Understanding the Perception of Incremental Robot Response in Human-Robot Interaction*

Lars Christian Jensen[1] Rosalyn Melissa Langedijk[1] and Kerstin Fischer[1]

Abstract— **Incremental feedback, i.e. the timely response to human behavior while it is happening, has previously been found to potentially speed up human-robot interactions, but it is unclear how people evaluate incremental robots. In this study, we show that the evaluation of incremental robot response depends on the actual success of the incremental feedback; that is, if the feedback leads to increased efficiency, people evaluate the robot as more competent and more credible. If the robot does not use incremental feedback, no interaction between evaluation and efficiency can be found. Thus, incremental feedback draws people's attention to interaction success.**

I. INTRODUCTION

Human interaction is characterized by incremental processing and, correspondingly, high mutual responsiveness and a fast pace. That is, people process their interaction partners' behaviors while they are happening and plan their responses at the same time so that response times are brief and can be provided online, while an action is still in progress. Concerning processing, incrementality concerns the online analysis of incoming input in a piecemeal fashion, which enables fast response times, interruptabilty and flexible dialog planning, whereas concerning production, incremental feedback means that a robot can provide the user with feedback while the action is happening. Recent work in human-robot interaction [1, 2] has begun to implement such processing principles in robots.

However, while one might expect that a human-like capability is always evaluated positively in human-robot interaction, as it has been shown in so many other areas of human behavior (e.g. gaze behavior [3], approaching behaviors [4, 5], and in the application of politeness principles [6]), incremental robot responses have not always received highest ratings [7, 8, 9]. In this paper, we explore the relationship between the perception of incremental robot response and task efficiency. Thus, we aim to shed light on the effects of incrementality in human-robot interaction.

II. PREVIOUS WORK

Previous work shows that incremental speech processing can decrease a system's response time since it will begin production before it has finished processing information relevant to that production [1, 2]. This may increase a system's efficiency; for instance, Kennington, Kousidis, Baumann, Buschmeier, Kopp, and Schlangen [10] implement incremental speech in a dialog system for a car simulator. Their study shows that participants who interacted with the incremental system perform better at driving-related tasks than participants who interacted with a non-incremental speech system. Ghigi, Eskenazi, Torres, and Lee [11] implement incremental speech processing in an information retrieval system. Their study shows that although dialogues become longer, the success rate is higher in the incremental condition. In other words, people are more likely to get the information they request, and they experience fewer problems.

There is already some work on implementing incremental speech processing in robots, though many studies do not provide experimental evaluations of their systems in live human-robot interactions. For example, Manuvinakurike, Paetzel, Qu, Schlangen, and DeVault [12] incrementally classify utterances into 18 different dialogue acts in their dialogue segmentation system, based on word-for-word processing of the speech input. Similarly, Carlmeyer, Schlangen, and Wrede [13] present a dialogue system for use in HRI in which users can provide feedback and correction to the robot. However, the work is only presented as a proof-of-concept and not tested experimentally in interaction. In contrast, Chromik, Carlmeyer, and Wrede [7] show that people interacting with a robot in an object-fetching task perform better when instructions are given incrementally than when given all at once. Thus, participants in the incremental condition will better be able to find the correct items.

Also when providing feedback, incremental dialog systems tend to outperform systems with non-incremental feedback. Generally, incremental speech systems are perceived as more polite and efficient [2, 14], more natural [15], as well as more responsive, enjoyable and attentive [16] than non-incremental systems. Kok, Hough, Hülsmann, Botsch, Schlangen, and Kopp [9] present a virtual coach that provides online feedback as participants do exercises, such as squats. Here, the feedback comes in the format of "watch your neck", and "go a little deeper" as the system detects errors in participants' behavior. Evaluations of the system show that incremental instructions were correlated with higher ratings of perceived intelligence, helpfulness, responsivity, humanlikeness and clarity, but that the robot is also perceived as tiring. The behavior was generated on the basis of analyses of a corpus of interactions between a human exercise coach and experiment participants [17].

