Multimodal breathiness in interaction: from breathy voice quality to global breathy "body behavior quality"

Liliya Tsvetanova, Véronique Aubergé, Yuko Sasa

LIG-lab, CNRS UMR 5217, Grenoble Alpes University, France
{Firstname}.{Lastname}@univ-grenoble-alpes.fr

Abstract

The face-to-face interaction is a complex dynamic process in which the interactants are mutually, continuously and reciprocally sharing with each other vocal and non-vocal information. This multimodal information has been defined as relevant to the speakers' interpersonal intimacy degree in the relation process ("socio-affective glue"). Otherwise, the human multimodal intimacy cues dynamics is a crucial aspect of the analysis of the human socio-affective behavior. To address this point, our study is based on the EEE corpus involving spontaneous dialogs between an elderly and a smart home control robot whose vocalizations are primitive pure prosodic expressions. The gradually increasing socio-affective gluing (intimacy) effect of the robot's vocalizations has been shown in previous studies. The utterances produced by the subjects are imposed commands (smart home orders). However, the elderly's vocal and non-vocal behavior changes gradually throughout the experiment by varying in breathiness, commands paraphrasing, and non-vocal cues, which we suppose could be meaningful for the nature of the human-robot relation. In this study, the communication dynamics is observed as an overall behavior. Accordingly, we suggest that the breathiness dynamics could be superposed, in the dialog time, to the dynamics of both the morpho-lexico-syntactical style and the proxemic cues (as postural proximity and gaze direction).

Index Terms: breathiness, socio-affective "glue," multimodal interaction, Human-Robot Interaction (HRI)

1. Introduction

Nowadays, one of the crucial aspects in the field of human-robot interaction is to provide the robot with acceptable and ethical interactional behavior so that it could be a part of the human social environment. In order to address this point, one of the approaches used in the research field is to focus on the robot recognition (detection and analysis) of the human speaker vocal and non-vocal behavioral cues observed during the interaction process (see [1] for an overview).

In terms of vocal behavior, it is well known that an important quantity of information about the speakers' affective state is likely to be discerned throughout the one's vocalizations and this phenomenon is observed in both human and animal species [2]–[4]. Accordingly, one utterance (even as small as the burst "eh") could be vocalized differently to express quite a different meaning [5], and so, quite a different affective state [6]. Acoustically, this kind of subtle affective vocal information is given by the expression style, namely the speech prosody [2]. Moreover, regarding social relations, the prosodic cues have been reported in the literature as an important vocal aspect indicating the speaker's attitude toward his or her interactional partner [5]. Thus, the prosody has been shown as a relevant factor in the establishment of an interpersonal connection between the interactants and this process of connectedness, which is called socio-affective "glue" [7], is based on an altruistic bond built according to the principles of mutual social grooming [8]. Nonetheless, the affective function of the prosodic expressiveness is unlikely to be related to its phonetic parameters [9], [10], but to a 4th prosodic dimension known as voice quality [11]. According to the findings, the voice quality refers to a specific folds' functioning/vibration (more or less folds openness than for modal speech), which seems to be led by the speaker affective state [12] and is reported as having social signaling functions [9]. One particular voice quality (see [13] for an overview) – the breathy voice quality (also known as breathy phonation) – is reported in the literature as the vocal manifestation of the interpersonal *intimacy* and *caring* [12]–[17], namely the highest degrees of altruistic gluing between the interactants.

According to the literature, a variety of non-vocal behaviors are likely to indicate heightened involvement in the interaction process. These behaviors have been studied in the field of proxemics [18], according to which interactional partners in a close relationship tend to unconsciously show their "closeness" physically during the interaction process. The proxemics is related to the notion of physical intimacy (also known as physical closeness or physical distance) [19], and it is observed throughout some non-vocal cues such as the postural proximity [20] and the gaze direction [21]. Moreover, recent studies suggest that the hand gesture could also be related to the relational gluing process [22], [23]. Thereby, according to the proxemic analysis, the whole body appears as an instrument able to express the intimacy level or, as we will refer to it in this paper, the "glue level."

