ERC Proof of Concept Lump Sum Pilot Grant 2019
Part B

Section 1: The idea – Excellence in Innovation potential

ERC StG CREAM (2014-2019) is a cognitive neuroscience research program aiming to understand how speech and music create emotions. The CREAM team has leveraged recent advances in the domain of Artificial Intelligence to create new research tools that are able to control the emotional aspects of speech in real-time conversations: for instance, transform the recording of a voice to make it sound brighter and smiling[1], or to inspire trust or awe[2]. The integration of these techniques in everyday communication devices, vocal servers or personal vocal assistants, opens important societal and commercial perspectives, which the present ERC PoC proposal ACTIVATE (Augmenting the value of Conversations with Voice Transformations) aims to identify, measure and facilitate.


Science Home News Journals Topics Careers

By Tania Rabesandratana | Mar. 30, 2018 , 8:00 AM ISTOCKPHOTO/ STEVE DEBENPORT

Want to sound like someone people can trust? This new software could help

a. Succinct description of the idea to be taken to proof of concept:
a.1- The problem

"Our brand is conversation. Proximity and Service are the two operative words in customer relation: the closer the operators are to the client, the better they understand what the motive is behind the call, and the better service we can deliver" (Frank Guilloteau, VP Customer relation, Mutavie)

In our modern, industrialized, multicultural societies dominated by the service industries, major commercial and societal value lie in conversations happening between stakeholders (clients and businesses, employee and employer, citizen and agencies, patient and caregivers, etc.). Even in an era of email/text messaging, much of such exchanges are spoken conversations (270 millions of conversations per year in US-market call-centers[3]). In the call-center industry, an operator's tone of voice is generally assumed to have a direct, measurable impact on client satisfaction and commercial outcomes[4]. In non-profit sectors such as 911 emergency medical triage, bad communication means increased volume of medical errors and suboptimal use of healthcare resources[5]. With growing pressure to scale such services at ever lower costs, ensuring not only the technical conditions but also the communicative value of conversations is becoming increasingly difficult. The pressure takes an important toll on the workforce, associated with poor work satisfaction, decreased call volume and greater employee turnover.

a.2- The solution

Our proposed solution is to bring to market real-time voice-transformation technologies based on ERC CREAM's research in emotion neuroscience, which can augment the value of spoken conversations by adding business-relevant control on emotional expressivity. For instance, in the context of a call-center conversation, our real-time voice transformations may make an angry client's voice 10% less aggressive, reducing employee fatigue at the end of the day, or make an operator's voice 10% more trustworthy, augmenting customer satisfaction after the call. It is this kind of business hypotheses, made credible by experimental evidence from our scientific research, that ERC ACTIVATE intends to test.

Because they are **transformations,** our solutions do not *synthesize* artificial sounds, but can be applied to any pre-existing voice, either human (e.g. a real call-center operator) or artificial (e.g. the vocal output of a personal assistant), to make them e.g. more relaxed or convincing. Because they are **real-time**, they can be applied transparently to any ongoing conversation with no noticeable delay and can be modified to react

1 https://www.vox.com/science-and-health/2016/1/17/10777304/sound-of-your-own-voice-affect-mood

2 http://www.sciencemag.org/news/2018/03/want-sound-someone-people-can-trust-new-software-could-help

3 Trade magazine En-Contact, March 2016.

4 Rafaeli, A., & Sutton, R. I. (1987). Expression of emotion as part of the work role. Academy of management review, 12(1), 23-37.

5 https://www.france24.com/en/20180509-france-buzyn-womans-death-mocked-phone-emergency-call

immediately to context, e.g. the voice of an assistant can become more or less smiling depending of how the user interacts with them. Finally, because they are **expressive**, they allow to augment conversations with appropriate, scientifically-validated emotions or attitudes that have a direct impact on their commercial or societal outcomes, e.g. customer satisfaction, retention rate, amount of information exchanged, duration, etc.

Value Proposition (1): in service industries such as call-centers which rely on conversation quality between clients and employees for their commercial bottom-line, the innovation solves the user's problem of training and supporting employees in the straining task of actively displaying positive emotions to maximize customer engagement (because these can now be assisted/augmented technologically), or confronting overly negative reactions (because these can now be smoothed out). This solution is potentially valuable for the industry because it is cheaper and easier to deploy than alternative solutions involving employee recruitment and training procedures (estimated at $12k per frontline employee[6]), and because it has (testable) potential to deliver outcomes with more consistency and alleviate employee fatigue.

