AAPG2021	ACCENT	PRCE
Coordinated by:	Karim MIGNONAC, Toulouse School of Management	48months
8.13. Industrie et usine du futur : Homme, organisation, technologies		

The Augmented Call-Center:

A laboratory for the cognitive, ethical and organizational impact of voice technology in new service industries

I. Pre-proposal's context, positioning and objective(s)

Project's objectives and research hypotheses

The global health and economic crisis triggered by the COVID-19 pandemic, and the social distancing policies that ensued, have compelled many manufacturing and service organizations to switch part or all of their activities online and to rely on distant voice-based operations (Mckinsey, 2020). This recent development amplifies a more general trend across all industries (ranging from online selling to telemedicine) towards the use of distant conversation systems and call-center-like operations for all organizational interactions, both internally (e.g. work meetings, employee HR services) and externally (e.g. B2B and B2C relations). This **ever-increasing reliance on distant voice-based operations** (Groth et al., 2019) has potentially important, yet largely uncharted, implications for the performance of organizations and the well-being of their members.

A deeper understanding of the impact of voice-based operations is more needed than ever given that the expansion of such operations is accompanied by an **unprecedented development of artificial-intelligence (AI)-based voice technologies** (Huang & Rust, 2018), which creates irresistible opportunities for voice synthesis (Oord et al. 2016) or real-time transformations of an employee's tone of voice, accent, sex or even language (Arias et al. 2020). Organizations are therefore faced with a situation where (1) an increasingly important part of their value is generated by distant, voice-based operations, (2) these occur in a technological landscape where potentially any conversational parameter can be optimized and personalized with voice technology, and yet (3) **almost nothing is known about how these conversational parameters impact the outcome of the conversation** for call-center employees and their organization.

ACCENT brings together an interdisciplinary group of researchers in management science, human & social sciences and engineering, together with a major player in the customer-relation industry (Comdata) and an emerging voice-technology startup (Alta Voce) to build an '**augmented call-center**' and use it as a laboratory to explore the human and organizational impact of voice technologies in the service industry.

In a first-of-its-kind initiative, the project will support the **deployment in a real call-center of a novel voice technology** able to transform the emotional tone of voice of employees and customers to make them sound, for instance, more smilely (Arias et al., 2020) or less aggressive (Liuni et al. 2020). We will then use this technological innovation to conduct **high-volume, controlled A/B experiments** testing the impact of voice parameters on productivity, customer satisfaction and employee well-being, in both real and simulated calls. Finally, we will clarify the societal implications of deploying such voice technologies by **conducting experiments in behavioral ethics** among stakeholders of the organization and in the general public, the results of which will inform future managerial decisions and public policy.

Project ACCENT is therefore a cross-disciplinary proposal with an industrial purpose that takes into account both the technological and human factors. It aims at answering key unanswered questions and making a significant breakthrough in the design of the service systems of the future.

Position of the project in relation to the state of the art

Previous research offers relatively disconnected disciplinary insights into the role of voice parameters in service operations. The field of management science, notably, has a long history of studying factors that facilitate positive outcomes in voice-based operations, but it typically does so by relying either on simulated environments or on observational studies. For instance, to study customer reactions to operators'

emotional displays, researchers may design a simulated call-center in a university (e.g., Groth, Hennig-Thurau & Walsh, 2009) where fake operators/participants are instructed to control their amount of smiling while they interact with other participants pretending to be customers. When they study real operations, authors typically rely on cross-sectional studies, in which customer aggressivity is correlated to employee's performance on the timescale of months (e.g., Sliter, Sliter & Jex, 2013). These approaches have either questionable relevance to real business situations, or, because they are only correlational, **fail to provide direct insights regarding the impact of voice-associated variables** in a call.

Conversely, while the field of cognitive science uses more direct interventional methods to investigate vocal communication, it tends to study only immediate perceptive judgements in a non-real-time context and in relatively small cohorts of participants (Boidron et al. 2016). For instance, Arias et al. (2020) manipulate recordings of 10 short spoken sentences with an acoustic effect designed to simulate the 'sound' of smile, and found that a group of 30 listeners judged these sentences as happier and reacted to them by smiling. However, without incorporating voice transformations at the scale of a real call-center, it is **unclear how these results transfer to real-business operations**, with real employees and customers making decisions over the period of several hours or days.

