Activity Report 2021

Team EXPRESSION

Expressiveness in Human Centered Data/Media

D6 – Media and Interactions

1 Team composition

1.1 Composition

Head of the team

Pierre-François Marteau, Professor, Université Bretagne Sud

Administrative assistants

Anne Idier, Université Bretagne Sud ($>$ sep. 2020)

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 Barreaud, 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

$\leftarrow$ Gwénolé Lecorvé, Associate professor, Université de Rennes 1

Damien Lolive, Associate 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

$\rightarrow$ Lauren Thevin, Université Catholique de l'Ouest, since June 2021

Non-permanent members

$\leftarrow$ Nazanin Dehghani, Post-doctoral researcher, Université de Rennes 1 (from August 2019 to August 2021)

Quentin Di-Fant, Engineer, Université de Rennes 1 (from November 2018 to October 2020)

$\leftarrow$ Hassan Hajipoor, Research Engineer, Université de Rennes 1 (from August 2019 to March 2021)

Simon Giddings, Engineer, Université de Rennes 1 (from November 2019)

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)

⇄ Betty Fabre, PhD Student, University of Rennes 1 (September to December 2021)

⇄ Valentin Durand de Gevigney, PhD Student, University of Rennes 1 (October to December 2021)

Aghilas Sini, Post-doctoral researcher, University of Rennes 1 (from November 2020 to December 2021)

→ Antoine Perquin, Post-doctoral researcher, University of Rennes 1 (from February 2021)

PhD students

→ Danrun Cao, Université Bretagne Sud, CIFRE OctopusMind (from October 2021)

Alexis Blandin, Université Bretagne Sud, CIFRE UNEEK (from January 2020)

Esso-Ridah Bléza, Université Bretagne Sud, CIFRE, 2nd year

Tiago Brizolara, Université Bretagne Sud, ARED/CD56, Cesure year

Mansour Tchénégnon, Univerisité Bretagne Sud, CDE/ARED 2nd year

← Valentin Durand De Gevigney, Université Bretagne Sud, DGA/CDE, defended in December 2021

Betty Fabre, Université de Rennes 1, CIFRE Orange Labs, 3rd year

Somaye Jafaritazehjani, Université de Rennes 1, CDE/TU Dublin, 3rd year

Jade Mekki, Université de Rennes 1, ANR TREMoLo, 3rd year

← Antoine Perquin, Université de Rennes 1, CDE, defended in February 2021

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

Master students

⇄ Romann Yvinec, University of Mac Gill and Université Bretagne Sud

⇄ Thibault Gaudier, University of Rennes 1 (from February 2021 to July 2021)

1.2   Evolution of the staff

The permanent staff has been stable during the year. The number of PhD students is almost stable with 2 PhD defense, against 1 newly hired PhD student. The number of Non permanent members is slightly decreasing too with on incoming member for two leaves.

Notably, Lauren Thevin joined the team as an external member since June 2020. Since then, Lauren has been reorienting her research activity from her past research topics to collaborative projects with Expression team members:

- Organization of the UBS & UCO-BS / Social Sciences & Technical Sciences Seminar on June 24, 2021 (12 participants);

- The workshop article at ACM ISS'2021 (rank A conference) aims to articulate research in HMI with problems of the Expression team, namely different representations of the same thing (remote or co-located participants, objects) on different supports (e.g., AR, VR);

- Project "avatars" accessible with visual impairments with Caroline Larboulette. The participatory design was carried out from September 2021 to December 2021. The experiment, where visually impaired and blind students collaborate to request crossing assistance, takes place the week of February 7, 2022, which will result in a joint publication. This first collaboration is the opportunity to continue a project of accessible and communicating avatars during 2022, coupled with a corpus of French sign language on animal.

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 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