Value Proposition (2): In services or products that rely on synthesized voices such as personal vocal assistants or augmentative and alternative communication devices (AAC), the innovation addresses the user's demand for highly reactive interactions in a variety of contexts (because transformed emotions can be turned on, off, and everything in between) as well as keeping repetitivity as low as possible (because a single recording can be transformed into a large number of alternative, all appropriate pronunciations). This solution is potentially valuable for the industry because it is an improvement over existing solutions with just a few pre-recorded identities to choose from, because emotional expression in speech synthesis is currently not adaptive in real-time, and because the technology can easily integrate to a variety of speech synthesis solutions by simply modifying their audio output.

British AAC user and standup comedian The Lost Voice Guy, demonstrating the problem of not being able to convey emotions with a speech synthesizer. https://www.youtube.com/watch?v=qNXL4wzvz54&t=5m46s

b. Demonstration of Innovation Potential

The **expected output** of the project is a commercially-viable and societally-acceptable technical solution to increase customer engagement and optimize emotional communication in computer-mediated spoken conversations. Our proposal is **innovative** with respect to existing speech technology in several respects. First, contrary to solutions based on speech synthesis (e.g. digital voice assistants like Amazon Echo), speech transformations do not require to learn voices in advance and can be applied to any vocal identity, natural or artificial, even if they are heard for the first time. Second, contrary to solutions based on pre-recorded speech (e.g. AAC devices like Tobii Dynavox), speech transformations do not require to record voices in different emotional versions, and they can be adapted dynamically depending on conversation context. Third, contrary to all of the above, speech transformations can be applied to conversations in any language (be it French, English or Chinese), without having to redevelop a system from scratch when a new market is being targeted.

Our technology will be used to **improve several types of products**. First, it will be used to improve B2B and B2C services that deliver voice communication over the internet (e.g. Skype-like internet voice calls, call centers, 911 telephone dispatch) allowing them not only to transmit voices but also to transform their emotions in real-time in order to optimize for conversation outcomes that are relevant for their commercial or societal bottom-line. Second, it will be used to improve voice-synthesis systems for voice digital assistants (e.g. Amazon Echo) or care system (e.g. Tobii Dynavox), adapting the output of the speech synthesizer to the user's communication needs.

6 https://qatc.org/winter-2015-connection/exploring-call-center-turnover-numbers/

Section 2 – The Expected Impact:

a. Identification and description of any effect or benefit to the economy, society, culture, public policy/services.

Because of the prevalence of vocal communication in society, our proposed innovation will impact a wide variety of sectors and actors. In **society**, real-time speech transformations have the potential to impact individuals using vocal digital assistants (e.g. we will test the hypothesis that users have a more immersive and enjoyable experience when voices from these products are more emotionally adaptive). The innovation will also impact communications between customers and companies (e.g. we will test the hypothesis that callers have higher customer satisfaction) and between citizens and public-sector agencies.

The emerging market of vocal digital assistants is expected to grow from 390mn in 2015 to 1.8bn users by the end of 2021, with revenue growing from US$1.6bn in 2015 to US$15.8bn in 2021. In 2012, the global market of internet voice calls registered 98.9bn subscribers, for a revenue of US$ 43.27bn, and is expected to increase to 204.8bn users and a revenue of US$ 86.20bn by 2020. *Source: Tractica, 2017*

In the **service industry**, emotion-controlled interactions will impact client-relation front-line employees (we will test the hypothesis that employees will have better work quality and less emotional fatigue).

The global market for call-centers is growing at a CAGR of 8.54% and projected to reach US$407bn by 2022, employing more than 3.4 million people in the EU alone. *Source: Technavio, 2018*

In the **care** sector, real-time speech transformations will impact individuals calling emergency medical services (e.g. we will test the hypothesis that patients will receive more attentive service from call operators) and telephone triage operators (e.g. we will test the hypothesis that medical operators will engage in more effective information seeking with patients). The innovation will also benefit non-verbal patients using AAC devices, e.g. after brain stroke or with conditions like Stephen Hawking's amyotrophic lateral sclerosis, who will be able to communicate with caregivers and family with a more personal and expressive tone of voice (we will test the hypothesis that this will result in increased quality of life and interaction with close ones).

