# Preliminary study to evaluate Emi emotional interaction with two young children

Sebastien Saint-Aimé, Brigitte Le Pévédic and Dominique Duhaut
*Valoria Laboratory*
*University of Bretagne Sud*
*56 000 Vannes, France*
{*sebastien.saint-aime,rigitte.le-pevedic,dominique.duhaut*}*@univ-ubs.fr*

***Abstract*— This article will focus on research in the field of emotional interaction, for the EmotiRob project, to maintain non-verbal interaction with chidren from 4 to 8 years old. The studies carried out on perception and emotional synthesis have allowed us to develop an experimental stuffed robot, Emi, using an emotional model, iGrace, allowing for an emotional reaction based on the speech of the user. Our iGrace model can also be used with all kinds of systems which have emotion in the environment. We begin the article with a presentation of the MAPH and EmotiRob project. Then, we will quickly describe the different hypotheses we have used for the iGrace emotional model and the Emi robotics conception. We conclude with a preliminary study we began with 2 children to evaluate interaction with Emi.**

***Index Terms*— Interaction model; emotion; companion robot.**

## I. Introduction

A new challenge in Robotics is to create systems capable of behaviour enhancement due to their interaction with humans. Research work in psychology has shown that facial expressions play an essential role in the coordination of human conversation [1] and constitute an essential modality in human communication. Robotherapy, a field in robotics, attempts to apply the principles of social robotics to better the psychological and physiological state of the ill, the secluded, or those with physical or mental handicaps. It seems that robots can play a role of both companionship and stimulation. They must, however, be designed with a maximum of communication capacities for such a purpose. One of the first experiments in this field of robotics was carried out with elderly people in a retirement home and Paro [2]. These experiments clearly showed that companion robots could give a certain amount of moral and psychological comfort to those that are most vulnerable.

In this context, we began experiments [3] using the Paro robots to check whether or not the reaction / interaction with the robots was dependent on cultural context. These experiments showed us two principal directions in which it would be interesting to work. The first one deals with mechanical problems: the robot must be very light, easy to take and to handle, easier than Paro; moreover, it must have a great deal of autonomy. The second one leads is changing man-machine interaction: the psychological comfort that the robot can provide is related to the quality of the emotional tie that the child has with it.

## II. MAPH and EmotiRob project

The MAPH project (see synoptic in figure 1) objective is to design an autonomous stuffed robot, which may bring some comfort to vulnerable children (eg, children in long hospital stay). However, a too complex and too voluminous robot is to be avoided. The EmotiRob project, which is a subproject of MAPH, aims to equip the robot with the perception and understanding capabilities of natural language so that it can react to the emotional state of the speaker. EmotiRob also includes the conception of a model for the emotional states of the robot and its evolution.

