

# Journées Jeunes Chercheurs en **A**udition, **A**coustique musicale et **S**ignal audio

8<sup>EME</sup> EDITION

5, 6 ET 7 DECEMBRE 2012

MARSEILLE



CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE  
LABORATOIRE DE MECANIQUE ET D'ACOUSTIQUE  
UPR – 7051  
31, CHEMIN JOSEPH AIGUIER  
13402 MARSEILLE, CEDEX 20  
FRANCE



## MERCI A TOUS LES PARTNAIRES DES JJCAAS



AINSI QU'A TOUS LES CONFERENCIERS  
ET PARTICIPANTS

## PROGRAMME

---

### MERCREDI 5

- 10h30 – 11h15      Accueil des participants
- 11h30 – 12h      Présentation du LMA et des thématiques de l'équipe Sons par **CHRISTOPHE VERGEZ**, Chercheur au LMA et responsable de l'équipe Sons
- 12h – 13h      Déjeuner
- 13h15 – 14h      **Richard KRONLAND-MARTINET** – Chercheur au LMA – CNRS – Marseille

#### *Contrôle perceptif de la synthèse sonore*

- 14h – 14h45      **DAMIEN HENRY** – Présentation d'**AUDIOGAMING**, entreprise spécialisée dans la synthèse sonore pour les jeux vidéos



- 15h – 17h30      Session Posters I
- 17h30 – 18h30      **YVES RIESEL** – Présentation de **qobuz**, plateforme de streaming en ligne haute fidélité



### JEUDI 6

- 9h – 9h30      Accueil des participants
- 9h30 – 12h      Session Posters II
- 12h – 13h      Déjeuner
- 13h15 – 14h15      **DANIELE SCHÖN** – Chercheur à l'Institut de Neurosciences des Systèmes - INSERM – Marseille

#### *Musique et plasticité cérébrale*

14h30 – 17h Session Poster III

19h – 23h Dîner de gala

## VENDREDI 7

9h – 9h30 Accueil des participants

9h30 – 11h30 Session Posters IV

11h30 – 12h **PATRICK BOUSSARD** – Présentation de **GENESIS**, Leader en France sur les simulateurs audio et les systèmes de réécoute 3D



12h – 13h Déjeuner

13h15 – 14h15 **STEFAN BILBAO** – Chercheur au département de recherche en musique à l'Université d'Edimbourg

*Synthèse numérique par modèle physique*

14h30 – 17h Session Poster V

17h – 17h30 Clôture des JJCAAS

## CONFERENCIERS

---

**Richard Kronland-Martinet** – Chercheur au Laboratoire de Mécanique et d'Acoustique de Marseille – CNRS UPR 7051

Contrôle perceptif de la synthèse sonore

La synthèse numérique des sons a plus de cinquante ans. Elle permet aujourd'hui de reproduire des sons et des timbres de nature très variée en s'appuyant sur des modèles physiques des sources ou sur des conformations morphologiques du signal. Si les méthodes de synthèse ont, pour l'essentiel, atteint un degré de maturité qui en font aujourd'hui des outils précieux pour l'ingénierie et la recherche, le contrôle de

ces processus reste, quant à lui, un des grands enjeux actuels. En effet, les sons de synthèse sont le plus souvent définis par les paramètres propres aux modèles utilisés, ce qui les rend difficiles à manipuler, du moins en terme d'évocation. Dans cette présentation, nous verrons comment le concept d'objet sonore peut se décliner en un paradigme simple de caractérisation perceptive des sons, basé sur la notion de source résonante et de geste producteur. L'hypothèse d'existence d'invariants perceptifs propres à chacun de ces attributs conduit à la construction de systèmes de contrôles basés sur des descripteurs acoustiques dont l'impact perceptif est naturel. Des exemples de tels descripteurs seront donnés dans le cadre de la caractérisation du matériau constitutif de l'objet résonant et pour deux actions et gestes particuliers : rouler et frotter. Cette notion d'invariant nous conduira naturellement au concept de métaphore sonore et à de nouveaux paradigmes de contrôle pour les sons environnementaux et musicaux.

**Daniele Schön** - Chercheur à l'Institut de Neurosciences des Systèmes de Marseille - INSERM U 1106

Musique et plasticité cérébrale

Pendant toute notre vie, l'interaction avec l'environnement affecte d'une manière durable notre comportement, principalement en modifiant les connexions entre neurones existant. Lorsque cette rencontre entre notre organisme et environnement est motivée par un désir d'interaction et qu'elle est suffisamment « intense », ces nouvelles connexions se stabilisent, donnant lieu à nos nouveaux acquis, connaissances, futurs souvenirs... La plasticité devient dès lors primordiale, et l'état stable devient un artefact pour satisfaire les besoins des scientifiques.

Le cerveau du musicien est particulièrement intéressant pour l'étude de la plasticité car le musicien commence souvent très tôt la pratique musicale et continue avec une grande régularité et entêtement. Ainsi certaines aires cérébrales sont structurellement modifiées par la pratique musicale, c'est à dire que l'anatomie même du cerveau du musicien est différente. D'autres structures sont fonctionnellement modifiées, c'est à dire que leur fonctionnement est affecté par la pratique musicale, sans qu'on puisse remarquer des changements visibles au niveau anatomique.

L'aspect qui m'intéresse de plus, au delà des changements qui ont lieu au niveau du système moteur et du système auditif primaires (i.e. les entrées auditives et les sorties motrices) sont les effets de la pratique musicale sur certaines compétences du langage. En d'autres termes, la possibilité que les modifications induites par la musique aient un effet sur des compétences et mécanismes généraux tels que la perception de sons complexes, leur organisation dans une structure temporelle, leur mémorisation et bien d'autres fonctions cognitives qui sont communes à la musique, au langage et peut être aussi à d'autres domaines d'expertise.

**Stefan Bilbao** - Chercheur au département de recherche en musique à l'Université d'Edimbourg

### Synthèse numérique par modèle physique

Parmi les méthodes souvent utilisées dans des applications de synthèse sonore par modèles physiques (mieux connus sont les guides d'onde numérique, et les approches modales), les techniques basées sur des maillages spatio-temporels sont celles qui permettent la modélisation avec le moins d'hypothèses, ce qui est important surtout dans le cas des systèmes fortement non linéaires. En même temps, elles sont normalement légèrement plus coûteuses que d'autres techniques, et le problème d'assurer la stabilité numérique devient difficile à surmonter. Plusieurs systèmes intéressants provenant de l'acoustique musicale seront présentés comme point de départ pour l'élaboration des enjeux algorithmiques. Enjeux investigués dans le cadre d'un nouveau projet ERC (NESS) qui aura lieu à Edimbourg entre 2012 et 2016. Des exemples sonores seront présentés.

---

**Session I - Mercredi 5 – 15h – 17h30**

<b>Maïté Castro</b>	Les effets bénéfiques de la musique familière sur le traitement cognitif des patients dans le coma
<b>Jérémy Danna</b>	Projet MetaSon : Sonifier l'écriture manuscrite pour améliorer le diagnostic et la remédiation de ses troubles
<b>Marion David</b>	Cocktail-party dans les salles : ségrégation séquentielle de la parole, indices binauraux et réverbération
<b>Idir Edjekouane</b>	La sonie des signaux non-stationnaires et sonie binaurale. Application à la téléphonométrie
<b>Patrice Guyot</b>	Reconnaissance automatique de son d'eau, une approche acoustique et perceptive
<b>Paul Luizard</b>	Les volumes couplés : comportement, conception et perception dans un contexte de salle de spectacles
<b>Jérôme Nika</b>	Intégration de contrôles harmoniques et gestion de la pulsation dans un processus informatique d'improvisation musicale
<b>Thibaut Meurisse</b>	Contrôle actif appliqué aux instruments à vent
<b>Mélodie Monteil</b>	Effets des procédés de fabrication sur la dynamique des steelpans. Etape 1 : l'embouti
<b>Arthur Paté</b>	Vibro-acoustique de la guitare électrique
<b>Camille-Eva Rakovec</b>	Synthèse sonore interactive pour la réalité virtuelle et le multimédia
<b>Etienne Thoret</b>	Approche <i>énactive</i> des relations son/geste pour le contrôle de la synthèse sonore et le guidage

<b>Sebastian Basterech</b>	Évaluation de la qualité d'un flux de voix transmis sur l'Internet en utilisant des machines de calcul avec de réservoir de neurones
<b>Alexis Benichoux</b>	Séparation de sources audio en milieux échoïques
<b>Louis Delebecque</b>	Modélisation physique de plosives bilabiales
<b>Charles De Paiva Santana</b>	La pièce musicale comme événement d'un système complexe, ouvert et modélisable: un cas de musicologie assistée par ordinateur
<b>Kaoutar El Ghali</b>	Figures sonométriques : catégorisation, segmentation, fouilles de données sonores
<b>Alexandre Hyafil</b>	Modèles computationnels du rôle des oscillations corticales dans la segmentation de la parole
<b>Balbine Maillou</b>	Caractérisation et identification de non linéarités de systèmes acoustiques
<b>Julian Palacino</b>	Outils de création de contenus audio spatialisés pour les terminaux mobiles grands publics
<b>Gaëtan Parseihian</b>	Localisation et mouvement de pointage vers des sons dans l'espace péripersonnel
<b>Alberto Torin</b>	Modèles physiques d'instruments à percussion
<b>Karim Youssef</b>	Perception binaurale pour l'analyse de scène auditive en robotique. Localisation binaurale de sources sonores

