# The EEE corpus: socio-affective "glue" cues in elderly-robot interactions in a Smart Home with the EmOz platform

**Véronique Aubergé[1], Yuko Sasa [1], Nicolas Bonnefond [1], Brigitte Meillon [1], Tim Robert [1], Jonathan Rey-Gorrez[1], Adrien Schwartz[1,2], Leandra Antunes[1,3], Gilles De Biasi [1], Sybille Caffiau [1], Florian Nebout[4]**

1 LIG-lab CNRS, Grenoble University, France

2 Floralis Company, Gières, France

3 University of Ouro Preto, Brazil

4 Awabot Company, Lyon, France

E-mail: Veronique.Auberge@imag.fr

**Abstract**

The aim of this preliminary study of feasibility is to give a glance at interactions in a Smart Home prototype between the elderly and a companion robot that is having some socio-affective language primitives as the only vector of communication. The paper particularly focuses on the methodology and the scenario made to collect a spontaneous corpus of human-robot interactions. Through a Wizard of Oz platform (EmOz), which was specifically developed for this issue, a robot is introduced as an intermediary between the technological environment and some elderly who have to give vocal commands to the robot to control the Smart Home. The robot vocal productions increases progressively by adding prosodic levels: (1) no speech, (2) pure prosodic mouth noises supposed to be the "glue's" tools, (3) lexicons with supposed "glue" prosody and (4) subject's commands imitations with supposed "glue" prosody. The elderly subjects' speech behaviours confirm the hypothesis that the socio-affective "glue" effect increase towards the prosodic levels, especially for socio-isolated people. The actual corpus is still on recording process and is motivated to collect data from socio-isolated elderly in real need.

**Keywords:** socio-affective "glue", human-robot interaction, socio-affective prosody, elderly, Smart Home, spontaneous corpus

## 1. Introduction

It is supposed here that whatever the social role created or borrowed for a robot entering the social sphere of the human, its role competences can be integrated in the human social space only if the relational link allows the dialog architecture by building the relevant "socio-affective glue" in a co-construction processing. The hypothesis underlying this work is that the material of the "socio-affective glue" is sufficiently non-lexical sounds and mimicry with "glue" prosody (Aubergé, 2012). Non lexical sounds - non phonological but prosodically relevant items - produced during or outside the talk turn, like onomatopoeias, interjections, fillers, grunts, bursts have been studied for their emotional functions [affects bursts (Scherer, 1994; Schröder 2003)], or for their pragmatic functions in dialog (Fonagy & Target, 1997). Non-lexical sounds have been observed both in listener feedbacks in backchannel (see Humaine D6d works) and in the feedbacks of the speaker, implied in a human/human or human/machine interaction (Morlec, 2001; Mairesse & al., 2007; Morency, 2010). They can express emotions, intentions, attitudes and cognitive/mental states and processing (like concentration, hesitation about an answer, etc.) that we name *Feeling of Thinking – FoT* (Aubergé, 2012). From a large spontaneous corpus (Aubergé, Rilliard, Audibert, 2005) some such functional lexical words have been selected (Vanpé & Aubergé, 2011) and perceively measured (De Biasi, 2012; Sasa & al. 2013). These non lexical sounds have been perceively classified in increasing "glue" competence: in order to long term develop an application of a "socio-affective glue trainer" for a robot with relationally isolated human, an experiment is presented is that shows that these selected sounds and mimicry, given to a butler robot in smart home, build a strong socio-affective glue with some isolated elderly.

## 2. Elderly situation

Gerontechnology emerged from a society challenge due to the demographic evolution of elderly (Bouma & al., 2007). The number of aged people living at home becomes higher every year in Western countries: over 80 years, 6/10 people live at home (4/10 in nursing homes), 25% of them with low dependency while only 2,5% are strongly dependent (Harrington & Harrington, 2000). The Smart Homes are often presented as convenient (and economical) issues to help elderly to stay longer at home. In a such socioeconomic situation, one main vector of elderly frailty is now the socio-affective isolation: it was observed in many studies [see the last ISG http://www.gerontechnology.info conferences] that the affective and organizational dimensions of isolation have direct and very strong consequences on the physical and mental health [6,30], which allows to keep elderly living at home. The main cue pointed by all those studies [see IAAG http://www.iagg.info/index.php and IUGMS http://www.eugms2014.org Congresses] is the isolated ones' socio-affective interactional competences degradation. It means that socio-affective interactional "coaching" would be a main issue, that starts to be taken into account by some professionals of elderly caregiving [www.bienalamaison.com]. The socio-affective interactional degradation occurring for elderly can appear in other societal areas, like the hikikomori syndrome described in Japan for young people (Furlong, 2008). Of course, it becomes a central issue for the pathologies including communication diseases, like Alzheimer or autistic syndrome.