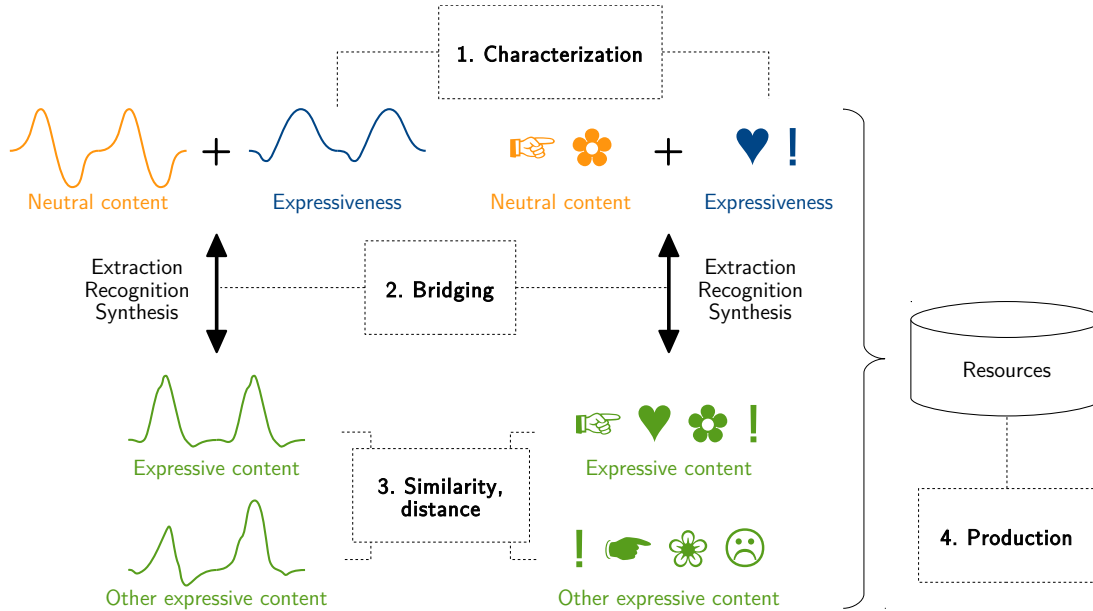Four main challenges will be addressed by the team.

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

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[1] (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

---

[1]Given two metric or similarity spaces (X, d) and (X', d'), a map $f : (X, d) \to (X', d')$ is called an embedding.

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.

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 2021 are listed into the following key issues items defined above for the team:

Anomaly detection

Random Partitioning Forest for Point-Wise and Collective Anomaly Detection - Application to Network Intrusion Detection [3] We have proposed, DiFF-RF, an ensemble approach composed of random partitioning binary trees to detect point-wise and collective (as well as contextual) anomalies. Thanks to a distance-based paradigm used at the leaves of the trees, this semi-supervised approach solves a drawback that has been identified in the isolation forest (IF) algorithm. Moreover, taking

into account the frequencies of visits in the leaves of the random trees allows to significantly improve the performance of DiFF-RF when considering the presence of collective anomalies. DiFF-RF is fairly easy to train, and excellent performance can be obtained by using a simple semi-supervised procedure to setup the extra hyper-parameter that is introduced. We first evaluate DiFF-RF on a synthetic data set to i) verify that the limitation of the IF algorithm is overcome, ii) demonstrate how collective anomalies are actually detected and iii) to analyze the effect of the meta-parameters it involves. We assess the DiFF-RF algorithm on a large set of datasets from the UCI repository, as well as two benchmarks related to intrusion detection applications. Our experiments show that DiFF-RF almost systematically outperforms the IF algorithm, but also challenges the one-class SVM baseline and a deep learning variational auto-encoder architecture. Furthermore, our experience shows that DiFF-RF can work well in the presence of small-scale learning data, which is conversely difficult for deep neural architectures. Finally, DiFF-RF is computationally efficient and can be easily parallelized on multi-core architectures.

Motion synthesis and editing for the generation of new sign language content [4]   Existing work on the animation of signing avatars often relies on pure procedural techniques or on the playback of Motion Capture (MoCap) data. While the first solution results in robotic and unnatural motions, the second one is very limited in the number of signs that it can produce. A data-driven motion synthesis technique has been modeled and implemented to increase the variety of Sign Language (SL) motions that can be made from a limited database. In order to generate new signs and inflection mechanisms based on an annotated French Sign Language MoCap corpus, we rely on phonological recombination, i.e. on the motion retrieval and modular reconstruction of SL content at a phonological level with a particular focus on three phonological components of SL: hand placement, hand configuration and hand movement. We propose to modify the values taken by those components in different signs to create their inflected version or completely new signs by (i) applying motion retrieval at a phonological level to exchange the value of one component without any modification, (ii) editing the retrieved data with different operators, or, (iii) using conventional motion generation techniques such as interpolation or inverse kinematics, which are parameterized to comply to the kinematic properties of real motion observed in the data set. The quality of the synthesized motions is perceptually assessed through two distinct evaluations that involved 75 and 53 participants respectively.

