THE VOICE OF CREATIVITY: EFFECTS OF PITCH RANGE IN THE VOICE OF A ROBOT FACILITATOR

Kerstin Fischer[1], Oliver Niebuhr[2], Ali Asadi[1]

[1]Department of Design and Communication, [2] Centre for Industrial Electronics
University of Southern Denmark, Sonderborg
[kerstin,olni,aa]@sdu.dk

Abstract: Previous work has identified a considerable role of the facilitator of creativity workshops; in this study, we investigate to what extent the observable differences between facilitators can be attributed to facilitators' speaking style. Specifically, we study the effects of pitch range. That pitch range may have an influence on listeners' creativity can be expected from the fact that it is associated with higher engagement. In order to eliminate effects of speaker personality, we use a robotic facilitator. In an online questionnaire, videos of a robot provided the participants with the necessary instructions on a creativity task, where the pre-synthesized speech of the robot was manipulated globally concerning pitch range. Dependent variables were the originality, fluency, flexibility and degree of elaboration of participants' ideas, as well as their subjective ratings of the robot facilitator. The results show significantly higher facilitator ratings and creativity-task performances in connection with the robot that used the larger pitch range. We discuss our results in terms practical applications and further relevant voice features.

1 Introduction

Previous work has identified a considerable role of the facilitator of creativity workshops, yet usually the observable differences between facilitators are attributed to aspects of confidence, team leadership and process management skills as well as of knowledge of creativity tools, rather than speaking style, see, for example, [1]. At the same time, previous work has documented that speaking style has an effect on persuasion [2-3] and on student performance [4-6], and thus it is possible that speaking style, and not (just) the traits mentioned above, influences the extent to which people perform in creativity tasks. In order to eliminate effects of the facilitator's personality, we use a robotic facilitator, which ensures that participants are exposed to identical stimuli, with the exception of the robot's speaking style, the independent variable under consideration.

That speaking style is likely to have an influence on participants' creativity was found by Fucinato et al. [7], yet that study used several different prosodic manipulations at the same time. Specifically, they investigated the effects of a voice that had a larger pitch range, more accented words and a lower center of gravity than the other voice. While the study's results suggest that style has an effect on participants' creativity, it does not allow us to identify what exactly the effect is due to. In the current study, we therefore concentrate on only a single prosodic feature. That pitch range may have an influence on listeners' creativity can be expected from the fact that it is associated with higher engagement [6]. Our hypothesis is therefore that an increased pitch range has a significant influence on the number and originality of participants' ideas.

2 Previous Work

Previous work concerns findings on the role of characteristics of the respective facilitator on participants' performance in creativity tasks; prior work on the effects on pitch range on listeners' performance; and prior studies on robots facilitating creativity.

2.1 The Role of Facilitator Characteristics in Creativity Tasks

That creativity workshop facilitators may vary in the extent to which they encourage workshop participants to be creative has already been addressed by Wallgren [1], who interviewed thirty-seven facilitators with respect to their own perceptions of success criteria for creative problem solving. The characteristics of a good facilitator suggested include leadership skills, the ability to include the right people and to involve them in the right way, the knowledge of tools to lead the dynamic process and of tools to encourage creativity. That is, the skills characteristic of a good creative problem-solving facilitator concern group leadership, process management and knowledge of a large range of supporting tools and methods, as well as some people skills; speech characteristics are not mentioned. However, more recently, Stewart [8, p.436] identified very similar facilitator characteristics, yet including "very clear verbal communication with good use of words, language and tone" as an important facilitator characteristic. While she does not get more specific than this, it is unlikely that she actually refers to speech characteristics.

Johnsson et al. [9] suggest that a facilitator should have experience in practical innovation work in accordance with the uncertain innovation process, be socially competent, opportunity driven, and flexible; and be able to establish trust. Another feature was put forth by Wróbel et al. [10], who argue that neutrality is central to creativity facilitation, and that impartiality, equidistance and fairness are three important dimensions of neutrality and apply to the facilitator's attitude towards the people, the process and the product alike.

Johnsson [11] carried out a long-term ethnographic study of the practices in three design teams, including interviews. He finds that the relevance of certain facilitator characteristics differs depending on the phase in the innovation process. However, among the 40 characteristics he identified, none concerns the facilitators' speech characteristics.