<b>Roman Auvray</b>	Etude acoustique et hydrodynamique des instruments de la famille des flûtes : influence du système amont sur la production du son
<b>Victor Benichoux</b>	A new model of sound localization in the cat based on realistic binaural cues
<b>Sophie Bouton</b>	Morphing des transitions formantiques de la parole naturelle à l'aide la méthode LPC
<b>Sébastien Denjean</b>	Influence du retour sonore sur la perception du mouvement automobile en accélération longitudinale
<b>Yo Fujiso</b>	Etude aéroacoustique des écoulements d'air à nombre de Reynolds modéré dans les voies aériennes supérieures chez l'humain - application à la production de sons de consonnes fricatives non voisées
<b>Sami Karkar</b>	Morphing des transitions formantiques de la parole naturelle à l'aide la méthode LPC
<b>Marc Michau</b>	Méthode des impédances mécaniques virtuelles pour le contrôle actif vibroacoustique d'un panneau aéronautique
<b>David Poirier-Quinot</b>	Navigation guidée par bonification des signaux GSM dans un contexte de secours
<b>Emmanuel Ponsot</b>	Processus temporels perceptifs et cognitif influençant la sonie des sons non stationnaires : application aux sons environnementaux
<b>François Rigaud</b>	Analyse des sons de piano par factorisation en matrices non-négatives avec contrainte d'inharmonicité
<b>Pierre Stahl</b>	L'implant cochléaire: Effet de la cadence de stimulation sur la perception de durée chez le sujet normo-entendant
<b>Michael Vannier</b>	Sonie des sons stationnaires en situation dichotique

---

**Session IV - Vendredi 7 – 9h30 – 11h30**

<b>Youssef Adel</b>	Modélisation de l'intelligibilité de la parole dans le bruit pour les implantés cochléaires
<b>Simon Bennachio</b>	Contrôle actif modal appliqué aux instruments de musique à cordes
<b>Alexis Guilloteau</b>	Etude de l'influence de la géométrie d'un trou de note sur la puissance rayonnée par un instrument à anche simple
<b>Brian Hamilton</b>	Méthode des différences finies appliquée à la modélisation de l'acoustique des salles
<b>Thibaud Leclere</b>	Effet de précedence et restitution multicanal : mise en place d'un modèle 3D de masquage de réponses impulsionnelles
<b>Gabriel Sargent</b>	Estimation de la structure des morceaux de musique par analyse multi-critères et sous contrainte de régularité
<b>Jean-François Sciabica</b>	Interaction entre la perception du bruit moteur et la dynamique du véhicule dans un habitacle automobile.
<b>Soizic Terrien</b>	Instruments de la famille des flûtes : apport des méthodes de continuation numérique à l'interprétation d'observations expérimentales
<b>Marc Evrard</b>	Reproduction de la personnalité vocale d'un acteur

---

**Session V - Vendredi 7 – 14h30 – 17h**

<b>Louis Bigo</b>	Représentations symboliques de structures musicales à l'aide du calcul spatial
<b>Ning Chu</b>	Robuste bayésienne super-résolution approche par a priori parcimonie pour l'imagerie aéroacoustique
<b>Charles Hudin</b>	Application du retournement temporel à la restitution tactile
<b>Simon Conan</b>	Recherche des invariants acoustiques liés à l'évocation d'événements sonores particuliers. Une application à la synthèse et au contrôle des sons de roulements
<b>Gilles Courtois</b>	Predict phantom source localization within stereophonic reproduction
<b>Ange Ebissou</b>	Intelligibilité en espace ouverts : étude de la gêne en présence de plusieurs sources de parole
<b>Benjamin Elie</b>	Caractérisation vibratoire et acoustique des instruments à cordes - Application à l'aide à la facture instrumentale
<b>Augustin Lefevre</b>	Méthodes d'apprentissage de dictionnaire pour la séparation de sources mono-canal
<b>Olivier Perrotin</b>	Chironomie et interfaces pour les instruments musicaux virtuels
<b>Quentin Mesnildrey</b>	Simulations acoustiques de l'implant cochléaire : Modélisation de la configuration géométrique des électrodes et effets sur la perception de la parole
<b>Camille Vauthrin</b>	Technique de jeu et production du son dans les instruments de musique à sons entretenus : application aux flûtes

# Résumés des présentations

---

Session I - Mercredi 5 – 15h – 7h30

## **Maité Castro - Les effets bénéfiques de la musique familière sur le traitement cognitif des patients dans le coma. Les effets bénéfiques de la musique familière sur le traitement cognitif des patients dans le coma**

La musique en tant qu'élément familier à forte charge émotionnelle a des effets bénéfiques sur le fonctionnement cognitif normal et également sur celui de certains patients cérébro-lésés. Par exemple, il a été montré qu'une courte exposition musicale augmente l'attention visuelle dans le champ négligé de patients héminégligents. L'objectif de notre étude est de déterminer si la musique familière peut augmenter le traitement cognitif et le niveau de conscience des patients dans le coma. L'électroencéphalogramme de patients dans le coma et de participants 'sains' a été enregistré en réponse à 10 séquences de prénoms, contenant 7 prénoms non-familiers et le propre prénom du patient ou participant. Ces séquences étaient précédées soit d'une musique familière choisie par les participants ou sélectionnée par des proches pour les patients (5 séquences), soit d'un bruit de caractéristiques fréquentielles proches (5 séquences). Les résultats confirment la possibilité d'évoquer une composante P300 en réponse au propre prénom du patient, reflet d'un processus cognitif préservé malgré une altération de la conscience. Ils montrent aussi que, chez certains patients dans le coma, la présentation préalable de musique familière augmente les chances d'évoquer une composante P300 de caractéristiques proches de celles observées chez les participants 'sains'. En effet, la composante P300 a une amplitude plus importante et une latence plus faible lorsqu'une séquence de musique familière précède la série de prénoms par rapport à une séquence de bruit. L'hétérogénéité des résultats entre patients pourrait être expliquée par la diversité étiologique des patients. Les résultats seront également discutés au regard des hypothèses qui proposent que la musique préférée induirait un état émotionnel positif lié à une augmentation d'éveil et qu'elle stimulerait des réseaux neuronaux liés à la mémoire autobiographique.

## **Jérémy Danna - Projet MetaSon : Sonifier l'écriture manuscrite pour améliorer le diagnostic et la remédiation de ses troubles.**

L'écriture est une activité silencieuse, si on excepte le léger crissement du stylo qui n'informe que très marginalement sur la qualité du mouvement produit. La transformer en une activité audible peut paraître curieux en première analyse. En y réfléchissant mieux, justement parce que la modalité auditive est disponible pendant l'écriture, et parce qu'elle est mieux à même de permettre la perception de différences fines au plan temporel et dynamique, utiliser des sons pour renseigner sur ce mouvement semble judicieux. L'objectif est donc de comprendre comment

l'ajout d'informations auditives permettrait d'améliorer le diagnostic et la remédiation de troubles affectant l'écriture (par exemple la dysgraphie développementale).

### **Marion David - Cocktail-party dans les salles : ségrégation séquentielle de la parole, indices binauraux et réverbération**

L'analyse de scènes auditives permet de dresser une représentation utile et fidèle de notre environnement sonore, en regroupant les éléments correspondant à chaque source afin de les identifier. Par exemple, dans un contexte de cocktail-party, un auditeur cherche à comprendre le message délivré par une personne cible alors qu'une autre personne parle en même temps. L'enjeu est alors de pouvoir séparer en deux flux distincts les sons provenant alternativement de la cible et de la source concurrente pour que la source cible soit intelligible. Dans les salles, la réverbération entraîne des distorsions temporelles et spectrales des signaux (coloration). Ce filtrage opéré par la salle va induire des différences spectrales fines sur les signaux acoustiques. Les sons provenant d'une source acoustique à une position donnée vont également être colorés par les propriétés de diffraction et de réflexion de la tête, de la forme des oreilles et du torse (Head Related Transfer Function). Dans cette étude, il était question d'évaluer l'influence de la coloration des sources sur la ségrégation des flux auditifs. Pour cela, trois expériences ont été réalisées mettant en jeu tout d'abord la contribution de l'ensemble tête-salle, puis celles de la tête et de la salle séparément. Les résultats ont montré que les différences spectrales induites par la coloration des sources par la tête et la salle ont un effet significatif sur la ségrégation.

### **Idir Edjekouane - La sonie des signaux non-stationnaires et sonie binaurale. Application à la téléphonométrie**

Le premier objectif de ces travaux est de proposer un outil adapté aux signaux non-stationnaires, en particulier la parole. Le second est de proposer un outil qui permette d'estimer la sonie binaurale (signal perçu par les deux oreilles). En effet, les outils existant aujourd'hui permettent d'estimer la sonie perçue par un utilisateur utilisant un combiné téléphonique sur une seule oreille (monaural) alors que les utilisateurs de terminaux utilisent de plus en plus les modes mains libres, casque ou téléconférence (contexte binaural donc) qui modifie énormément le ressenti. Ces deux sujets liés à la sonie sont en train d'émerger et sont des sujets de recherche très actuels.

### **Patrice Guyot - Reconnaissance automatique de son d'eau, une approche acoustique et perceptive**

La détection acoustique de sons d'eau est aujourd'hui utilisée pour la reconnaissance d'activités dans le contexte d'assistance aux personnes âgées. Une des difficultés de cette tâche réside dans le fait que les sons d'eau peuvent prendre des formes

acoustiques très différentes, même s'ils peuvent être modélisés par les mêmes phénomènes physiques. Un système de détection de sons d'eau basé sur des modèles physiques sera présenté, ainsi qu'une étude effectuée à l'IRCAM et visant à établir des catégories de son d'eau validées perspectivement.