That is why the present study prior goal is to collect a large spontaneous corpus of ecological situations implying elderly and a companion robot in order to further design technologies of human-robot interactions specifically devoted to elderly living in a Smart Home. This socio-interactional robotic technology (Interabot Project[1]) will be built to train the elderly to communicate (socio-affective prosthesis) with a robot while this tool is presented as the Smart Home's butler.

Some theoretical objectives motivate this study too: using a robot is here a method to evaluate some hypotheses on the interactions primitives that build what we call the socio-affective "glue". Prosody carries emotional, socio-affective and interactional information where each language has its own values (Decety, 2007). This communicative information appears in different prosodic levels as in non-lexical sounds. Those can be non-phonetic sounds like grunts, affect bursts or mouth noises (Schröder, 2006, Poggi, 2008), phonological as fillers, mind markers or interjections (Amecka, 1992), or onomatopoeia, widely studied in Japanese (Shibatani, 1990). These sounds that we can consider as pure prosodic tools, were studied for a specific and supposed emotional (Aubergé, 2012) and pragmatic (Fonagy & Target, 1997) functions, as well as moods, emotions, intentions, attitudes, cognitive processes and mental states also known as "Feeling of Thinking" (Aubergé, 2012). Moreover lexicons, sentences and paraphrases prosodic form also support various socio-affective values (Decety, 2007). These cues can be extended from simple sounds to sentences produced in a same context, which have been tested in synthesis (Morlec, 2001; Mairesse & al., 2007; Morency, 2010). Lately, the prosody carrying this communicative information was introduced as a way to develop "socio-affective glue" (Aubergé & al., 2013) that allows interlocutors to build dynamically the communicative channel depending on the interaction context. Furthermore, imitation has been studied as a basic process to create the same kind of "glue" in children language acquisition (Tomasello & al., 2005) or as a primitive of robots learning (Schaal, 1999).

By the way, since the 90's, Affecting Computing and multidisciplinary communities have been focusing their work on the face-to-face interactions, especially on facial, gestural and vocal expressions using virtual agents and robots as in various studies in social computing (Schaal, 1999; Breazeal & al., 2002). It is interesting to see that when a robot is not explicitly humanoid, human creates by himself a socio-affective relationship with this device toward its « pet » stance (Sasa & al. 2012). Because all these different prosodic levels have not been studied together particularly to see their functions in the "socio-affective glue" building, our work will test them progressively thanks to a robot interacting with elderly towards gradual vocal productions: (1) no speech, (2) pure prosodic mouth noises supposed to be the "glue's" tools, (3) lexicons with supposed "glue" prosody and (4) subject's commands imitations with supposed "glue" prosody. This will be tested with the EmOz wizard of Oz platform developed for this project (Aubergé & al. 2013) in the experimental Living Lab Domus of the LIG lab. Domus is a completely authentic Smart Home with hidden equipment and control room (Niitamo & al. 2006). A complex script is held to collect comparable data for many senior subjects (more than 40 are under recording), in the increasing levels of "glue", for the EMOX robot (developed by Awabot www.awabot.com/) playing the role of Domus' butler. The resulting corpus is EEE (Elderly EmOz Expressions).

[1] The Interabot project is financed by the French Industry ministry (Investissements d'Avenir) and held together by some industrial companies and academic partners,

## 3. The EEE script with EmOz

### 3.1 Experimental tools: EmOz – a Wizard of Oz

#### 3.1.1 Domus: a Smart Home prototype

The LIG developed a living-lab into the Multicom Platform where we can record high quality sounds in a specific room (*see A on figure 1*), film a recording set which looks like a meeting room but can also be arranged to look like every other kind of environments for experiments (*see B on figure 1*), and which all the devices can be controlled from a control room (*see C on figure 1*).

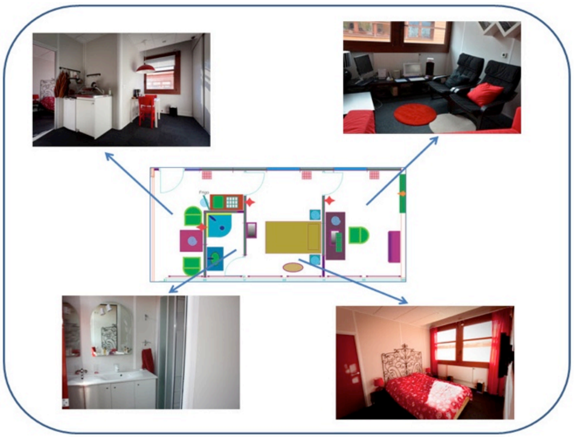

Figure 1: Multicom Platform of the LIG-lab.

In this platform, Domus (*see D on figure 1*), where our study takes place, is designed like a 40m$^2$ flat with a kitchen, a bedroom and a living room equipped with two cameras and two microphones in each room, and a shower room with a microphone (*see figure 2*). It has sensors and actuators conforming to the automation standards KNX (Konnex) that group a heterogeneous set of protocols exchanging information outside the building through an OSGI gateway.

