Nicolas OBIN

Associate Professor
Ircam, Sorbonne Université
Est. since 1980

Ircam
Analyse et Synthèse des Sons
1, place Igor Stravinsky 75004 Paris
☎ *(+33) 01 44 78 48 90*
✉ Nicolas.Obin@ircam.fr
in nicolasobin

I am associate professor (2013) at Sorbonne Université and research scientist in the Sound Analysis and Synthesis (A/S) team at the Science and Technology for Sound and Music laboratory (Ircam, CNRS, Sorbonne Université, Ministère de la Culture). I received in 2011 a PhD. thesis in computer sciences on the modeling of speech prosody and speaking style for text-to-speech synthesis for which I have been awarded by La Fondation Des Treilles. Through the years I have developed a strong interest in the behavior and communication between humans, animals, and robots. My main research area is the generative modeling of complex human productions with various applications in speech synthesis and transformation, multi-modal virtual agent animation, and humanoid robotics. As part of my artistic commitment to Ircam, I am actively promoting digital science and technology for arts, culture, and heritage, and have collaborated with renowned musicians and artists, such as: Georges Aperghis, Philippe Manoury, Pascal Dusapin, Alexander Schubert, Roman Polansky, Philippe Parreno, Eric Rohmer, Leos Carrax, and André Dussolier.

Education and Professional Experience

[2022-present] Head of the Master in Engineering of Intelligent Systems, 80-90 students, Sorbonne Université, Paris, France.

[2013-present] Associate Professor, STMS (Ircam, CNRS, Sorbonne Université, Ministry of Culture) and Sorbonne Université, Paris, France.

[2011] Prix de thèse de la fondation Des Treilles, section télécommunication

[2006-2011] PhD. thesis, Me$\mathcal{L}$os: Analysis and Modelling of Speech Prosody and Speaking Style, Université Pierre et Marie Curie. Supervision: Xavier Rodet (Ircam) and Anne Lacheret (Université Paris Ouest Nanterre - La Défense)

[2005-2006] Master in Engineering ATIAM (Acoustics, Signal Processing and Informatics Applied to Music), Université Pierre and Marie Curie. Major of the promotion.

[2006-2008] Master in Musicology (Musicology, Arts, Aesthetics), Université Paris 8 - Vincennes Saint-Denis . *Pierre Boulez, Franz Kafka: L'écriture de l'irrationnel*. Supervision : Ivanka Stoïanova.

Highlights

Research activities

Supervision	Since beginning	Since 2013	Since 2019
Post-doc	3	3	2
PhD Thesis	9 (+1)	9 (+1)	9 (+1)
Master Thesis	14	13	7

Table 1: Summary of supervised research (since 2008) (+1) = unofficial

Publications	Since beginning	Since 2013	Since 2019
Journal (int.)	8	7	4
Journal (nat.)	2	1	1
Book	0	0	0
Chapter	5	4	3
Conference (int.)	52	29	18
Conference (nat.)	13	7	1

Table 2: Summary of publications (since 2008) Int. = international, Nat. = national

[2022] International French-Japan Forum Intelligence Augmentation and Amplification + Society (CNRS, EHESS, JST), Paris, France, 2022. https://www.eu-japan.eu/news/ai-and-social-sciences This Franco-Japanese summit brings together researchers in artificial intelligence, robotics and social sciences, I was invited to participate as a researcher on the theme "Working Living in a Cyber Physical AI world: which interactions between human being and machine?".

[2013-present] Member of organization and scientific committee for the international Workshop VIHAR (Vocal Interactivity in-and-between Humans, Animals and Robots, 2021), SIVA (Socially Intelligent human-like Virtual Agents, 2023, founding member), SSW (Speech Synthesis Workshop, 2023) and Blizzard Challenge (2023)

[2013-present] Regular reviewer for the main international conferences and journals in speech and signal processing, e.g., IEEE ICASSP, ISCA Interspeech, EUSIPCO, ISCA Speech Prosody (4 reviews per conference per year in average) and IEEE TASLP, Elsevier Speech Communication, IEEE TAC (2 reviews per year in average)

[2013-present] Regular reviewer for the French National Research Agency (ANR): one project every two years in average.

Teaching activities

Teaching Unit	Since 2013	Since 2019
Creation / Transformation	5 TUs	5 TUs
Responsibility	6 TUs	6 TUs
Lectures	7 CM / 6 TD	5 CM / 3 TD
	8 TP / 2 PR	5 TP / 2 PR

Table 3: Summary of teaching activities (since 2013)
CM = lecture, TD = tutorials,
TP = practical work, PR = project

[2022-present] Head of the Master in Engineering of Intelligent Systems, 80-90 students, Sorbonne Université, Paris, France.

[2022-present] Head of communications for the Masters in Engineering Sciences (SdI) at Sorbonne Université

[2021 - present] Creation and organization of "Deep Voice, Paris", an annual 2/3-day event bringing together students, the general public, and professionals around voice technologies, combining exchanges, round tables, and technical training on the latest developments in the field of speech and natural language processing.

[2021] Founding member of "SOphIA: association des étudiants de SOrbonne université en Intelligence Artificielle", in collaboration with the Sorbonne Center for Artificial Intelligence (SCAI).

[2018 - present] Co-creation, responsibility, and facilitation of the "Deep Learning by Practice" professional continuing education course at Sorbonne University.

Funding

Principal Investigator

[2024-2029] Project PostGenAI. Head: Sorbonne Université, IA Cluster : pôles de recherche et de formation de rang mondial en intelligence artificielle. Scientific coordination for Ircam: Axel Roebel, Nicolas Obin. Total funding: 35 M€

[2021-2023] Project EMERGENCE Sorbonne Université ReVOLT: REvealing human bias with real Time VOcal deep fakes. Real-time neural voice conversion and its application to the study of human cognition and biases. Total funding: 61.5 k€

[2019-2022] Project Ph2D/IDF MoVE: MOdelling of speech attitudes and application to an expressive conversationnal agent. In collaboration with PSA/Stellantis. Total funding: 110 k€

[2017-2021] Project ANR PRCE TheVoice: voice design for the creative industry. In collaboration with: Laboratoire d'Informatique d'Avignon (LIA), Dubbing Brothers. Neural voice identity conversion, voice casting. Total funding: 565 k€

[2015-2016] Project EMERGENCE Sorbonne Université ROUTE: RObot à l'écoUTE, Binaural sound source localization. In collboration with: Institut des systèmes intelligents et de robotique (ISIR). Total funding: 98 k€

[2011-2015] Project FEDER Voice4Games: speech technologies for video games. In collaboration with: ExeQuo, Voxler, Whitebirds. Speech levelling and voice casting for acted speech. Total funding: 628 k€

[2006-2008] Projet PEPS Automatic learning of a generative model of speech prosody.

Scientific coordination (for Ircam)

[2023-2027] Projet ANR EVA : Explicit Voice Attributes. Head: Orange. Speech generation and manipulation conditioned on the representation of attributes. Total funding: 761 k€

[2022-2024] Projet ASTRID DeTOX : Lutte contre les videos hyper-truquees de personalites francaises. Head: EURECOM. Generation of audio-visual deep-fakes with synchonization of audio and video modalities. Total funding: 300 k€

[2022-2026] Projet ANR PRC BRUEL: Rethinking evaluation methodology of speaker authentification systems. Head: Laboratoire d'Informatique d'Avignon (LIA), in collboration with: EURECOM, CEA, and Institut National de Police Scientifique (INPS). Generation of adversarial attacks using voice cloning. Total funding: 698 k€

[2021-2024] Projet ANR PRCE EXOVOICES: Virtual story-telling for Children - Expressive and Cognitive aspects of text-to-speech synthesis. Head: Lunii, in collaboration with: Laboratoire de sciences cognitives et psycholinguistique (LSCP). Expressive and context-aware neural text-to-speech synthesis. Total funding: 740 k€

[2007-2011] Project ANR Rhapsodie: A reference prosodic corpus of spoken French. Responsible for speech to text alignment and prosodic labelling.