Finally, there is a substantial amount of work engaging with the ethics of AI (Devillers et al. 2020; Bonnefon, Shariff & Rahwan 2020), notably with experimental approaches attempting to crowdsource solutions to its ethical dilemmas (Awad et al. 2018). These approach are often limited by the availability of the technology under scrutiny, such as fully autonomous self-driving cars (Bonnefon, Shariff & Rahwan, 2016) or brain stimulation (Medaglia et al. 2019). Without direct access to this technology, researchers and participants can only design and take part in low fidelity experiments — which **do not always confront the most urgent and realistic moral dilemmas, with the most direct industrial and organisational implications**.

Innovative nature of the project, ambitiousness and originality

Project ACCENT departs from this state of the art by being a truly interdisciplinary endeavour and not the mere juxtaposition of separate disciplinary insights. Rather, it is built on the fundamental premise that the dialogue between disciplines leads to the most exciting discoveries and enhances communication between the scientific community and society at large. First, by using novel voice technologies to directly manipulate how a given voice is heard by customers and operators, project ACCENT will improve on current research in management science and provide **direct causal evidence of the influence of voice on human and organisational outcomes**. Second, by incorporating voice transformations into real-world call-center operations, thanks to a research partnership with one of the major players in the customer relation industry, project ACCENT will improve on current research in cognitive science and collect data on **100,000s of calls, 10,000s of hours of real-business operations, and 100s of organization employees**, thus promising inferences of unprecedented statistical power and organisational relevance. Finally, by designing surveys based on realistic moral dilemmas stemming from actual operational situations, project ACCENT will improve on current research in behavioral ethics and build a method for incorporating ethical insights into organizational decision-making, legislation and public policy.

Methodology to reach the scientific objectives

WP1. Development and integration of voice transformations in a real call-center: Work in WP1, led by partner AltaVoce, consists of incorporating real-time voice transformation in the telephony system used in the call-center, in a way that will allow the transformation of either the outgoing voice of operators when they talk to customers or the incoming voice of customers. The consortium has already conducted **proof-of-concept tests showing that this was feasible**: (1) a real-time algorithmic transformation of vocal smile already exists (a research outcome of previously-funded ANR project REFLETS – Arias et al. 2020), and has been licensed by CNRS to partner Alta Voce; (2) a minimal version of the effect has already been tested in operations at partner Comdata. Work will include software development to facilitate the installation and management of the voice transformations in hundreds of call-center workstations, development of a web-based dashboard to manage the programmable activation/deactivation of effects for A/B testing of WP2, as well as research to optimize the existing transformations and design an additional voice transformation effect able to reduce customer aggressivity (Liuni et al., 2020).

WP2. Controlled experiments in real-life call-center operations: Work in WP2, led by project coordinator Toulouse School of Management, consists of using the technical development of WP1 to conduct experiments testing the impact of voice technologies in real-life call-center operations. They will be conducted on samples of several 100s of call-center operators, tested in A/B tests with and without voice transformation over periods of several weeks. Data will include standard call-center performance metrics (e.g., sales, customer satisfaction) and custom measures of employees' attitudes and wellbeing (e.g., employee exhaustion, job satisfaction). The consortium has **already collected pilot data** testing the impact of making operators' voice more smilely in an outbound sale campaign, **suggesting medium-to-high effect sizes** on sales rates. Importantly, resources are provisioned in the project budget to facilitate the participation of personnel at partner Comdata (e.g. compensating lost bonuses, etc. for time spent participating in research) and to develop ethically appropriate ways to collect participant's informed consent, especially in experiments involving outbound calls to real customers (e.g. automated vocal messages asking for consent prior to calls).

WP3. Laboratory experiments in a simulated call environments: Work in WP3, led by partner FEMTO-ST, consists of using the technical developments of WP1 to build an open-source tool able to conduct video calls between remote research participants, while their voice is being manipulated by real-time transformations. The tool will be used to conduct laboratory experiments extending the semi-experimental studies of WP2. In addition to mitigating risk due to the operational constraints on WP2, advantages of doing so are substantial, incl. (1) simulating organizational contexts not covered by WP2 (e.g., job interviews, telemedicine; Boidron et al. 2016), (2) full experimental control on participant recruitment (e.g. non-Western populations) and (3) possibility to add physiological measurements (eg. participants' heart-rate variability in response to stress – Massaro & Pecchia, 2019).