The cited studies state that both vocal and non-vocal behaviors could reflect a kind of intimacy between the interactional partners. Thus, it seems that in the case of intimacy, the interactional partners create together a mental resonation [19], which could merge into a more global process as the interpersonal synchrony [24], [25], during which the speakers create an interactional dynamics resulting in temporal coordination of their behaviors. However, the dynamics (whether individual or interpersonal) of the observed interactional modalities is rarely taken into account in the context of interaction between a human and a robot [1], [26].

The purpose of this preliminary study based on the observation of spontaneous dialogs between elderly and a robot is to show that the dynamics of the vocal breathiness (a tangible intimacy indicator) is related in the dialog time to the dynamics of the glue building. Moreover, as the prosody dynamics could be transmitted and perceived throughout some non-vocal cues such as hands and face gesture [27], [28], we suppose that the

breathiness dynamics could be directly related to the dynamic changes in other vocal and non-vocal intimacy cues such as the linguistic form of the commands addressed to a robot, the postural proximity and the gaze direction. The affective communication could thus be considered as more global behavior characterized by the multimodal breathy dynamics.

2. Background: EEE spontaneous dialogs data corpus

The data used in this study are from the EEE (Elderly Emox Expressions) corpus [29] involving spontaneous dialogs between the non-anthropomorphic robot Emox (Awabot company) and socio-isolated elderly.

1.1. Study contextualization

In order to observe how the altruistic relation emerges in a micro-social system as elderly, a set of primitive sounds supposed to remain a fundamental tool for the mutual building of the socio-affective glue has been implemented in the Emox robot. Those vocalizations are gradually ordered according to their supposed gluing straight (see [29] for more details) in the following range: (1) no speech, (2) pure prosodic mouth noises, (3) interjections and lexicons with supposed gluing prosody, and (4) some subjects' commands imitations with supposed glue prosody, knowing that the imitation, or the so-called chameleon effect [30], has a high potential of establishing relationships. Those vocalizations notified as the lifeblood of the intimacy establishment between interactional partners were observed in a dialog context with socio-isolated elderly for whom the creation of this dynamic relational process seems to be more difficult [31]. In fact, the rate of social isolation increases with aging [32] and with the absence of intimate interactions [19], which seriously affects the elderly's communication skills, which are an essential part of creating and preserving the elderly's social relationships. In other words, the elderly's abilities to create the relational gluing process are damaged [29].

1.2. Collecting ecological data with a Wizard of Oz experimental scenario and "glue level" retrieval

The natural elderly-robot dialogs corpus has been collected using an experimental Wizard of Oz scenario (see [29] for more detailed information), which took place *in situ* in the Living Lab Domus (Computer Sciences Laboratory of Grenoble, France) arranged as a smart-home prototype. In this study, elderly subjects were invited to use a smart home control robot to carry out an imposed list of 31 home automation commands. The experiment was followed by an auto-annotation session, which took place a few weeks later. During this session, each participant was prodded to review, in order to involve his or her autobiographical memory [33], the whole experiment's video recording and to qualify their mental state at every moment of the interaction with the robot. The auto-annotation session aims are multifold, as follows: (1) to define the socio-affective glue value by the participant himself / herself while avoiding a possible incorrect "expert" interpretation of the collected data, (2) to observe the glue global and progressive transformation through the interaction process, and also (3) to detect the breaking points determining the border lines of each glue level.

Regarding the collected data, the EEE corpus comprises a video and audio data captured by the six ceiling cameras and six ceiling microphones in Domus; the participants were also equipped with a headset microphone allowing high-quality speech sounds collection. Concerning the headset microphone use, to avoid every suspiciousness about the fact that speech data are recorded, we let the participants think that the Emox' sound capturing sensor was broken and the only way to give commands to the home control robot was by using the robot's microphone. The auto-annotation session was also audio recorded. All the captured data were temporarily aligned in the ELAN software program, and all the speech was orthographically transcribed.