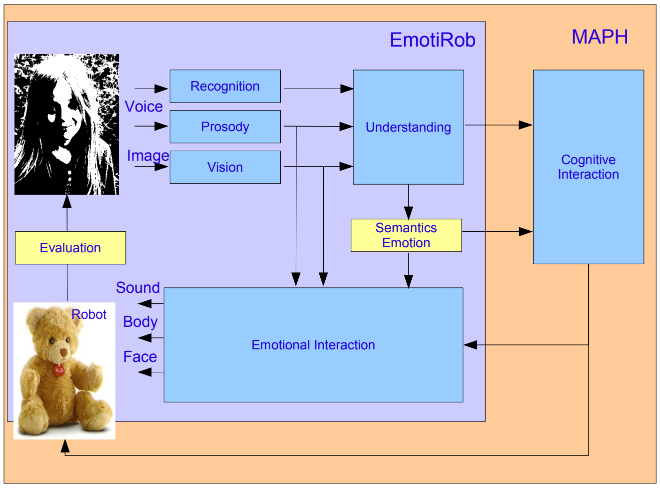



Fig. 1. Synoptic of MAPH project, including EmotiRob sub-project

The experiments we conducted on perception and emotional synthesis allowed us to determine the most appropriate way to express emotions and to obtain an acceptable recognition rate for our public. And thereby deduce the minimum degrees of freedom required to express the 6 primary emotions. The second step was the definition and description of our emotional model iGrace [4], instance of the generic model GRACE [5], to have a nonverbal emotional reaction to the utterance of the speaker. The experiments conducted allow for a hypothesis validation of the model that will be integrated on the robot: Emi. To avoid a repetition phenomenon of Emi reaction, a study on behaviour dynamics and its evolution has been undertaken. As Emi was under construction during this study, a virtual avatar, Art-e, was created to represent Emi's conscience. Art-e had to display the same facial expressions as the robot, but they did not have the exact same constraints. Art-e helped with work and tests on emotional dynamics, the good results of which will be used by the robot. Art-e has five expressive components: eyebrows and mouth - which are the same as the robot - and eyes, head and trunk. Each component has a role when an emotive experience is displayed. It is based on the six primary emotions, as well as neutrality.

## III. iGRACE – Computational Model of Emotions

iGrace computational model [6], [4], instance of the generic model GRACE [7], [5], to have a nonverbal emotional reaction to the speech of the speaker. It receives input information, processes it and determines behaviour that Emi can have for the discourse. iGrace is composed of 3 principal parts allowing for information processing:

- "Input" Module: results from the understanding module, it informs "emotional interaction" about the speech of the user.
- "Emotional interaction" Module: with information given by "Input" and the cognitive state of Emi, it generates an emotional state representing by a list of emotional experiences.
- "Expression of emotions" Module: expresses the emotional state of Emi calculated by "Emotional Interaction".

### A. Input

This module represents the interface for communication and data exchange between the understanding module and emotional interaction module. The parameters taken into account are the following:

- Video signal
- Audio signal
- 7-uplets of understanding module:
  - Actions "for the child"
  - Concepts "for the child"
  - Act of language
  - Coherence
  - Tense
  - Phase
  - Emotional state

### B. Emotional Interaction

This process objective is to generate the emotional state of Emi with information from the discourse given by "Input" and its internal cognitive state. It builds a list of emotional experiences for Emi based on a taxonomy of emotions: emotional experiences [8]. "Emotional Interaction" is composed of four main modules that produce lists $L_i$ of pairs $(eemo, C_(eemo))$ involving an emotional experience $eemo$ with an influence coefficient $C_(eemo)$ in four steps (see Fig. 2):

- Moderator: represents the cognitive internal emotional state of Emi. It builds a list $L_{mod}$ of emotional experiences functions of Emi personality and mood.
- Emotional experience selector: represents the emotional state of Emi for the discourse. It builds a list $L_{sel}$ of the emotional experience functions of the words of the discourse.
- Emotional experience generator: represents the emotional state of Emi, functions of the cognitive internal emotional state of Emi, emotional state of the user and for the discourse. First, it builds a list $L_{etat}$ of emotional experience functions of the emotional state of the child. Then it builds a list $L_{gen}$ by fusion of all lists: $L_{mod}$, $L_{sel}$, and $L_{etat}$.
- Behaviour: chooses the reaction of Emi among those proposed by *emotional experiences generator*. It extracts the best emotional experiences from the list $L_{gen}$ to $L_{comp}$.

