Artificial Empathy in Social Robots: An analysis of Emotions in Speech

Conference Paper · August 2018

DOI: 10.1109/ROMAN.2018.8525652

3 authors, including:

Jesin James
University of Auckland

Catherine I Watson
University of Auckland

Proceedings of the 27th IEEE International Symposium
on Robot and Human Interactive Communication,
Nanjing, China, August 27-31, 2018

Artificial Empathy in Social Robots: An analysis of Emotions in Speech

Jesin James[1] Catherine Inez Watson[1] Bruce MacDonald[1]

Abstract— **Artificial speech developed using speech synthesizers has been used as the voice for robots in Human Robot Interaction (HRI). As humans anthropomorphize robots, an empathetically interacting robot is expected to increase the level of acceptance of social robots. Here, a human perception experiment evaluates whether human subjects perceive empathy in robot speech. For this experiment, empathy is expressed only by adding appropriate emotions to the words in speech. Also, humans' preferences for a robot interacting with empathetic speech versus a standard robotic voice are also assessed. The results show that humans are able to perceive empathy and emotions in robot speech, and prefer it over the standard robotic voice. It is important for the emotions in empathetic speech to be consistent with the language content of what is being said, and with the human users' emotional state. Analyzing emotions in empathetic speech using valence-arousal model has revealed the importance of secondary emotions in developing empathetically speaking social robots.**

I. INTRODUCTION

Robotic technology development has enabled robots to be implemented in human-centered environments, interacting with people, and involving in activities. Human-Robot Interaction (HRI) research focuses on how robots react to humans in a social and engaging way. Such social robots are used for applications like education and healthcare. Healthcare robots are the focus of this paper. The healthcare robots - *Healthbots* [1, 2] under development at the University of Auckland, New Zealand are an application of HRI that provide support and care for older people. When developing robots that interact with humans, their acceptance is a primary concern. Some factors that enhance robots' acceptance are their appearance, humanness, personality, expressiveness and adaptability [3]. Currently, roboticists build robots that look like humans to improve their acceptance. However, users are disappointed by the lack of reciprocal empathy from robots [4]. Due to this gap, a concept called Artificial Empathy (AE) has been introduced in HRI [5, 7, 8]. AE is the affective response portrayed by the artificial intelligence in companion robots. Currently, robots and other artificial agents do not have the ability to "feel" as a human would, hence AE is a programmed reaction modeled on robots. The aim of current research is to incorporate a model for empathetic behavior in social robots [5]. This paper studies the role of speech modality in the expression of empathy in robots. Firstly, a human perception test was conducted to understand whether an empathetically sounding robot is acceptable to people. This test explores into the role that the type of synthesised voice has on people's perception of empathetic voice. Following this, an analysis of the emotions portrayed in the robots' dialogs is conducted based on models used in human-human interaction.

*This work is supported by Seed funding from Centre for Automation and Robotic Engineering Science (CARES), University of Auckland. The authors thank the participants of the perception test for their time and effort.

[1]Jesin James, Catherine Inez Watson & Bruce MacDonald are with Dept. of Electrical & Computer Engineering, University of Auckland, NZ, jjam194@aucklanduni.ac.nz, c.watson@auckland.ac.nz, b.macdonald@auckland.ac.nz

II. MOTIVATION

A. Impact of Robotic Speech on Anthropomorphism

Robots that interact in social situations are novel to people who use them due to their limited experience in actually interacting with robots [9]. People rationalize this novelty by anthropomorphism, i.e. projecting familiar human-like characteristics, emotions and behavior onto robots. Speech is a primary mode of interaction between robots and humans. Examples of social robots that use speech to interact with people are Kismet [17], a story teller robot [18], robot guide [19] and reception robots [20]. People's anthropomorphism of robots is impacted by the type of voice used by robots to converse, and this also affects the robots' acceptance. People make judgments about robots' personalities based on their voice [9, 10]. It was observed that the acceptance for the robot was enhanced when a human-like voice was used compared to a robotic voice [12]. Another experiment investigated the impact when affect was added to a robot's speech when the robot was instructing a team to complete a task [13]. A different team performed the same task instructed by a robot with a neutral voice. Here, performance of the team that did the task under the expressive robotic voice was better than the other team. Also, [14] shows that more playful speaking style produced more willingness to perform a task, compared to a serious speaking style. This shows that changes in the synthesized voice of the robot have an impact on the way people anthropomorphize robots.