2.2 The Effects of Pitch Range on Listener Performance

Pitch range is a prosodic feature which has been associated with charismatic speech (e.g. [2]), and it is one of the prosodic features that were identified as core characteristics of charismatic speech, for example, by Strangert & Gustafson [12] and Niebuhr & Skarnitzl [13].

In experimental studies using robots, pitch range was one of the speech characteristics used by Fischer et al. [3] to make robots more persuasive and by Fischer et al. [5] to influence learner performance. In particular, Fischer et al. [3] found that if a robot uses charismatic speech, which comprises higher pitch range as well as a raised average pitch level, pause duration, number of high-pitched and emphatic accents, number of hesitations, as well as differences in tempo and energy level (cf. [14-15]), people followed the robot's suggestions more. Fischer et al. [5] used the same kinds of manipulations to employ robots with a charismatic speaking style when providing the instructions to a prosodic production task; the results of this study show that language learners produce significantly more accurate utterances when previously instructed by a robot that uses charismatic speech, which comprises an increased pitch range. However, these studies used a broad range of prosodic manipulations, and thus the effects cannot be attributed to the increased pitch range alone.

2.3 Robots as Creativity Facilitators

Lubart et al. [16] argue that robots are not only generally well suited as creativity workshop facilitators, but also much more suited than computers, due to the fact that robots are embodied and share the context with the creativity team. Correspondingly, Kahn et al. [17] found that a robot creativity facilitator is more effective than the same encouragement delivered by an interactive PowerPoint presentation.

Hu et al. [18] and Alves-Olivera et al. [19] studied the effects of social behaviors by robots when serving as creativity facilitators, and both studies report an increased effect of the degree of sociality displayed by the robot on the extent to which people are creative. Ali et al. [20] study the effects of a robot facilitator across three different creativity tasks and show that a robot that demonstrates creativity itself is more effective in eliciting creativity in children than a robot that only tries to elicit creativity from the children. Thus, the robot facilitator's behavior has an impact on the extent to which children are creative.

In our own work [7], we had a robot lead teams of students through a creativity task, in one condition with a charismatic voice and in the other with a not so charismatic voice. The results show a clear effect of the robot's speech characteristics such that the more charismatic robot elicits more original, more flexible and more elaborate ideas. However, since several different speech manipulations were carried out, we do not know which speech characteristics the effects found are due to.

To conclude, robots may be principally suited to function as creativity facilitators, yet little work has addressed facilitator characteristics beyond basic behaviors like using social cues or exhibiting creativity itself. At the same time, the robot's speech characteristics may be expected to have an influence on participants' creativity, yet what causes these effects is as yet unknown.

3 Methods

The study was carried out based on an online survey system with a between-subject design. We created an online questionnaire in which a robot guided the participants through a creativity task. Videos of the robot provided the participants with the necessary instructions. In the videos in one condition, the robot's synthesized speech was manipulated globally with praat [21] to increase the pitch range of the robot's utterances, whereas in the other, the pitch range of the same synthesized robot utterances was decreased.

3.1 Procedure

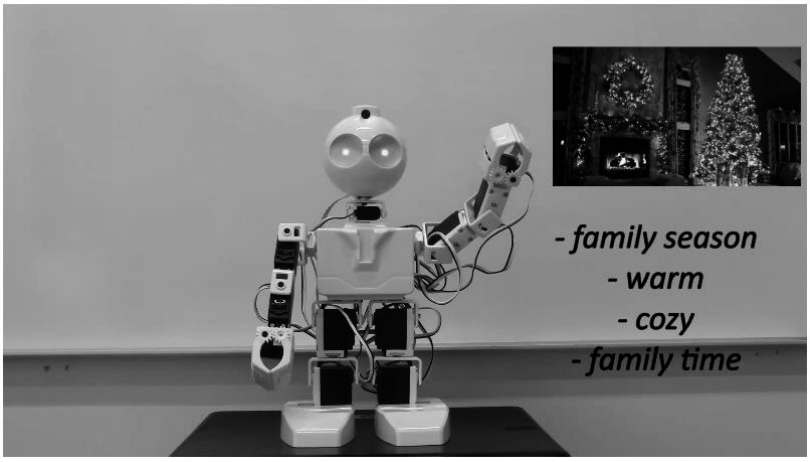

Figure 1 - The robot during the instructions for the brainstorming phase 1.

