

# RÉTROACTION MUSIQUE - PHYSIOLOGIE : UNE APPROCHE SELON LE PARADIGME DES ÉMOTIONS

*Pierre-Henri Vulliard, Joseph Larralde, Myriam Desainte-Catherine*

Univ. Bordeaux, LaBRI, UMR 5800, Talence, France

CNRS, LaBRI, UMR 5800, Talence, France

{phvulliard, joseph.larralde}@gmail.com

myriam.desainte-catherine@labri.fr

## RÉSUMÉ

Il est un fait avéré que l'écoute de musique induit des réactions physiologiques particulières chez l'auditeur, et l'étude de ces inductions constitue un terrain d'études encore vaste. Lorsque l'on veut analyser les signaux physiologiques mesurés sur une personne écoutant de la musique, il faut définir des modèles pour savoir quelles informations rechercher dans ces signaux. Inversement, lorsque l'on cherche à générer de la musique à partir de signaux physiologiques échantillonnés, on cherche en fait à créer une induction inverse de celle qui a lieu naturellement, et il faut pour cela définir des modèles afin d'être capable de contrôler tous les paramètres d'un système de musique générative à partir des quelques signaux physiologiques à disposition, et ce de manière cohérente. La notion d'émotion, en plus de sembler toute indiquée dans le contexte, se révèle être une notion pivot très pratique pour faire correspondre des modèles musicaux avec des modèles physiologiques.

Nous proposons dans cet article un système temps-réel expérimental visant à étudier les interactions et rétroactions entre musique et physiologie, basé sur le paradigme des émotions.

## 1. INTRODUCTION

Le système expérimental présenté dans cet article, baptisé MuZICO, permet d'aborder les interactions entre musique et physiologie selon deux approches : l'une, de bas niveau, se base sur la sonification des signaux physiologiques issus du système nerveux autonome, créant ainsi un lien direct entre l'expression inconsciente d'un état physiologique et la production de vibrations acoustiques. L'autre approche, de haut niveau, introduit une couche d'abstraction prenant en charge la reconnaissance d'états de vigilance (vigilance, relaxation, sommeil) et/ou émotionnels de l'auditeur, et permet de piloter des paramètres de génération musicale de haut niveau en fonction de ces états.

MuZICO utilise actuellement pour la capture de signaux physiologiques un système de biofeedback Thought Technologies ProComp5 infinity. Il s'agit d'un boîtier à 5 entrées sur lesquelles on peut connecter différents types de capteurs.

Les capteurs utilisés avec MuZICO sont les suivants :

- capteur de respiration abdominal / thoracique
- capteur de conductivité de la peau
- électrocardiographe
- électroencéphalographe

Par la suite, nous décrivons l'évolution de nos travaux sur la sonification et l'interaction entre les états de vigilance et la synthèse musicale, puis nous proposons une généralisation des états de vigilance aux états émotionnels et sa validation dans le cadre d'une collaboration avec une équipe effectuant des recherches sur la reconnaissance d'émotions exprimées corporellement. Enfin, nous étudions les perspectives offertes par la synthèse de ces expériences pour l'étude des interactions et rétroactions entre musique, émotions perçues, et physiologie.

Ces travaux se sont déroulés dans le cadre du projet SCRIME, conventionné par l'Université Bordeaux 1, l'Institut Polytechnique de Bordeaux et le Conservatoire de Musique de Bordeaux, et financé par la DGCA du Ministère de la Culture.

## 2. TRAITEMENT DE SIGNAUX PHYSIOLOGIQUES SELON L'APPROCHE ÉNERGÉTIQUE

### 2.1. Sonification : mise en correspondance directe

La sonification joue dans notre cas un rôle de catalyseur de l'attention de l'auditeur et de sa compréhension du lien entre les manifestations de son système nerveux autonome, habituellement involontaires, et la musique qu'elles génèrent. L'énergie déployée par les organes internes et échantillonnée grâce aux capteurs est traduite directement en énergie sonore, créant ainsi un phénomène de perception synesthésique.