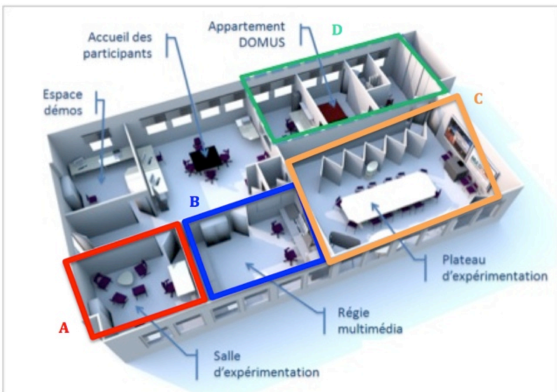


Figure 2: Illustrations of Domus Smart Home.

In our work, we selected few possible actions to execute into DOMUS and we proposed 30 vocal commands (*see Table 1*) that we can simulate from the platform control room.

| | |
|---|---|
| **Kitchen** | Monter/descendre/arrêter les stores *To up/down/stop blinds*<br>Allumer/éteindre la lumière *To turn on/off the light*<br>Mettre/éteindre la bouilloire *To turn on/off the kettle* |
| **Bedroom** | Monter/descendre/arrêter les stores *To up/down/stop blinds*<br>Ouvrir/fermer les rideaux *To open/close curtains*<br>Allumer/éteindre la lumière *To turn on/off the light*<br>Allumer/éteindre les lampes *To turn on/off lamps*<br>Mettre la lumière verte/bleue /jaune<br>*To turn on the green/blue/yellow light*<br>Allumer/éteindre cette lumière *To turn on/off this light*<br>Allumer/éteindre la télé *To turn on/off*<br>Moins/plus fort la télé *Lower/louder TV* |
| **Living Room** | Monter/descendre/arrêter les stores *To up/down/stop blinds*<br>Moins/plus fort la radio<br>To lower/louder the radio |

Table 1: The vocal commands available in each room.

### 3.1.2 Emox: a non-anthropomorphic robot

For this study we chose a non-anthropomorphic robot, Emox (*see Figure 3*), develop by Awabot Company. It used an Urbi system. Ethically, the fact that this robot neither look like a human nor an animal avoids the induction of the way people picture the device and would create artefacts that cannot be controlled or be misinterpreted. Its voice is also non-human, a choice that was motivated from a previous study (Sasa, Aubergé, Franck, Guillaume, Moujtahid, 2012), which tested different types of aesthetics for the robot voice by only changing the Fundamental Frenquency (F0). At the same time, we asked people which voice they prefer and checked if the information carried by some "mouth noises" (non phonetic nor phonologic sounds; e.g. laughs and various vocalizations or breaths) were recognized. Finally, the robot has a voice pitch increased by 1.52 from the original female speaker's F0, using Voxal software[2] for voice conversion. That gives robot a "cartoon-like voice", reducing the anthropomorphism to the minimal information carried by the speech.

Figure 3: The Emox robot – Awabot company.

### 3.1.3 EmOz : an interface for non-programmer to control Emox and Domus

In this study, we created a Wizard of Oz interface to control both Domus and Emox, using java-programming language. In order to facilitate the use by non-programmer researchers, this interface generates buttons based on excel files in which you fill in simple parameters as sounds file name, basic moves characteristics and Domus automation actions (*see figure 4 for instance*). One excel file corresponds to one button on the interface. Each time you create a button on the interface, it is possible to drag and drop it wherever you want, and the last positions of all the buttons are saved which allows displaying previous versions of the interface.

| | | | | |
|---|---|---|---|---|
| **Auteur** | | | | |
| Auteur du script | Yuko | | | |
| **Description** | | | | |
| Description du script | | | | |
| | | | | |
| **Paramètre** | **Valeur** | | | |
| AFFICHER | 1 | | | |
| ZONE | 2 | | | |
| | P3.4 MonterStores | | | |
| **Groupe** | **Action** | **P1** | **P2** | **P3** |
| 1 | moveFront | 25 | 15 | |
| 2 | domus | cuisine | store | up |
| 2 | speaker | 2-ok2 | | |
| 2 | moveBack | 25 | 15 | |

Figure 4: Example of an excel script to create a button named "P2.4 MonterStores" to move Emox forward and backward, up the blinds in the kitchen while he is playing the "2-ok2" sound.

That is how it is possible to create buttons in A and D zone of Figure 5 which shows the final aspect of the interface. In A, we placed our Emox stimuli in a specific order to graduate different levels of prosody while we follow with accuracy the script that carries our hypothesis. In D, there are some complex stimuli associated to some moves or moves and sounds. The B zone generates automatically all the audio stimuli that we use while we have to do some improvisation, depending on the subjects reactions. The tool in C allows us to record our voice in live, increased the F0 to have the same voice aesthetics as the other sounds and play it on Emox if needed, because some reactions are unpredictable during the experiment. Finally, we can control all Domus automation in the E zone.

