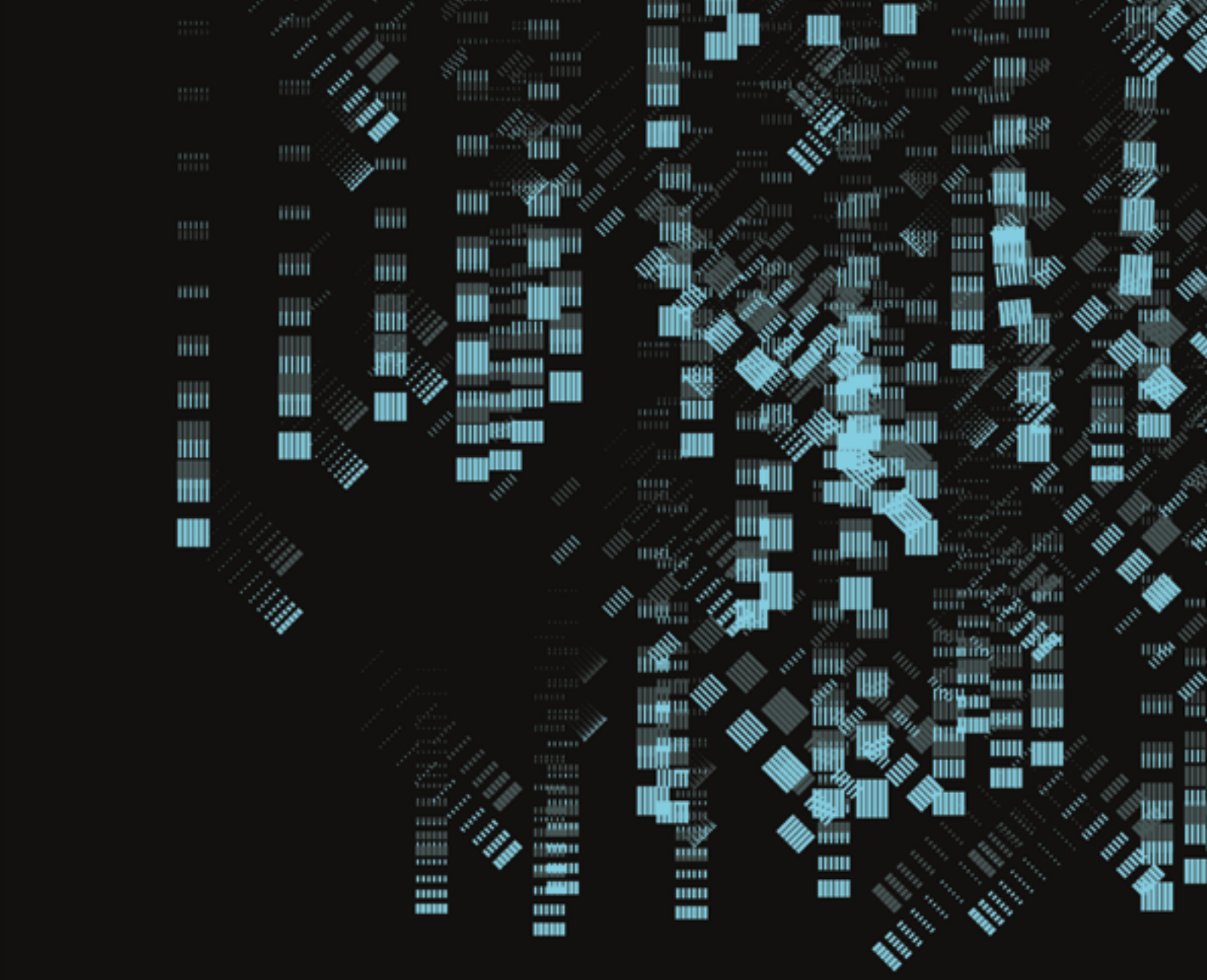


# SCIENCE AND TECHNOLOGY OF MUSIC AND SOUND LABORATORY

## UMR 9912 – STMS

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
Ministère Culture

ircam Centre Pompidou

cnrs

SORBONNE UNIVERSITÉ

# INNOVATION: RECENT SUCCESS STORIES

### THIBAUT CARPENTIER, LAUREATE OF THE CNRS 2018 CRISTAL MEDAL

Tools for spatialization developed and disseminated by this research engineer in the Acoustic and Cognitive Spaces team since 2009 are a vitrine for the team's technology and the entire laboratory. Their exceptional impact on artistic, scientific, and industrial communities has made a vital contribution to the excellence of our laboratory.

### SPAT REVOLUTION, NEW MASSIVE MULTICHANNEL VERSION, LAUREATE OF THE NAMM 2017 TECH AWARD

Developed by the Acoustic and Cognitive Spaces team in partnership with Flux::, Spat, a software program that spatializes sound in real-time, is designed as a library and lets musicians or sound engineers control the sound spatialization for different sound rendering systems.

### PATRICK SUSINI, THE HEAD OF SOUND PERCEPTION AND DESIGN TEAM RECEIVED THE 2016 CHAVASSE AWARD

bestowed by the Société Française d'Acoustique.

### SUCCESS IN THE EUROPEAN H2020 PROGRAM

Over the past two years of the H2020 program (2014-15), IRCAM, despite its small size, is among the top 15 French beneficiaries in the domain of information and communication technologies, and the top beneficiary in creative industry fields.

### DECIBEL D'OR PRIZE

The Conseil national du bruit (French National Council for Noise) awarded the Decibel d'Or 2014 prize to IRCAM's Sound Design and Perception team for its unique research in sound design that combines scientific work in auditory perception and cognition on environmental sounds with applications in the domains of sound quality and conception.

### CREAM: THE FIRST ERC PROJECT

Obtained at the end of 2013 by Jean-Julien Aucouturier, CNRS researcher in the STMS laboratory, the CREAM project (Cracking the Emotional Code of Music) was selected by the prestigious European Research Council and aims to characterize the emotional mechanisms induced by listening to music.

### THE OPENMUSIC SOFTWARE

Free and open source, OpenMusic is the reference for computer-assisted composition and mathematics applied to music.

### THÉORIES DE LA COMPOSITION MUSICALE AU XX$^{E}$ SIÈCLE

This monumental work made up of 1840 pages, 67 chapters, and 2 volumes published by Éditions Symétrie, edited by Nicolas Donin and Laurent Feneyrou, brings together contributions by today's top specialists, providing an overview of compositional theories of the last century. Published at the end of 2013, this work is the fruit of an unprecedented editorial project carried out over more than a decade.

### ANTESCOFO: WINNER OF THE 2013 AWARD FOR INDUSTRIAL INNOVATION AND THE 2011 SPECIAL AWARD FROM THE FRENCH MAGAZINE *LA RECHERCHE*

This real-time software simulates a musician's intelligence, capable of listening to other musicians and accompanying them according to their performance. It is the result of Arshia Cont's PhD, winner of the Gilles Kahn award.

### POLYPHONIC TRANSCRIPTION AUDIO2NOTE INTEGRATED IN LIVE 9

Version 9 of the software Live, launched in March 2013, includes the pioneering technology Audio2Note by IRCAM, the result of a collaborative project with Ableton in Berlin, making automatic transcription in MIDI of polyphonic audio recordings possible.

### NEW MASSIVE MULTICHANNEL AUDIO SYSTEM

Inaugurated in November 2012, this unequalled system is comprised of 345 loudspeakers installed in the Espace de projection – a room with modular acoustics – and offers high-resolution spatialized sound reproduction through the combination of the Wavefield Synthesis and High Order Ambisonics systems.

### SUPERVP ACCLAIMED BY THE AUDIO INDUSTRY

Software for high-quality sound processing, SuperVP is the international leader in its category and has been licensed to numerous software publishers and audio services. SuperVP is the engine for the AudioSculpt program.

### SOUND DESIGN OF CONCEPT AND PRODUCTION CARS BY RENAULT

UMR has worked closely with car manufacturers, including Renault, for the conception of the sounds heard in the passenger compartment as well as the sounds made by electric cars (e.g. the DeZir and Symbioz Renault concept cars).

# KEY FIGURES

**7**
**research teams**
work in various scientific fields related to music: acoustics, signal processing, modeling, control, computer science, design, psychology, musicology.

**110**
**contributors**
(researchers, engineers, technicians, coordinators), & 40 internships and guest researchers per year.

**30%**
**of staff**
are of foreign nationality.

**200**
**publications per years**
50 of which appear in scientific journals or as chapters in books.

**90**
**annual partnerships**
with 60 labs and public organizations and 30 with private companies.

**22**
**software environments**
developed and distributed to
**10 000**
**professional users**
via IRCAM Forum (software user group).

# A WORD FROM THE DIRECTOR

IRCAM hosts a range of research in the joint research unit on the sciences and technologies of music and sound (UMR 9912-STMS) bringing together IRCAM, the CNRS, Sorbonne Université, and the French Ministry of Culture. The period that starts with the upcoming five-year program (2019-2023) confirms the organization of the research structures present at IRCAM is consistent with that found throughout academia. This evolution responds to a number of challenges: the integration of artistic research in university structures and a proliferation of art-science themes (with, for example, the creation of doctoral programs in the arts), the restructuration of the Parisian research scene, the legibility and increasing appeal of the laboratory, the evolution of the French innovation ecosystem, and the renewal – 40 years after its fruition – of Pierre Boulez' groundbreaking vision.

This new organization, in which the joint research unit advances the institute's scientific project, relying on IRCAM's new department of Innovation and Research Means (headed by Hugues Vinet), was made possible through the completion of a number of large-scale projects carried out from 2012 to 217 by Gérard Assayag and Hugues Vinet (director and deputy director until 2017) and by Carlos Agon Amado (interim director until January 1, 2018). This coordination was concretized by a partnership agreement with the joint research laboratory's founding institutions; a contract I intend to implement during the next five years.

**Brigitte d'Andréa-Novel**
Director of the UMR 9912, June 2018

## LABORATORY HIGHLIGHTS

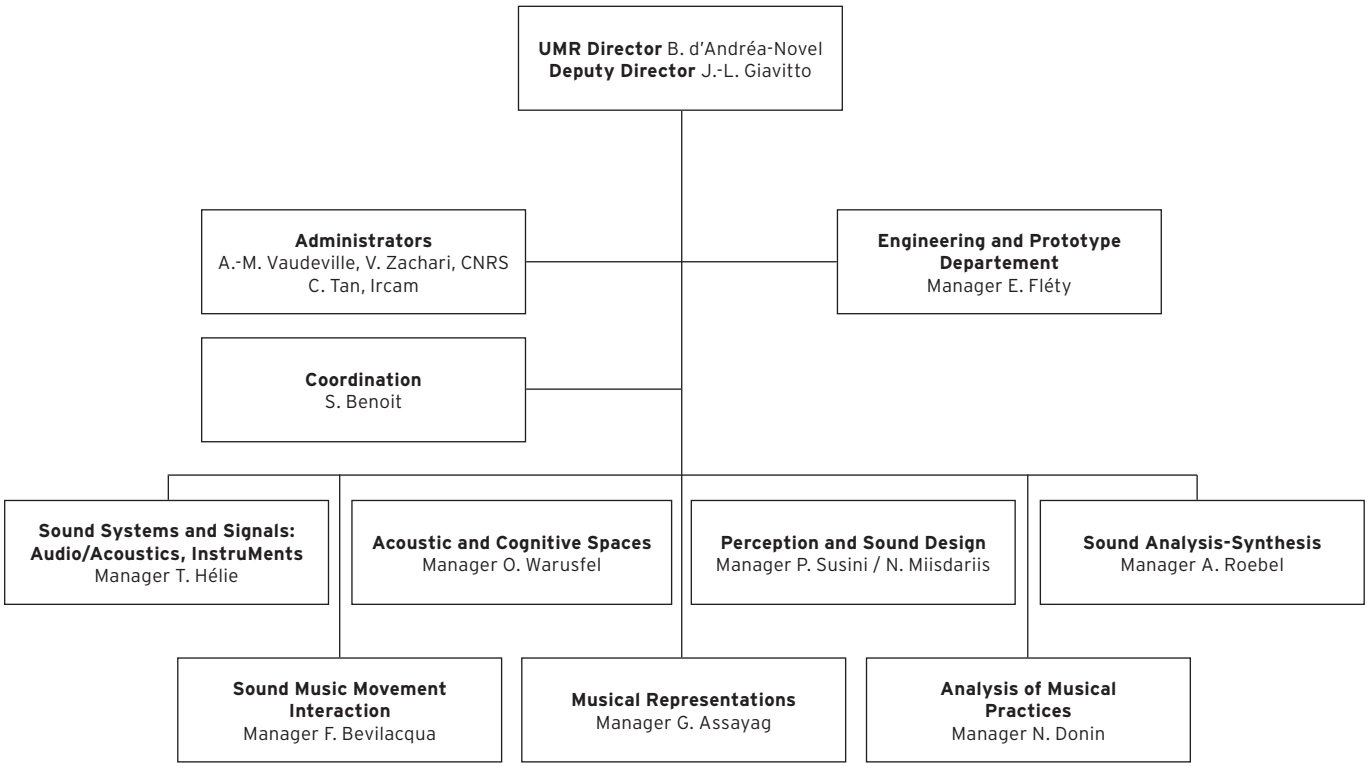
- **Creation of the new team S3AM** (Sound Systems and Signals: Audio/Acoustics, InstruMents) **in January 2017**
- Launch in March 2017 of a new annual event entitled **Mutations/Creations** organized by the Centre Pompidou and IRCAM—held at the Centre Pompidou— associating expositions, symposiums, and artistic events with a focus and debates on the evolutions of artistic practices in the contemporary scientific and technological ecosystem.
- **The European project VERTIGO** (H2020 ICT36-2016), a part of the STARTS initiative (Science, Technology, and the ARTS) founded by the DG Connect at the European Commission, aims at coordinating European collaborations between artists and R&D technology projects for innovation, via a large-scale program of residencies, the development of a collaborative Web platform, and an annual public presentation of the community's work.
- **IRCAM's success in the European H2020** call reinforcing its position among the 15 French actors involved in the H2020 ICT program with the selection of Vertigo and the i-MuSciCA projects on educational applications of sound synthesis via physical modeling and FuturePulse on marketing analytics in the music industry.
- **Numerous prizes awarded to the researchers in the joint research unit:** the SFA's (Société française d'acoustique) 2013 Prix Chavasse to Olivier Warusfel for the work of the ACS team (Acoustic and Cognitive Spaces); the SFA's 2016 Prix Chavasse to Patrick Susini, head of the PSD team (Perception and Sound Design); the SFA's 2016 Prix Rocard to Emmanuel Ponsot (PDS team); the CNRS's 2018 Cristal award to Thibaut Carpentier, engineer in the ACS team.
- **Organization of several public events from 2015 to 2018:** renewed success of open house afternoons in the "Studio 5, Live" Saturday program, symposia "Art Music and Pop Music, Articulations", WAC'15, "New Technologies for Music Notation and Representation", TCPM'15, "Improtech Paris-Philly 2017" bringing together the entire human-machine co-improvisation community in Philadelphia with the Omax/ImproteK/DJazz projects, an important international conference "Jean-Claude Risset – Interdisciplinarity" through the Collegium Musicae and Sorbonne Université.

# CONTENTS

# 1

# THE LABORATORY

# ORGANIZATIONAL CHART



# MISSION STATEMENT

The fundamental principle of IRCAM is to encourage productive interaction among scientific research, technological developments, and contemporary music production. Since its establishment in 1977, this initiative has provided the foundation for the institute's activities. One of the major issues is the importance of contributing to the renewal of musical expression through science and technology. Conversely, specific problems related to contemporary composition have led to innovative, theoretical, methodological, and applied advances in the sciences with ramifications far beyond the world of music. Focused on artistic production, in all its particularity and sensitivity, this highly original research dynamic addresses modes of representation and of extended manipulation of sound and music, thus eliciting scientific and technological applications that touch an increasingly wide-ranging public including sound and music professionals, industrial concerns, academics, music devotees, etc. Central to the interaction between research and musical production is the development of software tools for musicians, composers, performers, and musicologists using the models and prototypes created by the research teams in various music-related domains including information technology (e.g. languages, human-computer interaction, real-time, and databases), signal processing, acoustics, auditory perception and cognitive psychology, and musicology.

**The work carried out in the STMS joint research lab (Science and Technology of Music and Sound) is supported by the CNRS, the French Ministry of Culture, and Sorbonne Université.**

**Research**, the source of innovation, seeks to expand our knowledge of music related issues. Because of its essentially cross-disciplinary character, a great deal of research is carried out in collaboration with laboratories in France and abroad, academic institutions, institutional partners, and private partners. The integration of graduate and post-graduate students enables research training. The skills developed at IRCAM can be applied to domains beyond the musical arena and are frequently included in projects carried out with industrial partners or in the context of French, European, or multinational research projects.

**The development** pole seeks to adapt the knowledge, models, and prototypes produced by the research teams into software environments. The principle applications are used to create tools that assist musical creation by putting in place open and programmable environments, adaptable to a variety of aesthetic approaches and that allow for the continued integration of models and their updates. This configurable design makes it easier to develop simplified versions of the software programs intended for the general public: Ircam Tools (for audio professionals), IrcaMax (for live electronic music), IrcamLab (for a broad range of users). Central to software development is the design and evaluation of specific human-computer interaction interfaces and the permanent integration of the rapidly evolving technology from the computer industry. The IRCAM Forum—an IRCAM software user group—encourages the distribution of the institute's software within an international community of music and sound professionals (e.g. composers, multimedia and visual artists, sound engineers and designers, researchers, teachers, etc.) with an estimated 10,000 users since its beginnings. The IRCAM Forum also provides companies with licenses for either internal use or distribution purposes. These licenses can apply to either complete applications (e.g. the Max program, used by tens of thousands worldwide, is licensed to the company Cycling '74 in California) or to specific functional models (e.g. audio analysis and processing, indexing, spatialization, synthesis) that are included in commercial environments.

# ORGANIZATION AND FIELDS OF STUDY

The laboratory is structured in **seven teams** defined by their fields of study and theoretical approaches:

1. Sound Systems and Signals: Audio/Acoustics, InstruMents uses applied mathematics (signal, system, control, differential geometry, digital analysis), physics and mechatronics to explore, understand, reproduce, or invent multi-physical systems that produce sound signals (musical instruments, voices, etc.).
2. Acoustic and Cognitive Spaces focuses on acoustics, signal processing, cognitive listening psychology, and neurosciences for the capture, analysis, and reproduction of sound fields as well as the understanding of the notion of spatial perception in multisensory contexts.
3. Sound Perception and Design addresses the perception of vocal, musical, and environmental sounds, their emotional treatment, and sound design applications via psychoacoustics, physiology, musical cognition, and neurosciences.
4. Sound Analysis & Synthesis studies digital sound signals using sound signal processing theories and techniques, statistics, learning, and shape recognition to extract various characteristics, to synthesize sounds according to specified attributes, or the transformation of sounds based on compositional or other needs.
5. Sound Music Movement Interaction focuses its research on embodied interaction between humans and sonic and musical medias: motion capture, reactive real-time programming, learning, tangible and multimodal interfaces, sound synthesis techniques.
6. Musical Representations investigates the formal structures of music and creative environments for composition, analysis, and human-machine interaction using architectural and language approaches, algebraic formalizations, and artificial intelligence techniques such as learning optimization.
7. Analysis of Musical Practices develops interdisciplinary musicology taking into account historical, sociological, anthropological, psychological, and philosophic points of view favoring the use of creative tools, often in connection with other teams at IRCAM.

The laboratory's scientific project relies on the 7 teams, organized around **3 fields of interaction:**

- **Sound Workshop**
- **The Musical Body**
- **Creative Dynamics**

**• SOUND WORKSHOP**

Here, sound is understood as a physical phenomenon (mechanical, acoustic) as well as digital data involving mathematical and computer science (signal processing) techniques to process data audio streams and musical or multimedia scenes.

Digital audio signals are addressed in as sources and fields. We characterize or synthesize digital audio sources with the accuracy, subtlety, and quality required to meet the demands of contemporary creation. Our research focuses on the specificities of the musical signal and has led us to develop more precise models employing physics, biomechanics, perception and cognition, and even symbolic musical data. The extraction of more a symbolic, semantically pertinent description and the control of models via high-level qualitative data (e.g. prosody in the case of the synthesized singing voice) are two of the laboratory's objectives in this domain. As for sound fields, our studies aim at understanding their spatial representations and couple them with time/frequency representations, at capturing them in high resolution, and at reconstructing them realistically.

**• THE MUSICAL BODY**

The domain focused on here is that of music and sound, perceived and produced by humans who interact with an environment, a machine, or other human beings. Sound is not reduced to a simple vibration, but becomes a vehicle enabling interaction produced by a biological body and operating on individuals and behaviors. This domain is addressed through four issues—gesture, voice, the spatial perception of sound and its multi-sensorial integration—while examining the connections among sonic or musical contents and emotional valency.

## • CREATIVE DYNAMICS

While music is a phenomenon produced and perceived by a body that capitalizes on physical objects and organizes a sound world, it is also thought, imagined—or dreamt— and occasionally formalized before being realized. The question of creativity emerges today as a scientific[1] issue and a useful working paradigm for the conception of intelligent systems. Contrary to what a superficial analysis of the new directions taken by AI[2] could suggest, the challenge seems to be more about mentoring and companionship than imitation: what ideas and tools are necessary to inspire, guide, and assist an artist in their creative process and make the machine a partner? We are therefore opposed to an approach that aims at replacing humans, making autonomous machines and that which aims at augmenting human capacities. We can therefore talk about *augmented intelligence*.

In the artistic field, the first approach could possibly lead to understanding mechanisms that are replaced, but the second offers novel creative dimensions that could radically renew the science-music relationship. Using one of our major assets—our intimacy with musical creation—the laboratory's work endeavors to apprehend and document composers' creative process; to formalize, analyze, calculate, and produce symbolic musical structures in interaction with the creators' goals; to develop new tools for writing that address contemporary issues of notation; and to capture systems that assist collective creation.

These three fields ensure the legibility of our research and cover all our activities, they cross the physical and digital dimensions of the sound and musical world as well as its perception and production by humans in connection with issues surrounding creation and creativity. Among these research subjects, we can cite a few examples of subjects we intend to investigate more thoroughly in the upcoming years:

- Multiphysical modelling
- Expert expressive gesture
- Cyber physical & human systems
- The contribution of neurosciences, cognitive sciences, and connections with health
- AI and musical creativity
- New dimensions of musical notation

Despite the diversity of the projects carried out, the laboratory has a strong identity and visibility in terms of the subjects of its studies and its scientific position. **Three important principles characterize the laboratory's scientific policy contribute to this state of affairs**, compensate for the small number of researchers, and enable the laboratory to produce prominent scientific contributions and develop innovative tools.

## 1. RESEARCH-CREATION AFFILIATION

The first principle is the a fruitful affiliation of research themes with contemporary musical creation through interaction[3]. This affiliation ensures a large diffusion of research findings. Research-creation interaction is supported by several recurrent mechanisms. We will cite four.

- The joint laboratory teams participate in IRCAM's artistic productions, be it directly or via the activity of computer music designers who serve as a means of transmission among artists, prototypes, and tools developed in the laboratory. Their role is essential in adapting the long time necessary for research to the inevitably sorter time available for artistic productions.
- The Research/Creation Interfaces projects-residencies provide an opportunity for the teams to interact with an artist who has proposed a research project and is in residence at IRCAM for a 3-month period. These projects, removed from production contexts and constraints, make it possible to anticipate emerging needs and forecast new developments. Since 2012, 31 residents have been selected via an annual international call on the ULYSSES[4] platform. The majority of the artists in residence are not French, and half the projects involve at least two research teams.
- Created in 2013, the system of "Doctorate in Art" enables the development of interaction over a long period with an artist working full-time on their thesis. This type of doctoral program makes it possible to combine artistic practices and academic approaches with research and creation.
- Finally, research/creation interactions are grounded in encounters with individuals with unique issues. In order to develop a more open dialogue with the community of composers, and to systematically identify emerging themes IRCAM (the Research/Creation Interfaces and Production departments as well as the joint research lab) tested in 2017 during days set aside for encounters and discussions. These days, entitled "Méridien" will be renewed.

1. See the new international conferences "Digital Intelligence (#DI)", "Musical MetaCreation (MUME)", "International Conference on Computational Creativity (ICCC)", in which the joint research laboratory has participated regularly since their creation. See also the European initiative FET "Creative ICT" that supported 5 projects in 2016.
2. Like the Magenta project developed by Google: " [how to develop] algorithms that can learn how to generate art and music, potentially creating compelling and artistic content on their own" https://goo.gl/MF5foi
3. This model of interaction, in which IRCAM was a pioneer, is now recognized and used with, for example, the H2020-ICT project Vertigo (Adding socio-economic value to industry through the integration of artists in research and open innovation processes) begun at the end of 2016, coordinated by IRCAM: http://vertigo.ircam.fr/
4. http://www.ulysses-network.eu

## 2. A BACK-AND-FORTH BETWEEN FINALIZED APPLICATIONS AND ACADEMIC REFLECTIONS

The second principle that characterizes the scientific strategy of the joint research laboratory is that of consistently validating theoretical reflections through developments used beyond the teams and, vice versa, substantiate applicative developments via an epistemological or theoretical reflection. In this dialogue, the software or hardware application is used in the laboratory as a necessary tool for experimentation. The prototypes developed by the researchers are often developed within the framework of collaborative projects, meaning they are used by our partners as well as in artistic productions, associating laboratory results with the demands of a composer and their experienced ears. Beyond experimentation and via IRCAM's artistic projects, these tools reach a broader public through the IRCAM Forum[1] that provides transversal coordination at IRCAM (over 10,000 professional members worldwide).

## 3. AN IMPORTANT PRESENCE IN FRANCE AND ABROAD

The third principle is that of a laboratory with a strong presence in high-level French and international research scenes, in all the scientific communities with which its teams are connected. This point is covered in the following paragraph.

1. https://forumnet.ircam.fr

# INSTITUTIONAL LINKS

The joint research laboratory relies on the natural, diverse connections among its research teams, but also relies on the presence of its members at the heart of regional, national, and international communities. Proof of this can be seen in the large numbers of invitations extended to the laboratory's researchers (seminars, guest speakers) and in their participation in steering committees or in seminars and conferences they have either initiated or organized.

On a national level, the joint research laboratory at IRCAM is the only laboratory in the 2016-2020 Culture-CNRS framework agreement in the field of music[1].

On a regional level, the laboratory is a partner in the **ATIAM** Master's program (Acoustics, Signal Processing, and Computer Science Applied to Music) hosted and organized by IRCAM as a part of the Master's degree in Sciences and Technology at Sorbonne Université in collaboration with Télécom ParisTech. The laboratory is also a partner in the **Sound Design** Master's degree organized with the EPCC – École supérieure des beaux-arts Tours Angers Le Mans, l'université du Maine and l'ENSCI – Les Ateliers. These two programs are unparalleled in France. Abroad, these programs are usually part of a "Music Technology" class in a music school, intended for music students. Members of the laboratory also teach occasionally in other Master's degree programs (Grenoble, Strasbourg, Marseille).

The joint research laboratory hosts the doctoral schools in its domain from Sorbonne Universités; particularly EDITE (Paris doctoral school for computer science, telecommunications, and electronics), SMAER (mechanical sciences, acoustics, electronics, and robotics) and ED3C (brain, cognition, behavior).
A new doctoral school was launched with the new academic year in 2014 via the ComUE Sorbonne Universités and the Idex SUPER: Doctorate in Music – Research and Composition. IRCAM is a founding member[2] of the **Collegium Musicæ**, an institution that structures the activities of the ComUE in the field of music. Three themes are covered: analysis and creation, instruments and performers, and archives and heritage.

The joint research laboratory participates in two LabEx (**laboratories of excellence**):

- The LabEx SMART[3] that focuses on human-machine interaction, bringing together 5 laboratories from Sorbonne Université as well as Télécom ParisTech and the LUTIN (laboratory on uses associated with the Cité des sciences et de l'industrie).
- The LabEx CAP[4] via the Analysis of Musical Practices team. CAP studies the arts, creation, and heritage to understand and accompany contemporary societal mutations in the context of cultural and economic globalization.

The laboratory is also associated with the Fondation Sorbonne Université **Excellence Chair**, MouVIE, on mobility and standard of living in urban areas, supported until 2020 by PSA/Renault.

On an international level, the joint research laboratory maintains connections with today's most prestigious international research centers on computer music such as CCRMA at Stanford University, CNMAT at UC Berkeley, the MIT Medialab in Boston, CIRMMT at McGill University, or in with the MTG at the UPF Barcelona, the SPA department at Aalto University in Helsinki, and C4DM at Queen Mary University (guest researchers, doctoral exchanges, organization of common projects[5]). The laboratory has also connected with other large universities through workshops organized at the universities of Singapore, Columbia, NYU New York and Abu Dhabi, Campinas, Sao Paulo, Shanghai, and at EPFL.