Concrètement, nous calculons les vitesses de variation des données provenant de tous les capteurs, excepté des électrodes EEG. Ces vitesses de variation sont traduites en énergie par le contrôle direct du volume, du timbre ou de la fréquence d'un son. En ce qui concerne les capteurs cardiaques et de respiration, la fréquence du cycle est évaluée (fréquence des battements de cœur ou des cycles inspiration / expiration), et les variations de cette fréquence sont également exprimées par des variations de volume,

de timbre ou de hauteur. En ce qui concerne les signaux EEG, l'évaluation de l'évolution des rapports énergétiques entre des plages de fréquence caractéristiques de certains états de vigilance [6] est sonifiée de la même manière.

## 2.2. Interaction vigilance / synthèse musicale

Outre la production directe du son à partir des signaux physiologiques (approche orientée instrument), l'approche énergétique passe également par un plus haut niveau d'abstraction en analysant l'état d'éveil de l'auditeur, pour générer un flux musical à partir de règles de composition harmoniques, rythmiques et timbrales. L'idée est ici de contrôler le dynamisme de ce flux musical en fonction de l'état d'éveil de l'auditeur.

Les données physiologiques sont donc intégrées à un seul paramètre continu représenté par un axe et représentant l'état d'éveil de l'auditeur, lui-même étant ensuite redistribué sur l'ensemble des paramètres de génération musicale : il s'agit d'un mapping  $M \rightarrow 1 \rightarrow N$ .

Plusieurs travaux menés par Pierrick Legrand et Frédéric Faïta Ainseba [8] montrent le lien entre certains paramètres de composition musicale, l'énergie évoquée par la musique résultante, et le niveau de vigilance induit par son écoute chez un auditeur.

Ces paramètres, dans leur mise en œuvre au sein de l'environnement MuZICO, ont suivant les cas des valeurs discrètes ou continues, définies par des adjectifs ou des valeurs numériques. Par exemple le timbre d'un son peut aller de étouffé à brillant, ce qui correspond à une grandeur énergétique mesurable dans son signal, liée à son centroïde spectral. Les autres paramètres liés à l'énergie sont le tempo de la musique, l'attaque de l'enveloppe dynamique des notes, la hauteur des notes, le nombre d'instruments simultanés, le côté plus ou moins "naturel" des sons instrumentaux, la nature percussive ou entretenue des sons, ainsi que le nombre de notes simultanées [5].

Tous ces paramètres sont caractérisés par des intervalles bornés utilisés par MuZICO pour la génération musicale.

Soient  $V_{Emin}$  et  $V_{Emax}$  les intervalles de valeurs de paramètres musicaux correspondant respectivement à des énergies évoquées par la musique faibles et élevées :

Paramètre (unité)	$V_{Emin}$	$V_{Emax}$
tempo (bpm)	[0.1 ; 60]	[144 ; 300]
timbre	étouffé	brillant
attaque des notes (ms)	[100 ; 1000]	[1 ; 30]
hauteur des notes (Hz)	[110 ; 220]	[440 ; 880]
sons instrumentaux	naturels	synthétiques
nature des sons	percussive	entretenus
nombre de voix	1	beaucoup
densité de notes	faible	élevée

Les paramètres ayant une unité indéfinie se traduisent au sein de MuZICO dans différentes unités en fonction de l'algorithme de production sonore permettant leur implémentation (synthèse, lecture d'échantillons, contrôle d'effet). Par exemple, on pourra traduire la brillance d'un son



Figure 1. Pierre-Henri Vulliard au Marché de Lorme.

par un réglage de l'index de modulation dans un algorithme de synthèse FM, ou par le contrôle de la fréquence de coupure d'un filtre passe-bas sur un lecteur d'échantillons.

