Brain neurophysiology to objectify the social competence of conversational agents

Thierry Chaminade, Birgit Rauchbauer, Bruno Nazarian, Morgane Bourhis, Magalie Ochs, Laurent Prévot

▶ **To cite this version:**

Thierry Chaminade, Birgit Rauchbauer, Bruno Nazarian, Morgane Bourhis, Magalie Ochs, et al.. Brain neurophysiology to objectify the social competence of conversational agents. Human-Agent Interaction, Dec 2018, Southampton, United Kingdom. 10.1145/123_4 . hal-01973552

HAL Id: hal-01973552
https://hal.science/hal-01973552v1

Submitted on 22 Jan 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Brain Neurophysiology to Objectify the Social Competence of Conversational Agents

Thierry Chaminade
Institut de Neurosciences de la Timone
CNRS – Aix-Marseille Université
France
thierry.chaminade@univ-amu.fr

Birgit Rauchbauer
Institut de Neurosciences de la Timone
CNRS – Aix-Marseille Université
France
birgit.rauchbauer@univ-amu.fr

Bruno Nazarian
Institut de Neurosciences de la Timone
CNRS – Aix-Marseille Université
France
bruno.nazarian@univ-amu.fr

Morgane Bourhis
Institut de Neurosciences de la Timone
Aix-Marseille Université
France
morgana.bourhis@univ-amu.fr

Magalie Ochs
Laboratoire Informatique et Systèmes
CNRS – Aix-Marseille Université
France
magalie.ochs@univ-amu.fr

Laurent Prévot
Laboratoire Parole et Language
CNRS – Aix-Marseille Université
France
laurent.prevot@univ-amu.fr

ABSTRACT

We present an approach to objectify the social competence of artificial agents using human brain neurophysiology. Whole brain activity is recorded with functional Magnetic Resonance Imaging (fMRI) while participants discuss either with a human confederate or an artificial agent. This allows a direct comparison of local brain responses, including deep brain structures invisible to other neuroimaging techniques, as a function of the nature of the interlocutor. The present data (9 participants, artificial agent is the robotic conversational head Furhat controlled with a Wizard of Oz procedure) demonstrates the feasibility of this approach, and results confirm an increased activity in the hypothalamic region when interacting with a human compared to an artificial agent.

CCS CONCEPTS

• Human-centered computing → Human computer interaction

KEYWORDS

Humanoid Robot, Social Interaction, Neurophysiology

ACM Reference format:

Thierry Chaminade, Birgit Rauchbauer, Bruno Nazarian, Morganie Bourhis, Magalie Ochs, Laurent Prévot. 2018. Brain Neurophysiology to Objectify the Social Competence of Conversational Agents. In Proceedings of ACM Human Agent Interaction HAI '18. ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3284432.3287177

1 INTRODUCTION

While artificial agents are increasingly present in contact with humans, the assessment of how they are accepted as social partners by their users is very poor. The social acceptance of robots is usually addressed with questionnaires, for example the Negative Attitude towards Robot Scale (NARS, Nomura *et al.*, artificial agents through the comparison of behavioural and physiological responses between conditions in which one naïve person has similar natural interactions with a fellow human or an artificial agent. This allows us to distinguish not only behavioural, but also, and more importantly, neurophysiological features that are specific to interacting with a human to those that are preserved when the interacting agent is artificial (see also Wykowska, Chaminade, Cheng, 2016). The experimental approach has been presented in details in a recent publication describing its foundations and implementation for the behavioral experiment (Chaminade, 2017), and a forthcoming publication will describe its adaptation to the MRI environment (Rauchbauer et al., in revision). We will focus here on the added value of functional magnetic resonance imaging (fMRI) to objectify the social competence of artificial agents with this experimental approach.

In a nutshell, a naive participant is recorded in the MRI scanner with an experimental setup using a videoconferencing to support conversation with another agent, a human or a robot, located outside of the scanner room. We use a cover story to provide credible, but spurious, explanations for the experimental set-up (as in Chartrand and Bargh, 1999). A conversational robotic head form Furhat robotics (Al Moubayed *et al.*, 2012) was chosen for its ability to produce embodied conversation.

2 METHODS

Nine native French speakers, with no history of neurological or psychiatric disease, volunteered for this experiment. The experiment was presented as a study for an upcoming advertising campaign. They would be presented with campaign images and were instructed to discuss them with a fellow human (actually a confederate) and a conversational robotic head (Furhat controlled by the confederate through a Wizard of Oz procedure) located outside the MRI. For this WoZ procedure, prescripted conversational utterances based on actual human conversations of the same images (recordings described in Chaminade, 2017) were launched by clicking buttons on a touch screen.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
HAI '18, December 15–18, 2018, Southampton, United Kingdom

Two sets of each three stimuli were used in this experiment. Each trial started by one image, followed by a 3-second black screen. They then had one minute of conversation during which they talked alternatively either with the human or with the robot. Overall, 3 images from one dataset were presented once before discussing with each of the two agents in each recording run, resulting in a total of 6 minutes of conversation. 4 runs were recorded consecutively (details in Rauchbauer *et al.*, in revision).
MRI data was collected with a 3T Siemens Prisma. Blood oxygen level-dependent (BOLD) sensitive functional images were acquired using an EPI sequence in the 4 sessions. Analysis was performed with SPM 12 running on Matlab, according to standard procedures. Preprocessing entailed correcting for delays in slice acquisition ("slice timing"), realignment and unwarping of the magnetic field distortion, and smoothing with a 5-mm full-width half maximum kernel. First level intra subject statistical analysis was performed in each participant's native space. Normalization of beta estimates was based on DARTEL procedure. Second level analysis used ANOVA with factors Subject, Session and Agent.