1. See https://goo.gl/jmr9BE Two other joint research laboratories host programs supported by the French Ministry of Culture. The ministry is not a co-tutor of either of these laboratories. The LaBRI is host to the "Studio de création et de recherche en acoustique musicale (SCRIME)" and the LAM is home to the program for documents and sound archives as well as the work of the GIS-SPADON.
2. Along with 8 other major actors in education and musical research: IReMus, Institut de recherche en musicologie (UMR 8223 - CNRS - Paris-Sorbonne - BnF - MCC) ; LAM, Lutheries, Acoustique, Musique - Institut Jean le Rond d'Alembert (UMR 7190 - CNRS – Sorbonne Université – ministère de la Culture) ; Musée de la musique, Cité de la musique - Philharmonie de Paris : Centre de recherche sur la conservation (USR 3224 - CRCC - LRMH - ECR Musée de la musique) ; Unité systématique et catégorisation culturelles (UMR 7206 - CNRS - MNHN - Paris 7) ; Pôle Supérieur Paris-Boulogne-Billancourt (PSPBB) ; Chœur & Orchestre Sorbonne Universités (COSU) ; UFR Musique et musicologie (Paris-Sorbonne) ; UFR d'Ingénierie (Sorbonne Université).
3. The SU laboratories at LabEx SMART are ISIR, the LIP6, the laboratoire Jacques Louis Lions, the L2E and STMS: http://www.smart-labex.fr/
4. The LabEx CAP http://labexcap.fr/ covers 17 university and higher-education (from: EHESS, EPHE, CNAM, ENC, ENSCI, ESCP Europe, ENSAPLV, INHA, INP, IRCAM, LCPI ParisTech, the université Paris 1 Panthéon-Sorbonne) and 8 museums and national establishments (BnF, Musée du Louvre, Centre Pompidou, musée du quai Branly – Jacques-Chirac, Musée des Arts décoratifs, Musée des Arts et Métiers, Sèvres – Cité de la céramique, La Cité de l'architecture et du patrimoine).
5. These projects are funded in a variety of ways. For example, the bilateral contract and INRIA international funding, a Fullbright scholarship, an international Canadian program, European projects, etc.

# INNOVATION AND RESEARCH RESOURCES

At the heart of societal and economic issues crossed with those of culture and information technologies, the research hosted by IRCAM are presented in the international research landscape as an interdisciplinary benchmark for sciences and technologies relative to sound and music constantly exposed to society's new needs and uses. This mediation assumes a wide range of forms according to the targets and spheres of activity and is supported by the department of Innovation and Research Resources at IRCAM. This department is in charge of IRCAM's supervision of the joint research laboratory as well as IRCAM's internet and information technology teams.

Distinctive to IRCAM is this gathering of a broad spectrum of scientific and technological skills that find applications in numerous fields centered on the production and diffusion of music and sound, digital arts, cultural industries, and throughout the fields with a sonic dimension: sound design (automotive, transportation, urban environments), multimodal interfaces, virtual and augmented reality, simulation, etc. With several decades of experience in the management of innovation and collective creativity, IRCAM is consulted concerning these themes (innovation, marketing, human resources, communication, etc.) on a regular basis by large groups.

IRCAM's art-innovation model has been extended to all technological and artistic fields in the framework of the STARTS Residencies program in the European project VERTIGO coordinated by IRCAM. This program organizes artist residencies in conjunction with technology research projects throughout Europe with the support of the European Union's STARTS initiative[1] (Science, Technology and the Arts). This model is presented during the annual Forum Vertigo, an interdisciplinary symposium held at the Centre Pompidou during the Mutations / Créations event, bringing together artists, researchers, engineers, and actors of innovation focusing on current issues in artistic creation in its relationship with sciences and technologies.

Services provided by the teams at IRCAM for private partners are eligible for fiscal advantages through the French Ministry of Research and Further Education's research tax credit (CIR). Connections with industry can be found in numerous collaborations: French and European R&D projects, research findings in the joint research laboratory (20 currently underway, 1/3 of which are coordinated by the joint research laboratory), research services, co-tutoring of CIFRE dissertations, accompaniment of start-ups (Phonotonic, Mogees, Niland [bought by Spotify in 2017], Antescofo, and HyVibe are the most recent examples), capital holding, and licensing. IRCAM is particularly active in this final aspect with several dozen ongoing licensing agreements with large groups as well as small businesses (primarily in Europe and North America) for technological building bricks in areas in which it is a leader (sound synthesis and processing, music indexing, sound spatialization, interaction performance/sound, etc.).

Professional technological development within the institute's research teams leading to the production of directly transferrable finalized, functional, optimized modules is a part of IRCAM's culture. More generally, this development activity can be found throughout a variety of realizations adapted to different targets. Over 20 software environments are developed by the teams (see Part 4), perfected incrementally in rhythm with the latest advances in research, used by other departments at IRCAM for artistic productions or training (composers, computer music designers, sound engineers, musicians, dancers, etc.), and disseminated by the IRCAM Forum to an international community of over 10,000 professional users. Applications with simplified uses can be found in a range of software bouquets.

1. https://starts.eu

# RESEARCH COLLABORATIONS AND RECENT PARTNERS

3D for All Kft (Hongrie)
A-Volute
Ableton (Allemagne)
Acapela
Arkamys
Arte
Artipolis
Artshare (Belgique)
Athena Research Center (Grèce)
Athens Technology Center (Grèce)
b<>com
Bass Nation (Finlande)
Bayerischer Rundfunk (Allemagne)
BBC (Royaume-Uni)
Bmat (Espagne)
Buffet-Crampon
Cabrilog (France)
Centre de psychiatrie et neurosciences (Inserm-université Paris Descartes)
Centre de recherche de l'Institut du cerveau et de la mœlle épinière (UMR 7225)
Centre hospitalier universitaire de Nice
Centre interfacultaire des sciences affectives (Suisse)
CHU La Conception-Marseille
CHU-Liège (Belgique)
CNES
CNMAT-UC Berkeley
CNRS Telecom ParisTech
Conservatoire national de musique et de danse de Paris
CTEL
Culture Tech
Dassault Systèmes
Deezer
Deutsches Forschungszentrum für Künstliche Intelligenz (Allemagne)
Dualo
Dubbing Brothers
École de musique Schulich (Canada)
École polytechnique fédérale de Lausanne (Suisse)
ECR-Musée de la musique
EHESS
elephantcandy (Pays-Bas)
Ellinogermaniki Agogi (Grèce)
EMI
EMPAC (États-Unis)
Ensad
esba TALM
Eurecom Sophia Antipolis
Eurescom
FINCONS SpA
Flux::
Flying Eye (Allemagne)
France Télévisions
Fratelli Piacenza SpA (Italie)
Fraunhofer (Allemagne)
GENESIS (Genesis Acoustics)
Goldsmiths' College (Royaume-Uni)
Grame
Haute École de musique de Genève (Suisse)
HearDis! Corporate Sound GmbH (Allemagne)
I3S
ID Scenes
Inova+ (Portugal)
Inria
Inserm-Sorbonne Université-faculté de médecine
Institut Jean Le Rond d'Alembert-CNRS
Instituts de recherche technologique
INTEGRAL Markt - und Meinungsforschungsges.m.b.H. (Allemagne)
IreMus
ISAE SUPAERO
ISIR
IUAV (Italie)
Kainos (Royaume-Uni)
Kantar Media
KTH (Suède)
Laboratoire de neurophysique et physiologie (UMR 8119, université Paris-Descartes)
Laboratoire des sciences, des procédés et des matériaux de l'université Paris-13
LaBRI
LAGEP
Lambde Limited (Roli)
LATMOS
Libelium Comunicaciones (Espagne)
LIMSI
LMA
Lovemonk S.L. (Espagne)
LPL
LRI
Magix (Allemagne)
Maha
Makemusic
Mathematics for more (Espagne)
Mogees
Musée du Quai Branly – Jacques-Chirac
Musimap (Belgique)
Niland
No Design
Nokia (Finlande)
Novelab
Novespace
Open University (Royaume-Uni)
Orange
Orbe
OSU Ecce Terra
Parisson
Parrot
Pathé
Philharmonie de Paris
Phonotonic
Playground (Suède)
Plus – Wireless Biosignals SA
PSA Peugeot Citroën
Qwant
Radio France
Reactable Systems (Espagne)
Renault
Sigma-Orionis
Sky-Deutschland (Allemagne)
SNCF
Somethin'Else Sound Directions Ltd (Royaume-Uni)
Soundtrack Your Brand (Suède)
StreetLab
Stromatolite Innovation Lab (Royaume-Uni)
Stupeflix
Supelec
Technicolor
Technische Universität Berlin (Allemagne)
Technische Universität Wien (Autriche)
Testaluna (Italie)
Trinity College Dublin (Irlande)
Trinnov-Audio
Ubisoft
UC Limburg (Belgique)
Universal (Royaume Uni)
Universal Music Publishing Classical
Université Ben Gurion (Israël)
Université de Bretagne occidentale
Université de Fribourg (Suisse)
Université de Gênes (Italie)
Université de Thessalonique (Grèce)
Université de Zaragoza (Espagne)
Université Kiel(Allemagne)
Université McGill (Canada)
Université Pompeu Fabra (Espagne)
UVI Sounds and Software
Velti (Grèce)
Vi-live
Vizion'r
Xtranormal

# 2

# THE TEAMS

## SOUND SYSTEMS AND SIGNALS: AUDIO/ACOUSTICS, INSTRUMENTS

**Head Researcher: Thomas Hélie**

### ACTIVITIES

–

The team Sound Systems and Signals: Audio/Acoustics, InstruMents creates theoretical, technological, and experimental tools based on multiphysic systems and the sound signals they produce. The team is interested in exploring, understanding, reproducing, and inventing sound objects in audio, acoustics, focusing on musical instruments and the production of the human voice.

More precisely, the objectives are to model, simulate, identify, and optimize these systems (voice, musician/instrument, loudspeakers,

audio electronic effects, etc.) using physics to reveal and benefit from their intrinsic structures. The team creates methods for analysis, transformation, control, and simulation of sounds as well as tools to assist conception in virtual, real, or hybrid paradigms.

This global approach, "systems and signals" creates a synergy in the conception of these tools, connecting—not merely juxtaposing—several disciplines and scientific fields: physics, theory of systems and control, differential geometry, numerical analysis, signal processing, computer science, electronics, mechatronics, and robotics.

Target applications concern scientific, artistic, educational, and healthcare domains.

### INTEREST AREAS AND ASSOCIATED PROJECTS

–

- Physical modeling of musical instruments and voice
- Sound synthesis based on physical modeling
- Experimental robotic platforms: robotic artificial mouth to play brass instruments, robotic apparatus at a 1:1 scale
- Identification of non-linear systems
- Control of non-linear systems
- Animaglotte project (system of artificial animation of the larynx ex vivo, p. 41)
- Infidhem project (interconnected systems of infinite dimension for heterogenous environments, p. 28)
- iMuSciCA project (p. 24)
- Ondes Martenot project (p. 43)
- UMN project (Unfold Mechanics Network)

### SPECIALIST AREAS

–

Acoustics, mechanics, non-linear systems, control, signal theory, differential geometry, digital analysis, experimentation, mechatronics, real and virtual instrument-making, sound synthesis

### COLLABORATIONS

–

Athena-RIC (Greece), Cabrilog SAS, Cambridge University (United Kingdom), C2RMF-Louvre, C2RMF et Louvre-Lens, Centre Bernoulli- EPFL (Switzerland), CHU Liège, EPCC-Edinburgh Univ. (United Kingdom), GIPSA-lab, IJLRA-Sorbonne Université, IMJPRG-Sorbonne Université, Imperial College London (United Kingdom), ISAE-SUPAÉRO, LAGEP-université Lyon-1, LaSiE-université de La Rochelle, LEOPOLY (Hungary), LMA-CNRS, LMD-ENS, LPL, Mines ParisTech, Musée de la musique – Philharmonie de Paris, NTNU (Norway), Thalès Group.

### TEAM

–

Researchers: B. d'Andréa-Novel (Sorbonne Université), J. Bensoam, H. Boutin (Sorbonne Université), R. Caussé (emerita), T. Hélie (CNRS), D. Roze (CNRS)

Engineers: Q. Lamerand, R. Piéchaud

Technicians: D. Chalabi (apprentice, CNRS)

Doctoral Students: D. Bouvier, T. Lebrun, R. Muller, M. Wijnand

## ACOUSTIC AND COGNITIVE SPACES

**Head Researcher: Olivier Warusfel**

### ACTIVITIES

–

The Acoustic and Cognitive Spaces activity of research and development centers on the reproduction, the analysis/synthesis, and the perception of sound spaces. The team's scientific disciplines are signal processing and acoustics for the elaboration of spatialized audio reproduction techniques and methods for the analysis/synthesis of a sound field.

In parallel, the team devotes a large percentage of its time to cognitive studies on multisensorial integration for a rational development of new sonic mediations based on body/hearing/space interaction. The scientific activities described below are combined with the development of software libraries. These developments build on the team's expertise, and its academic and experimental research activities and are the major vector of our relationship with musical creation and other application domains.

The work carried out concerning spatialization techniques are concentrated on models based on a physical formalism of the sound field. The primary objective is the development of a formal framework for the analysis/synthesis of the sound field using spatial room impulse responses (SRIR). The SRIRs are generally measured using spherical arrays featuring several dozen transducers (microphones and/or loudspeakers). The principal application concerns the development of convolution reverberators using these high spatial resolution SRIRs to faithfully reproduce the complexity of a sound field.

The technique of binaural spatialization using headphones is also a focus of our attention. The evolution of listening practices and the democratization of interactive applications tend to favor listening with headphones through smartphones. Taking advantage of this sonic immersion, binaural listening has become the primary vector of tridimensional listening. Based on the exploitation of head-related transfer functions (HRTFs), it is the only approach that currently ensures a precise and dynamic reconstruction of the perceptual cues responsible for auditory localization. It has become the reference tool for experimental research in connection with spatial cognition in a multisensorial context and for virtual reality applications.

These 3D audio spatialization techniques associated with a tracking system that captures the movements of a performer or a member of the audience, constitute an organologic base essential for addressing questions on "musical, sound, and multimedia interaction". They offer an opportunity to reflect on the "cognitive foundation" related to the feeling of space, in particular on the coordination necessary among various sensory modalities for the perception and cognition of space. More specifically, we wish to highlight the importance of the processes of integration between idiothetic cues (related to our motor actions) and the acoustic cues (localization, distance, reverberation, etc.) used by the central nervous system to create a spatial representation of the perceived environment.

On the musical level, our ambition is to provide models and tools that enable composers to include sounds in a given space throughout the compositional process: from writing to concert. This contributes to making spatialization a parameter of musical writing. In the arts, this research also applies to post-production, to interactive sound installations, and to dance via the questions related to sound/space/body interaction. The incorporation of sound spatialization in virtual reality environments creates the opportunity for scientific applications to be used in neuroscience research, therapeutic systems, or transportation simulators.

### MAJOR THEMES AND ASSOCIATED PROJECTS

–

- Sound Spatialization: analysis of sound scenes via spherical arrays of transducers (p. 27), Hypbrid reverberation and directional impulse responses (p. 31), synthesis of sound fields by high density spatial networks (p. 33), Orpheus project (p. 35), binaural listening and the Bili project (p. 36), distributed spatialization (p. 31)
- Cognitive Foundations: auditory spatial cognition (p. 42), Audioself project (p. 46), spatial integration and emotion (p. 42), Entrecorps project (p. 39)
- Software: Spatialisateur (p.75), ToscA (p. 78), Panoramix (p. 74), ADMix Tools (p. 64)

### COLLABORATIONS

–

ARI-ÖAW (Austria), Bayerischer Rundfunk (Germany), BBC (United Kingdom), b<>com, Ben Gurion University (Israel), Conservatoire national supérieur de musique et de danse de Paris, CNES, elephantcandy (Netherlands, France Télévisions, Fraunhofer ISS (Germany), Hôpital de la Salpêtrière, HEGP, Hôpital universitaire de Zurich (Switzerland), IRBA, IRT (Germany), L-Acoustics, LAM, McGill University (Canada), Orange-Labs, RWTH (Germany), Radio France, RPI (United States).

### TEAM

–

Researchers: M. Noisternig, I. Viaud-Delmon (CNRS), O. Warusfel
Engineers: T. Carpentier (CNRS)
Postdoctoral Researchers: L. Hobeika
Doctoral Students: J. Moreira, V. Martin, P. Massé, F. Zagala
Residents/Guests/Long-Term Collaborators: N. Schütz, C. Suied, M. Taffou

## SOUND PERCEPTION AND DESIGN

**Head Researcher: Patrick Susini**
**Deputy Head Researcher: Nicolas Misdariis**

### ACTIVITIES

–

The team's research focuses on the perception and cognition of sounds, bringing together knowledge from psychoacoustics, sound signal processing/synthesis, psychology, and cognitive neurosciences. More generally, the team's research revolves around creation and educational activities in the field of sound design.

Research focuses primarily on environmental sounds. This rarely studied subject provides access to different levels of cognitive representation in connection with the surface characteristics of sounds and the sources that produce them. The team's work therefore concerns the perceptive characterization of these sounds and of the cognitive mechanisms used to identify them.

Environmental sounds are often produced in a context of interaction combined with movement when manipulating an object. One of the major themes in the team is to continue to study environmental sounds in a process that combines perception and action. This is a vital and unique gateway for research in auditory cognition and for interactive sound design applications.

The team's research topics have been extended to vocal imitations as well as musical sounds and to the voice in connection with environmental sounds. This creates new perspectives for better understanding the cognitive processes involved by the identification of a sound source and by the emotional processing of sounds.

Applications for the team's work are found primarily in sound design: we create the rules, the tools, and the methods—derived from our fundamental work—with the aim of facilitating the creation of functional and pleasant sounds for our everyday lives. The societal impact of our work was awarded in 2014 with the Décibel d'Or Recherche prize granted by the French Comité National du Bruit, and two awards from the Société Française d'Acoustique in 2016.

Creation in sound design is carried out combining the "knowledge" with "know-how" of the team's researchers and associated composers (A. Cera, S. Gaxie, A. Sigman, R. Rivas) in different projects. There are three types of creative projects carried out by the team: industrial projects in automotive with Renault, and the luxury industry with Krug, for example; research projects for the creation of sound signaling; and educational projects for a variety of workshops on subjects that range from sports equipment to mass transit.

Finally, the team has associated disparate educational actions:

- Organization of applicative workshops in partnership with industrial companies and collectivities on subjects ranging from sound signaling for tramways for the city of Le Mans to reinforcing the well-being of patients in the Sainte Anne hospital in Paris
- Organization of professional training courses that combine a project with classwork in connection with IRCAM's department of Education and Cultural Outreach

### MAJOR THEMES AND CURRENT PROJECTS

–

- Sonie (loudness) of non-stationary and multi-source sounds: MoUVie (p. 26)
- Perception of multisource sound scenes: MoUVie (p. 26)
- Analysis and perceptive characterization of complex sound scenes and environments: Houle (p. 62), MoUVie (p. 26)
- Sound, imitation, and identification: Skat-VG (p. 45)
- Sound and interactions: Legos (p. 47)
- Sound and emotions: CREAM (p. 40), Reflets (p. 39)
- Sound signaling and HNI
- Tools and methods: Skat-VG (p. 45)

### SPECIALIST AREAS

–

Psychoacoustics, experimental psychology, cognitive neurosciences, acoustics, sound design, data analysis

### COLLABORATIONS

–

TU-Berlin (Germany), J. Gutenberg Universität (Germany), McGill University (Canada), University of Canberra (Australia), university IUAV of Venice (Italy), Politecnico di Torino (Italy), Carnegie Mellon University (United States), KTH (Sweden), Lunds universitet (Sweden), ZHdK (Switzerland), LVA-INSA, LMA, CR-ICM, LAM, LAPPS , Ebam TALM, ENSCI, université de Strasbourg, Centre d'études techniques de l'équipement, IFSTTAR, Renault, PSA, SNCF, EDF, Klaxon, LAPS-Design, Genesis, Aximum, Univers Sons, Région Île-de-France.

### TEAM

–

Researchers: J.-J. Aucouturier (CNRS), M. Ayari (université de Strasbourg), O. Houix, M. Liuni, N. Misdariis, P. Susini
Engineer: L. Seropian (CNRS)
Postdoctoral Researchers: L. Goupil, R. Leiba, E. Ponsot, M. Vannier, L. Zattra
Doctoral Students: P. Arias, L. Rachman

## SOUND ANALYSIS & SYNTHESIS

**Head Researcher: Axel Roebel**

### ACTIVITIES

–

The Sound Analysis & Synthesis team carries out research and development activity in sound analysis, transformation, and synthesis of sound signals.

The analysis of sounds includes methods enabling the permanent extraction or automatic structuring of diverse sorts of information given off by the signal, such as the fundamental frequency or the spectral evolution determining the pitch and timbre of a perceived sound. Information outside what is strictly musical is also taken into consideration, notably concerning industrial acoustics, sound design and multimedia as well as the automatic indexing of recorded sounds. The methods used are based on signal processing, statistical analysis, learning techniques and deep learning, recognition techniques, but also on knowledge of auditory perception.

Techniques of transformation and synthesis are initially created to respond to the needs of musicians in the production of new sounds and new music. A typical example of this is the computer synthesis of a virtual choir for an opera. However, the work is frequently applied to the fields of mobile telecommunications, video games, and virtual reality.

Analysis and synthesis is based on the conception of both signal models (the modeling of sound effects produced via their signals) and physical models (the acoustical modeling of the source of sound production).

These models are integrated into pieces of computer software for PC or Macintosh (Mac OS X, Windows, or Linux) equipped with graphical interfaces specifically designed for musicians, but adaptable for sound engineers, acousticians and amateurs.

### SPECIALIST AREAS

–

Signal processing, statistics, information theory, learning techniques, form recognition, digital analysis, modeling

### MAJOR THEMES AND CURRENT PROJECTS

–

- Signal models: processing with a phase vocoder (p. 34)
- Sound characterization: automatic indexing (p.28), projects Skat-VG (p. 45), 3DTVS (p. 37)
- Analysis, transformation, and synthesis of the voice and speech, models of the voice, conversion of speakers: Chanter project (p. 44)
- Analysis of sound scenes: projects 3DTVS (p. 37) and Houle (p. 62)
- Software: SuperVP (p. 76), Audiosculpt (p. 66), IRCAMLab TS (p. 68), The Snail-Absolute Tuning (p. 77)

### COLLABORATIONS

–

Aalto University (Finland), Ableton (Germany), Acapela Group, Acoustic Research Institute (Austria), AudioGaming, Arte, Artipolis (Luxembourg), Dualo, Mines ParisTech, Exalead, ExeQuo, France Télécom R&D / Orange, Flying Eye (Germany), Game Audio Factory, Genesis SA, HHI Berlin (Germany), Idol, INESC, ISAE-SUPAERO, ISIR/Sorbonne Université, IUAV (Italy), Kantar Media, KTH (Sweden), LIMSI, LMA, LIA, Laboratoire Parole et Langage - Hôpital La Conception, NuHag (Austria), OFAI, ParisTech, Queen Mary University of London (United Kingdom), SCREAM National Cheng Kung University (Taiwan), Sky Deutschland (Germany), Smartlog, Smartsound, Sony Music France, Stupeflix, Ubisoft, UPF/MTG (Spain), université Huddersfield (Royaume-Uni), université Lumière-Lyon-2, université Paris-8, Univers Sons (UVI), Viddiga, Vizion'R, Voxygen, Yacast.Deutschland (Germany), Smartlog, Smartsound, Sony Music France, Stupeflix, Ubisoft, UPF/MTG (Spain)

### TEAM

–

Researchers: N. Obin (Sorbonne Université), G. Peeters, A. Roebel
Engineers: F. Cornu, C. Picasso
Postdoctoral Researchers: D. Fourer, R. Mignot, P. Lafitte, D. Basaran
Doctoral Students: H. Caracalla, A. Cohen Hadria, G. Doras, R. A. Ferro Mendes, H. Fouroughmand, C. Jacques, P. C. Li, Y. J. Lin, G. Meseguer Brocal
Residents/Guests/Long-Term Collaborators: J. Smith

## SOUND MUSIC MOVEMENT INTERACTION

**Head Researcher: Frédéric Bevilacqua**

### ACTIVITIES

–

The Sound Music Movement Interaction team (previously known as the Real-Time Musical Interactions team) carries out research and development on interactive systems dedicated to music and performances.

Our work relates to all aspects of the interactive process, including the capture and multimodal analysis of the gestures and sounds created by musicians, tools for the synchronization and management of interaction, as well as techniques for real-time synthesis and sound processing. These research projects and their associated computer developments are generally carried out within the framework of interdisciplinary projects that include scientists, artists, teachers, and designers and find applications in creative projects, music education, movement learning, or in digital audio industrial fields.

### MAJOR THEMES AND CURRENT PROJECTS

–

• Modeling and Analysis of Sounds and Gestures

This theme covers the theoretical developments concerning the analysis of the sound and gesture flow, or more generally, multi-modal temporal morphologies. This research concerns diverse techniques for audio analysis, the study of the gestures of performing musicians or dancers (see gesture analysis and recognition, p. 41).

• Technologies for Multimodal Interaction

This theme concerns our tools for analysis and multimodal recognition of movements and sound; tools for synchronization (gesture following, for example) and visualization.

• Interactive Sound Synthesis and Processing

This focuses essentially on synthesis and sound processing methods based on recorded sounds or large collections of sound (corpus-based concatenative synthesis, p. 32)

• Systems for Gesture Capture and Augmented Instruments

This theme focuses on the developments the team has made in terms of gestural interfaces and augmented instruments for music and performances (augmented instruments, p. 29).

### NATIONAL AND EUROPEAN PROJECTS

–

- Legos (p. 47)
- COSIMA (p. 59)
- SkAT-VG (p. 45)
- Rapid-Mix (p. 58)
- Musicbricks (p. 44)
- Wave (p. 61)

### SPECIALIST AREAS

–

Interactivity, real-time computer science, human-computer interaction, signal processing, motion capture, modeling sound and gesture, statistical modeling and automatic learning, real-time sound analysis and synthesis

### COLLABORATIONS

–

Atelier des feuillantines, BEK (Norway), CNMAT Berkeley (United States), Cycling'74 (United States), ENSAD, ENSCI, GRAME, HKU (Netherlands), Hôpital Pitié-Salpêtrière, ICK Amsterdam (Netherlands), IEM (Austria), ISIR-CNRS Sorbonne Université, Little Heart Movement, Mogees (United Kingdom/Italy), No Design, Motion Bank (Germany), LPP-CNRS université Paris-Descartes, université Pompeu Fabra (Spain), UserStudio, CRI-Paris université Paris-Descartes, Goldsmiths University of London (United Kingdom), université de Genève (Switzerland), LIMSI-CNRS université Paris-Sud, LRI-CNRS université Paris-Sud, Orbe.mobi, Plux (Portugal), ReacTable Systems (Spain), UCL (United Kingdom), Univers Sons/Ultimate Sound bank, Universidad Carlos III Madrid (Spain), université de Gênes (Italy), McGill University (Canada), ZhDK (Switzerland).

### TEAM

–

Researchers: F. Bevilacqua, D. Schwarz,
Engineers: R. Borghesi, J.-P. Lambert, B. Matuszewski
Doctoral Students: P. Antoniadis, H. Scurto, Y. Zeitoun
Residents/Guests/Long-Term Collaborators: B. Caramiaux (CNRS), M. Suarez Cifuentes, J. Françoise (CNRS)

## MUSICAL REPRESENTATIONS

**Head Researcher: Gérard Assayag**

### ACTIVITÉS

–

The Musical Representations team works on the formal structures of music and creative environments for composition and musical interaction. This work finds application in computer-assisted composition (CAC), performance, improvisation, performance and computational musicology. Reflection on the high-level representation of musical concepts and structures supported by original computer languages developed by the team, leads to the conception of models which can be used for musical analysis and creation.

On the musicology side, tools for representation and modeling enable a truly experimental approach that significantly rejuvenates this discipline.

On the creation side, the objective is to design musical companions that interact with composers, musicians, sound engineers, etc. throughout the musical workflow. The software developed has been distributed to a large community of musicians, materializing original forms of thought connected to the particular characteristics of the computer supports they represent (and execute): the final score, the score's different levels of formal elaboration, its algorithmic generators, and live interaction during a performance.

For the past few years, the team has worked on symbolic interaction and artificial creativity in its work via projects on artificial listening, synchronization of musical signals and score following (a subject that led to the creation of an INRIA team-project), orchestration assistance (using the analysis of temporal series and deep learning techniques), and engineering intelligent agents capable of listening, learning, and musical interaction in improvised contexts.

The team has a long history of collaborations with composers and musicians both from IRCAM and elsewhere. Archives of this work can be found in three volumes of the OM Composer's Book, guaranteeing its international dissemination and continuity.