## 2.3. Applications

Le système MuZICO a déjà été présenté au public dans des cadres artistiques (utilisation en concert comme accompagnement musical génératif) et de vulgarisation scientifique (présentation sous forme d'atelier ou de borne interactive). Ses applications peuvent donc être artistiques, pédagogiques, ou thérapeutiques.

Dans le cadre artistique, MuZICO a été utilisé dans deux configurations : l'une, impliquant fortement l'un des auteurs en tant que musicien saxophoniste, consiste à générer un accompagnement à partir de capteurs de respiration et cardiaques ou eeg en tant que support d'improvisation pour le musicien. Concrètement, cette configuration a été utilisée lors de concerts de musique électroacoustique organisés par le SCRIME (au Marché de Lorme de Bordeaux (voir figure 1), et à l'Hôtel de Région d'Aquitaine), ou de performances (au musée scientifique Cap Sciences et au club i-Boat à Bordeaux). L'autre a été mise en place dans le cadre du projet ANR Care et permet à un danseur (Gaël Domenger) muni d'une combinaison de capture de mouvements de piloter lui-même la génération de la musique sur laquelle il improvise (voir figure 2). Une représentation publique a eu lieu au Casino de Biarritz en mars



Figure 2. Gaël Domenger utilisant eMotion et MuZICO.

2011.

Dans un cadre plus pédagogique, des séances de dé-cryptage du système MuZICO ont été présentées au public de manière ponctuelle : des ateliers ont eu lieu en durant les journées Eurêka de la Fête de la Science 2011 à l'Hôtel de Région d'Aquitaine et pendant la Semaine du Son 2012 à Cap Sciences. D'une manière plus suivie, une borne interactive permettant d'essayer le système a été mise en place à Cap Sciences pendant toute la durée de l'exposition Numériquement Vôtre (mars - décembre 2011). Le public pouvait interagir avec MuZICO grâce à un capteur de conductivité de la peau intégré à la borne.

Dans le domaine thérapeutique, des investigations sur l'apport possible du système aux techniques de relaxation sont menées parallèlement depuis le début du projet par un des auteurs auprès de professionnels du milieu tels que des masseurs ou des hypnotiseurs.

### 3. INTERACTION ÉMOTIONS / SYNTHÈSE MUSICALE

#### 3.1. La représentation des émotions

Nous proposons ici une généralisation des états de vigilance aux états émotionnels dans le but d'augmenter les possibilités d'expressivité de la génération musicale, en ajoutant à l'axe de la vigilance un deuxième axe. Il y a beaucoup de définitions et de représentations des émotions dans le champ d'étude de la psychologie, mais peu d'entre elles sont actuellement utilisées en informatique. Nous avons choisi d'utiliser le même modèle que dans [5] et [9], modèle couramment employé en psychologie sous le nom de roue des émotions de Plutchik [7]. Il s'agit d'un espace bi-dimensionnel défini par deux axes correspondant à l'excitation (que nous assimilons à la vigilance dans le cadre de la génération musicale), et à la valence (plaisir - déplaisir), et dans lequel nous pouvons placer les émotions primaires d'Ekman [3] (voir figure 3).

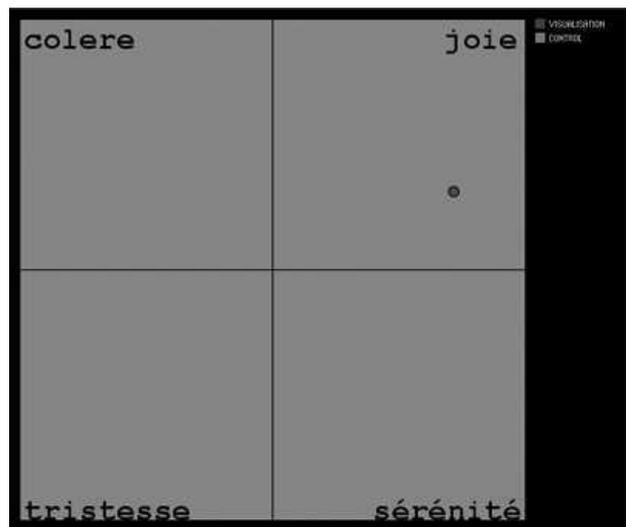


Figure 3. placement des émotions primaires d'Ekman dans l'espace 2D de représentation des émotions.

