie a été repris par de nombreux compositeurs après Pierre Schaeffer.

Le second aspect de la morphologie développé par l'auteur, d'une égale influence sur ses contemporains, est plus tardif et moins explicite. Il concerne la problématique de segmentation d'un flux sonore. Dans le cas de la musique instrumentale, la note constitue l'élément minimal de segmentation. S'il n'est pas toujours pertinent (un *grupetto* doit-il être considéré comme une suite de notes davantage qu'un seul élément ? Un trille est-il la succession de deux notes ou est-il une seule entité ?), il offre en revanche un statut d'objectivité qui n'existe pas dans le cas de sons enregistrés. Si la musique « de notes » interroge la problématique du regroupement d'unités minimales, la musique « de sons » interroge celle de la segmentation d'un flux. Cette problématique est d'ailleurs pertinente pour l'ensemble des musiques fixées sur support, et l'on peut la mettre en parallèle avec les nombreuses recherches en cognition auditive (travaux de Stephen McAdams [8]) ou le développement récent des techniques de segmentation automatique du flux sonore.¹

3. PROLONGEMENT DE LA RÉFLEXION AU G.R.M. : FRANÇOIS BAYLE, L'ACOUSMOGRAPHE, LES UNITÉS SÉMIOTIQUES TEMPORELLES.

¹Cf. par exemple les recherches menées au *Laboratory for the Recognition and Organization of Speech and Audio (University of Columbia, New-York)* ou à la *School of Computer Science (Carnegie Mellon University, Pittsburgh)*.

3.1. François Bayle

François Bayle, directeur du G.R.M. de 1966 à 1997, poursuit la réflexion de Pierre Schaeffer en la développant sur un double terrain. D'une part, d'un point de vue théorique, le compositeur s'inspire des travaux réalisés dans le domaine de la topographie. Il reprend ainsi à son compte la théorie des catastrophes du mathématicien René Thom. La théorie s'intéresse aux limites (appelées catastrophes) d'une morphologie, c'est-à-dire, d'un point de vue mathématique, aux limites au-delà desquelles une fonction se modifie brusquement, comme si un bord avait été atteint. D'un point de vue musical, c'est bien de la problématique de la segmentation dont il est question, mais généralisée à d'autres paramètres que l'intensité (point de vue schaefferien). Un flux peut ainsi être segmenté en fonction des changements de timbre, de masse ou de sémantique.

Le concept de morphologie pour François Bayle est également l'occasion de replacer la forme sous son aspect visuel, et plus particulièrement pictural. Le compositeur est en effet amateur de peinture, et les références aux artistes plasticiens sont fréquentes dans ses écrits. Les titres des œuvres reflètent cette importance accordée à la vue. Ce n'est par ailleurs pas le seul aspect visuel que François Bayle développe : l'idée d'« image de son », ainsi que le recours à des métaphores graphiques, sont récurrents.

3.2. L'Acousmographe

L'Acousmographe est un logiciel d'écoute et de représentation du signal sonore. Il permet de dessiner, au-dessus d'une représentation physique du son (sonagramme ou forme d'onde), toutes sortes d'objets qui peuvent faciliter le repérage, l'annotation et la description d'un enregistrement [3]. Il a été conçu au G.R.M. à partir de 1991, notamment par Olivier Koechlin, Hugues Vinet et Didier Bultiauw.

La majorité des objets disponibles par défaut dans le logiciel sont des formes correspondant à la réalité physique (le dessin) du sonagramme, et non des symboles. La forme d'attaque (triangle dont la pointe est tournée vers la droite) est ainsi la même que celle observée sur un sonagramme. Il en est de même pour les lignes ou les polygones pouvant également rendre compte de phénomènes observables. Il s'agit bien d'une notation graphique, mettant la *représentation* au centre des préoccupations, plutôt qu'une écriture symbolique.² Il s'agit d'un choix purement esthétique, une notation symbolique pouvant exister, comme le propose Lasse Thoresen dans *Spectromorphological Analysis of Sound Objects* [17]. En se basant sur les recherches de Pierre

²Seuls le cercle et la spirale, que l'on ne trouve pas sur un sonagramme, pourraient correspondre à des symboles, mais leur sens n'est pas explicite. Ils ne constituent donc que de simples éléments de repérage.

Schaeffer entreprises dans le *Traité des objets musicaux*, l'auteur établit une première distinction entre trois types de son : tonique (*pitched*), ditonique (*dystonic*), et complexe (*complex*). Il attribue à cette première classification trois symboles : le cercle, le losange et le carré. Il s'agit bien ici de symboles, puisque ces trois éléments graphiques n'ont rien de commun avec le son ou l'une de ses représentations. Le choix d'une représentation graphique au G.R.M. témoigne d'un cadre de pensée liée aux morphologies sonores.

3.3. Les Unités Sémiotiques Temporelles

Le Laboratoire Musique et Informatique de Marseille (MIM), fondé en 1984 par Marcel Frémot, a mis au point à partir de 1991, en collaboration avec le G.R.M., un programme de recherche sur la problématique du temps. Ce travail a donné naissance au concept d'unité sémiotique temporelle. La problématique est toujours celle de la segmentation d'un flux, mais abordée ici sous l'angle du sens donné aux fragments. Les différentes unités portent un nom évoquant leur intention sémantique (« Chute », « Trajectoire inexorable », « En flottement »), et sont décrites en utilisant des données de durée, de phase (moments successifs et différents constituant une unité), de répétition, de matière (manière d'occuper le champ des hauteurs) et de déroulement temporel [7]. La conception s'écarte de la distinction schaefferienne morphologie/matériau, au profit d'une forme sonore pensée comme archétype, et susceptible d'être appliquée à toute musique. L'unité « En flottement » peut ainsi s'appliquer aussi bien à certaines pages de Mozart, de Debussy ou de Stockhausen. Cette conception de la morphologie comme archétype est partagée par Iannis Xenakis, qui envisage également des croisements interdisciplinaires.

4. IANNIS XENAKIS : NOTATION GRAPHIQUE, TECHNOLOGIE ET ARCHÉTYPES MORPHOLOGIQUES

Les schémas préparatoires de Iannis Xenakis, tracés sur papier millimétré (ensuite transcrits sous forme de notation traditionnelle) mettent en évidence le lien entre morphologie et son. L'écriture d'un son devient par exemple un segment de droite plus ou moins long (durée) et plus ou moins incliné (*glissando*, variation de la hauteur). Le regroupement de ces segments donne bien naissance à des formes, qui trouvent un équivalent du point de vue musical. Dans le même ordre d'idées, le C.E.M.A.Mu (Centre d'Études de Mathématiques et Automatiques Musicales), centre de recherche fondé par Iannis Xenakis, a construit dans les années 1970 la première version de l'U.P.I.C. (Unité Polyagogique Informatique du C.E.M.A.Mu), permettant de dessiner la micro et la macro-structure de la musique. Le compositeur pouvait tracer les formes d'ondes ou les enveloppes d'intensité, ainsi que des arcs, chacun associés à un oscillateur avec sa propre trajectoire de fréquence, d'intensité ainsi que sa propre forme d'onde.

La saisie des arcs était faite soit par le dessin, avec une tablette graphique, soit à l'aide d'une souris. La macro-structure (répartition des sons sur plusieurs minutes) était également dessinée grâce à la tablette graphique. Le parallèle entre morphologie sonore et forme graphique est ici clair. Les continuités ou discontinuités tracées sur l'U.P.I.C. vont résulter en continuités ou discontinuités sonores.

Néanmoins, le concept de morphologie dépasse chez Iannis Xenakis la simple représentation graphique. Il remarque ainsi : « dans chaque domaine de l'activité humaine, il y a une sorte d'écume qui est celle de la forme. J'ai remarqué des figures, des formes qui appartiennent, soit au domaine de la spéculation abstraite comme les mathématiques, comme de la logique, soit aux spéculations plus matérielles comme celles de la physique, avec ses phénomènes soit subatomiques, soit atomiques, ou comme celles des expressions géométriques de la génétique ou des réactions de ses molécules chimiques. Or, ces figures, ces formes, qui appartiennent à tant de domaines disparates, ont des similitudes ou des diversités passionnantes et qui peuvent éclairer d'autres domaines, tels que ceux des activités artistiques. » [19] Selon Iannis Xenakis, les morphologies se rapportent donc à un petit nombre de modèles qui peuvent être considérés comme universels, relevant de domaines aussi différents que la musique, les arts plastiques ou les mathématiques. L'interface technologique permet alors de transposer un phénomène d'une discipline à l'autre.

5. DENIS SMALLEY : SONAGRAMMES ET SPECTROMORPHOLOGIES

Denis Smalley a établi une réflexion importante sur la morphologie en développant le concept de spectromorphologie [13]. Celui-ci s'appuie sur les recherches de Pierre Schaeffer et les développe, en considérant la morphologie d'un son comme évolution spectrale de ce dernier au cours du temps. Le parallèle avec l'idée schaefferienne de forme est clair : il s'agit dans les deux cas de s'intéresser à l'évolution de paramètres au cours du temps. L'aspect spectral d'une morphologie permet cependant d'envisager de nombreuses caractéristiques en même temps. Il rend compte en effet à la fois des variations de localisation de masse, de son épaisseur/harmonicité et de l'intensité du son. Le travail avec les sonagrammes, rendu possible par l'informatique, permet d'exploiter ces données de manière précise.

L'auteur entreprend ensuite une typologie des morphologies. L'auteur s'appuie sur trois figures archétypiques simples, repérable de manière visuelle sur un sonagramme : attaque-impulsion (*attack-impulse*), attaque-chute (*attack-decay*) et continuité graduelle (*graduated continuant*). Différentes variations de ces morphologies de base apportent davantage de précisions. Ainsi, le mouvement « unidirectionnel/ascension » (*uni-*

directional/ascent) décrit une spectromorphologie consistant en un *glissando* ascendant du son. Le mouvement bidirectionnel/divergence (*bi-directional/divergence*) décrit une spectromorphologie dont une partie du spectre est un *glissando* ascendant et l'autre partie est un *glissando* descendant. Le mouvement « dilatation » (*dilation*) décrit un sonagramme s'épaississant au cours du temps (cas par exemple d'un son de flûte attaqué très pur, puis progressivement teinté du souffle de l'instrumentiste).

L'approche de Denis Smalley se révèle particulièrement riche pour l'analyse. La prise en compte de la perception, aidée des sonagrammes, permet une grande précision. Par ailleurs, l'idée de spectromorphologie permet – tout comme les sonagrammes – de s'adapter à différentes échelles de temps. Ceci permet de dépasser la simple description morphologique pour entrevoir la pièce d'une manière plus large. Une œuvre réussie, selon Denis Smalley, doit exploiter à la fois plusieurs de ces échelles de temps. Ce qui relève du détail doit pouvoir être mis en regard de ce qui relève des durées plus importantes. Là encore, une technologie particulière influe sur un mode de conception de la musique électroacoustique et, *in fine*, constitue une donnée esthétique.

6. LA RÉFLEXION SUR LA CONTINUITÉ/DISCONTINUITÉ : DONNÉE TECHNOLOGIQUE ET ESTHÉTIQUE CONSTITUTIVE DE LA MUSIQUE ÉLECTROACOUSTIQUE

Parmi les nombreux questionnements soulevés par l'emploi de nouvelles technologies dans le domaine musical, la problématique du rapport et de l'expression entre continuité et discontinuité tient une place majeure. Elle apparaît comme indissociable de la composition ou de l'analyse de la musique électroacoustique, quel que soit son genre [10].

En ce qui concerne la musique occidentale jusqu'au XX^{ème} siècle, la problématique de la continuité/discontinuité semble s'être manifestée principalement dans l'organisation temporelle. Le cycle musical, par exemple, propose une série de pièces autonomes (discontinuité), pourtant reliées d'un point de vue plus large (continuité du cycle). En revanche, l'aspect continu/discontinu des paramètres du son lui-même (hauteur, timbre, intensité) était relativement fixe. Ainsi la hauteur était de l'ordre du discontinu : la gamme procède par notes successives discontinues. Certes le choix du tempérament, qui vise à déterminer la manière dont ces degrés sont établis, est une donnée variable. Pour la musique du XVI^{ème} siècle par exemple, le choix d'une tonalité jouée sur un certain tempérament offrait par lui-même une palette expressive. Néanmoins, si le choix d'un type particulier de discontinuité de la hauteur était possible, il n'était pas question d'un choix entre discontinuité et continuité. En outre, le choix n'est

possible que d'une manière globale, sur l'ensemble de la pièce. Le tempérament est un cadre général, il n'a pas le même rôle compositionnel que le choix des notes ou des harmonies par exemple. Jusqu'au XX^{ème} siècle, la hauteur reste donc un paramètre fixe de type discontinu. L'utilisation de *glissandi* en parallèle avec des hauteurs fixes, ce qui permet un jeu véritable entre continuité/discontinuité de hauteur, est relativement récente dans l'histoire de la musique occidentale. Le timbre était également un paramètre traité de manière discontinue, les différents instruments de l'orchestre étant perceptibles individuellement. Les orchestrations destinées à brouiller la perception individuelle des pupitres, ou les orchestrations dont la couleur varie progressivement sont également relativement récentes. Seule l'intensité semble avoir eu la possibilité d'être traitée de manière discrète (*sforzando* par exemple) ou continue (*crescendo/decrescendo*). Dans les autres cas, les paramètres sonores ont été traités préférentiellement de manière discontinue. Le développement de la musique électroacoustique, qui s'affranchit des instruments de musique traditionnels, a apporté de nouveaux éléments de réflexion en généralisant la problématique de la continuité/discontinuité à tous les paramètres sonores.

Les possibilités techniques offertes par le matériel audio permet en effet de dépasser les limites instrumentales et de remettre en cause le traitement continu ou discontinu traditionnellement associé à chaque paramètre sonore. Si la continuité de hauteur est relativement complexe à mettre en œuvre sur certains instruments (*glissandi*), voire impossible (cas du piano), les techniques électroacoustiques permettent aussi bien un traitement continu que discontinu de ce paramètre. Il en est de même pour le timbre, dont la continuité instrumentale exige d'importants moyens d'orchestration, alors que cette continuité s'obtient relativement facilement en studio. La continuité/discontinuité devient ainsi un choix de composition, et n'est plus une donnée fixée *a priori* par l'instrument. Dans *Computer Suite for Little Boy* (1968), Jean-Claude Risset présente ainsi des sons semblant se déplacer infiniment dans le grave, cette illusion auditive créant une solution de continuité entre le grave et l'aigu.

D'autre part, le développement de la technologie a rendu aisé, dans un contexte musical, la réalisation d'un son de durée virtuellement infinie. Les limites physiques du souffle d'un instrumentiste ou de sa fatigue sont dépassées, et le son peut se prolonger tant que l'appareil est en marche. Les pièces de François Bayle évoquant la transparence jouent fréquemment sur l'emploi d'un seul son modulé de manière continue. *Transparence du purgatoire* met ainsi en jeu un seul son obtenu par synthèse analogique. Dans ce contexte, la décision d'arrêter le son prend toute son importance car elle constitue alors un élément de discontinuité. La réflexion continuité/discontinuité se trouve à nouveau posée.