However, there are also studies that suggest that incremental feedback may be received negatively. For example, while Baumann and Lindner [14] find their simulated robot to be

*This work was carried out in the framework of the Smooth Project, supported by the Danish Innovation Foundation

[1]Lars Christian Jensen, Kerstin Fischer and Rosalyn Melissa Langedijk are all with the Department of design & Communication, University of southern Denmark, Universitetsparken 1, 6000 Kolding, Denmark. larscj@sdu.dk, kerstin@sdu.dk, rla@sdu.dk

perceived as more polite and natural, Chromik, Carlmeyer, and Wrede [7] report that their simulated robot is rated as less natural when using incremental speech, and Carlmeyer, Schlangen, and Wrede [8] report that their simulated robot is rated as less likable when using incremental speech.

To sum up, incremental speech interfaces are generally promising, but have so far only rarely been used on embodied robotic systems. It is therefore an open question how many of these findings also apply to the interaction with embodied robots. Furthermore, research on incremental speech processing in HRI indicates that incrementality can sometimes contribute to efficiency, and there is some evidence that suggests that participants perceive the robot more positively, although there is also evidence of the contrary.

Another open issue is whether incrementality contributes to a robot's persuasiveness. In particular, incremental feedback enables users to have online access to what contextual information a robot is attending to. For instance, by saying "go a little deeper," the robot indicates that it perceives the user's actions. Incrementality can thus function as a means to personalize an interaction by indicating that one takes the partner into account; it is therefore possible that incremental feedback has an effect on persuasion.

III. Hypotheses

Our literature review suggests that incrementality may positively affect task efficiency. We therefore predict:

- H_{1a}: Participants in the incremental condition complete tasks faster than participants in the control condition.

The literature review also suggests that incrementality may positively influence ratings of the robot's competence, such as intelligence [9] or efficiency [2, 14]. We therefore assume that the actual time people need to fulfill a task will be correlated with their perception of the robot. We predict:

- H_{1b}: a positive relation between task time and rated robot competence for participants in the incremental condition.

We furthermore assume that increased efficiency will influence people's judgments about the robot's awareness of the situation; we predict:

- H_{1c}: a positive relation between task time and rated robot awareness for participants in the incremental condition.

Some studies (e.g. [7, 8]) report on a trade-off between the efficiency of incremental speech processing on the one hand and affective ratings on the other. Therefore we predict:

- H_{1d}: a negative relation between task time and rated robot warmth in the incremental condition

and:

- H_{1e}: a positive relation between task time and rated robot discomfort in the incremental condition.

Only very little research deals with how incremental speech processing affects an agent's persuasive capabilities. For example, Kobberholm, Carstens, Bøg, Santos, Ramskov, Mohamed, and Jensen [18] found no relation between the two. However, we do know from human psychology [19] that people are more easily persuaded by people they like and by authority figures (such as doctors and scientists). Since an incremental robot signals more to users that it takes their actions into account, we expect it to me more persuasive. In our experiment, we measure the robot's persuasiveness by means of the extent to which participants follow the robot's advice and drink water after the experiment (see below). With regards to the relationship between persuasion (water intake) and subjective rating we therefore predict:

- H_{2a}: A positive relation between water intake and rated robot competence in the incremental condition.
- H_{2b}: A positive relation between water intake and rated robot awareness in the incremental condition.
- H_{2c}: A positive relation between water intake and rated robot warmth in the incremental condition.

and:

- H_{2d}: a negative relation between water intake and rated robot discomfort in the incremental condition.

Our final hypothesis concerns directly the relationship between persuasion and incrementality:

- H_3: Participants will drink more after the experiment in the incremental than in the non-incremental condition.

We carried out two studies in which the robot provides incremental feedback, one in the lab with students and members of the staff as participants and one at the municipality's LivingLab, a fully furnished flat to showcase technical innovations in the healthcare domain. Participants in this study were healthcare personnel.

IV. Method

A. Experiment Overview

In the two studies carried out, a robot guides participants through a large room or flat, indicates to them what objects to collect and offers to carry them for the participant. This task provides opportunity for incremental feedback when people are trying to find relevant objects.

What people are instructed to collect are objects needed to set a table, which they discover only over the course of the interaction. When they are instructed to pick up a napkin, for instance, the robot does, or does not, use incremental feedback to help them find the napkins in a box on the shelf. Similarly, when instructing participants to find the candle in a cupboard, the robot either does or does not use incremental feedback to help them locate the candle in one of the drawers. The robot furthermore informs participants about the health benefits of water intake when it instructs them to pick up a glass.

B. Experimental Conditions

The experiments were carried out in a between-subject experimental design with two conditions that differ with respect to whether the robot provides incremental feedback or not.

