



# Activity Report 2022

## Team EXPRESSION

Expressiveness in Human Centered Data/Media

D6 – Signal, Image, Language





## 1 Team composition

### 1.1 Composition

#### Head of the team

Pierre-François Marteau, Professor, Université Bretagne Sud

#### Administrative assistants

Anne Idier, Université Bretagne Sud

Martine Milcent, Université Bretagne Sud

Angélique Le Pennec, Université de Rennes 1

Joëlle Thépault, Université de Rennes 1

#### Permanent members

Pierre Alain, Research Engineer, Université de Rennes 1

Nelly Barbot, Associate professor, Université de Rennes 1

Vincent Barraud, Associate professor, Université de Rennes 1

Nicolas Béchet, Associate professor, Université Bretagne Sud

Giuseppe Bério, Professor, Université Bretagne Sud

Jonathan Chevelu, Associate professor, Université de Rennes 1

Arnaud Delhay-Lorrain, Associate professor, Université de Rennes 1

Sylvie Gibet, Professor, Université Bretagne Sud

Caroline Larboulette, Associate professor, Université Bretagne Sud

Damien Lolive, Professor, Université de Rennes 1

Gildas Ménier, Associate professor, Université Bretagne Sud

Jeanne Villaneau, Associate professor (emeritus), Université Bretagne Sud

#### External collaborator

Elisabeth Delais-Roussarie, Senior researcher, CNRS/LLF

Farida Said, Associate professor, Université Bretagne Sud

Lauren Thevin, Université Catholique de l'Ouest

#### Non-permanent members

⇒ Simon Giddings, Engineer, Université de Rennes 1 (from February 2022 to September 2022)

Gaëlle Vidal, Engineer, Université de Rennes 1 (from April 2019)

Rashedur Rahman, Post-doctoral researcher, Université de Rennes 1 (from May 2020)

← Felipe Verdugo, Post-doctoral researcher, University of Mac Gill, Université Bretagne Sud (July 2019 to June 2022)

- ⇔ Aghilas Sini, Post-doctoral researcher, University of Rennes 1 (until January 2022 and from October 2022)
- Antoine Perquin, Post-doctoral researcher, University of Rennes 1 (from December 2021)
- Clément Reverdy, ATER, Université de Bretagne Sud (since November 2022)
- David Guennec, contractual assistant professor, Université de Rennes (since September 2022)

### PhD students

- Danrun Cao, Université Bretagne Sud, CIFRE OctopusMind 2nd year
- Alexis Blandin, Université Bretagne Sud, CIFRE UNEEK 3rd year (from January 2020)
- ← Esso-Ridah Bléza, Université Bretagne Sud, CIFRE, defended in December 2022
- ← Tiago Brizolar, Université Bretagne Sud, ARED/CD56
- Mansour Tchénégnon, Université Bretagne Sud, CDE/ARED 3rd year
- ← Betty Fabre, Université de Rennes 1, CIFRE Orange Labs, defended in September 2022
- Somaye Jafaritazehjani, Université de Rennes 1, CDE/TU Dublin, 3rd year
- ← Jade Mekki, Université de Rennes 1, ANR TREMoLo, defended in September 2022
- Clémence Mertz, Université Rennes 1, CDE/LTC, 3rd year
- Lily Wadoux, Université de Rennes 1, CDE 2nd year
- Olivier Zhang, Université de Rennes 1, CIFRE Orange Labs, 2nd year
- Hoan My TRAN, Université de Rennes 1, CD22+CDE-IA, since november 2022

### Master students

- ⇔ Rania Bennetayeb, Master in Sciences and Technologies of the of the Faculty of Electrical Engineering of the University of Science and Technology Houari Boumediene, Alger, Algeria

## 1.2 Evolution of the staff

The permanent staff has been stable during the year. The number of PhD students is decreasing with 3 PhD defenses, against 1 newly hired PhD student.

## 2 Overall objectives

Expressivity or expressiveness are terms which are often used in a number of domains. In biology, they relate to genetics and phenotypes, whereas in computer science, expressivity of programming languages refers to the ability to formalize a wide range of concepts. When it comes to human expressivity, we will consider the following reading: expressivity is the way a human being conveys emotion, style or intention. Considering this definition, the EXPRESSION team focuses on studying human language data conveyed by different media: gesture, speech and text. Such data exhibit an intrinsic

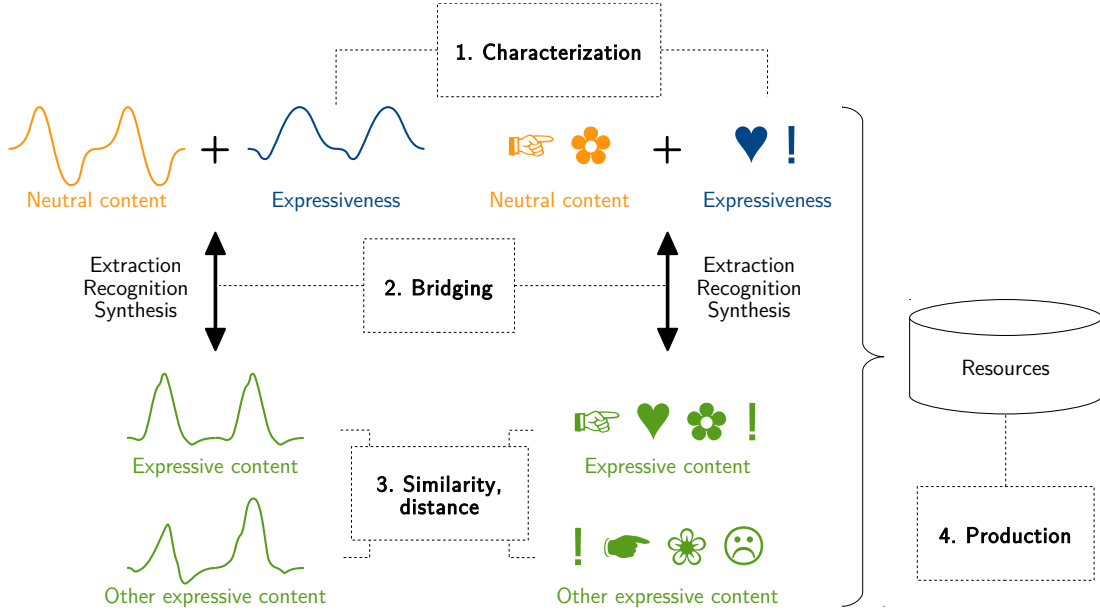


Figure 1: Overview of the main challenges considering both on continuous numerical (left) and discrete symbolic (right) data.

complexity characterized by the intrication of multidimensional and sequential features. Furthermore, these features may not belong to the same representation levels - basically, some features may be symbolic (e.g., words, phonemes, etc.) whereas others are digital (e.g., positions, angles, sound samples) - and sequentiality may result from temporality (e.g., signals).

Within this complexity, human language data embed latent structural patterns on which meaning is constructed and from which expressiveness and communication arise. Apprehending this expressiveness, and more generally variability, in multidimensional time series, sequential data and linguistic structures is the main proposed agenda of EXPRESSION. This main purpose comes to study problems for representing and characterizing heterogeneity, variability and expressivity, especially for pattern identification and categorization.

The research project targets the exploration and (re)characterization of data processing models in three mediated contexts:

1. Expressive gesture analysis, synthesis and recognition,
2. Expressive speech analysis and synthesis,
3. Expression in text and language.

## 2.1 Main challenges addressed by the team

Four main challenges will be addressed by the team.