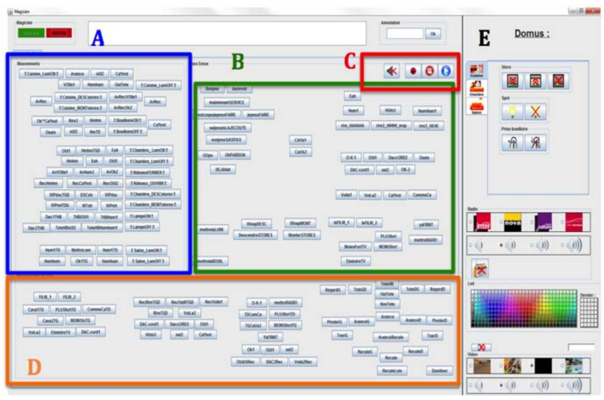


Figure 5: EmOz interface illustration.

In the left and up corner of the interface, we have got start and stop buttons generating (1) a form where you can fill in the subject's characteristics to anonymize the data; and (2) create a .csv format file with a timestamp, saving all the tracks of actions you did on the interface during the experiment.

[2] www.nchsoftware.com/voicechanger

### 3.2 The Elderly EmOz Expressions script

#### 3.2.1 Communication lack appearing with aging

Aging process depends on a physical, neuropsychological, social and environmental factors (Markle-Reid & Browne, 2003) and differs among individuals. In fact, some studies showed that our age is far different from biological and cognitive age (Anstey & Smith, 1999). Generally we talk about « elderly » over 75 years old, but their frailty or non-frailty is not equivalent and not related to their age, which is difficult to focus on the subjects who are interesting to observe. This kind of persons who likely become socio-isolated by loosing progressively their role in society, communicate less frequently to finally find more and more difficulties to engage in efficient interactions with their kinfolks or other persons they are in contact with. Each times their interactions fail, the elderly loose confidence which strengthening their lack of communication abilities (Segrin, 1994), while diseases and physical problems, directly related to communication failures appears. Thereby this loneliness and the loss of social relationships are strongly related to mortality (Holt-Lunstad, Smith & Layton, 2010; Luo, Hawkley, Waite & Cacioppo, 2012). To find really needed elderly, we asked a partnership with a national home caregivers company, Bien à la Maison, to help us find people who are still living in their own but start being frail and isolated. This choice is based on the opinion of the caregivers (mostly women) who visit regularly the elderly and who accepted to be the experimenters' accomplices in our study. The company measures the frailty with their own tools and that allowed us to base on person who are scaled GIR 5 or 6, a French standard to illustrate elderly frailty and dependency (Coutton, 2001). Once the caregiver or the organism manager find a subject corresponding to our criteria, an experimenter visits the elderly for a first interview.

#### 3.2.2 Pretext task to bring the subjects into Domus

During the first interview, the goal is to know better the subjects and to motivate them to come in Domus, our living-lab. The experimenter who is doing the recruitment introduce himself as a gerontechnology student who wants to know how people over 75 years old live and what kind of opinion they have got on technologies over a questionnaire. This gives an overall knowledge on the subjects' profile. As transition on technologies, the student says that some works have been done on a Smart Home prototype to study how we can allow seniors (who start having some physical but not too serious problems), to live as long as possible in their own house. He continues telling that in order to ease elderly's life, some researcher created technologies associated to the Smart Home but which cannot be tooled up yet at their home. So we expect the elderly to test these technologies in our flat prototype. However, in "previous studies" we observed some difficulties: when elderly change their living environment (e.g. move into a retirement/nursing home or a hospital) they mostly have trouble to accustom to this new place. Moreover, when there are technologies in this environment, people get completely confused. One of our "so-called hypotheses" to avoid this phenomena consist to ask people to bring some personal items (e.g. books, trinkets, decorative objects…etc.) and to arrange the new environment they have to handle with these items, so they can get used to the place more easily. This justification follows the idea of transitional objects sometimes used to help Alzheimer patients to be less lost (Habernas & Paha, 2002; Loboprabhu, Molinari, & Lomax, 2007). Finally, if the elderly do agree to come to the Smart Home, the student asks them to bring around ten items they care about and place their objects in Domus while evaluating it and its technologies. To help them choosing their objects, the experimenter gives a sheet where the elderly have to fill in the items they want to bring. As subjects, they will spend about two or three hours in the Smart Home. If they accept to be accompanied, we also ask them to come with their caregiver, which can ensure security. At last, the day they come to the living-lab, the student proposes to give a lift to the elderly and their caregiver, creating a situation that will facilitate the experimental scenario. He also tells that he has not visited yet the Smart Home as it was his adviser who took care of reserving Domus, so he will discover it at the same time as the subject.