In the **incremental condition**, the robot modifies its speech incrementally based on participants' non-verbal conduct. As participants are looking for certain hidden items, the robot can direct their search by producing utterances like "more to the right" and "yes, a little more". In the **non-incremental condition**, the robot provides complete descriptions of where to find the respective object and only repeats its previous utterance in cases in which participants are not finding what they are looking for.

C. Subjective Measures

Participants were presented with a questionnaire before and after the experiment. Demographic information and previous experiences with robots were elicited in the questionnaire given prior to participating, while participants' ratings of the robot were elicited in the post-experimental questionnaire. The post-experimental questionnaire consists of the RoSAS scale [20], a standardized instrument for measuring social attributions to robots.

The scale consists of the three indices warmth, competence, and discomfort. Each index contains a collection of adjectives with which the participants are asked to rate the robot. All items are rated on a 5-point Likert scale where 1 is 'not at all' and 5 is 'very much'. Warmth includes the adjectives *happy*, *feeling*, *sociable*, *compassionate*, and *emotional*. Competence includes the adjectives *capable*, *responsive*, *interactive*, *reliable*, *competent*, and *knowledgeable*. The discomfort scale includes the adjectives *scary*, *strange*, *awful*, *awkward*, *dangerous*, and *aggressive*.

Finally, in addition to RoSAS we also included an awareness scale, which consists of the following three question:

- to what extent do you think the robot took you into account?
- to what extent do you think the robot responded to your actions?
- to what extent do you think the robot perceived you?

These questions allow us to identify the extent to which incremental feedback is taken as an indicator of increased awareness of the human participant.

D. Objective Measures

The objective effects of our manipulations concern water intake on the one hand and the time participants needed to find a given object on the other. That is, the effect of incremental speech is evaluated by measuring the time it takes participants to find the two objects that are concealed from view. Time is measured from when the robot issues the instruction until participants takes hold of the object. Water intake is measured in milliliters; in particular, we measured how much water was missing from the 1 liter jug and from participants' glasses after they were done filling out the post-experimental questionnaire.

E. Robot and Software

Fig. 1. Robot

The robot used for this experiment is a Turtlebot 2 on a Yujin Kobuki mobile base. The robot is equipped with an Orbbec Astra 3D camera and is controlled by an Intel NUC running Canonical Ubuntu 16.04 LTS and ROS Kinetic. The robot moves autonomously from point to point. However, target locations are set by a remote wizard using RViz. The autonomous navigation is enabled by SLAM map building [21]. The robot's speech is presynthesized using IVONA TTS and ttsmp3.com (voices 'George' (English) and 'Mads' (Danish)). A remote wizard controls, via a collection of shell scripts, when the robots produces its speech and, to a limited extent, what it says (most actions are predefined, see below). Cameras were placed around the room on walls and on ceilings and are live-streamed to a PC in an adjacent room. The robot's design is a low-fidelity prototype, coated in styrofoam and equipped with a pair of eyes made from bottle caps (see Figure 1).

F. Speech Management

The robot's speech is presynthesized, and all of the robot's verbal actions are controlled via a series of shell scripts, thereby limiting the options available to the wizard at any given time during the experiment. This was to decrease the cognitive load of the wizard who already had to monitor several aspects of the participants' behavior and point the robot in the right direction. The script runs in the command line, and the wizard selects the next utterance by clicking the appropriate numerical key. This is demonstrated in Sample Shell Script below (Source Code 1).

Here, the '1.Continue' is printed to the screen indicating that the only option available to the wizard is '1', which will progress the interaction. Thus, in the example the wizard is not able to decide what the robot is going to say, merely when it is going to do so. The script ignores all other inputs than the numerical key press ‘1’. There are however also situations in which the robot needs to adjust to the participant’s behavior. Such situations are resolved by giving the wizard a small list of possible actions to perform.

```
while [ "$control" = "0" ]
do
echo "1: User puts glass on robot"
echo "2: User keeps the glass"
echo "3: User does not pick up glass"
echo " "
while true; do
read -rsn1 input
if [ "$input" = "1" ] && [ "$place" = "0" ]; then
echo " "
echo "You are welcome to put everything on my
        tray."
let "place++"
echo " "
play LivingLabAudio/napkins6.mp3
control="1"
break
elif [ "$input" = "1" ] && [ "$place" > "0" ]; then
echo " "
echo "Great"
echo " "
play LivingLabAudio/great.mp3
control="1"
break
elif [ "$input" = "2" ]; then
echo " "
control="1"
break
elif [ "$input" = "3" ]; then
echo "please pick up a glass"
play LivingLabAudio/glass5.mp3
break
fi
done
done
```