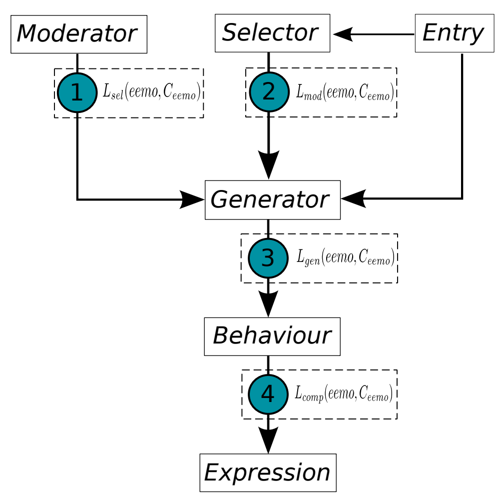
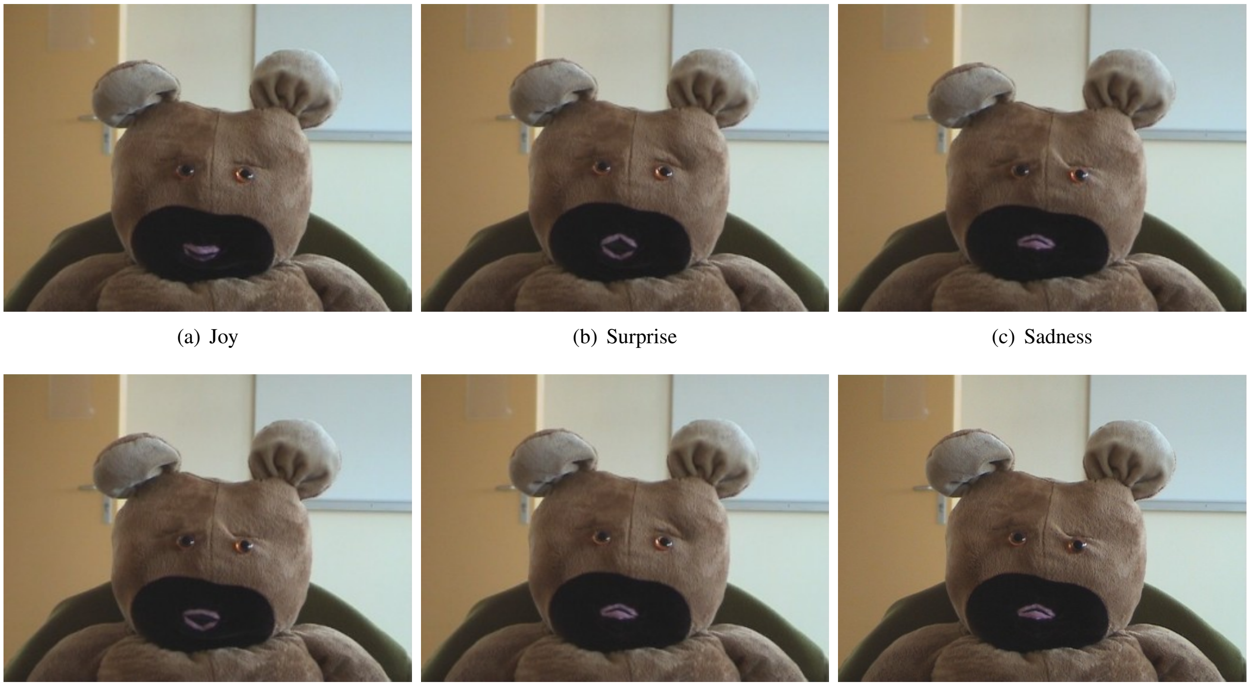



Fig. 2. Processing of event from understanding module in 4 steps

### C. Expression of emotions

This process objective is to express the emotional state of Emi by building a list of triplet <tone,posture,facial state>. Because Emi can only express the six primary emotions of P. Ekman, we convert emotional experiences from the list



(a) Joy (b) Surprise (c) Sadness

(d) Anger (e) Fear (f) Disgust

Fig. 3. Exemples of faciales expressions for the third version of Emi

$L_{comp}$ into facial expressions. A matrix $EM$ defines the relation between emotions and emotional experiences from the list $L_{comp}$. These emotions allow us to choose a facial expression in a large panel.

The expression of the emotional state implements the functions of Emi's features: buzzer and motors. This information can be converted into actions that can be execute by Emi. Facial expressions and postures are converted into motor movements and tone into music notes. Facial expressions for an emotion is expressed with 6 action units. We use a simple simplification of EMFACS system [9], [10].

## IV. Robotics conception

After an advanced research study on perception and emotional synthesis, we determined the most appropriate way to express emotion and obtain good recognition of expression with our users. The first steps of the project [11] allowed us to determine the degrees of freedom required to express the six primary emotions of P. Ekman [12] and then start the robotic designing process.