B. Empathy in Robotic Speech

The positive benefits of users anthropomorphizing robots can be achieved only if the robot's capabilities to behave and look like a human are matched with its communication modalities. Studies have shown that robots with an empathetic voice received positive ratings for likeability, trustworthiness, and were also perceived as supportive [11]. Empathetic robots reduce stress among users, improve their comfort and performance in task achievement [21]. The positive effects of showing empathy are produced in users only when the robot's expressions are in congruence with the users' affective state [21]. Hence, good modeling of empathy and expressivity is required to avoid a mismatch between the robots' behaviour and communication via speech. Due to empathy in robots being a relatively new concept, the tendency for humans to anthropomorphize robots is used

978-1-5386-7980-7/18/$31.00

as the key to derive a definition here. Empathy in Human-Human Interaction is the behavior that enables one human to experience what another human feels and respond to it. It is an emotional response that is automatically evoked by one's understanding of the other human [22]. When the companion is a robot, we define empathy in HRI (Artificial Empathy) as the programmed affective reaction of the robot to the behavior of the human that it can sense according to the technology embedded in it. The empathy portrayal by humans involve various communication modalities, such as facial, vocal (non-verbal and verbal) [4]. For robots, these communication modalities exist. There have been attempts to model empathetic behavior in social robots using modalities such as facial expressions, gestures and speech [6-8,34,37]. But the focus on this work is to use the vocal channel (speech) to express empathy. Speech has two components [31]: the verbal component and the prosody component. Verbal component focuses on the words alone and is defined using combinations of linguistic symbols. Most of the factors that make human speech more natural compared to monotone speech can be summarized under the prosody component. Emotions are expressed by variations in prosody component (like varying intonation, speech rate, stress) [31,39]. Empathetic behavior via speech can be depicted by a proper choice of words which is the verbal component and the emotions portrayed by the speaker which is the prosody component. The choice of words determines the lexical features, and emotions govern the acoustic features pertaining to prosody [30]. Often empathy is incorporated into synthesized speech by the inclusion of words that convey an affective response (called dialog modelling). This study examines how the emotions in the speech signal affect the perceived empathy in users that interact with robots. To understand if people can perceive that a robotic companion is empathetic to them when empathetic behaviour is expressed only by the emotions in speech, a human perception test was conducted as explained in the next section.

III. HUMAN PERCEPTION EXPERIMENT

A. Experiment Design

The Human Perception test was conducted to address the following research questions:

1: To understand if people can perceive empathetic behavior from a robot when only the emotions(i.e. prosody component) in its speech are used to express empathy.

2: To understand if people prefer empathetic voice from robots compared to a standard non-empathetic robotic voice.

3: To understand which factors of speech can be related to an empathetic voice.

The robot used for the perception experiment is the Heathbot. The Healthbots were chosen as they are a good exemplar for a social robot designed to interact with people, provide medication reminders, measure vital signs and provide entertainment instructions. A New Zealand English voice has been incorporated in the Healthbots and studies have been conducted to evaluate the naturalness of the voice [23, 24]. As the experiment is to evaluate robots in social