Participation

[2019-2023] Projet ANR ARS: Analysis and tRasnformation of singing Style. Singing style modelling and transformation. Head: Ircam, in collaboration with: Lutheries-Acoustique-Musique (LAM), laboratoire Passages Arts Littératures, and Flux.

[2014-2017] Project ANR ChaNTeR: Real-time controlled digital singing. Singing voice synthesis. Head: Laboratoire d'Informatique pour la mécanique et les sciences de l'ingénieur (LIMSI), in collaboration with: Acapela Group and Dualo.

[2005-2008] Project RIAM VIVOS: Creation of expressive voices : multimedia applications. Prosody modelling for speech transformation and synthesis.

PhD supervision, progress and defense committee

Supervision

[2024-2027] Diego Andres Torres Guarrin, Conversion neuronale des attributs de la voix, project ANR EVA, Ecole doctorale informatique, télécommunications et électronique (EDITE).

[2024-2027] Balthazar Bujard, Modèles de couplages entre signaux temporels pour le contrôle de la synthèse sonore, bourse du ministère, Ecole doctorale informatique, télécommunications et électronique (EDITE).

[2023-2026] Téo Guichoux, Génération multimodale du comportement et transfert de style pour l'animation d'un agent virtuel, bourse du ministère, Ecole doctorale informatique, télécommunications et électronique (EDITE).

[2023-2026] Mathilde Abrassart, Conversion neuronale pour la création d'attaques à l'usurpation d'identité, projet ANR BRUEL, Ecole doctorale informatique, télécommunications et électronique (EDITE).

[2023-2026] Théodor Lemerle, Expressive text-to-speech synthesis for storytelling performance, projet ANR EXOVOICES, Ecole doctorale informatique, télécommunications et électronique (EDITE).

[2017-2025] Lisa La Pietra, The independence of the voice from the XXIème siecle. Ethic et aesthetic of the vocal interpretation from Belcanto to artifical intelligence, co-encadrement avec Giordano Ferrari, Université Saint-Denis, 2017.

[2019-2022] Clément Le Moine, Neural conversion of social attitudes in speech signals, collaboration with Stellantis, programme doctoral Ph2D/IDF, Ecole doctorale informatique, télécommunications et électronique (EDITE), 2019. **Defended on February, 27, 2023**

[2019-2022] Mireille Fares, Multimodal expressive gesturing with style, STMS/ISIR, programme doctoral AI @ Sorbonne Université, Ecole doctorale informatique, télécommunications et électronique (EDITE), 2019. **Defended on February, 15, 2023**.

[2019-2022] Killian Martin, Cognitive control of Rooks' vocalizations, STMS/UMR Physiologie de la Reproduction et des Comportements, ED 549 Santé, Sciences Biologiques et Chimie du Vivant, Université de Tours, 2019. **Defended on December, 13, 2022**.

Supervision (unofficial)

[2013-2016] Olivier Migliore, Analyzing the musical prosody of French punk, rap, and ragga (1977-1992) à l'aide de l'outil informatique, Université Montpellier 3 (PhD in musicology, ED 58 Langues, littératures, cultures, civilisations).

Member of follow-up and defense PhD committee

[1] Marco Fiorini (2024). Instruments co-créatifs pour l'interaction musicale improvisée (2023-2026), direction : Gérard Assayag (STMS, Ircam), Ecole doctorale informatique, télécommunications et électronique (EDITE) (ED 130). Member of follow-up committee.

[2] Nicolas Dahan (2024). Evaluer la qualité de la traduction des documents scientifiques (2023-2026), direction : François Yvon (MLIA, ISIR, Sorbonne Université), Ecole doctorale informatique, télécommunications et électronique (EDITE) (ED 130). Member of follow-up committee.

[3] Sahar Husseini (2024). Reconnaissance des locuteurs à partir de la dynamique faciale (2022-2025), direction : Jean-Luc Dugelay (EURECOM), Ecole doctorale informatique, télécommunications et électronique (EDITE) (ED 130). Member of follow-up committee.

[4] Alexandre Libourel (2024). Lutte contre less vidéos hyper-truquées de personnalités françaises (2023-2026), direction : Jean-Luc Dugelay (EURECOM), Ecole doctorale informatique, télécommunications et électronique (EDITE) (ED 130). Member of follow-up committee.

[5] Sewade Ogun (2024). Generating diverse synthetic data for ASR training data augmentation (2021-2024), direction : Emmanuel Vincent et Vincent Colotte (INRIA, Université de Lorraine), Ecole doctorale Mathématiques, Télécommunications, Informatique, Signal, Systèmes, Électronique - Université de Lorraine (ED 77). **Reviewer**.

[6] Olivier Zhang (2023). Méthodes neuronales d'analyse et de synthèse multi aspects et leurs applications à la voix (2020-2023), direction : Nicolas Gengembre Olivier Le Blouch (Orange), Ecole doctorale Informatique Automatique Eletronique Electrotechnique Mathématiques (IAEM) - Université de Rennes (ED 601). **Reviewer**.

[7] Daniel Molina-Villota. Effets Audio Numériques Interactifs et Style Vocal, thèse de doctorat (2020-2024), financement projet anr ArS, direction : Christophe d'Alessandro (LAM), Ecole doctorale Sciences mécaniques, acoustique, électronique et robotique de Paris (ED 391). Member of follow-up committee. **Jury member**.

[8] Mathias Quillot (2022) Un premier pas vers la caractérisation de l'information véhiculée par les voix actées, Laboratoire d'Informatique d'Avignon, thèse de doctorat, Ecole Doctorale Agrosciences et Sciences (ED 536). **Jury member**.

[9] Matthieu Kirchmeyer. Domain generalization for classification and spatio-temporal dynamics, thèse de doctorat (2019-2022), financement : CIFRE Criteo, direction : Patrick Gallinari (LIP6) et Alain Rakotomamonjy (Criteo), Ecole doctorale Informatique, Télécommunications et Electronique de Paris (ED 130). Member of follow-up committee.

[10] Paul-Gauthier Noé. Learning an interpretable speaker representation for attributes manipulaition, thèse de doctorat (2019-2023), financement projet ANR - JST VoicePersonae, direction : Jean-François Bonastre (LIA), Driss Matrouf (LIA), Ecole Doctorale Agrosciences et Sciences (ED 536). Member of follow-up committee.

[11] Théis Bazin. Spectrogram Inpainting for Interactive Generation of Instrument Sounds, thèse de doctorat (2019-2023), financement : CIFRE Sony CSL, direction : Mikhail Malt (STMS), encadrement : Gaëtan Hadjeres (Sony CSL), Ecole doctorale Informatique, Télécommunications et Electronique de Paris (ED 130). Member of follow-up committee.

[12] Andrea Vaglio Automatic lyrics transcription, thèse de doctorat (2018-2021), financement : CIFRE Deezer, Télécom ParisTech et Deezer, direction : Gaël Richard (Telecom ParisTech), Romain Hennequin (Deezer), Ecole Doctorale de l'Institut Polytechnique de Paris (ED 626). Member of follow-up committee.

[13] Giacomo Valenti (2018) Efficient and secure recognition of multiple speakers for mobile devices, authentication, personalisation and contextualisation. CIFRE PhD thesis, Eurecom and NXP Semiconductor. **Jury member**.

