



OPEN

## Robots facilitate human language production

Olga A. Wudarczyk<sup>1✉</sup>, Murat Kirtay<sup>2</sup>, Doris Pischedda<sup>3,4</sup>, Verena V. Hafner<sup>2</sup>, John-Dylan Haynes<sup>3,4,5</sup>, Anna K. Kuhlen<sup>1,6</sup> & Rasha Abdel Rahman<sup>1,5,6</sup>

Despite recent developments in integrating autonomous and human-like robots into many aspects of everyday life, social interactions with robots are still a challenge. Here, we focus on a central tool for social interaction: verbal communication. We assess the extent to which humans co-represent (simulate and predict) a robot's verbal actions. During a joint picture naming task, participants took turns in naming objects together with a social robot (Pepper, Softbank Robotics). Previous findings using this task with human partners revealed internal simulations on behalf of the partner down to the level of selecting words from the mental lexicon, reflected in partner-elicited inhibitory effects on subsequent naming. Here, with the robot, the partner-elicited inhibitory effects were not observed. Instead, naming was facilitated, as revealed by faster naming of word categories co-named with the robot. This facilitation suggests that robots, unlike humans, are not simulated down to the level of lexical selection. Instead, a robot's speaking appears to be simulated at the initial level of language production where the meaning of the verbal message is generated, resulting in facilitated language production due to conceptual priming. We conclude that robots facilitate core conceptualization processes when humans transform thoughts to language during speaking.

Recent developments in artificial intelligence have introduced autonomous and human-like robots into numerous aspects of everyday life. Natural social interactions with robots are however still far from expectations, emphasizing the need to advance human-robot social interaction as one of the currently most pressing challenges of the field of robotics<sup>1</sup>. In this study we focus on the increasingly more prevalent domain of interaction with robots: verbal communication<sup>2,3</sup>. We assess the extent to which a social robot's verbal actions, in social interaction with humans, are simulated and predicted, or in other words *co-represented*, and explore the consequences of robot verbal co-representation on human language production. We focus on a social humanoid robot (Pepper, Softbank Robotics). Social robots, as physical agents, in contrast to other robots (e.g. service robots), have been developed specifically for interaction with humans<sup>4</sup>.

**Co-representation of human and robot task partners.** In humans, co-representation is a central mechanism underlying human social interaction (e.g., Ref.<sup>5</sup>). Co-representation refers to the ability to represent the partner's action alongside one's own actions<sup>6</sup>. This supports the understanding of the partner's intentions and enables the achievement of shared task goals. In social settings, humans seem to automatically co-represent the actions of their interaction partners (for reviews see, e.g., Refs.<sup>7-9</sup>). By observing a task partner's action, corresponding motor representations automatically activate in the observer, allowing predictions of the partner's behaviour and supporting coordination of the partner's action with one's own action<sup>9</sup>.

To achieve smooth human-robot interactions, humans and robots should be able to mutually predict and simulate each other's behaviour<sup>10</sup>. Despite early findings suggesting that humans treat social technologies similarly to real people<sup>11</sup>, the intricacies of robot co-representation present a more complex picture (for a review see Ref.<sup>10</sup>). Early investigations on this topic predominantly focused on motor co-representation of non-human agents and suggested that the shared representational system is predominantly tuned to other humans<sup>12,13</sup>. Yet subsequent studies, exploring the additional role of beliefs about the interaction partner, revealed that motor

<sup>1</sup>Department of Psychology, Neurocognitive Psychology, Humboldt-Universität zu Berlin, Berlin, Germany. <sup>2</sup>Department of Computer Science, Adaptive Systems Group, Humboldt-Universität zu Berlin, Berlin, Germany. <sup>3</sup>Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Bernstein Center for Computational Neuroscience, Berlin, Germany. <sup>4</sup>Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin Center for Advanced Neuroimaging, Berlin, Germany. <sup>5</sup>Faculty of Philosophy, Berlin School of Mind and Brain, Humboldt-Universität zu Berlin, Berlin, Germany. <sup>6</sup>These authors jointly supervised this work: Anna K. Kuhlen and Rasha Abdel Rahman. ✉email: wudarczyo@hu-berlin.de

co-representation of non-biological agents emerges if participants attribute *intentionality* to the agent, that is, the belief that the agent is acting in an active and intentional way<sup>14–17</sup>. Investigations in which participants completed real-life joint tasks with *specifically* humanoid robots corroborated these conclusions revealing that humans can co-represent a humanoid robot in a joint task (Social Simon Task<sup>18</sup>)<sup>19,20</sup>.

Co-representation, however, is not limited to lower levels of action control such as movements but extends to higher cognitive levels such as verbal communication. For instance, while our partner speaks, we covertly simulate and predict our partner's utterances<sup>21</sup> which in turn, affects our own verbal behavior<sup>22–26</sup>. The understanding of the cognitive underpinnings of a social robot's verbal co-representation is particularly important for the successful introduction of social robots across the range of verbal contexts, including daily interactions with social robots via speech (e.g., Ref.<sup>27</sup>), clinical contexts (e.g., Ref.<sup>28</sup>), learning environments (e.g., Ref.<sup>29</sup>) and social robots as social companions (e.g., Ref.<sup>30</sup>). Previous studies have not yet directly assessed the question of a social robot's verbal co-representation in a joint naming task setting. However, studies on interactive situations with artificial agents provide important insights into language processing and language prediction. Firstly, while it was shown that humans communicate differently when they believe that their interaction partner is a computer (and not another human<sup>31</sup>), such differences seem to vanish with additional social cues<sup>31</sup>, and when interacting with human-like avatars in virtual reality settings (see Ref.<sup>32</sup>, for a review). Further, in virtual reality settings, sentence processing<sup>33</sup> as well as the prediction of upcoming words were comparable in interactions with another human and human-like virtual agent<sup>34,35</sup>. With regards to humanoid robots, it was shown that humans adapt to their robot interaction partners by choosing words for everyday objects that match those of their artificial partner, a phenomenon referred to as lexical entrainment, reflecting a shift in language representations due to partner's verbal behaviour<sup>36–38</sup>. This shift in language representation was found to persist even after the human–robot interaction ended (Ref.<sup>38</sup>, also note other forms of linguistic alignment for other types of artificial agents, e.g. Refs.<sup>39–42</sup>). At the same time, other studies report differences between human and social robot (specifically, the NAO robot) interaction partners with regards to peer-pressuring speakers to commit morphologically primed grammatical mistakes (i.e. morphological alignment<sup>43</sup>).