### MAJOR THEMES AND CURRENT PROJECTS

–

- Computer-assisted composition: assisted composition (p. 52), orchestration assistance (p. 30)
- Control of synthesis and spatialization, creative systems to write for time, sound, space, and interaction
- Mathematics and music (p. 56)
- Computer languages for music: Open Music (p. 72), Antescofo (p. 65)
- Modeling style, dynamics of improvised interaction: improvised musical interactions (p. 55), DYCi2 (p. 51)
- New interfaces for composers and teaching
- Musicology and computational analysis
- Efficient search of temporal series (p. 30)
- Writing synchronous time (p. 54)
- Projects EfficacE (p. 60) and Inedit (p. 61)
- Augmented paper for CAC (p. 43)

### SPECIALIST AREAS

–

Computer-assisted composition and analysis, computer musicology, cognitive musicology, artificial intelligence, computer languages, algebraic and geometric methods, symbolic interactions, languages for synchronous time and tempered time, executable notations

### COLLABORATIONS

–

Bergen Center for Electronic Arts (Norway), CIRMMT/ McGill University (Canada), City University London, Conservatoire national supérieur de musique et danse de Paris, Columbia University, CNMAT/UC Berkeley, Electronic Music Foundation, Gmem, Grame Lyon, HEM Genève, École normale supérieure Paris, EsMuC Barcelone, Harvard University, Inria, IReMus – Sorbonne Paris-4, Jyvaskyla University, université de Bologne, USC Los Angeles, université Marc Bloch Strasbourg, Pontificad Javeriana Cali, université Paris-Sud Orsay, université de Pise, Sorbonne Université, UCSD San Diego, Yale University, Minnesota University, Washington University.

### TEAM

–

Researchers: C. Agon Amado (Sorbonne Université), M. Andreatta, (CNRS), G. Assayag, M. Ayari (université de Strasbourg), G. Bloch, J. Bresson, P. Esling (Sorbonne Université), J.-L. Giavitto (CNRS), K. Haddad, F. Jacquemard (Inria), M. Malt,

Postdoctoral Researchers: E. C. Cella, J. Nika

Doctoral Students: J. Blondeau, D. Cabanzo, T. Carsault, A. Chemla-Romeu-Santos, L. Crestel, K. Deguernel, P. Donat-Bouillud, J.-M. Fernandez, G. Genuys, D. Ghisi, F. Le Bel, L. Morciano, C. Poncelet, M. Prang, A. Ratoci, P. Talbot

Residents/Guests/Long-Term Collaborators: A. Farhang, M. Magalhaes, C. Rueda

## ANALYSIS OF MUSICAL PRACTICES

**Head Researcher: Nicolas Donin**

### ACTIVITIES

–

The Analysis of Musical Practices team carries out research on the works and scholarly practices that form IRCAM's heart: composition, improvisation, performance.

To this end, the team is developing an empirical and technological musicology, whose goal is to describe (or even to modify) knowledge outside of the realm of music analysis, through novel methods of data collection.

The goal is to characterize the phenomena under examination without reducing them a priori to their apparent structure (formalizable procedures, canonical texts), but in addressing them in all their dynamic, creative, contextualized, social, and cultural complexities. This can be applied to the study of contemporary activities and situations (e.g. a performer's work at home, in rehearsal, in concert, a composition using computer-music technology, the production of a multimedia musical analysis for instructive purposes, etc.) or to the study of past practices.

The various projects carried out offer a wide range of technological and social interdisciplinary configurations, depending on the study's focus. Thus, a study of the compositional activity carried out by the composer Philippe Leroux from 2001 to 2006 demanded a connection be made between the artist's manuscripts and cognitive anthropology: it coincided and interfered with the elaboration of a new work by the composer. Finally, it takes place in relationship with the development of a hypermedia document that traces the creative process and the way in which the composer listens to his own piece. Or again: the study of musical analysis in France (end of the 19th century, beginning of the 20th century) required the creation of a relationship between a micro-historic investigation into writing conditions as well as the reception of these texts, with a critical appraisal of the segmentation and categorization procedures of the 19th and 20th centuries and led to the creation of a computer tool for musical analysis.

These activities are disseminated through publications and symposiums within the scientific communities concerned (primarily historic musicology, musical analysis, cognitive ergonomics, sociology, genetic criticism, and music technology); they are also made public in other ways: the creation of computer tools and hypermedia documents, short documentary films, teaching, and conferences.

- Composition analysis: project Gemme (p. 60)
- Performance practices
- Contemporary listening practices
- Listening practices and musical analysis: a historical approach
- Musicology and the humanities/synthesis work: theories of musical composition of the 20th (p. 57)
- Multimedia publication tools for musicology de publication (p. 56), Wave project (p. 61)
- Wasabi project (p. 25)
- MICA project (p. 48)

### COLLABORATIONS

–

Centre de recherches sur les arts et le langage (EHESS, CNRS), CIRMMT/McGill University (United States), Conservatoire national supérieur de musique et de danse de Paris, Oxford University Faculty of Music, Haute École de musique/conservatoire de Genève, Institut des textes et manuscrits modernes (CNRS-ENS Ulm), OICRM/université de Montréal.

### TEAM

–

Researchers: A. Bonardi, C. Canonne (CNRS), N. Donin, L. Feneyrou (CNRS), F.-X. Féron (CNRS)

Postdoctoral researchers: F. Ribac (CNRS delegation), A. Robert (Sorbonne Université), L. Zattra

Residents/Guests/Long-Term Collaborators: A.-S. Barthel-Calvet, A. Fryberger

# 3

# THE PROJECTS

# //////////////////////// SOUND WORKSHOP

## FUTUREPULSE

–

Team Involved: Sound Analysis & Synthesis
Funding: H2020 ICT-19-2017
Calendar: September, 2017—August, 2020
Partners: Bmat (coordinator, Spain), Athens Technology Center (Greece), Musimap (Belgium), Playground (Sweden), Bass Nation (Finland), Soundtrack Your Brand (Sweden)

Music is one of the fastest evolving media industries, currently undergoing a transformation at the nexus of music streaming, social media and convergence technologies. The availability of huge music catalogues and choices has rendered the problems of recommendation and discovery as key in the competition for audience, while the continuous access to multiple sources of music consumption have resulted in a dynamic audience, characterized by a highly diverse set of tastes and volatility in preferences which also depend on the context of music consumption.
To serve the increasingly complex needs of the music ecosystem, FuturePulse will develop and pilot test a novel, close to market, music platform in three high-impact use cases:

- Record Labels
- Live Music
- Online Music Platforms

The project will help music companies leverage a variety of music data and content, ranging from broadcasters (TV, radio) and music streaming data, to sales statistics and streams of music-focused social media discussions, interactions and content, through sophisticated analytics and predictive modelling services to make highly informed business decisions, to better understand their audience and the music trends of the future, and ultimately to make music distribution more effective and profitable.

IRCAM's role in the project consists of the adaptation and perfection of the institute's technologies for indexing, enabling automatic extraction of information from musical recordings.

## IMUSICA

–

Team Involved: S3AM
Funding: European Commission Call H2020 ICT-22-2016
Calendar: January 2017 — June 2019
Partners: Athena Research Center (coordinator, Greece), UC Limburg (Belgium), Ellinogermaniki Agogi (Greece), 3D for All Kft (Hungary), Cabrilog (France), Mathematics for More (Spain), University of Fribourg (Switzerland).

iMuSciCA is an educational project that intends to improve the acquisition of scientific skills in secondary-level education. iMuSciCA offers new learning methods using innovative "discovery" technologies—both personalized and collaborative—in order to make certain scientific subjects, in particular math and physics, more appealing. The joint research laboratory brings its skills in virtual instrument making via Modalys, giving life to 3D instruments created by students. Some of these instruments feature parameters that can be tested and modified (geometry, material, etc.) in real-time.
The iMuSciCA environment, implemented through a collaboration among different industrial and educational partners, will be tested and assessed during a pilot program with a large number of students and teachers in three European countries: Belgium, France, and Greece.

## WASABI

**Web Audio Semantic Aggregated in the Browser for Indexation**

—

Teams Involved: Analysis of Musical Practices, Sound Analysis & Synthesis, IRCAM Resource Center
Funding: ANR, programme générique
Calendar: October 2016 — March 2021
Partners: I3S (coordinator), Parisson, Deezer, Radio France

The goal of the WASABI project is to define an optimized methodology for indexing music for the Web for large databases by linking metadata from audio analysis, Semantic Web techniques, and the analysis of textual data such as song lyrics using natural language analysis and to compare this methodology with case scenarios to develop unique services and applications using Web Audio technologies.

The project entails using algorithms to extract musical information and Semantic Web techniques to produce more consistent musical knowledge bases for streaming services and music databases. Services using Web Semantic data like LastFM, MusicBrainz, or DBPedia use the extraction of structural data, connecting works to metadata such as the producer, the recording studio name, the composer, the release year, or the subjects in the lyrics for example. The data in free text like the lyrics are also analyzed to determine the musical context of the piece. Web Audio technologies make it possible to explore these musical spaces improved with analyses such as high-level musical indexation: detecting emotion and plagiarism, detecting and characterizing the singing voice, detecting the structure and separating the different sources.

Open source software bricks and "open data" online services will be proposed at the end of the project for:

- The visualization of audio metadata and listening to unmixed tracks in a browser as well as using the latest Web Audio API technologies (mixing in real-time, audio effects)
- Automatic processing of lyrics, recognition and merging named entities, collaborative annotation and correction
- Access to a Web service with an API offering an environment in which to study musical similarities from audio and semantic analyses

These software bricks will be used in the development of formalized demonstrators with our partners and collaborators (journalists and composers), using the new Web Audio API standard making it possible to develop musical applications accessible to the general public via a Web browser.

## MOUVIE

**Mobility and Quality of Life in Urban Areas**

—

Team Involved: Sound Perception and Design
Funding: Chaire d'Excellence — Fondation Sorbonne Université
Calendar: April 2014—April 2019
Partners: OSU Ecce Terra, Laboratoire Atmosphère, Milieux, Observations Spatiales (LATMOS), Inserm, Sorbonne Université — Institut Jean le Rond d'Alembert (MPIA Team)

**The Impact of Atmospheric and Noise Pollution on the Health of Urban Populations**

The chair in mobility and quality of life in urban areas will measure the impacts of pollution by gases, particles and noise on the quality of life and health of the residents of a city. In response to this societal issue, the car manufacturers PSA Peugeot Citroën and Renault have realized the importance of supporting the sponsored research project that Sorbonne Université is running together with France's Institute of Health and Medical Research (INSERM), Atmosphere, Media and Spatial Observation Laboratory (LATMOS), Ecce Terra, the Jean le Rond d'Alembert Institute and the Institute for Acoustics and Music Research and Coordination (IRCAM). "By taking the many factors involved in chemical air pollution into account, as well as unconventional sources of pollution such as noise, the chair can bring a new dimension to the concept of the urban environment, which may be promising for the future," commented Laurence Eymard, Chair.

**An Inter-Disciplinary Research Team of Experts**

A chair headed by Laurence Eymard, coordinating an inter-disciplinary team of experts in acoustics, air quality and health. Laurence Eymard will coordinate a scientific community of more than 1,000 people specialized in earth system and environmental sciences, ranging from biodiversity to health, and from deep-earth science to planetary science, via climatology. Régis Marchiano, a professor at Sorbonne Université, works on linear and non-linear acoustic wave propagation in complex media at the Sorbonne Université/CNRS Jean le Rond d'Alembert Institute. Nicolas Misdariis is joint head of the Sound Perception and Design team at the Institute for Acoustics and Music Research and Coordination (IRCAM). Sébastien Payan, professor at Sorbonne Université and a research fellow at LATMOS, researches various issues on the boundary between molecular physics and atmospheric physics. In health, Isabella Annesi-Maesano, head of the INSERM/Sorbonne Université Epidemiology of Allergic and Respiratory Diseases team, runs research projects to understand the etiology of rhinitis, asthma, chronic obstructive broncho-pneumopathy, and identify subjects at risk. Annick Clément, professor of pediatric pneumology at Sorbonne Université, runs the pediatric pneumology ward at Trousseau Hospital. She also coordinates a European research group on rare lung diseases in children and sits on numerous scientific boards in France and around the world.

**Major Issues for Private-Sector Sponsors of the New Chair**

Sylvain Allano, Scientific and Future Technology Director at PSA Peugeot Citroen: "We thought a sponsored chair was an appropriate way to support high-quality inter-disciplinary research in areas with a strong societal dimension, such as air quality and noise pollution. The research results should provide the kind of solid, recognized scientific basis that often is missing from societal and environmental debates." Rémi Bastien, Head of Innovation Engineering at Renault: "Society expects mobility, which is both a necessity and a fundamental freedom, to contribute to a higher quality of life, especially in urban areas. Current forms of mobility have increasingly unacceptable impacts. As manufacturers, we are part of the problem, so we want to be part of the solution. The best way is to cooperate with high-level inter-disciplinary scientists. The new chair offers us a unique opportunity to pave the way for sustainable, reliable solutions backed by in-depth scientific expertise." Claire Martin, Vice President of CSR at Renault and Managing Director of the Renault Foundation: "The better we understand impacts in their complexity, the more able the government and industry will be to develop appropriate responses. It is also legitimate and necessary to analyze the technical, economic, social, health and other aspects of phenomena in order to send clear, objective messages to all audiences."

## 3DR3

**Room Impulse Response Renderer**

–

Team Involved: Acoustic and Cognitive Spaces, Production Department

3DR3 brings together the Acoustic and Cognitive Spaces research team with the Production department's sound team to work on the development of a generic environment for multichannel mixing (see panoramix, p. 74), using a unique model of hybrid reverberation recently developed by the Acoustic and Cognitive Spaces team. This innovative model takes advantage of convolution reverberators that accurately reproduce the "natural" sound of the reverberation measured in a room compared to feedback delay network reverberators. This extends the principle of spatialized responses (see Analysis of a Sound Scene through Spherical Transducer Arrays) while associating the flexibility of Spat~ and its control via perceptive factors. Sound sources may be spatialized using the room impulse responses measured in the concert venue where the performance takes place. This allows for a seamless integration between the virtualised sound sources and the live main recording.

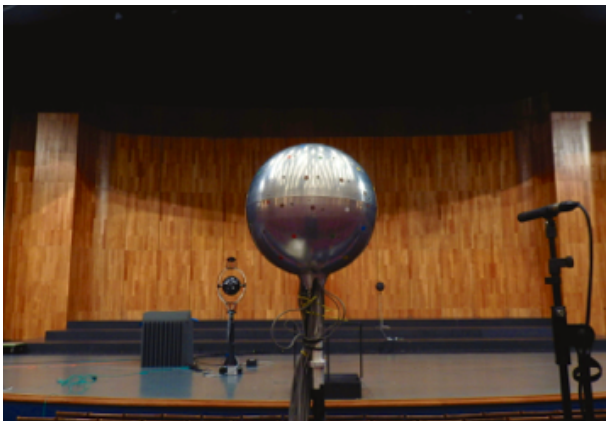
## ANALYSIS OF A SOUND SCENE THROUGH SPHERICAL TRANSDUCERS ARRAYS

–

Team Involved: Acoustic and Cognitive Spaces

Partners: Ben Gurion University, Supelec

The spatial properties of a sound scene are determinant for the perception of sounds in a room. These spatial properties are usually studied through separate investigations of either a microphone array, or a loudspeaker array. Measures that take into account both the microphone and the loudspeaker arrays, called MIMO (multiple-input multiple output), make it possible to augment the description of the sound scene. By using MIMO systems, the impulse responses of the room can be represented using a matrix and the analysis of the sound scene's spatial properties can call upon linear algebra tools. For example, the rank of the matrix and its kernel reveal interesting information such as the number of significant reflections, the direction of their incidences on the microphone, and the direction of their emission from the loudspeakers. This project is the object of theoretical work on the formalization of the problem and the optimization of the design of the spherical transducer arrays and also includes an experimental aspect dedicated to the room acoustic analysis.



Acoustic measures at *Grosses Festspielhaus* in Salzburg combining a spherical microphone array (64 capsules) developed at IRCAM and a spherical loudspeakers array developed at RWTH (Aachen).
Photo © Olivier Warusfel

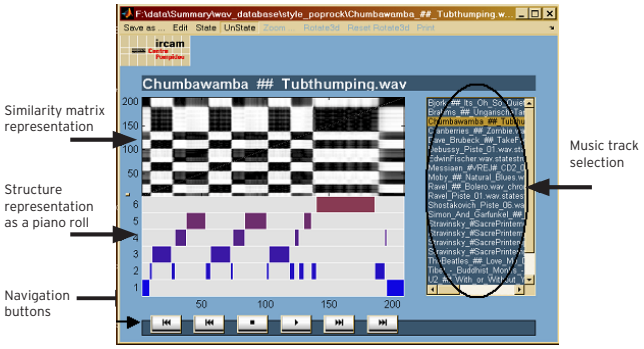
## AUTOMATIC MUSIC INDEXING

–

Team Involved: Sound Analysis & Synthesis

During projects presented hereinafter, the following subjects were addressed:

- Methods for the automatic extraction of musical descriptors for a piece of music such as the tempo, location of beats, metrical, tonality, or a temporal grouping for a chord. These descriptors facilitate the automatic classification of a piece and can be used for content-based searches in sound databases.
- Musical excerpt recognition methods, designed to automatically identify excerpts from pieces of music using reference databases. These methods are based on a compact sound signature (fingerprint) encoding the essential information. These algorithms compare each fragment of sound under investigation, with those in the database.
- Methods for the estimation of the temporal structure of a piece of music in terms of the repetition of a section being listened to and enabling browsing within the temporal structure of the given musical piece.
- Methods for the automatic creation of audio summaries making it possible to quickly pre-listen to the contents of a given musical piece via its key points.




Representation of the temporal structure of a musical piece via signal analysis – Cuidado, SemanticHIFI, and Ecoute projects

## INFIDHEM

**Interconnected inFInite-Dimensional systems for Heterogeneous Media**

–

Team Involved: Sound Systems and Signals: Audio/Acoustics, InstruMents

Funding: ANR / DFG

Calendar: February 2017—January 2020

Partners: Deutsche Forschungsgemeinschaft, ISAE SUPAERO (Toulouse), LAGEP (Lyon), Kiel University

Motivated by recent technological progress in mechanical, aeronautical, energy systems and chemical engineering and novel computational tools, the analysis and control of infinite-dimensional systems became a field of major interest during the last decades. The basic concepts of classical systems theory have been progressively generalized to infinite-dimensional systems with contributions stemming from the mathematical as well as from the engineering community.

More recently, scientists became interested in understanding systems composed of distributed-parameter subsystems (described by systems of partial differential equations, PDEs) which interact in networks. Most of the existing literature on modeling, system analysis and control, deals with networks of homogeneous systems such as trusses of elastic rods or the heat conductivity properties of metal foams.

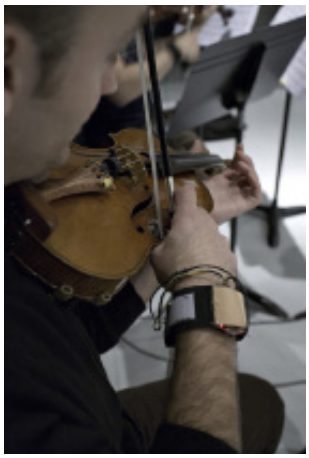
## AUGMENTED INSTRUMENTS

–

Teams Involved: Sound Music Movement Interaction, S3AM

Augmented instruments are acoustic instruments that have been fitted with sensors so that information concerning gestural parameters can be transmitted in real-time. These instruments are specifically adapted for mixed acoustic and electronic music. This work is carried out in close collaboration with composers and performers, the goal being the integration of this technology in their works. These developments are also a part of IRCAM's research into instrumental gesture.
This project began with the development of the augmented violin, which then took off due to the interest of various composers. We are now working on all the string quartet instruments, certain percussion instruments, and the piano.

Diverse computer techniques have also been developed to analyze, recognize, and follow instrumental gestures. For example, a list of musical elements such as playing styles or musical phrases could be defined by the composer and could be used as a foundation for interaction with sound processes including synthesis or spatialization. Finally, it is important to note that these augmented instruments are used in musical creations, but also in the educational arena.



Augmented Violin
© Florian Kleinefenn

## AUGMENTED INSTRUMENT-MAKING

–

Team Involved: S3AM

This project is at the heart of the research on musical instruments. Prototypes of musical instruments, or components of musical instruments, that can be used in concert have been produced in this project.  Work carried out includes a broadening of the range of mutes for brass instruments, studying a system for the automatic tuning of a kettledrum, the creation of a clarinet mouthpiece with a variable volume, and the creation of a bow and mouthpiece instrumented via several sensors. In addition, IRCAM has designed and owns an instrumentarium of SmartInstruments (guitars, string quartet, bass clarinet, brass mutes) equipped with sensors and actuators with acoustic properties that a musician or composer can modify, and can diffuse sounds from a variety of sources without an external loudspeaker. This led to the realization of COALA, an embedded low-latency active control system for SmartInstruments.

## COMPUTER-ASSISTED COMPOSITION (ORCHIDS)

–

Team Involved: Musical Representations

A unique and recent case of computer-assisted composition, this project addresses the question of orchestration via an automatic search of instrumentation and layering instruments approaching, depending on different acoustic similarity criteria, a target defined by the composer. Current research endeavors to make this dynamic orchestration paradigm heard, according to the targets with sonorous characteristics that vary with time.

Realized after the Orchidée software suite, Orchids is the first complete system for temporal computer-assisted orchestration and the optimization of timbre combinations. It provides an ensemble of algorithms making it possible to recreate any sound target that changes over time through a combination of instruments or samples, according to the psycho-acoustic criteria. This can help composers obtain unique timbre colors providing a multitude of effective solutions to best recreate the sound target. Through a large selection of functions, Orchids can also recreate the evolutions and abstract forms of spectral movements. Its results provide multiple orchestra scores that can be organized intuitively in order to quickly construct orchestral and musical ideas. This system provides several approximation algorithms that make it possible to conjointly optimize several timbre features. The advantages of the Orchids system lies in the fact that this approximation can be carried out separately on temporal forms, values, mean values or standard deviations (or any combination of the three) of each psycho-acoustic descriptor. In addition, users can also manually define a temporal deformation and carry out a multi-target search within several sound segments, making it possible to create full orchestral works in just a few seconds.

## EFFICIENT SEARCH OF TEMPORAL SERIES

–

Team Involved: Musical Representations

Searching for sounds can be a painful and tedious task when dealing with large-scale databases. Even when meta-information is available, query results are often far from the mental image imagined by the user. Today, there is no system that transforms the intuitive projection of a sound idea into an effective search; sound samples do not let users extract high-level information such as melody or lyrics from songs.

Beginning with this observation, we have developed a code that makes it possible to carry out effective searches on temporal forms and take into account the multidimensional nature of sound perception. This makes it possible to carry out searches based on the temporal form of the descriptors rather than on mean values. These descriptors are modeled to obtain their average, standard deviation as well as the form of their temporal evolution via a symbolic representation enabling both compact storage and an effective search. However, it was essential that the comparison of temporal series that make it possible to obtain a similarity based on perceptive criteria for objects that could possibly be very different mathematically. Using an approach derived from Dynamic Time Warping (DTW), we have developed a robust measure of similarity following non-linear distortions such as range, noise sound, and unique values. Thanks to a new algorithm for indexing, it is possible to obtain the best element from a database containing several million sound samples almost immediately.

Our study then opened to the implementation of higher-lever interactions. We studied the possibility of a query that is pertinent to several temporal curves simultaneously, going beyond the framework of the simple consideration of often less relevant criteria. Thanks to new heuristics, we have carried out the first precise multi-objective search algorithm for temporal series.

These techniques apply to all fields of scientific research due to the ubiquity of the temporal information. Multi-objective searches of temporal series are open to numerous applications in fields ranging from medical analysis to robotics.

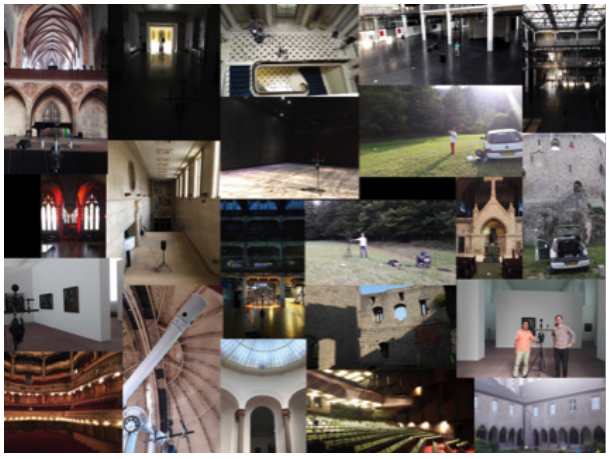
This also enables the installation of a system of request by vocal imitation based on multiples of spectral descriptors. These advances have been implemented in an interface using iPad multi-touch technology.

## HYBRID REVERBERATION AND SPATIALIZED IMPULSE RESPONSES

–

Team Involved: Acoustic and Cognitive Spaces

Most artificial reverberators use either feedback delay networks (FDN) or convolution reverb. These approaches differ in their computation cost, the "authenticity" of the results, the flexibility of use and adaptability vis-à-vis the system used for audio rendering. For several years, the Acoustic and Cognitive Spaces team has explored a "hybrid" approach. Using measured impulse responses (IR), the early reflections are reproduced by convolution while the later part of the reverberation is synthesized by a FDN that approximates the energy decay relief (EDR) of the IR. The transition from one mode to the other occurs at the so-called "mixing time" and the FDN is calibrated to guarantee the continuity of the EDR, assuring audio rendering that is perceptively indistinguishable from the original IR.
This hybrid technique has been extended to spatial room impulse responses (SRIR), measured with spherical microphone arrays (see photo) in order to recreate the spatial distribution of the room reflections.
During Pedro Garcia-Velasquez and Augustin Muller's residency in artistic research, their artistic work focused on the notion of "sound places" leading to the creation of a library of 3D impulse responses for acoustic simulation of remarkable sites. Rather than catalogue the fingerprints of concert halls with fairly homogenous acoustics, this library favors more atypical places both in terms of their acoustic signature as by their evocative power.



Assortment of remarkable sites used for 3D acoustic fingerprints.
Photo © P. Garcia-Velasquez, A. Muller

## DISTRIBUTED SPATIALIZATION

–

Teams Concerned: Acoustic and Cognitive Spaces,
Sound Music Movement Interaction

The generalization of mobile devices with motion capture or sound diffusion functions makes it possible to imagine new applications for spatialization relying not on a group of loudspeakers around the audience, but a network of telephones spread out throughout the audience. With this kind of network, we could imagine that the spatial diffusion is no longer centrally supervised but operates through propagation mechanisms. The spatial network made of mobile devices creates a sort of propagation medium; each device carries out an elementary function.

Different scenarios using distributed spatialization were tested and led to the development of a procedure for adaptive synchronization among the different nodes in the network (see the project COSIMA, p. 59).

A programming environment was developed to manage a park of mobile devices dispersed throughout a space and was used during artistic projects. Each system connects to a Web page published by a server and downloaded on to one or several devices that carry out elementary audio functions (reading a file, granular synthesis, convolution, etc.). A control interface was developed in Max making it possible to send messages to the server and pilot the network. Different uses were implemented such as the management of trajectories, echo effects, or reverberation within this virtual propagation medium.

Control interface for the Nü environment for a group of mobile devices set up throughout a space

## SOUND FIELD SYNTHESIS WITH HIGH-DENSITY LOUDSPEAKER ARRAYS

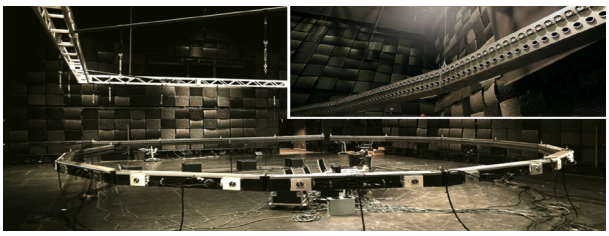
–

Team Concerned: Acoustic and Cognitive Spaces
Partner: EMPAC (Troy, United States)

Sound field synthesis with high-density loudspeaker arrays (HDLAs) remains one of the main research topics for the Acoustic and Cognitive Spaces team, which naturally finds its applications in musical production. It also represents an outstanding vehicle for the promotion of our developments and the numerous collaborations with other research institutions and prestigious musical creations.

As a part of a research collaboration with Rensselaer Polytechnic Institute, the concert hall in the Curtis R. Priem Experimental Media and Performing Arts Center (EMPAC) was equipped with a large scale HDLA. This new system consists of a linear 512-loudspeaker array (16 modules of 31 loudspeakers associated with a subwoofer) for Wave Field Synthesis (WFS) and a dome of 100 loudspeakers for Higher-Order Ambisonics (HOA). The distance in between the loudspeakers of the WFS array is 5.8 cm. Different geometric layouts were tested; each making it possible to take full advantage of different methods of sound field synthesis. For example, positioning two arrays on top of each other and offsetting them by half the on-center distance between drivers results in a horizontal distance of 2.9 cm between speakers and a total array length of 15 m. In doing so the aliasing frequency can be shifted up to around 6 kHz, which results in a striking impressive holophonic sound field reproduction. Arranging the speaker modules in a circle, on the contrary, allows for a 2D HOA reproduction of orders up to N=250.

