Trust in artificial voices: A "congruency effect" of first impressions and behavioural experience

Ilaria Torre
Trinity College Dublin
Electronic & Electrical Engineering
Dublin, Ireland
torrei@tcd.ie

Jeremy Goslin
Plymouth University
School of Psychology
Plymouth, U.K.
jeremy.goslin@plymouth.ac.uk

Laurence White
Plymouth University
School of Psychology
Plymouth, U.K.
laurence.white@plymouth.ac.uk

Debora Zanatto
Plymouth University
School of Computing, Electronics and Mathematics
Plymouth, U.K.
debora.zanatto@plymouth.ac.uk

ABSTRACT

Societies rely on trustworthy communication in order to function, and the need for trust clearly extends to human-machine communication. Therefore, it is essential to design machines to elicit trust, so as to make interactions with them acceptable and successful. However, while there is a substantial literature on first impressions of trustworthiness based on various characteristics, including voice, not much is known about the trust development process. Are first impressions maintained over time? Or are they influenced by the experience of an agent's behaviour? We addressed these questions in three experiments using the "iterated investment game", a methodology derived from game theory that allows implicit measures of trust to be collected over time. Participants played the game with various agents having different voices: in the first experiment, participants played with a computer agent that had either a Standard Southern British English accent or a Liverpool accent; in the second experiment, they played with a computer agent that had either an SSBE or a Birmingham accent; in the third experiment, they played with a robot that had either a natural or a synthetic voice. All these agents behaved either trustworthily or untrustworthily. In all three experiments, participants trusted the agent with one voice more when it was trustworthy, and the agent with the other voice more when it was untrustworthy. This suggests that participants might change their trusting behaviour based on the congruency of the agent's behaviour with the participant's first impression. Implications for human-machine interaction design are discussed.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**;
• **Applied computing → Psychology**;

KEYWORDS

Voice, Human-Machine Interaction, trust, behavioural measures

ACM Reference Format:
Ilaria Torre, Jeremy Goslin, Laurence White, and Debora Zanatto. 2018. Trust in artificial voices: A "congruency effect" of first impressions and behavioural experience. In *Proceedings of APAScience '18: Technology, Mind, and Society (TechMindSociety '18).* ACM, New York, NY, USA, 6 pages. https://doi.org/https://doi.org/10.1145/3183654.3183691

1 INTRODUCTION

Trust is a fundamental aspect of social interactions and human communication [e.g., 3]; without trust, neither one would function. Since the spoken channel is our main mode of communication, it seems plausible that we have developed means to signal our trustworthiness, and to detect someone else's, on the basis of voice. Indeed, evidence shows that people are able to make fast and accurate trustworthiness judgments upon hearing a voice for the first time [17]. There is a substantial literature on explicit evaluations of different voices. For example, trustworthiness can be inferred from vocal features, such as accent [16], prosody [18], or emotional expressions [24]. However, often the emphasis of such studies has been on certain indexical information that someone's voice conveys (for example, gender or place of origin), but rarely has the effect of vocal characteristics on trust been studied *per se*. The information conveyed in the voice might be a good measure of *person* perception, where a person's voice interacts with all sorts of other features, such as physical appearance or attire. However, not all interactions which require trust happen face-to-face, and this phenomenon can only increase with the advent of powerful and inexpensive voice-based technologies [see 22]. Also, speaking machines are becoming a reality for an increasing number of people, and researchers have been also studying *agent* perception [e.g., 20]. As technology advances, studying the effect that different agent characteristics have on human decision-making is paramount to ensure a smooth transition to a world where humans and machines cooperate in society.