### A. Mechanical architecture

The Emi robotic platform we built is a stuffed animal with a pleasant texture that can emotionnaly react by using facial expressions and body movements. Research work on emotional synthesis provided information on the different elements that make up the face and number of degrees of freedom necessary for the facial expressions corresponding to the 6 primary emotions of P. Ekman. The face of the robot is composed of the following elements:

- 1 mouth: 4 degrees of freedom.
- 2 eyebrows: 2 degrees of freedom (1 per eyebrow).

To these elements of the face, we have added:

- 2 immobile ears.
- 2 immobile eyes.
- 1 camera at nose level to follow the face and potentially for facial recognition. The camera used is a CMUCam 3.

(a) Used of cables system with the motors (b) Emi entier

Fig. 4. Realization of the third version of Emi

The material used for the skeleton of the head is made with epoxy resin allowing for better resistance. Movement of facial elements is made through a cable system (see Fig. 4) and springs to improve the expressiveness of the robot(see Fig. 3, videos can be downloaded from the Website of the project[1]). The system used for the mouth is a spring system surrounded by a very elastic fabric.

To increase the expressiveness of the robot, we associate body movements to facial expressions. The architecture allows for the following movements:

- For the head: 2 degrees of freedom for movement right – left and up – down.
- For the waist, 2 degrees of freedom for movement right – left and up – down.

The skeleton of the torso is made of aluminium and allows the robot to turn its head from left to right, as well as up and down. It also permits the same movements at the waist. There are a total of 4 motors that create these movements.

The motors used for the head are AX-12+ and for the torso RX-24 (powerful). These motors enable numerical communication with the computational architecture created for the project. Emi currently weighs about 2.8 kg.

So that Emi may play an emotional song, we have substituted one of the motors (AX-12+) of the face by an AX–S1. This motor incorporates new features like temperature and infrared sensors, a buzzer for some music notes, etc.

### B. *Computer architecture*

Currently, communication with the robot is done through a distant computer directly hooked up to the motors. This computer integrated a process called *iGrace*, like the computational model of emotion we have developed. This process was developed with the C++ language and uses a FTDI library for communication with motors. This library makes it possible to send (and receive) instruction packets with identification of the motors, action to do and parameters for this action. The link between motors and the computer (see Fig. 5) was done with wireless USB built by Hama. The 6 motors for the head use a TTL connection and these for the torso RS-485. We need to use a USB2Dynamixel convertor for USB⇔TTL and USB⇔RS-485.

## V. Experiments

The experimentation objective is to evaluate expressivness of Emi in its last version and its interaction with children. To evaluate these points, we asked 2 boys, 8 and 9 years old (see figure 6), to meet Emi and interact with it. Then, they had to fill an evaluation grid and answer some questions.

For the interaction between the child and the robot, we did not use the iGrace model for reaction to the user's discourse. The different reactions the robot could have were remotely controlled by a technician using a graphic tablet (see figure7) we developed in Flash which allowed us to choose the emotion Emi was to express. We used the Wizard of Oz technique to make the child believe Emi was understanding what he/she said and was talking about.