- C1:** The characterization of the expressiveness as defined above in human produced data (gesture, speech, text) is the first of our challenges. This characterization is challenging jointly the extraction, generation, or recognition processes. The aim is to develop models for manipulating or controlling expressiveness inside human or synthetic data utterances.
- C2:** Our second challenge aims at studying to what extent innovative methods, tools and results obtained for a given media or for a given pair of modality can be adapted and made cross-domain. More precisely, building comprehensive bridges between discrete/symbolic levels (meta data, semantic, syntactic, annotations) and mostly continuous levels (physical signals) evolving with time is greatly stimulating and nearly not explored in the different scientific communities.
- C3:** The third challenge is to address the characterization and exploitation of data-driven embeddings<sup>1</sup> (metric or similarity space embeddings) in order to ease post-processing of data, in particular to reduce the algorithmic complexity and meet the real-time or big-data challenges. The characterization of similarity in such embeddings is a key issue as well as the indexing, retrieval, or extraction of sub-sets of data relevant to user's defined tasks and needs, in particular the characterization of expressiveness and variability.
- C4:** The fourth challenge is to contribute to the production of resources that are required, in particular to develop, train and evaluate machine learning (statistical or rule-based) models for human language data processing. These resources are mainly corpora (built from speech, text and gesture time series), dictionaries, and semantic structures such as ontologies.

All the addressed challenges are tackled through the development of models, methods, resources and software tools dedicated to represent and manage gesture, speech or textual data. Thus we consider a complete processing chain that includes the creation of resources (corpus, thesaurus, semantic network, ontology, etc.), the labeling, indexing and retrieval, analysis and characterization of phenomena via classification and extraction of patterns (mostly sequential).

These challenges also target multi-level aspects, from digital tokens to semantic patterns, taking into account the complexity, the heterogeneity, the multi-dimensionality, the volume, and the nature of our temporal or sequential data.

We are aiming at addressing these challenges in terms of development and exploitation of machine learning and pattern discovery methods for clustering, classification, interactive control, recognition, and production of content (speech signals, texts or gestures), based on different levels of representation (captured or collected data but also knowledge that is specific to the media or the considered application). Finally, both objective and subjective (perceptive) evaluations of these models are a key issue of the research directions taken by the EXPRESSION team.

---

<sup>1</sup>Given two metric or similarity spaces  $(X, d)$  and  $(X', d')$ , a map  $f : (X, d) \rightarrow (X', d')$  is called an embedding.

## 2.2 Main research focus

Five thematic lines of research are identified to carry out this research.

**RF1: Data acquisition** – Gesture, speech or text data are characterized by high levels of heterogeneity and variability. Studying such media requires high quality data sets appropriate to a well defined and dedicated task. The data acquisition process is thus a crucial step since it will condition the outcomes of the team research, from the characterization of the studied phenomena, to the quality of the data driven models that will be extracted and to the assessment of the developed applications. The production of high quality and focused corpora is thus a main issue for our research communities. This research focus addresses mainly the fourth challenge;

**RF2: Multi-level representations** – We rely on multi-level representations (semantic, phonological, phonetic, signal processing) to organize and apprehend data. The heterogeneity of these representations (from metadata to raw data) prevents us from using standard modeling techniques that rely on homogeneous features. Building new multi-level representations is thus a main research direction. Such representations will provide efficient information access, support for database enrichment through bootstrapping and automatic annotation. This research focus contributes mainly to the second, third and fourth challenges;

**RF3: Knowledge extraction** – This research addresses data processing (indexing, filtering, retrieving, clustering, classification, recognition) through the development of distances or similarity measures, rule-based or pattern-based models, and machine learning methods. The developed methods will tackle symbolic data levels (semantic, lexical, etc.) or time series data levels (extraction of segmental units or patterns from dedicated databases). This research focus contributes mainly to the first and third challenges.

**RF4: Generation** – We are also interested in the automatic generation of high-quality content reproducing human behavior on two modalities (gesture and speech). In particular, to guarantee adequate expressiveness, the variability of the output has to be finely controlled. For gesture, statements and actions can be generated from structural models (composition of gestures in French sign language (LSF) from parameterized linguistic units). For speech, classical approaches are data-driven and rely either on speech segment extraction and combination, or on the use of statistical generation models. In both cases, the methods are based at the same time on data-driven approaches and on cognitive and machine learning control processes (e.g., neuromimetic). This research focus contributes mainly to the first and fourth challenges since generation can be seen also as a bootstrapping method. As parallels can be possibly drawn between expressive speech and expressive movement synthesis, the focus also contributes to the second challenge;

**RF5: Use cases and evaluation** – The objective is to develop intuitive tools and in particular sketch-based interfaces to improve or facilitate data access (using different modes of indexing, access content, development of specific metrics, and graphical interfaces), and to integrate our aforementioned models into these tools.

As such, this focus contributes to the first challenge and has a direct impact on the fourth challenge. Furthermore, whereas many encountered sub-problems are machine learning tasks that can be automatically evaluated, synthesizing human-like data requires final perceptive (*i.e.* human) evaluations. Such evaluations are costly and developing automatic methodologies to simulate them is a major challenge. In particular, one axis of research directly concerns the development of cross-disciplinary evaluation methodologies. This research focus contributes also to the second challenge;

### 3 Scientific achievements

#### 3.1 New Results by Key Issues

In accordance with the Team Project, the main outcomes for 2022 are listed into the following key issues items defined above for the team:

Use cases  
and eval-  
uation,  
Knowledge  
extraction

**Discrete-Events Simulation for Teaching Statistics in Industrial Engineering [3]** This study presents a discrete events simulation tool developed to support undergraduate students in their Statistics and Data Analysis course. Although the use of modern smart technologies in the industry contributes to a profusion of data, very few enterprise datasets are freely available, resulting in a serious lack of open real-world data for research and education. To overcome this difficulty, we designed a tool that simulates scheduling scenarios in a manufacturing environment. The generated data may be used to put statistical concepts and methods into practice to design cost-effective strategies for optimizing key performance indicators, such as reducing production time, improving quality, eliminating wastes, and maximizing profits. Keywords: industrial datasets, teaching statistics, discrete events simulation

Use cases  
and eval-  
uation,  
Knowledge  
extraction

**COSMOS: Experimental and Comparative Studies of Concept Representations in Schoolchildren [16]** COSMOS is a multidisciplinary research project investigating schoolchildren's beliefs and representations of specific concepts under control variables (age, gender, language spoken at home). Seven concepts are studied: friend, father, mother, villain, work, television and dog. We first present the protocol used and the data collected from a survey of 184 children in two age groups (6-7 and 9-11 years) in four schools in Brittany (France). A word-level lexical study shows that children's linguistic proficiency and lexical diversity increase with age, and we observe an interaction effect between gender and age on lexical diversity as measured with MLR (Measure of Lexical Richness). In contrast, none of the control variables affects lexical density. We also present the lemmas that schoolchildren most often associate with each concept. Generalized linear mixed-effects models reveal significant effects of age, gender, and home language on some concept-lemma associations and specific interactions between age and gender. Most of the identified effects are documented in the child development literature. To better understand the process of semantic construction in



children, additional lexical analyses at the n-gram, chunk, and clause levels would be helpful. We briefly present ongoing and planned work in this direction. The COSMOS data will soon be made freely available to the scientific community.