While it is clear that vocal cues influence trustworthiness attributions, how such initial impressions will change upon repeated interaction with this voice, and with the behaviour of its owner,

TechMindSociety '18, April 5–7, 2018, Washington, DC, USA

remains unclear. How do initial judgments change with exposure to a speaker's behaviour? While first impressions, such as those that questionnaires typically measure, provide examples of someone's conscious perception at a certain point in time, they do not reveal anything about any subsequent attitude development [2]. A study looking at the trustworthiness of Embodied Conversational Agents partially addressed the question of trustworthiness attributions over time: [11] collected trustworthiness ratings at different points during a staged interview with an agent. However, the agent randomly changed its appearance (female or male, smiling or neutral) after each rating, so it is possible that participants simply rated what they perceived to be a new agent from scratch, rather than the same one over time. Another open question is: are some voices more trustworthy in certain contexts, and other voices in other contexts? For example, [1] showed that the accent of a robot tour guide affected participants' responses differently in different contexts: participants believed that robots with an Arabic regional accent were more credible when they showed they were knowledgeable, whereas robots with an Arabic standard accent were more credible when they had little knowledge. This suggests that the "behavioural context" of an interaction might influence its outcome as well. However, little research has been dedicated to experimentally manipulating the behavioural characteristics of an agent in order to examine their effect on the trust development process. Research in Human-Machine Interaction aimed at automatically predicting personality traits from multimodal cues partially addressed this issue. [10] and [13] asked participants to judge pre-recorded conversations between a human and an avatar which exhibited different emotional and personality characteristics, like sadness or happiness. They found that this situational context affected people's perception and ratings, although the authors did not mention exactly how the judgments changed over time. [4] also specifically manipulated an agent's behaviour: participants played a map task with a virtual agent which was programmed to have different levels of cooperativeness, in order to elicit different reactions from the participants. Surprisingly, they found that the different cooperativeness levels did not elicit different personality attributions, nor did they help to better predict them. Thus, the effect of an agent's behavioural context on human-agent interactions remains unclear.

As previously mentioned, societies need trust to function, so human-machine societies will need trust to function as well. In particular, what is needed is a way of inferring a machine's trustworthiness, in order for humans to accept interacting with them in the first place. There is evidence that people unconsciously apply to human-machine interaction the same social rules that they apply in human-human interactions [14, 15, 21], and that they attribute personality traits to machines as well. For example, [20] showed that even approximate representations of personality traits in computer agents were recognised, and acted accordingly upon, by participants; the same holds for personality cues represented in a synthetic computer voice [19]. Several studies have also shown that personality traits are ascribed to robots [e.g., 23, 26] and even navigator systems [14] as well. However, while we might perceive other people as trustworthy by default [e.g., 12], the same cannot be said for agents, especially in high-risk contexts, where a misplaced trust in a robot could endanger a human's safety. Thus, it is up to the manufacturers to make sure that the machines are safe to work with, but it is also up to the designers to design machines that inspire trust. Since many of the machines that collaborate with us are disembodied (for example, navigator systems or mobile personal assistants), their voice should also be designed as to elicit trust. In the case of robots, which do have a body, the voice design should be an integrative part of the body design. For example, we might expect small companion robots to have high-pitched voices, compatible with a small vocal tract (what we may call a "physical" voice context); similarly, robots working in children's learning environments could have children-like voices to facilitate the children's acceptance of the robot as their peer (a "work" voice context). But how about a "behavioural" voice context? For example, if a navigator system gave us wrong directions, would we be more or less annoyed at it if it had a trustworthy-sounding, or untrustworthy-sounding voice?

In this paper, we address the issue of trust development through behavioural experience by presenting results from several studies investigating the effect of voice on the perception of an agent's trustworthiness. As we shall see, our results indeed suggest that initial trustworthiness attributions based on a voice develop in different ways over time, according to the behavioural context in which the interaction takes place.

2 METHOD

All the experiments examined trust attributions to different voices either in computerised agents or humanoid robots. The experimental procedure was always the same, an "iterated investment game". This methodology, which stems from game theory, was originally devised by [6], and it has been extensively used to study the effect of various characteristics (such as gender, facial attractiveness, attire) on implicit trust attributions [e.g., 9].

2.1 The investment game

In our version of the investment game, a human participant receives a monetary endorsement, and can decide to invest some, all or none of it with player B – which in our case is either a computer agent or a robot. The experimenter triples the amount that player B receives, so that player B can return to the participant any amount between 0 and the tripled investment. Thus, the invested amount is a measure of trust, because if player B is trustworthy, both the participant and player B will end up with more money than they had at the beginning of the game round. Furthermore, we programmed player B to be either trustworthy or untrustworthy – always returning to the participant more money than they invested (*generous* condition) or less (*mean* condition). The behaviour of player B was a manipulation that was consistent across all experiments. The other manipulations, described more in detail below, concerned the voice of player B.