3. Methods

For this study, we choose from the EEE spontaneous dialogs corpus the multimodal data of five elderly subjects, all women and French native speakers, from 69 to 89 years old and with none or low dependency (corresponding respectively to the 6th and 5th grades according to the French elderly dependency scale AGGIR [34] grouping the elderly from 1 – very dependent to 6 – no dependent). The data represent a total of 226 minutes (approximately 3.8h) of audio and video records with full speech transcription, and a total of 340 commands addressed to the robot.

The audio record from the headset microphone was used to analyze the breathiness level of all vocal commands extracted from the selected corpus. A number of solutions for automatic calculation of the breathiness exist, and their functioning is based on different acoustic properties reported in the literature, such as H1-A3 (difference between the amplitudes of the first harmonic and the third formant) [35], NAQ (Normalized Amplitude Quotient of the glottal waveform and its derivative waveform) [36], HNR (Harmonics-to-Noise Ratio) [37] or the inverse NHR (Noise-to-Harmonics ratio) [38], GNE (Glottal-to-Noise Excitation ratio, which needs inverse filtering to avoid the problem of high-pitched voices) [39], F-aperiodic (boundary frequency between harmonic and aperiodic components) [40] and even F1F3syn (synchronization of the amplitude envelopes of the first and third formant frequency bands) [41]. However, those measures either could not be used for spontaneous speech analysis or are not adapted for elderly high-pitched female voices, which are known as naturally breathier (due to the muscular slackening, which increases with aging [42]). For that reason, in this study, we proceed by an expert labeling of the voice quality (with emphasis on the breathiness level) during the vocal production of the commands.

The elderly's non-vocal behavior has been analyzed on video recordings, focusing on proxemics cues as posture (subject's body position), physical proximity (relative to the Emox' position) and gaze direction (head and eyes cast direction). These modalities have been annotated according to a list of labels (cf. Table 1 below) and only those performed during the command time have been considered in our analysis.

Table 1: *List of labels used to annotate the subject's posture, proximity to the robot and gaze direction.*

Modality	**Labels**
Posture	Standing
	Sitting
	Crouching down
	Laying on the bed
	Leaned forward
Physical Proximity	In other room

	In the same room Close (50cm) Close+ (25cm) Close++ (touch)
Gaze direction	Emox Commands list Object concerned by the command Object that is different from the object concerned by the command Human interlocutor

4. Results

Global analysis of the data from elderly, concerning the voice quality during the commands announcement, showed that the commands are produced in either modal or breathy voice. However, the observation of the breathiness dynamics throughout the experiment revealed that each vocal command could be seen as lying along a continuum of breathiness. On this continuum, the command voice quality varies globally from modal tense (no breathiness) to breathy lax (high breathiness level). Specifically, the modal voice (voice without breathiness) is associated with the lowest glue levels (when the relation between the elderly and the robot is not yet established), and the breathy voice is associated with the highest levels of socio-affective glue (at the end of the experiment when the robot's vocalizations are the most charged in glue). The breathiness scale varies in accordance (or even in response) to the robot's gluing vocalizations. In this way, the breathiness dynamics is following, in a progressive fashion, the socio-affective glue dynamics as it is illustrated in Figure 1 below.

An analysis consolidating the subjects' multimodal behavioral data compared to both the glue level and the previous breathiness observations showed some general tendencies in the variations of the linguistic style and the proxemic modalities. In accordance with the glue potency variation, every separate modality seems to evolve gradually on its continuum as shown in Figure 1 below. Figure 1 also illustrates the moment of emergence of some proxemic cues, which coincided with the first robot's vocalizations and the appearance of the vocal breathiness in the commands. The emergence of the cues revealing lower physical and vocal distance was the beginning of the closeness manifestation in the other modalities. At this moment, the linguistic form started to change from the imposed infinitive form into a "we" form manifesting a kind of togetherness or "we-ness," and then, into an "you" form that seems to be a characterization of the robot like a different entity (which is confirmed by the auto-annotations noting that at this moment, the robot is "like a child," "like another"). This temporal boundary line is also the beginning of the proxemic cues in a manner, which also showed lower distancing. Thus, postural proximity and gaze direction seem to be dynamically interrelated: the increase of the physical distance (decrease of the physical closeness) increases the gaze occurrences.