According to the World Health Organization (WHO) Disability & Health report published in November 2016, over a billion people - 15% of the world's population - have some form of disability. The global elderly and disabled assistive devices market was valued at US$12.3bn in 2012 and is predicted to reach US$19.6bn by 2019, expanding at 6.90% CAGR. The segment of text-to-speech assistive technologies was valued at US$1.30bn in 2016 and is expected to reach US$3.03bn by 2022, at a CAGR of 15.21%. *Source : Transparency Market Research, 2018*

b. Outline of the value creation process

Our proposed process to generate the above-listed benefits is the creation of a start-up company (henceforth, the Company), based on a first patent obtained last year in the context of ERC CREAM (EP2018/053433) as well as other IP to be generated in project ACTIVATE. The context of a start-up is considered an appropriate strategy, depending on the outcome of the market analysis and user testing, because the project's core technology is unique, differentiated, and hard to reproduce; the startup will capitalize on its own R&D and be able to adapt the technology to address the needs of a variety of clients.

b.1. Plans to assess and validate the effectiveness of the project's outcomes

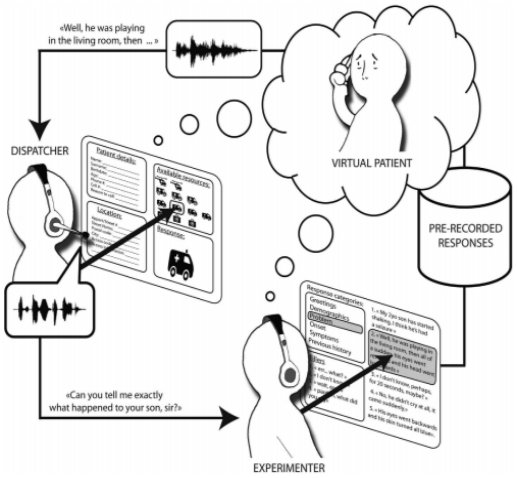
The innovation is based on voice-transformation technologies that are **currently at Technology Readiness Level (TRL) 4**: the basic technological components exist as stand-alone prototypes, and have been validated to work in laboratory conditions. For instance, our technology DAVID[7] is able to modify the height of a spoken voice in real-time to make it sound happier or more relaxed, and was validated to have an emotional effect on 20 native speakers of the English, French, Swedish and Japanese languages. Similarly, our technology ZIGGY[8] is able to simulate the sound of smile in speech, and was validated to trigger automatic 'response' smiles on a sample of 35 native-French listeners. Our proposed objective for the project is to **take these technologies to TRL6**, taking an experimental-science approach to testing our business hypotheses. We will (1) conduct market analyses and interviews of industrial players to identify relevant conversational situations in which the technologies can be tested (e.g. *for a call-center, a customer calling to resiliate their contract*), (2) identify precise conversation outcomes that have market value (e.g. *retention rate after the call*), (3) measure the impact of the voice transformation on these outcomes in a simulated test environment (using the Simulphone® software, in collaboration with software provider Anthropi http://www.anthropi.fr)

7 http://forumnet.ircam.fr/product/david

8 European Patent 2018/053433 - Real-time software amplification of smiling in speech

that is near the desired configuration in terms of performance and user performance and (4) use the technology's measured impact on relevant variables (e.g. *a X% increase of retention rate*) to estimate the value of the Company's minimum viable product (MVP) and write a business plan. TRL6 testing will be documented in the form of technical reports.

The Simulphone® test environment simulates a conversation between a user and a virtual interlocutor (illustrated here, a 911 operator responding to a virtual patient). The user's interaction is free, but the virtual interlocutor's responses are pre-recorded actor voices, selected in real-time by a human experimenter. Project ACTIVATE will use Simulphone to measure the impact of voice transformations on user in relevant business contexts.



b.2. Plans to clarify IPR position and strategy or knowledge transfer strategy

IPRs for the project already include patent EP2018/053433 covering real-time "smile" transformations of speech, created in ERC CREAM by the PI and held by CNRS and co-beneficiary IRCAM and Sorbonne University. More IP is expected to be created during the testing stages of project ACTIVATE : for instance, a new algorithm that, at right moment, automatically triggers a 'smile' transformation in an ongoing conversation to maximize client satisfaction. All IP created during the project will be jointly filed by CNRS, Sorbonne University and IRCAM and exclusively licensed to the Company upon its creation at the end of the project. CNRS Innovation, the knowledge transfer branch of CNRS and a third-party beneficiary of the project, will supervise knowledge transfer to the Company.