2.1.1 *Experiment 1.*

In the first experiment, we looked at computer agents which had either a prestigious accent, previously associated with high trust attributions – Standard Southern British English (SSBE) – or a regional accent, previously associated with low trust attributions – Liverpool English [7]. We recorded two male native speakers of these accents, whose voices were used for the virtual agents used in the game. From each speaker, we recorded one different sentence for each round of the game, so that participants would hear one sentence before making each investment decision. The sentences were all about strategies to follow in the game, apart from the first one, which served as an introduction ("Hello, let's get started with the investment game"). The accent of the speakers was crossed with the behavioural manipulation (*generous* vs. *mean* condition), in a 2 by 2 counterbalanced design. Each participant played 2 games of 20 rounds each, each with one of the 4 possible accent-behaviour pairs, and was given £8 at the beginning of each round of the game. The games were played on a computer, where the participants heard the initial sentences via good quality headphones and indicated the amount of money they wished to invest by pressing the corresponding key on the keyboard. 44 native British English speakers participated in this experiment. The full procedure and analyses of this experiment are described in [27].

2.1.2 Experiment 2. In the second experiment, the computer agents had either an SSBE accent or a Birmingham accent, the latter having also been previously associated with low trust attributions [7]. This time, we recorded two female speakers of each accent. The accent and behaviour or the virtual agents were counterbalanced within participants, while the speaker identity was counterbalanced between participants. Participants played 2 games of 20 rounds each, with an £8 fund available at the beginning of each round. The procedure was exactly the same as in the first experiment. 110 native British English speakers participated, although the data of 2 had to be excluded from the analyses due to technical errors. The full procedure and analyses of this experiment will be described in a manuscript in preparation.

2.1.3 Experiment 3. In the third experiment, we looked at trusting behaviours towards a humanoid robot having either a natural voice or a synthetic voice. The natural voice was recorded from two female SSBE speakers. This voice was then resynthesised in order to obtain a "robotic" voice. Firstly, the mean f0 value for each utterance was extracted using the software for phonetic analyses Praat [8]. The mean f0 value for each speaker was then obtained by averaging the value of all the utterances from the same speaker. Then, the f0 of each utterance was flattened to this mean f0 value. Finally, a comb filter was applied to each flattened utterance using the software Audacity. The 4 voices thus obtained were played on a Nao robot. The participants played with the robot through a Sandtray touchscreen computer [5], and they were endorsed with 10 "Experimental Currency Units" (ECU) at the beginning of each of the 20 rounds of the game (Figure 1). The robot was also manipulated to either display joint attention (following the participant's face and making deictic arm movements) or not, in a 2 (attention: joint or non-joint) by 2 (behaviour: *generous* or *mean*) within-subject design and a 2 (voice: natural or synthetic) by 2 (speaker identity: speaker 1 or 2) between-subject design. While the full procedure and analyses of this experiment are described in a manuscript in preparation, in this paper we only report a subset of the results on the voice condition. 120 native British English speakers participated in this experiment, playing 2 games each.

Figure 1: Setup of Experiment 3. Image credits: ©Ilaria Torre

3 RESULTS

The data from each experiment were analysed fitting mixed-effects linear models, using participants' investment as dependent variable, the various agent manipulations (behaviour, voice and game round) as independent variables and participant id and sentence id as random effects. While the full analyses of these experiments goes beyond the scope of this paper, and is in preparation for publication elsewhere, here we will focus on one result in particular that is common to these experiments, a "congruency effect" between participants' first impressions and actual behavioural experience.

3.1 Experiment 1

The mixed-effects linear model revealed a main effect of behaviour ($\chi^2(1) = 984.25, p < .001, N = 44$): as expected, participants invested more money with the *generous* virtual player (average = £5.8) than the *mean* virtual player (average = £2.6). Interestingly, there was also a significant interaction between accent and behaviour ($\chi^2(1) = 7.36, p = .007, N = 44$): in the *generous* condition, participants invested more money with the SSBE-accented virtual player (average = £6.34) than the Liverpool-accented one (average = £5.30), while in the *mean* condition, participants invested more with the Liverpool-accented virtual player (average = £3.06) than the SSBE-accented one (average = £2.14), as can be seen from Figure 2.