However, a more detailed data analysis of the behavior of every elderly person showed that the glue obtained from the auto-annotations arranges the subjects in three distinctive groups: (a) those who did not glue (one subject), (b) those who moderately glued (two subjects) and (c) those who glued strongly (two subjects). Moreover, the different gluing type seems to induce different dynamics variations in terms of linguistic behavior (breathiness and morpho-lexico-syntactical style) and proxemic behavior (postural proximity and gaze direction). By taking into account all cited modalities, the three profiles could be summarized as follows:

(a) *Low gluing profile*: the majority of the produced commands are in modal (tense or lax) voice quality (91%), and there are no commands emitted with either breathy or breathy lax voice. The infinitive structure of the command is maintained throughout the whole experiment. In terms of proxemic modalities, the subject maintains a close distance to the robot (labeled as "close"), but she is rarely looking in the direction of the robot; her gaze is more often oriented in the direction of the list of commands and the object concerned by the command.

(b) *Medium gluing profile:* these subjects' voice quality varies from modal tense to breathy lax, with a high percentage of modal lax and breathy tense commands (mean value of 52%). There are very few commands forms modifications. The physical distance with the robot decreased progressively from "close" at the begging through "close/close+" to "close+" at the end of the experiment when the gluing level is the highest. The gaze direction also changes in a similar way: at the experiment's beginning the preferred gaze targets are the list of commands, the object of the command, and Emox, while at the end of the experiment, the preferred gaze targets are the command object and the environment.

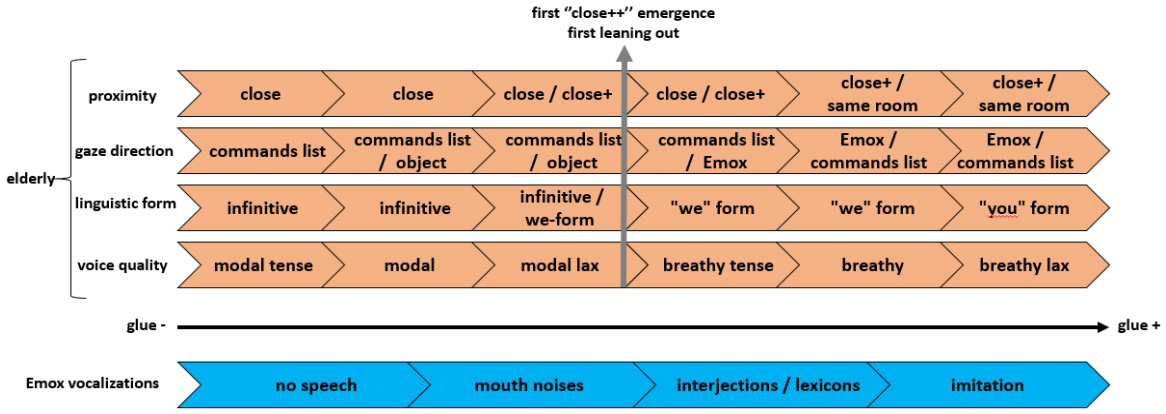

Figure 1: *The subjects' multimodal behavior variation in accordance with the robot's vocalizations and the socio-affective "glue level" (experiment's time)*

(c) *High gluing profile*: the voice quality variation begins from modal to breathy lax, with the highest percentage of breathy and breathy lax commands (on average 60%). The commands form varies accordingly to the perceived breathiness in a progressive manner from the infinitive to a form with the pronoun "we" (e.g. "We turn on the lights" / "On allume la lumière") and then into a form with the pronoun "you" (e.g. "You turn on the lights" / "Tu allumes la lumière"). The proxemic behavior is labeled as "close+" all along the experiment, the first noticeable closeness appears with the "close++" labels in accordance with some body leanings forward in the direction of the robot. The modal voice quality is associated to highly frequent glances in the direction of the list of commands, and the breathy voice quality is associated to glances in the direction of Emox.