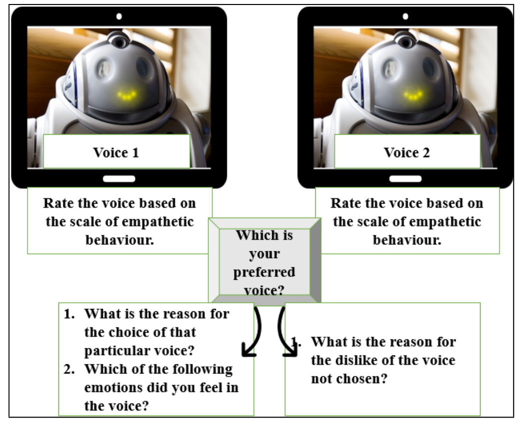

Fig. 1. Human perception test flow diagram

situations, the scenario presented to the participants is that of a Healthbot and a patient interacting with each other. There are 3 different situations in which the robot speaks to the patient - greeting the user, providing medicine reminders and guiding the user to use the touch interface. Dialogs were framed for each of these situations including dialogs already used by the Healthbots (more details about dialogs are in Section IV). Each situation had 20-25 dialogs. The dialogs were produced by a professional voice artist in two variations. One variation used a monotone voice, just like a standard robotic voice. There was no variation in prosody features like intonation and intensity. The second variation was spoken like a nurse speaking empathetically to a patient, with changes in emotions. A professional voice artist was used instead of using synthesised robotic and emotional voices as the current synthesised voices used in the Healthbot lack in naturalness and quality as they are still under development. This can cause the participants decisions being affected by the limitations in the naturalness of the voice. Also, robotic voices that speak empathetically have not been synthesised yet. This study is a pre-requisite to understanding what type of voices are preferred by human participants. Due to these two constraints, a professional voice artist was used. Both the variations used the same verbal content with variation in only the prosody component and the robot had a neutral facial expression (for testing *Research Question 1*). The dialog of the patient was spoken by a speaker in the same manner regardless of the variation in the robot's voice. The professional speaker's voice was used as the voice for the robot in the study. An example for a Healthbot dialog was "It seems like you are taking a long time to take your medicine.", for which the patient responds "I am lately very slow in all the tasks I do!". The flow diagram of the experiment with the robot used is shown in Fig. 1. The participants could see and hear the Healthbot speaking but the patient speaking to the robot was not shown to them. The patient was not shown to enable the participants to feel as if the robot was speaking to them and hence they could rate the robot's interaction to them. Each participant was given one scenario (greetings, reminders or instructions) with the two variations (standard robotic voice and empathetic voice). Both variations were shown to each participant to provide a scenario to judge *Research Question 2* of the experiment. The dialogs lasted for almost 1-2 minutes for each of the variations. The participants were then asked questions related to the speech of the robot to obtain the results for *Research Question 3* of the experiment.

The perception test was completed by 120 participants, aged 16-45. Based on their self-reporting, all participants had above average hearing ability, with 50 participants being first language New Zealand English speakers and 70 were bilingual speakers. All participants completed the test with 20% and 80% of them using loudspeakers and headphones respectively. Each participant took approximately 15 minutes for the test. An online survey platform was chosen to maximize the number of participants, which helps to generalize the findings. A generalized participation was chosen as the Healthbots will be used in applications where the users may not have any knowledge about rocotics.

B. Scale of Empathy Questionnaire

The questions asked to the participants for *Research Question 1* were based on the Empathy measuring scale from the Motivational Interviewing Treatment Integrity (MITI) module [15, 16] which defines 5 scales to rate a clinician's empathy. The 5 point scales used in MITI and in the experiment are shown in Table I. A score of 1 is the least empathetic. Column 2 gives the characteristics of a clinician evaluated for empathy rating in human-human interaction. Here the main factor for high empathy rating is the active interest that the clinician shows towards the patient. Column 3 (Table I) has been derived from Column 2 by avoiding the factors in the MITI scale that evaluate the understanding of the clinician and retains only the ones that evaluate the way the clinician responds. This is for extending the MITI scale to HRI for analyzing empathy portrayal via emotions in speech alone. The inclusion of parameters to evaluate "understanding" by the robot will require a study of the decision making ability of the robot, which is beyond the scope of this experiment. Also the role of the clinician has been replaced with that of the robot. Hence, Column 3 is a set of parameters that can be used to measure the level of perceived empathy that users feel from robots that interact with them via speech. The robotic voice and empathetic voice were randomly named as voice 1 and voice 2 (The flow of the experiment is shown in Fig. 1). The dialogs were not randomized as the they were framed as a conversation between the robot and the patient. First, the participants saw and heard the robot speaking voice 1. They were then asked to rate the voice based on the scale given in Table I Column 3. They then listened to voice 2 and rated it based on the same scale. The participants were also asked to choose their preferred voice (For *Research Question 2*). For evaluating *Research Question 3*, they were asked reasons for choosing their preferred voice. Also, they were asked which emotions they could feel when listening to each of the voices.