#### 3.2. Projet Care - logiciel eMotion

La collaboration dans le cadre du projet Care [2] avec une équipe effectuant des recherches sur la reconnaissance d'émotions basées sur le mouvement, nous a permis de tester la génération de la musique en utilisant les deux axes de l'espace des émotions.

Dans le cadre du spectacle de danse servant de conclusion à ce projet, nous avons généré un flux sonore et musical en utilisant l'approche énergétique décrite précédemment pour associer les règles de composition musicale aux étiquettes d'émotions envoyées par le logiciel de reconnaissance d'émotions eMotion. Ce logiciel effectuant ses analyses à partir des informations de position et de mouvement envoyées par une combinaison munie de capteurs, nous avons pu aussi utiliser l'approche instrumentale pour associer cette fois-ci des sons directement aux mouvements du danseur.

L'analyse de questionnaires distribués au public lors de la représentation au Casino de Biarritz a ensuite permis une réflexion a posteriori sur l'expérience.

#### 3.3. Validation du système par le public

La deuxième représentation du spectacle de restitution du projet Care s'est faite devant un public de 120 personnes. A leur arrivée, ces personnes se sont vues remettre chacune un questionnaire dans le but pour l'équipe de collecter les retours des spectateurs [1]. 97 questionnaires ont été remplis et rendus, soit approximativement 80% de l'ensemble des questionnaires distribués. Pour un non-chercheur, le monde scientifique peut parfois paraître quelque peu hermétique. Le spectacle présentait les résultats du projet Care sous une forme à la fois artistique et ludique, les scientifiques étaient sur scène, ce que le public a apprécié. 90% des personnes qui se sont exprimées sur le questionnaire ont cité l'aspect innovant et magique du spectacle comme le meilleur point. Cependant, 75% de ces

mêmes personnes ont trouvé que la musique générée évoquait de manière trop évidente l'émotion exprimée par le danseur.

Cette critique est intéressante car elle valide les paramètres choisis pour la génération musicale.

### 3.4. Valence et complexité musicale

D'après [5], les paramètres musicaux liés à la valence des émotions évoquées par la musique sont la tonalité et la complexité. Nous proposons ici des modèles génératifs harmoniques, rythmiques et mélodiques, ainsi qu'une description de leurs paramètres sous l'angle de vue de la complexité musicale.

La manière dont ces modèles (implémentés sous forme de modules dans MuZICO) communiquent entre eux est explicitée en appendice.

#### 3.4.1. Échelles de hauteurs

Nous proposons un modèle de génération d'échelles de hauteurs basé sur la suite :

$$U_{n+1} = \left( U_n \times \omega^{\frac{V_n \bmod (m+1)}{\theta}} - U_0 \right) \times \omega^{-p}$$

$$\text{où } U_0 = f_f, \text{ et } p \in \mathbb{N}^+ \text{ tel que } U_0 < U_{n+1} < \omega \times U_0 \quad (1)$$

où  $\omega$  est le rapport d'octave,  $\theta$  le nombre par lequel on divise l'octave pour obtenir le tempérament,  $f_f$  la fréquence fondamentale de la gamme générée,  $V$  un vecteur de dimension  $m$ , et  $V_{i,i \in [1..m]}$  la  $i^{\text{ème}}$  composante de  $V$ .

On considère que la complexité d'une échelle de hauteurs dépend du nombre d'itérations nécessaires à sa génération, et des valeurs de  $\omega$ ,  $\theta$ , et  $V$ .

Par exemple, on se place dans le cas où  $\omega = 2$ ,  $\theta = 12$  (musique tonale occidentale).