Source Code 1. Sample Shell Script

Note that the wizard is not given options for what to do, but rather a list of possibilities that the participant could be doing. Thus, the wizard does not need to evaluate the ‘right’ course of action, but merely respond to what he or she is observing. The different possible participant behaviors are derived from pilot studies. The incremental speech is implemented in a similar fashion. That is, the script will keep looping the options available to the wizard until the participant reaches the target object. This creates the illusion that the robot incrementally adjusts its own verbal output to the participant’s nonverbal conduct, for example, by saying ‘higher’ when a participant needs to look higher up on a shelf. The scripts are run remotely from an adjacent room to the Turtlebot NUC via SSH.

G. Participants and Settings

Both studies were carried out in facilities in which the robot could show participants around and indicate things for them to pick up, and in both cases, the tour ended with participants setting their own table.

In study 1, the experiments were carried out in our human-robot interaction lab, where participants were students, university lecturers and other staff, as well as some members of the public who happened to be on campus. 52 participants took part; one interaction had to be interrupted because the participant tried to sit on the robot, which leaves 51 valid interactions (26 in the incremental, and 25 in the non-incremental condition), mean age 28.2 (SD= 11). 31% were women, thus men were overrepresented.

In order to gain more credibility and to approach the target audience of our drink-serving robot, which is designed to address dehydration in elderly care facilities [22], we carried out a second study. In study 2, the setting was the municipality’s LivingLab, a large, multi-room facility that exhibits innovations in the healthcare domain. The experiment took place in the living room and the kitchen area of the LivingLab. Participants were exclusively healthcare staff, most of whom were women; there are only two male participants. 46 participants took part in the experiment. Mean age of participants is 42 (SD= 11). 24 participants are in the non-incremental condition and 22 are in the incremental condition.

H. Procedure

The following interaction protocol is identical for both studies, despite them taking place in two different locations.

Participants were greeted in the hallway and led to a separate room to fill out a consent form and the pre-experimental questionnaire. Then they were taken to the lab, introduced to the robot and told that the robot would lead them through the experiment. They were then left alone in the room with the robot.

The robot first greeted the participant and then moved towards a shelf, where it instructed the participant to pick up an item which was hidden in another container (i.e. napkins in a box in study 1 and matches in a plastic jar in study 2). In the incremental condition, the robot uses incremental speech to direct the participant. Then, the robot moves across the room and instructs the participant to pick up a glass and comments on the benefits of drinking enough fluids during the day (the manipulation to address the persuasiveness of the robot). The next interaction concerns a placemat, where the robot explicitly displays its situation awareness by commenting on the placemat the participant picks up. For example, if a participant picks up the green placemat, the robot says ‘Ah the green one, that’s my favorite too!’. If no items have yet been placed on the robot, it offers to carry them for the participant. Next, the robot moves to another shelf and instructs the participant to pick up a snack. Here, participants can choose between a cookie or a fruit. Again, the robot displays situation awareness by commenting on their choice (however, without judging their choice). After driving to yet another location, participants are asked to pick up a candle. The candle is hidden away in a drawer, so for participants in the incremental condition, the robot directs them using incremental speech. Finally, the robot asks people to set a table, to have a seat, to enjoy their snack and to fill out the post-experimental questionnaire, which is prepared for them on a tablet.

Construct		Item
Awareness	aware1	The robot responded to your actions
	aware2	The robot took you into account
	aware3	The robot perceived you
Discomfort	disco1	The robot is scary
	disco2	The robot is strange
	disco3	The robot is awful
	disco4	The robot is dangerous
	disco5	The robot is aggressive
	disco6	The robot is awkward
Competence	comp1	The robot is reliable
	comp2	The robot is competent
	comp3	Knowledgeable
	comp4	Interactive
	comp5	Responsive
	comp6	Capable
Warmth	warm1	Organic
	warm2	Sociable
	warm3	Emotional
	warm4	Compassionate
	warm5	Happy
	warm6	Feeling
Performance Construct		
Time	time1	Time measurement task 1
	time2	Time measurement task 2

TABLE I
LIST OF CONSTRUCTS

I. Data Analysis

The proposed model and hypothesis testing is performed using PLS analysis with [23] using the package [24]. This approach allows us to perform a factor analysis along with hypothesis testing. In line with our research questions, we will build, present and evaluate two different models, one for each data set, to highlight how people in each of these groups act differently to the same stimuli. The analysis of each model is twofold: We assess the reliability and validity of the measurement mode before we proceed to the hypothesis testing proper.