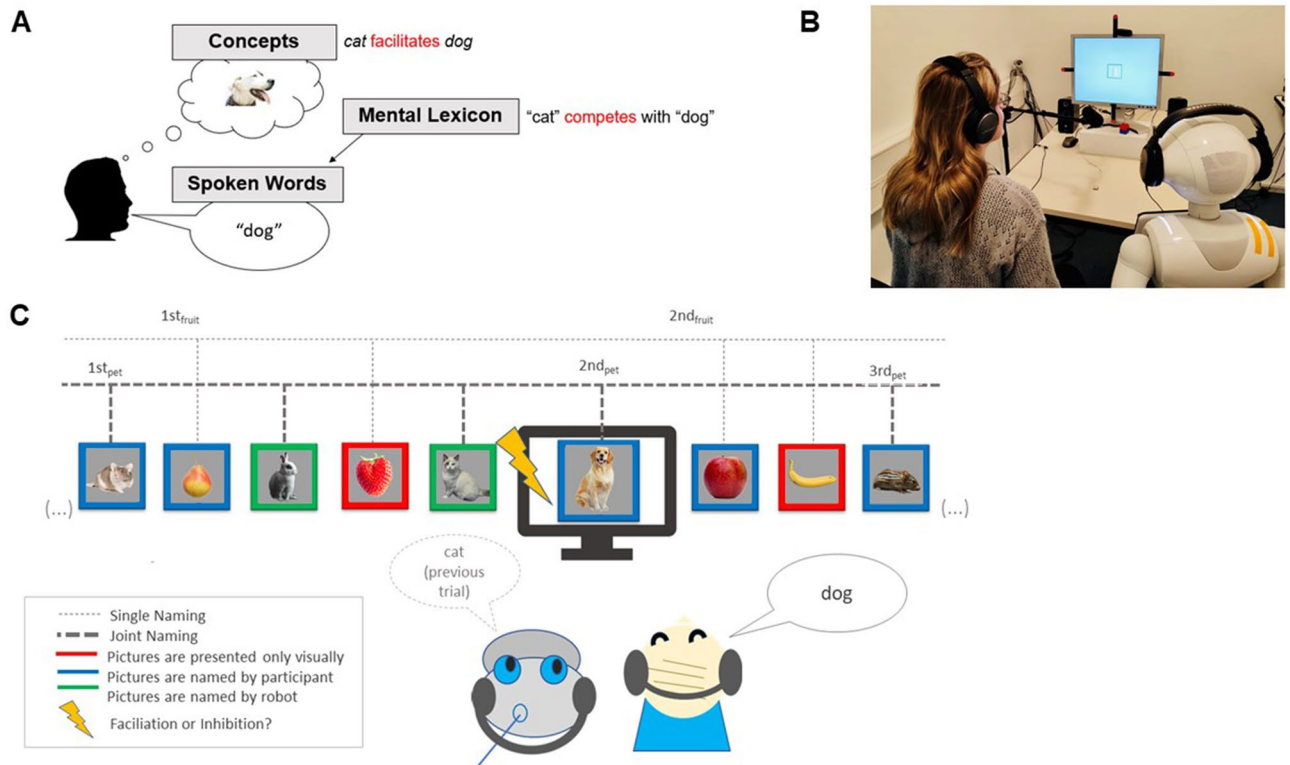
These lines of research provide the first evidence that a robot interaction partner can have an influence on a human's verbal behavior. What is not yet clear, however, is the extent to which such adaptations reflect a simulation of the partner's verbal behavior via co-representation, as well as the nature of such a robot co-representation. Hence, in the current study we specifically address the question of whether and how a social robot's verbal behaviour is simulated and thus co-represented in a joint language production task setting, and we explore the consequences of a robot's verbal co-representation on human's language production.

**Transforming thoughts to spoken words: the mechanisms of human language production.** To assess the question of robot verbal co-representation, in the current study we build on theoretical and empirical insights from psycholinguistic research on speaking. In the following paragraphs, we describe how preverbal thoughts are transformed into articulated speech that makes our thoughts accessible to others, highlighting the core mechanisms of language production targeted in the current study.

Speaking begins with thinking—the conceptualisation of the speaker's communicative intention with its meaning facets and associations<sup>44–48</sup>. The concepts and meaning facets are mapped onto corresponding words stored in the mental lexicon, from which co-activated words that best express the meaning of the intended message are selected. This process of meaning-based lexical selection is at the core of any act of speech production. Subsequently, the message is phonologically encoded and articulated.

While different theoretical alternatives have described lexical selection, they all share two assumptions: (1) speaking includes conceptualization at the message level and selection at the lexical level, and (2) conceptually related words are co-activated at the lexical level before the best-fitting words are selected (e.g., Refs.<sup>46,48–52</sup>). To illustrate this process, consider the course of a conversation about pets during which a speaker may activate the concept *dog* and related concepts (e.g., *cat*, *rabbit*), resulting in the co-activation of words representing the concepts of *dog* and other furry mammals such as *cat* and *rabbit*, among which the most appropriate word (“dog”) is then selected (see Fig. 1a). Facilitation (i.e., faster word production) emerges as a consequence of semantic priming during conceptualization (e.g., thinking about *cats* facilitates thinking about *dogs*). Yet, interference (i.e., slower word production) emerges as a consequence of competition among related and co-activated words in the mental lexicon (e.g., the word “cat” competes with the word “dog” for selection; e.g., Refs.<sup>46,48</sup>; for alternative views, see Refs.<sup>51,52</sup>) (Fig. 1a). Depending on the context and task demands, conceptual facilitation or lexical interference may dominate, resulting in overall facilitated or inhibited speaking latencies<sup>53,54</sup>.

In the present study, we employed the Continuous Picture Naming (CPN) Paradigm, a principal task that provides insights into these mechanisms, underlying the retrieval of words for speaking from the mental lexicon<sup>55–58</sup>. In this task, participants name pictures while the speeds at which they initiate speaking are closely monitored (see “Materials and methods”). While the pictures are presented in a seemingly random sequence, the critical context manipulation consists of the presentation of related objects belonging to a common semantic category, separated by a variable number of unrelated objects (see Fig. 1c). The repeated presentation of related objects should result in facilitation at the level of conceptualization due to semantic priming (e.g., Ref.<sup>57</sup>). However, due to competition for selection among related words at the level of the mental lexicon, this should also result in interference and therefore slower naming. For instance, if the cat has been named before and is co-activated upon naming a dog, the word “cat” strongly competes with the word “dog” for selection. This competition is increased by each additional member of the category of animals being named, resulting in an increase of naming times with each new category member named (i.e. with each Ordinal Position of the category's member). As a consequence, the repeated naming of objects from the same semantic category results in increasingly more effortful lexical selection, a robust empirical effect referred to as cumulative semantic interference.



**Figure 1.** (a) The cognitive model of lexical access during speech production explains how concepts activate entries in the mental lexicon that are then translated into speech. Facilitation and competition can occur during speech production on different levels (conceptual vs. lexical). (b) We operationalized the cognitive model in our experimental setup. A participant and the robot (Pepper, Softbank Robotics) sat next to each other in front of the computer screen, displaying pictures, some of which were semantically related. We assessed whether the robot's verbal behaviour is simulated, and at which processing level (conceptual vs. lexical). To exclude a possible confound of additional linguistic input during Shared Naming with the robot, both the participants and the robot wore headphones that masked the partner's naming. (c) Pictures were continuously displayed on the computer screen. In the single naming condition, half of the pictures within a given semantic category (e.g., fruits) were named by participants, the other half were presented only visually. In the joint naming condition half of the pictures within a given semantic category (e.g., pets) were named by participants, the other half by the robot partner. Ordinal positions 1–5 were counted for those pictures named by the participants.