TABLE I
EMPATHY SCALE IN MITI AND ITS EXTENSION TO HRI

Scale	Human-Human Interaction	Human-Robot Interaction
1	Clinician has no apparent interest in client's worldview. Gives little or no attention to the client's perspective.	The robot has no interest in the patient.
2	Clinician makes sporadic efforts to explore the client's perspective. Clinicians' understanding may be inaccurate or may detract from the client's true meaning.	The robot shows some interest in what the patient is saying, but makes no efforts to engage the patient.
3	Clinician is actively trying to understand the client's perspective, with modest success.	The robot shows great interest in what the patient is saying,but makes no efforts to engage the patient.
4	Clinician shows evidence of accurate understanding of client's worldview. Makes active and repeated efforts to understand client's point of view. Understanding mostly limited to explicit content.	The robot shows great interest in what the patient is saying and tries to engage the patient, but the robot's response could be better.
5	Clinician shows evidence of deep understanding of client's point of view, not just for what has been explicitly stated but what the client means but has not yet said.	The robot shows great interest in what the patient is saying, engages the patient well and responds appropriately to the patient.

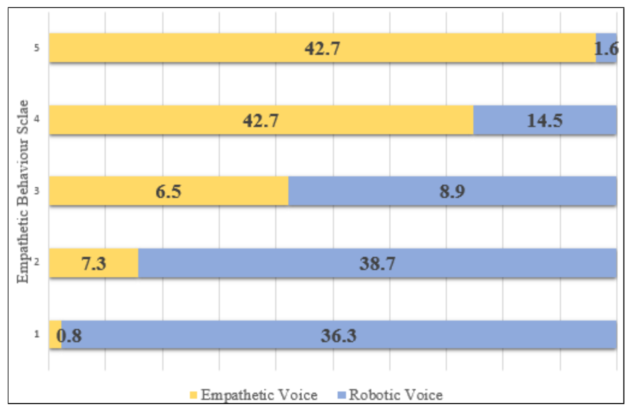

Fig. 2. Participants' rating of 2 voice types.

C. Experiment Results

Fig. 2 shows the empathetic behavior rating given by the participants to the two voice variations based on the scale given in Table I. Almost 85% of the participants felt that the robot with the empathetic voice showed great interest in the patient and tried to engage with them (they gave a score of 4 or 5). Half this group felt that the robot responded well, while the other half felt that it could do better. The reason people felt that the robot could do better may be related to people's inhibitions that a robot cannot actually "feel" the patient's situation, rather it is just being "programmed" to respond accordingly. Conversely, 75% of the participants felt that the robot with the robotic voice had little interest in the patient (given a rating of 1 or 2). Curiously, two participants (1.6%) felt that the standard robotic voice without any emotions still showed a high level of empathy. As the empathy rating scale decreases from 3 to 1, it can be seen that only less than 15% of the participants have given a lower rating for the empathetic voice. At the same time, for the robotic voice, most of the participants have given a rating of 1 or 2 on the scale. This indicates that standard robotic speech with appropriate words alone is not sufficient for people to

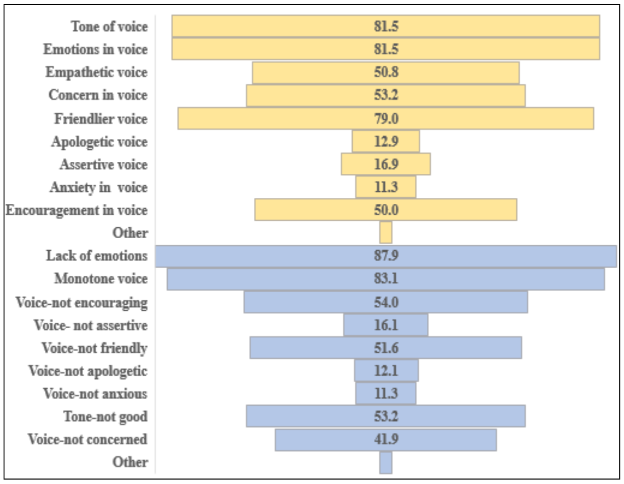

Fig. 3. Participants' reasons for choosing the empathetic voice (in blue) and not preferring the standard robotic voice (orange).

perceive an empathetic behavior from the robot.