Embedded in a questionnaire, the participants carried out a creativity task, about which they were instructed by a robot using a synthesized voice that exhibited either high or low pitch range. The creativity task used is visual synectics [22], a creativity technique in which participants first collect associations concerning a visual stimulus and use those associations then to

brainstorm about solutions for a novel product (cf. [23]). The task consists of two phases: one brainstorming phase in which an image serves as source of inspiration, and a second phase in which the participants apply the ideas from the first phase to create a new product.

The image used for the association task shows a tropical beach - while the choice of the visual stimulus can be random and has no direct impact on the results, we chose the beach to evoke positive associations. The innovation task then was to invent a new type of chocolate, in accordance with the example discussed in Brem & Brem [23], since most people like chocolate and have some experience with it.

Participants saw two videos[1] of a small humanoid robot, the JD Humanoid from EZ-robots (see Figures 1 and 2), in which the robot first introduces the visual association task and then the product innovation task. In both conditions, the same two videos were used, yet the robot voices associated with the videos differed according to condition, i.e. with respect to the extent of the pitch range. In addition, participants were asked about their age and gender. After the creativity task, participants were also asked to rate the robot with respect to measures related to speaker charisma and some additional questions (see below).

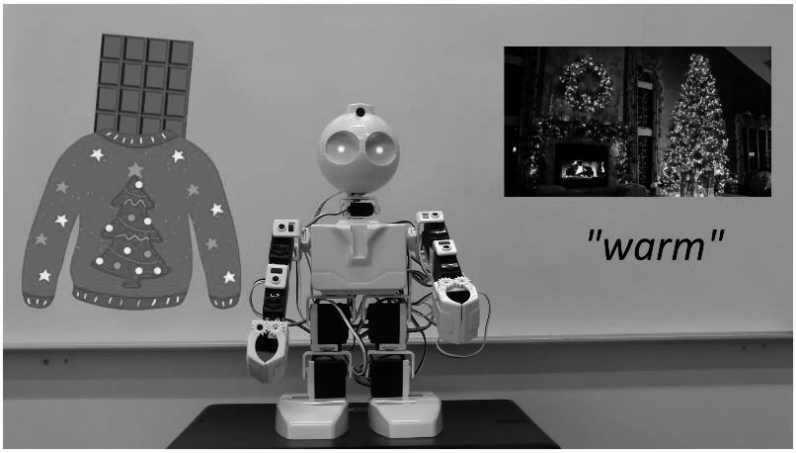

Figure 2 - The robot during the instructions for the innovation phase 2.

3.2 Stimuli

For the stimuli, first two texts by means of which the robot instructed the participants to carry out first the association task and then the innovation task were created:

These instructions were synthesized using the free text-to-speech system wideo.co, voice Jack Bailey (US). This voice simulates a male voice and uses a slight US American accent.

The pitch range of the robot's synthesized utterances were then manipulated in Praat [21] by expanding and reducing the pitch range of the original synthesized file by 40%. We used an overall manipulation only in order to make the manipulation as objective as possible and to keep our own contribution as small as possible. In this way, two versions of the robot's instructions were created, which differed only with respect to the pitch range of the robot's voice. The voice with the narrowed pitch range extends from 81 to 180 Hz and thereby spans about one

[1] Condition 1 video 1: https://youtu.be/8QBAn6leswk
Condition 1 video 2: https://youtu.be/uo5FnCR-AtY
Condition 2 video 1: https://youtu.be/QWvCHEhJAcM
Condition 2 video 2: https://youtu.be/IAZefvCeUFo

octave, i.e. 12 semitones, which is typical of matter-of-fact read speech in Western Germanic languages (cf. [24]). In contrast, the voice with the enlarged range extends from 101 to 294 Hz, which corresponds to about 1.5 octaves or 18-19 semitones and which was found to be typical for charismatic speech, for instance, of Steve Jobs or Benjamin Netanyahu (cf. [15], [25]). These two manipulation conditions are henceforth referred to as ‘narrow range’ and ‘large range’ and constitute the independent variable in our experiment.

The resulting audio files were then connected to two robot videos, one for the first task and one for the second, in which the robot used some co-speech gestures. Furthermore, in the first video (see Figure 1), the robot points to an image and provides some example attributes, while in the second video (see Figure 2), it gestures towards an innovative product based on those attributes. Thus, both videos are identical across the two conditions, with the only difference being the pitch range of the robot's voice in these videos.