The independently controlled loudspeakers are driven by a cluster of computers that compute the sound field synthesis in real-time. One of the major issues in this project was the implementation of multi-zone sound field control.



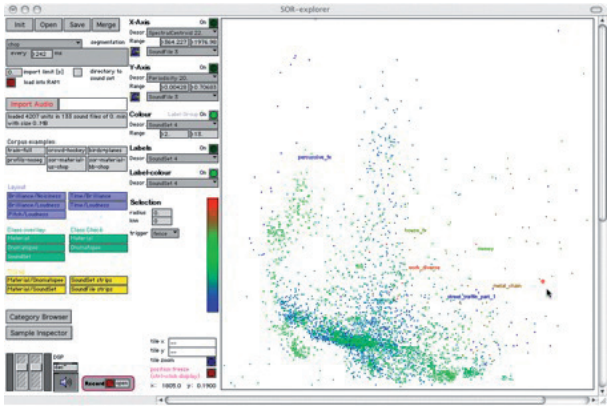
WFS – HOA system for spatial audio rendering in the Curtis R. Priem Experimental Media and Performing Arts Center at Rensselaer Polytechnic Institute de Troy (US). The system consists of 16 linear arrays of loudspeakers spaced at 5.8cm, configured here for bi-dimensional HOA diffusion.
Photo © Markus Noisternig

## CORPUS-BASED CONCATENATIVE SYNTHESIS

–

Team Involved: Sound Music Movement Interaction

Corpus-based concatenative synthesis uses a database of recorded sounds and a unit selection algorithm that chooses the segments from the database that best suit the musical sequence that we would like to synthesize by concatenation. The selection is based on the characteristics of the recording obtained through signal analysis and match, for example, the pitch, energy, or spectrum. The habitual methods for musical synthesis are based on a model of a sound signal, but it is very difficult to establish a model that conserves the entirety of the details and delicacy of the sound. However, concatenative synthesis—that uses real recordings—preserves these details.
When this new approach of concatenative synthesis by corpus is used in real-time, it enables an interactive exploration of a sound database and a granular composition that targets specific sound characteristics. It also makes it possible for composers and musicians to reach new sounds. This principle is carried out in the CataRT system. This system makes it possible to display a 2D projection of the descriptor space that can be browsed using a mouse or external controllers. Grains are then selected in the original recording and performed by geometric proximity, metronome, in loops, or continuously. It is also possible to define a perimeter around one's present position that selects a sub-group of grains that are then played randomly. CataRT is used for musical composition, performance, and in various sound installations.
As this field of research is fairly young, several interesting research questions have been raised (or will be raised in the future) concerning the analysis and exploitation of the information found in the data of a corpus, the visualization, and real-time interaction.



System of visualization used for sound synthesis by corpus

## WFS AND AMBISONIC SYSTEMS IN THE ESPACE DE PROJECTION

–

Team Involved: Acoustic and Cognitive Spaces

Funding: Programme Sesame (Conseil général d'Île-de-France), CNRS, IRCAM, ANR, Sorbonne Université

The Wave Field Synthesis (WFS) technique is a type of holophonic reproduction process that enables, by analogy with visual holograms, to capture or synthesize a sound scene while conserving the spatial characteristics of distance and direction. This approach, initiated by the University of Delft, goes well beyond the limits of conventional systems in terms of reproduction fidelity in a wide listening zone. While traditional stereophonic techniques (e.g. stereo, 5.1) are a kind of trompe-l'œil and can only be truly appreciated when one is positioned in the center of the system, the goal of holophonic technology is to reproduce a sound field in which listeners can move freely while maintaining a coherent perception of the localization of the sound sources. IRCAM has gained experience in this type of reproduction through its participation in the European CARROUSO Project. These studies have made it possible to carry out productions in partnership with the Centre Pompidou such as the production of an installation at the end of 2005 for the DADA exhibit, an interactive installation in June 2006 (in collaboration with N+N Corsino), and an installation for the Samuel Beckett exhibit in 2007.
From 2008–2011, IRCAM was able to acquire and install a WFS system in the Espace de Projection with the support of the Île-de-France Région, the CNRS, and Sorbonne Université. This system is made up of a horizontal ring of 264 loudspeakers placed at regular intervals around the stage and the audience for WFS sound diffusion and of a dome of 75 loudspeakers for three-dimensional Ambisonic sound diffusion. This equipment is used to experiment with new methods of spatialization for musical creation and for scientific experiments with virtual reality and spatial cognition. The WFS component of this installation, inaugurated in 2008, was used in the musical creations by P. Schoeller in June 2009 (*Operspective Hölderlin*), M. Jarrell in June 2010 (*Le Père*), M. Garcia-Vitoria in October 2011 (Mimesis), and in theatrical productions at the festival of Avignon (Shakespeare's *Richard II*, directed by J. B. Sastre, July 2010). The final system was inaugurated on November 28th, 2012.

## BINAURAL REPRODUCTION TECHNOLOGY

–

Team Involved: Acoustic and Cognitive Spaces

The realistic reproduction of sounds through headphones requires the use of binaural spatialization. This technique is based on a dynamic filtering of the sound source using transfer functions (HRTF - Head Related Transfer Functions) measured on the head of a listener or model. Despite the limited nature of its current application to music (except in virtual reality or interactive sound installations), binaural technology is a very important method of reproduction in the lab. Associated with systems that follow your position, this is currently the only technique that allows the complexity of a sound scene to be reproduced without artifact. This technique is particularly useful in perceptive validation work. At IRCAM it is used in the study of spatial auditory cognition including listener navigation and the perception-action process brought about by the motion control of a localized sound source.

Virtual reality experiment in which the sound contents of the scene, recreated on headphones using the binaural technique and updated in realtime according to the head movements of the participant.
© CNRS Photothèque / F. Vrignaud

## PROCESSING BY PHASE VOCODER

–

Team Involved: Sound Analysis & Synthesis

The phase vocoder, one of the most effective techniques for the analysis and transformation of sounds, represents the foundation of the SuperVP software program. With the phase vocoder, it is possible to transpose, stretch, or shorten sounds; it is possible to apply a practically limitless number of filters to sounds. By the same token, the level of sound quality of the transformed signals is extremely high when applied to speech. Numerous improvements and extensions have been introduced, for example:

- Reassigned spectrum
- Estimation of the spectral envelope via 'true envelope' transposition with the preservation of the spectral envelope transposition with the 'shape invariant' model
- Generalized cross synthesis enabling the synthesis of hybrid sounds
- Several methods for estimating the fundamental frequency (pitch) of a signal
- Classification by nature of the spectral, sinusoidal (voiced) or non-sinusoidal (non-voiced sounds or noises) peaks segmentation of the time/frequency zones into transitory and non-transitory regions and the increase or decrease of transitory sections
- Processing the sinusoidal, non-sinusoidal, and transitory time/frequency zones
- The LF model of a glottal source, making it possible to transform a voice, etc.

These different modules of analysis, synthesis, and processing are used in several software programs on the market today.

## MUSIC AND HACKING

–

Team Concerned: Analysis of Musical Practices
Calendar: September 2016-November 2017
Partner: Musée du Quai Branly — Jacques Chirac

In general, hacking can characterize a group of activities that is intended to optimize the performances of an object or technological system—the search for the best solution to a given problem—to be transgressive (circumvention of legal or technological standards, misuse of common practices, disrespectful attitude towards objects), and hedonic (pleasure in finding a solution, ingenious manipulation, technical prowess, customization). In view of this, we can consider hacking to be a series of practices and customs configured by the conceptual categories of computer-science: the computer is an "open", modular, object adaptable to the changing needs of the user. Code is a support for information, and the network a structure for communication – practices and customs that can then be reused (and remodeled) beyond the computer-science context.

The Music and Hacking project aims to find points of intersection between the hacker culture and today's music, from the most explicit manifestations (periodical organization of music hack days and other hackathons, massive pirating of music industry productions, etc.) to implicit transfers (surfacing of new conceptions of the musical instrument, questioning the function-author, ethical dimension of certain experimental music, etc.), shedding light on new musical practices.

During 2016-2017, the team analyze the activities of instrumental hacking in the "Lutheries Urbaines" workshop in Bagnolet, completed by the production of a set of interviews with important players in the improvised music scene who came to create their own performance system.

In November 2017, the project closed with the conference "Music and Hacking: Instruments, Communities, Ethics" organized by the Musée du quai Branly and IRCAM, and the organization of a music hack day after the conference (November 10 & 11 at IRCAM) gave a concrete look at the vibrancy and wealth of different approaches that make up the world of musical hacking.

## ORPHEUS

**Object Based Broadcasting**

orpheus-audio.eu

–

Team Involved: Acoustic and Cognitive Spaces

Funding: H2020-ICT-2015

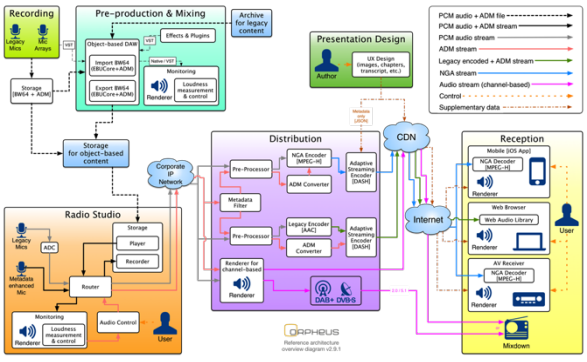
Funding Calendar: December 2015—May 2018

Partners: Fhg IIS, BBC, Bayerischer Rundfunk, IRT, b <> com, Magix, Trinnov-Audio, ElephantCandy, Eurescom

Object-based media is a promising approach for creating and deploying interactive, personalized, scalable and immersive content, by representing it as a set of individual assets (video and/or audio) together with metadata describing their temporal and spatial relationships and associations. This allows media objects to be assembled in groundbreaking ways to create new user experiences, offering various levels of interaction: adjustment of the balance between objects in the foreground and the background, spatial navigation, non-linear media usage, adaptation to listening contexts (public broadcast, home listening, headphone listening while moving, etc.).

The ORPHEUS project aims at developing and validating the whole audio production chain, from studio to listener, making it possible to create, transmit, and receive "object" contents on different platforms (audio-video receiver, smartphone, web browser). This chain relies on the Audio Definition Model (ADM) recommended and defined by several international authorities (European Broadcast Union and International Telecommunication Union).

At the heart of this research partnership, IRCAM focuses more specifically on aspects connected to the characterization, synthesis, and transmission of the effects of reverberation. IRCAM developed a software suite dedicated to recording, to reading, and rendering of ADM audio contents (see ADMix Tools in software, p. 64).




Architecture of treatment for the creation, broadcast, and reception of audio contents in object-based format

## ROUTE

**Robot à l'écoute**

–

Team Involved: Sound Analysis & Synthesis

Funding: Convergence de Sorbonne Universités — Emergence Programme 2014

Calendar: September 2015-September 2016

Partners: ISIR (Institut des systèmes intelligents et de robotique)

This project focuse on a common issue for two apparently distinct research domains:

- The automatic analysis of a sound scene, from a learning process using a dictionary. In audio signal processing, we find this aspect in several major applications for example computational auditory scene analysis (CASA), automatic indexing, source separation, detection and localization of sound objects.
- Artificial hearing, a recent field of study in robotics, for which the analysis of a sound stage gradually becomes a prerequisite for any modern application (e.g. monitoring the elderly, or studying human-robot interaction).

The central objective of the project is the design and development of a new method for the detection and localization of the main speaker in a sound scene. The method is intended to enable a robot to identify a vocal signal in the presence of noise and locate the main speaker's position, in the case where there are several speakers. The problem is closely related to the current field of computational auditory scene analysis (CASA).

The objective of CASA is to design automatic systems with a perception that mimics human hearing, considering its physical and psycho-acoustical aspects. This project takes a different approach, while the audio processing tools are comparable (machine learning, source separation), hearing is treated from the robot's point of view and the interest lies in the analysis of the audio scene.

## BILI

bili-project.org

–

Team Involved: Acoustic and Cognitive Spaces Funding: FUI
Calendar: January 2013 — June 2016
Partners: France Télévisions, Orange, Radio France, LIMSI-CNRS, Conservatoire de Paris, Arkamys, A-Volute, Trinnov-Audio

Personal devices (e.g. smart phones, tablets) now have several functions and represent one of the main means of accessing music, radio and television programs. As the public increases their use of these devices, binaural listening with headphones becomes more frequent and symbolizes the idea of a personal experience, providing access to a reproduction of 3D immersive sound scenes. However, the individual dependence of head-related transfer (HRTF) that is the basis of binaural technology has limited its dissemination for to the general public.

During the BiLi project, the software and hardware environment for measuring HRTFs in IRCAM's anechoic chamber was updated making it possible to considerably expand the spatial resolution as compared to comparable databases. The new spatial sampling (1680 directions) enables a high-order spherical harmonics decomposition, useful for diverse exploitation phases (spatial interpolation, HOA to binaural transcoding).

Sharing the HRTF database within the international community motivated the initiative to create a standard exchange format. Resulting from an international collaboration, the SOFA format (Spatially Oriented Format for Acoustics)—approved by the industry standardization authority, the Audio Engineering Society—permits storage of spatial acoustic data such as HRTFs or SRIRs. IRCAM implemented an OPenDAP server (Open-source Project for Network Data Access Protocol) that hosts a range of SOFA HRTFs databases and to which client applications (web applications, Matlab, etc.) can address download requests (a specific HRTF, a full head, etc.).

Several methods to individualize HRTFs were studied, notably estimations of HRTFs that do not require acoustic or morphological measures. These methods use databases now available to guide a user to select the most appropriate set of HRTFs. One of the approaches suggested relied on the blind analysis of binaural recordings made by the listener in unsupervised conditions (reverberant space, random signals and sources, and the listener in motion).

## CAGIMA

**Global Acoustic Conception of Reed Musical Instruments, In Tune and Homogeneous**
**http://cagima.ircam.fr**

–

Teams Involved: S3AM, Sound Analysis & Synthesis
Funding: ANR - programme blanc
Calendar: December 2011 — November 2015
Partners: Laboratoire de mécanique et d'acoustique de Marseille, Laboratoire des sciences, des procédés et des matériaux de l'Université Paris 13, and Buffet Crampon
Coordinator: LMA Marseille-CNRS

The Cagima project focused on flaws found in reed instruments in tuning, homogeneity of emitted sounds, and timbre from both the perspective of the musician and also that of the instrument-maker who endeavors to satisfy the specific demands of each musician. The project's ambition was to replace the historical incremental approach used by instrument-makers with a rational and global approach that aimed to design new "logical" instruments that minimized the identified flaws. This was a large step forward for the methodology and technology of instrument making. To begin, the constraints of sound production that are imposed by an instrument on a musician were evaluated via the measurement of a specific gesture (e.g. pressure in the mouth, pressing on the reed with a lip, configuration of the vocal tract) and were correlated to the acoustic flaws of the instruments in order to suggest novel methodologies for the general design for the bore and the lateral holes of reed instruments. For this, the study of global criterion was at the heart of this project; their understanding led to the fabrication of playable prototypes of musical instruments.

## 3DTVS

**Study of techniques for description of 3D audiovisual contents**

–

Team Involved: Sound Analysis & Synthesis
Funding: European Commission, ICT, FP7
Calendar: November 2011—January 2015
Partners: The Aristotle University of Thessaloniki (coordinator), Fraunhofer HHI, Flying Eye, Velti, Sky-Deutschland, Arte

The 3DTVs project was dedicated to the study of description techniques for the contents, automatic indexing, searches, and browsing in three-dimensional audiovisual contents available on both mobile and stationary platforms. The objective of this project was to improve the performance of the analyses carried out on the three- dimensional formation of audio and video signals. The project also aims to expand audio/video indexing through the multimodality of the approaches. The role of the joint research laboratory in this project focuses on the analysis methods for spatialized audio signals (multi-channel audio streams) in order to improve indexing and provide the localization of sound sources in three-dimensional space.

# ////////////////////// THE MUSICAL BODY

## ELEMENT

**Stimulate Movement Learning in Human-Machine Interactions**

–

Team Involved: Sound Music Movement Interaction
Funding: ANR, projet AAP
Calendar: November 2018 — October 2021
Coordinator: IRCAM
Partners: LRI, LIMSI-CNRS

The majority of interactions based on movement offer "intuitive" interfaces and trivial gesture vocabulary. While they facilitates the adoption of the system, they also limit the possibility of more complex, expressive, and truly embodied interactions.
We propose going from "intuitive" notions to notions of "learnability". Our project addresses computational problems of methodology and modelling.
Firstly, we must create methods to design movement vocabularies that will be easy to learn and compose in order to build rich and expressive phrases of movements. Secondly, we must design computational models capable of analyzing users' movements in real-time to provide diverse feedback mechanisms and multimodal guiding (for example visual and auditive).
This project raises three fundamental research issues:

1. How do we conceive movements and gestures, formed with components easy to learn while supporting techniques for complex interactions beyond simple commands?
2. How do we account for the sensory-motor learning with computational modeling of movement and interaction?
3. How do we optimize the feedback systems and computer guides in order to facilitate the acquisition of skills?

The long-term objective is to encourage innovation in multimodal interaction, from non-verbal communication to interaction with digital medias in creative applications.

## THEVOICE

**Voice design for creative industries**

–

Team Involved: Sound Analysis & Synthesis
Funding: ANR, programme Société de l'information et de la communication (DS07)
Calendar: January 2018—June 2021
Partners: LIA (Laboratoire d'informatique d'Avignon), Dubbing Brothers

TheVoice addresses voice design for the production of contents in the field of creative industries (films, series, documentaries). This is a particularly important field in terms of industrial possibilities while extremely demanding in terms of quality. The project stems from a simple observation: the production of voices is carried out exclusively by humans in a field that is practically entirely digital. The scientific and technological objectives of the project are to model the "vocal palette" of actors in order to create recommendations of voices by similarity as well as the creation of artificial voices capable of reproducing the vocal identity of an actor. The project will change today's uses through the production and industrialization of new technologies for the creation of natural and expressive vocal contents. The consortium, supported by a major player in the industry of creation of digital contents and made up of renowned research laboratories, aims to consolidate its position of excellence in digital research and technologies "made in France" as well as the promotion of French culture abroad.

## REFLETS

**Facial and Linguistic Emotional Retroaction in Traumatic Stress States**
**ANR Project "AAP GENERIQUE 2017"**

–

Team Involved: Sound Perception and Design
Funding: ANR, programme blanc "Technologies pour la santé"
Calendar: October 2017 — September 2021
Partners: CentraleSupélec (Rennes, France — coordinator), UMR 8257-Cognac-G (CNRS/Paris-Descartes/Service de Santé des Armées), Hôpital militaire Percy et Institut de recherche biomédicale des Armées (France), Dynamixyz (France), HumanEvo (France), Chanel (France)

The project REFLETS (Rétroaction Émotionnelle Faciale et Linguistique et États de Stress Traumatique) focuses on improving the care of those suffering from post-traumatic stress syndrome (PTSD) using a technological system that reacts with the sufferer's capacity to perceive and regulate their emotions. The project is based on technologies recently developed at IRCAM on the manipulation, in real-time, of indicators of smiles in the spoken voice and at CentralSupélec on the video manipulation of visual cues of a smile on a face.

The project focuses on three different disciplines. In the domain of information sciences, the project aims to develop "emotional mirror" technology in which the participant can see and hear themselves, but their emotional tone is manipulated algorithmically (in practice, participants will see and hear themselves with a smile while they express themselves using a neutral tone). In the domains of psychology and cognitive neurosciences, REFLECTS aims at studying mechanisms for perception of oneself and metacognition involved in this paradigm of false feedback. Preliminary studies have already established that hearing oneself speak with a more joyful tone has a positive effect on the speaker's emotions (Aucouturier et al., PNAS 2016). Finally, from a clinical point of view, the project intends to test the therapeutic impact of this type of system to treat patients suffering from deficiencies in self-perception and emotional control (alexithymia) in post-traumatic troubles. The project will lead to clinical trials on a group of PTSD patients selected in collaboration with the Percy Military Hospital.

## ENTRECORPS

–

Team Involved: Acoustic and Cognitive Spaces
Funding: Sorbonne Université — Programme Emergence
Calendar: September 2017-June 2018

Human interactions have been studied for several years by disciplines such as engineering and robotics, due to the growing impact of machines in communication among humans. Concurrently, neuroscience has begun to address the inclusion of humans in society and realize that it is difficult to establish social cognition neuronal mechanisms by studying isolated individuals. While we have long known that social context controls behavioral responses to sensorial stimuli, taking this factor into account is not common practice in the experimental protocol adopted in human behavioral neurosciences. With the introduction of new forms of communication, human interactions are made public and undoubtedly transformed.
This project aims to cross-reference knowledge from several disciplines to study the mechanisms of interpersonal coordination in public human interactions. Taking advantage of new collaborative applications, we would like to use them as experimental tools to provide knowledge to cognitive and social neurosciences. Inversely, the results of our experiments should be used to assess collaborative applications, questioning the sensory mechanisms they reveal as well as their social and artistic consequences.
The first example is that of personal space, studied this time in a situation of interaction. Over 50 years ago, the anthropologist Edward Hall created the idea of proxemics, defining different levels of territory around the body and their social function. Recently, this work was associated with research in neurophysiology establishing the neuronal standards for personal space. This personal space is coded on a cerebral level like a motor and multisensory interface between body and environment. Several studies have revealed the plasticity of the limits of this virtual space around the subject's body depending on the social and emotional situation in which she is placed.
A study was developed to investigate the impact of different social contexts on the size of personal space, in particular in contexts of either collaboration or competition between subjects. These studies confirmed the specificity of collaboration as a social context that controls the volume of personal space.

## MIM

### Enhancing Motion Interaction through Music Performance

–

Team Involved: Sound Music Movement Interaction
Funding: Program H2020-Marie Sklodowska-Curie Actions
Calendar: January 2016—December 2018
Partner: McGill University, Montreal Canada

This project focuses on Human-Computer interactions based on movement, leveraging a multidisciplinary approach between experimental psychology, music technology, and computational modeling. Initially, the project will look at sensorimotor learning mechanisms and expressive control in human movement. Computational models of these mechanisms will be developed based on experimental data gathered from the performers' movements. Then, the models developed will be applied to the domain of Digital Musical Instruments (DMI), creating new types of instruments based on sensorimotor learning mechanisms. The project contributes to two fairly uncharted research areas. Firstly, it contributes to the fundamental understanding of sensorimotor learning processes by considering complex human motion such as the movements of musicians. Secondly, it represents the development and assessment of unique interactive musical systems using computational models of expressive musical gestures.

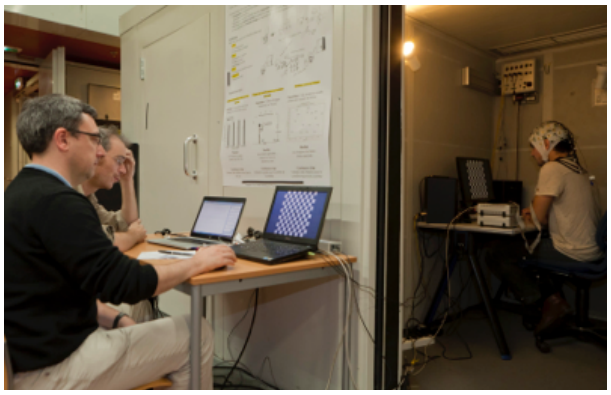
## CREAM

### Cracking the Emotional Code of Music

–

Team Involved: Sound Perception and Design
Funding: ERC Starting Grant 335536
Calendar: October 2014—September 2019
Partner: Centre de Recherche de l'Institut du cerveau et de la moelle épinière (CRICM UMR7225)

The CREAM project's objective is to produce technology and knowledge that make it possible to characterize which type of musical signal sets off which type of brain mechanism of emotional induction.
Until now, research in musical cognition has focused on correlations between relatively indistinct emotional reactions and loosely controlled musical stimuli: we know that music creates emotions, but we do not know how. The CREAM project suggests combining current neuroscience methods with a high technical level of signal processing to create, for the first time, musical stimuli capable of selectively causing or inhibiting certain cortical circuits involved in emotional processing so they may be studied in isolation. For example, we will suggest targeting centers of interpretation of prosody and speech by constructing musical sounds that "tremble" like an anxious voice or that "rejoice" like a happy voice.
These new experimental control techniques make it possible to expand our current understand of brain mechanisms of emotional induction, but also to conceive several clinical applications for therapy or diagnosis of depression or neurodegenerative illnesses. In other words, the CREAM project will turn music into a real clinical technology capable of stimulating specific neuronal circuits in a non-intrusive and non-pharmacological manner.



EEG measurements in an audiometric room at IRCAM for an auditory test

## GESTURE ANALYSIS AND RECOGNITION

–

Teams Involved: Sound Music Movement Interaction, Creation and Production departement

This research project replies to an increased interest for interactive musical systems based on gestural control. The applications concern not only music, but also performances such as dance or theater. The research carried out in the framework of this project is multi-disciplinary and includes engineering, physiology and biomechanics, cognitive sciences, and artistic domains. This work is carried out in synergy with the team's developments on gestural interfaces.

The first line of research focuses on the study of instrumental gesture and its relationship with both musical writing and the characteristics of the sound signal. Diverse methods of motion capture are used—from optical 3D measurement to wireless sensors (accelerometers, gyroscope). These methods allow us to measure and model musicians' movements, as in the case of piano strings.

Diverse issues are also addressed in this study: motor control, learning in the case of the gesture of an expert, characterization of playing styles taking into account sound and gesture parameters, and the modeling of the phenomena connected to the gestural co-articulation similar to those of speech.

Gesture capture and analysis
(© photo Hillary Giodell)

## ANIMAGLOTTE

**Artificial System of Ex Vivo Animation of the Larynx**

–

Team Involved: Sound Systems and Signals: Audio/Acoustics, InstruMents
Calendar: 2018-2019
Partners: LPL-CNRS (coordinator), LMA-CNRS, CHU La Conception-Marseille, CHU-Liège

The human voice has an extremely variable range of acoustic characteristics depending on whether it is a whispered, sung, spoken or pathologically effected voice. To produce these different types of voices, humans use the length of the vocal folds and their level of contact (delineating the glottis).
Our project develops a mechanized system of ex vivo animation for human larynxes excised post-mortem. Following measures taken in vivo of the dynamic (amplitude and speed) of the glottis' movements, we aim to reproduce them via an artificial mechanical controller.

## AUDITORY SPATIAL COGNITION

–

Team Involved: Acoustic and Cognitive Spaces

In the field of sound reproduction and communication, future audio technology will attempt to shift emphasis towards sensations of immersion and presence. Such notions are intimately linked to the spatial dimensions of a multimedia scene, and particularly in the field of sound, and are intensified in situations involving the participation of the listener. This participation may involve navigation within a scene or the gestural interaction of objects within it.

Under these conditions, made possible by binaural and holophonic technologies, the congruence and real-time updating of auditory spatial clues in accordance with the listener's movement or actions have a major impact on the sensation of presence. This new context led to the development of a set of experiments focusing on auditory spatial cognition, notably via the study of multi-sensorial integration processes, focusing on auditory and idiothetical modalities (clues induced by the subject's movements including balance and proprioception). Experimental methods based on the observation of a subject's performance, in terms of localization and navigation, when submitted to different exploratory contexts.

In this line of research, we are also interested in the relationships between multisensorial integration and emotional dimension. This could concern, for example, the effects of spatial conflicts between sound and vision on the subject's emotional reaction (see multisensorial integration and emotion, on the right), or the evaluation of the perception of numerosity (e.g. quantification of a crowd) in accordance with the sensorial modality and its connections with emotion.

## MULTISENSORIAL INTEGRATION AND EMOTION

–

Team Involved: Acoustic and Cognitive Spaces

Over the past 20 years, new types of trauma therapy using virtual reality technologies have been developed for the treatment and rehabilitation of emotional disorders, particularly of phobias. Virtual reality (VR) offers in-depth control of the presentation of sensory and spatial simulations, providing better control of the emotional impact of anxiety-inducing situations to which the patient is exposed. However, in order to take full advantage of VR, it is necessary to deepen our understanding of the connections between the characteristics of a stimulus that is presented to the patient and its emotional effect.

We have explored these connections in the framework of virtual crowds. The fear of crowds is a symptom found in several emotional troubles (agoraphobia, social phobias, fear of falling, etc.) and presents components that are both sensory (auditive and visual) and spatial, making a crowd the ideal stimulus for our studies.

A study carried out on two groups of participants (sensitive or not to the fear of crowds), immersed in a visual-auditory environment with virtual crowds. Participants had to indicate the intensity of their discomfort. The sensory presentation of the crowd was either visual, auditory, or both. The results indicate that the visual-auditory presentation amplifies negative feelings in participants sensitive to this fear of crowds. However, this effect is only observed when the crowd is close to the participant. This result reveals the interaction between spatial and sensory characteristics of the anxiety-inducing stimulus and the emotional impact.

This spatial dependency motivated research dedicated to the study of personal space. Humans do not perceive space in a homogenous manner: the brain codes the space close to the body differently from space further away. Personal space (close to the body), is coded by multisensorial neurons. We are trying to bring to light the plasticity of the limits of this space depending on the nature of sound and the emotional or social context.