About 95% of the participants preferred the empathetic voice over the robotic voice, which is a very strong result. The reasons for choosing the empathetic voice (blue color) and not preferring the robotic voice (orange color) are given in Fig. 3, along with the percentage of the participants who selected these reasons. The most influencing factors for preferring the empathetic voice was the tone and emotions in the the voice, closely followed by friendliness in the voice. People could also perceive empathy, concern and encouragement in the voice which also contributed to their choice. Looking at the reasons for not choosing the robotic voice, the lack of emotions and monotony in the voice are the most influencing factors, followed by lack of encouragement and concern in the voice. It is important to restate here that both the voices had the same verbal content. For example, the verbal content contained words that portrayed encouragement/concern. But people could not associate the choice of words to an active engagement from the robot (like encouragement/concern). Only when these words showing active engagement were spoken in an expressive manner with the appropriate emotions, participants could actually perceive the empathetic behavior of the robot. This is a clear indication that for developing empathetic artificial agents the interaction via speech plays a key role in influencing people's perception of robots. Even though other modalities like facial expressions are under research, speech synthesis needs to be developed to a more human-like quality. Speech communication is modeled by a proper dialog modeling along with the required emotion synthesis. From this test, it is very clear that proper dialog modeling alone is not enough. Here, the same set of dialogs were used for both variations of the voice. But participants perceived higher empathetic behavior only from the voice where the emotions matched the dialogs spoken by the robot.

IV. EMOTION ANALYSIS OF SOCIAL ROBOT

From the human perception test conducted it was seen that the addition of emotions to the verbal component

TABLE II

Healthbots DIALOG CLASSIFICATION BASED ON INTENT.

No.	Intent	Count
1	Routine Checkups	11
2	Providing Reminders	22
3	Giving Instructions	17
4	Providing Information	10
5	Greetings	13
6	Collecting Diagnosis information	20
7	Informing good/bad news	3
8	Communicating unprecedented situations	5
-	Total	101

is essential for people's perception of empathy in robotic speech. In order to synthesize speech that is empathetic, a proper modeling of emotions is essential to enhance the verbal component. But research to date has not identified the emotions associated with an empathetic voice. So, an emotion analysis of the *Healthbot* (typical example of a social robot) was done. Defining an emotional range that can be called as "empathetic" is the focus here, and also a pre-requisite for synthesizing empathetic voice. Currently, there are 101 dialogs designed for the *Healthbot* to suit the various functions that it performs (shown in Table II). From the Table, it can be seen that 13/101 dialogs are greetings to patients. A majority of the dialogs are for routine checkups and collecting information. A few of the dialogs are used to convey the results of tests or routine checkups, which can be good or bad depending on the result (3/101). From this classification, it is evident that a wide array of communication intents are portrayed in the dialogs and these intents match the usual nurse-patient interaction.

In order to understand the emotional range of the *Healthbots* an analysis of the dialogs was conducted. To achieve a realistic sense of how expressive the dialogs were, the entire set of 101 robot dialogs were spoken expressively by a female New Zealand English speaker and recorded. The speaker spoke as if in a healthcare environment, like a nurse talking to a patient. Each dialog was spoken with the emotion suitable for the context of the dialog. For example, the dialog *"Whoops, I think there is a problem with the network!"* was said apologetically. These recordings produced an expressive speech corpus based on the robot dialogs. Now the emotions in the dialogs were modelled using existing emotions models. As the results of the perception experiment clearly indicated the effect of emotions in the perception of empathy from speech, mimicking a wide array of emotions may be needed if an empathetic voice needs to be modelled. For this, a broader and diverse emotion model is used here. Russel's circumplex model (Fig. 4, [28]) depicts all emotions in the Valence-Arousal (VA) two-dimensional space. The valence level indicates the pleasantness of the voice from unpleasant (like sad) to pleasant (like happy). The arousal level specifies the level of reaction to a stimuli from inactive (like bored) to active (like surprise). The emotions in the recorded speech corpus were modelled using the Russel's circumplex model. The recordings of all the dialogs were listened to one by