5. Discussion

In this study, we tried to investigate the interactional dynamics in an elderly-robot interaction context. First of all, we supposed that as the breathiness is a proved intimacy vector, its dynamics follows the socio-affective glue dynamics. Secondly, we tried to show that the breathiness dynamics could characterize the rhythm of variation (or the dynamics) of the overall elderly's multimodal behavior. In an effort to check those affirmations, we analyzed the elderly's breathiness, linguistic styles and proxemic (postural proximity and gaze direction) behavior within the context of spontaneous interaction between the elderly and a butler robot implemented with socio-affective glue vocalizations.

As shown above, the point in the dialog time of the emergence of the first gluing vocalizations of the robot is corresponding to the beginning of the changes in the elderly's vocal and non-vocal behavior dynamics. On the one hand, the dynamics in the robot's vocalizations seems to influence the human vocal expression in a progressive way. Thus, more the robot's sounds are charged in glue (intimacy), more the elderly voice is breathier and more the level of reported glue is higher. A possible explanation of the dynamic modification of the human vocal behavior according to the robot's vocalizations could be the usual process of synchrony, evolving in the interaction process in whom the speakers share an important degree of intimacy. On the other hand, the analysis of both the breathiness dynamics (as socio-affective glue indicator) and the dynamics of the other modalities (command form, postural proximity, and gaze direction) reveals that there is no a complete temporal correspondence in the modalities changes. So, the granularity of our analysis allows us only to refer to the global tendencies but does not allow us to observe if two (or more) cues tend to change together dynamically. However, we could affirm that all the human modalities change in a manner, which shows vocally and non-vocally a general tendency to go towards more "close" (intimate) behavior.

The reported differences between the groups of subjects indicate that every subject reaches a different glue level, and not all the participants reach the highest glue level. The literature in the field of human-robot interaction often tends to explain the difference in human behavior by the personality traits, but we think that in our study the explanation is quite different. Our experience with the five elderly subjects lets us suppose that this difference in the elderly's gluing behavior could be explained by the "degree" of social isolation. Thus, the profile of elderly who glue less with the robot corresponds to the elderly who are less isolated.

Knowing that the elderly subjects who glued the most in the experiment modify the form of the commands from the imposed list of home automation commands as reported above, we expected to find a decreasing number of glances at the list of commands when the form is modified. Surprisingly, as shown by the results, the occurrences of gazes in the direction of the list remain high, even when the commands are in "we" form and "you" form. Even if the elderly continue to use the list of commands, it seems that this gaze behavior could be explained by a phenomenon of cognitive detachment of the list. This finding suggests that the speech recognition systems (which nowadays are based on lists of commands) have to take into account the state of the established relationship between the interactants. The first works in this direction have been implemented in the robot's dialog system called SARSI (Socio-Affective Robotics Speech Interaction), which is constructed accordingly to the socio-affective glue paradigm.

6. Conclusions

The acoustic breathiness dynamics is in high accordance with both the relation dynamics (the "glue life") and the global multi-dimensional elderly behavior. Thus, we observed the breathiness not only as a vocal quality but overall as a global interactional behavior quality, which could be seen as an indicator of the relation nature between the interactants. Some previous works point out the existence of body prosody as more holistic interactional (verbal and non-verbal) behavior, which is essential for the interaction. In this study, we referred to the intimate prosodic dimension – the breathiness – in order to show that its dynamics could be observed in the dynamic variation of the other intimate cues. This overall dynamics is important, not only for the interaction process but moreover, for the affective, relation process between interactants. In this case, the intimacy (namely the socio-affective glue) appears as a cognitive motor (as stated in the literature) for the discernable global breathiness behavior, which is materialized in what is said, how it is said and how it is shown by the proxemic cues as the postural proximity and the gaze direction.