3.2 Experiment 2

There was a main effect of behaviour, with an average investment of £5.63 to the *generous* virtual player and of £3.31 to the *mean* virtual player ($\chi^2(1) = 1135, p < .001, N = 108$). There was also a significant interaction between accent and behaviour ($\chi^2(5) = 34.91, p < .001, N = 108$): as can be seen from Figure 3, in the *generous* condition participants invested more money with the SSBE-accented player (average = £5.81) than the Birmingham-accented one (average = £5.46), while in the *mean* condition participants invested more money with the Birmingham-accented player (average = £3.56) than the SSBE-accented one (average = £3.05).

3.3 Experiment 3

We found a main effect of behaviour also in the third experiment ($\chi^2(1) = 90.17, p < .001, N = 120$), with participants investing more money in the *generous* robot (average = 8.13 ECU) than the *mean* robot (average = 4.5 ECU). There was also an interaction between behaviour, game turn and voice ($\chi^2(2) = 8, p <$

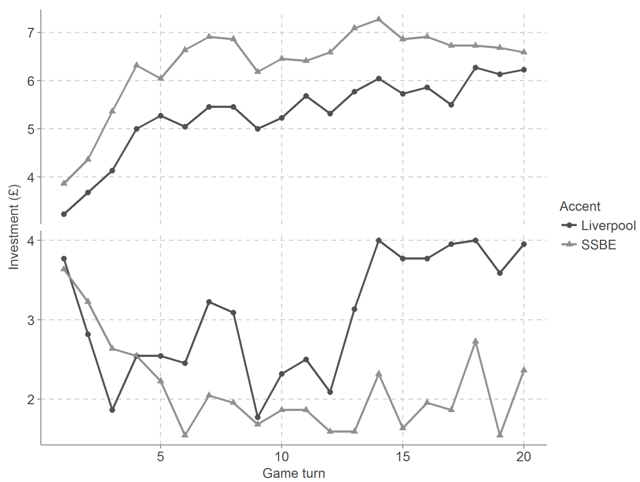

Figure 2: Average investments to the SSBE and Liverpool speakers in the *generous* (top) and *mean* (bottom) conditions in Experiment 1.

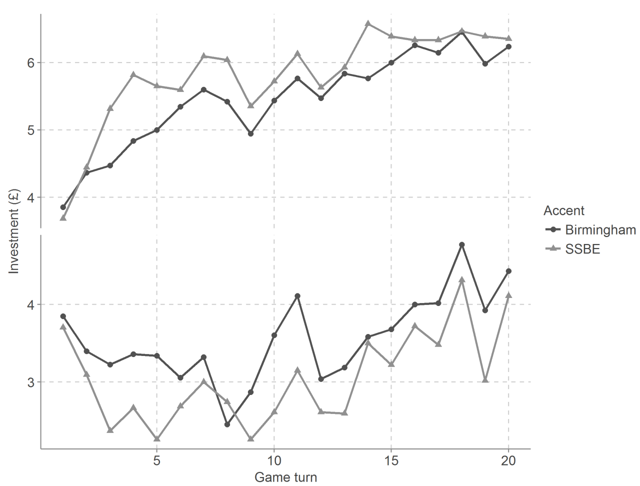

Figure 3: Average investments to the SSBE and Birmingham speakers in the *generous* (top) and *mean* (bottom) conditions in Experiment 2.

.02, $N = 120$). A post-hoc mixed-effects model was fitted to the *generous* and *mean* data separately to analyse this interaction. Investment was the dependent variable, game turn and voice were predictors, and participant and sentence id were random factors. In the *generous* condition, there was a main effect of game turn ($\chi^2(1) = 256.5, p < .001, N = 120$) and a marginally significant effect of voice ($\chi^2(1) = 3.32, p = .068, N = 120$). As can be seen from the top panel in Figure 4, investments increase over time, and are higher in the synthetic voice condition. In the *mean* condition, instead, there is no effect of game turn, and no effect of voice, but there is a significant interaction between game turn and voice ($\chi^2(2) = 16.33, p < .001, N = 120$). As can be seen from the bottom panel of Figure 4, investments decrease in the first half of the game