### **Paul Luizard - Les volumes couplés : comportement, conception et perception dans un contexte de salle de spectacles**

Les salles de concert basées sur le principe des volumes couplés sont des lieux dont l'acoustique est variable et sont en quelque sorte capables de s'accorder, à la manière d'un instrument de musique, en fonction du type de spectacle joué. Des échanges d'énergie acoustique ont lieu entre les différents volumes architecturaux, produisant des champs acoustiques complexes. Un modèle statistique de la réverbération couplée est proposé, comparé à des mesures sur maquette acoustique et en salle réelle ainsi qu'à des simulations numériques. De plus, l'aspect perceptif étant la finalité de la création d'espaces de spectacles, les résultats d'une étude de discrimination et préférence liées à la surface de couplage qui lie les volumes seront présentés.

### **Jérôme Nika - Intégration de contrôles harmoniques et gestion de la pulsation dans un processus informatique d'improvisation musicale**

Le système ImproteK propose d'intégrer un cadre rythmique et une structure harmonique sous-jacente dans un contexte d'improvisation. Dans la filiation du logiciel d'improvisation OMax développé à l'Ircam, il repose sur la structure d'oracle des facteurs pour tirer profit des propriétés particulièrement riches et pertinentes de cet automate dans un contexte musical. Au cours d'une session d'improvisation, ce système peut s'adapter à une pulsation régulière et proposer des phrases musicales suivant une progression harmonique donnée. ImproteK est conçu comme un instrument interactif dédié à la performance : ses improvisations sont construites à partir de la modélisation stylistique réalisée sur le jeu live d'un musicien ou sur un corpus offline. Associée à des techniques considérant le corpus comme une mémoire musicale, cette modélisation s'étend aux domaines de l'harmonisation et de l'arrangement dans un module d'interaction harmonique.

### **Thibaut Meurisse - Contrôle actif appliqué aux instruments à vent**

Les musiciens de tout temps ont été intéressés par faire évoluer leur instrument afin d'améliorer leur qualité et de pouvoir mieux répondre aux besoins des compositeurs. Il en est de même aujourd'hui où la fabrication des instruments non électroniques (non électroacoustiques) semble en retrait de l'évolution de la musique. Cette étude vise à modifier la qualité sonore et la jouabilité d'instruments à vent ainsi qu'à étendre leurs possibilités en utilisant le contrôle actif.

## **Mélodie Monteil - Vibrations non linéaires géométriques de coques minces : Effets des procédés de fabrication et réponses temporelles à des chocs - applications aux steelpans**

Ma thèse porte sur les vibrations non linéaires géométriques de coques minces ainsi que sur les procédés de leur fabrication. Ces deux parties du travail sont menées simultanément dans le but d'obtenir une modélisation permettant de qualifier et de quantifier l'impact des étapes de fabrication (déformation par martelage, façonnage, poinçonnage et chauffage) sur le comportement vibratoire linéaire et non linéaire.

Les coques minces, utilisées dans de nombreux domaines de l'ingénierie, sont souvent sujettes à de fortes sollicitations qui engendrent des vibrations de grande amplitude dont résultent des comportements typiquement non linéaires. En effet, il est reconnu que, pour les coques minces, des amplitudes vibratoires de l'ordre de l'épaisseur sont suffisantes pour donner lieu à des phénomènes typiquement non linéaires : coexistence de plusieurs régimes, phénomènes de saut, échanges d'énergie entre différents modes propres, dépendance de la fréquence de résonance à l'amplitude d'excitation... L'étude vibratoire se concentre sur les transferts d'énergie par couplage modaux via des résonances internes et a été développée de manière analytique, numérique et expérimentale. Le second thème de la thèse porte sur la modélisation des procédés de fabrication. Plus particulièrement, l'effet des grandes déformations plastiques (une centaine de fois l'épaisseur) lors de l'emboutissage d'une plaque plane transformée en coque sphérique, sur les caractéristiques vibratoires, a été mis à l'étude.

## **Artur Paté - Vibro-acoustique de la guitare électrique**

Le son rayonné par une guitare électrique provient essentiellement du système électro-acoustique amplificateur du mouvement de la corde. Pour étudier les différences sonores entre instruments, les études se sont légitimement focalisées sur les différents "microphones" (pickups) captant le mouvement de la corde.

Cependant le comportement dynamique de la structure de la guitare est important du fait du couplage des cordes au chevalet et aux frettes. Les éléments de lutherie (manche, corps, chevalet, jonction corps/manche, ...) jouent donc un rôle dans la vibration de la corde. Nous cherchons donc, dans le cadre de ce travail de thèse, à quantifier et à hiérarchiser ce rôle par le biais de modélisations et d'études expérimentales.

## **Camille-Eva Rakovec - Synthèse sonore interactive pour la réalité virtuelle et le multimédia**

Dans le cadre de la réalité virtuelle, un des enjeux scientifiques actuels réside dans la prise en compte des actions de l'utilisateur dans le processus de synthèse sonore afin d'augmenter la sensation d'immersion et de réalisme. Dans ce contexte, le travail engagé a pour objectif d'aborder les problématiques spécifiques aux sons produits

par la manipulation d'objets de l'environnement. Il s'agit dans un premier temps de mieux comprendre la perception auditive des caractéristiques physiques de ces sources (forme, taille, présence de cavité ...) et de définir des catégories perceptives propres à ces caractéristiques par des tests d'écoute. Les premiers résultats ont ainsi permis de mettre en évidence deux grandes catégories de matériaux perçus, parmi lesquels certains sont plus faciles à reconnaître que d'autres, mais aussi de montrer que la catégorisation des formes constitue une tâche relativement complexe pour les sujets. Suite à ces observations, l'enjeu consiste à déterminer des "invariants", ou "descripteurs acoustiques", responsables des évocations induites, en s'appuyant sur des considérations physiques et sur la caractérisation acoustique des signaux produits. La mise en place d'un modèle prédictif sur le principe de la régression logistique binaire, incluant une étape de validation et de calibration doit alors être engagée. Nous avons ainsi pu évaluer le niveau de pertinence des invariants retenus et déterminer les catégories perceptives nécessitant une description acoustique plus approfondie. En se basant sur ces résultats perceptifs et acoustiques, la finalité du travail consistera à mettre en œuvre des stratégies de contrôle perceptif et interactif, en vue de proposer des outils de synthèse sonore en temps réel pour la réalité virtuelle et augmentée (synthétiseur générique de sons d'environnement par exemple).

### **Etienne Thoret - Approche énaactive des relations son/geste pour le contrôle de la synthèse sonore et le guidage**

Dans le cadre de ce travail, nous nous intéressons aux relations acoustiques et perceptives qui existent entre les différents paramètres dynamiques d'un geste, et les morphologies sonores intrinsèques aux sons générés par ce geste. En s'appuyant sur des outils de synthèse sonore de bruits de frottements contrôlés par des paramètres dynamiques (profil de vitesse et pression), l'intégration sensori-motrice des différentes composantes d'un geste est étudiée par le biais de la modalité auditive. Par exemple, nous avons pu mettre en évidence l'importance du profil de vitesse et le fait que ce profil devait présenter des caractéristiques particulières pour l'évocation d'un geste humain naturel qui dessine. De plus, nous avons montré que dans la modalité auditive, cette information cinématique était également reliée à une information spatiale car elle permet, dans une certaine mesure, d'associer correctement des formes géométriques qui ont été dessinées aux sons générés, par la médiation du geste évoqué sous-jacent. Ces premiers résultats vont dans le sens d'une incarnation des caractéristiques cinématiques des mouvements humains à un très haut niveau cognitif. Cette hypothèse a déjà été suggérée par différentes études, en particulier dans Leman 2007. Ils viennent donc renforcer l'idée que la cinématique d'un geste humain est à la base des processus cognitifs mis en jeu pour se représenter (s'imaginer) « mentalement » un geste. Les perspectives de ces travaux préliminaires seront présentées. De plus, ils permettent d'envisager la possibilité d'évoquer ces caractéristiques par des morphologies sonores bien spécifiques. Outre les problématiques fondamentales soulevées par cette étude, des applications telles

que le développement d'outils pour le guidage d'un geste par le biais du son est envisagé notamment dans le contexte du projet ANR MetaSon. Les différentes applications seront présentées.

---

## **Session II - Jeudi 6 – 9h30 – 12h**

### **Sebastian Basterech - Évaluation de la qualité d'un flux de voix transmit sur l'Internet en utilisant des machines de calcul avec de réservoir de neurones**

Au cours de ces dernières années, un nouveau paradigme a été introduit dans le domaine de l'apprentissage automatique sous le nom de machines de calcul avec de réservoir de neurones (son nom en anglais: Reservoir Computing (RC)). Les modèles de RC consistent de deux parties: un réseau des neurones bouclés (possédant des circuits) qui ne participe pas à la procédure d'apprentissage, et un outil d'apprentissage automatique rapide et robuste (dont l'architecture n'a pas de circuit, ce qui lui permet de travailler vite). La croissance de ces méthodes a été rapide en raison de leur succès dans la résolution de problèmes d'apprentissage automatique et dans d'autres applications informatiques. Notre contribution est la proposition de nouvelles idées pour la mesure de la qualité perceptuelle. La qualité d'un flux audio ou de voix, transmit sur l'Internet, est fortement affectée par l'état du réseau. Nous avons développé un outil pour mesurer la qualité de la voix transmise par un système de télécommunications à l'aide des outils de RC. Ceci permet d'estimer la qualité des flux Voice sur IP de façon automatique, en temps réel et sans avoir recours aux signaux d'origine.

### **Alexis Benichoux - Séparation de sources audio en milieux échoïque**

La séparation de source aveugle en milieu échoïque nécessite l'acquisition d'une part de la réponse acoustique de salle et d'autre part des sources émises. Un approche consiste à formuler la solution comme le minimum d'une fonction de coût qui prend en compte l'attache aux données ainsi qu'un a priori sur les sources et sur la salle. On montre que dans un grand nombre de cas, une approche monocapteur ne permet pas d'obtenir la caractérisation de la séparation comme un minimum global, et on présente une approche alternative basée sur la recherche d'un minimum local.