[14] Neil Zeghidour, Toward unsupervised speech recognition. Learning of informative and invariant representations. CIFRE PhD thesis, LSCP, Paris Sciences et Lettre / Ecole Normale Supérieure and Facebook A.I Research, ED Cerveau, cognition, comportement. Member of follow-up committee.

[15] Olivier Migliore (2016). Analyzing the musical prosody of French punk, rap, and ragga (1977-1992) à l'aide de l'outil informatique, Université Montpellier 3 (PhD in musicology, ED 58 Langues, littératures, cultures, civilisations). **Jury member**.

Teaching

Since 2022, I am **Head of the Master of Engineering in Intelligent Systems** at Sorbonne Université (Master 1 = around 40 students, Master 2 = around 40 students). This includes students following initial or dual training.

Bachelor's degree (science et technologies, Sorbonne Université)

[2017-2022] Python programming, *Licence 1*, : TD, TP (38h)

[2015-2023] Orientation and Professional Integration, *Licence 1* (30h)

[2015-2020] Signals and systems, *Licence 3*, : TD (10h)

Master's degree (engineering, Sorbonne Université)

[2020-present] Biometrics (with IDEMIA compagny), *Master 2* : **responsible** (2020-), CM (2h), TP (4h).

[2018-present] Advanced machine learning, *Master 2* : **co-responsible** (2020-), CM/TP (14h), project (20h)

[2015-present] Advanced audio processing, *Master 2* : **responsible** until 2019, CM (10h), project (20h)

[2014-present] Collaborative student project, *Master 2* : **co-responsible** (2020-) project monitoring (10h)

[2013-2019] Information theory, *Master 1* : **responsible** (2013-2019), CM, TD, TP (30h)

[2013-present] Digital signal processing, *Master 1* : TD, TP (26h)

[2013-present] Audio signal processing, *Master 1* : **responsible** (2013-2019), CM, TP (22h)

Professional continuing education (Sorbonne Université)

[2018-present] Deep learning in practice: 3 days formation, **creator, co-responsible and instructor** (21h). Six sessions.

External interventions

[2016-2020] Music Technology Program, *Licence*, New-York University (1h)

[2015-2020] Job meetings, *Terminal S*, Lycée Saint-Cyr et autres (1h)

[2014-2020] Human voice and his challenges, *Engineering school*, Télécom ParisTech (1h)

Community engagement

[2022] Member and Ircam representative at Le Voice Lab, the French association for voice technologies. https://www.levoicelab.org/

[2021] Founding member of SOphIA, the Student Association in Artificial Intelligence of Sorbonne University, with the support of the Sorbonne Center for Artificial Intelligence (SCAI) https://sciences.sorbonne-universite.fr/associations/sophia-association-etudiante-intelligence-artificielle-de-sorbonne-universite

[2021] Member of the selection committee for the associate professor position at Ircam/ISIR-Sorbonne Université on "artificial intelligence for robotics or virtual reality" (section 61)

[2020-present] Founding member of the French community on "Culture and Artificial Intellingence", kick-off on 2020, March, 25, Paris

[2019-present] Member of GdR in TAL Traitement automatique des langues, STMS lab. representative

[2018] STMS lab. representative for the AI commission and strategy at Sorbonne Université

[2017-present] Member of the scientific council of the engineering UFR at Sorbonne Université

[2017] Member of the selection committee for the associate professor position at Ircam-UPMC on multi-physical systems and sound signals (section 60-61)

Awards and distinctions

[2011] Best PhD thesis award, Fondation Des Treilles.

Organization of conferences, seminars, workshops...

[2027] ISCA Interspeech, **organizing committee**, Sao Paolo, Brazil, general chairs: Plinio Barbosa and Sandra Madureira. Around 700 participants.

[2025] Fast Forward: Imagining the Future of Sound. **Organization and animation**. Supported by the Programme de Recherche Industries Culturelles et Créatives (PEPR ICCARE), France 2030, and labelled by the AI Action Summit, February, 12, 2025. Ircam, Espace de projection. Around 150 participants.
https://pepr-iccare.fr/events/fast-forward-imaginer-le-futur-du-son/.

[2024] S'éveiller par les sons - Repenser les médiations sonores à l'ère du numérique. **Organization** with: Emmanuelle Zoll and Margot Fuchs (action culturelle, IRCAM) and Frédéric Bevillacqua (ISMM). One half-day of conferences open to professional, and one half-day workshop open to families and children. Around 60 participants for the conference event. November, 13 and 23, 2024.
https://www.ircam.fr/agenda/s-eveiller-par-les-sons/

[2023] ISCA Speech Synthesis Workshop (SSW) and Blizzard Challenge on Text-to-Speech Synthesis, **organizing committee and technical chair**, Grenoble, France, in collaboration with the GIPSA lab (Grenoble) and the IRISA (Rennes). Around 80 participants.
https://ssw2023.org

[2023] IEEE workshop on Socially-Interactive Virtual Agents (SIVA), **organizer and program chair**, Hawaii, USA, as a satellite event of the IEEE International Conference on Automatic Face and Gesture Recognition (FC) and IEEE/CVF Winter Conference on Applications of Computer Vision (WACV).
https://www.stms-lab.fr/agenda/siva/detail

[2022] Deep Voice, Paris 2022. Second edition: diversity and inclusion in speech technologies. Annual scientific event on artificial intelligence and speech technologies. 3 days with 6 sessions (keynote, round table, discussion) involving around 20 international speakers (French and English), and 2 technical workshops on neural speech recognition and neural langage models, Ircam in collaboration with the Sorbonne Center for Artificial Intelligence (SCAI) and fully part as a scientific event during the Manifeste Ircam musical festival.
https://www.sorbonne-universite.fr/evenements/deep-voice-paris

[2022 -] Fast-Forward meetings between research, innovation, and sound design for cinema. **Founder and organizer**, in collaboration with the HAL-audio collective of French sound design artists, Paris France.
https://forum.ircam.fr/agenda/soiree-fast-foward

This event proposes a progressive format for inter-community exchanges between research scientists and engineers, and sound designers and artists in order to think and create the sound of the future. This event has a form of informal meetings and discussion, tutorial and masterclasses on the latest advances in sound research and technologies for creative purposes (3/4 times a year).

[2021] 3rd International Workshop on Vocal Interactivity in-and-between Humans, Animals and Robots (VIHAR), **organizing committee and technical chair**, Paris, France. About 30 participants.
https://vihar-2021.vihar.org/

[2021] Deep Voice, Paris 2021. First edition: diversity and inclusion in speech technologies. Annual scientific event on artificial intelligence and speech technologies. **Creation, and organization**, 2 days keynote,

round-table, and 2 technical workshops on neural speech synthesis, Ircam in collaboration with the Sorbonne Center for Artificial Intelligence (SCAI)

[2019] Sound Design Days (SDD 2019). **Co-founder and co-organizer**, 28-29 november 2019, Ircam, Paris, France. Around 120 participants.
https://www.ircam.fr/agenda/sound-design-days/detail/

[2018] Living Lab Supramuros, Avignon theater festival, TheVoice project, 2019 : supervision.

[2019] Speech, Music and Mind Workshop (SMM 2019) : member of the scientific committee.

[2018] Living Lab Supramuros, Avignon theater festival, invited speaker, TheVoice project, 2018.

[2018] "Workshop on Detection, Classification, Localization, and Density Estimation of Marine Mammals" (DCLDE 2018) : member of the scientific committee.

[2016] Conference "Journées d'Étude de la Parole" (JEP), 300 people, Paris : member of the organization committee, member of the scientific committee.

