Dialogue & Discourse 16(3) (2025) 8–24 doi: 10.5210/dad.2025.302

Laughter use by virtual agents increases task success

Bogdan Ludusan BOGDAN.LUDUSAN@UNI-BIELEFELD.DE
Phonetics Workgroup, Faculty of Linguistics and Literary Studies & CITEC
Bielefeld University, Germany

Petra Wagner PETRA.WAGNER@UNI-BIELEFELD.DE
Phonetics Workgroup, Faculty of Linguistics and Literary Studies & CITEC
Bielefeld University, Germany

Editor: Dimitra Gkatzia

Submitted 11/2023; Accepted 12/2024; Published online 12/2025

Abstract

Studies on human-robot interaction as well as on embodied conversational agents have revealed that the use of laughter by agents increases their perceived naturalness and their social presence. However, laughter plays a variety of functions in human interaction, and its effects on communication go beyond those previously investigated in the aforementioned fields. Taking into account that laughter has been shown to improve task performance in human-human interaction, we investigated here whether laughter use by a virtual agent increases task success also in human-machine interaction. A real-estate scenario was considered, in which an agent presented an apartment to an interested client. Both the presence of laughter and the nature of the agent (virtual or human) were varied in the experiment. We operationalized the task success as being the likelihood of participants recommending the apartment, while also examining the perceived rating of the agent. The results of an observer study showed that the use of laughter by a virtual agent translates in an increased task success, while also confirming previous findings regarding improvements in the social perception of the agent. Our results concerning the task success in the human agent condition were not in line with those of previous studies, most likely due to a reduced naturalness of the used laughter. This makes the findings pertaining to the virtual agent, where benefits were observed by the use of laughter in interaction, even more salient. Taken together, these results seem to suggest that in this case humans are less sensitive to a reduced laughter naturalness. We further discuss the need for better laughter integration with speech, as well as its automatic synthesis in order to better take advantage of these findings.

Keywords: laughter, human-machine interaction, virtual agent, human-human interaction, task success

1. Introduction

Robots employ a wide range of signals to interact with humans, including verbal and non-verbal ones. In particular, it has been shown that a robot's non-verbal expression may elicit emotional or behavioral responses from their interlocutor, as well as having an impact on the user's task performance (Saunderson and Nejat, 2019). Recent years have seen an increased research interest in this field in sound-based non-verbal expressions, as these can be straightforwardly realized also in robots that are constrained in their appearance or in their interactional capabilities (i.e., robots that cannot produce complex gestural or facial expressions, or have limited capacities for maintaining a dialog). An appealing characteristic of sound-based non-verbal expressions is that they are language-independent, quickly understood, and can crucially contribute to the perception of a robot's personality (Ritschel et al., 2019). Much research on robot's non-verbal behaviour focuses on artificially generated and not necessarily human-like sounds, for example MIDI-like tones. These often exploit universally understood, iconic relationships between sounds and meaning (such as high pitch representing "happy" and low pitch corresponding to "sad") (Zhang and Fitter, 2023). Here, we choose an alternative path by looking at a genuinely human, but nevertheless non-verbal signal – laughter, and we evaluate its usefulness for the interaction between a human and a virtual agent.

Laughter is a multi-faceted phenomenon playing various social and communicative roles in human interaction (Glenn, 2003). Besides the role it plays in humor (mirthful laughter), social (non-mirthful) laughter may be used to convey both positive or cooperative behaviours and negative ones. Among the positive aspects, laughter may indicate closeness or affiliation, or a feeling of pleasantness with regards to the conversational partner. It may evoke also negative emotions, with laughing at someone contributing to disaffiliation or hostility towards that person. Further functions of laughter may include the softening of previous affirmations, the remedy of possible interactional misunderstandings, or the handling of difficult situations (e.g., involving embarrassment or discomfort).

Considering the widespread use of laughter in human-human communication and the multitude of functions that it may have in conversation, laughter should be beneficial to the interaction, increasing its success one way or the other. This has been found also for interactions involving goal-oriented tasks, with previous work showing that the use of shared laughter (laughter initiated by one of the interlocutors and taken up by their conversational partner) may improve, for instance, task performance in workplace meetings (Kangasharju and Nikko, 2009), as well as increase the hiring chances of candidates at job interviews (Brosy et al., 2021). The above-mentioned characteristics of laughter have encouraged its implementation in human-machine interaction systems. This has been done either for general spoken dialogue systems (Maraev et al., 2018), or by integrating it in embodied conversational agents (Niewiadomski et al., 2013; Pecune et al., 2015; El Haddad et al., 2016; Mancini et al., 2017) or in robots (Becker-Asano and Ishiguro, 2009; Türker et al., 2017; Inoue et al., 2022). Some of these studies have focused exclusively on the required technical implementation of laughter in these systems (Maraev et al., 2018), while others only on the accuracy/naturalness of specific sub-systems (e.g., laughter prediction/synthesis El Haddad et al., 2016).