V. RESULTS

First we present a factor analysis of our measurement model, after which we proceed by testing our hypotheses.

A. Analysis of the Measurement Models

The construct *time* refers to the completion time for the two tasks in each interaction. In these tasks, participants had to find objects based on the robot's instructions using either incremental or non-incremental speech (depending on the condition). In addition to the *time* construct, the models also include the experimental condition (2 levels), questionnaire results (*awareness*, *competence*, *discomfort* and *warmth*), the amount of water participants consumed after the experiment, as well as moderation (interaction) effects between the experimental condition, water consumption and the time it took to complete the tasks.

The assessment of the measurement models relies on two measures of internal consistency, Cronbach's alpha and Dillon-Goldstein's rho index, as well as on a measure of unidimensionality, namely eigenvalues. Values for indices above 0.7 for both measures of internal consistency are generally accepted as indicators of homogeneity of the constructs. The unidimensionality assumes the first eigenvalue to be larger than one. Using these criteria, we can observe (see Table II) that awareness and competence fulfill all assumptions for both studies, while discomfort is fulfilled for study 2, but has a Cronbach's alpha value < 0.7 for study 1. Likewise, warmth has a Cronbach's alpha value < 0.7 for study 2. However, in order to maintain comparability, both models will include both discomfort and warmth. The construct *time* scores fairly low on Cronbach's alpha, but meets the assumption of the other criteria. A reason for this may lie in the low number of items or in the variance of those items. The two models will include the *time* construct, but any conclusion derived from it is subject to further investigation.

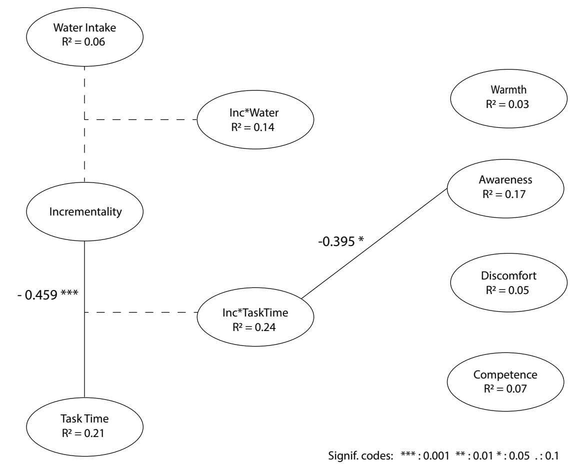

Fig. 2. Path Model for Study 1

B. Hypothesis Testing

We evaluate each of the hypotheses by observing the path coefficients. Statistically significant path coefficients are indicated by fully drawn lines in each of the models. The explanatory power is indicated by R^2 values. For study 1, we can observe a significant negative coefficient between the experimental variable and task time. That is, participants in the incremental condition complete tasks faster than participants in the non-incremental condition. Likewise, we can observe a moderation effect for task time between the experimental condition and the construct *awareness*. That is, participants find the robot to be more aware in the incremental condition – but the degree to which they find the robot aware is dependent on how fast they are able to solve tasks.

For study 2, we observe the same effect between experiment condition and task time. However, for these participants, all constructs based on questionnaire items are moderated by how much water they consume. That is, incrementality does impact these constructs positively, but the degree to which this applies is moderated by how persuasive the robot was in terms of explaining the health benefits of water intake.

Construct	α		DG Rho		Eigenvalue 1st		Eigenvalue 2nd	
	Study 1	Study 2	Study 1	Study 2	Study 1	Study 2	Study 1	Study 2
Awareness	0.77	0.88	0.87	0.93	2.06	2.44	0.30	0.61
Discomfort	0.67	0.70	0.79	0.80	2.31	2.43	1.61	1.29
Competence	0.83	0.73	0.87	0.92	3.22	2.63	1.05	1.10
Warmth	0.83	0.65	0.89	0.82	3.02	2.20	0.89	1.19
Time	0.32	0.49	0.75	0.80	1.19	1.33	0.81	0.67