[1]http://www-valoria.univ-ubs.fr/emotirob/, menu Robot –> Vido

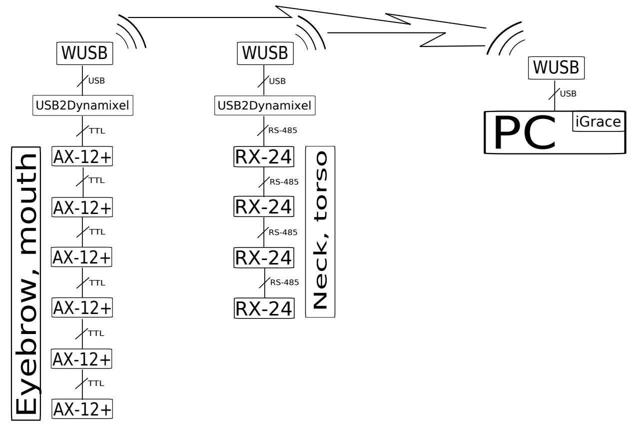



Fig. 5. Computer architecture for the third version of Emi

(a) Experiment 1 (b) Experiment 2

Fig. 6. Beginning of experimentation with 2 children

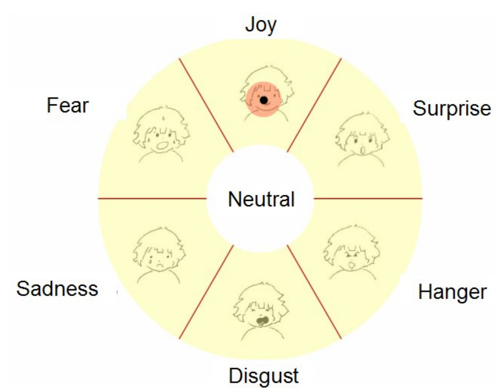



Fig. 7. Interface used to control Emi's expression

The evaluations were recorded with the agreement of all the parents. Explanations of evaluation were given the same day by an evaluator.

### A. *Protocol*

This experimentation took place in an isolated room (see Figure 8) in our university. The room was equiped with Emi on a chair, the 2 children equiped with cordless microphones, 2 supervisors and 2 cameras. The first camera was fixed on Emi and the second one on the child interacting with it. A second room were equiped with a technician who controlled Emi and could see the interaction with the children.

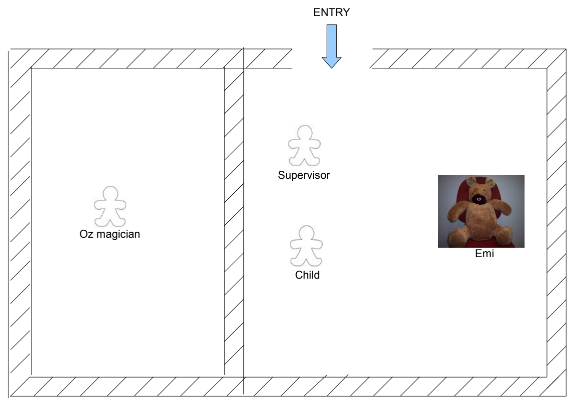



Fig. 8. Plan of the room we equipped for evaluation

Before beginning experimentation, the supervisor explained to the child what he had to do. We again specified that Emi could not speak or walk, but that it could move its body, head, mouth and eyebrows. The experimentation proceeded in 3 steps:

1) Interaction with predefined sentences: for this step, we needed the participation of the two children at the same time. Each of them, successively, had to read 7 sentences to Emi, with a total of 14 sentences. Each sentence corresponded to an emotion that Emi could express. At the end of the reading, the 2 children had to check, on the evaluation grid, the emotion they recognized. This method provided us with an evaluation for the actor and the observer.
2) Interaction in free discourse: this step is done with one child at a time. We asked the children to tell Emi a short story. We asked him to observe the robot's reaction and to answer some questions while he was telling the story. Indeed, some questions about emotions and reactions were asked while the story was being told, but others were asked after the interaction.
3) Imitation: this step is done with one child at a time. We asked the child to simulate one of the six emotions and to evaluate the quality of Emi's imitation.

### B. Evaluation grid

The evaluation grid was completed in 3 parts:

1) Emotional state of the child: given information about the emotional state of the child befor and after interaction. With this, we are able to give the impact of Emi on him. The scale used is the following:
   - Calm/neutral/agitated
   - Enthusiastic/neutral/little interested
   - Very talkative/neutral/impressed
   - Happy/neutral/not happy
2) Emi expressiveness: given expressions recognized by the child during interaction with Emi. It is composed of 14 lines of 8 cases, each line represents a sentence. The 7 primary cases have a facial expression for the 6 primary emotions (joy, sadness, anger, surprise, disgust, fear) + neutral. the 8th case, is an unknowed expression.
3) Emi physical aspect.