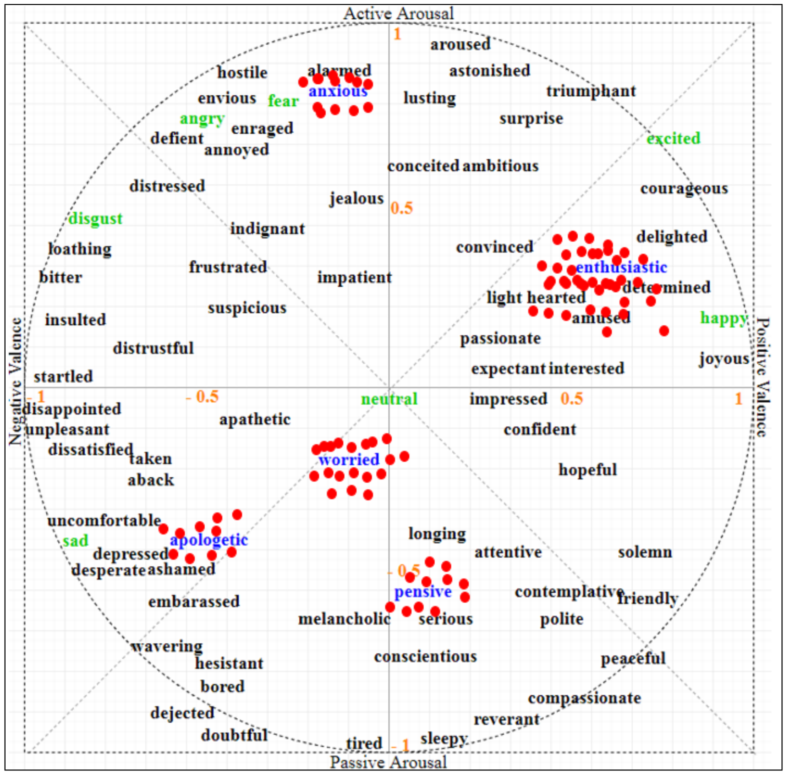

Fig. 4. Robot Dialog classification in the VA plane. Adapted from [28]. The emotions marked in blue are the chosen emotions for social robots.

one, and based on a subjective and perceptual analysis by the first author, a conclusion was drawn on the level of valence and arousal for each dialog. The VA levels are marked with values ranging between -1 and 1. For example, the dialog: *"Good job! All these exercises will help manage your COPD."* was recorded with an encouraging tone, which can be marked at a fairly positive valence level (value of 0.5 assigned) and a slightly active arousal level (value of 0.3 assigned). The point (0.5, 0.3) in the two-dimensional VA plane represents this dialog. All the dialogs were rated like this and a point corresponding to each dialog was plotted in the diagram (red dots in Fig. 4). As the *Healthbots* are a typical example of social robots, the emotional range analysis can be extended to other social robots as well.

The position of the points on the VA plane was based on the perceived valence and arousal levels associated with the rendering of each dialog. However, the labels obtained in this way are subjective and have been reported to be affected by inter-observer variations [29]. In order to ensure the repeatability of the marking process, once the levels were marked as depicted in Fig. 4, these points were assigned names of emotions that are closest to them in the Fig. 4. Based on the VA representation, the emotions nearest to these levels in the circumplex model were assigned to each robot dialog. The emotions identified were: *Enthusiastic, Apologetic, Pensive, Anxious* and *Worried*. It was observed that the dialogs that are designed for the social robot require emotions that do not fall under the category of the primary emotions (marked in green), but are rather variants of them which are the secondary emotions. Primary emotions are emotions that are innate to support reactive response behavior (Eg. angry, happy, sad, fear). Secondary emotions arise from higher cognitive processes, based on an ability to evaluate preferences over outcomes and expectations (Eg. relief, hope) [26,27]. There are various theories that define primary and secondary emotions [25], but here we will be looking at emotions that have been studied as part of HRI and speech synthesis studies. A lot of focus is given to synthesizing some well-defined primary emotions [33-38] which is used as the voice of the HRI agent to respond to stimuli. As important these primary emotions are in real life, there are nuanced secondary emotions used in human-human interaction. This study clearly shows that synthesizing these nuanced emotions are essential for an empathetic robot voice.