Les méthodes de travail et de composition induites par les outils informatiques conduisent elles-aussà une réflexion sur la continuité/discontinuité. Ainsi, le compositeur de musique électroacoustique écoute de nombreuses fois de courts fragments de sons, afin de percevoir leurs détails et de faire un choix parmi des centaines de variations. Travailler sur un court fragment peut durer des heures, voire des jours. Les plus petits détails, non perçus durant les premières écoutes, prennent alors une importance considérable. Ce qui apparaissait au premier abord comme une unité continue est alors perçu comme une succession de moments relativement autonomes. Où se situe alors la limite entre continuité et discontinuité ? C'est au compositeur lui-même de créer sa propre échelle de continuité/discontinuité en accord avec la technologie avec laquelle il travaille.

7. LA MUSIQUE ÉLECTROACOUSTIQUE COMME PLASTIQUE SONORE ?

L'emploi du terme de morphologie par les compositeurs de musique électroacoustique induit implicitement l'idée d'une comparaison avec les arts plastiques. Ces derniers se définissent en effet comme arts producteurs ou reproducteurs de formes. La musique électroacoustique rejoindrait-elle en ce sens des arts tels que la peinture ou la sculpture ? En laissant de côté les rapprochements purement esthétiques qui peuvent être menés entre les deux domaines, trois fondements conceptuels, liés habituellement aux arts visuels, peuvent être mis en évidence dans la musique électroacoustique.

Les morphologies sonores, comme les formes graphiques, sont pensées en termes de rapports les unes par rapport aux autres, davantage qu'en données absolues. Le début de *Petite polyphonie* de François Bayle en constitue un bon exemple. Le début de la pièce ne fait intervenir que deux hauteurs (732Hz, *fa#4*, et 847Hz, *sol#4* un peu haut), le timbre est uniforme, et l'ensemble du passage est *pianissimo*. Pourtant, aucun son n'est identique à l'autre. Le discours n'est donc pas basé sur la position absolue des sons sur l'échelle des hauteurs (limitées à deux) ni sur l'échelle des intensités (uniquement *pianissimo*) ni sur une échelle de timbre (timbre uniforme). C'est au contraire la combinaison des variations internes à chaque élément, telles la courbe d'intensité ou de hauteur, qui est variée pour chaque occurrence. Ainsi, la première morphologie de la pièce propose un *glissando* de la hauteur la plus grave à la plus aiguë, avec une intensité croissante, alors que la seconde par exemple est basée sur le cheminement inverse, ou que la troisième joue sur un aller-retour entre les deux hauteurs. En ce sens, le début de la pièce résiste à toute idée d'échelle, en mettant en jeu des modifications continues de différents profils. Ce travail de variation autour d'une valeur centrale répond à certaines contraintes techniques. Les possibilités de transpositions étant limitées (puisque'elles dénaturent le son), l'intensité sonore dépendant du matériel utilisé, il est difficile de

travailler avec des données absolues. En revanche, ces contraintes se prêtent particulièrement bien à un travail de variation autour d'un point central.

D'autre part, la distinction picturale entre fond et forme semble constituer deux pôles importants dans la musique électroacoustique, correspondant à deux types de sons fondamentalement différents. La première catégorie est constituée des sons de durée virtuellement illimitée, qui offrent peu de contrastes internes, et qui sont appréhendés comme un tout, comme le remplissage d'une durée. Les sons de la seconde catégorie sont au contraire plus courts, bien délimités, et créent un balisage temporel. La perception de chacun de ces types de son est radicalement différente, ce que Christiane Ten Hoopen résume sous le terme d'« écoute polarisée » [15]. En allant plus loin, on peut postuler que leurs proportions respectives, ainsi que la manière dont leurs rapports sont envisagés, constituent une donnée esthétique propre à chaque compositeur. Ainsi, les fonds complexes dans nombre d'œuvres de François Bayle ont peu de choses en commun avec le travail des morphologies presque sans fond de Bernard Parmegiani dans *De Natura Sonorum*.

Enfin, l'idée d'une musique pensée en tant que morphologie et fond, données statiques, soulève la question du déroulement temporel d'une telle pièce. À la manière d'une toile, le développement par accumulation ou augmentation de densité de la répartition des morphologies semble y jouer un rôle prépondérant [18]. Une autre possibilité de progression est celle consistant à se faire succéder différentes parties indépendantes, constituant autant de mise en œuvre de morphologies et de fonds. Ceci peut expliquer en partie l'importance du cycle musical constitué de courte pièce dans les productions du G.R.M.

Au regard des évolutions technologiques et des particularités esthétiques, l'idée d'une musique comme plastique sonore prend tout son sens. Pierre Schaeffer, en 1952, envisageait déjà cette possibilité : « Comment demander à la musique concrète [...] de se définir comme une nouvelle musique ou comme une anti-musique ? Peut-être fallait-il la baptiser du nom de musique plastique ou de plastique sonore ? » [11].

8. CONCLUSION

L'évolution des différentes technologies musicales s'est effectuée en parallèle avec une pensée du son – à travers le concept de morphologie ou forme sonore – en tant que rapports et composition de continuités et discontinuités. La réflexion semble s'être effectuée en allers-retours, la technologie influençant le résultat sonore, de même qu'un désir sonore a pu conduire au développement d'une technologie particulière.

Ainsi, le simple fait d'enregistrer un son soulève le questionnement de la continuité/discontinuité et de la

segmentation. L'enregistrement, considéré comme *un flux*, est matérialisé par un objet *continu* : le disque, la bande, puis de manière virtuelle la mémoire informatique. Silence et son sont présents et codés de la même manière sur le support. Néanmoins, l'écoute met en évidence des discontinuités dans ce flux continu. Bien que codé de manière similaire au son, le silence délimite des éléments distincts. Dans un processus inverse, le désir de travailler sur des morphologies a conduit Pierre Schaeffer et l'ingénieur Francis Poullin à l'invention du morphophone.

Par ailleurs, le développement des interfaces homme/machine aborde la même problématique sous un angle différent. En plaçant la notion de geste au centre du questionnement, la problématique de la continuité/discontinuité est transposée à l'espace. Un geste est en effet un mouvement considéré dans sa continuité, toute rupture constituant l'amorce d'un geste différent. S'intéresser à la captation de gestes, et à leur transmission au sonore, revient à s'intéresser à un domaine de continuité et à son analyse par le capteur.

Enfin, l'apparition de moyens informatiques suffisamment rapides pour produire des sonagrammes précis en temps réel joue un rôle clé dans une pensée morphologique du son. Celui-ci est en quelque sorte figé, et la représentation résultante permet de repérer des discontinuités de manière graphique. La morphologie sonore se rapproche alors véritablement de la forme graphique, une composition musicale tendant à se rapprocher de la composition d'une toile.

Ces éléments de réflexion liés aux technologies ont été présents dès les débuts de la musique électroacoustique. Karlheinz Stockhausen, pionnier de la musique électronique, propose dans les 1950 une réflexion sur la continuité/discontinuité formelle d'une pièce. La *Momentform*, telle que définie par le compositeur dans *Momentform : nouvelles corrélations entre durée d'exécution, durée de l'œuvre et moment* [14], met en jeu des instants indépendants dont le passage de l'un à l'autre peut être traité aussi bien sous l'angle de la continuité que de la discontinuité. Pour autant, ce questionnement, bien qu'intimement lié à la musique électroacoustique, dépasse ce contexte propre. Pierre Boulez propose ainsi dans *Penser la musique aujourd'hui* [1], autant pour la musique instrumentale qu'électroacoustique, une distinction entre temps strié et temps amorphe. Dans le premier cas, il y a perception d'un balisage du temps, c'est-à-dire perception de discontinuités dans le flux sonore. Dans le second cas au contraire, toutes les transitions sont continues, sans possibilité pour l'oreille de saisir un événement discret. C'est peut-être au développement de cette réflexion sur la continuité/discontinuité que contribue le plus la musique électroacoustique, et l'on peut citer ces mots d'Olivier Messiaen : « [...] au XX^{ème} siècle où il y a eu tant de choses, et tant de progrès scientifiques [...], il y en a une qui

frappe [...], c'est la musique électronique. Je crois que c'est la principale invention du XX^{ème} siècle, et c'est probablement celle qui a le plus marqué tous les compositeurs. Parce qu'il y a des compositeurs qui font de la musique électronique, comme Pierre Henry qui est spécialiste [...], mais presque tous les compositeurs ont subi l'influence de la musique électronique, même s'ils n'en font pas. » [9]

7. RÉFÉRENCES

- [1] BOULEZ, P., *Penser la musique aujourd'hui*, Gallimard, Paris, 1987.
- [2] CASTANET, P., « La théorie des catastrophes de René Thom et l'avènement du temporel dans les œuvres musicales de François Bayle et de Hugues Dufourt », *Analyse et création musicale* Paris, L'Harmattan, 2001.
- [3] COUPRIE, P., *La musique électroacoustique analyse morphologique et représentation analytique*. Thèse de doctorat, Université de Paris-Sorbonne (Paris IV), 2003.
- [4] GAYOU, É., *GRM, le groupe de recherches musicales cinquante ans d'histoire* Paris, Fayard, 2007.
- [5] KANDINSKY, V., *Point et ligne sur Plan: contribution à l'analyse des éléments de la peinture* Gallimard, Paris, 2000.
- [6] KLEE, P., *Théorie de l'art moderne* Gallimard, Paris, 1998.
- [7] *Les Unités sémiotiques temporelles éléments nouveaux d'analyse musicale*, Laboratoire musique et informatique de Marseille, Marseille, 1996.
- [8] MCADAMS, S., *La reconnaissance de sources et d'événements sonores. Penser les sons psychologie cognitive de l'audition* Paris, Presses Universitaires de France, 1995.
- [9] MESSIAEN, O., entretien télévisé avec Alain Duault. France 3, 10 décembre 1988.
- [10] ROY, S., *L'analyse des musiques électro-acoustiques modèles et propositions*. Paris, L'Harmattan, 2003.
- [11] SCHAEFFER, P., *À la recherche d'une musique concrète*, Éditions du Seuil, Paris, 1952.
- [12] SCHAEFFER, P. *Traité des objets musicaux: essai interdisciplines*, Paris, Éditions du Seuil, 1977.
- [13] SMALLEY, D., « Spectro-morphology and Structuring Prozesse, *The Language of Electro-acoustic Music* Macmillan Press, Londres, 1986, p. 61-93.
- [14] STOCKHAUSEN, K., « Momentform : nouvelles corrélations entre durée d'exécution, durée de l'œuvre et moment », *Contrechamps*, vol.9, 1988, p. 101-120.
- [15] TEN HOOPEN, C., « Polarized Listening strategies for Electroacoustic Music », *Analizzare la musica elettroacustica : Verso una definizione degli oggetti sonori*, secondo convegno europeo di analisi musicale, Università degli studi di Trento, 1992.
- [16] TERUGGI, D., « Technology and Musique Concrete The Technical Developments of the Groupe de Recherches Musicales and Their Implication in Musical Composition », *Organised Sound*, vol.12, n°3, p. 213-231, 2007.
- [17] THORESEN, L., « Spectromorphological Analysis of Sound Objects », actes du colloque Electronic Music Studies 2006, <http://www.ems-network.org/spip.php?article250>, 2006.
- [18] VAGGIONE H., « Objets, représentations, opérations », *Ars Sonora*, n°2, p. 28-47, novembre 1995.
- [19] XENAKIS, I., *Arts/Sciences. Alliages*, Casterman, Paris, 1976.

Session 8

Temps réel et instrumentalité

TAPEMOVIE : UN ENVIRONNEMENT LOGICIEL POUR LA CREATION TEMPS REEL INTERMEDIA

Tom Mays
Didascalie.net
CICM – Université de Paris VIII
CNSM de Paris
contact@tommys.net

Renaud Rubiano
Didascalie.net
contact@renaudrubiano.com

RÉSUMÉ

Nous voulons dans ce texte présenter l'environnement logiciel *tapemovie*, programmé avec Max/MSP/Jitter). C'est un environnement pour la composition et l'improvisation musicale temps réel, les dispositifs audio/visuel interactifs, et l'art numérique dans le spectacle vivant. Il est construit sur les fondamentaux de modularité, d'optimisation, de souplesse de configuration et d'interconnectivité. Son architecture est un hub (*tapemovie*) qui agglomère des groupes de modules et centralise les configurations des sous environnements *tape* (audio), et *movie*, (vidéo/GL). Une gestion de projet permet de changer rapidement de configuration et de préférences utilisateur. La partie *tape* met en jeu des modules de traitements et d'analyses sonores connectés par une matrice. La partie *movie* implémente des sources et des traitements vidéos à la fois en traitement matriciel (*jit.matrix*)¹ et en module OpenGL². Des mappers multi-destination permettent des connexions expressives avec des contrôleurs externes ou des analyses internes. Enfin, pour les événements, l'event manager centralise le stockage des données, leur état, leur séquençement et leur évolution.

1. INTRODUCTION

Tapemovie (<http://www.tapemovie.org>) est un environnement logiciel pour la création temps réel intermédia. Il a été écrit dans Max/MSP/Jitter de Cycling'74 et existe à la fois sous forme de patch et sous forme d'application standalone fonctionnant uniquement sous MacOSX – dans les 2 cas sous licence libre LGPL³. Ses champs d'action sont : la composition et l'improvisation musicale temps-réel, les dispositifs intermédia interactifs, et l'art numérique pour le spectacle vivant. C'est un outil d'écriture et de contrôle de média en temps réel, construit par des praticiens afin de répondre à leurs propres besoins. Il vise un utilisateur prêt à acquérir une certaine compétence en informatique musicale et visuelle, et qui cherche un environnement solide et puissant lui

permettant des réalisations exigeantes, sans avoir à programmer dans Max/MSP/Jitter. Il est le fruit de 20 ans de programmation temps réel de Tom Mays (<http://www.tommys.net>), de l'expérience intensive dans l'art visuel et le spectacle vivant interactif de Renaud Rubiano (<http://renaudrubiano.com/>), de la mise en situation très poussée et les documentations du réalisateur sonore Olivier Pfeiffer, et du soutien de la plateforme *didascalie.net* (<http://didascalie.net/>). Il est utilisé depuis 2007 dans les projets de ses auteurs, dans ceux de *didascalie.net* via les créations et les accompagnements des projets, ainsi que par divers étudiants en composition au CNSMDP⁴. Tapemovie participe également à l'évaluation des prototypes du projet de recherche Virage⁵ autour des interfaces de contrôle et d'écriture.

2. HISTORIQUE

A partir des premiers environnements intégrés de contrôle et de traitement de signal, comme Max/FTS à l'Ircam pour la NeXT et les cartes ISPW au début des années 90, en passant par PureData, puis MSP qui se rajoute à Max en 1997 [1], chaque utilisateur expert cherchait petit à petit à faciliter et optimiser leur travail de développement en créant des bibliothèques de fonctions (abstractions et objets) et en réutilisant des structures de patches globaux. L'auteur, ayant fait de nombreuses créations et réalisations avec Max MIDI, Max/FTS à l'Ircam et Max/MSP depuis 1991, a commencé à formaliser de cette manière son environnement de traitement audio temps réel au début des années 2000. Il y avait déjà l'architecture modulaire, la séparation des moteurs de traitement et leurs interfaces, une matrice centrale qui permettait de connecter tout à tout, et un séquençement d'événements en forme de qlist⁶. L'arrivée de Jitter en 2003 [1] a permis de créer un environnement vidéo basé sur les mêmes principes.