Gesture Generation

Proximal-to-Distal Sequences of Attack and Release Movements of Expert Pianists during Pressed-Staccato Keystrokes [5]   The aims of this study were to i) evaluate proximal-to-distal sequencing (PDS) in pianists' attack and release movements during pressed-staccato keystrokes, and ii) investigate if trunk motion facilitates PDS of upper-limb movements. Nine expert pianists performed a series of loud pressed-staccato keystrokes. Kinematic data was recorded with a 3D motion capture system. PDS was assessed by comparing temporal organization of peak velocities from the pelvis to the wrist. Evidence of PDS was found across the kinematic chain. Pianists' use of

Gesture Analysis

PDS differed between scapula and shoulder movements. Trunk motion facilitated PDS organization by increasing anticipatory shoulder movements and by preceding shoulder-girdle attack and release movements. Implications might relate to research on performance optimization and injury prevention strategies.

Knowledge extraction

Questionnement réflexif et méthodologique autour d'un corpus sensible recueilli en ligne [1]   We work on an approach to deal with a sensitive corpus. First, we conceptualize a multi-criteria textual analysis, with indices of variable nature and size, based on the principles of comparison and variation, and captured thanks to back and forth between linguistic and computer analyses. Our experiment on advice shows using specific machine learning model that the techno discursive characteristics of webchat can affect the reception of the act actually performed, in particular by potential confusion with the order.

Knowledge extraction

A 3-phase approach based on sequential mining and dependency parsing for enhancing hypernym patterns performance [2]   Abstract Patterns have been extensively used to extract hypernym relations from texts. The most popular patterns are Hearst's patterns, formulated as regular expressions mainly based on lexical information. Experiences have reported good precision and low recall for such patterns. Thus, several approaches have been developed for improving recall. While these approaches perform better in terms of recall, it remains quite difficult to further increase recall without degrading precision. In this paper, we propose a novel 3-phase approach based on sequential pattern mining to improve pattern-based approaches in terms of both precision and recall by (i) using a rich pattern representation based on grammatical dependencies (ii) discovering new hypernym patterns, and (iii) extending hypernym patterns with anti-hypernym patterns to prune wrong extracted hypernym relations. The results obtained by performing experiments on three corpora confirm that using our approach, we are able to learn sequential patterns and combine them to outperform existing hypernym patterns in terms of precision and recall. The comparison to unsupervised distributional baselines for hypernym detection shows that, as expected, our approach yields much better performance. When compared to supervised distributional baselines for hypernym detection, our approach can be shown to be complementary and much less loosely coupled with training datasets and corpora.

Knowledge extraction

TREMoLo-Tweets: a Multi-Label Corpus of French Tweets for Language Register Characterization [9]   The casual, neutral, and formal language registers are highly perceptible in discourse productions. However, they are still poorly studied in Natural Language Processing (NLP), especially outside English, and for new textual types like tweets. To stimulate research, this paper introduces a large corpus of 228,505 French tweets (6M words) annotated in language registers. Labels are provided by a multi-label CamemBERT classifier trained and checked on a manually annotated subset of the corpus, while the tweets are selected to avoid undesired biases. Based on the corpus, an initial analysis of linguistic traits from either human annotators or automatic

extractions is provided to describe the corpus and pave the way for various NLP tasks. The corpus, annotation guide and classifier are available on http://tremolo.irisa.fr.

Neural-Driven Search-Based Paraphrase Generation [7] We study a search-based paraphrase generation scheme where candidate paraphrases are generated by iterated transformations from the original sentence and evaluated in terms of syntax quality, semantic distance, and lexical distance. The semantic distance is derived from BERT, and the lexical quality is based on GPT2 perplexity. To solve this multi-objective search problem, we propose two algorithms: Monte-Carlo Tree Search For Paraphrase Generation (MCPG) and Pareto Tree Search (PTS). We provide an extensive set of experiments on 5 datasets with a rigorous reproduction and validation for several state-of-the-art paraphrase generation algorithms. These experiments show that, although being non explicitly supervised, our algorithms perform well against these baselines.

Generation, Use cases and evaluation

Style as Sentiment Versus Style as Formality: The Same or Different? [8] Unsupervised textual style transfer presupposes that style is a coherent and consistent concept and that style transfer approaches will generalise consistently across different domains of style. This paper explores whether this presupposition is appropriate for different types of style. We explore this question by comparing the performance and latent representations of a variety of neural encoder-decoder style-transfer architecture when applied to sentiment transfer and formality transfer. Our findings indicate that the relationship between style and content shifts between these different domains of style: for sentiment, style and content are closely entangled; however, for formality, they are less entangled. Our findings suggest that for different types of styles different approaches to modeling style for style-transfer are necessary.