#### 3.2.3 Scenario to introduce the Emox robot and Elderly

On the experiment day, the Smart Home engineer welcomes the student, the elderly and the caregiver in a reception room. They spend some time discussing about the study context to let the elderly calm down and feel comfortable. As the engineer pretends not knowing the student and the real purpose of his work, the student explains his "hypotheses" based on the personal items that allow elderly getting used to unknown and technologized environment more easily. Once the subject is ready and he is convinced of the pretext task, the engineer introduces the Smart Home and its different rooms. Very quickly, once everyone is in Domus, a third experimenter, waiting in the control room, calls the elderly's caregiver on her mobile phone, pretending he is the home help services company manager giving a mission to his employee that cannot be refused, as it is an emergency. At this time, the caregiver is aware of every details of the experiment because she passed a private interview with the experimenters before the experiment day in which she was told how to react precisely in each step of the experiment as accomplice. So she pretends having got an urgent mission very near the Smart Home that takes less than an hour and that she has to leave a moment. As she came by the student car, she asks for a ride because her mission is very important. The student understanding the emergency proposes to accompany her. He then asks to the engineer if it is possible for him to take care of the elderly subject for a while. The engineer says that he cannot stay all the time because he has got other work to do but that he will stay as long as necessary to explain how to use the Smart Home because its features are special. The student then demands the subject to start placing the personal items wherever the elderly wants to in the Smart Home, especially if both him and the caregiver have not returned yet after the engineer presentation. In addition, neither the student nor the caregiver know how the Smart Home works, so they ask to listen carefully to the engineer's explanation so the subject can describe all the operation while they will be back. Then, both student and caregiver leave Domus to go in the control room. Once they are gone, the engineer tells the subject that the Smart Home has not got any switches but it can be handle by vocal commands. At this moment he calls for "Emox", a robot he introduces as the butler of Domus and which will listen then execute the vocal commands. However, the engineer explains that at first, the robot has to learn the elderly voice for the effectiveness of the system (which is not true because both robot and Smart Home are controlled with a Wizard of Oz). The engineer then proposes a list of 30 possible vocal commands to the subject so he trains the robot to recognize his voice. The subject is asked to test at least once all the commands. When the elderly understood and starts giving the first commands, the engineer says that he has to go and he leaves the elderly to get into the control room, saying he will come back later to see if everything is fine.

#### 3.2.4 Scenario for Emox and Elderly interactions

In the control room, there are two or three Wizard of Oz experimenters who: (1) drive Emox with a joystick to follow the elderly while he is moving around Domus, (2) activate Domus automation while the subject is giving a vocal command, (3) play the vocal stimulus on Emox that carries our hypotheses on the "socio-affective glue". As the subject starts giving the first vocal commands, the Wizards are just executing Domus automation without speech from the robot. Then after three of four commands, we play some "mouth noises" that illustrate pure prosody, without any lexical information (Scherer, 1994; Campbell, 2004; Schröder & al., 2006) that we supposed to be the tools and selected from a database of noise collected (Aubergé, Rilliard, Audibert, 2005), described (Aubergé, Loyau 2006; Vanpé, Aubergé, 2011) and measured (Signorello, Aubergé, Vanpé, Granjon, Audibert, 2010; De Biasi, Aubergé, Granjon, Vanpé 2012; Sasa, Aubergé, Rilliard, 2013) from previous studies. We think these noises able to engage people in the glue process to converge with Emox. Then after some of these noises, we let Emox interact with lexicons as interjections (Ameka, 1992; Poggi, 2008), carrying also glue prosody. Finally, we introduce commands imitations, always with supposed glue prosody, to reinforce the eventual established relationship as described in the chameleon effect (Schaal, 1999; Decety, 2007). The Table 2 shows the 30 stimuli used and supposed to create and reinforce the "socio-affective glue" between the elderly and the robot.

These stimuli follow an accurate order in response to each Domus command, described in a script. Nonetheless, we sometime skip some sounds for more graduated form, whether because the elderly do not follow exactly the order of the commands list, or because very naively, as human, the wizards tend to react differently when they see some specific reactions from the subjects. If these modifications are objectively observed while analyzing the corpus, this could be a model that the robot has to follow and fit to make the "glue" with elderly as human do.

| | |
|---|---|
| **Mouth noises** | 3 types of laughs |
| | euh - hum1 - hum2 - humhum |
| | 2 onomatopoeia « woup »<br>(associated to the blinds up/down movements) |
| **Interjections** | Sounds to play before Domus actions - prosody1:<br>d'accord1 - ok1 - oui1 / *all right1, okay1, yes1* |
| | Sounds to play before Domus actions - prosody2:<br>d'accord2 - ok2 - oui2 / *all right2, okay2, yes2* |
| | Sounds to play after Domus actions :<br>ça va1 - ça va2 - comme ça - ça y est / *fine1-fine2-like this-that it* |
| | If Emox is mistaken : oups / *oops* |
| **Commands Imitations** | Infinitive form of the vocal commands + interjections used before |
| **Other sentences** | Some predictable minimal dialogue sentences used only if needed.<br>E.g. : je peux faire quelque chose / *can I do something*, mais c'est déjà fait / *but it is already done*, oh pardon / *oh sorry* |

Table 2: Emox robot audio stimuli.