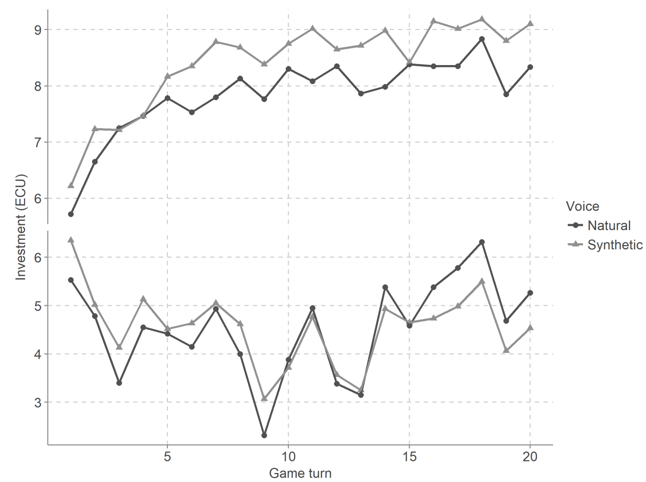

Figure 4: Average investments to the natural and synthetic voices in the *generous* (top) and *mean* (bottom) conditions in Experiment 3.

independently from the voice, while they increase in the second half of the game, and there are higher investments to the natural voice.

4 DISCUSSION

The pre-programmed behaviours of player B – which was either a computer agent or a robot – in the investment game influenced participants' trusting behaviours. Specifically, in Experiment 1 participants trusted the SSBE-accented virtual player more in the *generous* condition, and the Liverpool-accented virtual player more in the *mean* condition (see Figure 2). In Experiment 2, participants trusted the SSBE-accented virtual player more in the *generous* condition, and the Birmingham-accented virtual player more in the *mean* condition (Figure 3). In Experiment 3, participants trusted a robot with a synthetic voice more in the *generous* condition, and a robot with a natural voice more in the *mean* condition (Figure 4). These results suggest that participants might form an impression of an agent's trustworthiness upon hearing them for the first time – which in the investment game is simulated by the first round of the game. The actual trustworthiness of the speaker, in the form of the simulated *generous* and *mean* behaviours, then contributes to the further shaping of this impression, which is different based on whether the perceived and actual trustworthiness agree or not. For example, participants in Experiment 1 might have attributed relative trustworthiness to the SSBE voice when they heard it for the first time. In the *generous* behaviour, this impression of trustworthiness is confirmed, so their trust in this speaker is maintained and reinforced over time. In the *mean* condition, instead, the impression of trustworthiness is disconfirmed, so participants "punish" this incongruity by investing less money in this speaker. On the other hand, participants might have initially attributed a lower trustworthiness to the Liverpool voice, upon hearing it for the first time. In the *generous* condition, the actual behaviour of the speaker is at odds with first impressions, so participants might still be suspicious of this incongruent behaviour, resulting in lower investments. Instead, in the *mean* condition participants' implicit first impressions are confirmed, so their reaction to untrustworthy behaviour is less negative, because the Liverpool-accented computer agent is behaving congruently with participants' first impressions and expectations. Similarly, in Experiment 2, participants trusted the SSBE-accented computer agent more in the *generous* condition, and the Birmingham-accented virtual player more in the *mean* condition. Finally, this "congruency effect" seems to apply to voices embodied in robots as well. In Experiment 3, participants trusted a robot with a synthetic voice more in the *generous* condition, and a robot with a natural voice more in the *mean* condition, although the latter did not manifest itself from the very beginning.

This "congruency effect" corroborates [1]'s findings on perceived robot credibility based on its accent and context, and resembles results from [25]. In their study on face trustworthiness, participants played an iterated economic game in which they had to decide whether to borrow money from lenders with trustworthy or untrustworthy-looking faces. The lenders charged either no, moderate or high interests on the debt. In a post-game memory test, participants remembered the trustworthy-looking, high-interest lenders better than all the other categories, and the authors concluded that people have better memory for "wolves in a sheep's clothing than for the wolf in a wolf's clothing". This suggests that people might be particularly attentive to deceitful signals (such as appearing trustworthy while behaving untrustworthily). This coincides with results from the investment games, specifically that speaking agents that elicited particularly strong first impressions of trustworthiness might be "punished" when their behaviour is untrustworthy, and trusted even less than the speaking agents that were expected to be untrustworthy.