**Co-representation during joint language production.** More recently, the CPN task has also been employed to shed light on how human communication partners are co-represented during joint picture naming<sup>25,26</sup>. In these studies, two participants took turns naming pictures. The results of these studies revealed that naming times slow down not only for each semantic category member named by the participant (i.e. cumulative semantic interference, as reflected by a main effect of Ordinal Position), but also, crucially, for each semantic category member named by the partner (i.e. Ordinal Position  $\times$  Naming Condition interaction). This partner-elicited semantic interference suggests the occurrence of the simulation and co-representation of the partner's verbal behaviour at the level of lexical selection.

Crucially, variations of this effect allow insights into the partner's co-representation at the level of conceptualization or at the level of lexical selection. As described above, semantic co-activation leads to facilitatory priming effects during conceptualization and inhibitory effects of competition during lexical selection. While the interference caused by the partner's naming indicates that individuals co-represent their human partners by simulating their partner's verbal behavior down to the level of lexical access<sup>26</sup>, the critical question that remains to be answered is: Is a robot partner's verbal behaviour simulated in a similar way?

**Current study.** To answer this question thirty-six young adults were introduced to a socially interactive humanoid robot—Pepper (Softbank Robotics)—and together with the robot completed the CPN task<sup>26</sup>, see Fig. 1b. The participants named a stream of pictures, embedded in which were objects of the same semantic categories. Some of these categories were named together with the robot partner, others were named solely by the participant (see Fig. 1c). In both instances half of the objects presented within each category were named by the participant and the other half were not named by the participant. Instead, these were either named by the robot partner (Shared Naming) or presented visually, but neither named by the participant nor the robot (Single Naming). The robot was physically present and stood directly next to the participants throughout the experiment. Crucially, to capture a simulation of the partner's utterances (and not merely the representation of what

the partner is saying based on the perceptual evidence) participants wore noise-canceling headphones (Fig. 1b). Thus, participants were aware that the robot was speaking but they could not hear what the robot was saying (for a similar procedure, see Ref.<sup>26</sup>, Experiment 3). At the end of each session, participants completed questionnaires including ratings regarding perceived robot intentionality (see “Materials and methods”, and Supplementary Information for details).

We assessed whether a robot’s verbal behaviour is co-represented in the CPN task, and how it affects language production (“Robot as a task partner experiment”). We put forward the following hypotheses (pre-registered at <https://aspredicted.org/va9nr.pdf>): if the robot is co-represented similar to a human task partner, we expect participants’ naming latencies to increase not only in response to the number of semantically related pictures participants previously named themselves (Single Naming), but also, additionally, in response to semantically related pictures named by the robot (i.e. as in Ref.<sup>26</sup>, Naming Condition  $\times$  Ordinal Position Interaction during Shared Naming). The steeper increase for categories co-named with the robot would reveal that objects named by the robot elicit comparable lexical processes to naming the objects by oneself, and would thus reveal a co-representation of a robot partner down to the level of lexical selection (as in Ref.<sup>26</sup>). If, however, the robot’s verbal behaviour is not co-represented during the task, we expect no partner effects of the robot’s naming on the participants’ naming times.

In addition, though not previously predicted nor pre-registered by us, a main effect of Naming Condition (i.e. Single Naming vs. Shared Naming) would indicate that a robot’s naming is co-represented in the task, but differently than that of a human partner. For example, if a robot’s verbal co-representation occurs at a less deep and concrete verbal level, that is, only during the stage of conceptualization but not during lexical selection, facilitation of naming responses by the robot’s co-naming of semantic categories would be the result, as during conceptualization semantic priming instead of competition dominates (Refs.<sup>53,54</sup>, Fig. 1a).

As previous studies emphasized perceived robot intentionality as an important factor necessary for robot co-representation (for a review see Ref.<sup>10</sup>), we additionally explored whether the effects are modulated by perceived robot intentionality (“Robot as a task partner experiment: intentionality”). Lastly, we compared the current data where the robot acted as the naming partner with data where a human acted as the naming partner (Robot as a task partner and human as a task partner experiments). This analysis was achieved by pooling the current data with the data from Experiment 3 of Ref.<sup>26</sup>. This analysis was not previously pre-registered.

## Materials and methods

**Pre-registration.** The experimental procedures, hypotheses, data exclusion criteria, and analyses approaches were pre-registered prior to data collection (<https://aspredicted.org/va9nr.pdf>).

**Participants.** Thirty-six native speakers of German (26 females, 10 males) aged between 19 and 35 (mean = 25.6, SD = 4.6) were included in the data analyses. The sample size was determined via a-priori power analysis. Based on the effect sizes observed in the previous study (Ref.<sup>26</sup>, Experiment 3), we simulated the outcome of the anticipated LMM with 1000 iterations (R Simr package). With 36 participants, we reached a power estimate of 81.90% chance (95% confidence interval: 79.37, 84.24) for detecting an interaction between Ordinal Position and Naming Condition.

Participants with an error rate greater than 20% were excluded and replaced ( $n = 3$ ). Additionally, we excluded and replaced a participant ( $n = 1$ ) who, at the end of the experiment (see “Auditory input cancellation”), reported hearing a word named by the robot.

All participants gave written informed consent and were reimbursed or received course credit. The study was approved by the ethics committee of the Department of Psychology at Humboldt-Universität zu Berlin. The study was conducted in accordance with the Declaration of Helsinki (2013).

**Robot.** The robot used in the experiment as the task partner was Pepper (Softbank Robotics): a 1.2 m-tall humanoid robot designed for social interaction with humans.

**Stimuli.** The experimental stimulus-set of the continuous Picture Naming Paradigm (CPN) was derived from the previous study<sup>26</sup>. Namely, 320 photographs of natural or man-made objects mapping onto 32 different semantic categories (e.g., flowers, buildings, birds), containing 10 exemplars each, were used. In addition, 120 items, unrelated to the target categories, were used as fillers. The pictures were presented on a grey background and were 3.5 cm  $\times$  3.5 cm in size.

**Design.** Stimulus lists specifying the order in which pictures were named were derived from Ref.<sup>26</sup>. Lists were individually created for each participant considering the following rules: (1) The order of items within one category was randomly selected (by the program ‘Mix’<sup>59</sup>); (2), the items of one category were randomly separated by a minimum of two and a maximum of six unrelated items (i.e., items of a different category or filler items); (3) categories belonging to the same superordinate category did not overlap within the list (e.g., birds and hoofed mammals, merging to superordinate category animals were not used in the same list).

Participants named half of the items of a given category. The other half of the items were either named by the robot (Shared Naming), or by nobody, and displayed visually only (Single Naming). The assignment of trials in which the participant, the robot, or nobody named the objects was random, with the following exceptions: the first and the last items of the same category were always named by the participant; participant’s trials were separated by maximum three trials in which either the robot or nobody named the objects. All filler items were assigned to participants’ trials.

Each participant was presented with the complete set of pictures (440 trials) twice (two experimental blocks) thus, each participant completed 880 trials. Objects assigned to participants in the first block, were the same as in the second block, however, the individual exemplars appeared in a different ordinal position.