[2011] Workshop on "Advances in Speech Technologies", Paris : organizer.

[2010] Journées Jeunes Chercheurs en Audition, Acoustique musicale et Signal audio (JJCAS), Paris : member of the organization committee.

[2008] International conferences EMUS (Emotion in MUsic and Speech), Paris (France), Campinas (Brazil), Lyon (France) : member of the organization committee, member of the scientific committee.

Reviewing

[1] Reviewer for the French Research National Agency (ANR, 2008, 2015, 2017, 2021, 2022, 2023)

[2] Expertise pour l'Agence Nationale de la Recherche et de la Technologie (ANRT, 2018)

[3] Reviewer for the musical residency program at Ircam (2011-2015).

Journals

[1] IEEE/ACM Transactions of Audio, Speech, and Language Processing (2013, 2020)

[2] IEEE Transactions on Affective Computing (2015)

[2] Journal of the Audio Engineering Society (JAES, 2021)

[3] Elsevier Speech Communication (2015, 2018 x 2, 2020, 2021), 2023

[4] ACM Journal on Computing and Cultural Heritage (2017)

[5] PLOS ONE (2019)

[6] Journal of the Acoustic Society of America (2019)

[7] Acta Acustica (2017)

[8] International Journal of Multimedia Information Retrieval (2017)

[9] Elsevier Biomedical Signal Processing and Control (2015)

[10] Journal of Speech Sciences (2014)

[11] Springer Journal on Language Resources and Evaluation (2012)

[12] Frontiers in Digital Health (2020)

[13] Advanced Robotics (2023)

Conferences

I am a regular reviewer since 10 years for the following conferences.

[1] International Conference on Representation Learning (ICLR), since 2024

[2] IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)

[3] ISCA Interspeech

[4] European Signal Processing Conference (EUSIPCO)

[5] ISCA Speech Prosody

[6] Journées d'Etude de la Parole (JEP)

I also do on demand review for the following conferences: IEEE Automatic Speech Recognition and Understanding Workshop (ASRU, 2021), International Workshop on Detection, Classification, Localization, and Density Estimation of marine mammals using passive acoustics (2018), International Symposium on Music Information Retrieval (ISMIR, 2017 and 2012), International Conference on Digital Audio Effects (DAFx, 2017),

Implication in artistic projects

I am passionate and deeply committed in the promotion of speech technologies for creation, art, and culture. I have collaborated with renown artists and participated to many productions: composers, musicians, comedians, and film makers (Georges Aperghis, Philippe Manoury, Alexander Schubert, Roman Polansky, Philippe Parreno, Eric Rohmer, Leos Carrax, and André Dussolier). Main selection of artistic collaborations:

[2022 - 2023] *Bestiorobot*, in collaboration with designer Zoé Aegerter and PhD candidate in sciences of information and communication Clothilde Chevet (CELSA). The bestiorobot is an initiative to raise awareness and explain machines and artificial intelligence to classes of children.

[2023] *L'appel du 18 juin*, reconstitution de la voix du Général De Gaulle, en collaboration avec Le Monde.

[2022] *Anima*, by Alexander Schubert, Festival Manifeste, Centre-Pompidou (Paris).

[2022] *Annette*, film by Léos Carax, with sound designer Erwan Kerzanet

[2022] *L'Hôtel du Temps*, interviews with past personalities hosted by Thierry Ardisson using deep fake technologies. First episode : French comedian and singer Dalida.

[2019] *Lullaby experience*, by composer Pascal Dusapin, Frankfurter Positionen festival (Frankfurt) and Théâtre 104 (Paris).

[2018] *Les guerriers de l'ombre*, French documentary directed by Frédéric Schoendoerffer, Canal+. Realistic voice anonymization of secret agents.

[2014-2015] *Pourquoi j'ai pas mangé mon père*, French movie directed by Jamel Debbouze. Recreation of the voice of the French actor Louis de Funès, by using spectral voice identity conversion.

[2014-2015] *Juger Pétain*, TV documentary directed by Richard Saada, Planète, France 5. Recreation of the voice of Maréchal Pétain and other French personalities of the second world war, by using spectral voice identity conversion.

[2013] *Aliados*, opera by composer Sebastian Rivas, Ircam Manifeste festival.

[2012-2013] *Marilyn*, short movie directed by Philippe Parreno, Fondation Beyeler (2012) et Palais de Tokyo (2013). Recreation of the voice of Marilyn Monroe, by using intonation conversion.

[2011] *Luna Park*, by composer Georges Aperghis, Agora festival (Paris).

[2010] *Escort Girl*, first part of the documentary "L'Europe judiciaire" directed by Olivier Ballande, Maha Productions, France Télévision. Realistic speaker anonymization of individuals implied in criminal investigations.

[2009] *Hypermusic Prologue*, opera by composer Hector Parra, Agora festival (Paris).

Disseminating and popularizing scientific culture

Exhibitions

[1] "La Voix" exhibition, Cité des Sciences et de l'Industrie, Paris : member of the organization committee, responsible for the realization of three interactive terminals on voice casting and voice transformation, 2014.

Events

[1] ISCA Interspeech, comité d'organisation, Sao Paolo, Brésil, general chairs: Plinio Barbosa and Sandra Madureira, 2027. Environ 700 participants.

[2] Fast Forward: Imaginer le futur du son. Organisation et animation. Avec le soutien du Programme de Recherche Industries Culturelles et Créatives (PEPR ICCARE), France 2030, et la labélisation du sommet pour l'action sur l'intelligence articielle, 12 février 2025. Ircam, Espace de projection. Environ 150 participants.

[3] S'éveiller par les sons - Repenser les médiations sonores à l'ère du numérique, conférences professionelles et ateliers pour la famillle, co-organisation recherche et action culturelle avec Frédéric Bevilacqua, Emmanuelle Zoll, Marion Voillot, Conférence profesionnelle : Salle Stravinsky, 13 et 23 novembre 2024

[4] Deep Voice, Paris, 2022. Diversité et inclusion dans les technologies vocales. Deep Voice, Paris est l'événement parisien dédié à la voix et à l'intelligence artificielle dont l'objectif est de réunir expertises scientifique et technique, innovation et entrepreneuriat. Des ateliers techniques - "hands-on" - seront proposés aux participants pour développer leurs compétences dans les domaines du traitement de la parole et du langage naturel. Des moments de rencontres et de réseautage en fin de journée faciliteront les rencontres et l'échange entre les participants en toute décontraction. Deep Voice est un événement co-organisé par l'Ircam, Sorbonne Université, et SCAI dans le cadre du festival ManiFeste 2022, 15-17 juin 2022.

[5] International French-Japan Forum Intelligence Augmentation and Amplification + Society (CNRS, EHESS, JST), Invited research scientist, Paris, France, 2022. This Franco-Japanese summit brings together researchers in artificial intelligence, robotics and social sciences, I was invited to participate as a researcher on the theme "Working & Living in a Cyber Physical AI world: which interactions between human being and machine?".

[6] Deep Voice, Paris 2021. Assistants, robots affectifs et réalité virtuelle, Assistants, robots affectifs, et réalité virtuelle : Comment la simulation des émotions va-t-elle modifier notre rapport affectif aux machines et notre immersion dans un monde social artificiel ?, Co-organisé avec le soutien du Sorbonne Center for Artificial Intelligence (SCAI) et de l'Ircam, 23-24 juin 2021.