7. Acknowledgements

This work was partially funded by French grants Interabot, parts of BGLE no2 Investissements d'Avenir and it has been partially supported by the LabEx PERSYVAL-Lab (ANR- 11-LABX-0025-01) and the Major Program for the NSSF of China (13&ZD189). We would like to thank the Awabot Company (robotics), Bien A la Maison Company (caregiving services) and Roger Meffreys elderly housing for their collaboration in this study. Thank you to Romain Magnani, Natacha Borel and Ambre Davat for their support.

8. References

[1] A. Vinciarelli, M. Pantic, and H. Bourlard, "Social signal processing: Survey of an emerging domain," *Image Vis. Comput.*, vol. 27, no. 12, pp. 1743–1759, 2009.

[2] K. R. Scherer, "Vocal communication of emotion: A review of research paradigms," *Speech Commun.*, vol. 40, no. 1, pp. 227–256, 2003.

[3] R. M. Seyfarth and D. L. Cheney, "Meaning and emotion in animal vocalizations," *Ann. N. Y. Acad. Sci.*, vol. 1000, no. 1, pp. 32–55, 2003.

[4] P. Pongrácz, C. Molnár, and Á. Miklósi, "Acoustic parameters of dog barks carry emotional information for humans," *Appl. Anim. Behav. Sci.*, vol. 100, no. 3, pp. 228–240, 2006.

[5] N. Campbell, "Perception of affect in speech-towards an automatic processing of paralinguistic information in spoken conversation.," in *INTERSPEECH*, 2004.

[6] M. Schröder, "Experimental study of affect bursts," *Speech Commun.*, vol. 40, no. 1, pp. 99–116, 2003.

[7] V. Auberge, Y. Sasa, T. Robert, N. Bonnefond, and B. Meillon, "Emoz: a wizard of Oz for emerging the socio-affective glue with a non humanoid companion robot," in *WASSS 2013*, Grenoble, France, 2013.

[8] H. Nelson and G. Geher, "Mutual grooming in human dyadic relationships: an ethological perspective," *Curr. Psychol.*, vol. 26, no. 2, pp. 121–140, 2007.

[9] K. R. Scherer, "Vocal affect expression: a review and a model for future research.," *Psychol. Bull.*, vol. 99, no. 2, p. 143, 1986.

[10] T. Johnstone and K. Scherer, "The effects of emotion on voice quality," in *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco: University of California, Berkeley, 1999, pp. 2029–2032.

[11] N. Campbell and P. Mokhtari, "Voice quality: the 4th prosodic dimension," in *15th ICPhS*, Barcelona, Spain, 2003, pp. 2417–2420.

[12] C. Gobl and A. Ni Chasaide, "The role of voice quality in communicating emotion, mood and attitude," *Speech Communication 40 (1)*, pp. 189–212, 2003.

[13] J. Laver, "The phonetic description of voice quality," *Camb. Stud. Linguist. Lond.*, vol. 31, pp. 1–186, 1980.

[14] C. R. Rogers, *On becoming a person: A therapist's view of psychology*. Boston: Houghton Mifflin, 1961.

[15] A. Wichmann, "Attitudinal intonation and the inferential process," in *Speech Prosody 2002, International Conference*, 2002.

[16] N. Campbell, "On the use of nonverbal speech sounds in human communication," in *Verbal and nonverbal communication behaviours*, Springer, 2007, pp. 117–128.

[17] N. Audibert, V. Aubergé, and A. Rilliard, "When is the emotional information? A gating experiment for gradient and contours cues," in *Proceedings of ICPhS XVI Meeting. Saarbrucken*, 2007, pp. 6–10.

[18] E. T. Hall, "The hidden dimension," 1966.

[19] H. T. Reis and P. Shaver, "Intimacy as an interpersonal process," *Handb. Pers. Relatsh.*, vol. 24, no. 3, pp. 367–389, 1988.

[20] S. E. Scherer and M. R. Schiff, "Perceived intimacy, physical distance and eye contact," *Percept. Mot. Skills*, vol. 36, no. 3, pp. 835–841, 1973.

[21] M. Argyle and J. Dean, "Eye-contact, distance and affiliation," *Sociometry*, pp. 289–304, 1965.