## 4. The EEE Corpus

### 4.1 Overall description of the corpus

The subjects are, for now, from 68 to 92 years old (*see Figure 6*). It is still on recording processing, as we will try to collect around 40 subjects interactions. This corpus is composed by ten experiments lasting from an hour and a half to two hours each. For each subject, we have six videos (two per rooms) and an audio file collected by the subjects' lapel microphone. We have nearly 456 interactions between EMOX and the elderly (from 43 to 52 per subject), throughout the full experiment. Each interaction lasts about 10 to 50 seconds, showing a sequence of exchanges around one voice command. For the analysis we divided the results in three steps while the subjects: (1) are learning the commands with the engineer, (2) are alone with Emox, (3) are explaining how DOMUS and Emox work to the helper and then to the recruiter. We quote the commands forms used by the subjects, count and store them in the chronological order of appearance. Those commands are associated with punctual or gradual reactions of both the robot and the subjects, which illustrate the "socio-affective glue" degree between the robot and the elderly.

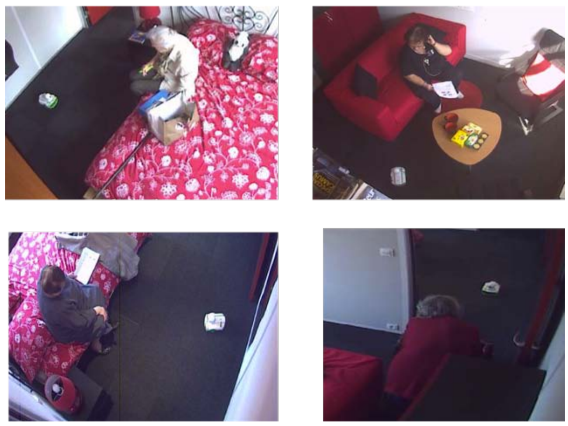

Figure 6: Some elderly subject interacting with Emox.

### 4.2 First analyses of EEE

There are of course some variations concerning the subject's behaviors during the experiment first steps, but some common main characteristics emerged as features of the "glue" building increasing steps: (1) declarative commands without paraphrasing; (2) the same original form commands but with a positive attitude prosody (in particular fundamental frequency arise which systematically appear at the end of the sentences, with a breathy voice); (3) commands paraphrased variations (used in synergy with a "we") with a globally high fundamental frequency and a great arise at the end of the sentences; and finally (4) multiple prosodic focuses of support terms with a higher fundamental frequency. These phenomena are observed as well as a voice quality becoming more and more breathy. This elderly's voice quality breathiness seems to vary particularly while the robot produced a feedback based on pure prosodic vocal micro-sounds.

The elderly's speech behaviors confirm that the effect of the socio-affective "glue" increases towards the prosodic levels, especially for socio-isolated people. Moreover, to allow a precise control of the robot reactions timing and order, we need an efficient interface so the cognitive effort of the Wizard of Oz experimenter is the same as the effort the robot "seems to produce" to execute the commands. Consequently an HRI technology will be specifically developed for the EEE situation, thinking of useful features add to the Smart Home. These technologies will then need important ethical considerations, leading to a functional system with a theoretical background focused on the practice of socio-affective interactions competences for socio-isolated people.

## 5. Conclusion

This work had both a theoretical and a technological goal: (1) to show that a strong "socio-affective glue" is built by carefully selected non lexical sounds and selected prosody on mimicry and that this glue is the base of any relation and ensures the relevance and the acceptability of the social role (here to control the smart home) (2) at least for isolated person, like elderly, whatever the role allowed to the robot, the really crucial expected role is to build a glue: the robot can train the human to relational performances and consequently help the isolated person and more efficient in the human-human communication. This has been validated both by the collected subjects expressions, the subjects' request and the professional of elderly car who assisted this experiment. The EEE corpus will be completed to a large panel of subjects, in order to build by machine learning, within hybrid system (rules on non lexical sounds hypotheses enriched and adapted by stochastic data learning). This will carry on a minimal dialog system for elderly in smart home that will be completed and augmented in active learning by telecare.by professional of care. It must be noted that the choice of a non-humanoid and non-animal like robot (to avoid the uncanny valley effect) is largely validated both by elderly and professional of care.