**Auditory input cancellation.** To exclude the possibility of a confound of receiving an additional linguistic (auditory) input in the Shared Naming Condition as compared to the Single Naming Condition, throughout the experiment, the participants wore noise cancelling headphones (Bose QuietComfort 25) continuously displaying pink noise (similarly to Ref.<sup>26</sup>, Experiment 3). To enhance the impression that the robot would also not hear the participant, the robot also wore headphones.

To not eliminate the impression that the robot is naming the words during the experiment, we aimed to have the participants hear that the robot speaks, but not have them hear the exact words that the robot is saying. To achieve this goal, the noise level was individually tested and adjusted for each participant. Before the start of the experiment, participants underwent a procedure in which the robot named seven words, one of which was an animal, and the participant had to indicate what animal was named by the robot. In case participants were able to hear an animal, the procedure was repeated with an increased volume of Pink Noise until they were no longer able to specify an animal named by the robot. The sound check was repeated once again in the middle of the experiment and 5 times at the end of the experiment. In case that participants were able to hear an animal named by the robot, in at least one of the final 5 trials at the end of the experiment, participants' data were excluded from the analyses ( $n = 1$ ) and replaced.

**Procedure.** The experimental task instructions are available at: <https://osf.io/qakbv/>.

At the beginning of each experimental session, the robot was introduced to the participants as their task partner. The robot was introduced as a humanoid and quite communicative robot. Participants were told that the robot can recognize them and follow them with its gaze. Consistently, they were shown how: (1) the robot can follow the *experimenter* with its gaze, (2) the robot can follow the *participant* with its gaze. Furthermore, as participants sat at the chair (and were encouraged to say, Hello' to the robot), the robot looked directly at the participant and said: 'Hello I am Pepper. I am learning German and I am learning to name different objects. We will now complete a task together'. The emphasis on learning and object recognition, was aimed at enhancing the impression of intentionality behind robot's behaviour, which consistently with previous research is linked to observation of social effects in experimental studies (for a review see Ref.<sup>10</sup>).

Following robot introduction, in a separate room, participants were given time (approximately 10 min) to familiarize themselves with the pictures that were going to be presented in the task and their target names. The pictures were presented on paper in an unsorted order.

During the experiment, participants were seated next to the robot in front of a computer screen displaying the pictures. The robot's speech and behaviour were generated via the NAOqi framework with custom scripts that communicate between the Pepper robot and the Presentation software that visually displayed the stimuli. Pictures appeared on the computer screen, one at a time. A colored frame around the picture indicated whose turn it was to name the picture: the participant's, the robot's or nobody's. The color assignment was counterbalanced across participants. Participants were instructed to name the pictures as fast and as accurately as possible in response to their ascribed color. In the remaining trials, participants were instructed not to take any action. Participants first completed a 15-trial practice session. Subsequently, participants completed two experimental blocks. Each block consisted of 440 trials, and participants were given a short break after 220 trials, and a longer break after the first block. To highlight robots' natural behaviour, during the breaks, the robot was set to respond to and to attend to its environment. During the CPN task, the robot exhibited subtle arm and head movements (instead of being fully static during picture naming).

Each trial began with a presentation of the fixation cross for 0.5 s. In the trials in which participants named the objects, the picture was presented until the naming response was initiated or for a maximum of 2 s. Participants' naming latency (reaction time) was recorded with the help of a voice-key from the onset of the picture presentation. In the trials in which the robot named the pictures, robot's naming response times were variable and were modelled after average response times for a given picture recorded in<sup>26</sup> and Experiment 1 of<sup>60</sup>. To keep the control condition (in which nobody names the pictures) comparable, the pictures in this condition were also modelled to disappear after the duration of average naming latencies for the given objects based on response times recorded in Ref.<sup>26</sup> and Experiment 1 of Ref.<sup>60</sup>. Following each Naming Condition, a blank screen of 2 s followed each picture presentation, after which the next trial followed.

Throughout the experiment, the experimenter coded failures of the voice key as well as erroneous trials (wrong picture naming, naming at the wrong turn).

At the end of the experiment, participants completed two questions pertaining to perceived robot intentionality (derived from Ref.<sup>19</sup>): "The robot acted intentionally" and "The robot decided actively when to respond to a stimulus" on a scale from 1 (not at all) to 11 (extremely). Additionally, they completed a battery of questionnaires including robot perception (Goodspeed Questionnaire<sup>61</sup>; RoSAS<sup>62</sup>; HRIES<sup>63</sup>), attitudes towards technology questionnaire (ATAI<sup>64</sup>) and three questions regarding previous experience with AI, Robots and Pepper specifically (see Supplementary Information for an overview). Finally, participants were debriefed and compensated for their participation.

**Statistical analyses.** *Confirmatory analyses.* Firstly, we assessed the effects of having a robot as a task partner (versus naming the pictures alone) on naming latencies to semantically related pictures.

LMMs, as implemented in the lmer function of the lme4 package for R<sup>65</sup> with random effects modeled for participants and items were applied to naming latencies treated with log transformation (applied following

the Box-Cox procedure<sup>66</sup>). Naming latencies for the trials in which the participants named the pictures were modelled as a function of the predictors: Naming Condition (Shared Naming vs. Single Naming), Ordinal Position, (ordinal position 1–5), and Experimental Block (1–2). The predictors Naming Condition and Block were contrast-coded using the sliding difference contrast. The predictor Ordinal Position was centered and entered as a continuous variable. Models were initially run with a maximum random effects structure allowed by the experimental design for participants and items<sup>67</sup>. Using singular value decomposition, the initial full random effect structure was simplified until the maximal informative and converging model was identified.

**Exploratory analysis.** To assess whether perceived robot intentionality affected the degree to which participants would be influenced by their robot partner's naming response, we additionally carried out an LMM adding the perceived robot intentionality score as a predictor to the original model. As in Ref.<sup>19</sup>, perceived intentionality score was calculated by averaging the responses to the questions: “The robot acted intentionally” and “The robot decided actively when to respond to a stimulus”, as the two items appeared to measure the same construct (Cronbach's  $\alpha=0.89$ ). The predictor Intentionality (1–11) was subsequently centered and added to the original model as a continuous variable.

**Additional analysis: comparing robot and human partner.** To compare the effects of naming the semantic categories together with a robot partner to naming the semantically related objects together with a human partner, we pooled the current data with the data obtained from Experiment 3 in Ref.<sup>26</sup>. In this experiment with human task partners, 36 participants between the ages of 18 and 36 (mean 25.14 years; 9 males, 27 females) participated<sup>26</sup>. The two experiments differed in the following aspects: the experiment with human partners<sup>26</sup> consisted of 1 block only (i.e., 440 trials). Thus, to make the two datasets comparable, for this analysis, we only included the data from block 1 of the current experiment (i.e., 440 trials). While the timing of the disappearance of the objects in the solo condition in the current experiment was modelled after naming responses observed in the previous studies utilizing the paradigm (Ref.<sup>26</sup>, Experiment 1<sup>60</sup>), the timing of disappearance in the experiment with human partners (Experiment 3<sup>26</sup>) reflected average reaction times recorded for Experiment 1 of the same study. The efficacy of the headphones was assessed 3 times in the current experiment (prior to experiment, after block 1, after block 2, whereas in the experiment with human partners, it was assessed only prior to the experiment. While in the experiment with human partners, participants who reported understanding the words named by the partner in the experiment were excluded from the analysis, in the current experiment we used an improved procedure to assess whether the participants were able to hear the words named by the robot to exclude participants based on auditory input (see “Auditory input cancellation”).