WP4. Behavioral ethics: Work in WP4, led by project coordinator TSM, consists in conducting experiments of behavioral ethics in which large numbers of online participants are asked to judge the moral acceptability of situations involving voice transformations in call-center operations, described with text vignettes (Bonnefon, Shariff & Rahwan, 2016). Vignettes will be designed to match situations and parameters measured in WP2 and illustrated with excerpts of actual conversations, where participants may be WP2 stakeholders (employees, customers) or the general public. Experiments will focus on documenting tradeoffs between organizational benefits and transparency (e.g. using accent-removal technology to increase workplace inclusion vs moral unacceptability of racial prejudice) and aim to inform organizational decision-making and public policies, notably by contributing to the French *Comité national pilote d'éthique du numérique* (CNPEN).

<u>Ability of the project to address the research issues covered by theme 8.13</u>

As Groth and colleagues underlined (2019), the service industry increasingly dominates the global economy and "*the provision of excellent customer service has become a competitive advantage and a core competency toward which most organizations strive*" (p. 90). To improve service encounters, organizations increasingly utilize new technologies – including assistive voice – to augment or replace the role of human employees. These technologies and their application within service delivery systems are increasingly blurring the boundaries between "*physical, digital and biological spheres*" (Huang and Rust, 2018, p. 155) and raises many scientific and organizational challenges; for example, to what extent does increased adoption of assistive voice technologies have on employee well-being and productivity? Will new assistive voice technologies accelerate business model transformation in services and have a long-term influence on demand for human employees? What is the social acceptance, resistance, or rejection of these technologies among the organization's stakeholders?

The ACCENT project aims to overcome these obstacles. In doing so, it strongly fits into the ANR Theme 8.13. "*Industry and the factory of the future: People, organisations, technologies*" which encourages researchers to pursue cross-disciplinary approaches, and calls for the integrated application of a diverse range of expertise and knowledge. Specifically, ACCENT aims to support, understand and boost the development of new digital technologies to meet the demand for innovative, personalized and optimized services, "*while reducing the cognitive and physical workloads of operators*."

II. Partnership (consortium or team)

The ACCENT consortium brings together two academic partners specialized in management science, social science and engineering, one emerging voice-tech startup and a major player of the call-center industry.

Coordinator Karim Mignonac is Professor of Organizational behaviour at Univ. Toulouse 1 Capitole, and director of Toulouse School of Management Research. His research interests include workplace commitment and well-being and his work has been published in journals incl. Strategic Management Journal, Journal of Management, and Journal of Organizational Behaviour, among others. He has been PI of several research projects, incl. ANRs FAIRHEALTH (2018–22), FRANBLE (2012–16) and FRANMIX (2008–11).

Partner 1 (TSM Research) in Toulouse (UMR5303, CNRS, Univ. Toulouse 1 Capitole) is one of the 3 French laboratories in management accredited by the CNRS (along with HEC, and Paris Dauphine University). Prof. **Karim Mignonac** and Dr. **Sarah Boujendar** (MCU, *Prix de thèse Association Francophone de Gestion des Ressources Humaines, 2019*) will supervise TSM's involvement in WP2, bringing expertise in the study of organizational behaviour, and vocal agressiveness in call-center operations. Dr. **Jean-François Bonnefon** (Directeur de recherche CNRS, *Médaille de bronze CNRS 2008*) will supervise TSM's involvement in WP4, bringing expertise in social sciences and behavioral ethics, as well as experience leading a well-known international research program in the ethics of autonomous vehicles which lead to publications in Science (2016), Nature (2019) and PNAS (2020).

Partner 2 (FEMTO-ST Institute) in Besançon (UMR6174, Université de Franche–Comté, CNRS, ENSMM, UTBM) is one of the country's largest engineering and system science research unit. Dr. **Jean-Julien Aucouturier** (Directeur de recherche CNRS, *Prix Emergence Scientifique de la Fondation pour l'Audition 2018*), previously from computer music institute IRCAM in Paris, will supervise FEMTO-ST's involvement in WP3, bringing expertise in real-time voice transformations and cognitive science. The project will also tap into the expertise of FEMTO-ST's RECITS team, specialized in the sociology of industrial change.