Public lectures and events

[1] Utilisation de l'IA en post-production audio, table ronde avec Jean-Peïc Chou et David Raichman, Les Monteurs Associés, Paris 18 décembre 2024

[2] Les usages de l'IA pour l'éducation musicale ?, à l'invitation de Thierry Rolando A-IPR en Education musicale du rectorat de l'académie de Paris, 16 octobre 2024

[3] Challenges and opportunities in audiovisual creation in the era of generative AI, FAIR, in the framework of +RAIN Film Festival, SONAR, Barcelona, 11-12 juin 2024

[4] Les enjeux de l'IA dans l'industrie cinématographique. Table ronde avec Mathilde Croze, Festival NUMOK, Bibliothèque François Truffaut, Paris, 2 mai 2024

[5] L'intelligence artificielle : quels impacts sur la création et les métiers du cinéma et de l'audiovisuel ? Table ronde avec Nicolas Becker, Pierre Lanchantin, Lex Dromgoole, et Erwan Kerzanet, La Femis, Paris, 2 avril 2024.

[6] Voix et sons à l'ère de l'intelligence artificielle. Table ronde avec Nicolas Muchielli et Erwan Kerzanet, La Cinémathèque française, Paris, 13 mars 2024.

[7] L'IA au service du sonore ? La semaine du Son, UNESCO, Paris, 18 janvier 2024

[8] Les usages médiatiques du deepfake : rendre la parole aux morts, une alternative aux archives audiovisuelles ?, Les Rendez-vous de l'histoire, Blois, 6 octobre 2023,

[9] Ethics and generative AI. Animateur de la table ronde avec Ambre Davat, Patrick Kuban, et Jeannette Gorzala, Speech Synthesis Workshop 2023, Machine that talks, Grenoble, 27 août 2023

[10] The Sound of Uncertainty, Round table on Music and AI with Nao Tokui and Yotam Mann, Sonar Music Festival, 16 juin 2023

[11] Entre nature et culture : l'intelligence artificielle pour interpréter un monde de sons, La nature et ses musiques, Exposition et Colloque, Philarmonie de Paris, Jeudi 19 janvier 2023

[12] La voix humaine: acoustique, signal, apprentissage - de la modélisation à la création, Association des Amis du Laboratoire Arago de Banyuls-sur-Mer, 2022

[13] La manipulation dans le discours politique : manipulation par le discours et manipulation du discours, Gouverner, Les Napoléons, Val d'Isère, 14 janvier 2022

[14] Deep fakes audio : entre manipulation et créativité, VoiceTech Paris, Paris, France, 23-24 novembre 2020.

[15] AI and creativity, Paris-Atlanta webinar, Paris, France, 22-23 octobre 2020.

[16] Sculpter la voix : algorithmes, IA, et créativité, Sound Design Days, Paris, France, 29 novembre 2019.

[17] Trucages, deepfake, clonage vocal : en 2020, ne croyez plus vos yeux et vos oreilles, Futurapolis, Toulouse, France, 16 novembre 2019. With Vincent Nozick, associate professor at Paris-Est Marne la Vallée University, and Guerric Poncet, journalist at Le Point.

[18] Voix artificielle et sound design : concepts, pratiques, enjeux (avec N. Misdariis, de l'équipe Perception et Design Sonore, Ircam), HyperVoix : enjeux de conception des interfaces conversationnelles, Journée d'étude organisée par la Fing, Le Square, Paris, 15 avril 2019.

[19] Creative industries and AI, invited speaker by Goldsmiths University of London, University of London Institute, France, 21 juin 2019.

[20] AI and creativity: can AI be creative?, Changing IP in Changing Europe, Trilateral Seminar of the French, German and Polish Groups of AIPPI, Paris, France, 05 avril 2019.

[21] Assistants, falsification, anonymisation : la voix à l'ère du numérique et de l'IA, CNIL, Paris, France, 15 février 2019.

[22] La voix et l'IA, Parcours Innovateur ignorant, Sciences Po, Paris, France, 19 octobre 2018.

[23] Marilyn: histoire de la recréation de la voix d'une icône, Ecole Normale Supérieure, Paris-Saclay, France, 9 octobre 2018.

[24] Les voix du numérique pour le design sonore et la création artistique, Festival Vino Voce, Saint-Emilion, France, 8-10 septembre 2017.

[25] Vox ex-machina. La voix et ses doubles numériques, présentation invité de la chaire de création artisique de Philippe Manoury, Collège de France, Paris, France, 6 juin 2017.

[26] La voix fantôme, Exposition Polyphonies, Centre Pompidou, Paris, France, Décembre 2016.

[27] Voix d'outre-tombe : recréer les voix du passé à partir d'archives sonores, Philarmonie de Paris, Paris, France, Novembre 2015.

[28] Le futur du traitement de la voix, Future en Seine, Paris, France, Juin 2014.

[29] Machine théâtrale et nouvelles technologies, Je ne suis pas scientifique, mais... Forum Science, Recherche, et Société - Conservatoire National des Arts et Métiers, Paris, France, Mai 2014.

[30] Les technologies de traitement de la voix, La Semaine Du Son, Paris, France, Janvier 2014.

Articles in newspaper

[1] IA et synthèse vocale : les nouvelles voix de la création. Dossier Science et IA : quelles (r)évolutions ? Revue de l'ANR #3, Janvier 2025.

[2] Nicolas Obin, Pierre Saint-Germier, Jean-Louis Giavitto, Frank Madlener, Axel Roebel, Frédéric Bevilacqua, Pour une intelligence artificielle responsable au service d'une création musicale, inventive et diverse, Tribune

publiée dans la revue Culture et Recherche 147 - Recherche et intelligence artificielle, p. 34-35, automne-hiver 2024.

[3] Réalité et fantasmes de l'IA au cinéma, Comment imaginer le cinéma dans 20 ans ?, Critikat, Décembre 2024.

[4] Clones, filtres et fakes…Éthique et IA, L'étincelle — Le journal de la création à l'Ircam #24, Manifeste 2024.

[5] Deep fake : la guerre des algorithmes. Dossier Espilon. Avril 2024

[6] Affaire Taylor Swift, Usbek & Rica, 6 février 2024

[7] Intelligence artificielle : «Aujourd'hui, on parvient à créer une voix de synthèse à partir de quelques secondes de votre voix», Libération, 12 octobre 2023

[8] Entretien avec l'Association Française du Son à l'Image (AFSI). 26 mars 2023

[9] L'intelligence artificielle pour modifier sa voix : « Pour un être humain, la différence va devenir impossible à détecter », Le Monde, 17 mars 2023

[10] La recherche à l'assaut des deep fakes vocaux, Isabelle Belin, Data Analytics Post, 15 décembre 2022

[11] Des voix toujours plus naturelles … et à la carte, Isabelle Belin, Data Analytics Post, 22 septembre 2022

[12] Déchiffrée, contrefaite, modifiée : les métamorphoses de la voix, Florence Rosier, Le Monde, 4 novembre 2019.

[13] À l'Ircam, la voix transformée. Dossier L'économie de la voix Usine Nouvelle, juillet 2019.

[14] Les prouesses des assistants vocaux. Dossier Intelligence Artificielle. Ça m'intéresse, février 2019.

[15] Comment fabrique-t-on une voix de synthèse ?, La Croix, 22 Janvier 2019.

[16] La voix sacrée, dossier voix et IA, Télérama, 16 Janvier 2019.

[17] Les chanteurs disparus ressuscités par la high-tech, L'express, 21 Décembre 2018.

[18] Des voix de synthèse presque humaines, Pour la Science, Décembre 2016, numéro 470, p. 54-62.

[19] Clonage vocal : avec l'Ircam. Juger Pétain : le son et l'image enfin réunis. SonoMag, Mai 2015

[20] On va pouvoir ressusciter la voix de Marilyn Monroe, Sciences et Vie, Juin 2015, p. 104-108.