Experiments to quantify the effect of agent laughter use on the perception of the agent by their human interlocutor or on the overall interaction have also been conducted. In Niewiadomski et al. (2013), the participants compared a laughing agent that produced shared laughter with them, to an agent not using laughter or employing laughs at predefined time instants. They perceived: 1) the video clips they were watching together with the agent as being more funny, and 2) the laughter of the agent more contagious, in the former case (the latter finding reflecting views found in the literature examining laughter in human communication Provine, 1996). Although no increased social presence of the agent was perceived, a subsequent study (Pecune et al., 2015) in which the embodied agent produced shared laughter with the user, by copying the latter's laughter intensity and trunk movements reported higher scores for agent social presence compared to the case when there was no shared laughter (but see Mancini et al., 2017 for the opposite result). Responding to the laughter of the human interlocutor also improved a number of measures related to user engagement in Türker et al. (2017). Lastly, a system employing a more complex decision on whether the agent should initiate shared laughter or not and which type of laughter to use, resulted in increased ratings (e.g., naturalness, human likeness) for some of the tested scenarios in Inoue et al. (2022), compared to a system responding to the laughter of the user every time. These studies have shown that the use of laughter by a virtual agent resulted in a higher rating given to that agent, in line to findings in human-human interaction which show that laughter use is beneficial to the relationship between interlocutors (Kurtz and Algoe, 2017) (see also Becker-Asano and Ishiguro 2009, in which the use of laughter has elicited also negative emotions towards the embodied agent, a normal reaction if the person perceives that they are being laughed at).

We have seen that a virtual agent that laughs is judged to be more natural or more human-like (Becker-Asano and Ishiguro, 2009; Niewiadomski et al., 2013; Inoue et al., 2022) and to have an increased social presence (e.g., be more empathetic and understanding Pecune et al., 2015; Inoue et al., 2022), while also increasing the engagement with its interlocutor (Türker et al., 2017). However, there is little work that proves that the use of laughter by a virtual agent is conducive to the success of the task performed with the user, beyond the improved social perception. While improved social characteristics may, indirectly, result in an increased success of the task, more direct evidence is needed.

We present a study investigating whether laughter use by a virtual agent increases the success of a task that agent is carrying out. The task consisted of a real-estate scenario, in which the agent was showing to a client an apartment for rent, with the participants to the experiments being observers of the task. The success of the interaction was operationalized as the degree to which the apartment would be recommended by the participants to their acquaintances. Furthermore, similarly to other studies, we examined the perception of different characteristics of the agent, as well as an overall rating of the agent, defined as the degree to which the agent would be recommended by the participants. We did not limit ourselves to considering a virtual agent that uses laughter versus one that does not laugh (Ludusan and Wagner, 2021), but we also compared it to how a human agent would be perceived at the same task, including the effect of laughter on task success. We extended the study in Ludusan and Wagner (2021) by testing three additional conditions: a human agent employing either no laughter, laughter, or "more natural" laughter, and performing a joint analysis of all conditions (virtual and human agent), with statistical models that better fit the distribution of the data.

2. Materials and methods

2.1 Stimuli

An interaction between a real-estate agent and a client visiting an apartment was designed for the experiment. We decided on using this particular situational setting because conversations involving real-estate scenarios have been shown to be useful for evaluating virtual agents (Cassell et al., 1999), while also providing a quantifiable measure of success for the interaction – whether the apartment was liked by the experiment participants.

The conditions considered in this study are illustrated in Figure 1. We varied two factors: the nature of the real-estate agent (human/virtual) and the use of laughter by the agent (no/yes), resulting in four conditions, representing a two-by-two crossed design. We also considered a fifth condition, the third in the case of the human agent, to evaluate the appropriateness of the laughter used in the other two laughter conditions.

We started by synthesizing the conversation lines belonging to the virtual agent, using a commercially-available state-of-the-art system[1], based on one of the WaveNet male German voices (de-DE-Wavenet-B). As we wanted to have an intonation more similar to conversational speech, rather than the default intonation employed by synthesis systems (trained mainly on read speech data), we tested several parameters (e.g., placement of punctuation marks, duration of pauses, the use of emphasis) during the synthesis in order to achieve this, while avoiding the introduction of any artefacts. The obtained stimuli were vetted by two native speakers and any reported inconsistencies were corrected. The resulting materials composed the speech of the virtual agent in the non-laughter condition.