**Pollen risk levels prediction from multi-source historical data [8]** Numerous studies show that meteorological conditions have an impact on the emission, dispersion and suspension of pollens in the air. Several allergenic species permanently threaten the health of millions of people in France and one can extrapolate that this is the case in most part of the world. Hence, preventive information on the risk of pollen exposure would become a real asset for allergy sufferers. The main objective of this article is to study, thanks to statistical learning techniques exploiting historical data and meteorological parameters of the day (T0), the ability to predict three days in advance (T0 + 3) the risk levels of pollen presence in the air on a given territory (in metropolitan France). We are interested in the prediction of risk -discretized in four levels- for three families of pollens which are among the most allergenic species (Ambrosia, Cupressaceae and grasses). The aggregation of binary logistic regression models for each level of risk by means of a random forest classifier allowed us to predict the level of pollen risk with performances in the range of 75% concerning mainly low and medium risk levels. The comparative study with other approaches shows that our model leads to the best predictive performances while remaining computationally efficient and explainable.

Use cases and evaluation, Knowledge extraction

**Graphical document representation for french newsletters analysis [7]** Document analysis is essential in many industrial applications. However, engineering natural language resources to represent entire documents is still challenging. Besides, available resources in French are scarce and do not cover all possible tasks, especially in specific business applications. In this context, we present a French newsletter dataset and its use to predict the good or bad impact of newsletters on readers. We propose a new representation of newsletters in the form of graphs that consider the newsletters' layout. We evaluate the relevance of the proposed representation to predict a newsletter's performance in terms of open and click rates using graph analysis methods.

Use cases and evaluation, Knowledge extraction

**DaFNeGE: Dataset of French Newsletters with Graph Representation and Embedding [6]** Natural language resources are essential for integrating linguistic engineering components into information processing suites. However, the resources available in French are scarce and do not cover all possible tasks, especially for specific business applications. In this context, we present a dataset of French newsletters and their use to predict their impact, good or bad, on readers. We propose an original representation of newsletters in the form of graphs that take into account the layout of the newsletters. We then evaluate the interest of such a representation in predicting a newsletter's performance in terms of open and click rates using graph convolution network models.

Use cases and evaluation, Knowledge extraction

**Chinese public procurement document harvesting pipeline [9]** We present a processing pipeline for Chinese public procurement document harvesting, with the aim of producing strategic data with greater added value. It consists of three micro-modules: data collection, information extraction, database indexing. The information extraction part is implemented through a hybrid system which combines rule-based and machine learning approaches. Rule-based method is used for extracting information with presenting recurring morphological features, such as dates, amounts and contract awardee information. Machine learning method is used for trade detection in the title of procurement documents.

Use cases  
and eval-  
uation,  
Knowledge  
extraction

**Open challenges for Machine Learning based Early Decision-Making research [2]** More and more applications require early decisions, i.e. taken as soon as possible from partially observed data. However, the later a decision is made, the more its accuracy tends to improve, since the description of the problem to hand is enriched over time. Such a compromise between the earliness and the accuracy of decisions has been particularly studied in the field of Early Time Series Classification. This paper introduces a more general problem, called Machine Learning based Early Decision Making (ML-EDM), which consists in optimizing the decision times of models in a wide range of settings where data is collected over time. After defining the ML-EDM problem, ten challenges are identified and proposed to the scientific community to further research in this area. These challenges open important application perspectives, discussed in this paper.

Generation,  
Use cases  
and evalua-  
tion

**A New Spatio-Temporal Loss Function for 3D Motion Reconstruction [14]** We propose a new loss function that we call Laplacian loss, based on spatio-temporal Laplacian representation of the motion as a graph. This loss function is intended to be used in training models for motion reconstruction through 3D human pose estimation from videos. It compares the differential coordinates of the joints obtained from the graph representation of the ground truth against the one of the estimation. We design a fully convolutional temporal network for motion reconstruction to achieve better temporal consistency of estimation. We use this generic model to study the impact of our proposed loss function on the benchmarks provided by Human3.6M. We also make use of various motion descriptors such as velocity, acceleration to make a thorough evaluation of the temporal consistency while comparing the results to some of the state-of-the-art solutions.

Generation

**A Greedy Algorithm for Generative String Art [10]** Computational String Art is a method of creating a non-photorealistic rendering using lines drawn between pairs of pins. Given a set of pins and an image to render, we propose a greedy algorithm to determine a good set of lines to imitate the input. Implementation details are discussed, as well as some possible optimization to speed up the algorithm. We also propose some improvements to the basic greedy approach, provide a stopping condition and allow for the use of non-opaque strings. We illustrate our work with various results.

**Voice conversion: characters conversion in audiobooks [13]** Audiobook readers play with their voices to emphasize some text passages, highlight discourse changes or significant events, or in order to make listening easier and entertaining. A dialog is a central passage in audiobooks where the reader applies significant voice transformation, mainly prosodic modifications, to realize character properties and changes. However, these intra-speaker modifications are hard to reproduce with simple text-to-speech synthesis. The manner of vocalizing characters involved in a given story depends on the text style and differs from one speaker to another. In this work, this problem is investigated through the prism of voice conversion. We propose to explore modifying the narrator’s voice to fit the context of the story, such as the character who is speaking, using voice conversion. To this end, two complementary experiments are designed: the first one aims to assess the quality of our Phonetic PosteriorGrams (PPG)-based voice conversion system using parallel data. Subjective evaluations with naive raters are conducted to estimate the quality of the signal generated and the speaker similarity. The second experiment applies an intra-speaker voice conversion, considering narration passages and direct speech passages as two distinct speakers. Data are then non parallel and the dissimilarity between character and narrator is subjectively measured. Results are promising and show that voice conversion methods are a track to explore further to convert indirect style to direct style. This is one of the rare studies to apply state of the art systems to voice conversion in the French language.

Use cases  
and eval-  
uation,  
Generation

**DISPEECH: a new dataset for speech disentangling [19]** Recently, a growing interest in unsupervised learning of disentangled representations has been observed, with successful applications to both synthetic and real data. In speech processing, such methods have been able to disentangle speakers’ attributes from verbal content. To have a better understanding of disentanglement, synthetic data is necessary, as it provides a controllable framework to train models and evaluate disentanglement. Thus, we introduce diSpeech, a corpus of speech synthesized with the Klatt synthesizer. Its first version is constrained to vowels synthesized with 5 generative factors relying on pitch and formants. Experiments show the ability of variational autoencoders to disentangle these generative factors and assess the reliability of disentanglement metrics. In addition to provide a support to benchmark speech disentanglement methods, diSpeech also enables the objective evaluation of disentanglement on real speech, which is to our knowledge unprecedented. To illustrate this methodology, we apply it to TIMIT’s isolated vowels.

Data ac-  
quisition,  
Multi-level  
representa-  
tions

**Voice Cloning Applied to Voice Disorders [17, 18]** Organic dysphonia can lead to vocal impairments. Recording patients’ impaired voice could allow them to use voice cloning systems. In the domain of speech synthesis, voice cloning is the process of producing speech matching a target speaker voice, given textual input and an audio sample from the speaker. It can achieve high-quality speech with only few data from the target speaker. However, dysphonic patients may only produce speech with specific

Generation

or limited phonetic content. To our knowledge, the impact of such constraints on a voice cloning system remains to be studied. This article presents the results of preliminary experiments on the matter, along with specifications about the models and datasets used.