The images used as stimuli for the creativity tasks are then different from those the robot uses, namely an image of a beach for the brainstorming task and an image of chocolate for the innovative product.

3.3 Questionnaire

A questionnaire was used to elicit some demographic information about the participants (age and gender). Furthermore, participants were asked to rate the robot in 7-point Likert scales according to the following attributes: enthusiastic, charming, convincing, engaging, boring, passionate, self-confident, uninspiring, and charismatic. These attributes had been found to be revealing in previous work on charismatic speech styles ([3], [5-6]).

In addition, we asked participants to rate themselves concerning the extent to which they perceived themselves to have been carried away by the task, were motivated to do a good job, were involved in the task and felt full of energy on a 7-point Likert scale. These questions were meant to give us an idea of the extent to which participants may have reached a flow state and felt engaged in the task.

3.4 Participants

Participants were recruited via the crowdsourcing platform Prolific and randomly and automatically assigned to one of the two conditions. Participants were paid 0.88£ for their participation, which lasted 7 minutes on average. The recruitment criteria were that participants were a) native speakers of English, to ensure that participants had no problems understanding the robot and the instructions, and b) being located in the US, to ensure that participants were familiar with the US American accent of the robot's voice. We also requested a gender-balanced sample. Correspondingly, 25 of the 5o participants identified as female, 25 as male, and none as other. They were on average 36.5 years old, with an age range between 19 and 78. This age range and gender distribution is typical for studies using crowdsourcing methodologies (cf. [26]).

53 participants were recruited, and their answers in the creativity tasks were checked for validity. 22 participants in one condition and 28 in the other condition had faithfully addressed the association and the innovation tasks (i.e. produced more than 3 ideas), leaving 50 participants whose data were included in the analysis.

3.5 Data Analysis

For the data analysis, we analyzed participants' associations to the picture and their product ideas based on Guilford's [27] categories of divergent thinking. This means that we coded the data according to four measures: Fluency, Flexibility, Originality, and Elaboration. Fluency concerns the quantity of ideas a participant generates, i.e. the number of ideas provided. In

contrast, flexibility is the number of different categories of relevant responses, i.e. the range of categories covered by the ideas. For instance, if there were two categories identified, the flexibility score would be 2. Originality is the relative novelty and uniqueness of each answer. To calculate originality scores, each response is compared to the total amount of responses from all participants. Responses that were given by only 1% of all participants are considered unique (1 point). Responses that were only given by that particular participant are considered one of a kind (2 points). Thus, higher total scores indicate more original thinking. Elaboration is the amount of detail in the responses. For example, "a blue chocolate" equals 0 points, but "a blue chocolate wearing sunglasses" would be 1 point. An additional point was given for any further details provided, such as what the chocolate tastes like or details in the design.

As the first step toward coding the data, the questionnaire results were downloaded into an excel file. Then, all the responses of the 50 participants who had successfully completed the survey were recorded and sorted according to the condition – i.e., decreased / increased pitch range – they were in. Next, for each participant, we omitted the repeated answer items and counted the number of their remaining ideas to calculate their fluency. Then, to measure flexibility, their responses were categorized into different groups and then the number of categories for each participant was counted. To illustrate how categorizations were done, the terms *swimming, snorkeling, partying,* and *dancing* were placed in 'activities' category and the words *warm, hot, tropical* and *sunny* were considered to belong in the 'weather' group. As for the third scoring criterion, we measured the originality of the responses according to the uniqueness of the ideas produced. For example, the word *tropical* was mentioned by at least 18 participants, so it was considered as a frequent word that led to no score. However, the term *Jack Sparrow* was stated by a single participant and hence regarded as unique. Finally, to measure elaboration, we investigated the amount of detail in participant's answers. For instance, the response 'white chocolate' was regarded as having not much detail, whereas the idea 'Chocolate-covered banana chips' bore more detail, leading us to allocate a score of zero and one to them, respectively. The response 'Island chocolate - made using cacao beans from the island, with a refreshing and fruity flavor' was considered detailed and received a score of three.

As for the statistical analysis of the data, we used a conservative multivariate Kruskal-Wallis test to analyze the dataset whose individual variables did not consistently show a normal distribution or met the assumption of homoscedasticity. In addition, a linear discriminant analysis (LDA) was conducted to determine the degree to which we can predict the participants' questionnaire ratings and creativity performances in the two pitch-range conditions, 'narrow range' and 'large range'.