VR system in IRCAM's studio 4. The visual stereo is projected on a large screen. Visual and auditory information is updated using a position-following system.

## ONDES MARTENOT

–

Team Involved: Sound Systems and Signals: Audio/Acoustics, InstruMents
Funding: Collegium Musicæ
Calendar: January—December 2018
Partner: ECR-Musée de la musique

This project has a dual focus: museums and science
The objective for museums in this project concerns the automatic generation of documentation for instruments and simulation code for "non-invasive listening"—important for preserving playable virtual forms—in addition to electronic cloning of ancient technological components as programmable analog components that can be inserted in circuits, an important factor for the preservation of instruments in a non-virtual form.
The scientific objective focuses on physical models of old non-linear electronic components, simulation in real-time with guaranteed passivity of electronic circuits, and automatic generation of documentation and simulation code (work based on Port-Hamiltonien systems, a very productive formalism for multi-physical systems). It also focuses on the conception of programmable analogical electronic components and the creation of a general methodology applicable to ondes Martenot and other families of analogical audio circuits.

## AUGMENTED PAPER FOR CAC

–

Teams Involved: Musical Representations, Analysis of Musical Practices
Partners: LRI Orsay, INRIA

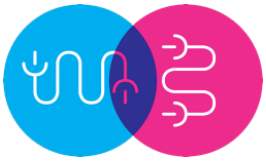
Through the technology present in augmented paper that contains a motif visible only to the small camera inserted in a special pen, the computer can detect what the composer is writing. This 'before' connection between paper and computer is completed by an 'after' connection from the computer to paper by printing on augmented paper, therefore creating an entirely new production cycle.
The goal of this project, extended upon request by the ANR, is to work in collaboration with artists and research workers to identify the possible musical applications that can come from the link between the computer support and augmented paper. Rather than replace paper, with its recognized properties of flexibility, comfort, and free association our approach consists of developing applications that combine the advantages of paper and electronic documents.
In the long run, we would like to develop prototypes of applications in three core fields:

- Notation
  In the field of musical composition, we are interested in both notation and annotation as both are complementary in their reflective nature for musical creation.
- Writing Synthesis
  Due to its flexibility of use and the subtlety of handwritten interaction compared to interaction via a mouse or even with a graphical tablet, paper remains the preferred support for the informal expression of artistic gestural ideas, and especially for writing down ideas for synthesis. In collaboration with the OpenMusic software program, this technology facilitates a natural evolution of various parameters for synthesis.
- Teaching
  The technology present in augmented paper is exceptionally interesting for musicologists and music teachers. By connecting a simple support, the printed score, to the MusiqueLab 2 software program, it is possible to work collaboratively and create new forms of interaction.

© Inria / Photo H. Raguet
Interfaced used for the premiere of *Quid sit musicus?* by Philippe Leroux in June 2014. This interface associates electronic musical processes with the calligraphy of Guillaume de Machaut's score.

## MUSICBRICKS

**Musical Building Blocks for Digital Makers and Content Creators**



–

Team Involved: Sound Music Movement Interaction
Funding: Program H2020- ICT18-2014-a
Calendar: January 2015—June 2016
Partners: Sigma-Orionis, Stromatolite Innovation Lab, Universitat Pompeu Fabra, Fraunhofer Institute IDMT, Technische Universität Wien
Coordinator: Sigma-Orionis

The goal of the European project MusicBricks is to facilitate the transfer of new musical technologies from major European research centers specialized in the domain to small digital creation companies. The project involves the creation of programming interfaces, of graphic and tangible user interfaces, the development of an ecosystem based on events during Music Tech Fest and in selected technology incubators to prepare access to the market. In this project, The joint research laboratory is in charge of the diffusion of technologies connected to tangible and graphic user interfaces (TUI and GUI). A particular focus of the project is to develop analysis modules of gestural data and mapping between movements and sounds.

## CHANTER

**Real-time controlled Digital singing**

–

Team Involved: Sound Analysis & Synthesis
Funding: ANR — programme Contenus et interactions
Calendar: January 2014—June 2017
Partners: LIMSI, Acapela, Dualo
Coordinator: LIMSI

The goal of this project is to create a high quality system for synthesizing song that can be used by the general public. The system must not only sing vowels, but must also say the words of a song. This type of system does not currently exist for the French language. The synthesizer imagined will work in two modes. The first is "song from text" mode where the user can enter a text to be sung along with a score (times and pitches), and the machine will transform it into sound. The second is the "virtual singer" mode in which the user controls the song synthesizer in real-time via specific interfaces; just like playing an instrument.

To create these tools, we suggest a combination of advanced voice transformation technologies, based on our expertise in synthesis by selection of units and in synthesis of singing by rules, and innovative interfaces for gestural control. The project focuses on capturing and reproducing a variety of vocal styles (e.g. lyrical/classical, popular/song). Beyond the quality tests ordinarily used for vocal synthesis the usability of the systems will be tested from the outset, paying particular attention to the creative aspects they allow (assessment through mini-concerts and mini composition projects that use the new control gestures, a virtual chorus, and/or a virtual soloist). Project partners will use the prototype of song synthesis that will be developed during the project to offer synthesized singing voice and singing instrument products that are currently lacking, or to improve the functions of currently existing products. The project will offer musicians and performers a new artistic approach to synthesized song, new means of creation that make interactive experiences with a sung voice possible.

## SKAT-VG

**Sketching Audio Technologies using Vocalizations and Gestures**

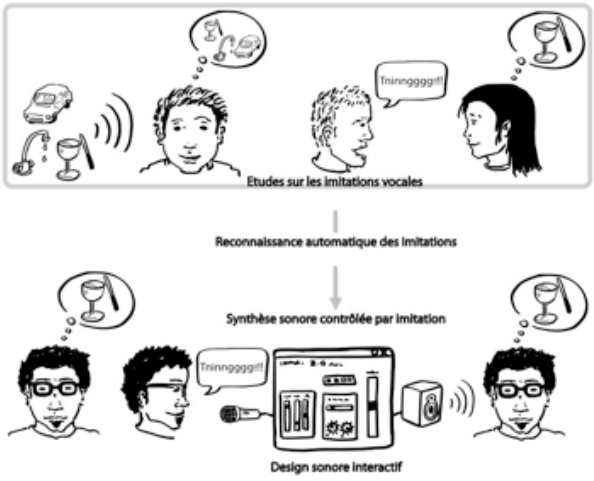
—

Teams Involved: Sound Perception and Design, Sound Music Movement Interaction, Sound Analysis & Synthesis
Calendar: January 2014—January 2017
Funding: Objective ICT-2011-C, FET-Open: Challenging current thinking FP7, European Commission
Partners: IUAV (Universita Iuav di Venezia), KTH (Kungliga Tekniska Högskolan), GENESIS (Genesis Acoustics)
Coordinator: IUAV (Universita Iuav di Venezia)

Sketching is at the root of any design activity. In visual design, hand and pencil are still the primary tools used to produce a large variety of initial concepts in a very short time. However, in product and media design the sonic behavior of objects is also of primary importance, as sounds may afford seamless and aesthetically pleasing interactions. But how might one sketch the auditory aspects and sonic behavior of objects, in the early stages of the design process? Non-verbal sounds, more than speech, are naturally and spontaneously used in everyday life to describe and imitate sonic events, often accompanied by manual expressive gestures that complement, qualify, or emphasize them.
The SkAT-VG project aims at enabling designers to use their voice and hands, directly, to sketch the auditory aspects of an object, thereby making it easier to exploit the functional and aesthetic possibilities of sound. The core of this framework is a system able to interpret users' intentions through gestures and vocalizations, to select appropriate sound synthesis modules, and to enable iterative refinement and sharing, as it is commonly done with drawn sketches in the early stages of the design process.
To reach its goal, the SkAT-VG project is based on an original mixture of complementary expertise: voice production, gesture analysis, cognitive psychology, machine learning, interaction design, and audio application development. The project tasks included case studies of how people naturally use vocalizations and gestures to communicate sounds, evaluation of current practices of sound designers, basic studies of sound identification trough vocalizations and gestural production, gesture analysis and machine learning, and development of the sketching tools.

The project has several scientific goals:

- Expand existing knowledge in the domain of perception and production of vocal imitations and expressive gestures
- Develop algorithms for automatic classification of vocal and gestural imitations, based on imitation itself, combining the analysis of a sound signal with the physiological mechanisms of vocal production
- Explore the effectiveness of sound and gestural sketches in the domain of interactive sound design, by taking advantage of automatic classification for selection and parameters of sound synthesis models
- Develop applications for intuitive sound creation and processing using voice and gesture



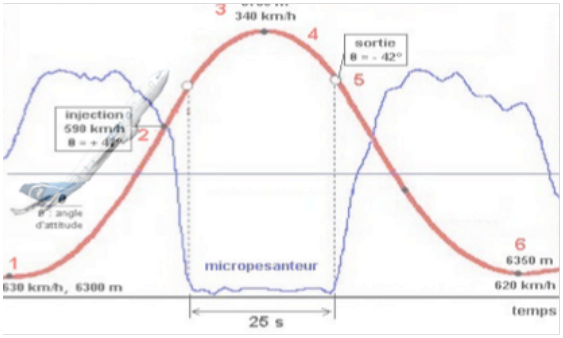

Sketch of the synopsis of the SkAT-VG project. (© O. Houix).

## AUDIOSELF

–

Team Involved: Acoustic and Cognitive Spaces
Funding: CNES
Calendar: 2013—2015
Partners: CNES, Novespace

Our perception is based on the impression that our "self" is located in the same position as our bodies; in other words, the feeling of embodiment, of living in a body, of being located within the physical confines of a body. This feeling is not governed by an isolated sensory organ, but depends on multisensory afferences. We study this feeling by looking at the contribution of audio and vestibular interactions in the perception of our own body, using zero gravity conditions created during parabolic flights. In zero gravity vestibular cues are naturally troubled and this may modify the feeling of unity between the feeling of "self" and one's own body. In other words, this may modify the perception of your body as a spatial reference.



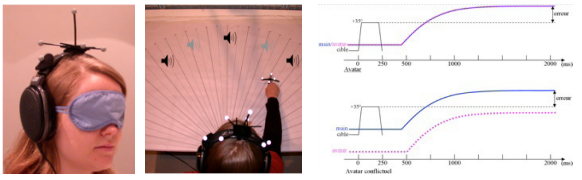

Experimental protocol in a microgravity situation. On the left: (blue line) changes in gravity, during the airplane's trajectory (red line). On the right: the subject is equipped with headphones and a reply box to indicate the spatial provenance of the binaural stimuli diffusion. The straps make it possible for the subject to float without drifting away in the aircraft.

## VERVE

–

Team Involved: Acoustic and Cognitive Spaces
Funding: FP7 ICT Call 7 iCt-2011.5.5 ICT for Smart and Personalized Inclusion, European Commission
Calendar: October 2011—September 2014
Partners: Trinity College Dublin (Ireland), Centre hospitalier universitaire de Nice, INRIA Sophia- Antipolis, CNRS Télécom Paristech (France), Testaluna (Italy), Kainos (United Kingdom), Univeristy of Zaragoza (Spain), and Deutsches Forschungszentrum für Künstliche Intelligenz (Germany)
Coordinator: Trinity College Dublin

The VERVE project aimed to improve the quality of life of people at risk of social exclusion due to the fear and apathy associated with ageing and neurological troubles. The consortium works on the specification, development, and testing personalized and populated virtual environments that can be used on different types of platforms from immersive rooms (CAVE) to smart phones. VERVE's efforts focused on three situations: fear of falling and Parkinson's disease; apathy related to cognitive decline and behavioral disturbances, in particular due to Alzheimer's Disease; and other emotional disturbances linked to anxiety. Although focusing on these areas initially, it is expected that the results of the research will be applicable to a much wider range of situations.
3D sound is used in immersive situations to study the connections between visual-audio integration and emotion. The understanding of the emotional impact of visual and audio rendering of a virtual environment, from an emotional perception point of view, must enable the implementation of new therapeutic strategies for troubles linked to anxiety.



Visionless pointing task to virtual sound sources. The patient is blindfolded. A sound feedback of the hand position is played back on headphones, either aligned or in conflict with the hand movements. Sound signal is spatialized in binaural mode and in real-time according to the head orientation and to the hand position.

## LEGOS

**Sensorimotor learning in gesture-based interactive sound systems**

http://legos.ircam.fr

legos project

Sensori-motor learning in gesture-sound interactive systems

–

Teams Involved: Sound Music Movement Interaction, Sound Perception and Design, Instrumental Acoustics, Acoustic and Cognitive Spaces

Funding: ANR — programme blanc

Calendar: November 2010—March 2015

Partners: Laboratoire de Neurophysique et Physiologie (UMR 8119), Université Paris Descartes

Coordinator: IRCAM

The central idea of the LEGOS project was to foster interdisciplinary expertise in sonic gesture control technologies with neurosciences, especially in regards to sensori-motor learning. We believe that these aspects are not sufficiently taken into account in the development of interactive sound systems. A better understanding of the sensori-motor learning mechanisms of the gesture / sound coupling is necessary to provide efficient methodologies for their evaluation and optimization. Such advances would significantly expand the usability of today's gesture-based interactive sound systems, often developed empirically.

The objective of the LEGOS project is to study systematically the coupling quality in gesture-sound systems using gestural interfaces. For this, we will evaluate the sensori-motor learning, and particularly its evolution over time, in various interactive devices. The aims are therefore to develop, evaluate, and compare interactive systems, with the mid-term goal of offering renewed paradigms for gestural interfaces control of digital media (including video games and sound design), as well as prototypes for medical applications such as rehabilitation.

The project will use extensively an experimental approach, considering these three perspectives:

- Sound Control: The first point corresponds to a case of sensori-motor learning where the goal is to produce a given sound through the manipulation of a gestural interface, as in the case of digital musical instruments. The sensori-motor learning is assessed in terms of the quality of the sound production.
- Learning Gesture with Audio Feedback: This second point corresponds to the sensori-motor learning where the goal is to make a gesture guided by an audio feedback. The sensori-motor learning in this case is assessed in terms of the gesture repeatability.
- Interactive Sound Design: This third point corresponds to sensori-motor learning in the case of tangible interfaces, where the goal is the proper handling of an object. The sensori-motor learning in this case is assessed through the quality of the objet manipulation for a given use.

# ////////////////////// CREATIVE DYNAMICS

## RASPUTIN

**Simulation of architectural acoustics for a better spatial understanding using immersive navigation in real-time**

–

Team Involved: Acoustic and Cognitive Spaces
Funding: ANR, programme AAP 2018
Calendar: November 2018 — April 2022
Partners: Institut Jean Le Rond d'Alembert (coordinator), Sorbonne Université, IJLRA-CNRS, Centre de Psychiatrie et Neurosciences (CPN UMR 894 INSERM-Université Paris Descartes), Novelab, StreetLab

RASPUTIN is a collaborative fundamental research project (PRCE) at the intersection of "digital sciences and technologies" and "psychology", that aims at reducing the cognitive complexity of navigation for the blind in a new interior environment through digital simulations and explorations using virtual reality with 3D audio. These explorations are also exercises that make it possible to mentally create cognitive maps of these environments not yet physically explored. The objective is to promote access to information by anyone from anywhere.
The RASPUTIN project targets the optimization of human-machine interaction through the development of auditory virtual environments that improve training and learning of spatial configurations for people with disabilities, as well as access to and understanding of the spatial architecture of public spaces.

## MICA

**Musical Improvisation and Collective Action**

–

Teams Involved: Analysis of Musical Practices, Sound Perception and Design, Sound Music Movement Interaction
Funding: ANR, programme JCJC
Calendar: March 2018 — February 2021
Partners: Sorbonne Université, EHESS
Coordinator: Ircam

The goal of the MICA project (ANR JCJC 2018-2021) is to look at collective action through the lens of musical improvisation, particularly in situations of so-called "free" improvisation. The project endeavors to understand :

1. How operational modes vary between the demands of coordination—found in situations of collective action—and the demands of innovation found solely in creative contexts
2. How the form of the coordination can be used as a support for communication among operators during collective action
3. How, during an action, the phenomena of low-level coordination (essentially motor) and the phenomena of high-level social cognition (attributions of intentions towards other operators, etc.) that seem necessary to put in place a collective action that is, in itself, simple.

To begin, the team working on the MICA project will collect complete and detail descriptions of collective improvised actions as well as the complex and spontaneous processes of coordination used in this type of action. This will be carried out through several longitudinal ethnographic studies of different improvisational collectives and the "Classe d'Improvisation Générative" at the Paris Conservatory. The crossover study of these various fields makes it possible to observe how the collective modus operandi is modified when the conditions of coordination among musicians changes.
This ethnographic work will create the foundation on which we can raise issues typically investigated on conjoint action, addressing the question of emerging coordination differently from the paradigm of synchronization. The next step is putting in place a series of experimental protocols, found in experimental psychology and social cognition, with the goal of shedding light on certain fundamental aspects of improvised didactic interaction. We also envisage protocols that make it possible to address this issue in three different time-frames: studying strategies used by agents to coordinate, almost instantaneously, their individual actions with optimal fluidity; the analysis of material, including indications of coordination and resources, is continuously integrated in the performance time permitting communication of complex intentions; highlighting markers of team cognition that have been appropriated after long periods of collaboration.

## EMODEMOS

**Musical interactions, empathy and development. Musical learning in the Demos project through the prism of cognitive sciences.**

–

Team Involved: Sound Music Movement Interaction

Funding: Philharmonie

Calendar: January 2018 — June 2019

Partners: Centre Interfacultaire des Sciences Affectives -Université de Genève (coordinator), Université de Gênes, Philharmonie de Paris

The primary objective of the EmoDemos project is to understand and make understood how orchestral practice proposed by DEMOS:

a) Reinforces technical skills—instrumental and sensorimotor acquired in individual courses
b) Develops all essential executive functions to regulate a child's behavior
c) This collective practice encourages the surfacing of soft skills—cognitive and emotional—central to the establishment of establishing relationships and being a part of a group.

Special attention will be paid to the phenomena of emotions, particularly empathy and altruistic behavior (ex. going toward others, listening, helping). The methods put into place during this study attempt to characterize cognitive and emotional development during the DEMOS training years as well as the following period.

## ALCOLL

–

Team Involved: Analysis of Musical Practices

Funding: Sorbonne Université, programme Emergences

Calendar: September 2017 — July 2018

Partners: IReMus

The project "Analyze the Collective in the Process of Musical Creation", ALCOLL, aims at understanding the creative process in music beyond the supposed opposition between the individual and the collective (and between musicology and the humanities and social sciences). Who is who in a given musical work? What is unique and what is general? What is specific to participation or to a creative action? This is what the ALCOLL project endeavors to measure by reexamining data by observing situations of musical creation (composition, improvisation, performance), the organization of seminars, and the realization of a ethnographical survey.

## MAKIMONO

–

Team Involved: Musical Representations
Funding: ANR, programme DS07
Calendar: 2017-2019
Coordinator: IRCAM
Partners: École de musique Schulich

Musical orchestration is the subtle art of writing musical pieces for orchestra, by combining instruments to achieve a particular sonic goal. Orchestration has been transmitted empirically and a true scientific theory of orchestration has never emerged.
This project aims to create the first partnership towards the long-term goal of a true scientific theory of orchestration by coalescing the domains of computer science, artificial intelligence, experimental psychology, digital signal processing, computational audition, and music analysis. To achieve this aim, the project will exploit a large number of orchestral pieces in digital form for both symbolic scores and multi-track acoustic renderings of independent instrumental tracks. Orchestral excerpts are currently being annotated by panels of experts, in terms of the occurrence of given perceptual orchestral effects. This library of orchestral knowledge, readily available in both symbolic and signal formats for data mining and deep learning approaches. Our objective is to evaluate the optimal representations for symbolic scores and audio recordings of orchestration, by assessing their predictive capabilities on given perceptual effects.
Then, we will develop novel models of learning and knowledge extraction capable of link musical signals, symbolic scores, and perceptual analyses by targeting multimodal embedding systems (transforming multiple sources of information into a unified coordinate system). These spaces can provide metric relationships between modalities that can be exploited for both automatic generation and knowledge extraction. The results from the models will then feed back to and be validated through extensive perceptual studies. By closing the loop between perceptual effects and learning, while validating the higher-level knowledge that will be extracted, this project will revolutionize creative approaches to orchestration and its pedagogy. The predicted outputs include the development of technological tools for the automatic analysis of musical scores, for predicting the perceptual results of combining multiple musical sources, as well as the development of digital media environments for orchestration pedagogy, computer-aided orchestration and instrumental performance in simulated ensembles.
This project implicates an international partnership with the Candaian CRSNG.

## VERTIGO

http://starts.eu/vertigo

–

Departments Involved: IMR, General Direction, Research/Creation Interfaces, IRCAM Resource Center, Communication
Funding: European Commission, H2020 Program, Call ICT36-2016
Calendar: December 2016—May 2020
Partners: Artshare (Belgium), Fraunhofer ICT Group (Germany), Inova+ (Portugal), Culture Tech (France), École polytechnique fédérale de Lausanne (Suisse), Libelium Comunicaciones (Spain)
Coordinator: IRCAM

The Vertigo project was written and selected in response to the first call during the European Commission's H2020 ICT program (Information and Communication Technologies) connected to the STARTS initiative (Science, Technology and the Arts). STARTS encourages synergies between artists and R&D projects in the field of ICT supporting innovation and Vertigo is committed to coordinate a range of actions carried out in this regard throughout Europe focusing on three areas:

- The organization of artistic residencies assessed by an international and multidisciplinary jury. A budget of 900 K€ is dedicated to supporting 3 annual cycles for a total of 45 artistic residency projects that focus on creating, via the production of unique artistic works, prototypes for innovative products using the technologies developed.
- The public presentation of different actions carried out during the new, annual event at the Centre Pompidou: Mutations / Créations. This event features exhibitions, symposia, and artistic events and aims to expose and debate evolutions in artistic practices within the contemporary scientific and technological ecosystem. The theme of the inaugural edition in March 2017, presented in connection with the exhibitions "Imprimer le monde (Printing the World)" and "Ross Lovegrove", is "Simulated Space or Digital Forms".
- The development of a communal online platform enabling different parties (scientists, artists, manufacturers, business incubators, investors) to contact each other and offer support for their actions (communication, organization of specific artistic residency programs, etc.). In addition to the consortium partners, the project oversees an international network of 20 correspondents in the cultural field, including Zentrum für Kunst und Media de Karslruhe, The Victoria and Albert Museum in London, the Venice Biennale, the festival Ars Electronica in Linz, etc.

## ABC_DJ

**Artist-to-Business-to-Business-to-Consumer Audio Branding System**

abcdj.eu

–

Team Involved: Sound Analysis & Synthesis
Funding: H2020 ICT19
Calendar: January 2016 — December 2018
Partners: Technische Universität Berlin, HearDis! Corporate Sound GmbH, FINCONS S.p.A., Lovemonk S.L., INTEGRAL Markt- und Meinungsforschungsges.m.b.H., Fratelli Piacenza S.p.A.
Coordinator: Technische Universität Berlin

The ABC_DJ project seeks to provide advanced technology to European creative agencies in the field of audio branding—the selection and diffusion of recorded music in relationship with the characteristics of products and brands (boutiques, communication, etc.). The primary goal is to enable these agencies to produce high-level contents so they can successfully compete with major actors in the market. The second goal is to open the audio branding value chain to European artists and independent recording labels by offering them new means of monetizing their works, notably through in-store music. The range of music that can be used by client agencies and brands will be considerably extended.
The scientific and technical approach employed in this project brings together research in engineering knowledge, sound signal processing, user studies, and market research to assist the creative process and the automatic broadcasting of music.
In this project, IRCAM is particularly involved in the creation of new algorithms for the analysis of audio musical contents (auto tagging by genre, emotion, instrumentation, estimation of the tonality and tempo) as well as new tools for automatic mixing (measuring the audio quality, segmentation, a complete hierarchic analysis of the structure, intelligent audio summaries, separation of audio sources).

## DYCI2

**Creative Dynamics for Improvised Interaction**

–

Team Involved: Musical Representations
Funding: ANR -programme générique
Calendar: October 2014—September 2018
Partners: Inria Centre Nancy-Grand Est, Université de La Rochelle

The project on creative dynamics for improvised interaction focuses on the creation, adaptation, and implementation of effective models for artificial listening, learning, and interaction as well as the automatic generation of musical contents enabling the creation of digital musical avatars that are creative and either capable of being incorporated in an interactive and artistically convincing manner within a range of human systems such as live performances, (post) production, or education or to contribute perceptive and communicative skills to cyber-physical systems. The project highlights improvised interaction, both as an anthropological and cognitive model of action and decision, as a schema of discovery and unsupervised learning and as a discursive tool for human– digital artifact interaction with the aim of modeling style and interaction. The objective is to produce creative agents that become autonomous through direct learning resulting from contact with live performances by human improvising musicians, creating a loop of stylistic retroaction via the simultaneous exposition of humans to the productions of digital artifacts that improvised themselves. This creates a situation of human-artifact communication that evolves in a complex dynamic of co-creativity. Off-line learning with archives can also be anticipated to systematically "color" the digital individuality of the agents or to situate the experience within different genres (jazz, classic, pop, etc.). The live performance situation could be extended to novel applications such as interaction with audio-visual archives dynamically resuscitated in artistic or educations scenarios of co-improvisation, as well as in the general situation of new narrative forms and virtual reality.
A collaboration with the EPFL and the Montreaux Jazz Festival provides access to the festival's audiovisual collection, an archive now inscribed in the UNESCO's Memory of the World register.
The goal is also to constitute procedural knowledge of music through this interaction and to produce a rich instantaneous human-digital experience, likely to provide aesthetic satisfaction for the user, to enrich his sound and musical production, to implement a dialog with him, to imitate or contradict him, and in general, to stimulate and revitalize the experience of collective performance.

## COMPUTER-ASSISTED COMPOSTION (OPENMUSIC)

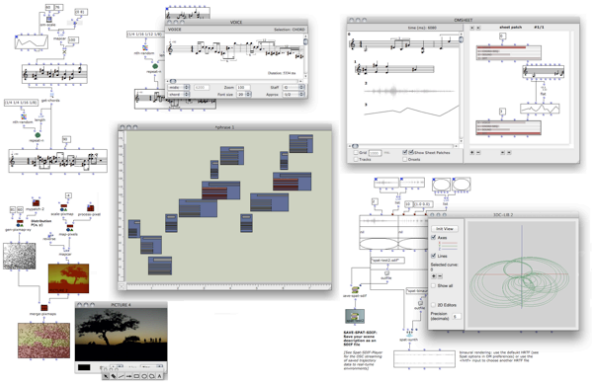
–

Team Involved: Musical Representations

The purpose of research in computer-assisted composition (CAC) is to study and design models and computer techniques adapted to the creative process, incorporating paradigms for calculations as well as musical interactions and representations. This approach favors symbolic orientation using programming languages for artistic creation and processing harmonic, temporal, and rhythmic data in addition to other aspects that come into play in the compositional process. Our work in this domain is articulated primarily around the OpenMusic environment, a visual programming language based on Common Lisp and dedicated to musical composition. Contemporary music composers have used this environment for the past 15 years. Today, is regarded as one of the principle references in computer-assisted composition. Taught in the top computer-music centers and dozens of universities in Europe and around the world, it has been downloaded by several thousand users from around the globe.

OpenMusic (OM) is a visual programming environment for composition or musical analysis assisted by computer. OM offers users a range of interconnected modules associated with specific functions, making up patches that enable the creation or transformation of musical data structures. OM also offers several editors to manipulate these data in addition to libraries in specialized sound analysis and synthesis, mathematical models, the resolution of constraint problems, etc. Unique interfaces like the maquette editor let users construct structures that include functional and temporal relationships among musical objects. The expressive richness of CAC environments is largely due to the junctions between temporal and functional relations within musical structures it enables.

Recently, a new calculation and programming paradigm was suggested for the OpenMusic environment, combining the existing demand-driven approach with a reactive approach inspired by event-driven, interactive real-time systems. Adding serial loops and calculations of musical structures permits the activation of reactive channels in visual programs increases the possibilities for interaction in a CAC environment: an event—a change or an action made by a user—made in a program or in the data it is made from, produce a series of reactions leading to an update (reevaluation). An event can also come from an outside source (typically, a MIDI port
or an open UDP and attached to an element of the visual program). Two-way communication can be established between the visual programs and exterior applications or systems. The CAC environment finds itself inserted in the temporality of a larger system, potentially governed by events and interactions produced by or in this system. This temporality could be that of the compositional process, or that of the performance.