Data acquisition,  
Multi-level  
representations

**Low-Cost motion capture corpus in French sign language [12]** The automatic translation of sign language videos into transcribed texts is rarely approached in its whole, as it implies to finely model the grammatical mechanisms that govern these languages. The presented work is a first step towards the interpretation of French sign language (LSF) by specifically targeting iconicity and spatial referencing. We have built the LSF-SHELVES corpus which contains utterances in LSF of this type. Our final goal is to use deep learning methods to circumvent the use of models in spatial referencing recognition. In order to obtain training material with sufficient variability, we designed a lightweight (and low-cost) capture protocol that enabled us to collect data from a large panel of LSF signers. This protocol involves the use of a portable device providing a 3D skeleton, and of a software developed specifically for this application to facilitate the post-processing of wrist movements and handshapes. The LSF-SHELVES includes simple and compound iconic and spatial dynamics, organized in 6 complexity levels, representing a total of 60 sequences signed by 15 LSF signers.

Data acquisition

**Large Breton speech corpus [11]** Breton is a minority language spoken in the Brittany region of France. Public initiatives are being undertaken in order to preserve the Breton language. As an effort toward that goal, we created a large Breton speech corpus and related automatic annotation tools. The corpus contains 20 hours of reading aloud for both a male and a female Breton speaker. Then, end-to-end text-to-speech synthesis systems are built. Subjective evaluation suggests that the systems are able to reproduce the voices of the original speakers faithfully.

Knowledge  
extraction,  
Generation

**A Processing Chain to Explain the Complexity of Texts for Children From a Linguistic and Psycho-linguistic Point of View [4]** This work addresses the issue of measuring the complexity of a text with respect to a target group of readers, namely children of reading age, through the implementation of a processing chain. This chain aims at extracting linguistic descriptors, mainly from psycholinguistics and readability studies, that can be used to understand and describe the complexity of a text. By applying it to a corpus of fiction texts, it allows us to study the correlations between various linguistic descriptors and the age ranges associated with the texts by the editors. The analysis of these correlations tends to validate the relevance of the age categorization by the editors. It thus justifies the mobilization of such a corpus to train a prediction model to recommend a text to the children of a target age on the basis of the editors' ages. Following these results, we have worked on sentence paraphrasing for the children of target ages. To rewrite a sentence suitable for children of target ages, we have explored constraint-controlled paraphrase generation as a seq-to-seq (encoder-decoder) approach. In this task, we define a single constraint, the

target age for paraphrase generation. Due to the lack of sufficient paraphrase dataset in french, this task begins with using ParaCota corpus consisting of synthetically generated french paraphrased sentence pairs. To train a paraphrase generation model with the age constraint, firstly, the paraphrased sentence pairs are labeled with ages by leveraging our age prediction model. Then we have studied two models: the LSTM-based encoder-decoder and the Transformer-based encoder-decoder model. Both models can generate paraphrased sentences to some extent but badly fails to meet our requirements according to a limited subjective evaluation. However, these two experiments help us to get some insights and to plan for further progress. Currently, we are exploring how to take advantage of the pretrained large language models to generate paraphrases for the target ages.

### Summary of the contributions

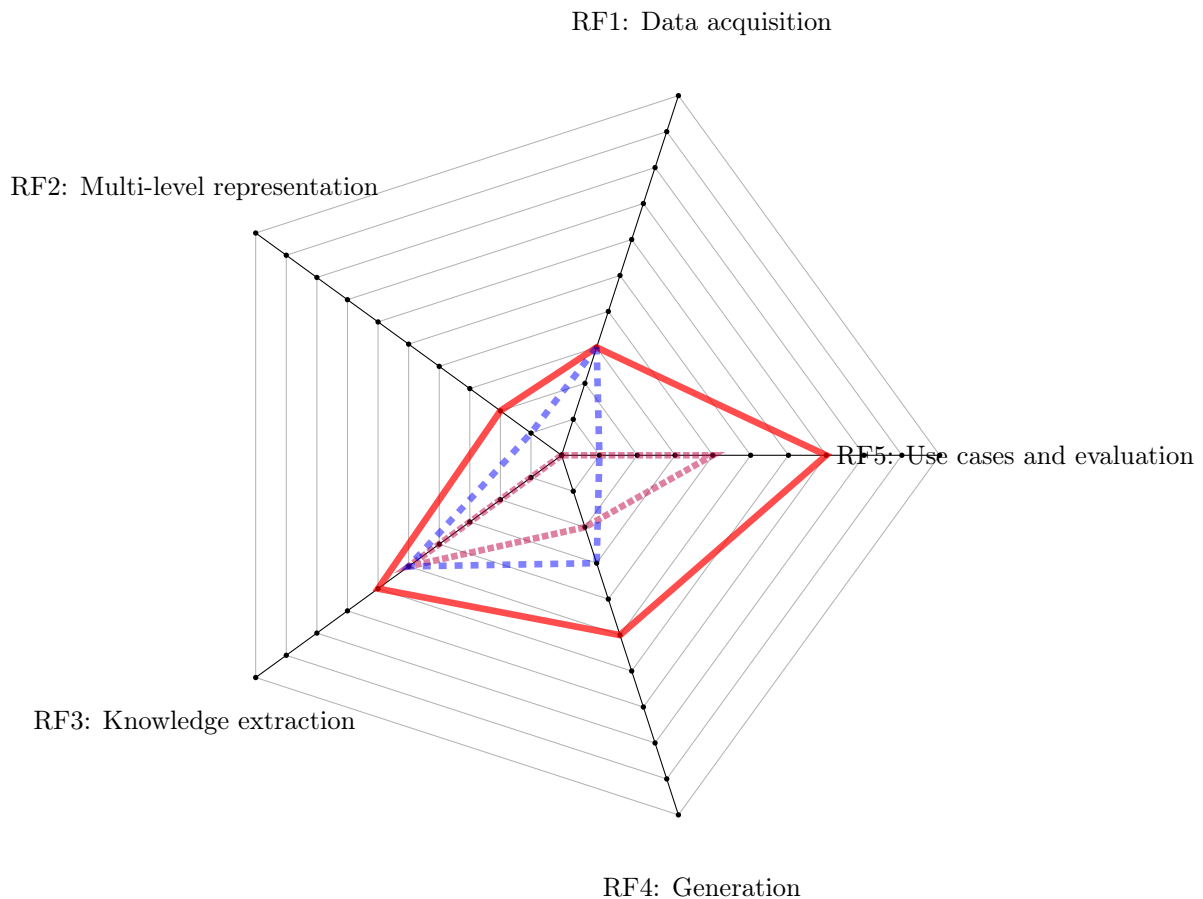


Figure 2: Contributions to each research focus of the team in 2021 (red, solid) compared to 2020 (purple, dotted) and 2019 (blue, dashed)

### 3.2 Defended PhDs and HDRs

- Esso-Ridah Bl'eza has defended his PhD on the 13th of December 2022.

- Jade Mekki has defended his PhD on the 8th of September 2022.
- Betty Fabre has defended his PhD on the 16th of September 2022.