## 6. Acknowledgements

This work has been partially supported by the LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01) granted to Yuko Sasa and the Interabot project (Investissements d'Avenir, DGCIS) collaboration with Awabot (robotics) company. We thank Bien à la Maison company (elderly personal services) for their active participation to collect the corpus. Thanks also to CAPES Foundation, Brazil, for the postdoctoral scholarship granted to Leandra Antunes – BEX 18020-12-7. This work has been partly supported by the Major Program for the National Social Science Found of China (13&ZD189).

## 7. References

Ameka, F. (1992). "Interjections: The universal yet neglected part of speech", Journal of Pragmatics, 18, 101-118, 1992.

Anstey, K. J., & Smith, G. A. (1999). Interrelationships among biological markers of aging, health, activity, acculturation, and cognitive performance in late adulthood. Psychology and aging, 14(4), 605.

Aubergé, V., Rilliard, A., & Audibert, N. (2005). De E-Wiz à E-Clone: méthodologie expérimentale pour la modélisation des émotions et affects authentiques. Proceedings of WACA Grenoble France.

Aubergé, V., Loyau, F. (2006) Expressions d'un agent humain entre ses tours de parole.

Aubergé V., Sasa Y., Robert T., Bonnefond N., Meillon B. (2013) "Emoz: a wizard of Oz for emerging the socio-affective glue with a non humanoid companion robot". In proceedings of WASSS 2013, Grenoble, France.

Aubergé, V. (2012) "Attitude vs. Emotion: A Question of Voluntary vs. Involuntary Control." In GSCP. Belo Horizonte, Brazil

Bayles K. A. and Kaszniak A. (1987) "Communication and Cognition: Normal Aging and Dementia", Little, Brown, Boston.

Bouma, H., Fozard, J.L., Bouwhuis, D.G., Taipale V.T. (2007) Gerontechnology in perspective. In Gerontechnology, Vol. 16, No 4, pp.190-216

Breazeal, C. and Aryananda, L. (2002) "Recognition of affective communicative intent in Robot-Directed speech", Autonomous Robots 12, pp 83-104.

Cacioppo, J. T., & Patrick, B. (2008). Loneliness: human nature and the need for social connection. New York: W. W. Norton & Company.

Campbell, N (2004). "Getting to the Heart of the Matter: Speech as the Expression of Affect; Rather than Just Text or Language", Languages Resources and Evaluation, 39, 109-118.

Chaby, L.; Chetouani, M.; Plaza, M.; Cohen, D. (2012) "Exploring multimodal social-emotional behaviors in autism spectrum disorders". In Proceedings of the 2012 ASE/IEEE International Conference on Social Computing and 2012 ASE/IEEE International Conference on Privacy, Security, Risk and Trust; IEEE Computer Society: Washington, DC, USA, 2012, pp. 950–954.

Coutton, V. (2001). Évaluer la dépendance à l'aide de groupes iso-ressources (GIR): une tentative en France avec la grille aggir. Gérontologie et société, (4), 111-129.

Darling K. (2012). “Extending Legal Rights to Social Robots”. We Robot Conference, University of Miami, April 2012.

De De Biasi, G, Aubergé V, and Granjon L (2012). “Perception of Social Affects from Non Lexical Sounds.” In GSCP. Belo Horizonte, Brazil.

Decety, J. (2007). “A social cognitive neuroscience model of human empathy”. In E. Harmon-Jones & P. Winkielman (Eds.), Social Neuroscience: Integrating Biological and Psychological Explanations of Social Behavior (pp. 246-270). New York: Guilford Publications.

Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., and Cohen, D. (2012). “Interpersonal Synchrony: A Survey of Evaluation Methods across Disciplines”. IEEE Transactions on Affective Computing 3(3), pp. 349-365.

Fonagy, P., & Target, M. (1997). “Attachment and reflective function: Their role in self-organization”. Development and Psychopathology, 9, 679-700.

Furlong, A. (2008). The Japanese hikikomori phenomenon: acute social withdrawal among young people. The Sociological Review, 56(2), pp. 309-325.

Greenberg Y., Tsuzaki M., Kato H. and Sagisaka Y. (2006) “A trial of communicative prosody generation based on control characteristic of one word utterance observed in real conversational speech”. In Proceedings of Speech Prosody 2006, pp. 37-40.

Habermas, T., & Paha, C. (2002). Souvenirs and Other Personal Objects: Reminding of Past Events and Significant Others in. Critical advances in reminiscence work: From theory to application, 123.

Harrington, T.L., Harrington, M.K. (2000) Gerontechnology : Why and How ? Eindoven : Shaker Publishing.

Ladd D.R., & Cuttler A. (1983). Models and measurements in the study of prosody. In A. Cutler & D. R. Ladd (Eds.), Prosody: Models and measurements (pp. 1-10). Heidelberg: Springer-Verlag.