## WRITING FOR SOUND AND SPACE

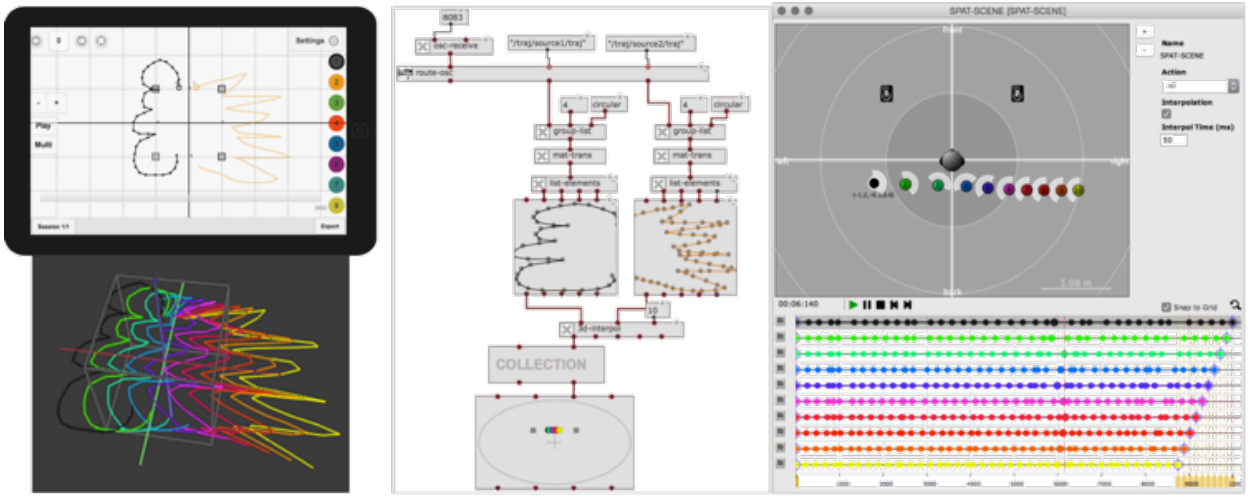
–

Team Involved: Musical Representations

The technologies used for sound synthesis and signal processing have made it possible to imagine new possibilities for musical composition. OpenMusic makes it possible include these technologies via several specialized libraries that connect programs created in a computer-assisted composition environment with processes for sound processing, synthesis, or spatialization (created by IRCAM tools such as SuperVP, Pm2, Chant, Modalys, Spat~, but also tools such as Csound and Faust). Bringing the domain of sound synthesis and computer-assisted composition together is a new approach to sound representation and processing through high-level programs and abstract data structures.

Developed in collaboration with Marco Stroppa, the OMChroma library makes it possible to control the processes of sound synthesis with the help of matrix data structures. OMChroma's extension for spatialization, OMPrisma, enables to implement "spatialized sound synthesis" processes, calling on spatialization (positions and trajectories, but also room characteristics, sound source orientation and directivity) when sounds are produced. Controlled in OpenMusic via an ensemble of graphical editors and operators, these tools propose a wealth of possibilities in the conjoint specification of synthesized sounds and spatialized spaces. The OM-Chant project also made it possible to bring back the technology of FOF (Formant Wave Functions) synthesis at the heart of the process of computer-assisted composition, and to create at the core of CAC processes synthesized sounds inspired by the model of vocal production.

The theme of space, and the collaboration with the Acoustic and Cognitive Spaces team, has been reinforced with the conception of innovative interfaces and software connections to control sound spatialization processes. OpenMusic objects and tools for 2D/3D spatial conception are based on a model of delayed data, enabling homogenous management. Connected to the CAC environment's sequencing motor, the graphical components and signal processors from the Spat library make it possible to render spatialized sound scenes and their evolution over time in an unprecedented way within the compositional process, specifications, and temporal transformations.



## WRITING TIMED INTERACTIONS AND MUSICAL SYNCHRONIZATIONS

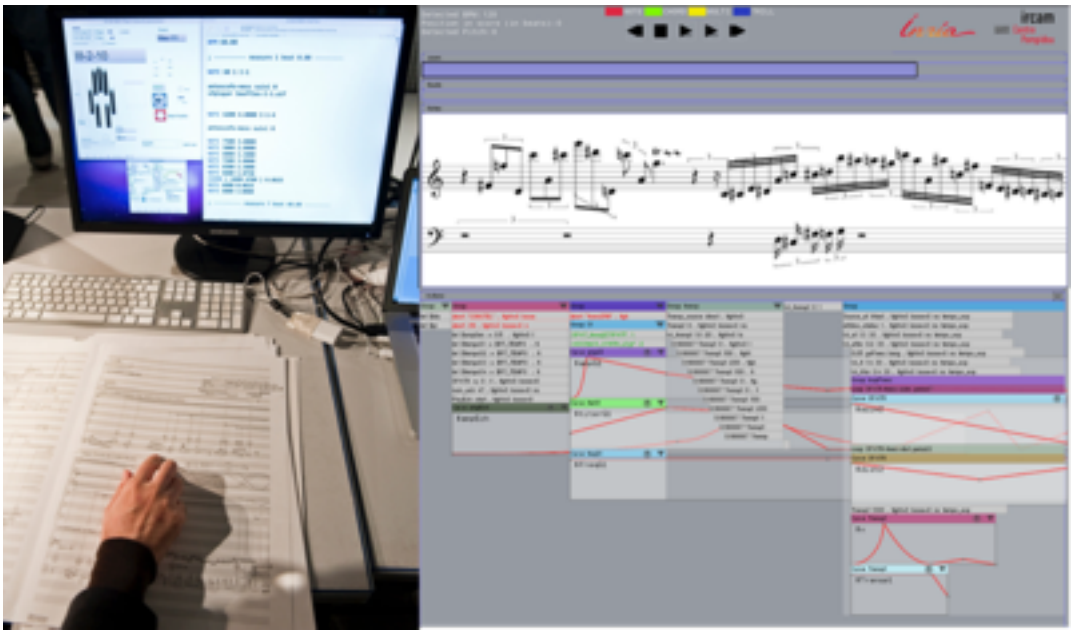
–

Team Involved: Musical Representations

In a context of the performance of contemporary works with instrumental and computer parts, score following is a technique often used for the synchronization of the electronic accompaniment with a solo performer. An augmented score is recorded in the computer using a specific format that contains both the essential elements of the instrumental score a specification for the computer part written in a dedicated musical language. During the performance, the analysis in real-time of the performer's sound and/or of the captured movement is aligned with the symbolic score. At every instant throughout the performance, the score follower's algorithm determines the location in the score and synchronizes the processes programmed in the electronic part of the work.

The score follower (Antescofo) now includes a real-time reactive, synchronous and timed programming language, enabling the definition of complex interactions between performers and live electronics. The use of this technology makes it possible to ensure coherent writing between the electronic and instrumental parts of the composition, and a synchronous and polyphonic performance of electronic modules with musicians in real-time. Its development will include an adaptation for different families of musical instruments and recognition of musical forms as well as the extension of dedicated expressive structures to address various paradigms for the electronic parts.

This system includes two important issues in computer music: the recognition and extraction of musical data in real-time (artificial listening) from an audio signal and reactive synchronous programming for writing time and interaction.



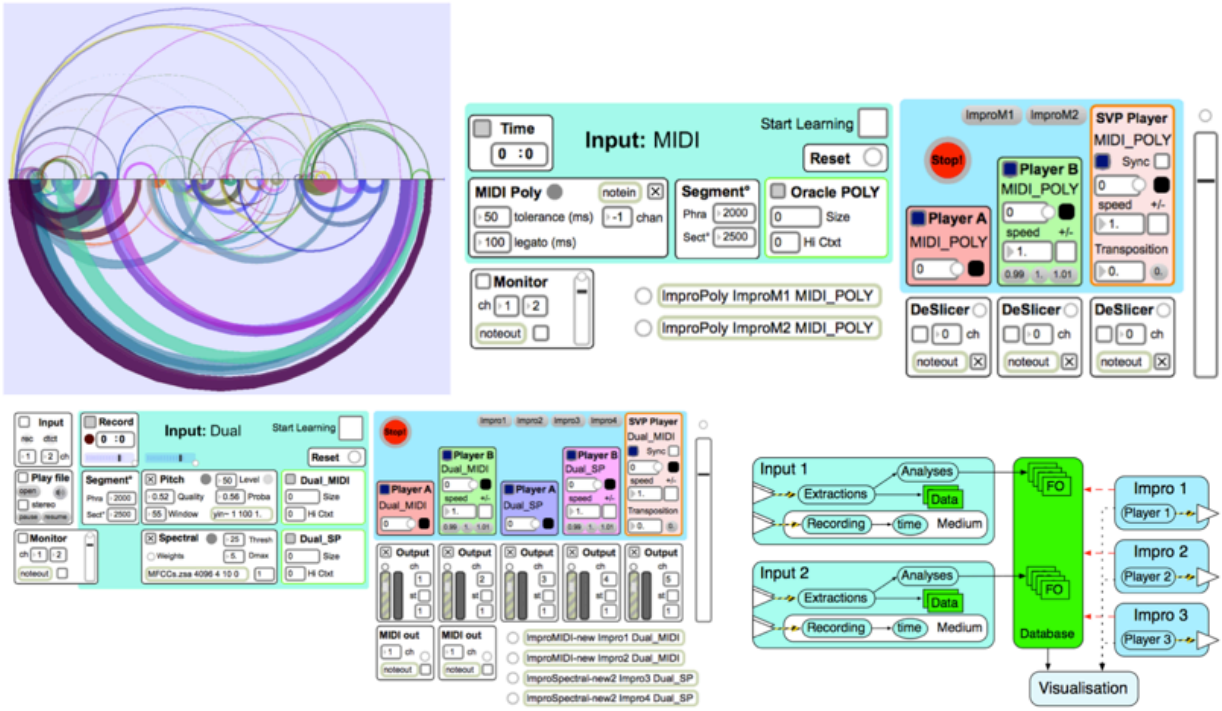
## IMPROVISED MUSICAL INTERACTIONS - OMAX & CO

–

Team Involved: Musical Representations

This project focuses on the development of improvised man-machine musical interactions. A new paradigm for interaction was invented at IRCAM and has been made available to the general public via the OMax software program.
Using machine learning techniques and formal languages, OMax learns in an unsupervised way from either a MIDI or an audio stream produced by a musician. The underlying process behind this interaction could be called "stylistic reinjection". The musician is continually kept informed by several sources providing complex feedback. He hears himself play, he listens to others while memorizing sound images that flow from the present towards the past. Using medium-term and long-term memory, these motifs, combined with even older images taken from the repertoire or musical culture, for example, can return to the present after undergoing several transformations, including one of the most common transformations in improvisation: formal recombination. OMax models this memory-based process and makes it possible to "reify" it, to make it heard on stage. It then re-injects musical figures taken from its short-term and long-term memory and reconstructs them in a manner that is both similar and innovative, providing the musician with stimuli that are familiar and stimulating.

This project has led to two new research projects: SoMax that explores the immediate reactivity of an artificial agent to its sound environment and Improtek (in collaboration with the EHESS) that explores the notion of guided improvisation in the framework of a specific scenario (e.g. a chord chart). The vocation of the ANR-funded project DyCI2 is to study the synthesis of these two different approaches with innovative artificial listening techniques.



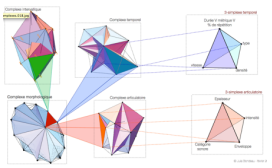
## MATHEMATICS AND MUSIC

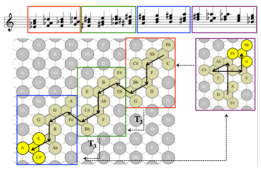
–

Team Involved: Musical Representations

The project "Modèles algébriques, topologiques et catégoriels en musicologie computationnelle (Algebraic Models, Topologies, and Categories in Computational Musicology)" was selected by the CNRS upon the creation of a director's position in the Musical Representations team. This project follows the MISA project (Computer Modeling of Algebraic Structures in Music and Musicology), broadening the spectrum of mathematical tools for computational musicology (tools from algebra and also from topology and category theory). At the same time, this project attacks the theoretical and practical articulations between classical and popular music (rock, pop, jazz, song). Four institutions support the project: the Society for Mathematics and Computation in Music (an international society for which the MISA project was a federating force), the GDR ESARS (Esthétique, Art et Science), a collaboration with Sorbonne Université'IReMus (during international encounters held at the Centre de Recherche on popular music), and a partnership with université de Strasbourg (in particular with IRMA and Labex GREAM). This project also supports the Mamuphi seminar on mathematics, music, and philosophy organized in collaboration with the école normale supérieure and the book collections: Musique/Sciences (Delatour France) and Computational Music Sciences (Springer).

The spatial program aims to model problems such as movements in a specific space or the transformation of spatial structures. It provides computer tools that enable the development of analyses in line with Set Theory. This work has already made it possible to explore the pertinence of topological tools for the representation and classification of musical objects such as calculating all-interval series, Neo-Reimannian harmonic theory, and the geometric representation of chord progressions.





The result is an experimental environment to assist analysis of musical sequences called HexaChord.

HexaChord is an environment that makes it possible to build spatial representations associated with a group of chords and analyzes them via several topological notions. The spatial representations proposed include divers Tonnetz and simplical complexes corresponding to groups of chords' pitches. The software offers 2D and 3D visualizations of the representations produced.

## MULTIMEDIA PUBLICATION TOOLS FOR MUSICOLOGY

–

Team Involved: Analysis of Musical Practices

This work aims to experiment with, and in some cases, standardize the original formats of multimedia publications on music. One can argue that any analytic discourse on music contains a potential 'multimedia' aspect, as it includes a text (the oral commentary) and indicators for musical information (e.g. citations of musical examples, references to movements or passages in a work, etc.). Even so, it is not until recently that musicologists became interested in the new expressive possibilities of multimedia publication, online and on other physical supports that were limited to educational and cultural CD-ROMs until recently. We have begun a two-step process in order to contribute to this process.

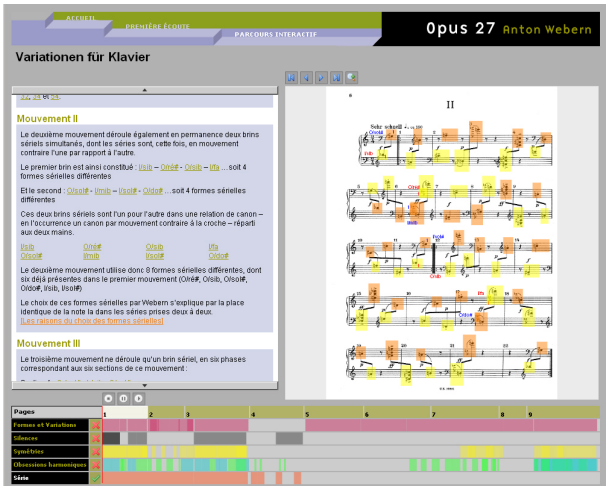
Firstly, the experimentation carried out consists of producing various musicological analyses in a multimedia environment, specifying as we go along the environment chosen according to the steps carried out in the analysis. This experimentation materializes itself via the development of analysis tools developed in collaboration with musicologists (e.g. Charting the Score: A Tool for "Segmented Listening" in collaboration with Jonathan Goldman) and also through the publication of multimedia documents, on DVD-ROM (such as those included in the Inouï, the IRCAM 2005 and 2006 journal) and in online magazines such as DEMeter (Université de Lille-3) or Musimédiane—an audiovisual and multimedia music analysis journal.

Secondly, using a process of comparison and abstraction we will carry out a formalization of the analytic operations brought into play. This enables standardization of publication formats using text, sound, and image—not only for the specific needs of musicologists, but also for disciplines that are close to music publication found within the institutional projects carried out at IRCAM (e.g. the revamping of the Brahms database, the creation of the IRCAM repertoire, software documentation). The issues are those of documenting: using a well-defined format (to facilitate the processing of the information) but one that is not overly restrictive for others; joining visual and sonorous musical excerpts together; making several types of publication possible (online, a CD-ROM, a DVD-ROM, paper). In order to do this, and given the sheer number of commands involved by these projects, it is unfeasible to depend on a manual layout of the data provided by authors for each publication. A tool dedicated to these types of publications must be used to integrate the aforementioned constraints. This was the focus of our participation in ECOUTE (Authoring environment for instrumental music listening, sound archive management, and multi-support publication) supported by the RIAM program and the RNTL Scenariplatform project.

The sum of these actions has already made it possible to

establish a publication format for musical analysis. This format was first tested in a model of online classes for the European project MusicWeb in 2003 and in a scientific publication on the analysis of a performance in 2005. The format was finally formalized with the production of a series of CD-ROMs associated with the Musique Lab 2 project featuring the Variationen op. 27 by Webern and Voi(rex) by Philippe Leroux (2006-2007) as a part of an agreement between IRCAM and the Provence Alpes Côte-d'Azur region. These CD-ROMs offered two different ways to address the work:

- The first introduces the work by letting the user read the score as they are listening. The score was annotated with a few contextual elements.
- The second offers several analytic points of view on the work. Each position leads, via a hypermedia link, to key passages in the score annotated by the author via the Musique Lab Annotation software program. The contents were produced entirely by the author using the Scenari-CHAIN authoring tool and the Musique Lab software package.




An interactive view of the *Variationen op. 27* by Anton Webern (by Emmanuel Ducreux with Nicolas Donin and Samuel Goldszmidt). CD-ROM designed and published by IRCAM and the PACA Region © 2007.

## THEORIES OF MUSICAL COMPOSITION OF THE 20TH CENTURY

–

Team Involved: Analysis of Musical Practices

Funding: CNRS, Réseau international d'études des écrits de compositeurs (Université de Montréal), Éditions Symétrie (Lyon)

This editorial project began in 2003 and aims to publish a substantial panorama of the main compositional theories of the 20th century. While certain theoretical corpora have been broadly published and commentated (e.g. Schoenberg, Xenakis, Boulez, Stockhausen), others, just as important for history and the musical aesthetic, remain relatively unknown due to a lack of available introductory texts. Today, neither musicians nor musicologists have a comprehensive reference source on the fertile relationship between composition and theory during the last century. Even so, the fertility of this relationship was a determining factor in the development of contemporary music, and particularly of computer-music.

Crossing historical musicology, discourse analysis and musical analysis, the book edited by N. Donin and L. Feneyrou will fill the current void with a collective reference publication on the compositional theories of past century—a century characterized by a profusion of writings by artists, aesthetic manifests defining musical techniques and technologies, and research carried out on compositional systems. Sixty international specialists have come together to treat, on one hand, compositional theories specific to an individual or a school—from Schoenberg to Hindemith to Carter to Rihm —and on the other hand, the categories and problems that surfaced throughout the century such as algorithmic music, musical theater, live electronics, or spectralism. Theoretical corpora unknown in French, sometimes unpublished, are presented alongside the most widely explored themes by musicologists during the past 40 years. The contributions offer a synthetic presentation, both technical and historical, of these notions and/ or doctrines, based on the most current knowledge about the articles, manifests, etc. on the composers studied. The level of diversity of the participants reflects different aspects of 20th century musicology, from Philippe Albèra to Elena Ungeheuer, from Angelo Orcalli to Richard Toop. Éditions Symétrie published this two 900-page volume work in 2013.

## ORCHESTRATION

–

Team Involved: Musical Representations
Funding: Conseil de recherches en sciences humaines du Canada (SSHRC)
Calander: April 2015—March 2018
Partners: McGill University (Canada, coordinator), Haute École de musique de Genève (Switzerland)

One of the most complex and mysterious aspects of musical composition, still scarcely studied in the scientific domain, concerns writing timbre, in particular orchestration techniques. The project is based on a comparison of the state of the art in musicology, in psychology of perception, and in computer music to create new tools to address problems connected to orchestration, its perception, and its instruction. The objective is to develop models that can be generalized, that facilitate the instruction and practice of orchestration, assisted by new technologies. The long-term goal is to create an interactive treaty on orchestration comprising knowledge on orchestration practices, the perception of orchestral effects, and digital tools to help resolve orchestration problems including writing electronic sections of mixed-music works. The project relies on a large quantity of annotations made on the corpus of classical music using a range of pertinent categories, intended to supply both perceptive experiences and educational materials as well as to establish practice data for automatic learning algorithms.
The joint research laboratory's role in this project focuses on computer music applications and artificial intelligence for computer-assisted orchestration tools based on the software program Orchids and new research on deep learning techniques.

## RAPID-MIX

–

Team Involved: Sound Music Movement Interaction
Funding: Program H2020- ICT22-2014-c
Calendar: February 2015—January 2018
Partners: Universitat Pompeu Fabra, Goldsmith's College, Plux — Wireless Biosignals S.A., Reactable Systems SL, Somethin' Else Sound Directions Limited, Lambde Limited (Roli), Orbe Sarl
Coordinator: Universitat Pompeu Fabra

Rapid-Mix's goal is to develop new modes of musical interaction by including multimodal data connected to movement and the body. From a methodology point of view, the project is based on user-centric evaluations and on the quick development of prototypes. Advanced techniques for signal analysis and automatic learning are also used. The transfer of these technologies to small business taking part in the project is an important aspect of the project. In this project, the joint research laboratory is responsible for the development of application prototypes (agile prototyping) in addition to its implication in the other aspects of the project.

## COSIMA

**Collaborative Situated Media**

www.cosima.ircam.fr

–

Teams Involved: Sound Music Movement Interaction, Acoustic and Cognitive Spaces

Funding: ANR — programme Contenus et interactions

Calendar: November 2013—April 2017

Partners: Orbe, EnsadLab, ID Scenes, ESBA TALM, No Design

Coordinator: IRCAM

In the past decade, we have seen a huge change in media and their relationship with bodies and space with tangible interfaces, augmented reality, and ambient Internet. The COSIMA project will explore the relationship between body, media, and space via new interfaces and tools for collaborative creation. The aim of COSIMA is to implement a platform for publishing and disseminating medias situated in space and time, combining several sensory modalities. These contents partner proprioception and perception, involving the entire body in their creation and reception phases. The principle of augmented reality combines a digital space with tangible space. The use of this kind of space makes it possible to edit collaborative situated medias. The COSIMA platform's applications include artistic projects, innovative public services, events, and communication.

- Implementation of an Accessible Platform
  The principle objective of COSIMA is to democratize this new type of media via an accessible, generic, open, and interoperable platform. The project aims at the implementation of a platform that uses standard and open formats, enabling relationships with a variety of mobile technologies and existing systems. The platform will make it easier to carry out this type of experience at a lesser cost and in a more accessible manner for the general public (cultural institutions, public services, artists, etc.). Putting in place an environment of multimodal expression in tangible space requires a system made up of different components: client applications installed on mobile terminals connected to a network, a web service that offers an open API, a database server, and an administration interface for the contents.
- Create new Participative Experiences
  The creation of a platform also requires studying different uses and validation. For this, prototypes are developed at each stage of the project to test different scenarios and to collect and analyze the results to improve the platform. The COSIMA project therefore defines a context for experimentation according to different themes: the creation of visual and sound media associated with information about movement and movement to create maps or paths in augmented reality; situated media is shared in a social space for staging pervasive games or zones of public expression; dissemination of situated media in real-time to a community of users makes it possible to create mobs, performances, or large group experiences.
- Supporting the Development of Communities for the Creation and Dissemination of Digital Media
  The combined development of Internet technology and mobile platforms defines new paradigms for the dissemination of medias. The COSIMA platform favors the surfacing and structuring of communities around these new media.

## EFFICAC(E)

**Extended Frameworks For 'In-Time' Computer-Aided Composition**

–

Teams Involved: Musical Representations, Sound Music Movement Interaction, Acoustic and Cognitive Spaces
Funding: ANR – JCJC Program
Calendar: October 2013—March 2017
Partner: CNMAT / UC Berkeley
Coordinator: IRCAM

The objective of this project is the development of tools for computer-assisted composition (CAC) exploring new relations between computation, time, and interactions in the OpenMusic environment and in other technologies developed at IRCAM and CNMAT. This project endeavors to take CAC processes beyond the traditional "offline" domain to include them in a structured interaction in their context be it performance, execution, or composition (in the processes that lead to the creation of musical material).
During this project, we will focus on a number of specific directions, such as:

- Reactive processes for computer-aided composition: communication and propagation of events and changes in compositional processes
- Management of temporal flows: Writing, dynamic scheduling, relations between musical execution time and offline computation processes; representation of temporal constructs
- Control, visualization and execution of sound synthesis and spatialization processes
- Gestural interactions in compositional processes

Several critical antagonisms in the domain of computer music are addressed in this project (signal versus symbolic approaches, offline versus real-time). By bridging high-level computer-aided composition systems with other disciplines and frameworks such as sound processing, spatialization and gestural integration, it includes control and interactions in abstract and expressive compositional models.

## GEMME

**Musical Gesture: Models and Experiments**

–

Teams Involved: Analysis of Musical Practices, Sound Music Movement Interaction
Funding: ANR — programme blanc shs
Calendar: November 2012 — June 2016
Partners: Université de Nice-Sophia Antipolis (CTEL: Centre transdisciplinaire d'épistémologie de la littérature)
Coordinator: IRCAM

In the beginning of the 21st century, gesture became a significant means of interaction with technology for the general public, and also within the musician community. While musical writing has put a strong emphasis on deconstructing and reshaping the acting body of the musician-performer for nearly 50 years, it is interesting to note that for the past 10 years there has been a strong multidisciplinary convergence on this research/creation topic, drawing the attention of composers, performers, and computer scientists as well as the domains of engineering, psychology, physiology, biomechanics, and cognitive sciences. This concept of gesture, commonly used in numerous domains, notably in the performing arts such as theater and dance, has only been the subject of embryonic research in the domain of musicology. The GEMME project offers an analysis of theoretical texts and musical works, and also carries out investigations before and after the premiere of a score: what theoretical and technical possibilities of the formalization of gestures are available to composers? What gestural procedures can they test on paper and during the performance of a work? What means of transmission of the gestural information are created not only during the collaboration between composer and performer, but also when the performance of the work is taught? This project endeavors to answer these questions via four main themes:

1. Tacit Theories of Gesture: genealogy of the compositional notion of gesture, its categorizations and periodization, the current state of the art
2. Gesture and Stage: study of a paradigmatic method—that of Kagel—where the musical idea is connected to its staged expression in the framework of musical and instrumental theater
3. Gesture and Instrument: study of a contrasting paradigmatic method—that of Lachenmann—where the composition calls upon a breakdown of the organological possibilities of sound production in relationship with a political and social criticism of expressive conventions
4. Gesture and Technology: a series of musical analyses of a group of seminal scores, from Ferneyhough's *Time and Motion Study II* to *Luna Park* by Aperghis, that offer a variety of technical and computing paradigms that formalize and/or accompany the instrumental gesture.

## WAVE

**Web Audio: Editing/Visualization**

–

Teams Involved: Analysis of Musical Practices, Sound Music Movement Interaction
Funding: ANR – Contenus et interactions
Calendar: November 2012 — October 2015
Partners: Eurecom Sophia Antipolis, Éditions Durand (Universal Music Publishing Classical), Vi-Live
Coordinator: IRCAM

The WAVE project aims to formalize new means of editing, visualization, and interaction with temporal audiovisual objects online. This project will lead to the design and development of software bricks concerning user interfaces and interactions, digital audio interfaces, client/server interfaces, and their data exchanges. These software bricks will come from the results of the analysis of expert musical practices compared to their ordinary use in web applications and standards, with the goal of offering innovative uses for them. The project's objective is to integrate these software bricks in existing software to develop new services and improve existing ones. New means of purchasing and sharing musical and audiovisual documents online, these services will be put in place during the project and made available to users by project partners using the W3C standards (particularly HTML5 and the new Web platform), completing them if necessary, and also proposing interaction via new terminals in order to offer coherent interfaces, accessible and innovative, and new user experiences adapted for consultation, interaction, annotation, transformation, and sharing of temporal objects. In a fluctuating web technology situation, the formalization and instanciation of these multiple interfaces in different technologies supported by WC3 standards and recommendations will contribute to the surfacing of a truly hypermedia and interactive Internet. In the context of the cultural industry's crisis, and particularly those connected to music, companies and institutions can make use of these cases and explore new markets via innovative systems for listening and consultation that are interactive and enhanced with temporal streaming from the Internet.
This project led to the organization of the first Web Audio Conference (WAC) in collaboration with Mozilla, Google, and the W3C at IRCAM on January 26, 27, and 28, 2015.

## INEDIT

**Interactivity in Writing of Interaction and Time**

–

A
MODELER
ACTION
R
INEDIT
TEMPS
MUSIQUE
R
INTERAGIR

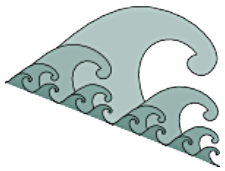
Team Involved: Musical Representations
Funding: ANR – Contenus et interactions
Calendar: September 2012—August 2015
Partners: Labri, Grame
Coordinator: IRCAM

This project aimed to enable the interoperability of sonic and musical creation, opening the path to new creative dimensions that couple writing for time and writing for interaction. Our motivations are reflected in the international efforts to bridge the current gap between compositional and performance aspects of existing tools for sonic creation.
This gap is apparent in the categorization of tools for computer-assisted composition and real-time. For a composer, performance-oriented tools (e.g. sound processing and synthesis, mixing, audio post-production) are often limited to the metaphors of production and sound-processing sources (notably oscillators, filters, and physical models). However, these metaphors should be included in a symbolic language like the final score, enabling real musical writing and favoring more powerful methods of representation, conceptualization, and communication. Inversely, composition-oriented tools still struggle to take into account a new problem, that of interaction: a sound system corresponds more or less to a work that is closed and entirely specified in advance, but must interact in real-time with an environment. This is the case, for example, for interactive musical works, video games, sound decors, or for multimedia systems and installations.
Our approach is based on a language vision: an interactive work is seen like a performer (in the computer sense) who must articulate locally synchronous low-level temporal streams (the audio signal) in a globally asynchronous event time (pertinent events for the composition process). The INEDIT project promotes a hybrid approach that enables cooperation among different tools and stylistic approaches at the heart of an environment that supports all the phases of the musical workflow, from composition to performance. New technologies must be developed to reach these goals: dedicated scheduler, mechanisms for cooperative coupling, on-the-go compilation, as well as innovative visualizations and tangible interfaces that enable the specification and control in real-time of musical processes in a natural process for the creator and designer.