Partner 3 (Alta Voce) in Paris (http://alta-voce.tech) is an emerging voice-tech startup specializing in real-time emotional voice transformations for the customer relation and entertainment industries, originated in computer music institute IRCAM in Paris, and supported by the CNRS Innovation program and BPI French Tech funding. Dr. **Marco Liuni** (Chief Design Officer) and Dr. **Gilles Degottex** (CTO) will supervise Alta Voce's involvement in WP1, bringing expertise in real-time voice transformations and web service development as well as prior experience working with partner Comdata to deploy technologies in the call-center environment.

Partner 4 (Comdata Group France) in Gennevilliers (https://france.comdatagroup.com). The Comdata Group is a leading global service provider in customer Interaction, employing more than 50,000 employees and operating more than 100 call-centers in 22 countries. Comdata France is the country's third biggest operator. **Vincent Premat** (Account manager) will supervise Comdata's involvement in the project, bringing expertise in call-center operations and innovation, and prior experience working with partner Alta Voce.

III. References related to the project

Arias, P., Soladié, C., Bouafif, O., Roebel, A., Séguier, R. & **Aucouturier**, JJ. (2020) Realistic manipulation of facial and vocal smiles in real-world video streams. IEEE Transactions on Affective Computing, Vol. 11(3) • Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., **Bonnefon**, J. F., & Rahwan, I. (2018). The moral machine experiment. Nature, 563, 59-64. • Boidron, L., Boudenia, K., Avena, C., Boucheix, J-M. and **Aucouturier**, J.J. (2016) Emergency medical triage decisions are swayed by computer-manipulated cues of physical dominance in caller's voice. Scientific Reports 6, 30219 • **Bonnefon**, J. F., Shariff, A., & Rahwan, I. (2020). The moral psychology of AI and the ethical opt-out problem. In S. M. Liao (Ed.) The Ethics of Artificial Intelligence. Oxford University Press. • **Bonnefon**, J. F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. Science, 352(6293), 1573-1576 • Devillers, L., Kawahara, T., Moore, R. K., & Scheutz, M. (2020). Spoken Language Interaction with Virtual Agents and Robots (SLIVAR): Towards Effective and Ethical Interaction. Dagstuhl Reports, Vol. 10, No. 1. • Groth, M., Hennig-Thurau, T., & Walsh, G. (2009). Customer reactions to emotional labor: The roles of employee acting strategies and customer detection accuracy. Academy of Management Journal, 52, 958 –974 • Groth, M., Wu, Y., Nguyen, H., & Johnson, A. (2019). The moment of truth: A review, synthesis, and research agenda for the customer service experience. Annual Review of Organizational Psychology and Organizational Behavior, 6, 89-113. • Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. Journal of Service Research, 21(2), 155-172. • **Liuni**, M., Ardaillon, L., Bonal, L., Seropian, L. & **Aucouturier,** JJ (2020) ANGUS: Real-time manipulation of vocal roughness for emotional speech transformation. arxiv 2008.11214 • McKinsey (2020) Coronavirus' impact on service organizations: Weathering the storm. McKinsey report, 29 April 2020 • Massaro, S., & Pecchia, L. (2019). Heart rate variability (HRV) analysis: A methodology for organizational neuroscience. Organizational Research Methods, 22(1), 354-393 • Medaglia, J. D., Yaden, D. B., Helion, C., & Haslam, M. (2019). Moral attitudes and willingness to enhance and repair cognition with brain stimulation. Brain stimulation, 12(1), 44-53 • Oord, Aaron van den, Sander Dieleman, Heiga Zen, Karen Simonyan, Oriol Vinyals, Alex Graves, Nal Kalchbrenner, Andrew Senior, and Koray Kavukcuoglu (2016) "Wavenet: A generative model for raw audio." arXiv preprint arXiv:1609.03499 • Sliter, M., Sliter, K., & Jex, S. (2012). The employee as a punching bag: The effect of multiple sources of incivility on employee withdrawal behavior and sales performance. Journal of Organizational Behavior, 33(1), 121-139.