Soit  $V = (7)$  (génération de l'échelle de hauteurs par parcours du cycle des quintes), et  $N$  le nombre d'itérations effectuées. On a alors :

- la gamme pentatonique ou l'accord Maj 6/9 pour  $N = 4$
- la gamme diatonique pour  $N = 6$
- la gamme chromatique pour  $N = 11$

Ces gammes ont une complexité harmonique croissante. Des échelles de complexités différentes peuvent aussi être créées en faisant varier  $V$  : soit  $V = (4)$  et  $N = 2$ , on génère alors une échelle de hauteur correspondant à un accord augmenté. Soit  $V = (3)$  et  $N = 3$ , on génère alors une échelle de hauteurs correspondant à un accord diminué. Soit  $V = (3, 4, 4, 3)$  et  $N = 6$ , on génère ainsi la gamme mineure mélodique ascendante.

#### 3.4.2. Suites d'accords

Dans MuZICO, les suites d'accords sont générées par une grammaire qui tient compte du contexte pour super-

viser les transpositions liées aux modulations. La complexité d'un accord dépend de plusieurs facteurs que nous avons identifié :

- le nombre de notes qui le constituent
- la complexité de l'échelle de hauteurs sur laquelle il est construit
- l'ordre d'apparition de ses notes dans la construction de cette échelle par itérations

La complexité d'une suite d'accords dépend également de plusieurs facteurs :

- la complexité des accords qui la constituent
- le nombre d'accords différents dans la séquence
- la complexité harmonique des enchaînements d'accords, en prenant en compte principalement les cadences et les résolutions dans le cadre de la musique modale et tonale, et les modulations dans le cadre de cette dernière.

#### 3.4.3. Motifs rythmiques

Les motifs rythmiques sont créés par superposition de plusieurs patterns rythmiques linéaires synchronisées à une même pulsation et de longueur définie, chaque pattern rythmique étant défini par un décalage par rapport à l'instant initial commun, et par une densité rythmique (nombre de pulsations entre deux coups du pattern). Cette superposition se fait par une opération de "ou" logique, si l'on considère les patterns rythmiques superposées comme des vecteurs de bits, 1 représentant un coup et 0 le silence.

Soient  $m$  la longueur des patterns rythmiques,  $q$  le nombre de patterns superposés,  $of_i$  le décalage par rapport à l'instant initial du  $i^{\text{ème}}$  pattern, et  $d_i$  la densité rythmique du  $i^{\text{ème}}$  pattern.

Si l'on considère qu'un vecteur de bits est équivalent à un ensemble fini de nombres entiers, chaque nombre de l'ensemble indiquant la présence d'un bit positif à l'indice lui correspondant dans le vecteur, un motif rythmique  $R$  constitué par superposition de patterns  $P_i$  peut alors être défini par :

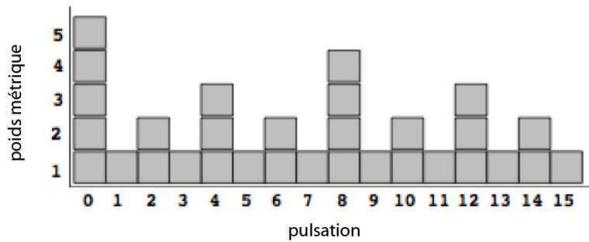
$$R = \bigcup_{i=0}^q P_i \quad (2)$$

$$\text{où } P_i = \bigcup_{n=0}^{E((m-of)/d)} (of_i + d_i \times n)$$

En ce qui concerne la complexité rythmique, nous utilisons un modèle basé sur la mesure de complexité de Toussaint ou autres techniques d'évaluation de complexité rythmique, qui attribue un poids à chaque pulsation (voir figure 4).