A. Discussion and Future Work

From Fig. 4 the overall trend of emotions in an empathetic robot's dialogs can be observed. Among the 101 dialogs, a majority of them fall under the enthusiastic emotion. This is followed by worried. There are fairly equal numbers for apologetic, pensive and anxious emotions. The number of dialogs in each category can vary depending upon the addition of functionalities to the application. From emotion analysis of the robot, the requirements for speech resources to be developed for HRI applications are:

1. Most of the emotions in the dialogs (red dots) fall in the areas of secondary emotions on the VA plane. This shows that the emotions that are required for a social robot application are not only the primary emotions like sad, happy and anger but also some secondary emotions that are variations of these. Resources to study these secondary emotions should be developed. An emotional corpus covering

the secondary emotions has been developed for New Zealand English [32] and acoustic analysis of the corpus is currently being conducted

2. The primary emotions are well apart in the VA plane. So, framing rules to synthesize them is easier. The emotions needed for social robots observed from this study are variants of the primary emotions. An inspection of the position of these emotions in the VA plane (Fig. 4) shows that they are not well separated on the VA plane as the primary emotions. This will be a major challenge when trying to model and synthesize these secondary emotions.

V. CONCLUSION

People's anthropomorphism of robots can reap positive benefits only if the communication modalities of robots are in congruence with their behaviour and appearance. This paper describes a human perception test to analyze the speech modality in social robots. Here, the same verbal content was used for a standard robotic voice (monotonous) and an empathetic voice for the robot. It was found that the empathetic behavior was perceived by participants only when the verbal content of speech was supported well by the emotional content. The participants preferred an empathetically speaking robot compared to the standard robotic voices used. Also, an analysis of the emotional range required for empathetic speech was conducted. An important finding is the dominant role of secondary emotions in empathetic speech. These emotions are not easy to model as they are not well separated on the VA plane. However, these emotions are the ones that people use to empathetically interact with each other in social situations and these emotions have to be modelled for HRI as well. This knowledge about the emotions that portray empathy in speech is a pre-requisite for synthesis of empathetic speech in social robots.

REFERENCES

[1] Broadbent E, Kerse N, Peri K, Robinson H, Jayawardena C, Kuo T, Jawalkar P, "Benefits and problems of health-care robots in aged care settings: A comparison trial",*in Australasian J on ageing*,35(1):23,2016.

[2] robotics.auckland.ac.nz/current-work/healthcare-assistive-technologies

[3] Broadbent, E., Stafford, R., MacDonald, B., "Acceptance of Healthcare Robots for the Older Population: Review and Future Directions", *Int J of Soc Robotics 1: 319*,2009

[4] P. Fung, D. Bertero, Y. Wan, et al., "Towards Empathetic Human-Robot Interactions", *arXiv preprint arXiv: 1605.04072,*2016.

[5] A. Stephan , "Empathy for Artificial Agents", *in Int J of Soc Robotics 7: 111*,2015.

[6] L. Damiano,P. Dumouchel,H. Lehmann, "Towards Human–Robot Affective Co-evolution Overcoming Oppositions in Constructing Emotions and Empathy", *in Int J of Soc Robotics 7: 7*,2015

[7] A. Lim, H. G. Okuno, "A Recipe for Empathy: Integrating the Mirror System", *in Int J of Soc Robotics 7:35–49*,2015

[8] M. Asada, "Towards Artificial Empathy: How Can Artificial Empathy Follow the Developmental Pathway of Natural Empathy?", *in Int J of Soc Robotics 7: 19*,2015

[9] S. R. Fussell, S. Kiesler, L. D Setlock, V. Yew, "How people anthropomorphize robots", *in Proc. ACM/IEEE International Conference on Human-Robot Interaction*, 2008

[10] Duffy, Brian R, "Anthropomorphism and the Social robot", *in Robotics and autonomous systems 42:3*,177-190,2003

[11] Nass C, Brave S, "Wired for speech: How voice activates and advances the human-computer relationship", in *Cambridge,MIT Press,* 2005.

[12] Eyssel, F, Kuchentrandt, D, Bobinger, S, de Ruiter, L, Hegel, F, "If you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism", *in Proc. International Conference on Human-Robot Interaction*, 2012.

[13] Scheutz M, Schermerhorn P, KramermJ, Middendorff C, "The utility of affect expression in natural language interactions in joint human-robot tasks", *in Proc. ACM International Conference on Human-Robot Interaction*,2006

[14] Goetz, J. Kiesler, S, Powers, A," Matching robot appearance and behavior to tasks to improve human-robot cooperation", *in Proc. 12th IEEE RO-MAN*, Vol. IXX,2003

[15] T. Moyers, T. Martin, J. Manuel, and W. Miller, "The motivational interviewing treatment integrity (MITI) code: Version 2.0", Unpublished.Albuquerque, NM: University of New Mexico, Center on Alcoholism, Substance Abuse and Addictions, 2008.