A partir de 2005 ces environnements ont commencé à servir dans les créations des spectacles intermédia de la Compagnie *Incidents Mémorables* (devenue *didascalie.net* en 2008). Les membres de

¹ le format des données vidéo qu'utilise Jitter de Cycling74

² le calcul de l'image sur la carte graphique

³ Lesser GNU Public Licence

⁴ Conservatoire National Supérieur de Musique et Danse de Paris

⁵ cf. site de Virage (<http://www.virage-platform.org>)

⁶ cf. *Max Object Reference*, www.cycling74.com

l'équipe expérimentant sur divers logiciels ont voulu partir de ce travail existant pour réaliser un environnement commun répondant à leurs différentes pratiques.

Il leur fallait donc un nom. Le premier nom par défaut, *l'environnement de Tom*, à été remplacé en 2007 par l'introduction des noms actuels pour les deux parties, audio et vidéo (avec une touche d'ironie) : *tape* (Tom's Audio Patch Environnement) et *movie* (MODular Video Environment). En 2008 les deux patches *tape* et *movie* et leurs préférences ont été réunis en tant que *plugins* à l'intérieur du patch *tapemovie* – en introduisant la notion de *projet* et celle de construction dynamique des instances des modules (fonction du *build*, cf. 3.3)

Tapemovie version 1.5, prévue pour une sortie officielle en avril 2010, est le sujet de cet article. Elle est programmée dans Max 4, bien qu'elle puisse tourner dans Max5. La version *standalone* permet un fonctionnement complet sans licence Max.

3. L'ARCHITECTURE GLOBALE

Que nous utilisons *tapemovie* dans sa version *standalone* ou dans sa version de *patch* le fonctionnement reste le même. Nous allons donc plutôt parler de la version *standalone*.

Quand on ouvre l'application *tapemovie*, on voit une fenêtre qui représente le *patch* principal (Figure 1).

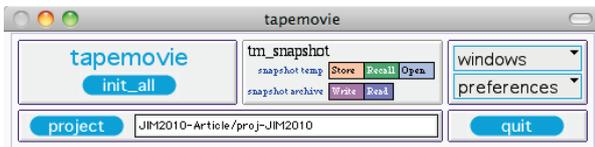


Figure 1. La fenêtre principale de tapemovie.

Ce *patch* est une sorte de « hub » qui sert à gérer le choix du *projet* (cf. 3.1 le *projet*), accéder aux réglages de toutes les configurations (cf 3.2 les configs), initier la construction par script des configurations choisies (cf 3.3 le *build*), faire un *snapshot* global de tous les paramètres, accéder aux diverses fenêtres de l'application, et également, quitter *tapemovie*. Ce patch est le niveau supérieur de l'arborescence. Il permet d'ouvrir des sous-environnements (*plugins*) *tape* et *movie*, ainsi que d'instancier divers modules communs aux deux environnements comme des contrôleurs externes, les objets de réseau, et l'event manager. (cf. 3.2.1 *tm_config*). On configure d'abord *tapemovie* puis *tape* et *movie* (si la configuration le nécessite), chacun ayant encore sa propre configuration avec ses propres modules audio et vidéo.

3.1. Le *project*

Le *project* est un dossier qui contient des sous-dossiers qui rassemblent toutes les préférences

utilisateur et les données associées à un « projet » spécifique.

Le dossier */config* contient toutes les données des choix des modules pour l'environnement, plus les données spécifiques à certains modules de traitement, ainsi que les fichiers temporaires (*snapshots*). Le dossier */events* contient uniquement des fichiers des événements et des séquences. Le dossier */instruments* contient le patch *instruments* qui est un espace de programmation « libre » interfaçable avec le reste de l'environnement (cf. 3.5). Le dossier */media* contient tous les fichiers des images et des sons par catégorie.

Par le biais du *project* nous pouvons entièrement changer les caractéristiques de l'environnement de manière souple et dynamique sans quitter le programme. Un même *project* est compatible à la fois avec la version *patch* et la version *standalone* de *tapemovie*.

3.2. Les *config*

Les *config* sont les pages de configuration de chaque partie de l'environnement – *tapemovie*, *tape*, et *movie*. Ils sont accessibles par le menu *preferences* de *tapemovie* en choisissant *tm_config*, *t_config*, et *m_config* respectivement.

3.2.1. *tm_config*

Le *tm_config* est édité dans la fenêtre *tm_configeditor*. Nous l'utilisons principalement pour choisir la quantité de chaque module disponible que nous souhaitons avoir dans notre *tapemovie*, spécifique à un projet. Les modules sont organisés par catégorie : *maps*, *osc*, *devices*, *controls* et *plugins* (Figure 2).

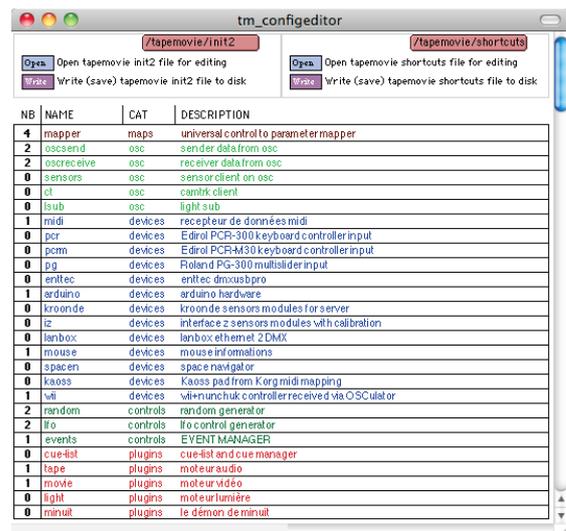


Figure 2. La fenêtre *tm_configeditor*.

Les *maps* sont les *mappeurs* centraux qui permettent de connecter n'importe quel contrôle (objet

send ou *forward* dans Max) à n'importe quel paramètre contrôlable en local ou sur le réseau (cf. 6). Les *osc* sont des modules d'envoi et/ou réception des messages OSC via UDP, comme le module *minuit*⁷ qui permet la communication avec le séquenceur Virage⁸. Dans cette catégorie sont aussi les clients « *sensors* » et « *camtrk* » qui reçoivent et distribuent les données provenant d'une application de gestion et d'analyse de capteurs et de suivi de caméra. Les *devices* implémentés dans *tapemovie* sont des contrôleurs comme *midi*, *arduino*, *kroonde*, *kaoss*, *wii* et *mouse*. Les *controls* sont des générateurs de *lfo* ou de *random*, ainsi que l'*event manager* qui est capable de stocker et d'organiser l'ensemble des paramètres. Les *plugins* sont *tape*, *movie* et *light* (gestion de lumière via DMX) On peut choisir jusqu'à 20 instances de modules normaux, mais uniquement une instance pour chaque *plugin*.

Dans le *tm_configeditor* nous pouvons également éditer une séquence d'initialisation personnalisée qui s'appelle *init2*, puis aussi un fichier de définition des *shortcuts* permettant d'ouvrir les fenêtres par une touche du clavier de l'ordinateur.

3.2.2. *t_config*

Le *t_config* fonctionne exactement sur le même principe que le *tm_config*, mais pour les modules *dsp* du sous environnement *tape* (cf. 4 pour plus info sur *tape*). Il y a, pourtant, un paramètre de plus: le choix global du dispositif de spatialisation (**Figure 3**). Il existe actuellement 3 dispositifs : *spat4* (4 haut parleurs en quadraphonie), *spat41* (*spat4* + un centre) et *spat441* (4 plus 4 plus un centre). (cf. 4.4 La Spatialisation)



Figure 3. Choix de dispositif de spatialisation

3.2.3. *m_config*

Le *m_config* fonctionne aussi sur le même principe que le *tm_config*, mais pour les modules du sous environnement *movie* (cf. 5 pour plus info sur *movie*). Les catégories sont *sources* pour les modules générant de l'image, *effects* pour les modules permettant un traitement de l'image et *outputs* pour ceux qui affiche une image, en local ou via le réseau. Les modules de la catégorie *effects* peuvent être de deux types, *jmatrix* pour les modules qui opèrent via le processeur de l'ordinateur (CPU), et *gl* pour les modules qui opèrent via le processeur graphique (GPU/OpenGL)⁹

⁷ <http://www.plateforme-virage.org/?p=1444>

⁸ cf. site de Virage (<http://www.virage-platform.org>)

⁹ cf. documentation de *Jitter* de Cycling74 (www.cycling74.com)

3.3. Le *build*

Une fois que nous avons réglé les configurations, nous lançons la construction de l'environnement par *script*. Cette fonction s'appelle un *build* et elle est lancée en cliquant sur le bouton *init* dans la fenêtre *tapemovie*. Cet étape peut prendre de 10 secondes à plus d'une minute selon les configurations.

Durant le *build*, les fenêtres *tape* et *movie* s'ouvrent (dans le cas où ils étaient choisis dans le *tm_config*). Le *build* a fini lorsque « *tapemovie ready* » s'affiche dans la fenêtre « *status* » (*standalone*) ou « *Max* » (patch *Max*). Le patch a été créé automatiquement suivant les choix de configuration de l'utilisateur.

3.4. La séparation des *moteurs* et des *éditeurs*

Un des principes forts dans *tapemovie* est la séparation des moteurs de traitement de leurs éditeurs qui ne sont pas essentiels. On peut donc ne pas les créer pour des raisons d'optimisation, si on souhaite lors d'un *build*.

Il est notre souhait dans le futur proche de s'inspirer de l'architecture modèle/vue/contrôleur pour améliorer ce système. « L'architecture *Modèle/Vue/Contrôleur* (MVC) est une façon d'organiser une interface graphique d'un programme [2]. Elle consiste à distinguer trois entités distinctes qui sont, le *modèle*, la *vue* et le *contrôleur* ayant chacun un rôle précis dans l'interface » [3]

On accède aux éditeurs par les menus *windows* de chaque partie de l'environnement.

3.5. Le patch *instruments*

Dans l'optique de faire de *tapemovie* un outil prêt à l'emploi bien que toujours ouvert à une utilisation experte, il fallait un « *endroit* » où l'on pouvait faire de la programmation *Max* librement. Le patch *instruments* se trouve dans le projet dans le dossier */instruments* (**Figure 4**). C'est un patch, pour ainsi dire, vide. On l'ouvre par le menu *windows* de *tapemovie*, ou bien automatiquement par le biais de l'*init2* – l'*init* « *user* » (cf. 3.2.1).

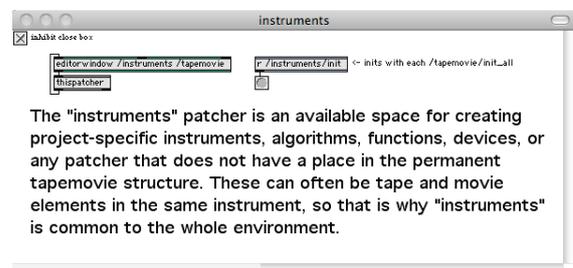


Figure 4. Le patch *instruments*

Il peut communiquer avec le reste de *tapemovie* par le moyens de *send/receive*¹⁰ traditionnel, et en audio par les *send~/receive~*¹¹ associés aux modules *aux* dans *tape* (cf. 4.2). Ainsi nous pouvons profiter de l'architecture de *tapemovie* tout en gardant la liberté et l'efficacité de programmer en parallèle dans Max pour ceux qui le souhaitent.

3.6. Le fonctionnement en réseau

Tapemovie incorpore les bases du protocole OSC¹² pour pouvoir contrôler les paramètres en dehors de l'application, ou bien en dehors de la machine. Chaque paramètre est capable également d'envoyer sa valeur en OSC par le réseau. Cet envoi de paramètres est très utile pour échanger des valeurs entre différents logiciels (compatible OSC) et aussi pour créer des interfaces sur des contrôleurs externes tels le Lemur¹³.

Il y a également une implémentation de Minuit¹⁴ par le module *minuit*. Celui ci à été créé pour pouvoir tester le Séquenceur Virage¹⁵.

3.7. Le SDK

Il est essentiel pour l'avenir de *tapemovie* qu'un utilisateur puisse incorporer ces patches dans l'environnement. Dans ce but, il existe un guide de développement qui documente le fonctionnement et la réalisation d'un module *tapemovie*. Ce SDK permet à un programmeur Max de créer ses propres modules *tapemovie*, *tape* ou *movie*.

4. LA PARTIE AUDIO : TAPE

Si le *plugin tape* a été choisi dans la *config* de *tapemovie* (cf. infra 3.2.1), une fois le *build* effectué, une fenêtre *tape* va s'ouvrir (Figure 5).



Figure 5. La fenêtre principale de *tape*.

Cette fenêtre ressemble beaucoup à la fenêtre *tapemovie*. Elle contient un bouton *init_tape* qui initialise ou construit uniquement la partie *tape*, un *snapshot* de tous les paramètres *tape*, un accès par menu (ou *shortcuts*) de tous les éditeurs *tape*, un menu pour accéder à la configuration de *tape* (*t_config*), un

bouton pour allumer/éteindre l'audio, un affichage de la cpu utilisé, puis enfin un contrôle de volume global de l'audio.

4.1. L'architecture

Les entrées et sorties audio (*adc*, *dac*) ainsi que tous les modules de traitement ou de synthèse, passent par une matrice centrale. Toutes les connexions et tous les paramètres sont manipulables par messages envoyés (cf. 7 event manager), ce qui permet une souplesse virtuellement infinie de configurations possibles.

4.2. Les modules audio

Les modules audio livrés en standard avec *tapemovie* représentent un certain choix de simplicité en se reposant sur le principe que nous pouvons construire des effets et des instruments complexes en connectant des modules entre eux par le biais de la matrice (Tableau 1. La liste des modules *tape*.), et en créant des mises en correspondance entre paramètres via les mappeurs (cf. 6).

adc	adc input
src	mono source
src2	stereo source
filt	standard filter
del	standard delay
fshift	freq shift/ringmod
harm	harmonizer module
gizmo	gizmo spectral domain harmonizer module
disto	tanh~ distortion module
mng	munger module
sfm	mono soundfile
sfst	stereo soundfile
sf4	quad soundfile
sf5	5 channel soundfile (L C R Ls Rs)
smp	mono one-shot sampler
smp4	mono 4 channel spat one-shot sampler with channel 5 to rev
vst	mono vst module
rev	reverb module based on gigaverb
ngate	noise gate
afol	amplitude follower
pt	pitch analysis
cntrd	spectral centroid analysis
zcross	zerocross analysis
delfib	8 tap delay with feedback based on fibonacci
fftilt	fft filter 253 bands
fftx	fft generalized cross-synthesis - 2 ins 1 out
fftsf	fft source-filter cross-synthesis - 2 ins 1 out
freeze	mono freeze
gran	granular synthesis
reson	model-based resonating filter bank
sndmass	soundmass fm synthesizer
synth	simple polyphonic fm synthesis
csynth	click synthesizer
nsynth	noise synthesizer
rewire	8 tracks rewire mixer
aux	auxiliary send/return (mono)
dac	dac output
Subb	dac output for sub

Tableau 1. La liste des modules *tape*.