## HOULE

**Hierarchical Object based Unsupervised Learning**

houle.ircam.fr



–

Teams Involved: Sound Analysis & Synthesis, Sound Perception and Design, Musical Representations

Funding: ANR – programme Jeunes Chercheurs

Calendar: September 2011—January 2015

In the context of the HOULE project, IRCAM suggests innovative unsupervised learning techniques to solve the problem of the structuring of audio scenes (Computational Auditory Scene Analysis: CASA); a problem that consists of identifying and describing sound objects that make up a scene. CASA is a very active domain of sound processing with numerous applications.

The points that block the current methods are:

- Without being able to draw up a hypothesis on the nature of the processed objects, we can not model them
- The objects combined in the scene can not be observed isolated from each other
- The structure of the objects is governed by numerous relationships that is difficult to prioritize

The characteristics that we will use for our approach are the hierarchal organization of audio scenes (atoms brought together in objects that are examples of classes such as "Piano A4", itself an example of "Piano note") and the redundancy present at all levels of this hierarchy. This redundancy enables us to identify reoccurring motifs on which a rich and robust representation can be based.

This new method of representing audio scenes has led to the creation of an unsupervised learning algorithm designed expressly to process this data. The system features two components: the MLG (multi-level grouping) that structures the data and the Supervisor (a reflexive adaptation module) embodies the learning aspect by optimizing the MLG function on-the-fly in reference to the stocked memories of past executions.

The originality of our proposition lies in its distance from traditional CASA approaches, beginning with the paradigm of representing scenes and objects. Innovation is primarily present in the unsupervised learning methods that we will develop, including applications that go well beyond the CASA framework.

# 4

# THE SOFTWARE

## ADMIX TOOLS

–

Design and Development: Acoustic and Cognitive Spaces Team

The ADMix tool suite, developed during the research project ORPHEUS (European Union's Horizon 2020 research and innovation program, p. 35), can be used for recording (ADMix Recorder) and reproduction (ADMix Renderer) of object-based audio contents.

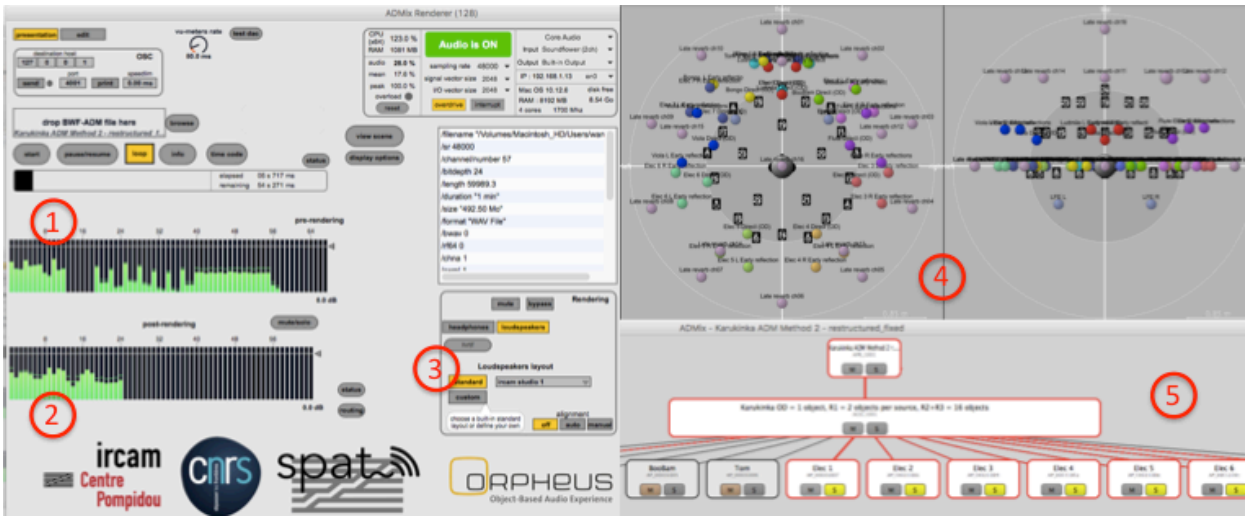
The object-based format follows the audio definition model (ADM) defined by international standardization bodies (European Broadcast Union and International Telecommunication Union). ADM data is typically stored in .wav audio files (BW64) in which a <axml> chunk embeds the metadata. The association of audio channels and the metadata is defined in the <chna> chunk.

ADM elements can be of different types: DirectSpeakers, HOA, binaural, Matrix, or Objects. DirectSpeakers elements make it possible to create a multi-channel bed for standard loudspeaker setup (stereo, 5.1, etc.). HOA elements are used to store a sound scene encoded as High Order Ambisonics. Object elements are the most generic and are defined by their location, as well as a their gain. These positions and gain values can vary over time.

An ADM file can further contain multiple audio "programs" that each exploit all, or a part, of the audio tracks and are associated with different metadata sets. This functionality can e.g. be used to transmit contents in different languages, where only the dialogues in the selected language are played while the musical components and ambient sounds are shared by all programs. Another application of multi-program contents could consist of transmitting a single sound scene, mixed and rendered according to different listening positions. In practice, the ADMix Recorder can connect with any digital-audio work station in order to convert automation tracks to ADM metadata. Communication between the DAW and ADMix is established via the ToscA plugin, and relies on the OSC communication protocol. ADMix Renderer is based on the Spat~ library and makes it possible to render contents on 2-or-3D reproduction systems, over headphones or any loudspeaker setup.

**Configuration**

The ADMix tool suite, downloadable from the IRCAM Forum website, is available for Mac OSX and Windows.




Main window of ADMix Renderer.
1) Meters corresponding to different objects in the audio file;
2) Meters of different output channels created depending on the rendering mode;
3) Choice rendering system (binaural with headphones or standard or ad hoc loudspeakers configurations;
4) Instantaneous visualization of the position of different objects;
5) Graph of the sound scene and activation of objects (mute/solo).

## ANTESCOFO

**http://support.ircam.fr/docs/Antescofo/manuals**

–

Design and Development: Musical Representations

Antescofo is a modular score following system as well as a synchronous programming language for musical composition. The module allows for automatic recognition of the player's position and tempo in a musical score from a real-time audio stream coming from a performer, thus making it possible to synchronize an instrumental performance with computer generated sounds. Antescofo brings together the description of instrumental and electronic parts of a work in a single score via a synchronous language designed for mixed music, with the goal of increasing the expressivity of the process of writing in real-time in a format adapted for the musical language. After changing the score, Antescofo is able to follow the position and tempo of the musicians in real-time and synchronizes the programmed actions for the computer-music portion of the score (setting off and controlling the electronic parts). Antescofo uses a polyphonic audio stream as input. Antescofo is also used as a programmable sequencer to control the processes of sound synthesis and complex spatialization in real-time.

**Applications**

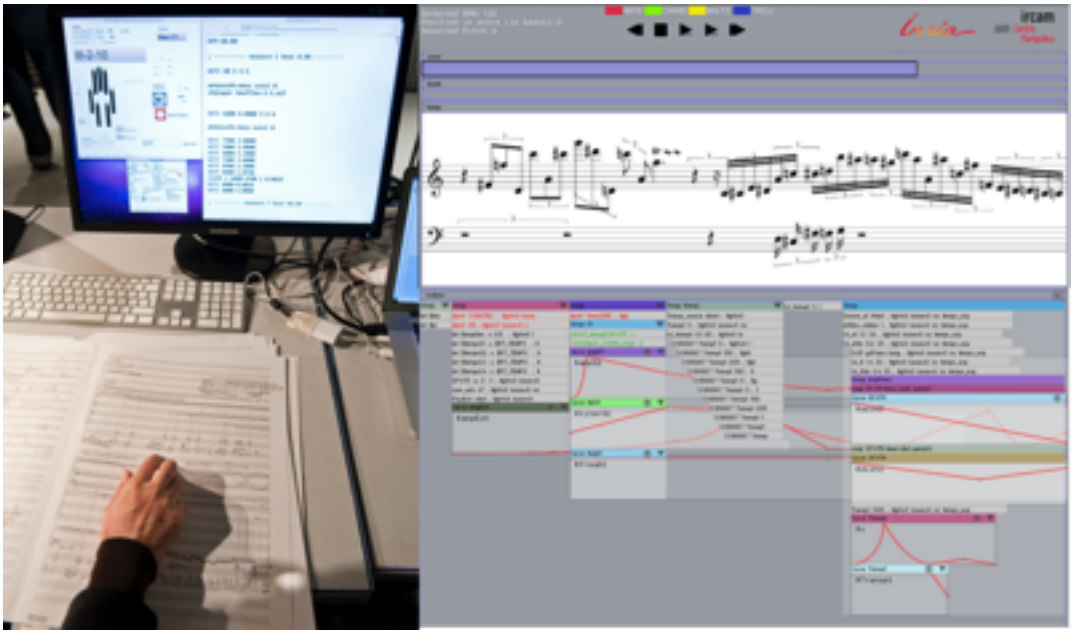
Interactive musical works, automatic accompaniment, mapping gestures/sound synthesis, spatialization control, music education

**Main Features**

- Recognition in real-time of the position and tempo of the polyphonic audio stream in a score
- Writing the electronic components of a score in relative, absolute, and rational time
- Management of events and continual chronometric time
- Interactive sequencing of multimedia scores
- Conversion of MIDI and MusicxML scores
- OSC communication, input/output in JSON format

**Platform**

Antescofo is a library of objects for Max (IRCAM/Cycling'74) that can be used with Mac OSX or Windows XP. Antescofo is also available as a PureData object for Max OSX, Windows, and Linux.



## AUDIOSCULPT

—

Design and Development: Sound Analysis & Synthesis Team

AudioSculpt uses a graphical approach to "sculpt" (to modify) a sound file. After a phase of analysis, the user modifies the analysis' result directly, using only a mouse to apply the desired changes to the sound. Several types of analysis of the spectral content are shown and the user can modify the sound several different ways. The main types of sound modification include filtering, moving segments of the sonogram in time and frequency, or applying one of the numerous high-quality transformations available.

**Applications**

Composition, sound design, post-production, movies, multi-media, acoustics, education, musicology, ethnomusicology, etc.

**Main Features**

- Sound representation – Sound file at a glance, zoom from overall amplitude envelope down to sample level, sonogram, spectral envelope, processing sequencer
- Sound analysis – LPC, True envelope, fundamental frequency, partial following, virtual pitch, estimation of time positions, tools for listening to specific parts of spectrum and showing harmonic relationships
- Annotation: text, segment, and MIDI notes can be added on the sonogram
- Sound processing – a sequencer makes it possible to group several tracks, each with different processing, and to listen to their effects in real-time before finalizing the result
- Filtering – directly on the sonogram using a pencil or eraser or by specifying with the mouse the points forming a polygon. Manual source separation and "denoising" is possible. Compression/expansion – allows modification of the length of a sound without changing its pitch or timbre and preserving the quality of the attack and transient
- Transposition – using a specific editor the user can transpose a sound without loss of data or quality of sound
- Denoising – spectral subtraction with estimated noise interpolation
- Cross synthesis – application of spectral data of one sound to another sound in order to make a spectral transition from one sound to another or an hybrid sound
- Partial synthesis – creation of a new sound from partials found by one of the analysis
- Collage – sound zones defined in terms of their time-frequency can be copied and pasted
- Listening – transformations can be heard in real-time. The scrub mode enables navigation in a sound at any speed
- Bpf curves – users can draw curves directly on the sonogram or on the waveform to control processing

**Technical Features and Configuration**

AudioSculpt is based on a vocoder Sound Analysis & Synthesis engine called SuperVP, which provides the means for performing most analyses and sound modifications. Sinusoidal Sound Analysis & Synthesis is performed in AudioSculpt using the Pm2 tool. AudioSculpt is the intuitive interface that allows access to SuperVP. The user can load text files containing commands addressed directly to SuperVP.

AudioSculpt is compatible with high-quality multi-track sounds (up to 32-bit integer of floating/192kHz) and uses SDIF format for analyses, which can be exchanged with other software programs.

**Platform**

Macintosh on Mac OSX (10.5 or higher). Compatible with AIFF/AIFC, WAV, and Sound Designer II and more.

## CATART

–

Design and Development: Sound Music Movement Interaction Team

Concatenative corpus-based synthesis makes use of a database of recorded sounds and an algorithm for the selection of units that makes it possible to choose the segments of the database in order to synthesize by concatenation a musical sequence. The selection is based on the characteristics of the recording that are obtained by an analysis of the signal and other features, such as the pitch, energy, or specter.
Using this new approach of sound synthesis by the CataRT system in real-time enables an interactive exploration of a sound base in real-time and a targeted granular composition with specific sound characteristics. This makes it possible for composers and musician to reach new sounds, and for sound designers to quickly explore a large sound corpus.
CataRT exists as a standalone application or as modular system implemented in Max with the MuBu extensions. Interaction relies on a simple interface made up of a 2D projection of the descriptors and a navigation through which grains can be selected and played through geometric proximity. This navigation can be controlled with a mouse, external controllers (multi-touch pads, joysticks, cameras), or with the analysis of the audio signal.

### Applications

Interactive musical and multimedia works, sound design, new instruments, searches in sound databases

### Main applications

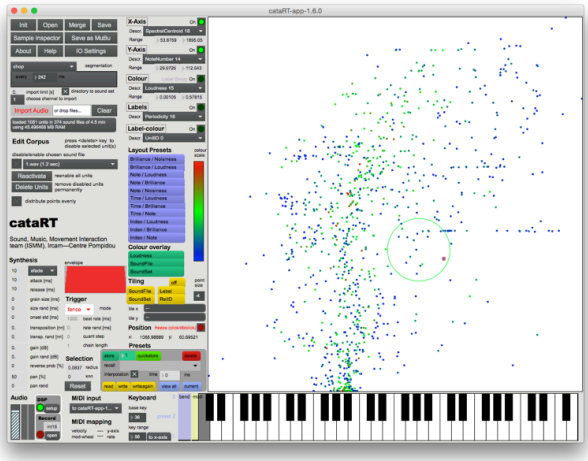
- Audio import with segmentation and automatic analysis
- Analysis, visualization, and sound interaction
- Analysis data export
- Composition via interactive navigation in a descriptor space
- Mixing and juxtaposition of massive amounts of sound samples
- Flexibility of granular synthesis
- MIDI compatibility
- Efficient interactive browsing of events or particularities in several or in long recordings, parasite noise separation

### Platform

- Application Version: Mac OS 10.6 or higher. Universal binary, compatible Mac Intel
- Modular Version: Mac OS 10.10 or higher or PC Windows with Max and the MuBu package

### Audio File Formats

- Up to 32-bit/192 kHz integer and floating point: AIFF, AIFC, WAV, FLAC, AAC, MP3
- SDIF files for importing segmentation markers from AudioSculpt
- MIDI files for importing segmentation and pitch annotation
- Text files for importing segmentation and label annotation (Audacity, for example)



## GESTURE AND TEMPORAL SHAPE FOLLOWING

–

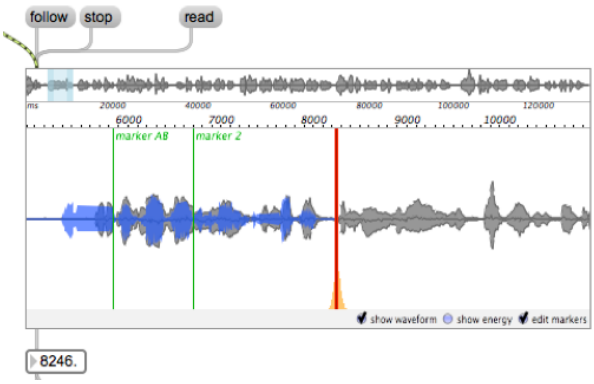
Design and Development: Sound Music Movement Interaction Team

Several Max objects are available to follow temporal morphologies from Markov models.
The voicefollower~ object shows in real-time the temporal progression of a performance in comparison with an audio recording. This enables the synchronization of a number of sonorous or visual processes with a pre-recorded text. This object was tested in a theatrical context with spoken and singing voices.
The motionfollower~ object is similar to the voice follower~ (they both use a similar interface) but functions with gestural parameters, like those that come from sensors. This makes it possible to follow augmented instruments or a beat with a motion sensor.
The gf object (for gesture follower), included in the MuBu modules, enables the recognition and following of the temporal evolution of gestural and/or sound morphologies, comparing them with recorded examples

**Applications**
Interactive musical and multimedia works, music, dance, gesture control of sound synthesis

**Main Features**
- Synchronization between gestures and digital media
- Continual gesture recognition from the beginning of the movement
- Similarity measurements between continual temporal morphologies

**Platform**
Logiciel Max, MacOSx, Windows




## IRCAMLAB TS

–

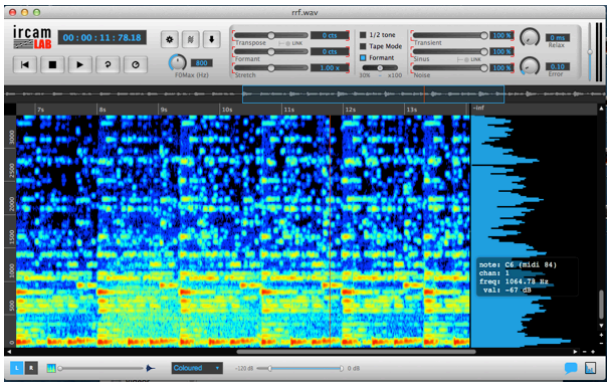
Design and Development: Sound Analysis & Synthesis Team

TS is a standalone software application for sound transformation and time stretching that uses SuperVP and elements of the AudioSculpt 3 interface.

**Applications**
Composition, sound design, post-production, teaching

**Main Features**
- Visualization – display of the sonogram, waveform, and spectrum. Interactive zoom from the temporal signal to sample, linear and logarithmic sonogram. FFT, LPC, and True Envelope analysis
- Compression/Expansion – modification of the length of a sound without changing its pitch or timbre and preserving the quality of the attacks and transitions. Presets for different transposition modes with the possibility of adjusting the max FO and spectral envelope
- Remix – remix the transient, sinus, or noise components of the sound signal
- Control – audio playback with standard controls. Results can be recorded in real-time or offline by modifying the parameters. Possibility of piloting and automation via a MIDI controller

**Platform**
Macintosh on MacOSX (10.5 or higher). Compatible with AIFF/AIFC, WAV, and other formats via the libsndfile library



## MAX

**The world standard for real-time sound interaction**

–

Max (formerly Max/MSP) is a visual programming environment for building real-time musical and interactive applications. It is currently the world's leading software program for interactive sound installations. Max is a bundling of Max (IRCAM/Cycling '74) for real-time control of interactive musical and multimedia applications through MIDI, of MSP, a large set of objects for real- time analysis, synthesis, and processing of audio signals, and of Jitter, a set of video, matrix, and 3D graphical objects for the environment. Max was developed for musicians, sound designers, and researchers who want to develop interactive real-time programs. Cycling '74 (United States) develops Max under an exclusive license from IRCAM.

**Composition**
Generation of musical structures using mathematical or random models.

**Live Performance**
Production of mixed works in which the electronic part reacts according to the instrumental performance using incoming instrumental capture, incoming MIDI signals, or direct modification of the instrumental sound as source material. Controlling the diffusion of sounds and audio files in time Max acts as a sampler. The dynamic control of complex programs that pilot musical signals, video, and lighting effects is also possible.

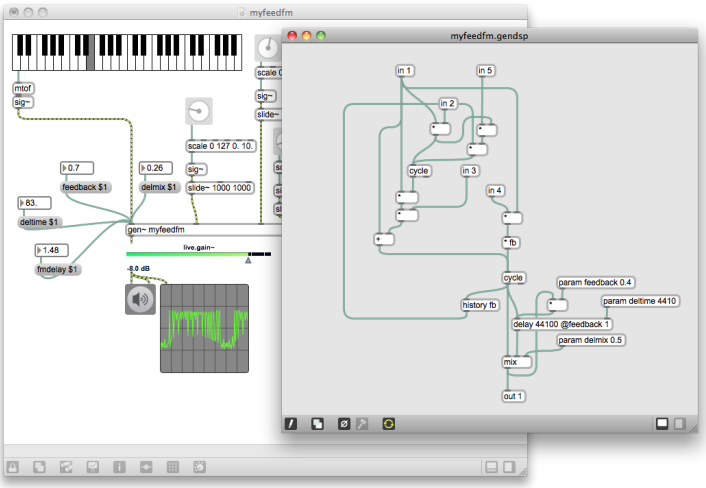
**Audio Post-Production**
Virtual control of several external devices such as the mixing console, direct-to-disk system, and synthesizers.

**Real-Time Video Processing**
A library of objects for video processing (Jitter) is included with the software. It provides users low-level control and the ability to program several applications.

**Education**
Max is widely used in universities, music schools, and conservatories as an educational tool. The documentation included in the software features numerous interactive tutorials that can be used to teach the fundamental aspects of MIDI systems, audio signals, or video.

**Research**
Max is used in research to create prototypes of synthesis or sound signal processing in real-time. New external objects can easily be written in JavaScript, Java, and C and dynamically linked to the program.

**Technical Features**
Several hundred objects for synthesis, control, processing, sampling, and analysis. The sound engine offers multi-processor support, 64-bit signals, and several improvements in sound quality.
Max includes a system of on-the-fly code generation and compilation based on patches, enabling the quick creation of new objects directly in Max.

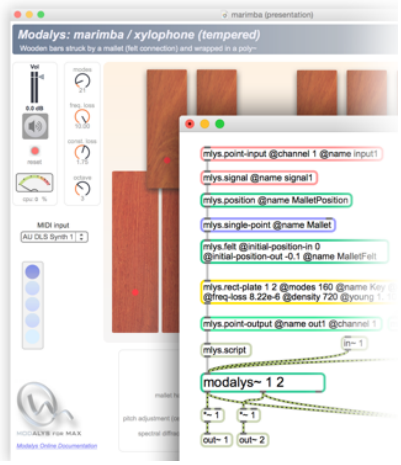


## MODALYS

–

Design and Development: Instrumental Acoustics Team

Modalys is an environment that lets users create unheard of virtual instruments based on simple physical objects such as strings, plates, tubes, membranes, plectrum, bows, or hammers, and making them interact. It is possible to construct objects with complex forms using a three-dimensional mesh or resulting from measurements. Modalys brings these virtual instruments to life by calculating how they vibrate when played. By combining different physical, resonating objects and by specifying the type of interaction (striking, rubbing, mouthpiece, bow, etc.), the user creates a virtual instrument and then decides how it will be played. OpenMusic and Max communicate directly with the Modalys synthesis engine via intuitive graphical interfaces.




### Applications

- Musical composition: contemporary, techno, electro, etc.
- Sound design: natural work on sound quality (metallic, woody, etc.)
- Education: modal theory in practice
- Cinema & video: sound design
- Scientific research & development: comparison of theoretical and measured modes

### Main Features

Modalys' synthesis engine makes it possible to hear virtual instruments made from simple physical objects or 3D mesh. The user can make an instrument and then decide how it will be played. To do this, the user must define the following in four steps:

- Objects which simulate the physical structures (tubes, strings, membranes, plates, complex inexistent forms, plectrums, reeds, hammers) which are calculated according to the default values given by the system or can be modified by the user
- Access to the objects (points on the object on which an action will take place)
- Connections which make it possible to define playing modes: hitting, blowing, bowing, etc.
- Controllers that will describe how the parameters of a connection will evolve over time

### Use

Chosen by a large number of composers and sound designers, Modalys for Max is the incarnation of Modalys in the real-time environment Max, used to construct sophisticated instruments graphically using strings, plates, membranes, tubes with any type of materials (metal, wood, diamond, etc.), size (unlimited), and interactions: hitting, plucking, rubbing, blowing, etc. This interface can be controlled live from Modalys in real-time. Modalys for Max requires Max 5 or higher (including 7) and is available for Mac OS X (10.6 or higher) and Windows (7 or more recent).
Modalys can also be used from ModaLisp, a textual programming interface that can be used to construct and play an instrument with a script. Using this approach, the user writes and executes Lisp programs to obtain the desired musical result. IRCAM's musical programming software, OpenMusic, also based on Lisp, enables a musical or graphical control of the models, for instance from a score and/or temporal envelopes.
For a scientific approach to model synthesis, Modalys can easily be used with MatLab through dedicated libraries.

### Platform

Mac OSX 10.5 or higher, Window 7 or higer (Modalys for Max only)

## MUBU FOR MAX

–

Design and Development: Sound Music Movement Interaction Team

MuBu (for "multi-buffer) is a set of modules for real-time multimodal signal processing (audio and movement), automatic learning, and sound synthesis via descriptors. Using the multimodal MuBu container users can store, edit, and visualize different types of temporally synchronized channels: audio, sound descriptors, motion capture data, segmentation markers, MIDI scores. Simplified symbolic musical representations and parameters for synthesis and spatialization control can also be integrated.
MuBu integrates modules for interactive automatic learning for recognition of sound or motion forms. MuBu also includes PiPo (Plugin Interface for Processing Objects) for signal processing.
The development of the MuBu and PiPo modules is the result of the experience acquired during the development of the FTM & Co libraries as well as development carried out on software components during research projects such as the MuBu, ZsaZsa, and RTA libraries.

### Applications

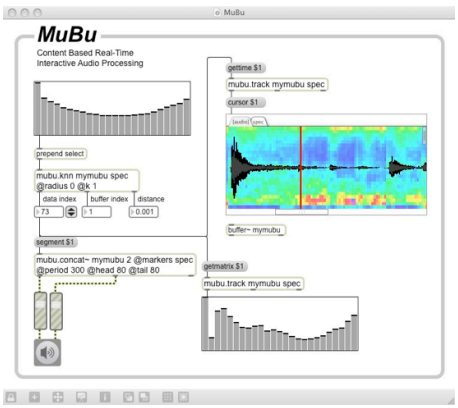
Musical works and interactive multimedia projects, interactive sound synthesis, music, dance

### Main Features

- Multi-buffer with a graphical interface
- Recording data streams and sequences
- Processing data streams and data store in the multi-buffer
- K-NN unit selection
- Concatenative synthesis, granular synthesis, PSOLA
- Additive synthesis

### Platform

Logiciel Max, MacOSx, Windows




## OMAX

–

Design and Development: Musical Representations Team

Omax is an environment for improvisation with a computer that analyzes, models, and re-improvises in real-time the performance of one or several musicians, in audio or in MIDI formats. OMax is based on a computer model called "Oracle Factors", a graph that connects all the motives, from smallest to biggest, and provides a map to the logic of the motives learned from the musician, leading to a large number of stylistically coherent variations. Omax bases its recognition on either notes (pitch following) or on timbres (spectral following). OMax 4 and higher are new versions rewritten in C and in Max.