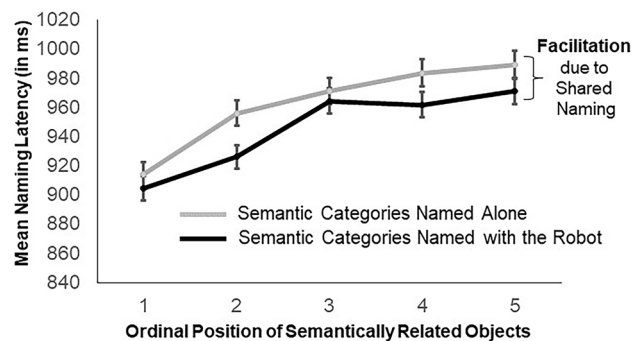
For LMM analysis, the predictor Experiment was contrast-coded using the sliding difference contrast and added to the maximal LMM model. Using singular value decomposition, the initial full random effect structure was simplified until the maximal informative and converging model was identified. This analysis was not previously pre-registered.

## Results

**Robot as a task partner experiment.** Participants' naming latencies in the CPN task increased with each ordinal position within semantically related categories by an average of 18 ms. A main effect of Ordinal Position confirmed the cumulative semantic interference effect, demonstrating that the paradigm produces the well-known effect also in the general presence of a robot partner. Crucially, we observed a 17.2 ms facilitation for semantic categories co-named with the robot (Shared Naming) as compared to the categories named alone (Single Naming). Specifically, naming latencies were faster for those semantic categories that were also named by the robot ( $M=945.8$ ,  $SD=282.1$ ) than for those semantic categories that were only named by the participant ( $M=963.0$ ,  $SD=293.4$ , Fig. 2). A main effect of Naming Condition confirmed the influence of having a robot as a task partner on facilitated naming. Lastly, participants named the pictures faster in block 2 ( $M=915.80$ ,  $SD=360.83$ ) as compared to block 1 ( $M=994.10$ ,  $SD=406.62$ ), as revealed by a significant main effect of Block. For an overview of Linear Mixed-effects Models (LMM) effects see Table 1.

**Robot as a task partner experiment: intentionality.** Participants' mean response to the question “The robot decided actively when to respond to a stimulus” was 6.6 ( $SD=3.5$ , range 1–11, 11 = very intentional) and to the question “The robot acted intentionally” it was 6.6 ( $SD=3.3$ , range 1–11, 11 = very intentional). Adding the average score of perceived robot intentionality to the LMM model did not result in a significant modulation of the effects.

**Comparing robot as a task partner and human as a task partner.** Across the pooled data set including the current experiment with a robot task partner and an experiment with a human task partner<sup>26</sup>, naming latencies increased with each ordinal position within semantically related categories by an average of 23 ms (see Fig. 3). The main effect of Ordinal Position confirmed the cumulative semantic interference effect in the two experiments. In addition, naming latencies were faster, in general, in the experiment in which the task partner was a robot ( $M=994.4$ ,  $SD=299.7$ , Block 1) as compared to the experiment in which the task partner was a human ( $M=1030.1$ ,  $SD=293.8$ ), as revealed by a main effect of the Experiment. Crucially, as revealed by a Naming Condition by Experiment interaction, participants were faster to name semantically related pictures with a robot partner ( $M=986.7$ ,  $SD=294.2$ ) compared to a human partner ( $M=1036.4$ ,  $SD=296.83$ ), while naming semantic categories alone was more comparable in the robot ( $M=1002.1$ ,  $M=305.24$ ) and the human ( $M=1023.8$ ,  $M=290.9$ ) experiments (see Fig. 3). The LMM analyses confirmed these effects (see Table 2 for a summary).



**Figure 2.** Naming latencies in the naming task with a robot partner. Mean naming latency and standard errors (in milliseconds) observed in the current experiment with a robot partner, broken down by ordinal position and naming condition. Faster naming latencies for semantic categories named with the robot (Shared Naming) as compared to semantic categories named alone (Single Naming) point to a facilitatory effect of sharing the naming task with the robot on speaking.

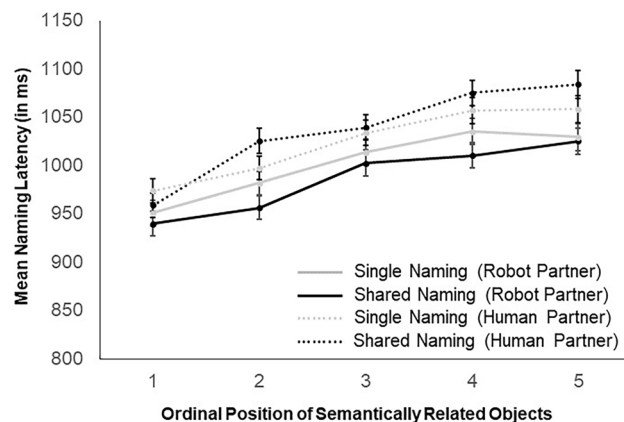
	Est/Beta	SE	df	t	p
<b>Fixed effects</b>					
Intercept	6.83	0.02	55.94	368.70	<0.001
Naming condition	0.02	0.00	296.9	3.60	<0.001
Ordinal position	0.02	0.00	10,080	11.61	<0.001
Block	-0.08	0.00	10,030	-19.19	<0.001
Naming condition × ordinal position	0.00	0.00	10,260	-0.48	0.632
Naming condition × block	0.01	0.01	10,120	0.76	0.447
Ordinal position × block	-0.01	0.00	10,120	-1.61	0.108
Naming condition × ordinal position × block	0.00	0.01	9806	0.50	0.614
	<b>Variance</b>	<b>Std. Dev</b>			
<b>Random effects</b>					
<b>Item</b>					
Intercept	0.02	0.15			
Naming condition	0.00	0.01			
<b>Subject</b>					
Intercept	0.01	0.10			
Residual	0.05	0.23			
<b>Model fit</b>					
Log likelihood	251.3				
Deviance	-502.7				
Model equation: $\log(\text{naming latencies}) \sim \text{naming condition} \times \text{ordinal position} \times \text{block} + (1   \text{subject}) + (\text{naming condition}   \text{item})$					

**Table 1.** LMM for the robot as a task partner experiment. Fixed-effect estimates, standard errors, degrees of freedom, t-values and p-values for the selected model of the Robot as a Task Partner Experiment's analyses, as well as estimates of the variance and square root (standard deviations) of the random effect structure and goodness-of-fit statistics.