### **Louis Delebecque - Modélisation physique de plosives bilabiales**

Ces travaux portent sur la modélisation mécanique et aéroacoustique des plosives bilabiales (/p/ et /b/). L'objectif est de différencier, dans la production, ce qui est dû au contrôle du locuteur de ce qui est dû à la physique de la parole.

La démarche adoptée consiste, dans un premier temps, à identifier, par le biais de mesures sur des locuteurs, les phénomènes physiques significatifs dans la production

d'une séquence voyelle-plosive bilabiale-voyelle. Ensuite, un modèle théorique pour décrire l'écoulement d'air dans les lèvres est proposé puis validé expérimentalement sur une maquette de l'appareil phonatoire. Enfin, la validation de ce modèle permet de réaliser des simulations numériques, basées sur une description bi-dimensionnelle de l'écoulement et sur un modèle à deux masses pour décrire le comportement mécanique des plis vocaux.

La confrontation entre la simulation et la mesure met en évidence la nécessité de modéliser l'expansion des joues lors de la production de plosives bilabiales.

### **Charles De Paiva Santana - La pièce musicale comme événement d'un système complexe, ouvert et modélisable: un cas de musicologie assistée par ordinateur**

Les compositions musicales peuvent être considérées comme des systèmes complexes et autonomes, et ainsi faire l'objet de modélisations informatiques avancées. Nous proposons de développer cette idée en réalisant des modèles formels permettant l'étude de pièces du répertoire musical contemporain, et de concevoir les outils informatiques correspondants. Nous nous appuyons notamment sur des notions inspirées de la programmation orientée objet afin de modéliser, produire et évaluer différentes instances et configurations de tels systèmes.

### **Kaoutar El Ghali - Figures sonométriques : catégorisation, segmentation, fouilles de données sonores**

Le travail de recherche que nous entreprenons s'articule autour de l'idée qu'il existe des figures temporelles sonométriques qui, au même titre que les figures géométriques, permettent à l'être humain de segmenter, caractériser, identifier et catégoriser une grande variété d'événements sonores à partir de l'organisation morphologique et dynamique du flux sonore. La « théorie des objets musicaux » de Pierre Schaeffer offre un terrain de réflexion à bon nombre de compositeurs, musicologues et chercheurs. Chez Schaeffer, l'objet sonore désigne « un phénomène sonore perçu dans le temps comme un tout, une unité, quels que soient ses causes, son sens, et le domaine auquel il appartient ». Parmi les travaux conduits dans ce sens, nous nous intéressons à la typologie sémantique des segments musicaux que sont les Unités Sémiotiques Temporelles (UST) du laboratoire Musique et Informatique de Marseille. Ces UST, qui sont au nombre de 19, constituent pour nous le point de départ à la détermination des figures sonométriques. Nous nous proposons de présenter cette problématique centrale ainsi que les axes de recherche qu'elle implique ; à savoir (i) la représentation des UST en figures sonométriques par le biais des caractéristiques sonores de ces premières, et (ii) la segmentation automatique d'œuvres musicales en UST. La modélisation des UST en figures sonométriques revient dans le cadre de notre étude à une catégorisation automatique de celles-ci ; cette catégorisation passe par l'étude des caractéristiques spectrales et temporelles des UST et par leur classification au vu de ces

caractéristiques. Le principal enjeu est donc de construire les descripteurs audio adéquats à cette tâche. La segmentation automatique en UST repose essentiellement sur le modèle des UST dégagé.

### **Alexandre Hyafil - Modèles computationnels du rôle des oscillations corticales dans la segmentation de la parole**

The question of how speech constituents (e.g. phonemes, syllables, words) are extracted from continuous acoustic input in the auditory cortex remains unsolved. The problem of categorization would be eased if the acoustic signal was parsed beforehand into relevant chunks (phonemic and/or syllabic). It has been recently proposed that delta/theta and gamma oscillations mediate nested chunking into respectively syllabic and phonemic bits. Here we present modeling work showing the feasibility of this scheme in biologically plausible neural networks. Theta and gamma oscillations are both generated by excitatory-inhibitory loops and receiving speech input from subcortical structures. We study the performance of theta oscillations in signaling syllable boundaries (comparing it to state-of-the-art offline algorithms), and in gamma oscillations in creating a robust neural code for latter phonemic categorization. We also relate the spectral and coding properties of the nested oscillatory network to a wealth of experimental findings in speech perception. Overall, our work aims at creating bridges between the research areas in human and automatic speech recognition.

### **Balbine Maillou - Caractérisation et identification de non linéarités de systèmes acoustiques**

Lors de la reproduction électroacoustique du son, l'influence du ou des hauts-parleurs peut être déterminante sur la qualité audio du signal rayonné, notamment en terme de distorsion harmonique. L'étude du haut-parleur électrodynamique montre que le comportement non linéaire des suspensions mécaniques intervient dans le phénomène de distorsion audio. Les travaux présentés traitent de la question de caractérisation expérimentale de ces non linéarités et de leur identification, paramétrique ou non.

### **Julian Palacino - Outils de création de contenus audio spatialisés pour les terminaux mobiles grands publics**

En vue de démocratiser le son 3D auprès du grand public, et parce que les terminaux mobiles sont devenus aujourd'hui un des principaux vecteurs de diffusion et d'accès aux contenus, cette étude/thèse vise à développer des outils de création de contenus audio 3D adaptés aux contraintes de ce type de terminaux, à la fois en termes de portabilité et de complexité. Pour la captation, la première solution étudiée repose sur des réseaux sphériques de microphones, selon le concept de la technologie Ambisonics. Pour pallier le faible nombre de capteurs imposé par le contexte des terminaux mobiles, des post-traitements de localisation de sources et

d'enrichissements de type upmix sont utilisés en complément de la prise de son. Pour la restitution, la première solution envisagée est la restitution sur casque du fait de sa portabilité et de sa compatibilité avec les terminaux mobiles. La technologie binaurale conjuguée à des pré-traitements associés, à la fois pour enrichir ou préciser l'information spatiale permet un rendu proche de l'écoute spatialisée naturelle. Les travaux présentés portent sur l'évaluation objective et subjective de différents dispositifs permettant une captation audio 3D en vue d'un rendu sur casque, et des pré et post-traitements associés.

### **Gaëtan Parsehian - Localisation et mouvement de pointage vers des sons dans l'espace péripersonnel**

Alors que de nombreuses études se sont intéressées à la localisation auditive pour des sources distantes, très peu ont étudié la perception de sources sonores situées à moins d'un mètre de l'auditeur. Cette étude examine la précision du geste de saisie vers des sons situés dans l'espace péripersonnel. Les performances de localisation sont évaluées pour des sources réelles et virtuelles et comparées aux performances de pointage vers des points lumineux.

### **Alberto Torin – Modèles physiques d'instruments à percussion**

Des modèles aux différences finies ont été utilisés dans le passé pour modéliser les vibrations non linéaires d'une plaque (équations de von Kármán). En synthèse sonore, ces modèles permettent de reproduire par exemple des sons des gongs et de cymbales. Une extension de ce système simulant explicitement l'interaction de ces structures avec l'air sera présentée. La preuve qu'il existe un schéma aux différences finies qui est énergétiquement conservatif, condition nécessaire pour la robustesse du code, sera également présentée. Plusieurs plaques non-linéaires peuvent être mises dans une *boîte* d'air et interagir les unes avec les autres. Dans le code implémenté, il est possible de changer tous les paramètres qui décrivent le système et de spécifier différents types d'excitation. De cette façon, on obtient un instrument virtuel qui peut être utilisé par un compositeur pour créer de la musique.

### **Karim Youssef - Perception binaurale pour l'analyse de scène auditive en robotique. Localisation binaurale de sources sonores**

Les robots capables de communiquer et d'interagir avec les hommes sont en train de devenir de plus en plus envisageables. Ces robots nécessitent des capacités auditives qui leur permettent de reconnaître, localiser et séparer les sources sonores. La localisation de sources sonores, une des capacités auditives indispensables pour de tels robots, a déjà été largement étudiée dans des contextes variés impliquant des antennes de microphones ou des systèmes binauraux. Ce travail s'intéresse au cas binaural et comporte deux parties. En premier lieu, une vue d'ensemble de différentes techniques d'extraction d'indices acoustiques binauraux est présentée. Ces techniques sont appliquées sur une base de données simulant une source

émettant des sons à partir de plusieurs positions par rapport à un récepteur binaural dans un environnement réverbérant. Les capacités des différents indices acoustiques calculés, en termes de discrimination des différentes positions des sources, sont analysées. En second lieu, un système de localisation binaurale de sources sonores est étudié. Il utilise les conclusions de l'étude précédente. Les indices interauraux de différences de phase et d'énergie dans plusieurs bandes fréquentielles sont exploités par un réseau de neurones afin d'estimer la position d'une source sonore dans une image captée par une caméra.

---

### **Session III - Jeudi 6 – 14h30 – 17h**

#### **Roman Auvray - Etude acoustique et hydrodynamique des instruments de la famille des flûtes : influence du système amont sur la production du son**

In flutes and flue organ pipes, the blowing pressure is often considered as a control parameter at time scales lower than the acoustic time scales. For instance, the typical time of a rise of pressure represents an objective descriptor to analyse attacks (typical time about 20 ms for fast attacks). It has been observed that the blowing pressure is also prone to oscillate at time scales of the order of the acoustic time scales. This might be due to the acoustic coupling between the instrument and the pressure reservoir. The present work investigates the influence of such a coupling on the sound production, and its pertinence from a musical point a view. In other words, can the ability of a musician (or an instrument maker) to control this coupling be regarded as a control parameter of the sound production ? This paper presents a preliminary experimental study focused on the effects of a pulsating blowing pressure on the sound production. The fluctuations of the blowing pressure are forced by using a loudspeaker within an artificial mouth. Different effects – such as modifications of the spectral enhancement or changes in the transients--, resulting from different supply and "coupling" conditions are presented.