Loboprabhu, S., Molinari, V., & Lomax, J. (2007). The transitional object in dementia: Clinical implications. International Journal of Applied Psychoanalytic Studies, 4(2), 144-169.

Loyau, F., Aubergé, V. (2006). “Expressions outside the talk turn: ethograms of the Feeling of Thinking”, 5th LREC, pp.47-50.

Mac D, Castelli E, Aubergé V. (2012). “Modeling the prosody of Vietnamese attitudes for expressive speech synthesis”. Workshop of Spoken Languages Technologies for Under-resourced Languages (SLTU 2012), Cape Town, South Africa.

Mairesse, F., A. Walker, M., R. Mehl Matthias and K. Moore, R. (2007) “Using linguistic cues for the automatic recognition of personality in conversation and text”. In Journal of Artificial Intelligence Research 30, pp. 457-500.

Markle-Reid, M., Browne, G. (2003) Conceptualizations of frailty in relation to older adults. In Journal of Advanced Nursing, Vol. 44, 58-68.

Morency, L.P (2010) “Modeling human communication dynamics”, Social Sciences. Signal Processing Magazine, IEEE, 27(5), pp. 112 -116.

Morlec Y., Bailly G. and Aubergé V. (2001). “Generating prosodic attitudes in French: data, model and evaluation”. Speech Communication: 357-371.

Niitamo, V.-P.; Kulkki, S.; Eriksson, M.; Hribernik, K. A. (2006) “State-of-the-art and good practice in the field of living labs”. In Proceedings of the 12th International Conference on Concurrent Enterprising: Innovative Products and Services through Collaborative Networks, Milan, Italy, pp.349-357.

Poggi I (2008). “The language of interjections”, In COST 2012 School: 170–186.

Renault S. (2004). “Du concept de fragilité et de l'efficacité de la grille AGGIR”, in Gérontologie et société. 2004, n° 109, pp.83-107.

Sasa Y., Aubergé V., Franck P., Guillaume L., Moujtahid S. (2012). “Des micro-expressions au service de la macro-communication pour le robot compagnon EMOX”, Actes du WACAI 2012, Grenoble, pp.54-59.

Sasa, Y., Aubergé V, and Rilliard A. (2013) “Audio-Visual Micro-Expressions within Japanese-French Contrast.” In WASSS 2013. Grenoble, France

Schaal, S. (1999) “Is imitation learning the route to humanoid robots?” Trends Cognit. Sci.3, 233-242.

Scherer, K.R., “Affect bursts” (1994). In S.H.M. van Goozen, N. E. van de Poll & J.A. Sergeant (Eds.), Emotions, Hillsdale (NJ, USA), Lawrence Erlbaum, 161-193

Schröder M., Heylen D., Poggi I. (2006) “Perception of non-verbal emotional listener feedback”. In Proceedings of Speech Prosody 2006.

Schröder, M., “Experimental study of affect bursts”, Speech Communication, 40(1-2), 99-116, 2003

Segrin, C. (1994). Social skills and psychosocial problems among the elderly. Research on Aging, 16(3), 301-321.

Shibatani. M. (1990). “The languages of Japan”.

Signorello, R., Aubergé, V., Vanpé, A., Granjon, L., & Audibert, N. (2011). Indices de langue et de culture dans les micro-événements audibles et visibles de l'interaction face à face. In 9ème Rencontres des Jeunes Chercheurs en Parole 2011 (RJCP).

Sun X., K. Truong, A. Nijholt, and M. Pantic, (2012) "Automatic Visual Mimicry Expression Analysis in Interpersonal Interaction," in Proc. IEEE Int’l Conf. Computer Vision and Pattern Recognition (CVPR-W’11), Workshop on CVPR for Human Behaviour Analysis, pp. 40-46.

Tomaka, J., Thompson, S., & Palacios, R. (2006). The relation of social isolation, loneliness, and social support to disease outcomes among the elderly. Journal of Aging and Health, 18(3), 359-384.

Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. (2005) “Understanding and sharing intentions: The origins of cultural cognition”. Behavioral and Brain Sciences 28(5), pp. 675–91.

Vacher M., A. Fleury, F. Portet, J.-F. Serignat, N. Noury (2010) Complete Sound and Speech Recognition System for Health Smart Homes: Application to the Recognition of Activities of Daily Living, New Developments in Biomedical Engineering, Intech Book, pp. 645-673.

Vanpé, A., & Aubergé, V. (2010). Prosodie expressive audio-visuelle de l'interaction personne-machine. Techniques et Sciences Informatiquess, 29(spécial Agents Conversationels Animés), 880-832.

Ward, N. "Non-lexical conversational sounds in American English", Pragmatics & Cognition, 14(1), 129-182 (54), 2006.

Wichmann A. (2000). "The attitudinal effects of prosody, and how they relate to emotion". In proceedings of ITRW on Speech and Emotion, Newcastle, Northern Ireland, UK.