5 CONCLUSION

Referring back to the hypothetical navigator system example mentioned in the introduction, the three experiments in this study add empirical grounds to the argument. The different levels of annoyance in our hypothetical reaction to the navigator's mistake could be explained by a "congruency effect": being deceived by an intrinsically trustworthy-sounding voice is unexpected and results in higher annoyance. On the other hand, an untrustworthy-sounding navigator system, which does not sound trustworthy to begin with, does not elicit as much a negative reaction when its information turns out to be wrong.

The research presented here suggests that even vocal characteristics alone can influence trusting behaviours, and that these behaviours are mediated by the speaking agent's actual trustworthiness. In particular, people trust "congruent" voice-behaviour pairs more than "incongruent" ones. This has applications in the field of Human-Machine Interaction, for example in the case of machines that work within a certain error margin. (Speaking) machines do not always work properly, for example in tasks such as speech or speaker recognition. The current results suggest that people implicitly trust machines speaking with a certain voice more when they are cooperative, and machines speaking with another voice more when they are non-cooperative. Designers should take this into account when creating voices for different machines. Furthermore, machines performing different tasks should have voices which sound trustworthy relative to that particular task. Thus, while the "work" and "physical" contexts are important for designing an appropriate and trustworthy machine voice, the results from the current experiments suggest that a "behavioural" context should be taken into account as well.

ACKNOWLEDGMENTS

This work was supported by CogNovo FP7-PEOPLE-2013-ITN-604764 (https://CogNovo.eu), a project funded by Marie Skłodowska-Curie Actions. The first author received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 713567 (http://edge-research.eu/), and from the ADAPT Centre for Digital Content Technology, which is funded under the SFI Research Centres Programme (Grant 13/RC/2016) and is co-funded by the European Regional Development Fund.

REFERENCES

[1] Sean Andrist, Micheline Ziadee, Halim Boukaram, Bilge Mutlu, and Majd Sakr. 2015. Effects of Culture on the Credibility of Robot Speech. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI '15*. ACM, ACM Press, 157–164. https://doi.org/10.1145/2696454.2696464

[2] Solomon E. Asch. 1946. Forming impressions of personality. *The Journal of Abnormal and Social Psychology* 41, 3 (1946), 258–290. https://doi.org/10.1037/h0055756

[3] Annette Baier. 1986. Trust and Antitrust. 96, 2 (1986), 231–260. https://doi.org/10.1086/292745

[4] Ligia Batrinca, Bruno Lepri, Nadia Mana, and Fabio Pianesi. 2012. Multimodal recognition of personality traits in human-computer collaborative tasks. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction*. ACM, Santa Monica, CA, USA, 39–46.

[5] Paul Baxter, Rachel Wood, and Tony Belpaeme. 2012. A touchscreen-based Sandtray to facilitate, mediate and contextualise human-robot social interaction. In *Proceedings of the 7th ACM/IEEE International Conference on Human-Robot Interaction*. IEEE, 105–106.

[6] Joyce Berg, John Dickhaut, and Kevin McCabe. 1995. Trust, reciprocity, and social history. *Games and economic behavior* 10, 1 (1995), 122–142.

[7] Hywel Bishop, Nikolas Coupland, and Peter Garrett. 2005. Conceptual accent evaluation: Thirty years of accent prejudice in the UK. 37, 1 (2005), 131–154. https://doi.org/10.1080/03740463.2005.10416087

[8] Paul Boersma and David Weenink. 2017. Praat: doing phonetics by computer (Version 6.0.27). (2017). http://www.praat.org

[9] Nancy R. Buchan, Rachel T. A. Croson, and Sara Solnick. 2008. Trust and gender: An examination of behavior and beliefs in the Investment Game. *Journal of Economic Behavior & Organization* 68, 3 (2008), 466–476.