Generation, Use cases and evaluation, Knowledge extraction

Style as Sentiment Versus Style as Formality: The Same or Different? [6] Email communication and newsletter campaigns remain a significant concern for companies. The main question addressed here is how to optimize the form and content of a newsletter so that it is not interpreted as spam or annoyance by the recipient. We address this question by analyzing the emotions and opinions conveyed by emails and evaluating how they affect their open and click rate performance. We first describe a new dataset of French newsletters, and then we use emotional embeddings to analyze the associations between emotions and email performance. We finally derive clues on how to write effective email campaigns.

Use cases and evaluation, Knowledge extraction

Voice conversion: characters conversion in audiobooks Audiobook readers play with their voices to emphasize text passages, highlight discourse changes or significant events in order to make listening easier and entertaining. In particular, 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,

Use cases and evaluation, Generation

these intra-speaker modifications are hard to reproduce with simple text-to-speech synthesis. We have investigated this problem through the prism of voice conversion. This study has been initiated during the Master internship of Thibault Gaudier and is continued by Aghilas Sini, Damien Lolive, Nelly Barbot and Pierre Alain.

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. The main contributions of this work are three-fold: (1) to tackle the feasibility of the intra-speaker voice conversion, we apply an any-to-one voice conversion model and propose to train models using amateur audiobooks ; (2) two datasets in French are presented and used in this work including a parallel dataset containing six different speakers for studying inter-speaker voice conversion, and a non-parallel dataset of single female speaker to address the intra-speaker voice conversion ; (3) results show that voice conversion methods can be used to convert indirect speech style to direct speech style using subjective evaluations. These first results are submitted to an international conference.
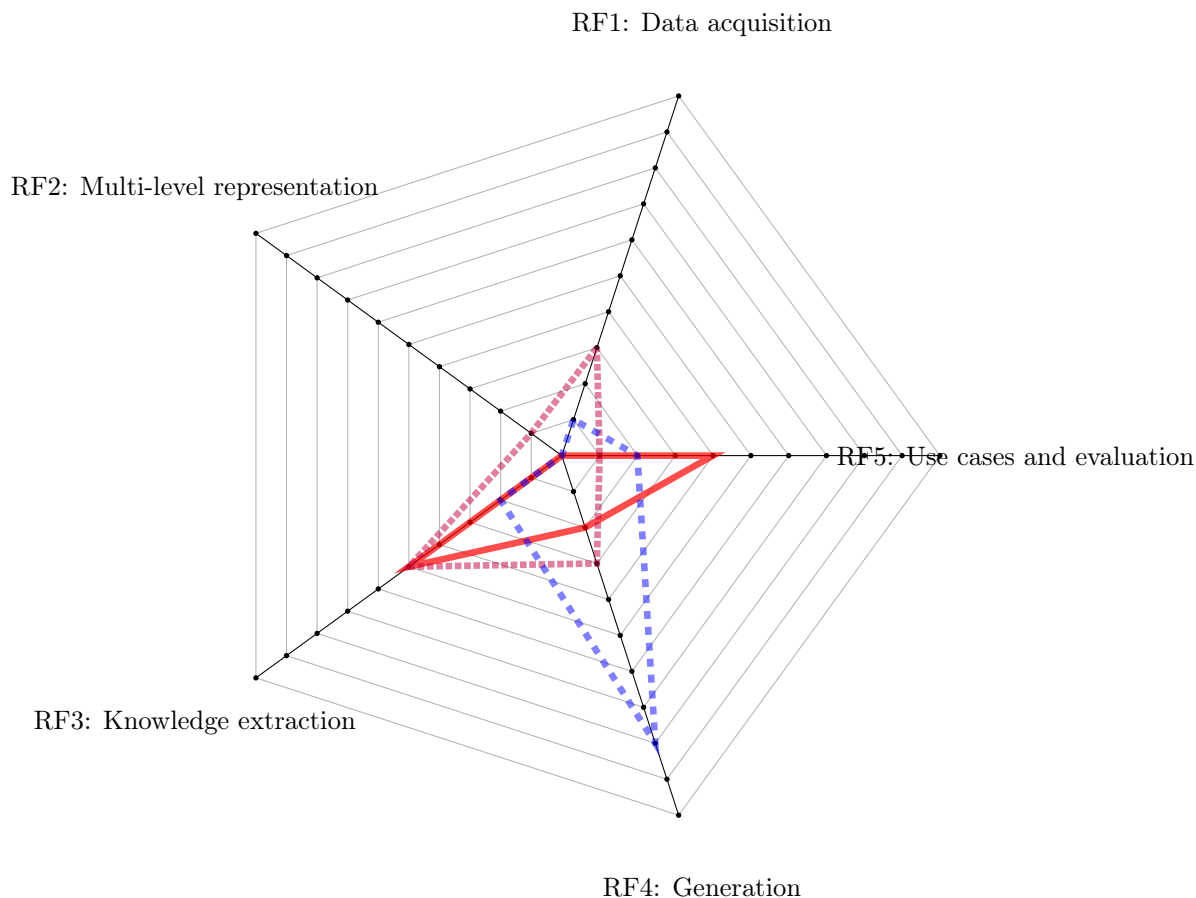
Summary of the contributions