TABLE II
ASSESSMENT OF THE MEASUREMENT MODEL

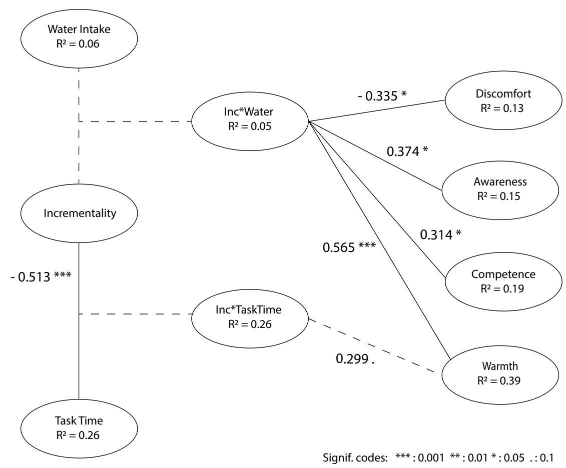

Fig. 3. Path Model for Study 2

VI. DISCUSSION

In this study, we aimed to reach a better understanding of the way people perceive a robot that uses incremental speech. Our first hypothesis (1a) predicted that participants in the incremental condition completed tasks faster than participants in the control conditions. This hypothesis is confirmed in both studies and aligns well with previous works (e.g. [11, 10, 14]).

The next set of hypotheses concerned a positive relation between performance (task time, 1b) and perceived robot competence and awareness (1c). We found no support for 1b in either study, however, we did find support for 1c in the first study (though not in the second).

The next two hypotheses suspected a relationship between individual task performance and perception of the robot in terms of warmth (1d) and discomfort (1e). We found no support for either hypothesis in any of the studies, although we did observe a marginal statistical effect on warmth in study 2.

With regards to persuasion, we found no support for the hypothesis (2a) that incrementality has a direct effect on persuasion (measured as water intake). However, for study 2 we observed that water intake served as a moderating factor on the affective factors competence (2b), awareness (2c), warmth (2d) and discomfort (2e). Thus, we find support for these hypotheses, but only in the second study. Like in the case of the ratings of the robot as competent and aware, which were mediated by the actual task time, also the perception of the robot was thus influenced by the extent to which it actually was persuasive.

VII. CONCLUSIONS

In this paper, we investigated factors that influence the perception of incremental robot responses to participant behaviors in two studies with two different populations. For both studies we observed that incremental speech processing has a direct effect on task time. The first study, which was conducted mostly with university students and academics, reveals task time as a moderating effect of incrementality on perceived awareness. The second study, which was entirely conducted with nurses, reveals a moderating effect of water intake on each of the four affective constructs. These results show that the perception of incrementality is mediated by the success of the incrementality in terms of task efficiency and robot persuasiveness. That the way incremental behavior is perceived is mediated by effectiveness (in terms of task completion time and persuasiveness) can explain the conflicting results of previous studies on the value of incremental processing.

REFERENCES

[1] G. Skantze and A. Hjalmarsson, "Towards incremental speech generation in dialogue systems," in *Proceedings of the 11th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, Association for Computational Linguistics, 2010, pp. 1–8.

[2] ——, "Towards incremental speech generation in conversational systems," *Computer Speech & Language*, vol. 27, no. 1, pp. 243–262, 2013.

[3] S. Andrist, X. Z. Tan, M. Gleicher, and B. Mutlu, "Conversational gaze aversion for humanlike robots," in *Proceedings of the 2014 ACM/IEEE int. conference on Human-robot interaction*, ACM, 2014, pp. 25–32, ISBN: 1450326587.

[4] Y. Kato, T. Kanda, and H. Ishiguro, "May i help you?- design of human-like polite approaching behavior," in *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, IEEE, 2015, pp. 35–42.

[5] K. Pitsch, H. Kuzuoka, Y. Suzuki, L. Sussenbach, P. Luff, and C. Heath, "“the first five seconds”: Contingent stepwise entry into an interaction as a means to secure sustained engagement in hri," in *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*, IEEE, 2009, pp. 985–991.

[6] Á. Castro-González, J. C. Castillo, F. Alonso-Martin, O. V. Olortegui-Ortega, V. González-Pacheco, M. Malfaz, and M. A. Salichs, "The effects of an impolite vs. a polite robot playing rock-paper-scissors," in *International Conference on Social Robotics*, Springer, 2016, pp. 306–316.