## Discussion

The aim of this study was to assess whether humans co-represent a social robot's verbal behaviour in a joint verbal communication task, and to identify the consequences of robot verbal co-representation on language production. To this end, human participants named semantically related pictures together with a robot task partner. Half of the semantic categories underlying these pictures were named by the participant only, the other half was named together with the robot partner. We found facilitatory effects of sharing the naming task with the robot on language production, pointing to: (1) co-representation of the robot partner during a shared verbal task, and (2) facilitatory effects of robot verbal co-representation at the level of conceptualization, where the meaning of the intended verbal message is generated. Crucially, this is not a mere presence effect, but an effect specific to sharing a task with a robot partner, only seen for semantic categories that were co-named together with the robot.

Contrary to previous findings with human partners<sup>25,26</sup>, we did not observe enhanced cumulative interference due to shared naming together with the robot partner. Instead, naming semantic categories together with



**Figure 3.** Naming latencies in the naming tasks with a robot and a human partner. Mean naming latency and standard errors (in milliseconds) observed in the current experiment with a robot partner (Block 1) and in experiment 3 from Ref.<sup>26</sup> with a human partner, broken down by ordinal position and naming condition.

	Est/Beta	SE	df	t	p
<b>Fixed effects</b>					
Intercept	6.89	0.02	65.36	432.54	<0.001
Naming condition	0.00	0.00	9812	0.350	0.7262
Ordinal position	0.02	0.00	9899	13.488	<0.001
Experiment	0.05	0.00	9845	9.543	<0.001
Naming condition × ordinal position	0.00	0.00	9819	- 1.207	0.2275
Naming condition × experiment	- 0.02	0.01	9812	- 2.517	0.0119
Ordinal position × experiment	0.00	0.00	9827	1.463	0.1435
Naming condition × ordinal position × experiment	0.00	0.01	9820	- 0.550	0.5826
	<b>Variance</b>	<b>Std. Dev</b>			
<b>Random effects</b>					
Item (intercept)	0.02	0.15			
Subject (intercept)	0.01	0.08			
Residual	0.06	0.23			
<b>Model fit</b>					
Log likelihood	- 261.0				
Deviance	522.0				
Model equation: $\log(\text{naming latencies}) \sim \text{naming condition} \times \text{ordinal position} \times \text{experiment} + (1   \text{subject}) + (1   \text{item})$					

**Table 2.** LMM for the combined robot and human datasets. Fixed-effect estimates, standard errors, degrees of freedom, t- values and p-values for the selected model of the Robot as a Task Partner and Human as a Task Partner Experiments; estimates of the variance and square root (standard deviations) of the random effect structure and goodness-of-fit statistics are also presented.

the robot led to facilitation, as revealed by faster naming of semantic category members co-named with the robot. This difference was corroborated in analyses directly comparing the data from experiments in which the task partner was a robot and in which the task partner was a human: while naming was facilitated with a robot partner, naming was not facilitated by human partners, thus revealing the specificity of the facilitatory effect to sharing the task with a robot partner.

These findings demonstrate that a robot's verbal behaviour is co-represented, yet differently from that of humans. As introduced above, language production models distinguish between the meaning-based level of conceptualization and the lexical level of word selection (e.g., Refs.<sup>47–52</sup>). The co-activation of concepts may facilitate the generation of the message, whereas, in parallel, the co-activation of alternative words may inhibit the selection of the best-fitting word due to competition from related alternative words. The observation of a facilitation effect induced by a robot task partner suggests that robots are not co-represented down to the level of lexical selection (as humans do<sup>22,26</sup>). Instead, robot's language appears to be simulated only at the conceptual level, resulting in a facilitatory effect on language production due to priming. This more 'shallow' co-representation is also in line with previous findings on linguistic alignment showing that lexical entrainment happens for both human and robot partners, yet that the robot's influence is weaker than that of the human partner<sup>38</sup>. Since the



Continuous Picture Naming task has been employed with human and robot partners, we conclude that robots are co-represented, but the simulation seems to include only conceptualization on their behalf, and not lexical selection. This meaning-based simulation of the robot partner helps speakers to generate their own verbal message, which in turn facilitates speaking.

A note should be made that a related line of research<sup>24</sup> claimed that speakers do not co-represent the content of their partners' utterances, but instead they represent merely the act of each other's naming. We complement this line of research by demonstrating that the representation of conceptual and/or lexical content of the partner's task can be observed. Just as the baseline effects of semantic effects in this task, it is the previous experience of having named a related object or the experience of simulating conceptual or lexical processing on behalf of the task partner that elicits these effects in the present task. Thus, in comparison to the study by Ref.<sup>24</sup>, the current paradigm includes (implicit) memory of processing objects either directly or on behalf of the task partner. We have built our predictions on the above-mentioned language production models specifically assuming interference via accessing semantically related items at the level of the mental lexicon and facilitation at the level of concepts (e.g. Refs.<sup>53,54</sup>). Consistently, the observed effects (i.e. partner-elicited inhibitory effects with a human partner<sup>26</sup>; and facilitatory effects with a robot partner, observed here) are in line with these predictions, suggesting co-representation of the *content* of the partner's speech. The exact nature of this co-representation is still nonetheless unknown and should be addressed in future studies by, for example, employing physiological measures, such as EEG, specifically targeting the underlying mechanisms (also see Ref.<sup>60</sup>).

Future research should also reveal further insights into the critical differences and potential commonalities between human and robot verbal co-representation, utilizing between or within-subjects designs. Since in the current study we did not find a modulatory influence of perceived robot intentionality on robot verbal co-representation, this factor does not seem to play a major role. Future studies shall also systematically investigate whether the specifics of a robot's behavior, (synthetic) voice or visual appearance, as well as the intensity of verbal interactions or enhanced knowledge about the robot's attributes may be relevant factors affecting robot verbal co-representation. Although in the current study we performed repeated checks to assure that participants could not hear the exact words named by the robot (and thus that the effects were driven by a simulation of robot's speech rather than linguistic processes), future studies should consider having artificial agents produce phonologically acceptable non-words instead of having them name real words during the experiment.

To conclude, not only can the verbal behaviour of robots be simulated during HRI, sharing a verbal task with a robot also facilitates speaking. These results offer important insights into contexts in which verbal communication with robots is prevalent. The findings might also be of special importance for clinical populations where speech delays and lexical retrieval difficulties prevail, and in educational contexts where language production enhancement is particularly desired: verbal interactions with robots in these contexts might be particularly advantageous in facilitating spoken language production.

Representation of robot partners by humans has previously been confirmed across information processing levels which include: motor (for a review, see Ref.<sup>10</sup>), social (for a review, see Ref.<sup>68</sup>), and higher-order cognitive levels<sup>69</sup>. Our results provide a crucial extension to the increasingly relevant domain of robot verbal co-representation by providing insights into the nature of how robots are simulated in shared tasks, and how it affects human behavior. This advances our path towards characterizing the intricacies of real-world human–robot social interactions, which currently remains one of the grand challenges of the field of robotics<sup>1,70</sup>. The mechanisms of co-representation and language facilitation revealed here might be integrated into a cognitive architecture for social robots to enhance their communicative skills in verbal interactions with humans<sup>10</sup>.