4.3. La matrice

Une partie de la souplesse de *tape* réside dans sa matrice (éditeur: *mtrx*). Toutes les entrées audio (*adc*),

¹⁰ cf. documentation de Max de Cycling74 (www.cycling74.com)

¹¹ *Ibid.*

¹² cf. site de « open sound control » (<http://opensoundcontrol.org>)

¹³ cf. site de JazzMutant (<http://www.jazzmutant.com>)

¹⁴ cf. le site du plateforme-virage et la page spécifique à *minuit*. (<http://www.plateforme-virage.org/?p=1444>)

¹⁵ cf. site de Virage (<http://www.virage-platform.org>)

les modules de traitement et d'analyse, ainsi que les sorties audios (*dac*) y passent. L'éditeur de la matrice affiche le nom d'entrée, la valeur en dB et le nom de sortie dans chaque case. Dans **Figure 6**, nous voyons que *adc.1* est connecté à *filt.1* avec une valeur de -5 dB, puis les autres connections sont toutes éteintes, à une valeur de -127 dB.

adc.1	adc.2	src.1
-127	-127	-127
src2.1R	src2.1R	src2.1R
adc.1	adc.2	src.1
-5	-127	-127
filt.1	filt.1	filt.1
adc.1	adc.2	src.1
-127	-127	-127
filt.2	filt.2	filt.2

Figure 6. Vue d'une petite partie de l'éditeur de la matrice

4.4. La spatialisation

Chaque instance de chaque module de la configuration choisie peut accéder à la spatialisation. Ceci est rendu possible grâce au passage de tous les modules par la matrice, qui peuvent ainsi rejoindre les sorties (*dacs*) et la réverbération. Le calcul des amplitudes envoie automatiquement les bonnes valeurs à la matrice pour effectuer une spatialisation. Pour un exemple d'un éditeur de spat, dans un dispositif de 441 (cf. 3.2.2 *t_config*), regardons dans le spat du module de l'*adc* (**Figure 7**).

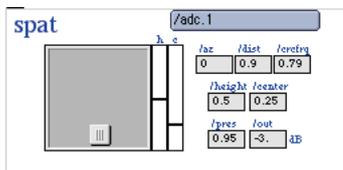


Figure 7. l'éditeur d'un spat441.

Il y a les paramètres azimuth et distance, plus un moteur de rotation (*crcfrq* = fréquence de circle), avec aussi un paramètre de hauteur (*height*) pour la quadriphonie haut, et centre pour le haut-parleur de centre. Le paramètre « pres » est pour la présence qui va effectuer le dosage du son direct et du son réverbéré.

4.5. Les modules d'analyse et le contrôle

Parmi les modules de *tape* il y a des modules d'analyse permettant le contrôle de paramètres. Il y a des modules indépendants ainsi que des modules dédiés aux *adc*. Par exemple, chaque module *adc* contient un détecteur de niveau (*noise gate*), une détection de hauteur (*pt*), un suivi d'amplitude (*afol*), un suivi du centroïde (*cntrd*), et un zérocross pour la détection du bruit (**Figure 8**).

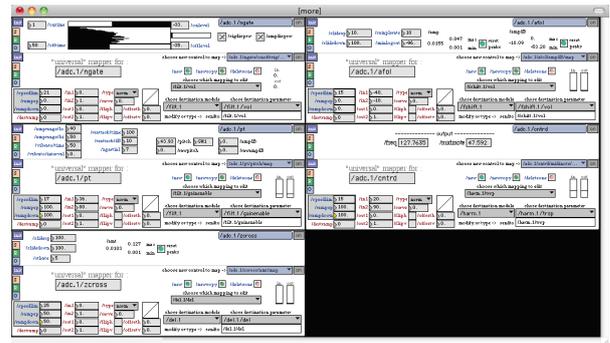


Figure 8. La fenêtre des modules d'analyse audio.

5. LA PARTIE VIDEO ET OPEN GL : MOVIE

Si le *plugin movie* a été choisi dans la *config* de *tapemovie* (cf. 3.2.1), une fois le *build* effectué, une fenêtre *movie* va s'ouvrir (**Figure 9**).

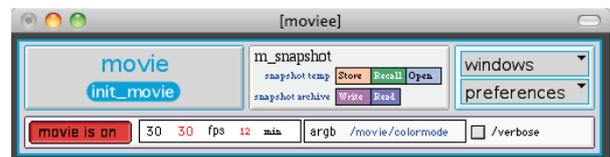


Figure 9. La fenêtre principale de *movie*.

Cette fenêtre ressemble également à la fenêtre *tapemovie*. Elle contient un bouton *init_movie* qui initialise ou construit uniquement la partie *movie*, un *snapshot* de tous les paramètres *movie*, un accès par menu (ou *shortcuts*) de tous les éditeurs *movie*, un menu pour accéder à la configuration de *movie* (*m_config*), un bouton pour allumer/éteindre le moteur des « frames », un affichage du *framerate* actuel, puis un affichage du *colormode*.

5.1. L'architecture

L'architecture est un peu plus complexe que *tape* car la chaîne de traitement vidéo passe entre calcul par *jit.matrix* (cpu) et calcul par *jit.gl.slab*¹⁶ (GPU). Il en résulte que les connections entre les modules se font par *textures* plutôt que via une matrice centrale. A part ce fait, l'architecture de *movie* fonctionne sur les mêmes bases que *tapemovie* ou *tape* permettant plusieurs espaces de rendu. Ainsi nous pouvons diriger les images en fin de chaîne vers de multiples projections ou écrans (cf. 5.4 *output*).

5.2. Les modules vidéo et GL/GPU

Les modules vidéo livrés en standard avec *movie* (**Tableau 2**. La liste des modules *movie*.) représentent le même choix d'efficacité que dans *tape*. La

¹⁶ cf. documentation de *Jitter* de Cycling74 (www.cycling74.com)

différence étant que par la nature *gl.slabs* (OpenGL) de *movie*, le matricage se fait par *textures* (cf. 5.3).

mov	jmatrix	qt movie player
cam	jmatrix	live video input
pict	jmatrix	picture player
noiz	jmatrix	noise generator
net	jmatrix	network video input
jrec	jmatrix	record matrix video to fichier Quicktime
rec	gl	record gl video to fichier Quicktime (EN COURS)
jnetsend	jmatrix	output jmatrix video to network (EN COURS)
grid	gl	gridshape module
handle	gl	handle give you position and rotation access of openGL outputs direct on screen with the mouse
boids	gl	boids generation
nurbs	gl	nurbs module
plane	gl	plane for video projection
liner	gl	lines moving with physics
mdl	gl	display 3D models in .obj format
mdlgrp	gl	display 3D models groups
mesh	gl	like a plane with crop/keystone - works with depthbuffer 1
render	gl	render space
jblur	jmatrix	blur effect
jbrcosa	jmatrix	brightness contrast and saturation corrections
jcroper	jmatrix	crop
jcrop	jmatrix	crop
jframediff	jmatrix	absolute frame difference
jrota	jmatrix	rotation correction
jslide	jmatrix	slide effect
jwake	jmatrix	wake effect
brcosa	gl	brightness contrast and saturation corrections
blur	gl	blur effect
buf	gl	dynamic video delay buffer
chromakey	gl	chroma keyer
composite	gl	compositing
crop	gl	crop and zoom
dilate	gl	dilate effect
emboss	gl	emboss effect
gaussian	gl	gaussian blur
keystone	gl	projection mapping
lumadisplace	gl	lumadisplace effect
mix	gl	binaries operators
outline	gl	outline maker
radial	gl	distorsion radial
rota	gl	rotation corrctction
sharpen	gl	sharpen effect
slide	gl	slide effect
texture	gl	texture
tllslo	gl	tllslo special module - a little bit like a channel
mask	gl	threshold maker
wobble	gl	wobble effect

Tableau 2. La liste des modules *movie*.

5.3. La matricage par texture

Il n'y a pas de représentation globale des connections entre les modules. En revanche, chaque éditeur de module de traitement *movie* contient un menu permettant de choisir une texture comme entrée (Figure 10). Nous créons ainsi l'enchaînement de modules nécessaire en choisissant la source comme paramètre pour chaque module.

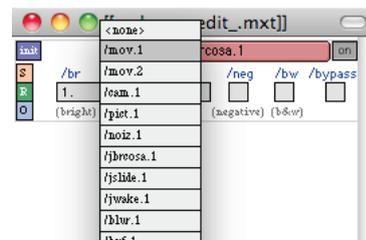


Figure 10. Le menu *texture* dans le module *brcosa*

5.4. Output

Le module de l'espace de rendu s'appelle *render* et il crée un espace de rendu (*render*) dans une sortie (*out*) qui est la fenêtre de sortie de l'image que nous pouvons placer sur un écran ou dans une vidéoprojection. (Figure 11).

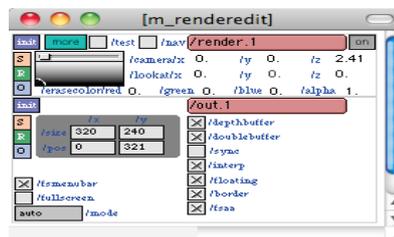


Figure 11. L'éditeur du module *render* + *out*.

6. LE CONTRÔLE ET LES MAPPEURS

Tous les *contrôleurs*, qu'ils viennent des entrées OSC ou MIDI, ou bien depuis les analyses internes (comme les suivi de hauteur ou d'amplitude se trouvant dans les *adc*) ont la possibilité de faire un « multi-mapping » (*one to many*) pour chaque paramètre de contrôle. Ces *mappeurs* ont été conçus et éprouvés par la pratique.

Ils sont situés à deux niveaux. Le premier est au niveau des modules *tapemovie*. Ce sont les *mappeurs centraux* et nous en choisissons le nombre dans la configuration pour *tapemovie* (Figure 12). Dans ces *mappeurs centraux* nous déterminons un contrôleur d'entrée (à *mapper*) puis un paramètre de destination. Dans les deux cas nous avons l'aide de menus listant les modules de contrôle ou de traitement disponible, puis un sous-menu avec les contrôleurs ou paramètres de ce module.

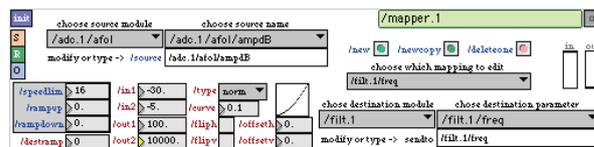


Figure 12. L'éditeur d'un *mappeur* central

Nous réglons évidemment l'échelle d'entrée et de sortie de chaque *mapping* mais également les choix de courbes entre exponentiel-linéaire-logarithmique,

symétrique autour de zéro et bosse/creux (*up/down*). Des fonctions de *invert* et *offset* permettent de rapidement configurer la courbe selon l'échelle.

Quand nous avons un *mapping* il suffit de cliquer sur *new* ou *new copy* pour en créer un autre avec le même contrôleur – vers une destination différente. En principe il n'y a pas de limite au nombre de *multi-mappings*.

Le deuxième niveau des *mappeurs* est le niveau associé aux contrôleurs directement. Dans chaque éditeur de contrôleur ou module d'analyse, il y a un éditeur unique de *mappeur* permettant de choisir dans un menu le contrôle spécifique du module en question. Le réglage de ces *mappeurs* se font de la même manière que les *mappeurs centraux*.

7. L'EVENT MANAGER

Pour stocker des états de paramètres en créant et en organisant des événements, on choisit au minimum une instance du module « événements » dans le *tm_config*. Ce module permet de créer un événement du groupe *tapemovie* (tous les paramètres), du groupe *tape* (paramètres d'audio) ou du groupe *movie* (paramètres du traitement de l'image). Les « events » sont stockés dans le dossier *events* du projet en tant que fichier texte, mais on peut aussi créer des sous-dossiers pour les différencier. Une fois créé, on peut faire « play », « edit », « rename » ou « delete » d'un événement. (Figure 13).

Pour en exécuter un, on le sélectionne dans le menu *play*. A ce moment, le fichier de l'événement sera lu depuis le disque. Puisqu'ils sont relus avant chaque lecture, on peut les modifier avec n'importe quel éditeur de texte à l'extérieur de *tapemovie*, et l'event manager va relire la version modifiée automatiquement depuis le disque.

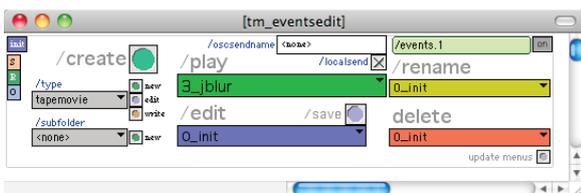


Figure 13. l'éditeur des *events*.

8. LA PÉRENNITE DES CRÉATIONS

Une des questions la plus importante dans la création avec les outils informatiques temps réel est comment rejouer l'oeuvre dans le futur. Les pièces du répertoire doivent subir des réactualisations (portages) régulièrement, sinon elles risquent de ne plus pouvoir être jouées à cause des supports de stockage, l'obsolescence des standards techniques, l'absence de soin dans la préservation des parties électroniques des oeuvres, les conditions pratiques de la performance, et la confection souvent empirique des programmes [4].

Tapemovie aborde le sujet de la pérennité des créations par son recours aux fichiers texte qui définissent les événements. Chaque événement est une description d'un état ou une action incluant tous les paramètres nécessaires à son exécution. Ces fichiers texte sont lisibles en dehors de l'application par un éditeur de texte.

Si on combine la lisibilité des événements avec la documentation de l'environnement de *tapemovie* on peut dire que dans plusieurs années il serait encore possible d'étudier la documentation de *tapemovie* puis déchiffrer les configurations et événements d'un *project* pour, avec les fichiers sons, re-crée l'oeuvre dans un environnement futur – sans jamais avoir à faire tourner *tapemovie*.

9. CONCLUSIONS

Tapemovie est utilisé dans la composition, l'improvisation et le spectacle vivant. Il peut servir à la fois le créateur individuel et toute une équipe avec des ordinateurs en réseau. Il a une capacité d'extension par sa nature open-source et son utilisation par différentes équipes dans des spécialités variées. En servant une pluralité de projets, *tapemovie* s'augmente aussi en terme de configurations possibles.

Les réflexions entre les développeurs ont commencé pour les futures versions créées en Max 5. La prochaine est en préparation en collaboration avec les développeurs de Jamoma¹⁷

Tapemovie est un environnement de création qui tend la main vers une utilisation intuitive, tout en gardant la spécificité d'un outil expert, solide et ouvert.