[21] J'ai cherché Audrey, la synthèse vocale de mon ordinateur, Rue 89, 5 mai 2015.

[22] De Funès dans le film de Debbouze ? Mais c'est la voix de qui ? Rue 89, 8 avril 2015.

[23] "Pourquoi j'ai pas mangé mon père" : l'Ircam refait parler Louis de Funès, Sciences et Avenir, 8 avril 2015.

Participation to broadcast programs

[1] Côté Club, avec Benjamin Epps, Uzi Freyja et Nicolas Obin, Côté Club, France Inter, animé par Laurent Goumarre, 10 février 2025.

[2] Histoire et IA, voix recrées, archives falsifiées ?, La marche du monde, animé par Valérie Nivelon, Radio France International (RFI), en direct de la Bibliothèque Nationale de France, dans le cadre de la semaine sur l'action pour l'IA, 8 février 2025.

[3] Musique et sciences : l'hymne à la voix, avec Nathalie Henrich-Bernardoni (GIPSA-lab) et Gaël Richard (Télécom Paris), France Culture, 30 décembre 2024.

[4] Journal de 07h30 Intervention en direct sur SONAR music festival à Barcelone, France Inter, 16 juin 2024

[5] Musique et IA, avec Whim Therapy (artiste) et Michael Turbot (Sony CSL) Côté Club, France Inter, animé par Laurent Goumarre, 4 mai 2023.

[6] Peut-on se laisser séduire par une voix artificielle comme dans HER?, podcast 7ème Science, avec Binge Audio, 13 juillet 2022

[7] Deepfake : le vrai du faux d'une technologie révolutionnaire, Le Meilleur des Mondes, France Culture, 13 Mai 2022, animé par François Saltiel

[8] A l'Ircam, la salle la plus silencieuse au monde dévoile ses secrets, journal télévisé, 20h, TF1, 9 septembre 2016.

[9] Donner de la Voix, FutureMag, Arte, 5 décembre 2015.

[10] Carte blanche d'Eva Bettan, A'live, France Inter, 18 Juin 2015.

[11] La Visite Musicale, Le Magazine de la Santé, France 5, 27 mars 2015.

[12] Pianotine, Visite de l'Institut de Recherche sur l'Acoustique et la Musique, Arte, 30 Novembre 2014.

[13] Petite Histoire de la Synthèse Vocale, France Culture, 12 Août 2014.

[14] L'Informatique cherche ses Voix, France Culture, 11 Août 2014.

[15] On a Volé la Voix d'André Dussolier, Le Monde, 7 Juillet 2014.

[16] Parler avec la Voix d'un Autre, On n'est pas que des cobayes, France 5, 2014.

[17] Voix et Identité, Universcience, Arte Creative, 2014.

[18] Vers une autre Voix, France Culture Plus, 2014.

[19] Les Intonations, France Culture Plus, David Christoffel, 2013.

Publications

My publication indexes: **h-index: 22**, **i10-index : 41** (source: google scholar, 02/2025)).

Articles in peer-reviewed journals (rank A)

[1] K. Martin, F. M. Cornero, N. S. Clayton, O. Adam, N. Obin, V. Dufour (2024). Vocal complexity in a socially complex corvid: gradation, diversity and lack of common call repertoire in male rooks, in Royal Society Open Science 11(1), January 2024. Impact factor : 3.653

[2] M. Fares, C. Pelachaud and N. Obin (2023). Zero-Shot Style Transfer for Gesture Animation driven by Text and Speech using Adversarial Disentanglement of Multimodal Style Encoding, in Frontiers in Artificial Intelligence, vol. 6, 12 Jun 2023. Inpact factor : 3.000

[3] L. Benaroya, N. Obin and A. Roebel (2023). Manipulating Voice Attributes by Adversarial Learning of Structured Disentangled Representations, in Entropy 2023, 25(2), 18 Feb 2023. Inpact factor : 2.738

[4] K. Martin, O. Adam, N. Obin, V. Dufour (2022). Rookognise: Acoustic detection and identification of individual rooks in field recordings using multi-task neural networks, submitted to Ecological Informatics. Impact factor: 4.498.

[5] L. Benaroya, N. Obin, M. Liuni, A. Roebel, W. Raumel, S. Argentieri (2018). Binaural Localization of Multiple Sound Sources by Non-Negative Tensor Factorization. IEEE/ACM Transactions on Audio, Speech and Language Processing, vol. 26, no 6, p. 1072-1082, February 2018. Facteur d'impact : 2.491

[6] N. Obin, A. Roebel (2016). Similarity Search of Acted Voices for Automatic Voice Casting. IEEE/ACM Transactions on Audio, Speech and Language Processing, vol. 24, no 9, p. 1638-1647, September 2016. Facteur d'impact : 1.877

[7] N. Obin, P. Lanchantin (2015). Symbolic Modelling of Speech Prosody: From Linguistics to Statistics. IEEE/ACM Transactions on Audio, Speech and Language Processing, vol.23, no. 3, p. 588-599, March 2015. Facteur d'impact : 1.877

Articles in peer-reviewed journals (others)

[8] M. Avanzi, N. Obin, A. Lacheret and B. Victorri (2011). Vers une modélisation continue de la structure prosodique. Le cas des proéminences accentuelles. Journal of French Languages Studies, Vol. 1 (21), p. 53-71, 2011.

Articles in peer-reviewed journals (French)

[9] O. Migliore, N. Obin, J.Bresson (2019). Outils et méthodes de transcription pour l'analyse de l'interprétation vocale dans le rap :l'exemple d'IAM concept. Revue Musimédiane [en ligne], Moreno Andreatta [dir.], 2019.

[10] G. Beller, C. Veaux, G. Degottex, N. Obin, P. Lanchantin and X. Rodet (2009). IrcamCorpusTools : plateforme pour les corpus de parole. Traitement Automatique des Langues, Vol. 3 (49), 2009.

Article in peer-reviewed conferences (international, rank A to B)
Source : http://www.conferenceranks.com

[1] T. Guichoux, L. Soulier, N. Obin, C. Pelachaud (2024) 2D or not 2D: How Does the Dimensionality of Gesture Representation Affect 3D Co-Speech Gesture Generation?, In ACM Intelligent Virtual Agents (IVA), Glasgow, Scotland.

[2] T. Lemerle, N. Obin, A. Roebel (2024). Small-E: Small Language Model with Linear Attention for Efficient Speech Synthesis, in ISCA Interspeech, Kos Island, Greece

[3] C. Le Moine, V. Rosi, L. Salais, P. Arias Sarah, N. Obin (2024) BWSNET: Automatic Perceptual Assessment of Audio Signals, in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

[4] W. Phokhinanan, N. Obin, S. Argentieri (2024). Auditory Cortex-Inspired Spectral Attention Modulation for Binaural Sound Localization in HRTF Mismatch, in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)

[5] M. Fares, N. Obin and C. Pelachaud (2023). TranSTYLer: Multimodal Behavioral Style Transfer for Facial and Body Gestures Generation, Submitted to ACM International Conference on Multimodal Interaction (ICMI), Paris, France, 2023

[6] W. Phokhinanan, N. Obin, S. Argentieri (2023). Binaural Sound Localization in Noisy Environments Using Frequency-Based Audio Vision Transformer (FAViT), in ISCA Interspeech, Dublin, Ireland, 2023.

[7] L. Salais, P. Arias, C. Le Moine, V. Rosi, Y. Teytaut, N.Obin, A. Roebel (2022). Production Strategies of Vocal Attitudes, in ISCA Interspeech, Incheon, Korea, 2022.