The commercialisation of the innovation by the Company will be made either in a software-as-service model (with voice processing time charged on a per minute basis, similarly to e.g. Google Cloud's Speech Recognition product), by licensing algorithmic implementations for integration in a speech-synthesis environment, or by installing software on a third-party web server deployed e.g. by a call-center when low latency and data privacy is required.

b.3. Plans to set up contacts with industrial partners, societal or cultural organisations, or any other potential users

The two markets targeted by ERC ACTIVATE are that of commercial service conversations, i.e. call-centers and digital voice assistants, and care, i.e. AAC devices and tele-medicine. For each market, the project will measure the impact of voice transformations in situations involving both human and synthesized voices:

Market:	Transforming human voices	Transforming synthesized voices
Commercial service	Call center: transform customer's voice and test the impact on operator outcomes (eg. emotional fatigue, work satisfaction), and transform operator's voice and test impact on customer's outcomes (eg. customer satisfation, intention to purchase)	Vocal digital assistants: transform the assistant's synthesized voice and test the impact on user outcomes (eg. satisfaction with information provided, number of daily interactions)
Care	Tele-medecine / telephone regulation : transform patient's voice and test the impact on caregiver outcomes (eg. number of questions asked, medical decision), and transform caregiver's voice and test impact on patient's outcomes (eg. compliance with care)	AAC: transform the patient's synthesized tone of voice and test the impact on caregiver and family outcomes (eg. time spent interaction with patient, quality of social affiliation) as well as on patient (eg. feeling of control, satisfaction with device)

In each market, several partners have been identified as targets for business creation (underlined: contact made prior to the application).

- Call-center market: call-center outsourcers (Webhelp, Teleperformance), call-center operators (Gucci, OUI SNCF), think-tank (INRC Lab).
- Vocal digital assistants: voice synthesis software providers (Acapela, Google Duplex, Amazon Alexa), digital assistant manufacturer (Amazon Echo), assistive call-center services (RogerVoice)
- Telemedecine: medical providers (CHU Lille, SAMU 21), teleconsultation operators (Acetiam, Feelae, Doctolib)
- AAC: synthesis software provider (Acapela), device manufacturer (Tobii Dynavox, Luminaud, Zygo)

Section 3: The proof of concept plan

a. Project-management plan including risk and contingency measures

a.1 - Organisational structure and decision-making

Project ACTIVATE will be implemented in the PI's research team at the Science and Technology of Music and Sound (STMS) lab in Paris, France, and integrate the same governance and infrastructure that has successfully managed the $1.5M ERC CREAM project since 2014. The project's HI is the French **CNRS** (Centre National de la Recherche Scientifique, http://www.cnrs.fr), which whom the PI is a tenured researcher. The main fundamental research organization in Europe, CNRS is a government-funded organization under the administrative authority of France's Ministry of Research. Like ERC CREAM, ERC ACTIVATE will be managed by the "Paris Villejuif" Regional Office of CNRS which is fully experienced in legal, financial and administrative aspects of European project management. Third-party beneficiaries for the project include **CNRS Innovation** (http://www.cnrsinnovation.com), the CNRS's subsidiary company in charge of innovation transfer, **IRCAM** (http://www.ircam.fr), a public research institute and co-operating body of the PI's CNRS laboratory, and **Sorbonne Université** (http://www.sorbonne-universite.fr), also co-operating body of the laboratory.

User tests with Simulphone® will be conducted at the INSEAD-Sorbonne Université Center for Behavioural Sciences (https://www.insead.edu/centres/insead-sorbonne-universite-lab-en), a high-throughput experimental platform for consumer behaviour freely available via third-party Sorbonne Université, which provides state-of-art user-experience facilities (e.g. eye-tracking), anonymous data management and ethical validation via its Institutional Review Board.