10. RÉFÉRENCES

- [1] Manning, Peter, *Electronic and Computer Music*, Oxford University Press, New York, 2004, p. 361-374
- [2] Reenskaug, Trygve, 2003, *The Model-View-Controller (MVC) Its Past and Present*, article publié par INTEGRATED DOMAIN SERVICES, http://heim.ifi.uio.no/~trygver/2003/javazone-jao/MVC_pattern.pdf
- [3] Le site du *Laboratoire d'Informatique Algorithmique*, la page sur l'architecture MVC : (<http://www.liafa.jussieu.fr/~carton/Enseignement/InterfacesGraphiques/MasterInfo/Cours/Swing/mvc.html>)
- [4] Alain Bonardi et Jérôme Barthélemy, 2008, «Le patch comme document numérique : support de création et de constitution de connaissances pour les arts de la performance», CIDE [En ligne], CIDE 10, Session Document culturel, mis à jour le 16/09/2008, URL : <http://172.16.128.67:50010/cide/index.php?id=298>

¹⁷ cf. site officiel Jamoma (www.jamoma.org)

MULTICURVE : UN NOUVEL OBJET GRAPHIQUE DE GESTION DE DONNEES DE CONTROLE DANS MAX/MSP.

Jean-François Baud

CICM-Centre de recherche Informatique et
Création Musicale
Maison des Sciences de l'Homme Paris Nord
Université de Paris8
cicm@univ-paris8.fr

RÉSUMÉ

Dans cet article, nous décrivons un objet Max/MSP proposant de nouvelles stratégies de visualisation, de stockage, d'édition et de "lecture" de données de contrôle. Ce développement logiciel s'appuie sur une recherche et des travaux concernant l'interfaçage du corps humain avec un environnement interactif temps réel. Cet objet trouve sa place dans une démarche interdisciplinaire de la sphère Art, Science, Technologie, visant à implémenter certains outils scientifiques sur une plateforme logicielle utilisable sur le terrain des Arts interactifs, telle que Max/MSP. Notre conclusion porte sur la consistance d'une telle démarche, l'utilité d'un tel objet et les perspectives de développement futures.

1. INTRODUCTION

Ce texte relate un développement logiciel au service d'une démarche artistique en interaction avec des exigences scientifiques et technologiques.

Cet objet Max/MSP a été développé en fonction de deux problématiques. En premier lieu proposer une solution répondant à l'absence d'un objet multi-courbes performant dans Max/MSP. En second lieu offrir une interface de contrôle très fine et supportant une fréquence d'échantillonnage proche de la milliseconde. Ce développement fait partie d'une démarche d'intégration des outils et des technologies scientifiques sur des plateformes logicielles utilisables sur le terrain des Arts interactifs.

J'exposerai tout d'abord le contexte général au sein duquel ce projet a été réalisé. Je présenterai premièrement l'approche théorique originale, celle-ci prend place dans un cadre de recherche en sciences humaines, sciences de l'art, spécialité musique, alliant la création artistique, les sciences et les technologies¹. On s'intéressera également au cheminement expérimental, on insistera sur la collaboration pluridisciplinaire et sur les précédents projets à l'origine du développement de

cet objet. Je ferai aussi un bref état des lieux des objets Max/MSP destinés à la représentation graphique de courbes de valeurs, j'insisterai en particulier sur les limites de ces objets et l'intérêt de développer un objet plus performant.

J'introduirai ensuite l'objet Multicurve. J'exposerai en premier lieu son architecture générale, ses fonctions graphiques et ses fonctions relatives à la gestion de données. Je m'attacherai ensuite à expliquer le mécanisme de certains algorithmes lui permettant de réaliser des performances intéressantes au niveau de la consommation des ressources du processeur. J'évoquerai finalement les résultats des tests effectués avec Multicurve afin de quantifier ses performances techniques.

L'article ci-dessous présente le développement d'un objet Max/MSP à travers le contexte d'expérimentation particulier dans lequel il trouve son origine puis à travers un angle plus technique et descriptif. Il conclut sur des questions d'interaction entre les outils technologiques scientifiques et artistiques, et la place d'un tel développement dans cette démarche.

2. LE CONTEXTE À L'ORIGINE DE CE DÉVELOPPEMENT LOGICIEL

2.1. Le contexte théorique

Dès 2008 je me suis intéressé au développement d'interfaces logicielles et physiques permettant la captation de signaux corporels et leur utilisation comme contrôleurs sonores² [1]. C'est dans le cadre de ce sujet que je me suis orienté vers les sciences cognitives. Cela m'a permis d'étudier de plus près le fonctionnement "sensori-moteur" de l'être humain qui est, à mes yeux,

¹ Licence, Master et Doctorat au département de musique de l'université de Paris 8, Ecole Doctorale "esthétique, sciences et technologies des arts." <http://artweb.univ-paris8.fr> et www.edesta.fr,

² Jean-François Baud, Anne Sèdes, « *Vers une approche physiologique de l'interface corporelle pour les arts interactifs* », prix du jeune chercheur de l'Association Française d'Informatique Musicale, JIM 2009.

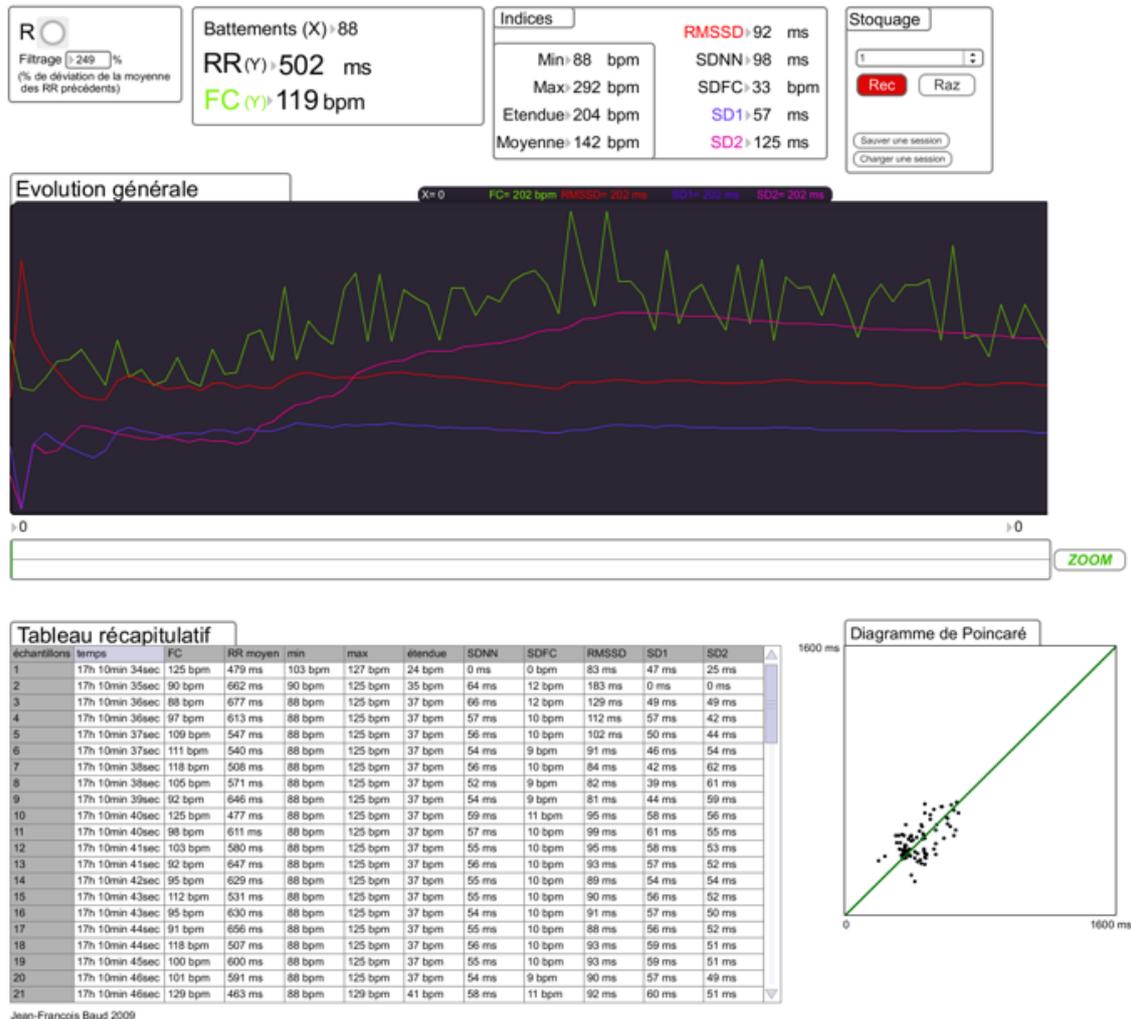


Figure 1: fenêtre principale du logiciel Varialog

l'élément principal au sein de tels dispositifs interactifs. Je me suis alors plus particulièrement tourné dans la direction des théories de Francisco Varela [2] concernant l'enaction et la corporéité (l'embodiment). Ces théories exposent le lien d'interaction particulier entre les processus cognitifs et leurs conséquences physiques ou physiologiques. Cette thématique est également reprise par Alain Berthoz qui y ajoute la problématique de la place des émotions au sein de ce processus. Dans son ouvrage *La décision* [3] il aborde la question du lien entre les émotions et la décision et plus précisément l'ancrage somatique des émotions. On y apprend que les premières conséquences corporelles de la perception seraient détectables dans la régulation des fonctions automatiques de l'organisme et seraient donc significatives de l'implication émotionnelle du sujet de la perception.

M'appuyant sur ce postulat, j'ai donc adopté comme idée directrice qu'un dispositif d'interfaces physiques et logicielles captant et analysant la régulation de certaines fonctions automatiques de l'organisme nous permettrait d'obtenir des informations sur l'état psychologique, émotionnel du sujet de la captation. Ces informations intéressantes seraient ensuite exploitées comme contrôleurs d'un dispositif multimédia.

2.2. Une collaboration pluridisciplinaire

L'analyse de la régulation des fonctions automatiques de l'organisme étant un objet d'étude important de la physiologie, j'ai alors entrepris une collaboration avec l'équipe de chercheurs de la plate-forme de recherche en Sciences et Techniques des Activités Physiques et Sportives (Staps) de l'UFR de Santé, Médecine, Biologie Humaine de l'Université de Paris 13, laboratoire "Réponses cellulaires et fonctionnelles à l'hypoxie". Cette équipe, animée par Aurélien Pichon, étudie plus particulièrement la variabilité cardiaque et son interprétation à travers le prisme de la physiologie et de la psychologie [4] [5].

C'est au sein de cette collaboration que j'ai réalisé le logiciel Varialog me permettant de m'approprier les outils d'analyse de la variabilité cardiaque des physiologistes et de les intégrer sur la plate-forme logicielle Max/MSP utilisable sur le terrain des Arts interactifs. Le logiciel fut finalisé en mars 2009.

2.3. Un premier prototype d'objet de visualisation de plusieurs courbes

Ce logiciel d'analyse temps réel de la variabilité cardiaque permet de calculer et d'enregistrer les valeurs des intervalles temporels entre chaque battement cardiaque et d'opérer une série d'analyses à l'aide de différents indices statistiques, l'évolution des quatre indices principaux est représentée en temps réel sous forme de courbes, ces valeurs sont ensuite stockées dans un tableau.

C'est lors de la programmation de la fenêtre de visualisation des quatre courbes que je me suis heurté à une difficulté technique, en effet Max/MSP ne dispose pas d'un objet efficace permettant la visualisation de plusieurs séries de valeurs qui s'implémentent en temps réel. Pourtant c'est une fonction principale présente dans la plupart des logiciels destinés à l'usage scientifique, tel que MatLab. Techniquement le logiciel réalise cela en augmentant automatiquement la taille de l'échantillon pour chaque nouvelle valeur reçue, la ou les courbes sont alors redessinées en ayant donc été incrémentées d'un nouveau point. Ce procédé permet visuellement d'avoir en permanence l'ensemble de l'échantillon représenté et donc de percevoir l'évolution globale des valeurs dans le temps.

Je me suis alors tourné vers les objets existants pour trouver une solution alternative.

2.4. Les limites des objets Max destinés à la représentation graphique de courbes de valeurs

Les objets principaux destinés à cet usage sont Multislidier, Itable et Function. Ils sont cependant limités à une courbe³⁴ et ne peuvent être superposés du fait de l'absence du mode "transparent". L'objet Lcd ne peut également pas répondre à cette problématique car il ne garde aucune mémoire de quelconques valeurs, il est seulement destiné à la représentation graphique.

Pour résoudre ce problème technique j'ai donc couplé plusieurs Multislidiers avec un Lcd, ce qui m'a permis d'utiliser la mémoire des Multislidiers conjointement à la capacité graphique du Lcd. Cette solution alternative, difficile à programmer mais opérationnelle, m'a permis de finaliser mon programme. Elle présente cependant des inconvénients majeurs :

Une sollicitation du processeur trop importante et une limitation du nombre de valeurs maximales utilisables :

³ L'objet Multislidier en mode "scroll" peut afficher plusieurs courbes simultanément cependant elles n'apparaissent pas dans le même repère spatial, il s'agit en fait d'une superposition de repères spatiaux ce qui ne répond pas à notre problématique technique. De plus le mode "scroll" n'augmente pas la taille de l'échantillon pour chaque nouvelle valeur reçue, mais efface les valeurs les plus anciennes.

⁴ Dans la bibliothèque FTM développée par l'IRCAM il existe l'objet Vectorsdisplay qui permet la visualisation de 5 courbes simultanément. Il fonctionne soit avec des *list* de *floats* soit avec les autres objets de FTM. Je n'ai pas retenu cet objet car ce système de *list* m'a paru inadapté à mon utilisation (une *list* est limitée à quelques centaines de valeurs), de plus la prise en main de cette bibliothèque est relativement complexe.

De manière générale, tous ces objets Max destinés à la représentation graphique de valeurs "consomment" un pourcentage élevé des ressources du processeur lors de leur fonctionnement. Cette consommation est évidemment proportionnelle au nombre de valeurs stockées dans la mémoire des objets.

En effectuant des tests sur ces objets, je me suis rendu compte des limites qu'ils présentaient réellement : éditer la courbe d'un Multislidier au maximum du nombre de ses sliders (32767) ou éditer une Itable de 10 000 valeurs ou bien encore écrire rapidement (toute les 10 ms) de nouveaux points dans un objet Function représente une consommation de 60% à 80% du processeur⁵. Cette consommation est trop importante pour de simples opérations de gestion de données de contrôle, de plus elle présente de réels risques de *crash* de Max/MSP.

On peut donc considérer que ces objets Max permettent une gestion et une visualisation efficace de données dans la limite d'une courbe de quelques milliers de valeurs au maximum. Ils ne sont donc pas compatibles avec un échantillonnage trop rapide ou avec une durée d'enregistrement trop longue⁶.

J'ai alors entrepris le développement d'un objet dépassant ces limites de fonctionnement et répondant aux exigences du contexte scientifique et artistique dans lequel se trouvent mes travaux.