[8] F. Bous, L. Benaroya, N. Obin, A. Roebel (2022) Voice Reenactment with F0 and timing constraints and adversarial learning of conversions, in European Signal Processing Conference (EUSIPCO), Belgrade, Serbia, 2022

[9] M. Fares, C. Pelachaud, N. Obin Transformer Network for Semantically-Aware and Speech-Driven Upper-Face Generation, in European Signal Processing Conference (EUSIPCO), Belgrade, Serbia, 2022

[10] C. Le Moine, N.Obin, A. Roebel (2021). Speaker Attentive Speech Emotion Recognition, in ISCA Interspeech, Brno, Czech Republic, 2021.

[11] C. Le Moine, N.Obin, A. Roebel (2021). Towards End-to-End F0 Voice Conversion based on Dual-GAN with Convolutional Wavelet Kernerls, in European Signal Processing Conference (EUSIPCO), Dublin, Ireland, 2021.

[12] R. Ferro, N. Obin, A. Roebel (2020). CycleGAN Voice Conversion of Spectral Envelopes using Adversarial Weights, European Signal Processing Conference (EUSIPCO), Amsterdam, Netherlands, 2020.

[13] M. Fares, N. Obin, C. Pelachaud (2020). Towards Multimodal Human-Like Characteristics and Expressive Visual Prosody in Virtual Agents, ACM International Conference on Multimodal Interaction (ICMI), Utrecht, Netherlands, 2020.

[14] C. Robinson, N.Obin, A. Roebel (2019). Sequence-To-Sequence Modelling of F0 for Speech Emotion Conversion, International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Brighton, UK, 2019.

[15] D. Bouvier, N.Obin, M. Liuni, A. Roebel (2016). A Source/Filter Model with Adaptive Constraints for NMF-based Speech Separation International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Shanghai, China, 2016.

[16] R. Gong, P. Cuvillier, N.Obin, A. Cont (2015). Real-Time Audio-to-Score Alignment of Singing Voice Based on Melody and Lyric Information. Interspeech, Dresden, Germany, 2015.

[17] X. Favory, N.Obin, G. Degottex, A. Roebel (2015). The Role of Glottal Source Parameters for High-Quality Transformation of Perceptual Age. International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Brisbane, Australia, 2015.

[18] G. Degottex, N.Obin (2014). Phase Distortion Statistics as a Representation of the Glottal Source: Application to the Classification of Voice Qualities. Interspeech, Singapore, Singapore, 2014.

[19] N.Obin, A. Roebel, G. Bachman (2014). On Automatic Voice Casting for Expressive Speech: Speaker Recognition vs. Speech Classification. International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Florence, Italy, 2014.

[20] N.Obin, F. Lamare, A. Roebel (2013). Syll-O-Matic: an Adaptive Time-Frequency Representation for the Automatic Segmentation of Speech into Syllables. International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Vancouver, Canada, 2013.

[21] N.Obin (2012). Cries and Whispers - Classification of Vocal Effort in Expressive Speech. Interspeech, Portland, USA, 2012.

[22] J. Lorenzo-Trueba, R. Barra-Chicote, T. Raitio, N. Obin, P. Alku, J. Yamagishi, J. M. Montero (2012). Towards Glottal Source Controllability in Expressive Speech Synthesis. Interspeech, Portland, USA, 2012.

[23] M. Avanzi, P. Dubosson, S. Schwab, N.Obin (2012). Accentual Transfer from Swiss-German to French: A Study of "Français Fédéral". Interspeech, Portland, USA, 2012.

[24] N. Obin, P. Lanchantin, A. Lacheret and X. Rodet (2011). Reformulating Prosodic Break Model into Segmental HMMs and Information Fusion. Interspeech 2011, Florence, Italy, 2011.

[25] N. Obin, P. Lanchantin, A. Lacheret and X. Rodet (2011). Discrete/Continuous Modelling of Speaking Style in HMM-based Speech Synthesis: Design and Evaluation. Interspeech 2011, Florence, Italy, 2011.

[26] N. Obin, A. Lacheret and X. Rodet (2011). Stylization and Trajectory Modelling of Short and Long Term Speech Prosody Variations. Interspeech 2011, Florence, Italy, 2011.

[27] M. Avanzi, N. Obin, Anne Lacheret-Dujour and B. Victorri (2011). Toward a Continuous Modeling of French Prosodic Structure: Using Acoustic Features to Predict Prominence Location and Prominence Degree. Interspeech 2011, Florence, Italy, 2011.

[28] N. Obin, A. Lacheret and X. Rodet. HMM-based Prosodic Structure Model Using Rich Linguistic Context. Interspeech 2010, p. 1133-1136, Makuhari, Japan, 2010.

[29] N. Obin, V. Dellwo, A. Lacheret and X. Rodet. Expectations for Discourse Genre Identification: a Prosodic Study. Interspeech 2010, p. 3070-3073, Makuhari, Japan, 2010.

[30] N. Obin, X. Rodet and A. Lacheret-Dujour. A Multi-Level Context-Dependent Prosodic Model Applied to Durational Modeling. Interspeech 2009, p. 512-515, Brighton, UK, 2009.

[31] N. Obin, A. Lacheret-Dujour, C. Veaux, X. Rodet and A.-C. Simon. A Method for Automatic and Dynamic Estimation of Discourse Genre Typology with Prosodic Features. Interspeech 2008, p. 1204-1207, Brisbane, Australia, 2008.

[32] N. Obin, X. Rodet and A. Lacheret-Dujour. French Prominence: A Probabilistic Framework. International Conference on Acoustics, Speech, and Signal Processing (ICASSP), p. 3993-3996, Las Vegas, USA, 2008.

Article in peer-reviewed conferences (international, others)

[1] C. Pelachaud, N. Obin, M. Fares (2023). Behavior Generation Model for Socially Interactive Agents, Gesture and Speech in Interaction (GeSpIn), Nijmegen, Netherlands, 2023.

[2] M. Fares, N. Obin, C. Pelachaud (2023). I-Brow: Hierarchical and Multimodal Transformer Model for Eyebrows Animation Synthesis, International Conference on Human-Computer Interaction (HCI), Copenhagen, Denmark, 2023.

[3] C. Le Moine, N. Obin, A. Roebel (2020). Att-HACK: An Expressive Speech Database with Social Attitudes, ISCA International Conference on Speech Prosody, Tokyo, Japan, 2020.

[4] L. Liu, A. Lacheret-Dujour, N. Obin (2019). Acoustic Modelling and Labelling of Speech Prosody: What's new with SLAM+? International Congress of Phonetic Sciences (ICPhS), Melbourne, Australia, 2019.

[5] N.Obin, J. Belião (2018). Sparse Coding of Pitch Contours with Deep Auto-Encoders. Speech Prosody, Poznań, Poland, 2018.

[6] O. Migliore, N.Obin (2018). At the Interface of Speech and Music: A Study of Prosody and Musical Prosody in Rap Music. Speech Prosody, Poznań, Poland, 2018.

[7] R. Gong, N. Obin, G. Dzhambazov and X. Serra (2017). Score-Informed Syllable Segmentation for Jingju a Cappella Singing Voice with Mel-Frequency Intensity Profiles. International Workshop on Folk Music Analysis (FMA), Malaga, Spain, 2017.

[8] N.Obin, J. Belião, C. Veaux, A. Lacheret (2014). SLAM: Automatic Stylization and Labelling of Speech Melody. Speech Prosody, Dublin, Ireland, 2014.

[9] A. Lacheret, J. Belião, A. Dister, K. Gerdes, J.-P. Goldman, S. Kahane, N. Obin, P. Pietrandrea, A. Tchobanov (2014). Rhapsodie: a Prosodic-Syntactic Treebank for Spoken French. Language Resources and Evaluation Conference (LREC), Reykjavik, Iceland, 2014.