Decisions concerning business interviews, valuation and business plan elaboration will be made at majority-voting by PI Jean-Julien Aucouturier (JJA), and team member Marco Liuni (ML) and Gilles Degottex (GD), based on inputs from third-party beneficiary CNRS Innovation, which has the necessary expertise and will be the valorisation unit of host institution CNRS upon creation of the Company. Decisions concerning user testing will be made by PI JJA, based on inputs from test environment software provider Anthropi.

a.2 – Risk management

Description of the risk	Assessment of risk likelihood
1. Failure to make contact with industry player ahead of user testing (e.g. can't find a call-center company willing to co-design a user test with us)	First contacts with key opinion leaders, journalist specialized in client relation and client satisfaction manager from key accounts OUI SNCF and WebHelp have shown a great interest in the technology.
2. Failure to recruit participants for user tests with required skills (eg. can't find enough medically-skilled operators to take part in a patient-calling scenario)	The team has extensive experience recruiting special populations for research, incl. 40 medical doctors for previous study of medical call-centers. Support from INSEAD-Sorbonne Université Center for Behavioural Sciences includes access to a substantial database of volunteers, with varied qualifications.
3. Unable to use software Simulphone for the tests	The team has already used Simulphone for a previous project testing the impact of a voice transformation in a medical call-center scenario, successfully measuring its impact on medical decisions (Boidron et al. 2016).
4. Test results raise ethical issues (eg. voice transformation irrationally increases the believability of news information delivered by a vocal digital assistant)	All user tests will be subjected to prior approval by the INSEAD Sorbonne Université Institutional Review Board. However, even if user tests are conducted ethically, results may suggest a risk for potential misuse of the technology. In case of issue, we will conduct appropriate measures (see a-3)

a.3- Plan for unforeseen events

In case of unforeseen events, the team will evaluate whether it is necessary to reorganize the project's workplan (Table 1.a), based on the decision process detailed above. Any amendment to the workplan will then be submitted to the Project's Officer at the ERCEA. Proposed plans related to risks listed above include:
Risk 1: Fallback on societal player with no commercial interest in the technology (eg. consumer association, patient-family association) and/or design with the best of publicly available knowledge (eg. trade magazines)
Risk 2: Design at least one scenario per market which can be tested on the general population (ex. test the impact of pre-recorded medical operator's voice on would-be patient, rather then the opposite)
Risk 3: Fallback on non-interactive tests where conversations between both actors are pre-recorded, transformed by the technology, and subjected to ratings of external observers (eg. *if you were the customer in this conversation, how would you rate your satisfaction with the call?*)

Risk 4: For all identified effect, gather a user panel to evaluate the societal acceptability of the technology, its possible impact and define constraints for its ethical, societally-acceptable use. Ethical panels will be supervised by the PI with inputs from the INSEAD-Sorbonne Université Ethical committee (see Table 1.a, **Activity 4**).

b. Description of the team

b.1 – Team achievements and experience.

The project's PI, **Jean-Julien Aucouturier** (https://www.linkedin.com/in/jjtokyo, PhD. University Pierre et Marie Curie, Paris 2006), is a computer scientist and cognitive neuroscientist, and a permanent senior researcher with the project's HI CNRS. From 2014 to 2019, JJA has been the PI of ERC CREAM, on which the current proposal is based. Besides academic research, JJ.A. has extensive experience with industrial R&D in the field of consumer electronics in both Europe and Asia, and is the co-inventor of 4 patents (EP2018/053433, US2008/040362, JP2006/106754, WO2006/037786) including one covering technology involved in the current project. Senior team members include **Marco Liuni** (https://www.linkedin.com/in/marco-liuni-50871217/, PhD. University Pierre et Marie Curie, Paris, 2012), a computer scientist, sound designer and researcher with a permanent position at project third-party beneficiary IRCAM; and **Gilles Degottex** (https://www.linkedin.com/in/degottex, PhD. University Pierre et Marie Curie, Paris, 2010), a self-employed computer scientist and expert in voice signal processing.

JJA, ML and GD are foreseen to be co-founders of the Company upon completion of the project.