3. PRÉSENTATION ET FONCTIONNEMENT DE MULTICURVE

3.1. Présentation générale et construction des courbes

Dans un patch Max/MSP, on peut lancer Multicurve en inscrivant son nom dans une boîte objet ou en le sélectionnant directement dans la palette des objets graphiques. Il se présente sous forme d'une fenêtre noire comprenant 5 entrées et 5 sorties. Les 4 premières entrées reçoivent 4 flux de données de contrôle qui correspondent aux 4 courbes représentées dans la fenêtre de l'objet (par défaut l'entrée 1 correspond à la courbe verte, le second à la courbe rouge, le troisième à la courbe bleue puis le dernier à la courbe violette⁷). Ces flux de valeurs sont échantillonnés en fonction de l'arrivée de nouvelles valeurs dans l'entrée 1 à la façon d'un objet Pack, autrement dit dès que l'entrée 1 reçoit une nouvelle valeur il "autorise" les 3 autres entrées à intégrer de nouvelles valeurs. Ce procédé permet de donner à ces 4 flux de données la même temporalité. Il convient donc de lier à l'entrée 1 le flux de données le plus rapide afin d'avoir la meilleure résolution temporelle.

⁵ Ces tests ont été réalisés avec un MacBook doté d'un processeur de 2.16 GHz, 1 Go de RAM, Mac 10.6.2 et MaxMSP 5.1

⁶ Les limites pour une consommation de CPU acceptable (aux alentours de 30%) se situent en réalité à 5000 valeurs maximales par objet et une écriture de nouveaux points tout les 200ms au minimum.

⁷ Les couleurs des courbes ainsi que celles de tous les repères visuels sont paramétrables dans l'inspecteur de l'objet.

Après l'échantillonnage, les quatre courbes sont donc incrémentées d'un nouveau point correspondant au "paquet" des 4 valeurs échantillonnées. Graphiquement, pour chaque nouveau "paquet" toutes les courbes sont redessinées et auto-ajustées à la taille de la fenêtre, ce système permet visuellement d'avoir en permanence l'ensemble de l'échantillon. L'axe des abscisses est donc ajusté dynamiquement, les ordonnées par contre correspondent à une échelle, un *range* paramétrable dans l'inspecteur.

La sortie 5 exprime le nombre total de points par courbe (soit le nombre total de "paquets" ou encore la taille de l'axe des abscisses).

L'entrée 5 quant à elle est un index de lecture des courbes. Elle considère chaque valeur reçue comme une abscisse et renvoie respectivement dans les sorties 1, 2, 3, 4 les ordonnées des quatre courbes, correspondant à cette abscisse.

3.2. Repères visuels

Lors du développement de l'objet, un critère m'a paru très important : le caractère intuitif de l'interface graphique. C'est dans cet esprit que j'ai implémenté l'objet de nombreux repères visuels dynamiques facilitant la lecture des courbes. Le principe est d'amener certaines informations de lecture des courbes à l'endroit où l'utilisateur focalise son attention.

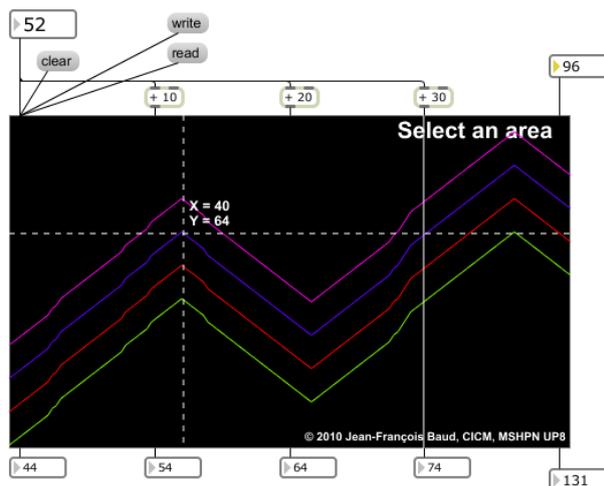


Figure 2: fenêtre de visualisation de Multicurve et repères visuels.

Par exemple lorsque la souris est à l'intérieur de la fenêtre, ses coordonnées sont exprimées dans l'échelle des courbes et suivent visuellement la trajectoire du curseur. Ce procédé permet de prendre connaissance des coordonnées d'un point simplement en plaçant le curseur sur celui-ci et donc sans se référer à un repère fixe tel un axe des abscisses. Notre repère de coordonnées est donc dynamique et plus intuitif pour l'utilisateur. Deux axes pointillés (un vertical et un horizontal) sont également associés à la trajectoire du curseur dans la fenêtre, ils facilitent la lecture des courbes et en particulier la

comparaison entre les courbes ou entre les différents "pics" des courbes.

Lors de la lecture des courbes, l'abscisse envoyée dans l'entrée 5 est matérialisée dans l'objet par une barre verticale ce qui permet de suivre visuellement cette lecture. Ceci est d'autant plus important lors d'une lecture rapide⁸.

Dans le cas d'une échelle des ordonnées comprenant des valeurs négatives et positives, un axe correspondant à $Y=0$ est matérialisé par une barre horizontale.

3.3. Modes de fonctionnement

Comme nous l'avons vu, Multicurve interagit avec la souris simplement lorsque celle-ci est à l'intérieur de la fenêtre de visualisation. Les "clicks" de la souris sont également pris en compte puisqu'ils donnent accès à deux modes de fonctionnement. Ces modes correspondent à deux fonctions principales de l'objet : la visualisation et l'édition précise de plusieurs courbes de données.

Lorsque le curseur entre dans la fenêtre le message "Select an area" apparaît en haut à droite de l'écran, il suffit alors de sélectionner une zone en cliquant pour opérer un zoom sur celle-ci. Sur cette fenêtre sélectionnée, il est alors possible d'effectuer deux opérations : zoomer ou éditer. Pour zoomer il faut procéder comme précédemment, en sélectionnant une zone. Il est possible de zoomer jusqu'à sélectionner une zone composée seulement de quelques points. Pour revenir ensuite à la fenêtre de visualisation de l'ensemble de l'échantillon, il faut simplement faire sortir le curseur de la fenêtre de l'objet.

Pour éditer une courbe il faut placer le curseur sur un de ces points, la courbe apparaît alors en surbrillance, il suffit alors de maintenir enfoncée la souris et de rectifier la courbe.

Le mode de fonctionnement dans lequel on se trouve est indiqué en haut de la fenêtre à droite : "ZOOM" ou "EDIT".

Il est possible de réaliser ces opérations à tout moment, même durant l'écriture ou/et la lecture des courbes.

⁸ Un simple objet Metro associé à un Counter lié à l'entrée 5 permet de réaliser un module de lecture à vitesse variable, contrôlable à la milliseconde près.

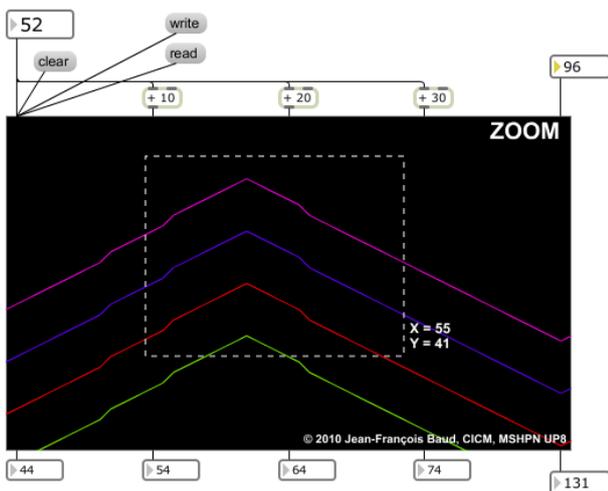


Figure 3: Multicurve en mode zoom.

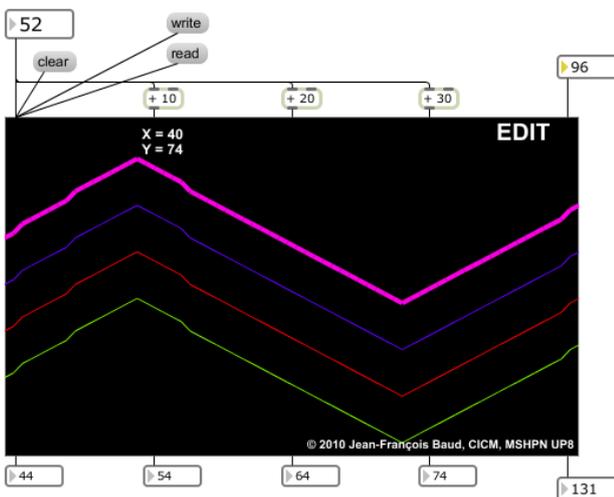


Figure 4: Multicurve en mode édition.

3.4. Messages et fonction de stockage

Multicurve est sensible à trois messages venant de Max/MSP. Le message “Clear” efface toutes les données implémentées dans l’objet et par conséquent les courbes. Le message “Write” active une fonction de stockage des courbes sous forme de fichier. Plusieurs fenêtres correspondant au nombre de courbes s’affichent successivement et demandent à l’utilisateur de nommer le fichier contenant les valeurs de la courbe puis de le placer à un endroit du disque dur. Le message “Read” effectue l’opération inverse en chargeant dans l’objet des courbes sauvegardées sous forme de fichier. L’utilisateur peut ainsi sauvegarder ses données facilement mais également charger dans l’objet des courbes appartenant à différentes “sessions d’enregistrement”, si les courbes sont de longueurs différentes, un ajustement sera fait en fonction de la plus grande.

Bien que le nombre d’entrées et sorties soit constant, il est possible de paramétrer le nombre de courbes dans l’inspecteur, dans la limite de 4 maximum. Ce paramètre est intéressant car il permet de réaliser un gain de CPU suivant le nombre de courbes.

3.5. Caractéristiques techniques, tests et performances de l’objet

Comme je l’ai précisé, l’intérêt d’un tel développement est d’obtenir un objet plus performant sur les critères d’économie de ressources processeurs, de fonctionnement avec un grand nombre de valeurs, un échantillonnage très rapide et la possibilité d’écrire plusieurs courbes.

L’optimisation de la consommation des ressources CPU est donc le critère central de ce développement. J’ai résolu ce problème technique grâce à deux implémentations. La première est un mécanisme de limitation de la fréquence de rafraîchissement de la fenêtre graphique de l’objet, cette fréquence est calculée dynamiquement et assure une faible consommation de CPU sans altérer la “réactivité graphique” de l’objet. La seconde implémentation est un algorithme qui s’active lorsque le nombre de points des courbes est plus important que le nombre de pixels de la fenêtre graphique, pour réduire alors la taille des courbes il effectue alors une “mise en paquets” des valeurs de la courbe puis effectue des moyennes de ces paquets qu’il propose alors comme les nouvelles courbes. La taille des paquets est définie dynamiquement dans un rapport de proportionnalité au nombre total des valeurs brutes. Cet algorithme permet un rendu visuel des courbes le plus fidèle possible à leurs valeurs brutes.

Grâce à ces implémentations Multicurve consomme peu de ressources processeur et répond aux critères de développement évoqués plus haut. J’ai effectué de nombreux tests tout au long du développement en le poussant dans ses limites de fonctionnement, d’une manière générale il ne dépasse quasiment jamais 30% de CPU. Plus précisément, pour un objet de quatre courbes et une fréquence d’échantillonnage à 1 ms, sa consommation se situe tout d’abord aux alentours de 20% de CPU puis elle dépasse les 30% lorsque l’objet contient un million de valeurs par courbe. J’ai donc borné l’objet à 1 500 000 valeurs par courbes pour rester sous les 40% de CPU.

Lors des opérations d’édition et de zoom, même avec un million de valeurs, la consommation ne dépasse jamais les 20%, et toujours dans ces mêmes conditions si on lit les courbes à une vitesse d’un X par ms le processeur consomme aux alentours de 25% de ses ressources.

4. CONCLUSION

Dans la suite de ce développement, je pense étudier certaines possibilités d’optimisation intéressantes afin de les implémenter à Multicurve. Une première idée d’optimisation serait de donner à chaque courbe sa

propre échelle, son propre axe des ordonnées. Cela permettrait visuellement de faire cohabiter des courbes ayant une échelle de valeurs très différente.

Il serait ensuite intéressant d'implémenter un mode "oscilloscope" permettant de visualiser les flux de données de manière dynamique.

Il conviendrait également de pouvoir "faire glisser" une fenêtre de zoom, afin de naviguer dans tout l'échantillon seulement avec cette fenêtre choisie.

La version actuelle de Multicurve est téléchargeable sur le site du CICM⁹.

Dans cet article, le développement de cet objet a été présenté au sein d'une démarche de récupération et d'adaptation des outils scientifiques à un contexte artistique¹⁰. Cette démarche se veut au service des Arts Interactifs, considérant que ces contributions scientifiques et technologiques nouvelles sont porteuses de consistances artistiques.

Dans cette approche, la captation n'est pas simplement considérée comme génératrice d'un flux de données de contrôle dynamiquement intéressant à exploiter. Elle est avant tout considérée comme un moyen d'interfacer un aspect de la réalité et d'en restituer, après une phase d'analyse, un flux de données de contrôle soulignant cette réalité. Par exemple une pulsation cardiaque exploitée telle quelle a un potentiel expressif limité car ses variations sont assez régulières et ont peu d'amplitude, cependant si on l'exploite avec des outils statistiques analysant la variabilité, on obtient alors un flux de données beaucoup plus expressif parce que soulignant, exprimant les moindres variations subtiles du rythme cardiaque brut.

Cette approche d'appropriation des outils scientifiques est donc au service de l'expressivité. Multicurve est un objet se situant dans cette démarche car il répond aux exigences de ces outils d'analyses scientifiques, en particulier sur le terrain de la précision et de la résolution temporelle.

Le développement de cet objet nous permet ainsi d'exploiter les relations fructueuses entre arts, sciences et technologies, dans un contexte profondément interdisciplinaire.

5. RÉFÉRENCES

- [1] Jean-François Baud, Anne Sèdes, "Vers une approche physiologique de l'interface corporelle pour les arts interactifs", JIM 2009.
- [2] Francisco Varela, Evan Thompson, Eleanor Rosch. *L'inscription corporelle de l'esprit*, sciences cognitives et expérience humaine, p. 235 Seuil, Paris, 1993.
- [3] Alain Berthoz : *La décision*, Odile Jacob, 2003, Paris.
- [4] Aurélien Pichon, Claire de Bisschop, Véronique Diaz, André Denjean : "*Parasympathetic airway response and heart rate variability before and at the end of methacholine challenge*", Chest, 2005.
- [5] Aurélien Pichon, Chapelot Didier, Nuissier Frédéric : "*Heart rate variability is an indicator of changes in psychological states over a year in university students*", psychophysiologie, 2008.

⁹ <http://cicm.mshparisnord.org/>

¹⁰ Cet objet sera d'ailleurs utilisé pour effectuer des analyses de variabilité cardiaque sur des instrumentistes en diverses situations d'improvisations, dans le cadre d'un projet dirigé par Cécile Vallet de la MSH Paris Nord.