### C. Results

The objective of this experimentation was to have children evaluate Emi's expressiveness. The Figure 9 presents the recognition rate of emotions for the 2 children. These values are not significant but enable us to validate the new mechanical design of Emi and to begin further experimentation with more children. However, we have to choose new facial expressions for anger because it was not well recognized.

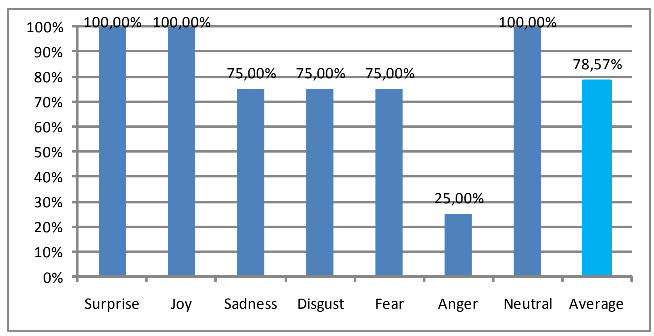



Fig. 9. Global recognition rate of emotions

Figure 10 presents the recognition rate between the actor and the observer, which show that the observer recognizes the emotions better than the actor. Indeed, the observer, who has no sentences to read, can remain more concentrated than the actor and can more easily recognize all the emotions expressed. Unlike the observer, the actor is not able to look at the beginning of Emi's expression or reaction. For future evaluations, a time lapse should be added after the utterance. The lapse must be long enough to allow the actor time to watch the entire reaction, but short enough not to be seen as a lack of expressiveness.

Figure 11 presents the satisfaction rate obtained for the imitation game, which like for expressiveness, was a good rate. Although, some discords could be observed between the children when evaluating negative emotions. These emotions are those with the lowest recognition rates is Figure 9. Creating new pattern for these facial expressions is essential in order to obtain the best possible results.

Figure 12 presents the satisfaction rate for Emi's physical aspect and interest for children. As for the previous evaluation, this rate is very high and Emi is very well accepted.