### 3.3 On going PhDs

1. Esso-Ridah Bl'eza has pursued his PhD according to his new CIFRE contract established with the start-up LifyAir. His research subject addresses the prediction of the risk of pollen emission all over the national territory. As such, Esso-Ridah has developed spatio-temporal prediction models based on machine learning approaches capable of effectively exploiting complex historical data. Esso-Ridah Bl'eza will defend his PhD the 13th of December 2022.
2. Alexis Blandin has pursued his PhD work (3rd year) in the context of a CIFRE contract with UNEEK Company. He has mostly focused his work on emailing campaign problems. More precisely, he has addressed the design of a labeled dataset extracted from the CRM of UNEEK company to evaluate the capability to predict the opening and click ratios of newsletters. The objective of his work is to help the design of newsletter in order to maximize its expected impact. His work has led to publications in TextMine/EGC 2022 workshop[5], TSD2022[6] and ACM DocEng 2022[7] conferences.
3. Mansour Tchénégnon has started his third year of PhD. He proposed a new loss function for training 3D motion reconstruction model from RGB videos. This loss function is based on the Laplacian representation of the motion as a 3D+t graph. From this, differential coordinates based on spatial and temporal connections in 3D poses can be computed. These results have been published and presented at European Conference on Computer Vision (ECCV 2022), Workshop on What is Motion for? [14]. He is currently developing an approach to correct 3D motion by using deep learning algorithms based on dual information – 3D poses and variational Laplacian coordinates –, and he will apply his results on motion of people with motor disabilities, where some data is not available.
4. Danrun Cao has started her PhD in the context of an industrial CIFRE contract with Octopus Mind, located in Nantes. He work will address information extraction (entities and relationships) in the context of business intelligence, while exploiting multilingual call for tenders. Danrun has published during her first year a conference paper in the ACM DocEng conference. She is currently developing state of the art Deep Learning approaches to Name Entity Recognition and semantic relation extraction for the French language.
5. Clémence Mertz has completed her third PhD year in 2022. She designed a corpus based on the concepts of spatiality and iconicity in LSF (see LSF-SHELVES) and built the corresponding dataset. This grammatical corpus, the corresponding capture protocol, and the motion processing was presented at the International Conference on Language Ressources and Evaluation (LREC 2022) [12]. She then initiated a reflection on the use of Deep Learning techniques applied to the translation of LSF in written Pivot language.

6. Somaye Jafaritazehjani has started her PhD in November 2018. She is supervised by Damien Lolive, Gwénolé Lecorvé, and John D. Kelleher (from TU Dublin, Ireland). Somaye works on style transfer in natural language processing using deep learning, that is the task to shift a text from one style to another. During her first year, she has built a bibliographical review of the domain, and started extending a state-of-the-art approach. By doing so, she has also investigated the issue of objective evaluation for this natural language generation task. From her second year she is working at TU Dublin. She is expected to defend her PhD beginning of 2023.
7. Lily Wadoux has completed her second PhD year during 2022. She works on “Voice preservation: adaptation of voice cloning to pathological cases”. During this year, she worked on the bibliographical part and also on the pipeline that should be set up for voice cloning. Indeed, she has got more familiar with the tools and the Jean Zay platform which is used to learn the models. She conducted first experimental evaluations of the pipeline (informal evaluation of the speech output). She participated to the Advanced Language Processing Winter School held in early January 2022. She pursued her work on voice cloning during 2022 and published two papers in conferences.
8. Olivier Zhang has started his PhD in October 2020 under a CIFRE contract funded by Orange Labs. The subject of his PhD is “Multi-aspect neural analysis and synthesis methods and their application to voice”. This PhD is directed by Damien Lolive and co-supervised by Olivier Le Blouch and Nicolas Gengembre from Orange Labs.
9. Hoan My Tran started his PhD in November 2022 on the detection of deep speech fakes in the context of voice conversion and speech synthesis.

## 4 Software development

**MappEMG pipeline** The *MappEMG pipeline* [15] aims to augment the traditional classical concert experience by giving listeners access, through the sense of touch, to an intimate and non-visible dimension of the musicians’ bodily experience while performing. The live-stream MappEMG pipeline produces vibrations based on muscle activity captured through surface electromyography (EMG). It allows the audience to experience the performer’s muscle effort, an essential component of music performance which is typically unavailable to direct visual observation.

**LSF-SHELVES** The *LSF-SHELVES* corpus is a French Sign Language Corpus developed during Clemence Mertz’s thesis (2021–2022). This corpus implements the grammatical mechanisms of spatial referencing and iconicity in French Sign Language (LSF), selected through examples of positioning objects in relation to each other on shelves. It proposes to produce utterances with increasing levels of difficulty. A novel, low-cost motion acquisition technique based on the use of a Microsoft Kinect was used to collect data sets from 14 speakers, whereas conventional protocols use motion capture studios. An additional editing step was required to incorporate manual configura-

tions and correct wrist orientation trajectories (software developed by Motion Up, see <https://www.motion-up.com/>).

**MiS** The *MiS* corpus has been recorded in June 2022 and has been designed to study the *style* in motion. It consists of a set of walks, a sensory path, jumps and waving movements, in a controlled setting with a choice of participants of various ages (from 8 to 69 years old), sex, weight and size. The originality of this dataset is the style variations on imposed movements and the diversity of subjects. The aim is to use this data to analyse the style in movements and study the foot-floor contacts.

**Text-To-Speech system** In the frame of several technology transfer projects, developments have been done in 2019 on many pieces of software involved in the team's text-to-speech system to make it usable in industrial environments. This has mainly consisted in shifting some tools from script languages (Python, Perl, shell) to C++, and setting up web services on a production server, compatibility for Android and ARM7 architectures. This has been mainly achieved in the frame of the SPAM project (maturation project funded by the transfer technology service of University of Rennes 1). Developments are going on from then in the frame of different projects such as NADINE or KALYGO-DYS.

**Visual Quantum Simulator [Gildas M nier]** Most of the human behaviors involve uncertainty and fuzzy inner decisions in a way or another : Expression can be seen as a individual by-product as the top level visible artifacts of some unknown inner process. Markov models, or statistical process can be used to mimics the observable results of observable behaviors. Quantum State evolution shares also many similar features with the obfuscated or hidden way of thinking and communicating.

Not only does this (relatively) new paradigm bring some ideas related to human decision's process, it also may be a key progress for Machine Learning - Quantum Machine Learning or QML -. Some intensive computation involved in ML may be sped-up using Quantum schemes.

We are investing some efforts in the study of Quantum Computing both as a computing potential asset and source of ideas for human inner process simulation.

A Scala DSL (domain specific) language has been developed to manage Quantum Algorithm Evaluations (<https://github.com/gmenier/VisualQuantumSimulator/wiki>)

The software has been registered by *APP* (Agence pour la Protection des Programmes) under the number Inter Deposit Digital Number IDDN.FR.001.300006.000.S.C.2021.000.10000 (06.56.11046).

This simulator is designed to allow a graphical view of the inner processing involved in Quantum computing, thus helping grasping its inner working.

It is also designed as a starting backbone for experimentation and is actually used as a support for education.



## 5 Contracts and collaborations

### 5.1 National Initiatives

#### 5.1.1 TREMoLo ANR project

**Participants:** Nicolas Béchet, Jonathan Chevelu, Jade Mekki,.

EXPRESSION is leader of the ANR project TREMoLo, which has been accepted in December 2016. The project studies the use of language registers and seeks to develop automatic methods towards the transformation of texts from a register to another. To do so, the project proposes to extract linguistic patterns which discriminate a register from another, and to integrate them into a probabilistic automatic paraphrase generation process. The language under study is French.

This project is mainly supported by IRISA, coordinated by Gwénoél Lecorvé and involves a member of MoDyCo (UMR 7114 Modèles, Dynamiques, Corpus), Delphine Battistelli.

Up-to-date information are available at <https://tremolo.irisa.fr>.

#### 5.1.2 TextToKids ANR Project

**Participants:** Nicolas Béchet, Jonathan Chevelu, Damien Lolive, Alexis Blandin, Rashedur Rahman.

The TextToKids ANR project is the continuation of the homonymous CNRS PEPS project, previously running in 2018. It aims to facilitate the writing and the filtering of texts for children, especially but not only in order to tell them about current events (e.g., presidential elections, Brexit, reception of migrants in France, etc.) in respect of their language skills. The targeted age group is that of young readers, that is, the 7-12 age group. The consortium, which brings together linguists, computer scientists and specialized journalists, will seek to characterize the linguistic constraints to be respected for such a purpose and to propose assistive tools (automated textual analysis, search engine, reformulation, good practices). In terms of benefits, the project works in the direction of a "children's Internet" and opens a way to other modalities (speech, images).