L'INSTRUMENT NUMÉRIQUE COMME OBJET D'ANALYSE DES MUSIQUES MIXTES

Frédéric Dufeu
MIAC/APP, Université Rennes 2
frederic.dufeu@gmail.com

RÉSUMÉ

L'intervention des technologies numériques dans les lutheries des xx^e et xxi^e siècles ayant contribué à déplacer une partie de l'écriture compositionnelle au sein même de l'instrument de musique, l'approche poétique des œuvres dites mixtes doit accorder une attention particulière à la constitution des dispositifs d'exécution. À partir d'exemples notoires de programmes informatiques inscrits dans deux domaines différenciés, les musiques écrites et l'improvisation, cet article s'attache à montrer comment les instruments numériques peuvent être considérés comme des instruments d'analyse musicologique. D'une part, l'examen de fragments de l'environnement informatique d'une œuvre importante du répertoire du temps réel, *Jupiter* de Philippe Manoury, permet de saisir l'interaction mise en place entre partie instrumentale et partie électroacoustique et de compléter ainsi la lecture de la partition et de ses écrits adjacents. D'autre part, le programme OMax, destiné à l'improvisation avec ordinateur, peut se prêter à des ajustements offrant, moyennant un travail de lutherie de la part du musicologue, des perspectives utiles pour l'analyse de l'improvisation avec dispositif numérique.

1. INTRODUCTION

Au siècle dernier, l'intervention des technologies numériques dans le développement de nouvelles lutheries a largement contribué à questionner la notion d'instrument de musique. Dans un article intitulé « De l'amplification à l'autonomie de l'instrument : ruptures électroniques et conséquences esthétiques » [7], Bruno Bossis remarque que

« la lutherie s'est non seulement enrichie d'interfaces gestuelles nouvelles, mais l'organisation et le comportement des instruments sont également l'objet d'expérimentations radicalement différentes. »¹

Les évolutions et ruptures technologiques constatées conduisent à un certain nombre de difficultés quant à la définition du statut même de l'instrument et de l'instrumentalité, comme le montre une recherche présentée par Caroline Cance et Hugues Genevois aux Journées d'Informatique Musicale de 2009 [9], et développée avec Danièle Dubois pour le colloque

¹ [7], p. 101.

CIM09 [10] :

« En informatique musicale, on rencontre une grande variété de dénominations pour désigner les objets matériels et logiciels qui sont développés et utilisés. Il n'est [...] pas aisé de s'y retrouver entre ces diverses terminologies, qui ne sont le plus souvent pas définies. Il faut noter ici que ce processus est propre à toute nouvelle technologie et pratique émergente, qui voit un besoin de création et de stabilisation de vocabulaire commun. »²

Cance et Genevois précisent que

« pour [Claude] Cadoz, l'ordinateur ne peut être considéré comme un instrument, mais plutôt comme une représentation d'instrument [8]. Il existe des positionnements alternatifs, tels que celui développé par [Kasper] Toeplitz, pour qui l'ordinateur est en passe de devenir un instrument de scène [23]. »³

En 2007, Bossis propose l'énoncé suivant :

« Un instrument de musique ne peut être considéré comme tel que s'il comprend la chaîne complète du geste à l'émission sonore ou au moins à un signal électrique analogue à cette émission. [...] Un ordinateur n'en est pas un, mais s'il est muni d'une entrée son ou d'une interface gestuelle, d'un logiciel de synthèse ou de traitement, alors l'ensemble du dispositif peut être considéré comme un instrument de musique. »⁴

Nous reprendrons ici une définition formulée en 2008 et reposant sur la proposition précédente :

« Un instrument de musique numérique peut être défini de manière très générale par sa constitution – la présence d'une interface d'accès gestuel, d'un environnement informatique et d'une interface de sortie sonore – et sa destination : l'exécution musicale. »⁵

L'objectif de cet article est de soulever l'intérêt, dans le cadre des musiques reposant sur des dispositifs numériques d'exécution en temps réel, de la considération de l'instrument de musique lui-même en tant qu'objet d'analyse. Comme l'affirme Bruno Bossis,

« contrairement à l'instrument traditionnel, l'"instrument composé" n'est pas entièrement défini avant le travail d'écriture par le musicien. Le compositeur prend ainsi en charge une partie du métier du luthier. »⁶

² [9], p. 137.

³ [9], p. 138.

⁴ [6], p. 2.

⁵ [13], p. 3.

L'écriture de l'instrument étant pleinement inscrite dans la démarche poétique des compositeurs, nous évoquerons d'abord le programme informatique permettant la mise en œuvre d'une pièce importante du répertoire, *Jupiter* de Philippe Manoury, en nous attachant à montrer l'utilité de son examen en complément à celui de la partition et de ses paratextes. Dans un tout autre domaine, les musiques improvisées reposant sur les technologies peuvent aussi bien être étudiées à partir des procédures informatiques générant les résultats électroacoustiques. Nous exposerons comment le programme OMax, développé à l'Ircam et destiné à l'improvisation avec ordinateur, peut être adapté à des fins d'analyse.

2. L'EXAMEN DE L'INSTRUMENT COMME COMPLÉMENT À L'ANALYSE DE LA PARTITION : L'EXEMPLE DE *JUPITER* DE PHILIPPE MANOURY

2.1. Sources disponibles pour l'étude de *Jupiter*

Créée en 1987, *Jupiter* pour flûte et électronique en temps réel est la première pièce du cycle *Sonus ex machina* de Philippe Manoury, entièrement consacré à l'interaction entre jeu instrumental et dispositif électroacoustique. Les enjeux liés à l'usage de la technologie dans cette pièce sont largement exposés par le compositeur lui-même dans de nombreux écrits théoriques [20] et textes de circonstance [16, 19]. Dans un article daté de 1990, il exposait ainsi les intentions maîtresses des débuts de son cycle :

« Mon principal but a été d'intégrer dans un même discours deux mondes qui auparavant communiquaient de façons plus parallèles que conjointes : la musique instrumentale et la musique électronique. Je ne veux bien sûr pas dire par là que [mes] œuvres sont les premières à avoir confronté ces deux genres, mais plutôt qu'elles inaugurent une démarche tendant à une unification de ces deux domaines, tant du point de vue théorique que de celui de leur propre mode de fonctionnement. En un mot, j'ai essayé d'unifier ces méthodes dans une vision commune en faisant bénéficier la plus récente des apports de la plus ancienne. »⁷

L'implémentation d'une reconnaissance de partition, effectuée par Miller Puckette à la suite de travaux de Barry Vercoe [24] et de Roger Dannenberg [11], permet au dispositif de déclencher automatiquement les opérations audionumériques appropriées par la comparaison en temps réel du jeu de l'instrumentiste avec une représentation simplifiée de la partition mémorisée dans l'environnement informatique, donnant à l'interprète le contrôle du *tempo* global de la partie électroacoustique. La lecture de la partition fait apparaître l'organisation générale et locale de l'interaction mise en place. Au-dessus de la portée dédiée à la flûte se trouve un système de portées sur lequel se trouve représentée, sous des formes diverses,

la partie générée par le système temps réel (STR). Les sections sont numérotées par des chiffres romains, les sous-sections par les lettres de l'alphabet latin (de IA à XIIIc) et cette numérotation est encadrée. Les événements électroacoustiques déclenchés par le flûtiste sont numérotés par ordre croissant pour chaque grande section et entourés (figure 1).

Figure 1. Fin de l'introduction et début de la section IA de *Jupiter* (version 1992)⁸.

En plus de la partition et de ses écrits adjacents, le musicologue peut disposer du programme informatique destiné à l'exécution de *Jupiter*⁹ pour l'étude de la pièce.

2.2. Utilité analytique du programme informatique d'exécution

La partition du début de la section IA (figure 1) montre qu'au premier *do dièse* joué par le flûtiste correspond un premier événement électroacoustique, déclenchant le traitement du signal issu de la flûte par un *harmonizer* générant le même *do dièse* et transposant le signal sur *sol*, le tout passant dans une réverbération infinie. Le deuxième *do dièse* est marqué par un second événement, mais aucune modification de la partie électroacoustique n'est visible. Le *ré* est ignoré par le dispositif, et au *do dièse* de la deuxième mesure correspond un troisième événement, entraînant une nouvelle transposition sur *sol dièse*. Dans Max/MSP, la reconnaissance de partition permet d'affecter aux notes concernées des numéros d'événements, auxquels sont associées des mises à jour de valeurs paramétriques écrites dans un objet *qlist* (figure 2). Sans détailler ici la syntaxe du contenu d'une *qlist* ni tous les paramètres utilisés, remarquons que les variables *trans0* à *trans3* sont fixées, lors du premier événement, aux valeurs -600 (transposition de 600 *midicents* vers le grave, soit un triton correspondant à la transposition de *do dièse* vers *sol*), 0, 0 et 0 (pas de transposition). Lors du deuxième événement, ces quatre variables sont fixées aux mêmes valeurs, confirmant l'information donnée par la partition, selon laquelle à l'événement détecté ne correspond aucune modification de la partie électroacoustique.

⁶ [6], p. 2.

⁷ [17], p. 43.

⁸ [18], p. 1.

⁹ [21].

```

0 1 ----- 10;
spatop 1;
hto2 60, 127 1000;
hto4 60, 127 1000;
oto2 60, 93 1000;
oto4 60, 93 1000;
rto2 60, 108 1000;
rto4 60, 108 1000;
amparpeg 0;
ampgliss 100;
pufamp -1, 127, 127, 138, 138;
dtoh 127;
htor 127;
dtr 0;
rgate 117;
revfb 127;
trans0 -600;
trans1 0;
trans2 0;
trans3 0;
950 rgate 0;
0 2 ----- 30;
trans0 -600;
trans1 0;
trans2 0;
trans3 0;
0 3 ----- 40;
amp5 137;
amp6 137;
rgate 117;
revfb 127;
trans0 -600;
trans1 -1700;
trans2 0;
trans3 0;
600 rgate 0;
0 4 ----- 60;

```

Figure 2. Trois premiers événements écrits dans la *qlist* de la section I de *Jupiter*¹⁰.

Lors du troisième événement, la variable `trans1` passe de 0 à -1700, correspondant à la transposition de *do dièse* vers *sol dièse* (une octave plus une quarte vers la grave). À part l'implémentation technique des intentions musicales et des traitements utilisés, l'examen du programme ne permet ici que de vérifier ce dont la partition a déjà informé.

Dans d'autres cas, la partition peut être assez précisément notée, mais laisser une ambiguïté sur le déroulement de la partie électroacoustique. Dans la section Ib, des motifs rythmiques sont appliqués à un *frequency shifter* et à l'*harmonizer*. La partition ne permet pas de savoir si ces motifs sont temporellement fixes ou dépendants d'une mesure du *tempo* joué par l'interprète jusqu'alors (figure 3). Le contenu de la *qlist* montre que ces motifs rythmiques sont absolument fixés avant l'exécution. Le jeu des notes *sol* par le flûtiste est reconnu comme devant déclencher les motifs, mais le déroulement temporel de ceux-ci est ensuite indépendant des paramètres de l'interprétation. Par exemple, le motif correspondant à l'événement 13 est toujours déclenché 500 millisecondes après la détection, et les durées de ses trois premières unités sont fixées à 375, 125, et 500 millisecondes.

The image shows a musical score for section Ib of Jupiter. It consists of three staves. The top staff is labeled '(Freq. sh.)' and contains a sequence of notes with a box labeled 'IB' above it. The middle staff is labeled '(Harm.)' and contains notes with circled numbers 13 and 14. The bottom staff contains notes with dynamic markings 'sfz sfzpp' and 'mp'. Above the staves, there are some square symbols and the text '(Freq. sh. → Harm.)'.

Figure 3. Début de la section Ib de *Jupiter*¹¹.

La dernière note est coupée par le jeu du *sol* associé à l'événement 14 (figure 4). Remarquons que si ces valeurs respectent les proportions rythmiques du motif écrit sur la partition, elles ne correspondent pas au *tempo* que celle-ci indique (72 à la noire), mais à 60 à la noire.

```

0 13 ----- 200;
fto4 0;
fto2 127;
hto2 113;
htof 127;
ftor 0;
500 trans0 -600;
fpos 127;
fsfre 79;
375 trans1 -600;
fneg 127;
fsfre 79;
fpos 0;
rtof 0;
ftor 0;
125 trans2 -600;
hamp1 0;
fpos 127;
fsfre 79;
fneg 0;
rtof 0;
ftor 0;
500 trans3 -600;
hamp2 0;
fneg 127;
fsfre 79;
fpos 0;
rtof 0;
ftor 0;
0 14 ----- 210;
291 trans0 -600;
hamp3 0;
fneg 127;
fsfre 79;
167 trans1 0;
hamp4 0;
167 trans2 -300;
hamp1 0;
166 trans3 -400;
hamp2 0;
167 trans0 -600;
hamp3 0;
0 15 ----- 220;

```

Figure 4. *qlist* des événements 13 et 14 de la section Ib de *Jupiter*¹².

Comme cela peut être le cas plus généralement avec les compositions pour dispositif temps réel, l'examen de l'environnement informatique constitutif de l'instrument d'exécution de *Jupiter* permet ainsi de connaître avec

¹⁰ [21].

¹¹ [18], p. 2.

¹² [21].

précision l'interaction des parties instrumentale et électroacoustique, ainsi que l'organisation locale de celle-ci, pour laquelle la partition seule n'est que rarement un outil d'analyse suffisant.

3. LE PROGRAMME OMAX ET SON ADAPTATION À DES FINS D'ANALYSE

3.1. Principe général de l'application du programme OMax

Développé depuis 2004 à partir de recherches en modélisation stylistique portées par Gérard Assayag et Shlomo Dubnov, et de recherches sur l'improvisation avec l'ordinateur menées par Assayag, Marc Chemillier et Georges Bloch, le programme OMax repose sur une communication par le protocole *OSC*¹³ entre les logiciels OpenMusic, principalement destiné à la composition assistée par ordinateur, et Max/MSP, orienté sur l'interaction et le traitement de signal en temps réel. Les trois auteurs du programme précisent que

« l'environnement permet l'interaction avec un ou plusieurs improvisateurs humains, un apprentissage stylistique à la volée, la génération d'une improvisation virtuelle, un alignement métrique et harmonique, un archivage des modèles stylistiques et leur hybridation. »¹⁴

Bien qu'OMax puisse fonctionner de manière autonome, sans intervention extérieure ajoutée au jeu du ou des instrumentistes, il est généralement contrôlé par un opérateur informatique qui régit la partie électroacoustique de l'improvisation. Le dispositif global, constitué de l'instrument source et de l'environnement informatique, est donc idéalement joué par deux instrumentistes.

Le principe général qui gouverne l'improvisation avec OMax est celui de « réinjection stylistique » [2]. L'algorithme retenu pour son application est l'Oracle des facteurs¹⁵, un modèle issu de recherches en reconnaissance de motifs au sein de chaînes de caractères présentées par Allauzen, Crochemore et Raffinot en 1999 [1]. Il permet de repérer dans une chaîne de caractères donnée des sous-séquences communes précisément appelées facteurs et de créer des liens entre elles. La navigation entre les facteurs permet de recombinaison le matériau fourni par l'instrumentiste de manière cohérente avec son style de jeu, lui permettant de dialoguer avec son double électroacoustique.