[10] N.Obin, M. Liuni (2012). On the Generalization of Shannon Entropy for Speech Recognition. IEEE workshop on Spoken Language Technology, Miami, USA, 2012.

[11] N.Obin, C. Veaux, P. Lanchantin (2012). Making Sense of Variations: Introducing Alternatives in Speech Synthesis. Speech Prosody, Shanghai, China, 2012.

[12] N.Obin, M. Avanzi, G. Bordal, A. Bardiaux (2012). Regional Variations of Speech Rhythm in French: In Search of Lost Times. Speech Prosody, Shanghai, China, 2012.

[13] G. Bordal, M. Avanzi, N.Obin, A. Bardiaux (2012). Variations in the Realization of the French Accentual Phrase in the Light of Language Contact. Speech Prosody, Shanghai, China, 2012.

[14] M. Avanzi, A. Lacheret, N.Obin (2012). Transcription of Prosody in Continuous Speech: Prosodic Encoding the Spoken French Corpus Rhapsodie. Advancing Prosodic Transcription for Spoken Language Science and Technology, LabPhon, Stuggart, Germany, 2012.

[15] P. Lanchantin, S. Farner, C. Veaux, G. Degottex, N. Obin, G. Beller, Stefan Huber (2011). Vivos Voco: a Survey of Recent Research on Voice Transformations at IRCAM. International Conference on Digital Audio Effects, Paris, France, 2011.

[16] N. Obin, M. Avanzi, and A. Lacheret (2011). Transcription of French Prosody: the Rhapsodie Protocole. Interface Discours et Prosodie 2011, Manchester, UK, 2011.

[17] M. Avanzi, G. Bordal and N. Obin (2011). Typological Variations in the Realization of French Accentual Phrase. International Congress of Phonetic Sciences (ICPHS), Hong-Kong, China, 2011.

[18] N. Obin, P. Lanchantin, M. Avanzi, A. Lacheret-Dujour and X. Rodet (2010). Toward Improved HMM-based Speech Synthesis Using High-Level Syntactical Features. Speech Proosdy 2010, Chicago, USA, 2010.

[19] N. Obin, X. Rodet and A. Lacheret-Dujour (2009). A Syllable-Based Prominence Detection Model Based on Discriminant Analysis and Context-Dependency. International Conference on Speech and Computer (SPECOM), p. 97-100, St-Petersburg, Russia, 2009.

[20] G. Beller, N. Obin and X. Rodet (2008). Articulation Degree as a Prosodic Dimension of Expressive Speech. Speech Prosody, Campinas, Brazil, 2008.

Article in peer-reviewed conferences (French)

[1] T. Guichoux, L. Soulier, N. Obin, C. Pelachaud (2024). Investigating the impact of 2D gesture representation on co-speech gesture generation, in Workshop Affects, Compagnons Artificiels et Interactions (WACAI), Bordeaux, France. **Prix du meilleur article**

[2] M. Quillot, L. Guillou, A. Gresse, R. Ferro, R. Roth, D. Malinas, R. Dufour, A. Roebel, N. Obin, J.-F. Bonastre, E. Ethis (2020). La voix actée : pratiques, enjeux, applications. Journées d'Etude de la Parole (JEP-TALN-RECITAL), Nancye, France, 4-8 juin 2018.

[3] N. Obin , P. Pham, A. Roebel (2018). Conversion d'Identité de la Voix Chantée par Sélection et Concaténation d'Unités Spectrales. Journées d'Etude de la Parole, Aix-en-Provence, France, 4-8 juin 2018.

[4] A. Roebel, L. Ardaillon, N. Obin (2018). Expressive Singing Synthesis using an Extended Source Filter Model. Congrès Français d'Acoustique, Le Havre, France, 23-27 Avril 2018.

[5] O. Migliore, N. Obin (2016). Analyser la prosodie musicale : l'exemple du rap français des débuts (1984-1991), La voix dans les chansons, Paris, France, 2016.

[6] O. Migliore, N. Obin, J. Bresson (2014). Rap, Ragga et Rock Alternatif Français des Années 1980 : Vers une Musicologie Computationnelle. Journées d'Analyse Musicale (JAM), Paris, France, 2014.

[7] A. Lacheret, J. Belião, A. Dister, K. Gerdes, J.-P. Goldman, S. Kahane, N. Obin, P. Pietrandrea, A. Tchobanov (2014). Rhapsodie: un Treebank annoté pour l'Etude de l'Interface Syntaxe-Prosodie en Français Parlé. Congrès Mondial de Linguistique Française (CMLF), Berlin, Germany, 2014.

[8] N.Obin, M. Avanzi, G. Bordal, A. Bardiaux (2012). A la Recherche des Temps Perdus : Variations sur le Rythme en Français. Journées d'Etude de la Parole, Grenoble, France, 2012.

[9] M. Avanzi, N.Obin, G. Bordal, A. Bardiaux (2012). La Variation Prosodique Dialectale en Français. Données et Hypothèses. Journées d'Etude de la Parole, Grenoble, France, 2012.

[10] N. Obin, Modélisation du Style en Synthèse de Parole par HMM, Journées Jeunes Chercheurs en Audition, Acoustique musicale et Signal audio, Paris, France, 2010.

[11] M. Avanzi, N. Obin, A. Lacheret-Dujour. Vers une modélisation continue de la structure prosodique du français : le cas des proéminences accentuelles. Colloque Conscila, Paris, France, 2009.

[12] N. Obin, X. Rodet and A. Lacheret-Dujour. Un modèle de durée des syllabes fondé sur les propriétés syllabiques intrinsèques et les variations locales de débit. Journées d'étude de la parole (JEP), p. 333-336, Avignon, France, 2008.

[13] N. Obin, J.-P. Goldman, M. Avanzi and A. Lacheret-Dujour. Comparaison de trois outils de détection automatique de proéminences en français parlé. Journées d'étude de la parole (JEP), p. 85-88, Avignon, France, 2008.

Chapters in edited monographs

[1] N. Obin, A. Lacheret, C. Veaux, B. Victorri (2019). Tonal Annotation: Stylization of Complex Melodic Contours over Arbitrary Linguistic Segments. Rhapsodie: a prosodic-syntactic Treebank for Spoken French, Benjamins, Amsterdam, 2019.

[2] M. Avanzi, N. Obin, G. Bordal, J. Sauvage (2019). The Annotation of Prominences and Disfluencies. Rhapsodie: a prosodic-syntactic Treebank for Spoken French, Benjamins, Amsterdam, 2019.

[3] A. Lacheret, P. Pietrandrea, O. Baude, N. Obin, A.-C. Simon, A. Tchobanov (2019). Collecting data for the Rhapsodie Treebank. Rhapsodie: a prosodic-syntactic Treebank for Spoken French, Benjamins, Amsterdam, 2019.

[4] N. Obin, C. Veaux, P. Lanchantin (2015). Exploiting Alternatives for Text-To-Speech Synthesis: From Machine to Human. Speech Prosody in Speech Synthesis: Modeling and generation of prosody for high quality and flexible speech synthesis Chapter 3: Control of Prosody in Speech Synthesis, p.189-202, Springer Verlag, February, 2015.

[5] X. Rodet, G. Beller, N. Bogaards, G. Degottex, S. Farner, P. Lanchantin, N. Obin, A. Roebel, C. Veaux and F. Villavicencio. Transformation et Synthèse de la Voix Parlée et de la Voix Chantée. Parole et Musique: Aux Origines du Dialogue Humain, Odile Jacob, Paris, 2009.