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

- Antoine Perquin defended his PhD on the 12th of February 2021. The goal of this thesis is two-fold. 1) To study whether it is possible alleviate the need for human linguistic expertise to build or modify a TTS system. 2) To study whether it is possible to produce speech corresponding to different speakers, with their respective tone and regionalism accent. This manuscript presents three contributions. First, we show that the embedding property of neural networks can be used to lower the amount of expertise in unit selection speech synthesis. Second, we show that character embeddings can remove all linguistic expertise for end-to-end systems. Finally, we attempt to explicitly model speaker and accent characteristics in order to build a multi-speaker multi-accent end-to-end speech synthesis system.

- Valentin Durand de Gevigney defended his PhD on the 10th of December 2021. His research led to the design of Deep Neural Network models for the multimodal detection of abnormal behaviours or anomalies. Here, anomalies, are defined has behaviours that are unexpected considering the context they occur in. His major results are the proposal of a Variational Interpolating Auto-Encoder architecture that allows to cope with the sequentiality of the data in a very efficient manner. This architecture has been validated on video surveillance and network traffic monitoring with very promising outcomes. This last work will be submitted in an international journal.

3.3   On going PhDs

1. Esso-Ridah Bleza has pursued his PhD in July 2021 while changing the research subject following the bankruptcy of his former company. A new CIFRE contract has been established with the start-up LifyAir. His new research subject concerns the prediction of the risk of pollen emission all over the national territory. As such, Esso-Ridah is required to develop spatio-temporal prediction models based on machine learning approaches capable of effectively exploiting complex historical data.

2. Alexis Blandin has pursued his PhD work (2nd 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 a publication in CENTRIC 2022 conference [6] and to submit at the end of this year two contributions, one in TextMine 2022 workshop and one in LREC2022 conference.

3. Tiago Brizolara has taken a *Cesure* year at the end of his 2nd year of PhD.

4. Mansour Tchénégnon has completed his first year of PhD. During this year, he experimented several deep learning architectures to reconstruct 3D-skeleton data from video data. To this end, he created an avatar-based dataset that contains

a large variety of motion features. Based on his dataset, he proposed a novel method focusing on temporal aspects of motion and showed first promising results. A DCNN-based model, taking as input the 2D pose estimations directly computed from videos, was trained. He had a paper accepted to the J.FIG Workshop (Journées Françaises d'Informatique Graphique) [10]. In this paper he presented his first training tasks and results obtained using the CVM-Net deep neural network model to directly estimate 3D poses. Three different models were trained using the same architecture applied on several configurations of the dataset. Using a small benchmark, the network architecture was evaluated.

5. 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.

6. Betty Fabre has started her PhD in October 2018. This CIFRE PhD work is jointly conducted in Orange Labs and IRISA/Expression in Lannion. Her research aims at reinforcement learning and structured prediction applied to paraphrase generation. We welcomed Betty in the team since September for her to prepare her manuscript. Her defense is planned to happen in February 2022.

7. Clémence Mertz has completed her second PhD year in 2021. She initiated a reflection on the use of Deep Learning techniques applied to the translation of LSF in French. A first corpus was designed on the basis of spatial referencing elements in LSF (see LSF-SHELVES). Data was collected from 15 signers with a light-weight capture device. LSF-SHELVES is the main subject of an article submitted to the International Conference on Language Ressources and Evaluation (LREC).

8. 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.

9. Lily Wadoux has started her second PhD year. 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 plateform which is used to learn the models. She conducted first experimental evaluations of the pipeline (informal evaluation of the speech output). She has also prepared her participaton to the Advanced Language Processing Winter School that will be help in early January 2022. A paper is being prepared for submission in January 2022.