This project is coordinated by Delphine Battistelli (MoDyCo lab), assisted by Gwénoél Lecorvé.

#### 5.1.3 Breton Synthesis

**Participants:** Damien Lolive, Antoine Perquin, David Guennec, Gaëlle Vidal.

We conducted experiments aiming to deepen and complement works on Breton language, as a continuation of the project funded by *OPLB - Office Public de la Langue Bretonne*, which led in 2021 to build speech resources and a speech synthesis engine

based on neural network modeling. In June 2022 an evaluation was conducted to measure how systems were able to reproduce the voices of the original speakers faithfully, and their overall quality[11]. Also initiated by OPLB, a new project ran from September to November 2022 to produce two new voices (female and male) using voice adaptation, with only 3 hours of recording for each.

#### 5.1.4 MSHB Project SADSui

**Participants:** Nicolas Béchet.

The SADSui research project consists of an analysis of suicide prevention chat corpus in France (corpus of 10 years) and in Japan (corpus in constitution). Analytical work is interdisciplinary, combining manual (language sciences), automated (computing) and content (psychology) analyses. It is a question of analyzing the dysfunctions and the conditions of happiness in these particular interactions.

#### 5.1.5 GIS Marsouin Project SADSui

**Participants:** Nicolas Béchet.

The SADSui project is an interdisciplinary analysis (language sciences - computer sciences - psychology - sociology) of a digital device for the prevention of suicide by cat by means of a tooled discourse analysis (lexicometry) of a large cat corpus ( 2005-2021), supplemented by surveys and interviews with users and volunteers of the digital prevention system.

#### 5.1.6 Kaligo+

**Participants:** Damien Lolive, Simon Giddings.

This project aims at proposing a set of digital application on tablets and using virtual reality tools (like Oculus Quest) to favor training children with oral and written troubles. In particular, our role is to work on the oral part of the problem by proposing Text-To-Speech technologies to make an oral feedback to users, and pronunciation analysis tools to develop pronunciation training methods. This project begun in 2022 and is expected to finish in August 2024.

#### 5.1.7 Koralie - Plan de relance

**Participants:** Damien Lolive, Antoine Perquin, Aghilas Sini.

This project aims at developing pronunciation analysis and speech synthesis tool for different language in the context of education. This project is funded by *France Relance* and is registered under the record number 2021\_000583. it will last 24 months and started the 1st December 2021.

## 5.2 National Collaborations

- We are collaborating actively with Marie Tahon, associate professor at LIUM in Le Mans, France. We shared common interests on emotions annotation, and generation in the context of speech synthesis.
- In the frame of the TREMoLo and TextToKids projects, the team works with members of the MoDyCo lab, especially Delphine Battistelli, full professor at University of Paris-Nanterre. 2 MSc interns have been co-supervised (Alexis Blandin, Aline Étienne), while Jade Mekki defended her PhD in September 2022.
- Nelly Barbot and Aghilas Sini have collaborated with Laurent Miclet (former) emeritus professor at University of Rennes 1 about the supervised classification based on analogical proportions.
- Caroline Larboulette collaborates with Laura Pouppeville, an artist from the "ateliers Montebello" in Lyon, on the project Motion in Style, that has been awarded a grant from the IRISA project "1 artiste - 1 chercheur". The aim of the project, that started in december 2021, is to collect and analyse data, and use this data to produce an art piece that shall be presented at the "fete de la science" or similar venues. They collected data in june 2022. The data is still being processed and analyzed.
- Since december 2021, Caroline Larboulette collaborates with Paul Richard (associate professor at Polytech' Angers, LARIS lab) and Jérémy Besnard (associate professor at Université d'Angers, laboratoire de Psychologie des Pays de la Loire) on the capture and synthesis of avatars, for the study of social interactions in Virtual Reality.

## 5.3 International Collaborations

- In 2017, we have developed a collaboration with Ingmar Steiner and Sébastien Le Maguer from Saarland University, Saarbruck, Germany. Notably, we recruited an internship to work together on the construction of a common interface for Speech synthesis systems enabling to visualize and interact with several systems, like Expression TTS systems and also MaryTTS. We have continued this collaboration since 2018 with Sébastien Le Maguer, who has moved to the ADAPT center, Trinity College Dublin.
- The collaboration started in 2018 with John D. Kelleher (TU Dublin) has continued with the PhD of Somaye Jafaritazehjani, co-supervised by Damien Lolive and Gwénolé Lecorvé.
- Since September 2019, Sylvie Gibet collaborates with the laboratory IDMIL from McGill University. She co-supervises the postdoctoral internship of Felipe Verdugo with Marcelo Wanderley. The research subject is about the haptification of the pianistic gesture: how an audience may perceive the muscle activities of the pianist, which reflects his expressivity during the performance [15]. She started a collaboration in September 2022 with Sofia Dahl (associate professor, Denmark) and Doga Cavdir on expressive musical gestures.

- Since October 2021, Caroline Larboulette collaborates with Ravi Dattatreya, researcher and Managing Partner of Neonyx Technology, New York, and a master student from ENS Rennes, Baptiste Demoussel, on the String Art project. Since October 2022, they collaborate on a new Art and Science project, still dealing with Computational Art: the Kolam project. The aim is to create a unified/general algorithm capable of drawing Indian Kolams and Celtic Knots using the same formalism. They work with a new student, Hugo Boulier, from ENS Rennes as well.

## 6 Dissemination

### 6.1 Involvement in the Scientific Community

- Sylvie Gibet has served as a reviewer for international conferences: New Interfaces for Musical Expression 2022 (NIME 2022), Language Resources and Evaluation (LREC 2022), Sign Language Translation and Avatar Technology workshop (SLTAT 2022), Motion in Game (MIG 2022). She also served as a reviewer for journals, including, Computer & Graphics, Universal Access in the Information Society. She also participated in the scientific committee of the Journées d'Informatique Théâtrale (JIT 2022) in Grenoble. She was a member of the thesis committees of Grégoire Locqueville in Paris-Sorbonne University in June 2022.
- Giuseppe Berio served as reviewer for national and international conferences: Extraction et Gestion de Connaissances (EGC2019), Practice of Enterprise Modelling (2019), Trends in Enterprise Architecture Research (TEAR2019), Research Challenges in Information Science (RCIS2019). He also acts as steering committee member of the International Conference on Advances in Semantic Processing (SEMAPRO).
- Jonathan Chevelu has served as a reviewer for the International Conference on Audio, Speech and Signal Processing (ICASSP). He served as an expert for the French research agency (ANR). He is an elected member of the 'Conseil scientifique' (Research committee) of École Nationale Supérieure des Sciences Appliquées et de Technologie (ENSSAT).
- Nelly Barbot has served as a reviewer for the International conference of the International Speech Communication Association (Interspeech 2022) and the IJCAI-ECAI 2022 workshop on Interactions between Analogical Reasoning and Machine Learning. She is an elected member of the 'Conseil scientifique' (Research committee) of ENSSAT.
- Arnaud Delhay is an elected member of the 'Commission Recherche' (Research committee) of the IUT of Lannion. He has served as a reviewer for 13th Conference on Language Resources and Evaluation (LREC 2022), and the International conference of the International Speech Communication Association (Interspeech 2022).
- Caroline Larboulette is a member of various program committees for international conferences (ISEA2022, MIG2022, NIME2022, CASA2022, SIGGRAPH

Asia 2022), a member of the editorial review board of the international journal of computer graphics and creative interfaces (IJCICG) and serves as a reviewer for various journals (Computer & Graphics, TVCG, CAVW, Journal on Multimodal User Interfaces). In 2021, she joined the Editorial Board of Frontiers in Computer Science as an Associate Editor, to participate to the launch a new section on Computer Graphics and Visualization in 2022. Since november 2019, she is co-director of the GTAS, the "Groupe de Travail Animation et Simulation" of the GdR IG-RV (Informatique Géométrique et Graphique, Réalité Virtuelle et Visualisation) of the CNRS INS2i and also member of the direction committee of the GdR.