[10] Oya Celiktutan and Hatice Gunes. 2016. Automatic Prediction of Impressions in Time and across Varying Context: Personality, Attractiveness and Likeability. *IEEE Transactions on Affective Computing* PP, 99 (2016), 1–1. https://doi.org/10.1109/TAFFC.2015.2513401

[11] Aaron C. Elkins and Douglas C. Derrick. 2013. The sound of trust: voice as a measurement of trust during interactions with embodied conversational agents. *Group Decision and Negotiation* 22, 5 (2013), 897–913.

[12] Gareth R. Jones and Jennifer M. George. 1998. The experience and evolution of trust: Implications for cooperation and teamwork. *Academy of Management Review* 23, 3 (1998), 531–546.

[13] Jyoti Joshi, Hatice Gunes, and Roland Goecke. 2014. Automatic prediction of perceived traits using visual cues under varied situational context. In *Proceedings of the 22nd International Conference on Pattern Recognition (ICPR)*. IEEE, Stockholm, Sweden, 2855–2860.

[14] David R. Large, Leigh Clark, Annie Quandt, Gary E. Burnett, and Lee Skrypchuk. 2017. Steering the conversation: a linguistic exploration of natural language interactions with a digital assistant during simulated driving. *Applied ergonomics* 63 (2017), 53–61.

[15] Kwan Min Lee, Wei Peng, Seung-A Jin, and Chang Yan. 2006. Can robots manifest personality? An empirical test of personality recognition, social responses, and social presence in human–robot interaction. *Journal of communication* 56, 4 (2006), 754–772.

[16] Shiri Lev-Ari and Boaz Keysar. 2010. Why don't we believe non-native speakers? The influence of accent on credibility. 46, 6 (2010), 1093–1096. https://doi.org/10.1016/j.jesp.2010.05.025

[17] Phil McAleer, Alexander Todorov, and Pascal Belin. 2014. How do you say Hello? Personality impressions from brief novel voices. *PLoS ONE* 9, 3 (2014), e90779.

[18] Norman Miller, Geoffrey Maruyama, Rex J. Beaber, and Keith Valone. 1976. Speed of speech and persuasion. *Journal of Personality and Social Psychology* 34, 4 (1976), 615.

[19] Clifford I. Nass and Kwan Min Lee. 2001. Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology: Applied* 7, 3 (2001), 171–181. https://doi.org/10.1037//1076-898x.7.3.171

[20] Clifford I. Nass, Youngme Moon, B. J. Fogg, Byron Reeves, and Chris Dryer. 1995. Can computer personalities be human personalities? *International Journal of Human-Computer Studies* 43, 2 (1995), 223–239.

[21] Clifford I. Nass, Jonathan Steuer, and Ellen R. Tauber. 1994. Computers are social actors. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 72–78.

[22] René Riedl and Andrija Javor. 2012. The biology of trust: Integrating evidence from genetics, endocrinology, and functional brain imaging. *Journal of Neuroscience, Psychology, and Economics* 5, 2 (2012), 63.

[23] Maha Salem and Kerstin Dautenhahn. 2015. Evaluating trust and safety in HRI: Practical issues and ethical challenges. In *Emerging Policy and Ethics of Human-Robot Interaction: A Workshop at 10th ACM/IEEE International Conference on Human-Robot Interaction*. ACM Press.

[24] Joanna Schug, David Matsumoto, Yutaka Horita, Toshio Yamagishi, and Kemberlee Bonnet. 2010-03. Emotional expressivity as a signal of cooperation. 31, 2 (2010-03), 87–94. https://doi.org/10.1016/j.evolhumbehav.2009.09.006

[25] Atsunobu Suzuki and Sayaka Suga. 2010. Enhanced memory for the wolf in sheep’s clothing: Facial trustworthiness modulates face-trait associative memory. 117, 2 (2010), 224–229. https://doi.org/10.1016/j.cognition.2010.08.004

[26] Dag Sverre Syrdal, Kerstin Dautenhahn, Sarah N. Woods, Michael L. Walters, and Kheng Lee Koay. 2007. Looking Good? Appearance Preferences and Robot Personality Inferences at Zero Acquaintance.. In *AAAI Spring Symposium: Multidisciplinary Collaboration for Socially Assistive Robotics*. 86–92.

[27] Ilaria Torre, Jeremy Goslin, and Laurence White. 2015. Investing in accents: How does experience mediate trust attributions to different voices?. In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS 2015)*. Glasgow, U. K.