#### **Marc Michau - Méthode des impédances mécaniques virtuelles pour le contrôle actif vibroacoustique d'un panneau aéronautique**

Afin de modifier la vibration d'une structure dans le but de réduire sa puissance acoustique rayonnée, on l'équipe de paires actionneurs/capteurs colocalisées. On peut alors contrôler localement l'impédance mécanique de la structure (masse, raideur, et amortissement virtuels). L'approche harmonique permet de considérer à chaque fréquence les impédances mécaniques virtuelles optimales à imposer localement. L'étude théorique a été faite dans le cas d'une plaque en appui simple. L'approche expérimentale consiste en deux étapes: 1) Recherche des impédances mécaniques virtuelles optimales expérimentales 2) Contrôle des impédances

mécaniques virtuelles.

### **Yo Fujiso - Etude aéroacoustique des écoulements d'air à nombre de Reynolds modéré dans les voies aériennes supérieures chez l'humain - application à la production de sons de consonnes fricatives non voisées**

L'air que nous respirons à travers notre conduit vocal est un ingrédient indispensable pour vivre et pour communiquer oralement. Malgré leur importance, peu d'études ont été faites sur les écoulements d'air à nombre de Reynolds modéré ( $100 < Re < 10000$ ) comme celui de la respiration humaine, contrairement aux écoulements à plus haut Reynolds typiques des applications industrielles (aéronautique, automobile...). Dans le cadre de l'étude de la production de sons de parole humaine, le comportement physique des écoulements d'air utilisés pour parler est en général considérablement simplifié pour modéliser la production de sons voisés (générés par la vibration des plis vocaux) tels que les voyelles. Or ces écoulements sont en réalité de nature très complexe car turbulents et sensibles aux conditions physiologiques environnantes naturellement fortement variables. Dans le cas de la production de sons non voisés comme certaines consonnes fricatives ([s] et [f]), une caractérisation plus fine des écoulements s'avère indispensable pour pouvoir modéliser correctement les phénomènes aéroacoustiques sous-jacents à la production de ces sons. A l'aide d'expériences sur maquettes in-vitro représentant de manière simplifiée certaines parties du conduit vocal humain, et de simulations numériques, l'objectif de cette thèse est de contribuer à la modélisation et la caractérisation aéroacoustique d'écoulements d'air à nombre de Reynolds modéré, avec application à l'étude de la production de sons de consonnes fricatives non voisées. Une attention toute particulière est portée à l'influence des conditions aux limites et de la turbulence sur les sons produits.

### **François Rigaud - Analyse des sons de piano par factorisation en matrices non-négatives avec contrainte d'inharmonicité**

Les méthodes d'analyse automatique de musique tendent de plus en plus à inclure des considérations musicologiques (modèle de tempo, de durée de note, ... ) et physique (harmonicité des sons par exemple) dans les modèles mathématiques de représentation de signaux. Dans le cas de la musique pour piano solo, une des difficultés réside dans la prise en compte de la nature inharmonique des sons. Ce phénomène est dû en grande partie à la raideur des cordes et se manifeste dans le spectre d'une note par une déviation des partiels par rapport à la série harmonique vers les fréquences aiguës, le spectre est "étiré". Nous présenterons une méthode d'analyse de sons de piano, prenant en compte cette inharmonicité, et basée sur les approches de type Factorisation en Matrices Non-Négatives (NMF). Celles-ci ont connu ces dernières années un vrai engouement car elles font espérer une décomposition automatique du spectrogramme, en spectres des notes jouées (ou atomes) et leurs activations dans le temps ; autrement dit, en la production

automatique d'un piano-roll (équivalent d'un fichier de contrôle MIDI). La particularité de notre approche repose sur l'introduction de l'inharmonicité par l'ajout d'une contrainte relaxée sur les fréquences de partiels des atomes du dictionnaire. L'application visée est l'estimation précise de la fréquence fondamentale et du coefficient d'inharmonicité de notes de piano à partir d'enregistrements de notes seules et d'accords (les notes sont supposées connues). Les résultats montrent une estimation précise et robuste sur une large partie de la tessiture, même dans le cas d'un degré de polyphonie élevé. A plus long terme, l'utilisation du formalisme de la NMF permet d'envisager des applications comme la transcription de musique pour piano solo et la séparation de sources/notes dans un mélange polyphonique.

### **Pierre Stahl - L'implant cochléaire: Effet de la cadence de stimulation sur la perception de durée chez le sujet normo-entendant**

Les implants cochléaires (IC) ont vu leur cadence de stimulation augmenter grandement ces 15 dernières années, dépassant les 5000 impulsions par seconde (ips) par canal. Cependant, aucune amélioration significative de reconnaissance de la parole n'a été observée pour des cadences supérieures à 500 ips. Une explication possible est que ces cadences de stimulation élevées induisent un effet d'adaptation dans le nerf auditif qui se traduit par une chute exponentielle du taux de décharge des fibres nerveuses auditives, ce qui empêche les neurones de suivre les fluctuations temporelles rapides des sons. Cet effet a déjà été observé, notamment en électrophysiologie par stimulation acoustique (Shackleton, 2008), électrique (Zhang, 2007) ou sur sujets équipés d'IC (Mc Cutcheon 2005). D'autre part, il a été prouvé qu'un signal possédant une enveloppe amortie (comparable au phénomène d'adaptation) est perçu comme étant plus court qu'un signal d'enveloppe stationnaire de même durée physique (Grassi and Darwin, 2006). Nous nous proposons de mettre en évidence cet effet par des mesures psychophysiques chez le sujet normo-entendant, par des tâches d'ajustement de durée de sons de cadence différentes: les sons de cadence élevée sont supposés être sous-estimés comparé à ceux de cadence faible. Contrairement à nos attentes, nous avons pu observer l'effet inverse, résultats que nous nous proposons de discuter par la suite.

### **David Poirier-Quinot - Navigation guidée par bonification des signaux GSM dans un contexte de secours**

- Navigation assistée par modalités acoustiques
- Environnement de simulation pour la Navigation dans un contexte de secours (CAVE)
- Chaîne Matlab - Max / MSP - Blender : de la simulation de propagation RF à la Navigation en Environnement simulé en passant par les métaphores de sonification

## **Michael Vannier - Sonie des sons stationnaires en situation dichotique**

L'objectif de l'étude est de contribuer à l'élaboration d'un modèle valide pour la prédiction de la sonie binaurale de sons stationnaires en situation fortement dichotique (situation faisant intervenir des différences interaurales), à partir d'enregistrements réalisés sur un mannequin acoustique. Sont présentées les mesures préliminaires d'HRTFs (Head Related Transfer Functions) réalisées sur 18 auditeurs volontaires et une tête artificielle, dans le but de réaliser la synthèse binaurale (individualisée ou non) de sources acoustiques au casque, pour différentes positions de l'espace ( $0^\circ < \text{Azimut} < 180^\circ$  ;  $0^\circ < \text{Elévation} < 90^\circ$ ). Des mesures de sonie, inspirées des travaux de V.P. Sivonen et W. Ellermeier, seront ensuite réalisées au casque à partir de ces sources virtuelles, permettant de tester les modèles existants (Moore et al. 1997 [J. Audio Eng. Soc., 45, 224-240]; Sivonen et Ellermeier 2006 [J. Acoust. Soc. Am., 119, 2965-2980] ; Moore et Glasberg 2006 [J. Acoust. Soc. Am., 121, 1604-1612]) avant de passer à des situations réalistes plus complexes.

## **Emmanuel Ponsot - Processus temporels perceptif et cognitif influençant la sonie des sons non stationnaires : application aux sons environnementaux**

L'influence du profil temporel d'un son sur sa sonie (intensité perçue) a été étudiée en psychoacoustique. En particulier, les sons croissants sont perçus plus forts que les sons décroissants (en niveau), et les causes de cette différence perceptive ne sont toujours pas clairement identifiées. En supposant que des pondérations temporelles perceptives distinctes entre ces deux types de sons sont à l'origine de la différence observée, une expérience psychoacoustique adoptant une approche « psychophysique moléculaire » a été conduite afin de déterminer les contributions respectives de différents segments temporels des sons croissants et décroissants sur leur sonie. Les stimuli employés, des sons purs d'une durée de 2 s, étaient constitués de 16 segments stationnaires juxtaposés, dont les niveaux variaient autour d'un profil temporel soit croissant [65-80 dB SPL], soit décroissant [80-65 dB SPL]. Pour chaque type de profil, une classification en deux groupes fort/faible a été réalisée dans une tâche de discrimination en sonie globale. Les réponses des auditeurs, corrélées aux différentes variations de chaque stimulus, ont permis d'extraire les « poids perceptifs » des 16 segments de chaque profil. Les premières analyses révèlent que ce sont globalement les 3 derniers segments des sons croissants, et à l'inverse, les 3 premiers segments des sons décroissants qui reçoivent des poids perceptifs significativement non nuls. Les patterns de pondération obtenus sont par ailleurs symétriques. Ces résultats suggèrent l'existence d'un phénomène perceptif commun de « saillance », où l'attention des auditeurs jugeant la sonie est exclusivement portée sur les segments de niveau sonore maximal, de la même façon pour les sons croissants et décroissants. L'approche moléculaire ne permettant pas d'expliquer les différences de jugement de sonie précédemment observées entre les sons croissants et décroissants, d'autres expériences doivent donc être menées afin d'identifier le(s) phénomène(s) responsable(s) de cet effet.