#### 3.4.4. Génération de mélodies

Pour générer la mélodie, MuZICO est doté d'un module prenant en entrée l'échelle de hauteurs courante et plusieurs motifs rythmiques générés selon les techniques expliquées précédemment. Chaque note a un poids correspondant à son ordre d'apparition dans la construction de l'échelle de hauteurs courante. Pour chaque coup du motif rythmique, une note est tirée au hasard selon une densité



**Figure 4.** poids des temps dans un pattern rythmique de taille 16.

de probabilité gaussienne variant en fonction du poids métrique correspondant. On obtient ainsi un profil mélodique initial qui pourra ensuite être développé musicalement par divers algorithmes tels que le PST (Probabilistic Suffix Tree). Ce profil est ensuite modifié en fonction des informations de modulation transmises par la grammaire générant les accords pour donner la mélodie qui sera jouée. La complexité d'une mélodie jouée de cette manière dépend des densités de probabilités gaussiennes et des paramètres de configuration du PST.

#### 4. CONCLUSION ET DISCUSSION

Le logiciel MuZICO a montré sa pertinence, d'une part dans la génération de musique en fonction des mesures physiologiques de la vigilance, d'autre part dans la génération de musique en fonction des émotions exprimées et représentées dans un espace à deux dimensions (énergie/valence). Une prochaine étape consiste donc à trouver à partir des données électroencéphalographiques un nouveau critère, soit par la mesure de l'évolution de la fréquence ou de l'amplitude des ondes cérébrales en un site correspondant à une activité cognitive, émotionnelle ou perceptive particulière (par exemple perception de la consonnance), soit par la mesure de l'évolution comparée des fréquences ou des amplitudes des ondes cérébrales de points situés sur chacun des hémisphères cérébraux. Nous entamons actuellement dans cette optique une collaboration avec des chercheurs en Interface Cerveau-Ordinateur. Nous pourrions ainsi générer la musique en fonction des émotions réellement ressenties et non plus exprimées par l'auditeur, et donc optimiser de manière significative la rétroaction, pour améliorer à la fois le pouvoir d'évocation de la musique et ses qualités esthétiques.

Une autre étape à venir sera d'affiner notre système au niveau de la synthèse sonore en faisant correspondre de manière plus précise aux émotions les paramètres musicaux de plus bas niveau (correspondant aux paramètres perceptifs des sons utilisés ou générés), par une meilleure gestion des bornes, en se conformant à un namespace au format url similaire à celui utilisé dans le protocole OSC. Une piste de validation alternative pour le projet est d'utiliser des algorithmes d'apprentissage par renforcement [4] afin de déterminer les combinaisons de paramètres les plus évocatrices d'émotions particulières. Des expérimentations de cet ordre sont en cours depuis le début du projet avec