In addition, the project will hire a sound/user-experience designer (2-5 years experience, 12 mm), to manage ideation during business interviews and subsequent user testing (Table 1.A, **Activity 1-4**).
CNRS Innovation, the CNRS's subsidiary company in charge of innovation transfer and project third-party beneficiary, brings to the project 25-year experience for market analysis (Dr **Camille Foussal**, Director of Scientific and Competitive Intelligence and Mme Clarisse Thibault, Scientific and Competitive Intelligence analyst), IPR management (Dr **Giovanni Altamura**, Transfer & Licensing Manager) and business plan elaboration (M. **Jules Meunier**, Start-up Coaching Unit leader).
ANTHROPI (http://www.anthropi.fr), a software service company, provides training courses for the public sector based on a proprietary telephone conversation software simulator (Simulphone®) that will be used in the project to measure the market value of the project's technologies. Access to Simulphone will be charged to the project by subcontracting. Dr **Laurent Boidron,** CEO, will provide assistance on using the technology.

b.2 – Roles, strengths and weaknesses

JJ.A. will supervise all projects operations, and conduct user testing and ethical panels (Table 1.a, **Activity 2-4**); his salary will be supported at 20% on the project. ML will supervise actor interviews, script writing and voice actor recordings (**Activity 1-2**); his salary will not be supported on the project. GD will supervise the technical integration of the project's technologies in the test simulator, with assistance from Anthropi (**Activity 2**); his salary will not be supported on the project. In addition, both ML and GD will be involved in business plan development (**Activity 5**). CNRS Innovation's CF, CT, GA and JM. will provide assistance on industry player interviews (**Activity 1**), business plan elaboration (**activity 5**) and knowledge transfer strategy (**Activity 6**).
Strengths: PI JJA and senior member ML have been involved in the development and laboratory testing of the project's core technologies in ERC CREAM, and these skills will be essential for the experimental approach to measuring the technology's market value in project ACTIVATE. In addition, GD has been involved in the C++ reimplementation of the project's core technologies, and so is ideally positioned to supervise their deployment in the test environment used in the project. JJA and Anthropi have already validated the use of Simulphone® in previous work to test the impact of voice transformations in the context of emergency medical services[9].
Weaknesses: As of October 2018, the three co-founders JJA, ML and GD lacked experience in business development. This situation has been addressed by participating in an intensive entrepreneurship training course (DeepTechFounders, Oct.18-Jan.19) and securing personalized coaching by CNRS Innovations' RISE program (http://www.cnrsinnovation.com/rise/ Jan.19-onwards). In addition, the participation in project ACTIVATE of team members CF, CT, GA and JM from CNRS Innovations ensures the needed expertise is available.

9 Boidron, L., Boudenia, K., Avena, C., Boucheix, J.M. & Aucouturier, JJ. (2016) *Emergency medical triage decisions swayed by manipulated cues of physical dominance in caller's voice*, Scientific Reports, 6, 30219. https://www.nature.com/articles/srep30219

c. Plan of the activities

c.1 – Resources: Grant takes the form of the standard 150,000 EUR lump sum pre-fixed by EC decision for the 2019 ERC PoC call. We confirm that: (1) subcontracts will be best value for money and free of conflict of interest and (2) all beneficiaries have followed their own accounting practices for the preparation of the budget and have included therein only costs that would be eligible for an actual costs grant, excluding costs that are ineligible under the H2020 rules.

c.2 – Description of the work:

Table 1.a : Action description

Action title	Proof of Concept

Objectives: Test the business hypothesis that real-time emotional voice transformation technologies can augment the commercial and societal value of conversations in the four markets of call-centers, digital voice assistants, AAC devices and tele-medicine, and elaborate a business plan and IPR strategy based on the measured indicators of value, for a startup company to be created upon completion of the project.

Description of the work and resources planned for each activity

Activity 1: Commercial player interviews (Months 1-4)

Make contact with potential business partners in the four identified markets of call-centers, digital voice assistants, AAC devices and tele-medicine (see Table p.4). In each market, conduct interviews and small-group expert panels to identify (1) one or two business-critical conversational situations in which the technologies can be tested (e.g. *for a call-center, a customer calling to resiliate their contract*) and (2) measurable conversation outcomes that have market value (e.g. *retention rate after the call*).

Resources planned: Host institution CNRS will report (1) personnel costs for one sound/user-experience designer (2-5 years experience, 12 mm, est. 40,000€), hired to manage ideation during small-group panels as well as subsequent user testing (**Activity 2-4**) and (2) travel expenses (est. 10,000€) including transportation and accommodation for team members to visit national and international business partners. Third-party CNRS Innovation will report (3) costs of 12 000€ for market analysis and organizing business interviews.