## **Sébastien Denjean - Influence du retour sonore sur la perception du mouvement automobile en accélération longitudinale**

Le développement de nouvelles motorisations électriques ou hybrides entraîne un profond bouleversement du retour sonore perçu par le conducteur à l'intérieur de l'habitacle automobile. La perte des informations portées par le bruit moteur peut changer le rapport à la conduite et modifier la perception du mouvement du conducteur. Nous cherchons à déterminer dans quelle mesure ces changements acoustiques modifient la perception de la dynamique du véhicule et le comportement des conducteurs. Pour cela, deux études ont été menées en simulateur de conduite, en accélération et en décélération. Les participants devaient accélérer ou ralentir pour atteindre une vitesse cible puis la maintenir constante, avec le compteur de vitesse masqué, dans trois conditions sonores : motorisation thermique, motorisation électrique ou sans aucun retour sonore. Ces études ont permis de montrer l'influence du retour sonore sur la perception des vitesses et des accélérations par les conducteurs. Sans retour sonore, les conducteurs ont tendance à produire une variation de vitesse plus grande qu'en présence de son. En présence de bruit moteur, les conducteurs présentent plus de facilités à maintenir une vitesse constante, grâce aux informations qu'il apporte sur les faibles accélérations.

## **Victor Benichoux - A new model of sound localization in the cat based on realistic binaural cues**

Mammals use mainly interaural time differences (ITD) at low frequencies to localize sound sources in the horizontal plane. Previous modeling studies have focused on the simple situation where this delay does not depend on the frequency of the input. In mammals, ITD sensitivity takes place in the binaural neurons of the MSO (Medial Superior Olive), where binaural inputs are first integrated, and further in the IC (Inferior Colliculus). In current models, binaural neurons compare the timing of two monaural inputs delayed by a fixed amount of time (e.g. due to axonal propagation delay). This is true of both the classical Jeffress model and the hemispheric model. However, this hypothesis contradicts the observation that the best delays of binaural neurons recorded in many mammalian species depend on input frequency. A simple way to measure the exact acoustical effect of the head and body on the incoming acoustical input is to record Head Related Transfer Functions (HRTFs), which yields a pair of filters for every position on a sphere around the subject. HRTF measurements on a variety of species show that ITD also significantly depends on frequency, and displays variations in the relevant low frequency range, as predicted by analytical results in spherical head models. Based on these observations, we propose an alternative hypothesis, in which each neuron encodes source position rather than a single ITD. This implies that neurons are selective to the interaural phase spectrum of their preferred position's HRTF. We demonstrate how this can be achieved with binaural neurons receiving inputs with mismatched preferred frequency and axonal

delay. Probing this model with inputs rendered through our measured HRTFs, we were able to predict the relationship between input frequency and best phase in binaural neurons. Finally, we show that these predictions agree with recordings in the cat's IC.

### **Sophie Bouton & Sami Karkar - Morphing des transitions formantiques de la parole naturelle à l'aide la méthode LPC**

La perception phonémique est un processus habituellement évalué en utilisant des stimuli qui varient sur une échelle physique continue. Par exemple, pour étudier la perception de phonèmes en fonction du lieu d'articulation (point où se produit l'obstruction dans le conduit vocal), on peut être amené à construire un continuum entre les sons /ba/, /da/ et /ga/. La plupart des études réalisées utilisent des syllabes produites par des synthétiseurs vocaux (Klatt, 1982) ou issues de parole naturelle modifiée par des techniques de morphing qui font varier simultanément l'ensemble des paramètres acoustiques (STRAIGHT, Kawahara et al., 1999). Ces solutions ne sont toutefois pas adaptées aux études concernant les corrélats neurologiques de la perception auditive, où l'on cherche à corrélérer chaque paramètre acoustique à un traitement cognitif donné, de manière indépendante. Dans cette perspective, nous proposons une méthode de morphing sonore originale, basée sur l'analyse/resynthèse LPC de sons de parole naturelle. La source issue de l'analyse LPC y est remplacée par une source synthétique identique pour tous les stimuli, ce qui permet ensuite de faire varier un unique paramètre acoustique (dans le cas présent, la transition du second formant) le long du continuum. On obtient ainsi un continuum sonore dont les extrêmes sont très proches des sons de parole naturelle originaux qui répond aux exigences d'une étude perceptive ciblée.

---

### **Session IV - Vendredi 7 – 9h30 – 11h30**

### **Youssef Adel - Modélisation de l'intelligibilité de la parole dans le bruit pour les implantés cochléaires**

Les utilisateurs de l'implant cochléaire (cochlear implant - CI) avec une stimulation électrique et acoustique dans la même oreille (electro-acoustic stimulation - EAS) montrent dans le bruit une intelligibilité augmentée de façon significative par rapport aux utilisateurs de CI conventionnel. Afin d'étudier l'influence de la basse fréquence acoustique, un modèle pour la reproduction de l'intelligibilité de ces groupes d'utilisateurs a été créé, en utilisant le test de Oldenburger et des signaux sonores qui simulent CI/EAS. Les taux de détection dans différentes fréquences acoustiques entre la stimulation acoustique et électrique ont été déterminés dans le silence et dans le bruit (Fastl, Oldenburger). Pour valider le modèle, les seuils de l'intelligibilité calculés sont comparés avec les résultats des tests d'audition avec les utilisateurs de CI et EAS. Les données du modèle sont en bon accord avec ces résultats.

## **Simon Bennachio - Contrôle actif modal appliqué aux instruments de musique à cordes**

Le projet IMAREV (Instruments de Musique Actifs à REglages Virtuels) est un projet ANR visant à appliquer les techniques actuelles de contrôle actif au domaine musical et plus particulièrement aux instruments de musique. Cette thèse s'inscrivant au sein de ce projet tend à contrôler en temps réel les caractéristiques des modes propres de vibration de tables d'harmonie d'instruments à cordes pour en changer le timbre.

## **Alexis Guilloteau - Etude de l'influence de la géométrie d'un trou de note sur la puissance rayonnée par un instrument à anche simple**

Sur des instruments à vent à trous latéraux on peut distinguer des propriétés de jeu essentielles, appréciés par les musiciens, la justesse, la facilité d'émission et l'étendue de la dynamique admise par l'instrument. Bien que ces propriétés soient corrélées entre elles, cette étude cherche à isoler les causes géométriques susceptibles de modifier de manière significative la puissance acoustique rayonnée par l'instrument. Ceci afin de prédire le dimensionnement des trous latéraux d'une clarinette de manière à respecter le critère d'étendue de la dynamique de l'instrument. En première approximation il est possible de considérer que l'essentiel de l'énergie acoustique est rayonnée par le premier trou ouvert du réseau de trou ce qui permet de prendre en considération des modèles simplifiés du résonateur. On part du constat théorique suivant : les pertes par rayonnement sont indépendantes du rayon de l'orifice ouvert pour une fréquence fixe. Ensuite, à l'aide de modèles analytiques et numériques de fonctionnement de l'instrument, la puissance rayonnée est modélisée pour une embouchure fixe du musicien et en fonction de son paramètre de contrôle principal de la dynamique, la pression dans la bouche. Ceci permet d'isoler les propriétés du résonateur imputables aux variations de puissance émise lorsque le rayon du trou de note est modifié. Une seconde étape expérimentale, en cours de réalisation, permettra de confronter les résultats théoriques obtenus.

## **Brian Hamilton - Méthode des différences finies appliquée à la modélisation de l'acoustique des salles**

Les travaux présentés portent sur une nouvelle famille de schémas aux différences finies appliqués à l'équation d'onde acoustique en deux dimensions sur un maillage hexagonal. Ces schémas aux différences finies utilisent un opérateur Laplacien discret paramétré sur 13 points. On peut montrer qu'une précision du quatrième ordre est obtenue avec un choix approprié de paramètres. Un autre choix de paramètres produit une erreur d'approximation isotrope à l'ordre huit, ce qui est utile dans le cadre des pré- et post-traitements appliqués aux techniques de distorsion de fréquence.

## **Thibaud Leclere - Effet de précedence et restitution multicanal : mise en place d'un modèle 3D de masquage de réponses impulsionnelles**

L'écoute de contenus multicanaux dans des locaux inadaptés pose des problèmes de fidélité de restitution dus aux réflexions acoustiques qui interagissent avec les ondes directes. Divers systèmes existants proposent des solutions d'optimisation qui, tenant compte des conditions d'écoute, corrigent dynamiquement le système de diffusion. Les travaux présentés, à travers la mise en place d'une expérience perceptive, proposent de déterminer une "fonction de seuil spatial" afin de pouvoir qualifier une réflexion de perturbatrice ou non (sur un plan perceptif). Cette fonction permet de sélectionner les réflexions à traiter et allège ainsi le calcul des corrections du système. Un test d'écoute en chambre anéchoïque a été réalisé pour permettre la mesure de seuils de détection (méthode 2 down /1 up) selon plusieurs conditions spatiales d'ondes directes et réfléchies. Les résultats montrent un effet significatif de la position et montrent ainsi le caractère spatial de la perception des réflexions. De plus, les résultats semblent indiquer que cette perception se baserait sur des indices binauraux. Une première fonction de seuil dans le plan horizontal a donc été élaborée et brevetée par Orange.

L'écoute de contenus multicanaux dans des locaux inadaptés pose des problèmes de fidélité de restitution dus aux réflexions acoustiques qui interagissent avec les ondes directes. Divers systèmes existants proposent des solutions d'optimisation qui, tenant compte des conditions d'écoute, corrigent dynamiquement le système de diffusion. Ce stage, à travers la mise en place d'une expérience perceptive, propose de déterminer une "fonction de seuil spatiale" afin de pouvoir qualifier une réflexion de perturbatrice ou non (sur un plan perceptif). Cette fonction permet de sélectionner les réflexions à traiter et allège ainsi le calcul des corrections du système. Un test d'écoute en chambre anéchoïque a été réalisé pour permettre la mesure de seuils de détection (méthode 2 down /1 up) selon plusieurs conditions spatiales d'ondes directes et réfléchies. Les résultats montrent un effet significatif de la position et montrent ainsi le caractère spatial de la perception des réflexions. De plus, les résultats semblent indiquer que cette perception se baserait sur des indices binauraux. Une première fonction de seuil dans le plan horizontal a donc été élaborée et brevetée par Orange.