10. 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.

4   Software development

LSF-ANIMAL   The *LSF-ANIMAL* corpus developed in 2018 was augmented with 30 min of MoCap data coming from another signer to add inter personal variations. The new content consists in animal descriptions, various isolated signs, hand configurations and grammatical mechanisms such as pointing gestures and proforms. The *LSF-ANIMAL* corpus was perceptually evaluated and validated by 50 people (deaf and hearing). The data was added to the *SGN library* which has been developed in the team for several years. Motion retrieval methods comes from *SGN* and new interpolation techniques were added in *SGN*.

LSF-SHELVES   The *LSF-SHELVES* corpus is a French Sign Language Corpus developed during Clemence Mertz's thesis (2021). This corpus implements the spatial referencing and iconicity mechanisms in LSF, selected through examples of positioning objects in relation to each other on shelves. It proposes to produce utterances with increasing levels of difficulty. An original low-cost motion acquisition technique based on the use of a Microsoft Kinect was derived to collect the data sets from 15 signers whereas classical protocols make use of motion capture studios. An additional editing step was required in order to incorporate manual configurations and correct wrist orientation trajectories.

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 AMR7 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 Ménier]   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: Gwénolé Lecorvé, Nicolas Béchet, Jonathan Chevelu, Nazanin Dehghani, Jade Mekki, Aline Étienne.

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 Gwénolé 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   PI IA Education

Participants: Gwénolé Lecorvé, Damien Lolive.

This project, coordinated by Learn&Go, aims at developing a software usable to learn the French language. Our contribution is on Speech Synthesis and language analysis for education applications. Notably, we have to produce an adapted feedback for children. This project is funded by the ministry of education. This project ended mid-2021.

5.1.3   TextToKids ANR Project

Participants: Gwénolé Lecorvé, Nicolas Béchet, Jonathan Chevelu, Damien Lolive, Alexis Blandin.

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 Gwénolé Lecorvé.

5.1.4   OPLB Breton Synthesis

Participants: Damien Lolive, Gwénolé Lecorvé, Gaëlle Vidal, Hassan Hajipoor.

This project aims at building speech resources and a speech synthesis engine for the Breton Language. It started in August 2019. This project is funded by *OPLB - Office Public de la Langue Bretonne* and lasts 18 months. Considering the pandemy and the delay in project realization, the end is postponed to end of March 2021. The outcome of the project is a corpus of Breton Speech including to voices (one female and one male) with more than 20h each. On the software side, an experimental but functional speech synthesis system has been delivered along with some tools such as a phonetizer, a syllabifier, a stress predictor, a text normalisation software.

5.1.5   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.

5.1.6   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.

5.1.7   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   International Initiatives

5.2.1   H2020 NADINE

Participants: Damien Lolive, Gwénolé Lecorvé, Arnaud Delhay, Waseem Safi, Cédric Fayet, Aghilas Sini, Antoine Perquin, Gaëlle Vidal.

NADINE project aim is to develop a novel way of integrating migrants and refugees through ICT-enabled solutions that will automatically adapt to the specificities of each person. The consortium agrees that one of the main enablers of migrants/refugees inclusion, in the host societies, is their ability to work. Hence NADINE's motto is "Give migrants and refugees their dignity back by giving them a decent job with a decent salary". Taking into account this important factor, NADINE will create an adaptable platform able to: 1) Provide functionalities for skill assessment, 2) dynamically create tailored suited training programs to adapt existing skills into host societies needed skills, 3) provide a digital companion that will suggest and assist the end-users through administrative tasks and 4) create a data lake available to public administration bodies for better organisation of migration flows. NADINE will innovate in several directions from novel training tools, adaptable to different learning setups, to novel ways of information flow handling for public administrations to work efficiently in both business as usual contexts and migration bursts ones. NADINE platform will create potential new markets in different market areas and also will provide novel open tools that will foster new innovation capacity to the EU area. This project ended in October 2021.

5.3   National Collaborations

- 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's PhD entered its second year.

- Nelly Barbot collaborates with Laurent Miclet (former) emeritus professor at University of Rennes 1, Henri Prade, emeritus research director of CNRS at IRIT Toulouse and Gilles Richard, full professor at IRIT Toulouse 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" in october 2022.

- 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.

5.4   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 continue 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.

- From September 2019 to June 2022, a collaboration with the laboratory IDMIL from McGill University has started with the postdoctoral internship of Felipe Verdugo, supervised by Marcelo Wanderley and Sylvie Gibet. The post-doc 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.