4 Results

Results show a clear advantage of the 'large range' over the 'narrow range' condition, see Figure 3. That is, the robot that moderated the two creativity tasks with about 1.5 octaves (18 semitones) instead of just a single octave (12 semitones) in between the lowest and highest pitch values in its utterances was rated by the participants to sound significantly more enthusiastic ($H[1]=4.27$, $p=0.03$) and more charismatic ($H[1]=4.65$, $p=0.02$). Additionally, participants who received the instruction from that robot provided higher self-assessment scores on the "I felt full of energy" scale ($H[1]=4.11$, $p=0.05$).

More importantly, we also see differences in the actual creativity task performances as a function of the pitch range used by the robot. Compared to the participants who did the creativity tasks based on the 'narrow range' instructions, the participants who received their instructions from the 'large range' robot produced higher levels of Flexibility ($H[1]=4.14$, $p=0.04$) and Elaboration ($H[1]=3.99$, $p=0.05$) in the association (i.e. brainstorming) task as well as a higher level of Originality in the innovation (product-creation) task ($H[1]=3.96$, $p=0.05$).

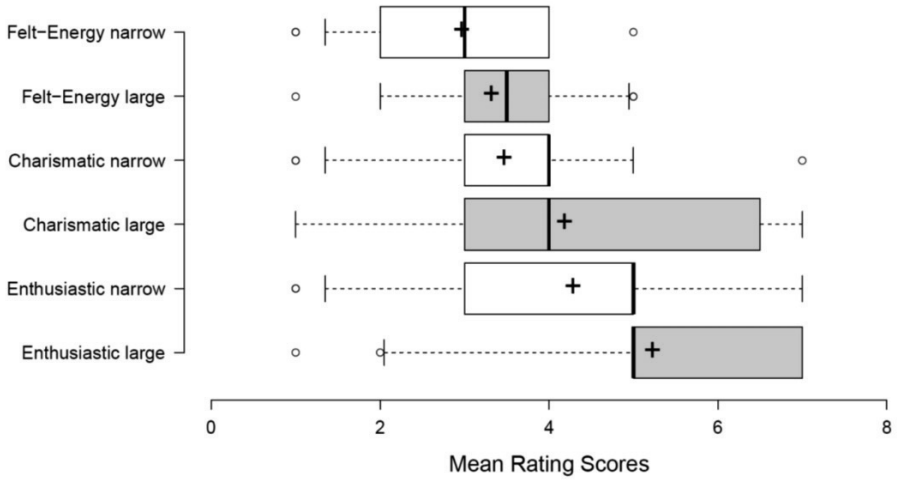

Figure 3 - Summary of significant differences (narrow vs large) in the questionnaire ratings.

More specifically, we see in Figure 4 that it was actually only the large-range condition that generated a significant amount of original product ideas at all. In the narrow-range condition, the originality level was practically 0, i.e. there were almost no unique or one-of-a-kind ideas within the sample of participants. The range of ideas (Flexibility) was also significantly lower in the narrow-range condition and, with some participants showing the minimum value of 1. A similar pattern emerged for the elaboration performance. The amount of additionally specified details on the product idea was close to 0 in the narrow-range condition, while many participants in the large-range condition invested the extra effort to substantiate their ideas with details. The latter finding fits with the statement "I felt full of energy", which participants in the large-range condition subscribed to to a greater extent than participants in the narrow-range condition. Thus, to sum up, the difference generated by the robot's intonation does not only concern different degrees of creativity, but rather the difference between the presence or absence of creativity.

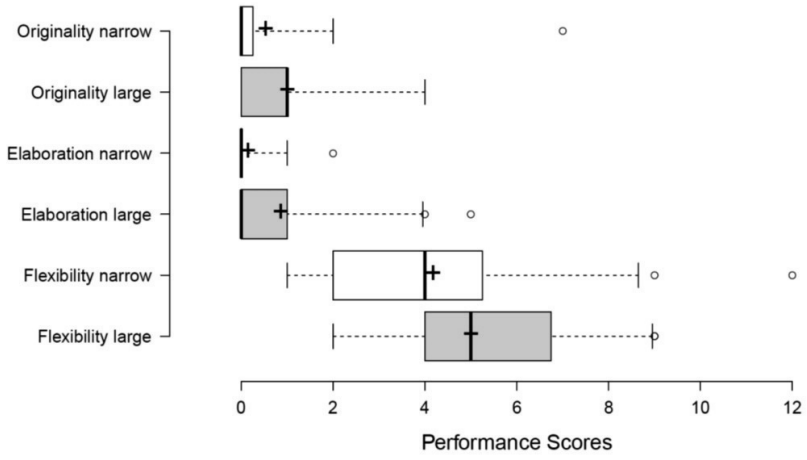

Figure 4 - Summary of significant differences (narrow vs large) in task performance.