## **Gabriel Sargent - Estimation de la structure des morceaux de musique par analyse multi-critères et sous contrainte de régularité**

J'introduirai la problématique de l'estimation de la structure des morceaux de musique et présenterai un critère morphologique pour l'estimation automatique des frontières structurelles.

## **Jean-François Sciabica - Interaction entre la perception du bruit moteur et la dynamique du véhicule dans un habitacle automobile**

Les tests perceptifs comme les tests de dissemblance sont traditionnellement utilisés pour étudier la perception du son et pour construire des espaces perceptifs ou espaces des timbres. Puisque la passation de ces tests est longue lorsque nous étudions un nombre important de sons, ce travail propose une méthode pour les réaliser automatiquement en combinant une représentation auditive, le cochléagramme, et le concept de masque temps/fréquence. Le cochléagramme donne en effet une distribution de l'énergie dans l'espace temps/fréquence qui peut être considéré comme une représentation perceptive du son. En introduisant la représentation auditive dans l'espace temps/fréquence, nous souhaitons améliorer la description du timbre par l'intermédiaire des masques temps/fréquence. Ces derniers contiennent en effet une information sur les différences dans le plan temps/fréquence entre deux signaux. Nous appliquons cette méthode au bruit moteur, un signal riche et complexe dont la perception dépend de l'évolution du bruit au cours du temps. Afin d'évaluer la robustesse de notre modèle, nous comparons notre espace des timbres obtenus avec un espace perceptif et un espace calculé par notre méthode.

## **Soizic Terrien - Instruments de la famille des flûtes : apport des méthodes de continuation numérique à l'interprétation d'observations expérimentales**

Les instruments de la famille des flûtes présentent des comportements complexes, pour certains encore mal compris. Si l'approche expérimentale, incluant à la fois des mesures sur musicien et sur bouche artificielle, permet une exploration du comportement de l'instrument, les résultats obtenus restent souvent difficilement interprétables. L'étude d'un modèle physique de l'instrument, vu comme un système à retard de type neutre, apparaît alors complémentaire de l'approche expérimentale. En complément des méthodes de simulations temporelles, les méthodes de continuation numérique permettent d'accéder à une vision plus globale de la dynamique du système (coexistence de plusieurs solutions périodiques, existence de solutions instables ...), et ainsi de mieux comprendre les comportements de l'instrument observés expérimentalement.

## **Charles Hudin - Application du retournement temporel à la restitution tactile**

L'absence de retour tactile dans de nombreux dispositifs d'interaction homme machine mène à une diminution de la précision et de la rapidité d'exécution d'une tâche. L'information tactile est ainsi couramment reportée sur d'autres sens, tels que la vision ou l'audition, souvent déjà largement sollicités. Les travaux présentés portent sur la focalisation d'ondes acoustiques comme moyen de stimulation tactile. La technique de retournement temporel est employée pour générer des stimuli tactiles localisés dans le temps et l'espace à la surface d'une plaque transparente. La répétition et la superposition de ces stimulations permettent de ressentir

simultanément différents rythmes en différents points de la plaque. Une application musicale est envisagée, combinant cette surface tactile à un écran et un système de restitution sonore. Une telle interface permettrait de parcourir rapidement une base de données musicale par exploration tactile multi digitale de motifs rythmiques associés.

### **Marc Evrard - Reproduction de la personnalité vocale d'un acteur**

Ces travaux sont réalisés dans le cadre d'un projet de création, gestion et exploitation de doublures numériques audio-visuelles pour le secteur du jeu-vidéo et du cinéma. Le but est de caractériser et reproduire l'expressivité d'une personnalité vocale. Les domaines principaux du travail concernent la synthèse vocale expressive et l'imitation.

---

## **Session V - Vendredi 7 – 14h30 – 17h00**

### **Louis Bigo - Représentations symboliques de structures musicales à l'aide du calcul spatial**

Les travaux présentés ici proposent différentes manières d'utiliser l'espace pour la manipulation de structures harmoniques. Notre approche est motivée par la recherche d'espaces symboliques faisant émerger la singularité d'une séquence musicale qui y est représentée. Cette motivation se situe au cœur du calcul spatial que nous appliquons à l'aide du langage dédié MGS. Les progressions harmoniques peuvent être représentées dans différents types d'espaces symboliques, réguliers ou pas. Parmi les espaces réguliers, nous présenterons les Tonnetze, ainsi qu'un exemple d'application en composition musicale à l'aide du papier interactif.

### **Ning Chu - Robuste bayésienne super-résolution approche par a priori parcimonie pour l'imagerie aéroacoustique**

L'imagerie aéroacoustique a, depuis des décennies, beaucoup attiré l'attention des chercheurs et ingénieurs en terme de localisation de sources aéroacoustiques et d'estimation de puissance. Récemment, les méthodes de déconvolution ont grandement amélioré la résolution spatiale que la méthode conventionnelle de formation de voie. Mais la plupart des méthodes classiques ne sont pas robustes au bruit de fond, elles ne fournissent une large gamme dynamique de l'estimation de la puissance, en particulier dans le faible rapport signal sur bruit (SNR).

Dans cet article, nous avons d'abord amélioré le modèle de propagation aéroacoustique en tenant compte du bruit de fond et l'incertitude système. Ensuite, nous proposons une bayésienne robuste super-résolution approche par a priori

parcimonie pour estimer la puissance de source et les positions dans pauvres SNR, et estimer conjointement la puissance de bruit et autres paramètres du modèle. Notre approche est comparée à quelques-unes des méthodes de l'état de l'art sur les données réelles simulées, et hybrides. Les principaux avantages de l'approche proposée sont la robustesse au bruit, une large gamme dynamique, super résolution spatiale, et la non-nécessité de la connaissance préalable du numéro de la source ou SNR. Il est possible de l'appliquer en l'imagerie aéroacoustique monopole et étendue basée sur le réseau de microphones non-uniforme pour des essais en soufflerie.

### **Simon Conan - Recherche des invariants acoustiques liés à l'évocation d'événements sonores particuliers. Une application à la synthèse et au contrôle des sons de roulements**

Un des enjeux de la synthèse sonore aujourd'hui est le contrôle perceptif (i.e. à partir d'évocations) des processus de synthèse. En effet, les modèles de synthèses sonores comprennent généralement un grand nombre de paramètres à bas niveau qui nécessitent une certaine expertise des processus. Avoir des contrôles perceptifs sur un synthétiseur, par exemple des sons du quotidien, offre beaucoup d'avantages : le contrôle du synthétiseur s'effectue directement par la façon dont nous percevons les sons et donc plus réaliste, et permet également à des utilisateurs non-experts de créer et contrôler des sons de manière intuitive. Cette approche est basée sur un certain nombre d'hypothèses concernant notre perception, notamment la présence de morphologies acoustiques ou « invariants » responsables de l'identification d'un événement sonore. Nos travaux se penchent actuellement sur la synthèse de sons d'interactions continues, et en particulier sur les sons évoquant le roulement, basé sur un paradigme {action/objet}. Un des buts de l'étude est de mettre en évidence des « invariants » liés à la perception de l'action « rouler ». L'étude d'un modèle basé sur la physique nous a permis de comprendre qu'un possible invariant semblait être lié à la force d'interaction non-linéaire entre la bille et la surface. Cette force est une suite de micro impacts ayant des propriétés statistiques particulières. Un modèle de signal comprenant ces propriétés a donc été développé, et est utilisable comme excitation associée au roulement dans un modèle source-filtre générique. On présentera ce modèle. D'autres invariants supposés être liés au roulement ainsi que les travaux à mener afin de les valider perceptivement seront également présentés. Enfin, les perspectives offertes en terme de contrôle de la synthèse seront discutées.

### **Gilles Courtois - Predict phantom source localization within stereophonic reproduction**

Two-channel stereophony is known as the most used playback standard for music reproduction. It enables the recreation of a virtual sound stage lying between the two loudspeakers. Auditory events may appear where no physical sound source is present; they are the so-called phantom sources. However, stereophony is not a robust technique as it requires a very precise and restrictive set-up. In particular, the

arrangement of the loudspeakers and the listener should be perfectly symmetric. Deviations from this symmetry result in a distorted auditory image. Signal processing algorithms aiming at counteracting such distortions would have to manage to predict the perceived sound stage from different listening positions. This master thesis introduces a model that achieves such a task. The main localization cues used by the human auditory system are simulated and incidence directions are derived from them. The comparison of computed results with data gained in subjective listening tests indicates that the developed model is able to predict the perceived incidence direction of single auditory events from different listening positions.

### **Ange Ebissou - Intelligibilité en espace ouverts : étude de la gêne en présence de plusieurs sources de parole**

Dans un bureau ouvert, de type open-space, l'absence de cloisons séparatrices en toute hauteur et la proximité des voisins contribuent à l'émergence d'un environnement sonore riche en bruit de parole. Si les niveaux atteints ne sont pas dommageables pour l'audition, le brouhaha permanent rend le travail plus difficile à accomplir. Il s'avère de plus que la gêne est d'autant plus fortement ressentie que les discours environnants sont intelligibles, c'est-à-dire qu'ils sont entendus suffisamment clairement pour être compréhensibles. Ce travail de thèse se donne pour objectif de quantifier la nuisance ainsi induite par la présence de signaux acoustiques de parole dans un contexte de travail. Il s'agit plus particulièrement d'étudier l'évolution de la gêne subie par un individu face à une tâche à réaliser en fonction de l'intelligibilité des voix présentes en sa proximité. L'adéquation des indices d'intelligibilité déjà existants est évaluée. Un indice propre, mieux adapté aux situations où plusieurs sources de parole sont présentes, est éventuellement à développer.