3 RESULTS

The main effect of conversation (Figure 1), including trials from both agents revealed activation clusters in the superior temporal gyrus (STG) bilaterally, extending well into the superior temporal sulcus (STS) and dorsally to the ventral premotor cortex (vPM); into the depth of the central sulcus; in the left middle temporal gyrus (MTG); in the extrastriate body area (EBA) and temporo-occipital junction (TPJ) in the right hemisphere; and finally in the left inferior frontal gyrus (IFG) corresponding to Broca's area, involved in language perception and production. The contrast Human *vs* Robot yielded increased response in the hypothalamus, the amygdala and left STG. The reverse contrast revealed a large number of activated clusters, in visual primary and associative areas as well as in parietal and frontal associations cortices.

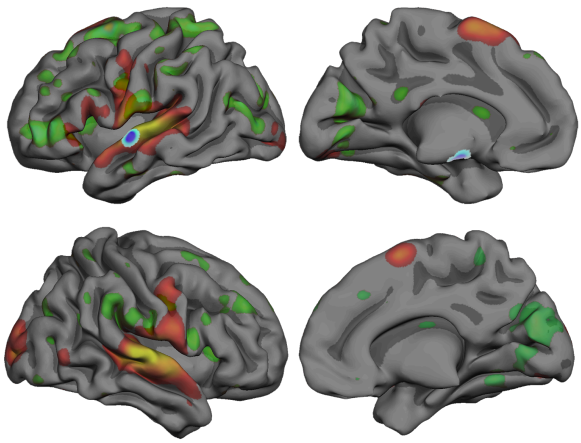

Figure 1: Lateral and medial renders of the activated clusters found in the main effect (Human *&* Robot, hot scale), in Robot *vs* Human (green) and Human *vs* Robot (blue).

4 DISCUSSION

It was largely expected, yet the finding of highly similar neural resources for processing the natural conversation with the human or the robot agent further justifies the use of anthropomorphism in the design of artificial agents. Sensorimotor loops in verbal interactions involve the temporal cortex (STG, STS, MTG) that comprises the primary auditory cortex as well as areas devoted to speech perception. Articulatory control is reflected in the activity of the ventral part of the primary motor cortex in the central sulcus and the frontally adjacent left vPM and IFG. Altogether these results justify the use of anthropomorphic appearance and behaviour for designing interactive artificial agents: sensorimotor processes during interaction rely on the same neural mechanisms than for natural human interactions.
Increased activity in when interacting with the Robot *vs* the Human is found in many brain areas. A number of these clusters can be explained by the appearance (including its conversational behaviour) of the robot. While this could appear paradoxical at first glance, it was reported when the human brain processes sensory information coming from artificial anthropomorphic agents and can be interpreted in terms of an error signal signalling breach of expectations (for instance, response of the fusiform face area reflects the fact that the face of the agent in imperfectly mimicking a human face, Chaminade *et al.*, 2010). The measure of this increased activity in areas involved in face perception - the same applies to other features such as the voice - when perceiving different artificial agents could be used to index the objective distance between these different renders and the intrinsic template of canonical human face in the perceivers' brain.
The hypothalamus and the amygdala were found in the contrast Human *vs* Robot. These subcortical responses can be thought of as reflecting an autonomic drive for social interactions, in contrast to conscious sensorimotor processing, illustrated by the numerous cortical clusters, involved when interacting with the robot.

5 CONCLUSION

In this paper, we report the first attempt to record the neurophysiology of a natural conversation with a fellow human or an artificial agent with fMRI. Results suggest that despite similarities in sensorimotor processing, expected given the anthropomorphism of interactive artificial agents, participants adopted different stances depending on the agent they were interacting with. Confederates experimenters (MB and TC) observed many differences in the participants' behaviour, and it is likely that the inclusion of descriptors of behaviour (ongoing work) will be useful characterize more accurately the relations between behaviours and local neurophysiology. This experimental approach paves the way for an objective assessment of artificial agents social competence during natural social interactions.

ACKNOWLEDGMENTS

Support from grants ANR-16-CONV-0002 (ILCB) and ANR-11-LABX-0036 (BLRI) and the Excellence Initiative of Aix-Marseille University - A*Midex, a French "Investissements d'Avenir" programme (AAP-ID-17-46-170301-11.1).

REFERENCES

[1] Al Moubayed, S., Beskow, J., Skantze, G., & Granström, B. (2012). Furhat: a back-projected human-like robot head for multiparty human-machine interaction. In Cognitive behavioural systems (pp. 114–130). Springer.

[2] Chaminade, T. (2017). An experimental approach to study the physiology of natural social interactions. *Interaction Studies*, *18*(2).

[3] Chaminade, T., Zecca, M., Blakemore, S. J., Takanishi, A., Frith, C. D., Micera, S., Dario, P., Rizzolatti, G., Gallese, V., & Umilta, M. A. (2010). Brain response to a humanoid robot in areas implicated in the perception of human emotional gestures. *PLoS One, 5*(7), e11577.

[4] Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of personality and social psychology*, *76*(6), 893

[5] Nomura, T., Suzuki, T., Kanda, T., & Kato, K. (2006). Measurement of negative attitudes toward robots. *Interaction Studies*, 7(3), 437-454

[6] Rauchbauer, B. Nazarian, B., Bourhis, B., Ochs, M., Prévot, L. & Chaminade, T. (in revision). Neural bases of social interactions investigated with an artificial agent. *Phil. Trans. R. Soc. B.*

[7] Wykowska, A., Chaminade, T., & Cheng, G. (2016). Embodied artificial agents for understanding human social cognition. *Phil. Trans. R. Soc. B*, *371*(1693)