- 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.

6 Dissemination

6.1 Involvement in the Scientific Community

- Pierre-François Marteau served as a reviewer in international journals (IEEE Trans. on NNLS, Journal of Soft Computing, Pattern Recognition, IEEE Trans. on Information Forensic and Security). He serves as an expert for French Ministry of Research (CIR/JEI). He was member of a thesis committee at XXXX , and participated in a few local PhD defense juries. He presented three seminars as a guest speaker: TS days INRIA/IRISA RENNES, 25th and 26th March 2019, SEC DAYS INRIA/IRISA RENNES, 9th January 2019, Data Science workshop at Bilbao Center for Applied Mathematics, Bilbao, 7th and 8th November 2019. He had co-organized the FDST2019 workshop dedicated to "Learning from Data Streams and Time Series" at Telecom ParisTech, 12 September 2019. He is member of the Strategic Orientation Committee at IRISA.

- Sylvie Gibet has served as a reviewer for national and international conferences, including NIME 2021, Motion & Computing (MoCo2021), Motion in Game (MIG 2021). She also served as a reviewer for journals, including Journal of Biomechanics, Machine Translation, Computer & Graphics, Journal of Multimodal User Interfaces. She was a member of the thesis committees of Théo Voillemin at IMT in Lille in October 2021 (reader), and of Fajrian Yunus at Paris-Sorbonne University in December 2021 (reviewer).

- 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).

- Nelly Barbot has served as a reviewer for the International conference of the International Speech Communication Association (Interspeech 2021).

- Arnaud Delhay is an elected member of the 'Commission Recherche' (Research committee) of the IUT of Lannion. He has served as a reviewer for the International conference of the International Conference on Audio, Speech and Signal Processing (ICASSP 2022).

- Caroline Larboulette is a member of various program committees for international conferences (ISEA2022, MIG2021, NIME2021, Eurographics Doctoral Consortium 2021), 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. She is a member of the ACM SIGGRAPH Specialized Conferences Committee that attributes the ACM SIG-GRAPH labels to conferences and supervises the budget of conferences sponsored by ACM SIGGRAPH and the liaison agent from this committee with the ACM SIGGRAPH Students Travels Grants committee. 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.

- Gwénolé Lecorvé is an elected member of the laboratory council of IRISA, and of the board of directors of the French speech communication association (AFCP). Since December 2019, he is also a nominated (deputy) member of the National Council of Universities (*Conseil National des Universités, CNU*) in the Computer Science section, and an elected member of the Scientific Council of ENSSAT. In 2019, he also served as a reviewer for conferences and journals (Interspeech, ICASSP, TALN, Journal of Applied Soft Computing).

- Damien Lolive is head of the 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 also regularly serves has an expert for the french research agency, ANR.

- 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 Barreaud 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: Algorithmic at Bachelor level (Python), an introduction of Digital Signal Processing and Machine Learning (1st year master level), and Movement and Artifical Intelligence (2nd year master level).

- Caroline Larboulette teaches character animation as part of an introductory lecture of computer science for freshmen and logic for undergraduates of the UFR SSI; C++ programming for ENSIBS graduate students; introduction to computer graphics (I2G) and simulation and interactive applications (SAI) at the master level (Master of Computer Science, AIDN (Interactive Applications and Data)).

- Gwénolé Lecorvé teaches the following computer science courses at École Nationale Supérieure des Sciences Appliquées et de Technologie (ENSSAT): distributed algorithmics; artificial intelligence; and machine learning in Master level. He also teaches automatic speech recognition and speech synthesis in Research Master program of University of Rennes 1, in Rennes.

- 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 is Head of the Computer Science department at ENSSAT and also responsible for the master 1 level of computer science.

- Pierre-François Marteau teaches programming languages, logics, introduction to cryptography and 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 Ménier teaches various computer sciences courses at the faculty of sciences, Université de Bretagne Sud.

6.3   Conferences, workshops, invitations

- Sylvie Gibet was an invited lecture at the CIRMMT Distinguished Lectures in the Science and Technology of Music, McGill University, Montréal – Gesture: A language to sense, express, control, April 29, 2021.