### **Benjamin Elie - Caractérisation vibratoire et acoustique des instruments à cordes - Application à l'aide à la facture instrumentale**

La qualité des instruments à cordes frottées, et plus particulièrement celle des violons, est abordée dans la littérature sous diverses formes. Les études se basent souvent sur des *a priori* et partent de postulats affirmant que les bons violons présentent certaines caractéristiques, telles que la présence d'une amplification de la mobilité et du spectre dans une zone allant de 2 à 3kHz. Une autre approche employée dans la littérature consiste à relier le vocabulaire généralement employé pour qualifier un violon à ses caractéristiques spectrales, ou vibratoires. Le son produit par un violon résulte en fin de compte de deux caractéristiques fondamentales et différentes: le geste instrumental et le comportement dynamique de l'instrument. L'approche développée dans les travaux présentés ne se base sur aucun jugement subjectif quant à la qualité des instruments. Elle vise à définir des corrélations entre des attributs acoustiques des sons de violons et des caractéristiques de la caisse. Elle vise donc à rechercher les traces de la signature

acoustique des violons dans les caractéristiques spectrales de leurs sons. L'étude comprend deux parties : la première concerne l'analyse du comportement vibratoire du violon, à travers l'étude de la mobilité au chevalet. La seconde vise à relier les caractéristiques identifiées pour la mobilité avec des paramètres spectraux caractéristiques des sons produits par les instruments. Ces travaux montrent que la mobilité latérale au chevalet des violons est présente de manière sous-jacente dans le spectre des sons de violons.

### **Augustin Lefevre - Méthodes d'apprentissage de dictionnaire pour la séparation de sources mono-canal**

La séparation de sources mono-canal pour les signaux sonores est un problème mal posé qui nécessite de faire des hypothèses sur les signaux sources que l'on entend récupérer. Si l'on dispose d'ensembles d'échantillons représentatifs (dits d'entraînement) de la diversité sonore de chaque source, et qu'à l'échelle de la vingtaine de millisecondes, un échantillon test constitué d'un mélange de chaque source peut être représenté comme une combinaison linéaire de quelques échantillons d'entraînement de chaque source, alors la séparation de sources se ramène à un problème de régression linéaire surdéterminé. Afin de travailler sur des spectrogrammes, il est commode d'utiliser une variation de l'apprentissage de dictionnaire avec des contraintes de positivité, la factorisation en matrices positives (NMF). Dans le cas où l'on ne dispose pas d'échantillons d'entraînement, nous montrons qu'il est possible d'utiliser des annotations dans le plan temps-fréquence directement sur l'échantillon test. Ces annotations indiquent, pour chaque point du plan temps-fréquence, la proportion de l'énergie partagée par chaque source. Nous proposons un programme inspiré de l'inpainting pour calculer une estimation de chaque source qui satisfait les annotations et telle que la somme des sources reconstitue correctement le spectrogramme. Cette extension permet de traiter la NMF comme une "boîte noire" pouvant dialoguer avec un utilisateur capable de fournir des informations intéressantes sur l'échantillon test. Nous expérimentons également plusieurs logiciels d'annotations automatiques "faits maison", le but étant d'ouvrir une porte de communication avec les méthodes de Computational Auditory Scene Analysis.

### **Olivier Perrotin - Chironomie et interfaces pour les instruments musicaux virtuels**

Les instruments musicaux virtuels comportent deux facettes : d'une part le moteur de synthèse, qui calcule le signal numérique à partir des paramètres de contrôle, et d'autre part l'interface de contrôle. Cette thèse dans le domaine de l'informatique vise à étudier ce second aspect : l'utilisation de la chironomie, mouvements expressifs de la main, et le développement d'interfaces pour le contrôle gestuel d'instruments de musique virtuels (voix, cordes, vents). L'objectif est de concevoir des interfaces peu coûteuses, faciles à utiliser, mais présentant aussi une marge de progression importante en offrant de nombreuses possibilités de jeu héritées des

instruments réels.

Ce travail est dans la continuité des travaux du LIMSI sur le contrôle gestuel de la synthèse vocale [1],[2]. Le but est d'étendre les systèmes de synthèse et les interfaces de contrôle actuels aux instruments de musique. Il se divise en trois parties : le développement du contrôle gestuel de nouveaux types de synthétiseurs (cordes, vents, percussions), le développement des interfaces de contrôle et le test d'utilisation des interfaces par des musiciens et non musiciens.

[1] Chironomic stylization of intonation Christophe d'Alessandro, Albert Rilliard, and Sylvain Le Beux

[2] Chorus Digitalis: experiments in chironomic choir singing Sylvain Le Beux, Lionel Feugère, Christophe d'Alessandro

### **Quentin Mesnildrey - Simulations acoustiques de l'implant cochléaire : Modélisation de la configuration géométrique des électrodes et effets sur la perception de la parole**

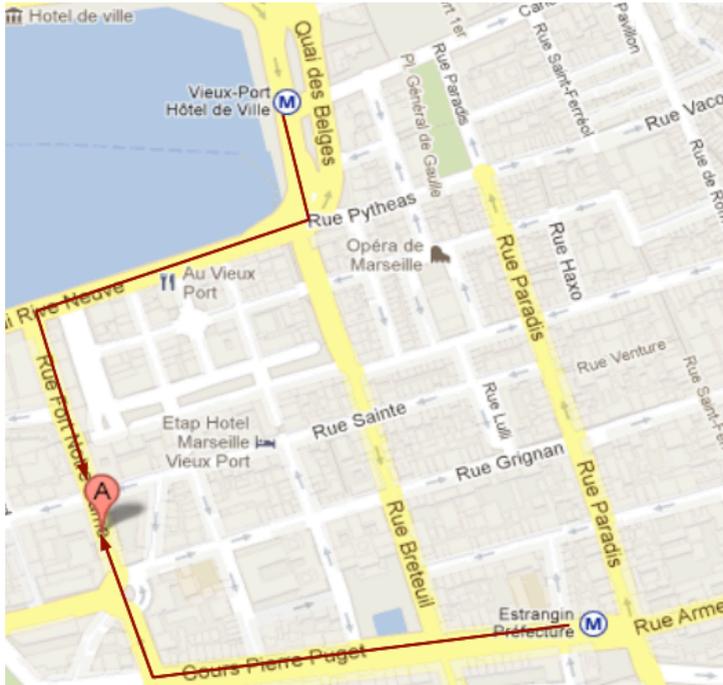
Travail sur les effets perceptifs de la configuration des électrodes de l'implant cochléaire et particulièrement des modes de stimulation bipolaires. Implémentation d'un vocodeur pour simuler ces effets et test sur des sujets normo-entendants.

### **Camille Vauthrin - Technique de jeu et production du son dans les instruments de musique à sons entretenus : application aux flûtes**

L'objectif de cette étude est de réaliser une quena chromatique en partenariat avec un luthier. La quena est une flûte des Andes. Le modèle sur lequel est basé notre sujet est une flûte chilienne, cet instrument joue dans une gamme 'occidentale', diatonique. On propose alors une aide au luthier, utilisant un modèle acoustique de l'impédance d'entrée de l'instrument afin d'en déduire des paramètres tels que les fréquences de jeu et facteurs de qualité.

## Le logement

Le logement est situé à l'hôtel Vertigo Vieux Port, au 38 rue Fort Notre-Dame. Le plus simple pour s'y rendre est d'y aller à pieds depuis la station de métro Vieux-Port ou Estrangin (Métro 1).

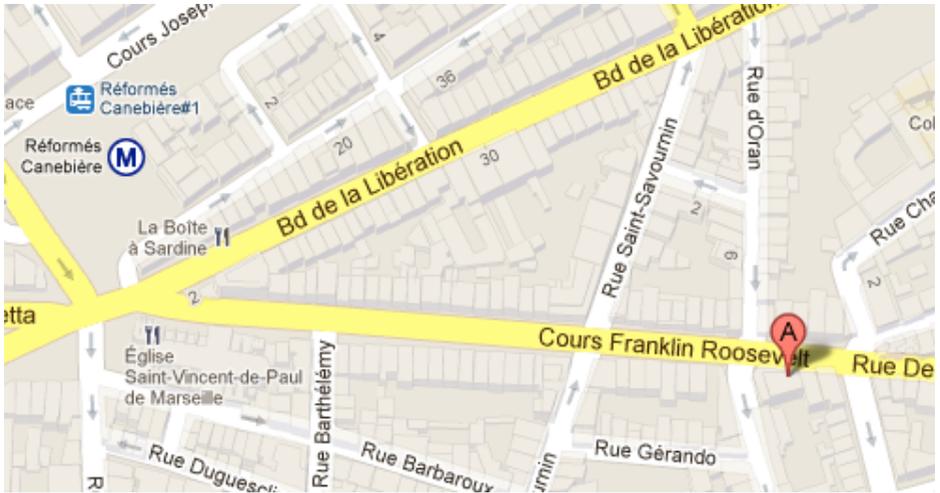


Le petit-déjeuner vous est offert et est servi à 8h à 10h à l'hôtel.

Le jour de votre départ, la chambre devra être libérée à 10h. Veuillez à la rendre propre, vaisselle faite. Merci !

## La soirée du jeudi soir

La soirée du jeudi soir aura lieu au Centre Social Tivoli, au 66 Cours Franklin Roosevelt. Pour s'y rendre, descendre à la station de métro Réformés (Métro 1) puis finir à pied.



En cas de problème, vous pouvez nous contacter :

Simon Conan : 06.95.21.52.51

Soizic Terrien : 06.81.17.98.34

Etienne Thoret : 06.48.06.84.80