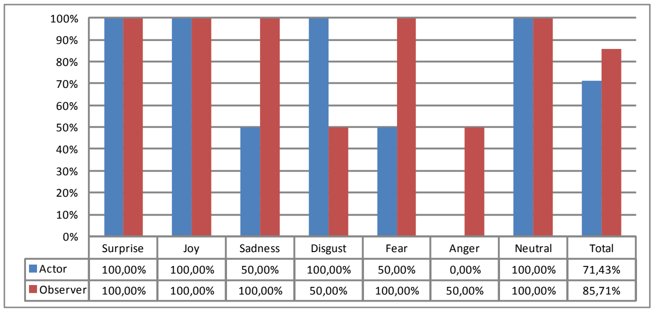



Fig. 10. Recognition rate comparison between actor and observer

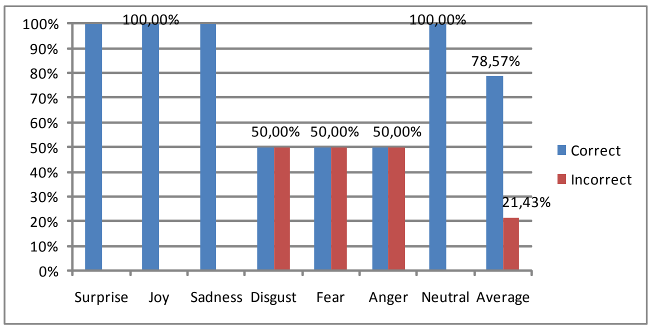



Fig. 11. Satisfaction rate for imitation

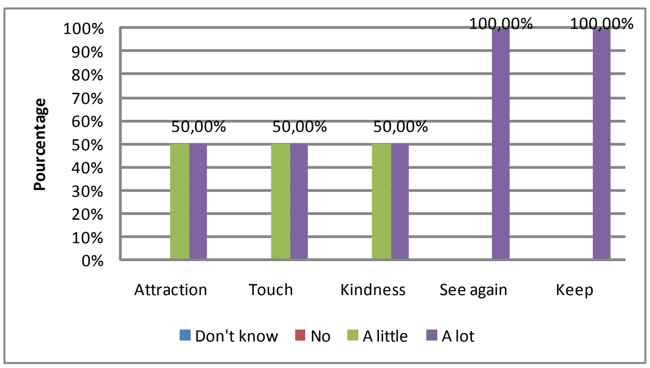



Fig. 12. Satisfaction rate for Emi acceptance

With regard to this evaluation, we can say that a new design choice can increase Emi's expressiveness. Even with a lower recognition rate for the anger emotion, globally, results are good. To validate and confirm other design choices, a new evaluation with more children has to be done.

## VI. Conclusion and perspectives

The experiment we are doing will allow us to know if the expressions of Emi are good. As we have conducted interaction experiments with only 2 children, we cannot say if these results are significant. Further evaluation with children will allow us to validate or not the robotics design, as well as the expressions of the robot. However, from these results, we can note that some modfications and adjustments must be done to improve expression and tactile interaction.

However, since reaction satisfaction obtained good results with the Wizard of Oz technique, we will start new experiments with more children for interaction evaluation.

## VII. Acknowledgments

EmotiRob is currently financed by the regional council of Martinique, for the development of the emotional synthesis, the regional council of Bretagne for language comprehension, and the ANR for the construction of the robot.

Most of all, we would like to thank the regional council of Martinique, as well as the ANR for their collaboration and the financing of future work.

The authors would also like to thank all participants for the time spent on this experiment and all constructive comments.

## References

[1] E. A. Boyle, A. H. Anderson, and A. Newlands, "The effects of visibility on dialogue and performance in a cooperative problem solving task," *Language and Speech*, vol. 37, no. 1, pp. 1–20, 1994.

[2] T. Shibata, "An overview of human interactive robots for psychological enrichment," *IEEE*, vol. 92, no. 11, pp. 1749–1758, 2004.

[3] B. Le-Pévédic, T. Shibata, and D. Duhaut, "Study of the psychological interaction between a robot and disabled children." 2006.

[4] S. Saint-Aimé, B. Le Pévédic, and D. Duhaut, *iGrace – Emotional Computational Model for EmI Companion Robot*. InTech Education and Publishing, 2009, ch. 4, pp. 51–76.

[5] T.-H.-H. Dang and S. Letellier-Zarshenas and D. Duhaut, "GRACE – GENERIC ROBOTIC ARCHITECTURE TO CREATE EMOTIONS," *Advances in Mobile Robotics: Proceedings of the Eleventh International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines*, pp. 174–181, September 2008.

[6] S. Saint-Aimé, "Conception et réalisation d'un robot compagnon expressif basé sur un modèle calculatoire d'émotions," Ph.D. dissertation, Valoria – Université de Bretagne Sud, Vannes, 9 juillet 2010.

[7] T.-H.-H. Dang, S. Letellier-Zarshenas, and D. Duhaut, "Comparison of recent architectures of emotions," in *Control, Automation, Robotics and Vision, 2008. ICARCV 2008. 10th International Conference on*, Decembre 2008, pp. 1976–1981.

[8] M. Larivey, *La puissance des émotions : Comment distinguer les vraies des fausses.*, de l'homme ed. Québec: Les éditions de l'Homme, 2002.

[9] P. Ekman and W. Friesen, "EMFACS facial coding manual," *Human Interaction Laboratory, San Francisco*, 1983.

[10] W. Friesen and P. Ekman, "Emfacs-7: emotional facial action coding system," *Unpublished manuscript, University of California at San Francisco*, 1983.

[11] S. Saint-Aimé, B. Le-Pévédic, and D. Duhaut, "Building emotions with 6 degrees of freedom," in *Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on*, Oct. 2007, pp. 942–947.

[12] P. Ekman, "Universal and cultural differences in facial expression of emotion," in *Nebraska Symposium on Motivation*, vol. 19, Nebraska University Press, 1972, pp. 207–283.