## Data availability

The datasets generated and analyzed during the current study are available in the Open Science Framework repository: <https://osf.io/qakbv/>.

Received: 8 April 2021; Accepted: 19 July 2021

Published online: 18 August 2021

## References

1. Yang, G.-Z. *et al.* The grand challenges of science robotics. *Sci. Robot.* **3**, eaar7650 (2018).
2. Mavridis, N. A review of verbal and non-verbal human–robot interactive communication. *Robot. Auton. Syst.* **63**, 22–35 (2015).
3. Taniguchi, T. *et al.* Language and robotics. *Front. Robot. AI*. <https://doi.org/10.3389/frobt.2021.674832> (2021).
4. Hegel, F., Lohse, M. & Wrede, B. Effects of visual appearance on the attribution of applications in social robotics. In *RO-MAN 2009—The 18th IEEE International Symposium on Robot and Human Interactive Communication* 64–71 (2009).
5. Knoblich, G., Butterfill, S. & Sebanz, N. Psychological research on joint action. *Psychol. Learn. Motiv.* **54**, 59–101 (2011).
6. Sebanz, N., Knoblich, G. & Prinz, W. How two share a task: Corepresenting stimulus–response mappings. *J. Exp. Psychol. Hum. Percept. Perform.* **31**, 1234–1246 (2005).
7. Curioni, A., Minio-Paluello, I., Sacheli, L. M., Candidi, M. & Aglioti, S. M. Autistic traits affect interpersonal motor coordination by modulating strategic use of role-based behavior. *Mol. Autism* **8**, 1–13 (2017).
8. Wenke, D. *et al.* What is shared in joint action? Issues of co-representation, response conflict, and agent identification. *Rev. Philos. Psychol.* **2**, 147–172 (2011).
9. Sebanz, N. & Knoblich, G. Prediction in joint action: What, when, and where. *Top. Cogn. Sci.* **1**, 353–367 (2009).
10. Kirtay, M. *et al.* Modeling robot co-representation: State-of-the-art, open issues, and predictive learning as a possible framework. In *2020 Joint IEEE 10th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* 1–8 (2020).
11. Reeves, B. & Nass, C. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People* (Cambridge University Press, 1996).
12. Kilner, J. M., Paulignan, Y. & Blakemore, S. J. An interference effect of observed biological movement on action. *Curr. Biol.* **13**, 522–525 (2003).

13. Tsai, C.-C., Kuo, W.-J., Hung, D. L. & Tzeng, O. J. L. Action co-representation is tuned to other humans. *J. Cogn. Neurosci.* **20**, 2015–2024 (2008).
14. Atmaca, S., Sebanz, N. & Knoblich, G. The joint flanker effect: Sharing tasks with real and imagined co-actors. *Exp. Brain Res.* **211**, 371–385 (2011).
15. Liepelt, R., Cramon, D. & Brass, M. What is matched in direct matching? Intention attribution modulates motor priming. *J. Exp. Psychol. Hum. Percept. Perform.* **34**, 578–591 (2008).
16. Liepelt, R. & Brass, M. Top-down modulation of motor priming by belief about animacy. *Exp. Psychol.* **57**, 221–227 (2010).
17. Müller, B. C. *et al.* When Pinocchio acts like a human, a wooden hand becomes embodied. Action co-representation for non-biological agents. *Neuropsychologia* **49**, 1373–1377 (2011).
18. Sebanz, N., Knoblich, G. & Prinz, W. Representing others' actions: just like one's own?. *Cognition* **88**, B11–B21 (2003).
19. Stenzel, A. *et al.* When humanoid robots become human-like interaction partners: Corepresentation of robotic actions. *J. Exp. Psychol. Hum. Percept. Perform.* **38**, 1073–1077 (2012).
20. Strait, M. *et al.* A three-site reproduction of the joint simon effect with the NAO robot. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction* (ed. Belpaeme, T.) 103–111 (2020).
21. Pickering, M. J. & Garrod, S. An integrated theory of language production and comprehension. *Behav. Brain Sci.* **36**, 329–347 (2013).
22. Baus, C. *et al.* On predicting others' words: Electrophysiological evidence of prediction in speech production. *Cognition* **133**, 395–407 (2014).
23. Brehm, L., Taschenberger, L. & Meyer, A. Mental representations of partner task cause interference in picture naming. *Acta Psychol.* **199**, 102888 (2019).
24. Gambi, C., van de Cavey, J. & Pickering, M. J. Interference in joint picture naming. *J. Exp. Psychol. Learn. Mem. Cogn.* **41**, 1–21 (2015).
25. Hoedemaker, R. S., Ernst, J., Meyer, A. S. & Belke, E. Language production in a shared task: Cumulative Semantic Interference from self- and other-produced context words. *Acta Physiol. (Oxf)* **172**, 55–63 (2017).
26. Kuhlen, A. K. & Abdel Rahman, R. Having a task partner affects lexical retrieval: Spoken word production in shared task settings. *Cognition* **166**, 94–106 (2017).
27. Mubin, O., Ahmad, M. I., Kaur, S., Shi, W. & Khan, A. Social robots in public spaces: A meta-review. In *Social Robotics* (eds Ge, S. S. *et al.*) 213–220 (Springer International Publishing, 2018).
28. Sartorato, F., Przybylowski, L. & Sarko, D. K. Improving therapeutic outcomes in autism spectrum disorders: Enhancing social communication and sensory processing through the use of interactive robots. *J. Psychiatr. Res.* **90**, 1–11 (2017).
29. van den Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., van der Ven, S. & Leseman, P. Social robots for language learning: A review. *Rev. Educ. Res.* **89**, 259–295 (2019).
30. Grigore, E. C., Pereira, A., Zhou, I., Wang, D. & Scassellati, B. Talk to me: Verbal communication improves perceptions of friendship and social presence in human-robot interaction. In *Intelligent Virtual Agents* (eds Traum, D. *et al.*) 51–63 (Springer International Publishing, 2016).
31. Bergmann, K., Branigan, H. P. & Kopp, S. Exploring the alignment space-lexical and gestural alignment with real and virtual humans. *Front. ICT* **2**, 7 (2015).
32. Peeters, D. Virtual reality: A game-changing method for the language sciences. *Psychon. Bull. Rev.* **26**, 894–900 (2019).
33. Heyselaar, E., Hagoort, P. & Segaert, K. In dialogue with an avatar, language behavior is identical to dialogue with a human partner. *Behav. Res. Methods* **49**, 46–60 (2017).
34. Heyselaar, E., Peeters, D. & Hagoort, P. Do we predict upcoming speech content in naturalistic environments?. *Lang. Cogn. Neurosci.* **36**, 440–461 (2021).
35. Eichert, N., Peeters, D. & Hagoort, P. Language-driven anticipatory eye movements in virtual reality. *Behav. Res. Methods* **50**, 1102–1115 (2018).
36. Pickering, M. J. & Garrod, S. The interactive-alignment model: Developments and refinements. *Behav. Brain Sci.* **27**, 212–225 (2004).
37. Iio, T. *et al.* Lexical entrainment in human robot interaction. *Int. J. Soc. Robot.* **7**, 253–263 (2015).
38. Brandstetter, J., Beckner, C., Sandoval, E. B., & Bartneck, C. Persistent lexical entrainment in HRI. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, 63–72 (2017).
39. Staum Casasanto, L., Jasmin, K. & Casasanto, D. Virtually accommodating: Speech rate accommodation to a virtual interlocutor. In *32nd Annual Meeting of the Cognitive Science Society (CogSci 2010)* 127–132 (2010).
40. Gijssels, T., Casasanto, L. S., Jasmin, K., Hagoort, P. & Casasanto, D. Speech accommodation without priming: The case of pitch. *Discourse Process.* **53**, 233–251 (2016).
41. Heyselaar, E., Hagoort, P. & Segaert, K. In dialogue with an avatar, language behavior is identical to dialogue with a human partner. *Behav. Res. Methods* **49**, 46–60 (2015).
42. Branigan, H. P., Pickering, M. J., Pearson, J., McLean, J. F. & Nass, C. Syntactic alignment between computers and people: The role of belief about mental states. In *Proceedings of the 25th Annual Conference of the Cognitive Science Society*, 186–191 (2003).
43. Beckner, C., Rácz, P., Hay, J., Brandstetter, J. & Bartneck, C. Participants conform to humans but not to humanoid robots in an english past tense formation task. *J. Lang. Soc. Psychol.* **35**, 158–179 (2016).
44. Dell, G. S. A spreading-activation theory of retrieval in sentence production. *Psychol. Rev.* **93**, 283–321 (1986).
45. Levelt, W. J. M. *Speaking: From Intention to Articulation* (MIT Press, 1993).
46. Levelt, W. J. M., Roelofs, A. & Meyer, A. S. A theory of lexical access in speech production. *Behav. Brain Sci.* **22**, 1–38 (1999).
47. Roelofs, A. A spreading-activation theory of lemma retrieval in speaking. *Cognition* **42**, 107–142 (1992).
48. Roelofs, A. A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition* **172**, 59–72 (2018).
49. Caramazza, A. Access of phonological and orthographic lexical forms: Evidence from dissociations in reading and spelling. *Cogn. Neuropsychol. (Print)* **14**, 1–2 (1997).
50. Dell, G. S., Burger, L. K. & Svec, W. R. Language production and serial order: A functional analysis and a model. *Psychol. Rev.* **104**, 123–147 (1997).
51. Mahon, B. Z., Costa, A., Peterson, R., Vargas, K. A. & Caramazza, A. Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *J. Exp. Psychol. Learn. Mem. Cogn.* **33**, 503–535 (2007).
52. Oppenheim, G. M., Dell, G. S. & Schwartz, M. F. The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition* **114**, 227–252 (2010).
53. Abdel Rahman, R. & Melinger, A. Semantic context effects in language production: A swinging lexical network proposal and a review. *Lang. Cogn. Process.* **24**, 713–734 (2009).
54. Abdel Rahman, R. & Melinger, A. Semantic processing during language production: An update of the swinging lexical network. *Lang. Cogn. Neurosci.* **34**, 1176–1192 (2019).
55. Navarrete, E., Mahon, B. Z. & Caramazza, A. The cumulative semantic cost does not reflect lexical selection by competition. *Acta Physiol. (Oxf)* **134**, 279–289 (2010).