In line with the clear results obtained from the multivariate Kruskal-Wallis test, the LDA also came out significant (discriminant function: Eigenvalue=1.56, canonical correlation=0.78, Wilks-Lambda=0.40, $\chi^2[21]=34.46$, $p=0.03$). That is, based on both the ratings and the task-performance indicators, it was possible to significantly predict whether a participant was a member of the narrow-range or the large-rage sample. The ratio of correct predictions was 90.1% for the large-range condition and 92.9% in the narrow-range condition, which is obviously considerably above the 50% chance level threshold in a two-condition setup. Only two participants in each condition were misassigned to the other respective condition. In addition to the significant dependent variables presented above, the LDA's prediction performance also benefited greatly (in terms of the standardized canonical discriminant coefficients) from the motivation ratings, which were on average lower in the narrow-range than in the large-range condition.

5 Discussion

In this study, we set out to explore whether a robotic creativity facilitator's speaking style has an impact on people's creativity and engagement. We tested the speaker's pitch range because previous work suggests that it may be related to listener engagement and speaker charisma. The results show that this simple global manipulation of the robot's synthesized utterances has a significant effect on how the robot is perceived, how energetic participants rate themselves and how original, flexible and elaborate their ideas are - even though the robot's contributions were only presented to the participants in prerecorded videos. Given the non-interactive nature of the stimuli and the restricted role of the robot during the task, these results are truly surprising. While Fucinato et al. [7] found similar results in a similar scenario, their study employed several different prosodic manipulations; that pitch range alone has such a strong effect is therefore remarkable. The effect found may be related to listeners' heightened engagement in the task, as evidenced by their self-reports on higher energy levels, which then led to higher performance.

The experiment was carried out in an online questionnaire format with participants recruited through a crowdsourcing platform. This situation is probably quite different from more common creativity workshop scenarios. However, given that the typical creativity workshop would rather involve more, and more responsive, contributions by the facilitator than in the current scenario, one may expect that the effect may even be more pronounced in such a context. We do not know, however, how long such an effect may last; the task investigated in the current scenario took only a few minutes. Future work will have to show whether the effect of speaking style on creativity wears off over time.

We can conclude that a facilitator's pitch range has a considerable influence on people's creativity in activities like the visual synectics task.

6 Acknowledgements

This paper was partially supported by the project *Facilitation Robotics for Teams Foundational Research on Group Processes in Human-Robot Ensembles with Social Robots (S-Factor)*, kindly funded by the Nordforsk foundation. Note the 2nd author (ON) is the founder and CEO of the speech-technology company AllGood Speakers ApS. Please visit the following link for a conflict-of-interest statement: https://oliverniebuhr.com/conflict-of-interest.html

7 References

[1] WALLGREN, M. K. (1998). Reported practices of creative problem solving facilitators. *The Journal of Creative Behavior*, *32*(2), 134-148.

[2] ROSENBERG, A. and HIRSCHBERG, J. (2009). Charisma perception from text and speech. Speech Communication 51:640-655.

[3] FISCHER, K., NIEBUHR, O., Jensen, L. C. and BODENHAGEN, L. (2020): Speech Melody Matters – How robots can profit from using charismatic speech. *ACM Transactions in Human-Robot Interaction 9, 1, Article 4*: 1-21.

[4] FOX CABANE, O. (2012). *The Charisma Myth: How Anyone Can Master the Art and Science of Personal Magnetism*. New York, NY: Penguin.

[5] FISCHER, K., NIEBUHR, O. and ALM, M. (2021): Robots for Foreign Language Learning: Speaking Style Influences Student Performance. *Frontiers in Robotics and AI* 8: 273ff, DOI=10.3389/frobt.2021.680509.

[6] NIEBUHR, O. (2021). Advancing higher-education practice by analyzing and training students' vocal charisma: Evidence from a Danish field study. Proc. 7th International Conference on Higher Education Advances, Valencia, Spain, 743-751.