3.2. Problèmes posés pour l'analyse de l'improvisation par l'état actuel du programme

OMax est destiné à l'improvisation et non à l'analyse musicologique de celle-ci. Un module d'enregistrement

¹³ *OpenSoundControl*.

¹⁴ « *This environment allows interaction with one or more human player, on-the-fly stylistic learning, virtual improvisation generation, metrical and harmonical alignment, stylistic model archiving and hybridation.* » ([3], p. 1). Notre traduction.

¹⁵ En anglais *Factor Oracle (FO)*.

de la production sonore est disponible dans le programme, mais l'information audionumérique de sortie est une ressource limitée et ne permet pas à elle seule l'examen de l'interaction entre jeu de l'instrumentiste et recombinaison des matériaux. Pour l'analyse, une autre difficulté vient du fait que la partie OpenMusic d'OMax, programmée en Lisp, n'est pas transparente pour l'utilisateur, contrairement à sa partie Max/MSP, qu'il est possible d'explorer et, éventuellement, de reprogrammer. Toutefois, il existe une version nommée WoMax¹⁶, réécrite par Benjamin Lévy des principales fonctions du programme initial entièrement dans Max/MSP/Jitter¹⁷. Cette version est augmentée d'un dispositif original de visualisation de l'information qui facilite par ailleurs l'approche musicologique (figure 5).

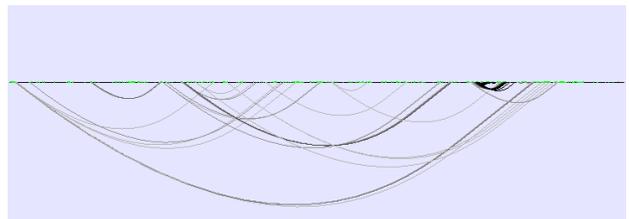


Figure 5. Représentation dans WoMax des transitions possibles entre les facteurs d'une séquence musicale.

Si cette possibilité de visualisation peut s'avérer très efficace pour les musiciens improvisateurs et leur compréhension des possibilités de recombinaison des matériaux dans le temps du jeu, l'analyse d'une improvisation particulière implique quelques adaptations de l'environnement informatique.

3.3. Adaptation de WoMax à une approche analytique

L'obtention et la mémorisation d'informations solfégiques, tant pour les matériaux fournis par l'instrumentiste que pour ceux générés par OMax, semblent plus adaptées à l'analyse que des informations audionumériques brutes. Il s'agit donc de repérer la partie du programme située après la conversion du signal audionumérique entrant en informations littérales et avant leur envoi à l'Oracle des facteurs. Dans WoMax, l'information la plus pertinente se trouve à la sortie du milieu de l'objet *YinGB* (figure 6). Cette abstraction a pour noyau l'objet *yin~*, développé pour Max/MSP par Norbert Schnell à partir de l'algorithme YIN de détection d'une fréquence fondamentale proposé par Alain de Cheveigné et Hideki Kawahara [12].

¹⁶ Téléchargé depuis [15].

¹⁷ WoMax signifie littéralement « Without OpenMusic OMax » ([14], p. 24).

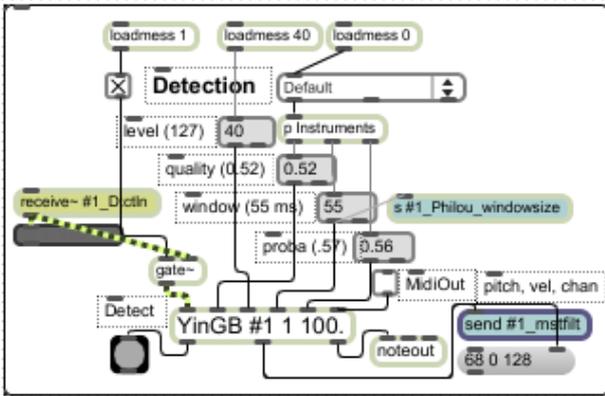


Figure 6. Objet *YinGB* dans la section de détection de *WoMax* (vue du programme en mode éditeur).

Une partition Midi des fréquences fondamentales jouées par l'instrumentiste peut être simplement obtenue en traitant cette information et en l'enregistrant dans un fichier lisible dans un éditeur de partition ou sous toute autre forme aisément déchiffirable (figure 7).

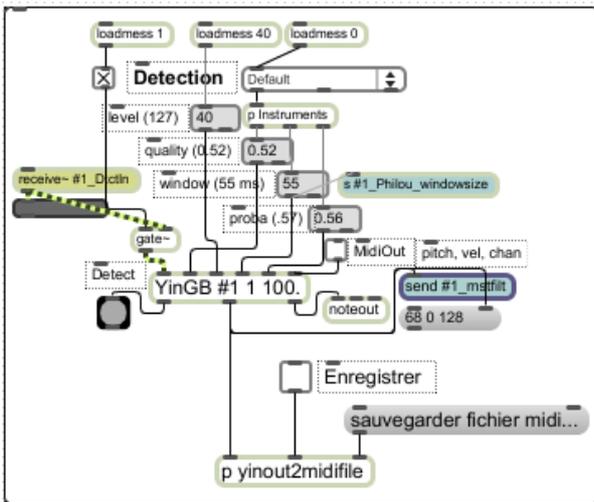


Figure 7. Adjonction d'un sous-programme de création d'un fichier Midi à la sortie de l'objet *YinGB* de *WoMax*.

Selon le même principe, il est possible d'obtenir une partition Midi des fréquences fondamentales des échantillons recomposés et envoyés à la sortie de *WoMax* en ajoutant, au niveau de l'ordonnanceur (*scheduler*) de la recombinaison, un autre sous-programme permettant la création d'un fichier de représentations musicales adaptées à la lecture.

Les informations musicales d'entrée et de sortie obtenues sous une forme symbolique et non audio numérique peuvent ainsi être mises en correspondance. Toutefois, sur une durée d'improvisation conséquente, le grand nombre d'événements vis-à-vis du petit nombre de notes possibles rend difficile la compréhension de la

recombinaison à un niveau local. La mise en œuvre d'un sous-programme associant à chaque événement de la partition de sortie son index d'état tel qu'il est fourni par l'Oracle des facteurs est alors souhaitable. De cette manière, chaque événement du matériau électroacoustique porte avec lui l'information relative à son origine exacte dans le matériau fourni par l'instrumentiste.

Le caractère programmable d'*OMax* et de *WoMax* en font un instrument adaptable à la diversité des musiques improvisées, mais également à l'examen même de ces pratiques. Les solutions proposées ici doivent ouvrir la perspective d'un ajustement du dispositif d'exécution afin que celui-ci puisse produire les traces utiles à l'analyse de l'improvisation.

4. CONCLUSION

À propos des musiques fixées sur bande et s'appuyant sur la notion de *faktura* comme outil conceptuel d'analyse introduite par Marc Battier [4], Olivier Baudouin affirme que

« l'analyse de la musique électroacoustique a pour objet un produit fini, généralement enregistré sur un support et dont les sources présentent la particularité de ne pas suivre les codes de la musique conventionnelle tout en s'insérant dans un dispositif lié à la technologie. Par conséquent, la compréhension du résultat obtenu à l'issue de la construction de l'œuvre passe – à l'instar d'une analyse menée avec une partition conventionnelle mais selon d'autres modalités – par sa déconstruction [...] »¹⁸

Transposant cette proposition aux musiques interprétables ou improvisées, l'étude des œuvres reposant sur des dispositifs numériques d'exécution peut impliquer des compétences de luthier pour le musicologue souhaitant s'en emparer comme les instruments de son étude. Dans son ouvrage *Du mode d'existence des objets techniques* [22], Gilbert Simondon opère la distinction suivante :

« [on peut entendre] par *outil* l'objet technique qui permet de prolonger et d'armer le corps pour accomplir un geste, et par *instrument* l'objet technique qui permet de prolonger et d'adapter le corps pour obtenir une meilleure perception ; l'instrument est outil de perception. Certains objets techniques sont à la fois des outils et des instruments, mais on peut les dénommer outils ou instruments selon la prédominance de la fonction active ou de la fonction perceptive. »¹⁹

À partir des quelques exemples décrits dans cet article, issus de ces deux domaines différenciés que sont la musique écrite et l'improvisation, apparaît pour les dispositifs d'exécution, en plus de leur fonction active, une fonction perceptive augmentée d'une fonction analytique. Le programme numérique situé au cœur de l'instrument de musique peut aussi bien être l'objet d'une observation éclairante pour l'étude de l'écriture d'une œuvre ou d'une performance improvisée que se

¹⁸ [5], p. 78.

¹⁹ [22], p. 114. C'est Simondon qui souligne.

prêter à des manipulations destinées à favoriser l'analyse des démarches poétiques dont il est à la fois le vecteur et le témoin.

5. RÉFÉRENCES

- [1] Allauzen, C., Crochemore, M., Raffinot, M., « Factor oracle: a new structure for pattern matching », *Proceedings of SOFSEM'99*, Milovy (République tchèque), du 27 novembre au 4 décembre 1999, p. 295-310.
- [2] Assayag, G., Bloch, G., Chemillier, M., « Improvisation et réinjection stylistiques », Actes des Rencontres Musicales Pluridisciplinaires 2006, Lyon, Grame, 17 et 18 mars 2006, p. 79-85.
- [3] Assayag, G., Bloch, G., Chemillier, M., « OMax-OFon », *Proceedings of Sound and Music Computing SMC'06*, Marseille, du 18 au 20 mai 2006. <http://articles.ircam.fr/textes/Assayag06b/> (lien vérifié le 14 février 2010).
- [4] Battier, M., « A Constructivist Approach to the Analysis of Electronic Music and Audio Art – Between Instruments and *Faktura* », *Organised Sound*, vol. 8, n° 3, décembre 2003, p. 249-255.
- [5] Baudouin, O., « La *Faktura*, “outil conceptuel d’analyse” – illustration avec *Stria*, de John Chowning », Actes des 14^{èmes} Journées d’Informatique Musicale, Grenoble, ACROE et laboratoire ICA, du 1^{er} au 3 avril 2009, p. 77-83.
- [6] Bossis, B., « Écriture instrumentale, écriture de l’instrument », Actes du colloque *Composer au XXI^e siècle. Processus et philosophies*, Montréal, du 28 février au 3 mars 2007. http://www.oiccm.umontreal.ca/doc/col_2007/textes_conferenciers/bossis/bossis_montrealcompos101env.pdf (actuellement inaccessible. Dernière consultation le 21 avril 2009).
- [7] Bossis, B., « De l’amplification à l’autonomie de l’instrument : ruptures électroniques et conséquences esthétiques », *Analyse musicale*, n° 61, 4^e trimestre 2009, p. 98-103.
- [8] Cadoz, C., « Musique, geste, technologie », in Genevois, H. et De Vivo, R. (dir.), *Les nouveaux gestes de la musique*, Marseille, Éditions Parenthèses, collection Eupalinos, 1999, p. 47-92.
- [9] Cance, C., Genevois, H., « Questionner la notion d’instrument en informatique musicale : analyse des discours sur les pratiques du méta-instrument et de la méta-mallette », Actes des 14^{èmes} Journées d’Informatique Musicale, Grenoble, ACROE et laboratoire ICA, du 1^{er} au 3 avril 2009, p. 133-142.
- [10] Cance, C., Genevois, H., Dubois, D., « What Is Instrumentality In New Digital Musical Devices? A contribution from cognitive linguistics & psychology », Actes en ligne du colloque CIM09 *La musique et ses instruments*, Paris, Université Pierre et Marie Curie Paris 6, Université Paris-Sorbonne Paris IV, Musée de la Musique et Musée du Quai Branly, du 26 au 29 octobre 2009. http://cim09.lam.jussieu.fr/CIM09-en/Proceedings_files/21A-Cance%26al.pdf (lien vérifié le 14 février 2010).
- [11] Dannenberg, R., « An On-Line Algorithm for Real-Time Accompaniment », *Proceedings of the 1984 International Computer Music Conference*, Paris, Ircam, du 19 au 23 octobre 1984, p. 193-198.
- [12] De Cheveigné, A., Kawahara, H., « YIN, a fundamental frequency estimator for speech and music », *Journal of Acoustic Society of America*, vol. 111, n° 4, avril 2002, p. 1917-1930.
- [13] Dufeu, F., « Temps réel et temps différé dans l’instrument de musique numérique », Actes en ligne du colloque EMS08 *Musique concrète, 60 ans plus tard*, Paris, Ina-GRM et Université Paris-Sorbonne Paris IV, du 3 au 7 juin 2008. <http://www.ems-network.org/ems08/papers/dufeu.pdf> (lien vérifié le 14 février 2010).
- [14] Lévy, B., *Visualizing OMax*, Mémoire de DEA en Acoustique, Traitement du Signal et Informatique Appliqués à la Musique (ATIAM), dir. Assayag, G. et Bloch, G., Ircam et Université Pierre et Marie Curie Paris 6, 2009. <http://articles.ircam.fr/textes/Levy09a/> (lien vérifié le 14 février 2010).
- [15] Lévy, B., *WoMax Project Homepage*. <http://recherche.ircam.fr/equipes/repmus/WoMax/> (lien vérifié le 14 février 2010).
- [16] Manoury, P., notes de programme de la création de *Jupiter*, 25 avril 1987. <http://brahms.ircam.fr/works/work/10482/> (lien vérifié le 14 février 2010).
- [17] Manoury, P., « La note et le son : un carnet de bord », janvier 1990, in *La note et le son. Écrits et entretiens. 1981-1998*, Paris, L’Itinéraire, L’Harmattan, collection Musique et Musicologie : les Dialogues, 1998, p. 43-57.
- [18] Manoury, P., partition de *Jupiter* pour flûte et électronique en temps réel (1987 – révision de 1992), partition, Paris, Amphion, dépôt de copyright : 1991.
- [19] Manoury, P., notice du disque *La partition du ciel et de l’enfer. Jupiter*, CD, Musidisc France, Accord, 465 307-2, 1997.
- [20] Manoury, P., *La note et le son. Écrits et entretiens. 1981-1998*, Paris, L’Itinéraire, L’Harmattan, collection Musique et Musicologie : les Dialogues, 1998.
- [21] Manoury, P., programme Max/MSP pour l’exécution de *Jupiter*, version 2006. <http://production.ircam.fr/doc/Jupiter-2006.zip> (lien

vérifié le 14 février 2010).

- [22] Simondon, G., *Du mode d'existence des objets techniques*, Paris, Aubier, première édition : 1958, édition consultée : édition augmentée d'une préface de J. Hart et d'une postface de Y. Deforge, 1989.
- [23] Toeplitz, K., « L'ordinateur comme instrument de concert – aussi une question d'écriture ? », Actes des 9^{èmes} Journées d'Informatique Musicale, Marseille, du 29 au 31 mai 2002, p. 199-207.
- [24] Vercoe, B., « The Synthetic Performer in the Context of Live Performance », *Proceedings of the 1984 International Computer Music Conference*, Paris, Ircam, du 19 au 23 octobre 1984, p. 199-200.