[22] L. Guillaume *et al.*, "HRI in an ecological dynamic experiment: the GEE corpus based approach for the Emox robot," presented at the IEEE International Workshop on Advanced Robotics and its SOcial impacts (ARSO), Lyon, France, 2015.

[23] M. Girard-Rivier *et al.*, "Ecological Gestures for HRI: the GEE Corpus," 2016. [Online]. Available: http://www.lrec-conf.org/proceedings/lrec2016/pdf/1040_Paper.pdf. [Accessed: 25-Sep-2016].

[24] E. Delaherche and M. Chetouani, "Multimodal Coordination: Exploring Relevant Features and Measures," in *Proceedings of the 2nd International Workshop on Social Signal Processing*, 2010, pp. 47–52.

[25] O. Weisman, E. Delaherche, M. Rondeau, M. Chetouani, D. Cohen, and R. Feldman, "Oxytocin shapes parental motion during father-infant interaction," *Biology letters*, 2013.

[26] L.-P. Morency, "Modeling human communication dynamics," *IEEE Signal Process. Mag.*, vol. 27, no. 5, pp. 112–116, 2010.

[27] D. Brentari and L. Crossley, "Prosody on the hands and face: Evidence from American Sign Language," *Sign Lang. Linguist.*, vol. 5, no. 2, pp. 105–130, 2002.

[28] Guellaï, Bahia, Langus Alan, and Nespor Marina, "Prosody in the hands of the speaker," *Frontiers in Psychology*, 2014.

[29] V. Aubergé *et al.*, "The EEE corpus: socio-affective' glue' cues in elderly-robot interactions in a Smart Home with the EmOz platform," in *5th International Workshop on EMOTION, SOCIAL SIGNALS, SENTIMENT & LINKED OPEN DATA*, Reykjavik, Iceland, 2014.

[30] T. Chartrand and J. Bargh, "The chameleon effect: the perception-behavior link and social interaction," *J. Pers. Soc. Psychol.*, vol. 76, no. 6, pp. 893–910, 1999.

[31] K. A. Bayles, A. W. Kaszniak, and C. K. Tomoeda, *Communication and cognition in normal aging and dementia*. College-Hill Press/Little, Brown & Co, 1987.

[32] J.-L. Pan Ké Shon, "Isolement relationnel et mal-être," 2003.

[33] H. L. Williams, M. A. Conway, and G. Cohen, "Autobiographical memory," *Mem. Real World*, p. 21, 2007.

[34] J. Belmin *et al.*, "Level of dependency: a simple marker associated with mortality during the 2003 heatwave among French dependent elderly people living in the community or in institutions," *Age Ageing*, pp. 298–303, 2007.

[35] H. M. Hanson, "Glottal characteristics of female speakers.," 1995.

[36] P. Alku and E. Vilkman, "Amplitude domain quotient for characterization of the glottal volume velocity waveform estimated by inverse filtering," *Speech Commun.*, vol. 18, no. 2, pp. 131–138, 1996.

[37] C. T. Ferrand, "Harmonics-to-noise ratio: an index of vocal aging," *J. Voice*, vol. 16, no. 4, pp. 480–487, 2002.

[38] D. Deliyski, "Effects of aging on selected acoustic voice parameters: Preliminary normative data and educational implications," *Educ. Gerontol.*, vol. 27, no. 2, pp. 159–168, 2001.

[39] D. Michaelis, T. Gramss, and H. W. Strube, "Glottal-to-noise excitation ratio–a new measure for describing pathological voices," *Acta Acust. United Acust.*, vol. 83, no. 4, pp. 700–706, 1997.

[40] T. Ohtsuka and H. Kasuya, "Aperiodicity control in ARX-based speech analysis-synthesis method.," in *INTERSPEECH*, 2001, pp. 2267–2270.

[41] C. T. Ishi, "A new acoustic measure for aspiration noise detection.," in *INTERSPEECH*, 2004.

[42] H. Hollien, "'Old voices': What do we really know about them?," *J. Voice*, vol. 1, no. 1, pp. 2–17, 1987.