[7] M. Chromik, B. Carlmeyer, and B. Wrede, "Ready for the next step?: Investigating the effect of incremental information presentation in an object fetching task," in *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, ACM, 2017, pp. 95–96.

[8] B. Carlmeyer, D. Schlangen, and B. Wrede, "Look at me!: Self-interruptions as attention booster?" In *Proceedings of the Fourth International Conference on Human Agent Interaction*, ACM, 2016, pp. 221–224.

[9] I. de Kok, J. Hough, F. Hülsmann, M. Botsch, D. Schlangen, and S. Kopp, *A multimodal system for real-time action instruction in motor skill learning*, Conference Paper, 2015. DOI: `10.1145/2818346.2820746`.

[10] C. Kennington, S. Kousidis, T. Baumann, H. Buschmeier, S. Kopp, and D. Schlangen, "Better driving and recall when in-car information presentation uses situationally-aware incremental speech output generation," in *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, ACM, 2014, pp. 1–7.

[11] F. Ghigi, M. Eskenazi, M. I. Torres, and S. Lee, "Incremental dialog processing in a task-oriented dialog," in *Fifteenth Annual Conference of the International Speech Communication Association*, 2014.

[12] R. Manuvinakurike, M. Paetzel, C. Qu, D. Schlangen, and D. DeVault, "Toward incremental dialogue act segmentation in fast-paced interactive dialogue systems," in *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 2016, pp. 252–262.

[13] B. Carlmeyer, D. Schlangen, and B. Wrede, "Towards closed feedback loops in HRI: Integrating inprotk and pamini," in *Proceedings of the 2014 Workshop on Multimodal, Multi-Party, Real-World Human-Robot Interaction*, ACM, 2014, pp. 1–6.

[14] T. Baumann and F. Lindner, "Incremental speech production for polite and natural personal-space intrusion," in *International Conference on Social Robotics*, Springer, 2015, pp. 72–82.

[15] H. Buschmeier, T. Baumann, B. Dosch, S. Kopp, and D. Schlangen, "Combining incremental language generation and incremental speech synthesis for adaptive information presentation," in *Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, Association for Computational Linguistics, 2012, pp. 295–303.

[16] V. Tsai, T. Baumann, F. Pecune, and J. Casell, "Faster responses are better responses: Introducing incrementality into sociable virtual personal assistants," in *Proceedings of the 2018 International Workshop on Spoken Dialog System Technology*, 2018.

[17] J. Hough, I. de Kok, D. Schlangen, and S. Kopp, "Timing and grounding in motor skill coaching interaction: Consequences for the information state," in *Proceedings of the 19th SemDial Workshop on the Semantics and Pragmatics of Dialogue (goDIAL)*, 2015.

[18] K. W. Kobberholm, K. S. Carstens, L. W. Bøg, M. H. Santos, S. Ramskov, S. A. Mohamed, and L. C. Jensen, "The influence of incremental information presentation on the persuasiveness of a robot," in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, 2020, pp. 302–304.

[19] R. B. Cialdini, *Influence*. A. Michel Port Harcourt, 1987, vol. 3.

[20] C. M. Carpinella, A. B. Wyman, M. A. Perez, and S. J. Stroessner, "The robotic social attributes scale (rosas): Development and validation," in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, ACM, 2017, pp. 254–262.

[21] A. Pajaziti, "Slam–map building and navigation via ros," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 2, no. 4, pp. 71–75, 2014.

[22] K. Fischer, J. Seibt, R. Rodogno, M. K. Rasmussen, A. Weiss, L. Bodenhagen, W. K. Juel, and N. Krüger, "Integrative social robotics hands-on," *Interaction Studies*, vol. 21, no. 1, pp. 145–185, 2020.

[23] R Core Team, *R: A language and environment for statistical computing*, Vienna, Austria: R Foundation for Statistical Computing, 2020. [Online]. Available: `https://www.R-project.org/`.

[24] G. Sanchez, L. Trinchera, and G. Russolillo, *Plspm: Tools for partial least squares path modeling (pls-pm)*, R package version 0.4.9, 2017. [Online]. Available: `https://CRAN.R-project.org/package=plspm`.

ACKNOWLEDGMENT

The work reported on here was partially funded by Innovationsfonden in the framework of the Smooth project. We are furthermore much indebted to Ellen Langager from the Sønderborg Kommune LivingLab.