[16] B. Xiao, P. G. Georgiou, S. Narayanan, "Analyzing the language of therapist empathy in motivational interview based psychotherapy", *in Proc. Asia Pacific Signal Inf. Process. Assoc.*, pp. 1-4, 2012.

[17] C. Breazeal,B. Scassellati, "A context-dependent attention system for a social robot", *in Proc. 16th Int'l Joint Conf. Artificial Intelligence*,vol. 2 pp. 1146-1151 1999.

[18] A. Chella, R.E. Barone, G. Pilato, R. Sorbello, "An emotional storyteller robot", *in Proc. AAAI Spring Symposium: Emotion, Personality, and Social Behavior*,2006

[19] S. Roehling, B. MacDonald, C. Watson, "Towards expressive speech synthesis in English on a robotic platform", *in Proc. Australasian International Conference on Speech Science and Technology*, 2006

[20] Mavridis, N, "A review of verbal and non-verbal human–robot interactive communication", *Robotics and Autonomous Systems*,22-35,2015

[21] Niculescu, A., van Dijk, B., Nijholt, A. et al., "Making social robots more attractive: the effects of voice pitch, humor and empathy", *in Int J of Soc Robotics 5.2*,171-191 , 2013

[22] Cuff Benjamin MP, "Empathy: a review of the concept", *Emotion Review 8.2*,144-153, 2016

[23] A. Igic, C. Watson, R. Stafford, "Perception of synthetic speech with emotion modeling delivered through a robot platform: an initial investigation with older listeners", *in Proc. SST*,2010

[24] Watson, C., Liu, W., MacDonald, B. , "The Effect of Age and Native Speaker Status on Intelligibility", *in Proc. ISCA Speech Synthesis Workshop*, 2013

[25] Kemper T.D, "How Many Emotions Are There? Wedding the Social and the Autonomic Components", *American J of Sociology,* Volume 93, Number 2,Sep., 1987

[26] Damasio, A., "Descartes' error, emotion reason and the human brain" ,Grosset/Putnam, 1994.

[27] Becker-Asano, C., and Wachsmuth, I., "Affective computing with primary and secondary emotions in a virtual human", *Autonomous Agents and Multi-Agent Systems*, 2010

[28] G. Paltoglou and M. Thelwall, "Seeing Stars of Valence and Arousal", *IEEE Transactions of Affective Computing* , 2013.

[29] AA. Rui, O. Martinez, X. Binefa, and F. Sukno, "Fusion of Valence and Arousal Annotations through Dynamic Subjective Ordinal Modelling", *in Proc. IEEE FG*, 2017

[30] Alam, F.; Danieli, M., Riccardi, G., "Annotating and modeling empathy in spoken conversations",*Computer Speech Language 50* , 2018

[31] P.Taylor, "Text-to-Speech Synthesis", *Chapter 13, Cambridge*, 2009

[32] J. James, L. Tian and C. Watson, "An Open Source Emotional Speech Corpus for Human Robot Interaction Applications", *in Proc. Interspeech*, 2018.

[33] Ochs, M., Sadek, D. Pelachaud, "A formal model of emotions for an empathic rational dialog agent" , *C. Auton Agent Multi-Agent Syst*, 2012

[34] H. Boukricha, I. Wachsmuth, M. N. Carminati, P. Knoeferle, "A computational model of empathy: Empirical evaluation", *Proc. ACII, pp. 1-6,* 2013.

[35] M. Schröder, "Emotional Speech Synthesis: A Review", *in Proc. Eurospeech,* 2001

[36] C. Breazeal, "Emotive qualities in robot speech", *in Proc IEEE/RSJ International Conference on Intelligent Robots and Systems,* 2001.

[37] Li X, Watson CI, Igic A, Macdonald BA, Expressive speech for a virtual talking head, *in Proc. Australasian Conference on Robotics and Automation,* 2009

[38] Crumpton, J. Bethel, "A Survey of Using Vocal Prosody to Convey Emotion in Robot Speech", *Int J of Soc Robotics 8: 271*, 2016

[39] Crumpton, J. Bethel, "Validation of vocal prosody modifications to communicate emotion in robot speech", in Proc CTS), 2015