6.4   Meetings, other dissemination

- Nelly Barbot, Jonathan Chevelu, Arnaud Delhay, Gwénolé Lecorvé and Damien Lolive were part of the organisation committee of the CNRS summer school ETAL[2] in collaboration with the GdR-TAL and GdR MaDICS. Valentin Durand de Gevigney, Antoine Perquin, Aghilas Sini, Lily Wadoux, and Angélique Le Pennec helped with the organisation, especially with the technical aspects and the running of this school. This school took place at ENSSAT (Lannion): initially, this school was scheduled for July 2020 and has been postponed to June 2021, due to the Covid pandemy. It lasted 5 days and brought together 25 participants and 7 teachers in person. This school aimed to introduce new advances in the joint treatment of oral and textual modalities and train to the main associated tools. Damien Lolive was responsible of the organisation committee and member of the scientific committee.

6.5   Graduate student and student internship

- Frédéric Le Bellour has done his Master level internship in Lannion and in collaboration with the LIUM in Le Mans, France. He has been supervised by Damien Lolive and has worked on the semi-supervised annotation of emotions in audiobooks.

- Aline Étienne has done his Master level internship in Nanterre (MoDyCo lab), supervised by Delphine Battistelli and Gwénolé Lecorvé. She has been studying linguistic cues of interest to explain how a child may understand or not a given text.

- Thibault Gaudier has done his Master level in Lannion. He has been supervised by Aghilas Sini, Damien Lolive and Nelly Barbot. He has worked on the voice conversion for the character changes in an audiobook.

7   Bibliography

Articles in referred journals and book chapters

[1] N. Béchet, N. Garric, G. Ledegen, F. Pugnière-Saavedra, "Questionnement réflexif et méthodologique autour d'un corpus sensible recueilli en ligne", *Terminal. Technologie de l'information, culture société*, 129, March 2021, https://hal.archives-ouvertes.fr/hal-03221366.

---

[2]https://gdr-tal.ls2n.fr/etal-2021/

[2] A. Issa Alaa Aldine, M. Harzallah, G. Berio, N. Béchet, A. Faour, "A 3-phase approach based on sequential mining and dependency parsing for enhancing hypernym patterns performance", *Knowledge Engineering Review 36*, September 2021, `https://hal.archives-ouvertes.fr/hal-03502370`.

[3] P.-F. Marteau, "Random Partitioning Forest for Point-Wise and Collective Anomaly Detection - Application to Network Intrusion Detection", *IEEE Transactions on Information Forensics and Security 16*, January 2021, p. 2157–2172, `https://hal.archives-ouvertes.fr/hal-02882548`.

[4] L. Naert, C. Larboulette, S. Gibet, "Motion synthesis and editing for the generation of new sign language content", *Machine Translation*, 2021, `https://hal.inria.fr/hal-03324348`.

[5] F. Verdugo, M. Begon, M. Wanderley, S. Gibet, "Proximal-to-Distal Sequences of Attack and Release Movements of Expert Pianists during Pressed-Staccato Keystrokes", *Journal of Motor Behavior*, 2021, p. 1–11, `https://hal.inria.fr/hal-03324360`.

Publications in Conferences and Workshops

[6] A. Blandin, F. Saïd, J. Villaneau, P.-F. Marteau, "Automatic Emotions Analysis for French Email Campaigns Optimization", *in: CENTRIC 2021*, Barcelone, Spain, October 2021, `https://hal.archives-ouvertes.fr/hal-03424725`.

[7] B. Fabre, T. Urvoy, J. Chevelu, D. Lolive, "Neural-Driven Search-Based Paraphrase Generation", *in: 16th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2021)*, Kiev, Ukraine, April 2021, `https://hal.archives-ouvertes.fr/hal-03540926`.

[8] S. Jafaritazehjani, G. Lecorvé, D. Lolive, J. D. Kelleher, "Style as Sentiment Versus Style as Formality: The Same or Different?", *in: Artificial Neural Networks and Machine Learning – ICANN 2021*, p. 487–499, Bratislava, Slovakia, 2021, `https://hal.archives-ouvertes.fr/hal-03541251`.

[9] J. Mekki, G. Lecorvé, D. Battistelli, N. Béchet, "TREMoLo-Tweets: a Multi-Label Corpus of French Tweets for Language Register Characterization", *in: RANLP 2021 - Recent Advances in Natural Language Processing*, Varna, Bulgaria, September 2021, `https://hal.archives-ouvertes.fr/hal-03331738`.

[10] M. Tchenegnon, T. Le Naour, S. Gibet, "CVM-Net: Motion Reconstruction from a Single RGB Camera with a Fully Supervised DCNN", *in: J.FIG*, Sophia Antipolis, France, November 2021, `https://hal.inria.fr/hal-03536041`.