[7] FUCINATO, K., NIEBUHR, O., NØRSKOV, S. and FISCHER, K. (submitted). Charismatic Robot Instructions Enhance Team Creativity.

[8] STEWART, J.-A. (2006). High-performing (and threshold) competencies for group facilitators. Journal of Change Management, 6, 417–439.

[9] JOHNSSON, M., EKMAN S., WIKTORSSON M. & KARLSSON T. (2010). A model-based process for developing environmental innovations: Four cases where the RAFT-model has been used at environmental innovations. In: IASP Conference XXVII, Daedak, South Korea.

[10] WRÓBEL, A. E., CASH, P., & LOMBERG, C. (2020). Pro-active neutrality: The key to understanding creative facilitation. Creativity and Innovation Management, 29(3), 424-437.

[11] JOHNSSON, M. (2018). The innovation facilitator: characteristics and importance for innovation teams. Journal of Innovation Management, 6(2), 12-44.

[12] STRANGERT E. and GUSTAFSON, J. (2008) What Makes a Good Speaker? Subject Ratings, Acoustic Measurements and Perceptual Evaluations. Proc. 9th International Interspeech Conference, Brisbane, Australia, 1688-1691.

[13] NIEBUHR, O., & SKARNITZL, R. (2019). Measuring a speaker's acoustic correlates of pitch–but which? a contrastive analysis based on perceived speaker charisma. In *Proceedings of 19th International Congress of Phonetic Sciences*.

[14] NIEBUHR, O., VO e, J., and BREM, A. (2016). What makes a charismatic speaker? A computer-based acoustic-prosodic analysis of Steve Jobs tone of voice. Computers in Human Behavior 64: 366-382.

[15] NIEBUHR, O., BREM, A., MICHALSKY, J., & NEITSCH, J. (2020). What makes business speakers sound charismatic? A contrastive acoustic-melodic analysis of Steve Jobs and Mark Zuckerberg. *Cadernos de Linguistica e Teoria da Literatura*, *1*(1). https://doi.org/10.25189/2675-4916.2020.V1.N1.ID272

[16] LUBART, T., ESPOSITO, D., GUBENKO, A., & HOUSSEMAND, C. "Creativity in Humans, Robots, Humbots." *Creativity. Theories–Research-Applications* 8, no. 1 (2021): 23-37.

[17] KAHN, P. H., KANDA, T., ISHIGURO, H., GILL, B. T., SHEN, S., RUCKERT, J. H., & GARY, H. E. (2016, March). Human creativity can be facilitated through interacting with a social robot. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 173-180). IEEE.

[18] HU, Y., FENG, L., MUTLU, B., & ADMONI, H. "Exploring the Role of Social Robot Behaviors in a Creative Activity." In *Designing Interactive Systems Conference 2021*, pp. 1380-1389. 2021.

[19] Alves-Oliveira, P., Arriaga, P., Cronin, M. A., & Paiva, A. "Creativity encounters between children and robots." In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, pp. 379-388. 2020.

[20] Ali, S., Devasia, N., Park, H. W., & Breazeal, C. "Social Robots as Creativity Eliciting Agents." *Frontiers in Robotics and AI* 8 (2021).

[21] Boersma, P. (2001). Praat: A system for doing phonetics by computer. Glot International 4, 341-345.

[22] Asanowicz, A. (2008). How to Find an Idea? - Computer Aided Creativity. Architecture in Computro [26th ECAADe Conference Proceedings], 735–742.

[23] Brem, A., & Brem, S. (2019). Die Kreativ-Toolbox für Unternehmen. *Ideen generieren und innovatives Denken fördern. Stuttgart*.

[24] Mennen, I., Schaeffler, F. & Docherty, G. (2007). Pitching it differently : a comparison of the pitch ranges of German and English speakers. Proc. 16th International Congress of Phonetic Sciences, Saarbrücken, Germany, 1769-1772.

[25] Niebuhr, O. & Silber-Varod, V. (2022). How versatility performance influences perception of charismatic speech – A study on two Israeli politicians. Interaction Studies.

[26] Lee, H.R., Cheong, E., Lim, C. and Fischer, K. (2022): Configuring Humans: What Roles Humans Play in HRI Research. *Proceedings of the IEEE Human-Robot Interaction Conference* 2022.

[27] Guilford, J. P. (1967). The nature of human intelligence. McGraw-Hill.