Activity 2: Design user testing (Months 4-6)

Follow the methodology of Boidron et al. (2016). Write a conversation script that instantiates the situations identified in **Activity 1** and co-designed with business partners (eg. *write all possible answers of a call-center operator responding to a customer calling to resiliate their contract*). For scenarios involving transforming a human voice (eg. *call center*), hire a voice actor to record all sentences, for use in software Simulphone®. For scenarios involving transforming a synthesize voice (eg. *interaction with a pre-recorded AAC user*), use business partner's speech synthesis software (or, alternatively, a freely-available state-of-art software) to generate recordings. Use the project's technology to produce two versions of these recordings, one with and one without voice transformation (eg. *operator's responses with or without a smile transformation*). Adapt the Simulphone user-interface so that the test participant can report ratings on the relevant measures identified as part of **Activity 1** (eg. *a 7-point Likert scale of how likely the customer is to not resiliate their contract at the end of the simulated call*).

Resources planned: Host institution CNRS will report (1) personnel costs for sound designer (same as **Activity 1**), (2) Voice actor fees (100 sentences, 200€ per scenario, est. 1000€), (3) subcontracting costs to software provider Anthropi for assistance in using software Simulphone® (est. 20000€). Third-party IRCAM will provide (3) free access to audio engineering facilities (studios, microphones, etc.) as part of the project's structure costs.

Activity 3: user testing / data collection (Months 6-9)

Collect N=20 participant responses to each of the simulated scenarios constructed in **Activity 1-2**. Participants may be recruited as part of the general population (eg. *for scenarios involving a customer's response to pre-recorded call-center operators*), or in a group with expert skills, possibly recruited among employees of business partners involved in **Activity 1** (eg. *for scenarios involving a call-center operator's response to pre-recorded customers*). All user tests will be conducted at the Sorbonne Université-INSEAD behavioral plateform, user participation will obey to EU ethical standards on human data collection (informed consent, anonymous data management, ethical validation). Data analysis to test the hypothesis of a positive impact of the voice transformation on the predefined conversation outcomes using statistical tests (eg. *a 10% increase in likelihood to not resilitate a contract after the call*). Write a report on the user test, to document reaching TRL6.

Resources planned: Host institution CNRS will report (1) personnel costs for sound designer (same as **Activity 1-2**), (2) Participant fees (20€/hr, est. 3000€) for participation in data collection. Third-party Sorbonne Université will provide (3) free access to the Sorbonne Université-INSEAD behavioral plateform facilities and service for data management and ethical validation, as part of the project's structure costs.

Activity 4: Ethical focus group (Months 9-10)

For each effect identified as part of **Activity 3,** gather an independent group of users and ethics experts to evaluate the societal acceptability of the technology, its possible impact and define constraints for its ethical, societally-acceptable use. Write a report with ethical recommendations, to be considered in Activity 5.

Resources planned: Host institution CNRS will report (1) personnel costs for sound designer (same as **Activity 1-3**), (2) Participant fees (20€/hr, est. 1000€) for participation in the panel.

Activity 5: Business plan elaboration (Months 9-12)

Use the technology's impact on relevant variables (e.g. a X% increase of retention rate) measured in Activity 3, as well as ethical recommendations of Activity 4 to estimate the value of the Company's minimum viable product (MVP) and write a business plan. The business plan will include a) a business, technology and team presentation; b) how the technology developed could solve a problem and impact a given market; c) a development plan and a risk analysis; d) a business model suitable to the markets targeted; e) the establishment of a three years budget and financing estimated plan, and f) a relevant financial strategy for the project, listing the potential grants and contests attainable and the main business angels and capital ventures partners which could be useful to mentor and potentially fund the project.

Resources planned: Third-party CNRS Innovation will report (1) costs of 4000€ for technology transfer and market valorisation roadmap.

Activity 6: Clarifying IPR and knowledge transfer (Months 9-12)

Identify possible targets for IPRs in work produced as part of the project, and draft a proposal to transfer knowledge to the Company, to be created upon completion of the project.

Resources planned: Third-party CNRS Innovation will report (1) costs of 4000€ for IP position analysis and drafting a knowledge transfer proposal.