### Applications

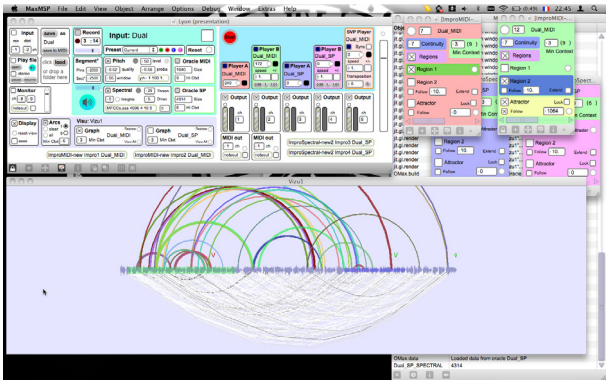
Improvisation with a computer, interactive musical works, automatic accompaniment

### Main Features:

- Artificial Listening: Extraction of the F0 descriptors, MFCC, MIDI input, automatic segmentation, adaptive learning of the symbolic alphabet of musical units
- Modeling: Incremental construction of a model of knowledge representing a full array of musical motives according to diverse similarity functions
- Real-time visualization of knowledge models, multiple selection functions, focalization on the visual representation
- Generation of sequences according to heuristic browsing through the model, taking into account its formal characteristics and musical criteria of continuity of motifs, rhythms, and energy
- Modular architecture that makes it possible to allocated improvisers, players, and to connect them as you like (e.g. several players connected to the same improviser can create a cannon with a rhythmic increase)

### Platform

Max 5 or higher and Jitter, Max OSX 10.5 or higher, storage in the JSon format



## OPENMUSIC

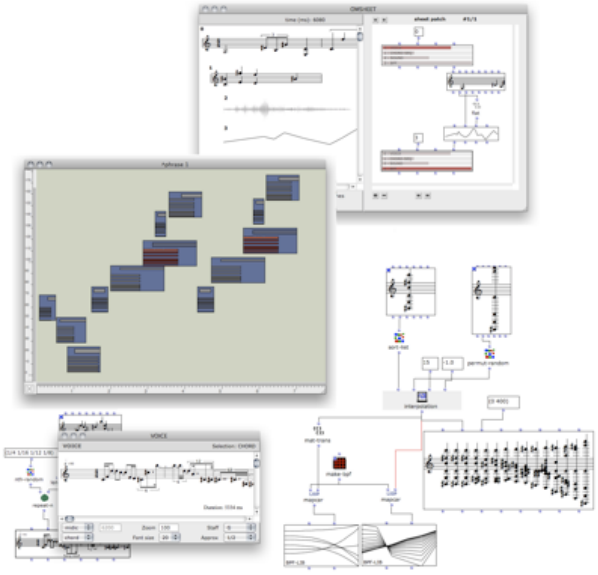
–

Design and Development: Musical Representations Team

OpenMusic (OM) is a visual programming environment for creating computer aided composition applications. OpenMusic offers users the possibility of graphical construction for processing procedures or generation of musical data with the assistance of numerous graphical modules and predefined functions, assembled in visual programs. Programs can then be integrated within each other, producing structures that are more and more elaborate. Today, OpenMusic is taught in the world's major computer-music centers and several universities in Europe and the United States.

**Main Features**

- Graphical construction of functional programs (abstractions/application modes, higher-order functions, conditional structures, iterations, etc.)
- High-level functions for the creation and transformation of musical structures (e.g. pitch, rhythm, etc.)
- Algorithmic data manipulation in MIDI, audio, and SDIF formats, communication with OSC
- Graphical editors: score editing functions for traditional and proportional music notation (with import and export to main software programs), 2D and 3D curves, and breakpoint functions
- Mathematical tools for analysis and composition
- "Maquettes": a graphical editor that makes it possible to give a temporal and hierarchical form to visual programs



**Specialized Libraries**

- LZ: statistical analysis and generation of improvised music
- Esquisse: spectral and frequency functions
- OMRC: programming rhythmic constraints
- OMCS: programming by generalized constraints
- Repmus: The Music Representation Team's function collection
- Profile: manipulates harmonic profiles and contours
- Morphologie: number and symbol sequence analysis functions
- OMChaos: use of fractal functions
- OMAlea: random functions
- OM2Csound: sound synthesis / generating scores for Csound
- OM-SuperVP: for sound analysis, processing, and synthesis with SuperVP
- OM-Pm2: additive analysis / synthesis with Pm2
- OM-Diph: creation of dictionaries and scripts for the Diphone synthesizer
- OM-Chant: control the Chant synthesizer
- OM-Spat: spatialize sound sources with Spat
- OMChroma: high-level structures to control sound synthesis

**Technical Features and Platforms**

Programming based on Common Lisp / LispWorks (http://www.lispworks.com/)
MacOSx, Windows, Linux
Open Source Distribution: LGPL license for the basic environment
Collaboration BEK (Bergen Center for Electronic Arts, Norway) [version OM Linux]

## ORCHIDS/ÉRIC

–

Design and Development: Musical Representations Team

Orchids/Éric is the first complete system for temporal computer-assisted orchestration and timbral mixture optimization. It provides a set of algorithms and features to reconstruct any time-evolving target sound with a combination of acoustic instruments, given a set of psychoacoustic criteria. It can help composers to achieve unthinkable timbral colors by providing many efficient sets of solutions that best match a sound target. With our extended set of features, it can now also reproduce abstract spectral descriptors evolutions (bypassing the need for a sound file). Its results provide multiple orchestral scores that can be arranged in an intuitive way in order to perform a fast sketching of musical ideas. The system provides several algorithms for approximating jointly several timbre properties. The advantages of the Orchids/Éric system is that this approximation can be made separately over temporal shapes, mean values or standard deviations (or any combinations of the three) of each psychoacoustic descriptor. Furthermore, users can also define a manual temporal warping, and even perform multi-target searches inside parts of a sound, quickly providing full orchestral pieces in seconds. The new system provides an extensive out-of-the-box database of orchestral possibilities but can also be unlimitedly extended, even with synthesis sounds, by a simple drag-and-drop operation. Finally, the software provides an intelligent time series querying system that allows to easily search for temporal shapes inside the sound database.

**Applications**

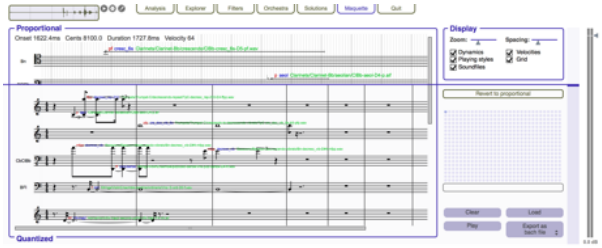
Orchestral composition, re-orchestration, timbral writing, music education

**Main Features**

- Generation of new combinations and orchestral combinations by optimization of a target
- Possibility of timbral writing through orchestral mixtures and optimization of the temporal evolution
- Definition of abstract targets and temporal evolution of spectral descriptors
- Simple interface enabling the spatialization of an orchestra or efficient work of the model of an orchestral work
- Simple extension of a knowledge base and efficient temporal searches
- Interaction and exports towards several systems of musical notation

**Platform**

Orchids/Éric is a stand alone software program that works with all Mac OSX operating systems. Orchids/Éric is also available as Max patches and as a C++ server.

## PANORAMIX

–

Design and Development: Acoustic and Cognitive Spaces Team

Panoramix is a post-production workstation for 3D-audio contents.
This tool offers a comprehensive environment for mixing, reverberating, and spatializing sound materials from different microphone systems: main tree, spot microphones, Higher Order Ambisonics capture.
Several 3D spatialization techniques (VBAP, HOA, binaural) can be combined and mixed simultaneously in different formats.
Panoramix also provides conventional features found in other mixing tools (equalizer, compressor/expander, grouping parameters, routing of input/output signals, etc.), and it can be controlled entirely via the Open Sound Control protocol.

### Architecture

Like a traditional mixing desk, the panoramix interface is designed as vertical strips (see Figure) which are organized in two main categories: input tracks and busses. Tracks are used to process signals from a recording session.
Panoramix supports several input formats: mono (typically from a spot microphone or an electronic track), multi (a multitrack is basically a grouping of several mono tracks), SMA (used to process recordings from a spherical microphone array like the EigenMike ®), or tree (used to process a main microphone tree).

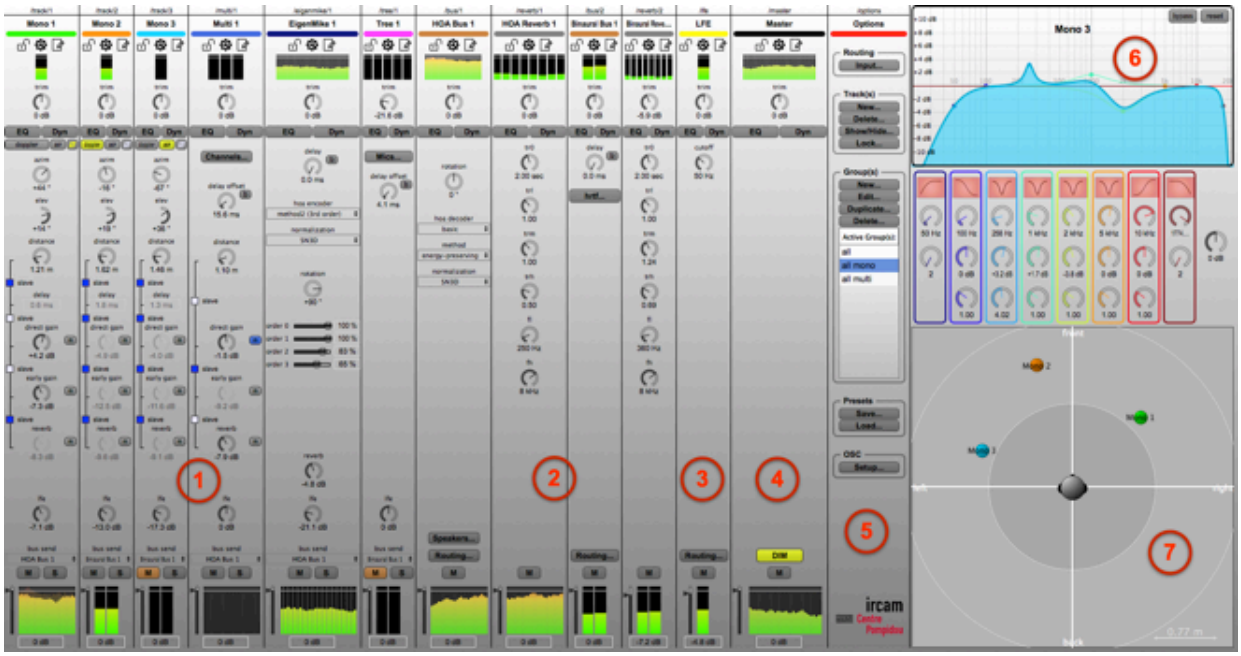
Three types of busses are available: a panning bus, a reverb bus, and the LFE (low frequency enhancement) bus.
The purpose of the panning bus is threefold: 1) it allows for the summation of multiple input tracks, 2) it determines the spatialization technique (VBAP, HOA, or binaural) being rendered, 3) it provides various settings for the encoding and decoding of the signals (e.g. HOA decoder settings or HRTF attributes in the case of binaural reproduction).
The reverberation busses create and control the late/diffuse sections of the artificial reverberation. The current version implements a feedback delay network.

Finally, all the output signals from all the busses are collected in the master strip and then sent to panoramix's physical outputs or bounced to disk. While each session only has one master track, it is possible to create mixes in several formats simultaneously.

### Configuration

Panoramix can be downloaded from the IRCAM Forum website and is available for the Mac OSX, Windows, and Linux environments.

The panoramix station console window 1a) Mono input channel; 1b) HOA input channel; 2a) Panning and HOA reverberation bus; 2b) Stereo bus; 2c) VBAP bus; 3) Session options (management of groups, import/export presents, etc.; 4) Module to create focus groups in HOA ; 5) Geometric positioning interface



## SPAT~

–

Design and Development: Acoustic and Cognitive Spaces Team

Spat~ is a library dedicated to sound spatialization in real-time. Originally designed as a library, it enables musicians and sound engineers to control the spatial sound processing for various sound broadcasting systems. Applications range from reproducing sound in a home setting (with headphones, a stereo system, or 5.1) to concert situations (2D or 3D loudspeaker distribution) to holophonic and interactive sound installations.

**Applications**

- Concerts and Real-Time Spatial Sound Processing
  Composers can map a particular room effect or sound position to each note or musical event in their score by using a score-following program. The Spat~ can be controlled via a sequencer, a score-following program, or by using a high-level control dedicated to composition (e.g. ListenSpace, OpenMusic, etc.).
- Post-Production
  Spat~ can be added to each channel of a mixing console or virtual mixing environment, enabling intuitive control of the position of each sound source and the effect of a virtual room.
- Simulation and Virtual Reality
  Spat~ is well suited for immersion through binaural rendering when wearing headphones. The effect is even more impressive if the system is associated with a head-tracking system. The sound sources maintain their localization according to the user's navigation.
- Holophonic Reproduction
  Spat~ has been adapted to the holophonic reproduction system (WFS), facilitating the reconstruction of an acoustic sound field while preserving the consistency of spatial characteristics in an extended zone for the audience. For interactive installations, this reproduction mode enables natural spatial sensations to be retrieved during the exploration of the sound scene.

**Technical Features**

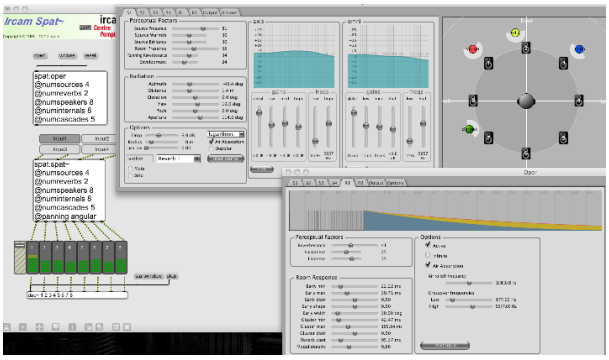
- Perceptual Controls: azimuth and elevation, distance, source presence, room presence, running and late reverberance envelopment, frequency balances.
- Low-Level DSP Controls: equalization, Doppler effect, air absorption, multi-channel reverberation, source direction according to of the reproduction system (e.g. binaural, transaural, stereo, 3/2 stereo, 4 to 8 HP multi-track, VBAP for 3D loudspeaker configurations, or Wave Field Synthesis).

**Platform**

Spat~ is a library of objects for Max (IRCAM/Cycling '74) that runs on Mac OSX or Windows XP. The Spat~ is also available as a collection of software plug-ins developed in collaboration with the company Flux:: that can be used in software for music edition and in major brand-name digital audio workstations (e.g. Protools, LogicAudio, Digital Performer, etc.).

**Technology and Partners**

France Télécom, Cycling'74, Flux::



## SUPERVP, TRAX, AND LIBRARIES

–

Design and Development: Sound Analysis & Synthesis Team

SuperVP is a signal-processing library that is based on an implementation of a sophisticated phase vocoder algorithm. The library can be used to perform a large number of signal transformations with outstanding sound quality (time stretching, transposition of pitch and spectral envelope, denoising, noise-sinusoids-transient remixing, amplitude scaling of the spectral envelope, generalized cross synthesis, source filter cross synthesis, etc.) and provides access to an extensive set of parameters that offer complete and fine grained control over the different algorithms. In addition to sound transformation algorithms, the library includes a large collection of signal analysis algorithms (F0, onsets, spectrogram, re-assigned spectrogram, spectral envelope, etc.). One of the by- products of SuperVP is the VoiceForger library that offers a large number of real-time SuperVP processing and provides high-level access (see the Sample Orchestrator, Vivos, and Affective Avatars projects).
The TRaX tools have been designed so that the majority of the library parameters have been condensed in a set of high-level user parameters that enable an intuitive control of the application.

**Main Features**
Signal Transformation:

- Time Stretching/Compression: with and without preservation of transients, special mode with waveform preservation for speech transformation
- Transposition: independent transposition of pitch and spectral envelope, sample rate transformation (re-sampling)
- Denoising: interpolation of multiple noise keys and independent processing of sinusoidal noise
- Noise-Sinusoids-Transient Remixing: detection of sinusoidal, noise and transient components with independent mixing controls
- Envelope Transformation: estimation of the spectral and time envelope with independent mixing controls
- Cross synthesis: phase vocoder or source filter based cross synthesis with various controls

**Signal Analysis:**

- Spectrogram: Short-time Fourier transform (STFT) with and without reassignment
- Spectral Envelope: LPC, True Envelope
- Fundamental Frequency (F0): precise analysis with very low CPU usage, specialized instrument presets for various instruments (i.e. violin, flute, xylophone, vibraphone)
- Attacks: versatile algorithm for the detection of note onsets and other transients
- Voiced/Non-Voiced: detection of the cut-off frequency boundary

**Technical Characteristics and Configuration**
The library can be controlled on different levels of abstraction. It supports file and memory based sound input/output as well as real-time control of many parameters. The library is used in a wide variety of environments. There is a command line application using file based sound input/output that gives access to all available sound transformations and analysis. It is used for example in the AudioSculpt and OpenMusic applications (see OM-SuperVP). In Max, objects enable sound transformation and analysis in real-time; the TRaX application is based on these objects. The library is programmed in C + + using advanced optimized methods, such as SIMD operations (SSE2, SSE, AltiVec) found in today's processors. The library runs on Mac OSX operating systems, Windows XP, and Linux on Intel and PowerPC. Its multithread implementation makes it possible to exploit the calculated capabilities of the multi- core and computer processors.

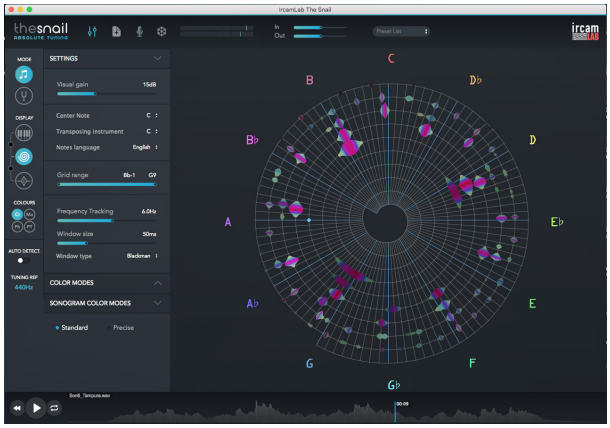
**Technologies and Partners**
MakeMusic (SmartMusic), Flux::, Univers Sons, Xtranormal, Voxler, MXP4, etc.

## THE SNAIL-ABSOLUTE TUNING

–

Design and Development: Sound Analysis & Synthesis Team

The Snail-Absolute Tuning is a sound frequency analyzer that uses an inventive representation with chromatic alignment. Compared to standard analyzers, this representation makes it possible to easily visualize active zones with light (like on a spectrogram) and to organize frequencies by note on a tempered scale, in tune with a diapason.




**Applications**

Musical visualization of a sound signal, education, tuning

**Main Features**

- Visualization – display of The Snail-Absolute Tuning's chromatic abacus simplified for a musical image of sound
- Interface – configurable analysis engine. Settings for different modes of visualization: phase, spectrum, or snail (tuner). Settings for the frequency region visible in MIDI notes. Diapason tuning

**Platform**

Macintosh on MacOSX (10.6 or higher), iOS (iPad). Compatible with all audio files (all formats supported by the native MacOSX library) or signal input via the audio input of the equipment used

## TOSCA

–

Design and Development: Acoustic and Cognitive Spaces Team

ToscA is a plugin that makes it possible to send and receive the automation parameters of a digital audio workstation towards or from other applications applications using the OSC protocol. Its application typically concerns the production of object-oriented spatialized mixes independently of the host software constraints.

**Main Features**

ToscA can be inserted into a DAW (digital audio workstation) for each track you want to spatialize. ToscA does not affect the audio signals, which are simply short-circuited. However, the plugin provides a certain number of parameters that can be automated. In order to ensure compatibility with a wide range of digital audio workstations, the plugin exposes a maximum of 32 parameters per track. During playback, active automation tracks are read by ToscA and the corresponding OSC messages are sent over UDP. When the automation tracks are armed for recording, ToscA accepts incoming OSC packets from remote applications and data can be written in the sequencer tracks.

Each ToscA instance has an ID that can be adjusted by the user (see illustration). This ID is the root pattern of the OSC messages sent and received by ToscA. Their syntax is as follows: "/ID/ParamaterName ParamaterValue" (e.g. "/3/azimuth 135.0"). For the sake of simplicity, the parameters in ToscA are all double precision floating-number points. Other types of data that can be included in an OSC message (such as integer numbers or text strings) are not currently supported.

ToscA is not connected to a particular spatialization or synthesis engine. In other words, the automation parameters are generic. By default, these 32 parameters are called param1, param2, and so on. The labels of the parameters can be modified. This is carried out via an XML mapping file that can be loaded into ToscA (see illustration).

The XML mapping file makes it possible to specify the automation scale for each parameter: inside the DAW environment, ToscA scales up the tracks in a [0 – 1] floating point range towards the [min – max] domain selected by the user. This scaling affects data output and the inverse scaling is applied to input data.

**Platform**

ToscA was developed with the Juce framework. Juce offers a plugin wrapper that makes it possible, using a common base code, to generate plugins in different formats (VST, AudioUnits, etc). Therefore, ToscA is available for MacOSX and Windows, in VST, VST3, AU, and AAX formats and in either 32 or 64 bit configurations. ToscA is distributed by the IRCAM Forum.


```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>

<tosca version="0.4">
    <parameter index="1" name="azim" min="-180" max="180" scaling="linear"/>
    <parameter index="2" name="elev" min="-90" max="90" scaling="linear"/>
    <parameter index="3" name="gain" min="-60" max="0" scaling="linear"/>
    <parameter index="4" name="aperture" min="0" max="180" scaling="linear"/>
    <parameter index="5" name="orientation" min="-180" max="180" scaling="linear"/>
</tosca>
```

A view of ToscA in ProTools and an example of a file for automation parameter settings
1. Plugin inserted in track 1
2. List of parameters exposed
3. Plugin window
4. Automation tracks

# THE IRCAM FORUM

IRCAM is committed to artistic research, serving music, visual art, digital, and performing arts creators. They are found in the IRCAM Forum community, a community that touches other professional sectors as well—sound design and engineering, teaching and performance—and has over 10,000 members worldwide.

## A COLLABORATIVE ONLINE PLATFORM
### FORUM.IRCAM.FR

In the fall of 2018, the IRCAM Forum will launch an updated collaborative platform featuring a host of new functions. In addition to the already existing professional networking tools, communication, and technology dissemination, the platform will include a new service that facilitates artistic project management and offers multiple views for each step of your project's elaboration. The IRCAM Forum simplifies group development of new ideas, favors sharing of practices, archiving, and indexing of knowledge, and offers an overview of the tools and innovations available to meet new challenges.

**PUBLIC** ARTISTS, RESEARCHERS, ENGINEERS, DESIGNERS, SOUND DESIGNERS, MAKERS, TEACHERS, SOUND AND MUSIC PROFESSIONALS

## THE IRCAM FORUM WORKSHOPS

Three days of meetings open to professionals and amateurs. Three days of conferences, workshops, demos, and exchanges on music and sound in a multidisciplinary spirit.

## THE IRCAM FORUM WORKSHOPS BEYOND IRCAM

This event is held once a year abroad in collaboration with universities, research and cultural centers the world over (Seoul 2014, Sao Paulo and Buenos Aires 2015, Taipei 2016, Santiago du Chili 2017).

## NEXT EVENT
**IRCAM FORUM WORKSHOPS**
**➜ March 26-29, 2019 IRCAM**

**CONTACT**

**01 44 78 49 62**
**ircam-forum@ircam.fr**

**FORUM.IRCAM.FR**

**RESEARCH: CREATION INTERFACES**
Manager **Grégory Beller**
**Karim Haddad, Stéphanie Leroy, Paola Palumbo, Raphaël Voyazopoulos**

# THE TEAMS FROM JUNE 2017 TO JUNE 2018

## ■ UMR 9912 – SCIENCES AND TECHNOLOGIES OF MUSIC AND SOUND

**Director:** Brigitte d'Andréa-Novel (SU)
**Deputy Director:** Jean-Louis Giavitto (CNRS)
**Directeur (acting):** Carlos Agon Amado (SU) (July-December 2017)

**Administration**
Sylvie Benoit, Carole Tan, Anne-Marie Vaudeville (CNRS), Vasiliki Zachari (CNRS)

**Head Researchers**
Sound Systems and Signals: Audio/Acoustics, Instruments: Thomas Hélie
Acoustic and Cognitive Spaces: Olivier Warusfel
Sound Perception and Design: Patrick Susini
Sound Analysis & Synthesis: Axel Roebel
Sound Music Movement Interaction: Frédéric Bevilacqua
Musical Representations: Gérard Assayag
Analysis of Musical Practices: Nicolas Donin

**Researchers and Engineers**
Carlos Agon-Amado (SU), Moreno Andreatta (CNRS), Luc Ardaillon (CNRS), Jean-Julien Aucouturier (CNRS), Mondher Ayari (univ. Strasbourg), Dogaç Basaran (CNRS), Joël Bensoam, Georges Bloch (univ. Strasbourg), Alain Bonardi (univ. Paris-8), Riccardo Borghesi, Henri Boutin (SU), Jean Bresson, Clément Cannone (CNRS), Baptiste Caramiaux (CNRS), Thibaut Carpentier (CNRS), Carmine Emanuele Cella, Frédéric Cornu, Ken Deguernel (univ. de La Rochelle), Franck Elisabeth, Philippe Esling (SU), Laurent Feneyrou (CNRS), François-Xavier Féron (CNRS), Emmanuel Fléty, Raphaël Foulon, Hadrien Foroughmand, Dominique Fourer, Annelies Fryberger (SU), Matthias Geier, Jean-Louis Giavitto (CNRS), Louise Goupil (CNRS), Karim Haddad, Thomas Hélie (CNRS), Lise Hobeika (CNRS), Olivier Houix, Marguerite Jossic, Pierre Laffitte, Jean-Philippe Lambert, Quentin Lamerand, Joseph Larralde, Raphaël Leïba (Sorbonne Université), Marco Liuni (CNRS), Yi-Ju Lin (National Cheng Kung University), Benjamin Matuszewski, Adelino Mendes Ferro, Gabriel Meseguer, Rémi Mignot, Nicolas Misdariis, Jérôme Nika (univ. de La Rochelle), Markus Noisternig, Nicolas Obin (SU), Geoffroy Peeters, Charles Picasso, Robert Piéchaud, David Poirier Quinot, Mathieu Prang, François Ribac (univ. de Bourgogne), Alexandre Robert (SU), David Roze (CNRS), Stéphan Schaub (univ. d'État de Campinas), Norbert Schnell, Nadine Schütz (ETH Zurich), Diemo Schwarz, Lou Seropian (CNRS), Jordan Smith (Centrale Supelec) Marine Taffou (IRBA), Vincent Tiffon (univ. Lille 3), Louise Vasa (CNRS), Isabelle Viaud-Delmon (CNRS). Anders Vinjar (CNRS), Gabriel Vogel (CNRS)

**Technicians**
Djellal Chalabi (CNRS), Camille Dianoux (SU), Arnaud Recher (CNRS)

**Documentalist**
Sandra El-Fakhouri

**Doctoral Students**
Pavlos Antoniadis, Pablo Arias, Baptiste Bacot, Adrien Bitton, Julia Blondeau, Damien Bouvier, Daniel Cabanzo, Hugo Caracalla, Tristan Carsault, Axel Chemla, Alice Cohen, Leopold Crestel, Ken Deguernel, Pierre Donat, Guillaume Doras, Jean François Ducher, José Miguel Fernandez, Grégoire Genuys, Daniele Ghisi, Lise Hobeika, Céline Jacques, Marguerite Jossic, Frédéric Le Bel, Tristan Lebrun, Raphaël Leïba, Jeanne Maisonhaute, Namur Matos Rocha, Lara Morciano, Remy Muller, Laura Rachman, Alessandro Ratoci, Kevin Sanlaville, Hugo Scurto, Maxime Sirbu, Marco Suarez Cifuentes, Pierre Talbot, Marc Wijnand

**Interns**
Achille Aknin, Cyran Aouameur, Victor Audouze, Daniel Bedoya, Paul Best, Damien Bestard, Elena Brunet, Francesco Cameli, Pierre Carré, Tristan Carsault, Nicolas Chartier, Eugénia Eremeeva, Godelive Forceville, Hadrien Foroughmand, Pierre-Amaury Grumiaux, Thomas Guennoc, Nicolas Haezebaert, Vincent Iampietro, Mélissa Jeulin, Sarah Jibodh-Jiaouan, Cyril Jost, Colin Lardier, Pierre Lavalette, Jonas Margraf, Lucie Marignier, Giacomo Martinelli, Pierre Massé, Judy Najnudel, Ward Nijman, Rosalie Ollivier, Romain Petit, Anh-Khoa Pham, Mathis Raibaud, Pierre Rampon, Carl Robinson, Felix Rohrlich, Paul Saint-Aubert, Victoire Siguret, Clément Tabary, Bavo Van Kerrebroeck, Victor Wetzel, Marc Wijnand, Frank Zagala

## ■ DEPARTMENT OF INNOVATION & RESEARCH RESOURCES

**Directeur:** Hugues Vinet

**Administration, Coordination, Project Management**
Sylvie Benoit, Louise Enjalbert, Clémentine Gorlier, Ghislaine Montagne, Francis Rousseaux

**Industrial Relations**
Manager: Frederick Rousseau
Charles Picasso

**Web Service**
Manager: Guillaume Pellerin
Cécile Drencourt, Emilie Zawadzki

**Information Technology**
Manager: Raphael Sallé de chou
Benjamin Boudoir, Rémi Desmonet, Mikhaël Gauthier, Olivier Labat

**Contact:** firstname.lastname@ircam.fr

**Publication Director** Frank Madlener **Publication Editor** Brigitte d'Andréa-Novel **Translation** Deborah Lopatin
**Layout** Olivier Umecker **Photocopying** Jean-Paul Rodrigues **Cover** BelleVille © photography Thinkstock

UMR 9912 – STMS
IRCAM | 1, place Igor-Stravinsky 75004 Paris, France | +33 (0)1 44 78 48 43


Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
Ministère
Culture

ircam
Centre
Pompidou

cnrs

SORBONNE
UNIVERSITÉ