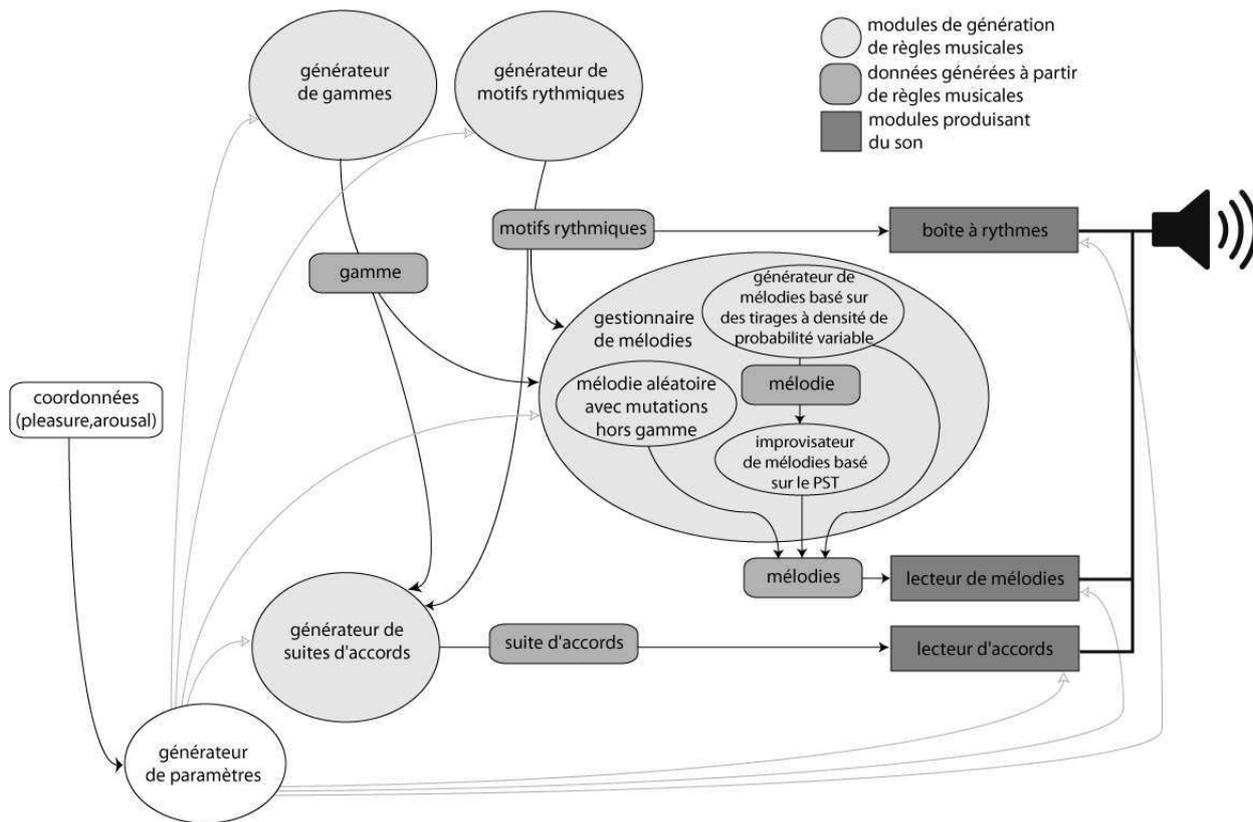
ce type d'outils logiciels.

## Remerciements

Nous remercions Pierre Héricourt, ingénieur au service système du LaBRI, pour le développement du driver permettant la liaison entre le boîtier ProComp5 et les différents logiciels utilisés (Pure Data pour la génération du son et l'analyse des données physiologiques, et OpenVibes, bibliothèque d'outils logiciels d'interaction cerveau-ordinateur, pour une analyse plus fine des signaux EEG).

## 5. REFERENCES

- [1] A. Clay, N. Couture, E. Decarsin, M. Desainte-Catherine, P-H. Vulliard, and J. Larralde. Movement to emotions to music : using whole body emotional expression as an interaction for electronic music generation. 2012.
- [2] A. Clay, G. Domenger, N. Couture, J.B. De La Rivière, M. Desainte-Catherine, et al. Spectacle augmenté : le projet care, un processus de recherche. 2011.
- [3] P. Ekman. An argument for basic emotions. *Cognition & Emotion*, 6(3-4) :169–200, 1992.
- [4] S. Le Groux and P.F.M.J. Verschure. towards adaptive music generation by reinforcement learning of musical tension. 2010.
- [5] S.R. Livingstone, R. Muhlberger, A.R. Brown, and W.F. Thompson. Changing musical emotion : A computational rule system for modifying score and performance. *Computer Music Journal*, 34(1) :41–64, 2010.
- [6] J. Paty. Les états de conscience. *Psychologie médicale*, 1985.
- [7] J.A. Russel. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6) :1161–1178, 1980.
- [8] L. Vezard, M. Chavent, P. Legrand, F. Faight-Ainseba, J. Clauzel, et al. Caractérisation d'états psychophysologiques par classification de signaux eeg. intégration de ces résultats dans le projet psi. 2011.
- [9] I. Wallis, T. Ingalls, and E. Campana. Computer-generating emotional music : The design of an affective music algorithm. In *International Conference on Digital Audio Effects*, pages 7–12, 2008.



Annexe. Schéma fonctionnel de la partie de MuZICO gérant les règles de composition musicale.