56. Howard, D., Nickels, L., Coltheart, M. & Cole-Virtue, J. Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition* **100**, 464–482 (2006).
57. Belke, E. Long-lasting inhibitory semantic context effects on object naming are necessarily conceptually mediated: Implications for models of lexical-semantic encoding. *J. Mem. Lang.* **69**, 228–256 (2013).
58. Costa, A., Strijkers, K., Martin, C. & Thierry, G. The time course of word retrieval revealed by event-related brain potentials during overt speech. *Proc. Natl. Acad. Sci.* **106**, 21442–21446 (2009).
59. van Casteren, M. & Davis, M. H. Mix, a program for pseudorandomization. *Behav. Res. Methods* **38**, 584–589 (2006).
60. Kuhlen, A. K. & Abdel Rahman, R. Joint language production: An electrophysiological investigation of simulated lexical access on behalf of task partner. *J. Exp. Psychol. Learn. Mem. Cogn.* (in press).
61. Bartneck, C., Kulić, D., Croft, E. & Zoghbi, S. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *Int. J. Soc. Robot.* **1**, 71–81 (2009).
62. Carpinella, C. M., Wyman, A. B., Perez, M. A. & Stroessner, S. J. The robotic social attributes scale (RoSAS) development and validation. In *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction*, 254–262 (2017).
63. Spatola, N., Kühnlenz, B. & Cheng, G. Perception and evaluation in human–robot interaction: The Human–Robot Interaction Evaluation Scale (HRIES)—A multicomponent approach of anthropomorphism. *Int. J. Soc. Robot.* 1–23 (2021).
64. Sindermann, C. *et al.* Assessing the attitude towards artificial intelligence: Introduction of a short measure in German, Chinese, and English Language. *Künstl. Intell.* **35**(1), 109–118 (2020).
65. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Soft.* **67**, 1–48 (2015).
66. Box, G. E. P. & Cox, D. R. An analysis of transformations. *J. R. Stat. Soc.: Ser. B (Methodol.)* **26**, 211–243 (1964).
67. Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J. Mem. Lang.* **68**, 255–278 (2013).
68. Henschel, A., Hortensius, R. & Cross, E. S. Social cognition in the age of human–robot interaction. *Trends Neurosci.* **43**, 373–384 (2020).
69. Spatola, N. *et al.* Not as bad as it seems: When the presence of a threatening humanoid robot improves human performance. *Sci. Robot.* **3**, eaat5843 (2018).
70. Cross, E. S. & Ramsey, R. Mind meets machine: Towards a cognitive science of human–machine interactions. *Trends Cogn. Sci.* **25**, 200–212 (2021).

## Acknowledgements

We thank Guido Kiecker for technical support; and Martina Erbach, Wiebke Rehn, Annika Just, and Barbora Wolf for help with data collection. Anna Kuhlen and Rasha Abdel Rahman have given permission for the use of data (from previous publication<sup>26</sup> for use in the supplemental analysis (Comparing Robot as a Task Partner and Human as a Task Partner) in the current report.

## Author contributions

Conceptualization: R.A.R., A.K.K., O.W., Methodology: O.W., A.K.K., M.K., D.P., R.A.R., Investigation: O.W., M.K., D.P., Writing—Original Draft: O.W., R.A.R., A.K.K., Writing—Review & Editing: M.K., D.P., V.H., J.H., Funding Acquisition: R.A.R., V.H., J.H., Resources, R.A.R., A.K.K., M.K., V.H., Supervision, A.K.K., R.A.R.

## Funding

Open Access funding enabled and organized by Projekt DEAL. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—EXC 2002/1 "Science of Intelligence"—project number 390523135.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-95645-9>.

**Correspondence** and requests for materials should be addressed to O.A.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021