- Damien Lolive is head of the Signal, Image, Language Department (formerly Media and Interaction Department) of IRISA lab from October 2021. He is part of board of directors of the French speech communication association (AFCP). He is also co-responsible for the working group on inter-modality and multi-modality of the GDR TAL (special interest group in NLP). He serves as a reviewer for the IEEE Transactions on Speech and Language processing, for the IEEE Access journal, for the *Traitement Automatique des Langues* journal, for the International conference of the International Speech Communication Association (Interspeech), the International Conference on Audio, Speech and Signal Processing (ICASSP), the international conferences LREC and Speech Prosody, as well as for the *Journées d'Études sur la Parole* conference. He regularly serves as an expert for the french research agency, ANR. He also participated to 3 Phd committees during 2022: 2 as a reviewer (Gabirele Chignoli, Ajinkyia Kulkarni) and 1 as examiner and president of the jury (Manon Macary).
- Nicolas Béchet is a member of program committees for international conferences International Conference on Natural Language & Information Systems (NLDB) and Language Resources and Evaluation Conference (LREC). He also has served as a reviewer for the Intelligent Decision Technologies Journal (IDT).

## 6.2 Teaching

- Giuseppe Berio teaches courses on design and implementation of decision information systems at IUT Vannes and faculty of sciences (master level), Université Bretagne Sud. A project management course is also taught at IUT Vannes.
- Nelly Barbot teaches the following mathematics courses at École Nationale Supérieure des Sciences Appliquées et de Technologie (ENSSAT): algebra and analysis basis, mathematical logic in Licence level, probability and statistics in Master level. She is responsible of the student admission processes at ENSSAT.
- Vincent Barraud teaches the following computer science courses at École Nationale Supérieure des Sciences Appliquées et de Technologie (ENSSAT): Web Programming, Web Services and Distributed Algorithms in Licence level. He has been director of studies at ENSSAT until mid 2021 and is now responsible for the master 1 level of computer science.

- Nicolas Béchet teaches various computer sciences courses at the Statistique et Informatique Décisionnelle department of IUT Vannes. He is director of studies at IUT in STID department.
- Arnaud Delhay teaches databases and web programming (server- and client-side) in Licence levels at IUT of Lannion, calculability and computational complexity of problems in Master level at École Nationale Supérieure des Sciences Appliquées et de Technologie (ENSSAT).
- Jonathan Chevelu teaches the following computer science courses at École Nationale Supérieure des Sciences Appliquées et de Technologie (ENSSAT): cybersecurity in Licence and Master level, operative systems in Licence level and natural language processing in Master level.
- Sylvie Gibet teaches the following Computer Science courses at the faculty of sciences, Université Bretagne Sud: Data Structures and Algorithmic to undergraduate students, an introduction of Digital Signal Processing and Machine Learning (1st year master level, AIDN), and Movement and Artificial Intelligence (2nd year master level, AIDN). She is responsible for the 2nd year of the Master in Computer Science (AIDN program).
- Caroline Larboulette teaches graphical user interfaces to undergraduate students as well as introduction to computer graphics (I2G) and simulation and interactive applications (SAI) at the master level (Master of Computer Science, AIDN (Interactive Applications and Data)) at the faculty of science of the Université Bretagne Sud. She also teaches a character animation lecture at PolyTech' Angers. Since october 2020, she has been deputy head of the MIS (Mathematics - Computer Science - Statistics) department at Université Bretagne Sud. She has been elected Head of this department in november 2022. She is also responsible of both Computer Science degrees, bachelor and master.
- Damien Lolive teaches the following computer science courses at École Nationale Supérieure des Sciences Appliquées et de Technologie (ENSSAT): object-oriented programming in Licence level, compilers architecture and formal languages theory in Master level, speech and language processing in Master level, and pattern recognition in Master level. Damien Lolive has been Head of the Computer Science department at ENSSAT (until November 2022).
- Pierre-François Marteau teaches programming languages, logics, introduction to cryptography, information retrieval, machine learning and intrusion detection courses in computer sciences License and Master levels, mostly at École Nationale Supérieure de Bretagne Sud (ENSIBS). He is responsible of the bachelor level for the computer science program at ENSIBS.
- Gildas Ménéier teaches various computer sciences courses at the faculty of sciences, Université de Bretagne Sud.

### 6.3 Conferences, workshops, invitations

- Sylvie Gibet was a co-organiser of the Journées d'Étude sur la Parole (JEP 2022) in Noirmoutier (June, 13-17, see the program at <https://jep2022.univ-nantes.fr/>) [1]. She participated in the working group *Journées du Groupe de Travail Animation et Simulation* organized by Caroline Larboulette. During this conference, she co-supervised a motion capture Workshop with her. She was invited to the International Seminar on Sonic Design in University of Oslo (May 5-6).
- Caroline Larboulette organized the "journées du Groupe de Travail Animation et Simulation" that took place in Vannes on the 5th and 6th of July 2022. The program can be found here: [https://gdr-igrv.fr/event/journee\\_gtas2022/](https://gdr-igrv.fr/event/journee_gtas2022/). During this national conference, she organized a motion capture workshop with Sylvie Gibet and gave a talk on her on-going research with student Victor Haguët and artist Laura Pouppeville.
- Nelly Barbot, Damien Lolive and Aghilas Sini were co-organisers of the Journées d'Études sur la Parole (JEP 2022) in Noirmoutier (June, 13-17, see the program at <https://jep2022.univ-nantes.fr/>) [1].

### 6.4 Meetings, other dissemination

### 6.5 Graduate student and student internship

- Rania Bennetayeb has done her Master level internship in Lannion as part of his Master in Sciences and Technologies of the of the Faculty of Electrical Engineering of the University of Science and Technology Houari Boumediene. She has been supervised by Aghilas Sini (Learn&Go), Damien Lolive and Arnaud Delhay-Lorrain. She has worked on a joint learning framework for a text-to-speech system and voice conversion in a deep learning context.
- Pierre Bonneau, student in computer science Master 1 at ENSSAT has done an internship, between 5 weeks, dealing on supervised classification based on proportions. He has been supervised by Aghilas Sini (Learn&Go), Nelly Barbot and Laurent Miclet.
- Mathis Leroux has done his internship in the team as part of his DUT (Diplôme Universitaire de Technologie - 2nd year bachelor level) in computer science. He has been supervised by Aghilas Sini (Learn&Go) and Arnaud Delhay-Lorrain. He has developed a platform that aims to provide the team's researchers with a demonstration tools of the speech technologies APIs designed by the team.
- Victor Dang, a License 3 student in computer science at ENSSAT, has done a 5-weeks internship. He worked on building a tool to split automatically text and audio samples at the sentence level, using publicly available deep learning alignment models. He was supervised by Antoine Perquin.
- Mattias Kockum, a Master 1 student in computer science at ENSSAT, has done a 4-weeks internship. He worked on automating the evaluation process of the

pronunciation analysis tool developed by EXPRESSION. He was supervised by Antoine Perquin and Aghilas Sini.

## 7 Bibliography

### Books and Monographs

- [1] O. CROUZET, E. DELAIS-ROUSSARIE, A. DUNIEC, L. LEPRIEUR, P. L. ROHRER, M. TAHON, J. WOTTAWA, M. BRABANT, H. RIGUIDEL, N. BARBOT, S. GIBET, D. LOLLIVE, A. SINI (editors), *Actes des XXXIVe Journées d'Études sur la Parole – JEP 2022*, ISCA, December 2022, <https://hal.science/hal-03958018>.

### Articles in referred journals and book chapters

- [2] A. BONDU, Y. ACHENCHABE, A. BIFET, F. CLEROT, A. CORNUEJOLS, J. GAMA, G. HEBRAIL, V. LEMAIRE, P.-F. MARTEAU, “Open challenges for Machine Learning based Early Decision-Making research”, *ACM SIGKDD Explorations Newsletter* 24, 2, nov 2022, p. 12–31, <https://doi.org/10.1145/2F3575637.3575643>.
- [3] F. SAÏD, I. EVENO, J. VILLANEAU, “Discrete-Events Simulation for Teaching Statistics in Industrial Engineering”, *Athens Journal of Technology & Engineering* 9, 1, March 2022, <https://hal.science/hal-03572881>.

### Publications in Conferences and Workshops

- [4] D. BATTISTELLI, A. ETIENNE, R. RAHMAN, C. TEISSÈDRE, G. LECORVÉ, “Une chaîne de traitements pour prédire et appréhender la complexité des textes pour enfants d’un point de vue linguistique et psycho-linguistique.”, *in: Traitement Automatique des Langues Naturelles (TALN 2022)*, Y. Estève, T. Jiménez, T. Parcollet, M. Zanon Boito (editors), ATALA, p. 236–246, Avignon, France, June 2022, <https://hal.science/hal-03701501>.
- [5] A. BLANDIN, F. SAÏD, J. VILLANEAU, P.-F. MARTEAU, “Analyse automatique d’émotions pour l’optimisation de campagnes d’e-mails en français”, *in: TextMine ’22 - Atelier sur la Fouille de Texte*, Conférence EGC (Extraction et Gestion des Connaissances) 2022, p. 67–77, Blois, France, January 2022, <https://hal.science/hal-03981781>.
- [6] A. BLANDIN, F. SAÏD, J. VILLANEAU, P.-F. MARTEAU, “DaFNeGE: Dataset of French Newsletters with Graph Representation and Embedding”, *in: Text, Speech and Dialog (TSD) conference, Lecture Notes in Computer Science, 13502*, Springer International Publishing, p. 16–27, Brno, Czech Republic, September 2022, <https://hal.science/hal-03966533>.
- [7] A. BLANDIN, F. SAÏD, J. VILLANEAU, P.-F. MARTEAU, “Graphical document representation for french newsletters analysis”, *in: DocEng ’22: ACM Symposium on Document Engineering 2022*, ACM, p. 1–8, San Jose California, France, September 2022, <https://hal.science/hal-03966518>.
- [8] E.-R. BLEZA, V. MONBET, P.-F. MARTEAU, “Prédiction des niveaux de risque pollinique à partir de données historiques multi-sources”, *in: EGC 2022 - Extraction et Gestion des Connaissances*, Blois, France, January 2022, <https://hal.science/hal-03614274>.



- [9] D. CAO, O. AHMIA, N. BÉCHET, P.-F. MARTEAU, “Chinese public procurement document harvesting pipeline”, in: *DocEng '22: ACM Symposium on Document Engineering 2022*, ACM, p. 1–4, San Jose California, France, September 2022, <https://hal.science/hal-03966541>.
- [10] B. DEMOUSSEL, C. LARBOULETTE, R. DATTATREYA, “A Greedy Algorithm for Generative String Art”, in: *Bridges 2022 ( Mathematics, Art, Music, Architecture, Culture )*, Aalto, Finland, August 2022, <https://hal.science/hal-03901755>.
- [11] D. GUENNEC, H. HAJIPOOR, G. LECORVÉ, P. LINTANF, D. LOLIVE, A. PERQUIN, G. VIDAL, “BreizhCorpus: a Large Breton Language Speech Corpus and its use for Text-to-Speech Synthesis”, in: *Odyssey Workshop 2022*, ISCA (International Speech Communication Association), ISCA, p. 263–270, Beijing, China, June 2022, <https://hal.science/hal-03944464>.
- [12] C. MERTZ, V. BARREAUD, T. LE NAOUR, D. LOLIVE, S. GIBET, “A Low-Cost Motion Capture Corpus in French Sign Language for Interpreting Iconicity and Spatial Referencing Mechanisms”, in: *Language Resources and Evaluation Conference*, Marseille, France, June 2022, <https://hal.science/hal-03966599>.
- [13] A. SINI, D. LOLIVE, N. BARBOT, P. ALAIN, “Investigating Inter- and Intra-speaker Voice Conversion using Audiobooks”, in: *Language Resources and Evaluation Conference*, Marseille, France, June 2022, <https://hal.inria.fr/hal-03698340>.
- [14] M. TCHENEGNON, S. GIBET, T. LE NAOUR, “A New Spatio-Temporal Loss Function for 3D Motion Reconstruction and Extended Temporal Metrics for Motion Evaluation”, in: *European Conference on Computer Vision (ECCV 2022), Workshop on What is Motion for?*, Tel Aviv, Israel, October 2022, <https://hal.science/hal-03966941>.
- [15] F. VERDUGO, A. CEGLIA, C. FRISSON, A. BURTON, M. BEGON, S. GIBET, M. WANDERLEY, “Feeling the Effort of Classical Musicians - A Pipeline from Electromyography to Smartphone Vibration for Live Music Performance”, in: *International Conference on New Interfaces for Musical Expression*, PubPub, June 2022, <https://hal.inria.fr/hal-03967243>.
- [16] J. VILLANEAU, F. SAID, “COSMOS: Experimental and Comparative Studies of Concept Representations in Schoolchildren”, in: *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, European Language Resources Association, p. 5251–5260, Marseille, France, June 2022, <https://aclanthology.org/2022.lrec-1.563>.
- [17] L. WADOUX, N. BARBOT, J. CHEVELU, D. LOLIVE, “Impact du contenu phonétique sur les plongements de locuteurs pour le clonage de voix : vers l’application aux pathologies vocales”, in: *XXXIVe Journées d’Études sur la Parole – JEP 2022*, ISCA, p. 814–822, Noirmoutier-en-l’Île, France, June 2022, <https://hal.science/hal-03975729>.
- [18] L. WADOUX, N. BARBOT, D. LOLIVE, J. CHEVELU, “Voice Cloning Applied to Voice Disorders: a Study of Extreme Phonetic Content in Speaker Embeddings”, in: *35th Canadian Conference on Artificial Intelligence*, Toronto, Canada, May 2022, <https://hal.inria.fr/hal-03697484>.
- [19] O. ZHANG, N. GENGEMBRE, O. L. BLOUCH, D. LOLIVE, “Dispeech: A Synthetic Toy Dataset for Speech Disentangling”, in: *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, p. 8557–8561, Singapore, Singapore, May 2022, <https://hal.science/hal-03975832>.