Then, we produced the lines corresponding to the virtual agent, in the laughter condition. Conversational laughter instances produced by male speakers were extracted from the recordings of the DUEL corpus (Hough et al., 2016). The corpus contains dyadic task-based interactions that encourage the use of conversational laughter. The chosen instances had the characteristics shown previously to be well integrated with synthetic speech (Trouvain and Schröder, 2004): being composed of one to three laughter syllables, either voiced or unvoiced (snort-like), and having low to medium intensity. Eight positions within the lines of the agent were identified as being appropriate to contain laughter and we spliced at those positions suitable laughter tokens (in terms of pitch, timbre, or expressed function). These employed laughs were either three-, two- or one-syllable long (two, four, and two instances, respectively), with two of the two-syllable long laughs being unvoiced, the remaining laughs being voiced.

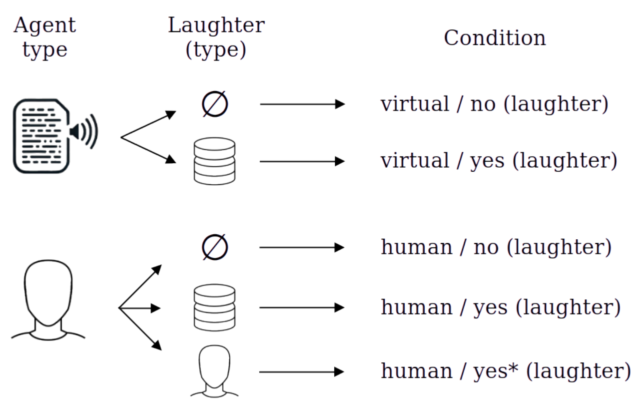

Figure 1: An illustration of the stimuli creation procedure. The lines for the virtual agent were created using a state-of-the-art speech synthesis system, while those corresponding to the human agent were recorded by a native German speaker. The agent used either no laughter (cross symbol), laughter extracted from corpora of conversational speech (database symbol) or laughter instances produced by the native speaker.

Next, we recorded the lines corresponding to the client in the recording studio of the university. A female native speaker of German familiarized herself beforehand with the lines of the client.

1. https://cloud.google.com/text-to-speech

During the recording session, the experimenter played through loudspeakers the line corresponding to the agent (laughter condition) and the speaker would respond with the appropriate line. In order to increase the naturalness of the recordings, the speaker was allowed to change the exact wording of the sentence, provided that its meaning stayed the same. Several takes of the entire conversation were performed and the most spontaneously-sounding lines were employed in the experiment.

Two versions of the stimuli were put together: for the non-laughter condition we combined the synthesized speech not containing laughter with the lines of the female client, while for the laughter case, the synthesized speech including the spliced laughs with the lines of the client. Thus, the only difference between the two obtained recordings were the absence/presence of laughter in the virtual agent's speech.

Finally, we created equivalent conditions involving a human agent. A male German native speaker recorded the lines of the agent in a similar setting to how the female client lines were recorded, hearing through the loudspeakers the sentences produced by the client and answering with the corresponding agent line. However, he was not allowed to rephrase the sentences, in order to have identical materials in the virtual and human agent cases. For the laughter condition, the speaker was asked to produce laughs which were as similar as possible to the ones employed for the virtual agent. Therefore, we had a design varying two types of agents (human/virtual) and two laughter conditions (absence/presence). As mentioned at the beginning of this section, we also considered a control condition for the human agent, in order to check the appropriateness of the laughter used. In that condition, the speaker was asked to laugh as he would normally laugh at that particular point of the interaction, instead of producing laughs similar to the ones used in the virtual agent condition. Three of the employed laughter instances were speech-laughs (defined as concurrent production of speech and laughter in which neither of the two components is dominant Nwokah et al., 1999), with the rest being laughs, having an overall lower intensity than in the human agent, laughter condition.

The audio stimuli[2] corresponding to the virtual agent were around 4 minutes 30 seconds long, while those of the human agent were circa 3 minutes 50 seconds long, due to differences between the speech rate of the synthesized speech of the virtual agent and that of the human speaker. The recordings of each of the five experimental conditions were combined with the images of an apartment, in order to create a video clip. The employed images did not contain any visual representation of the agent or of the client. At any time, the displayed image matched the location where the discussion was taking place (hallway, rooms, etc.) and the images changed at the same points in the interaction in all conditions. The resulting videos, having identical visual content and different audio components, depending on the condition, represented the stimuli used in this study.

2.2 Procedure

The experiment was run on the PsyToolkit platform (Stoet, 2010, 2017). The participants watched the video containing the slide show of the apartment and the recording between the agent and the client, having been informed of the nature of the agent (human/virtual) at the beginning of the experiment. After that, they were asked to judge the performance of the agent with respect to several dimensions: professionalism, communication, pleasantness, formality and spontaneity. Next, they were asked to rate the probability of recommending the apartment (our task success measure) and the agent to their acquaintances. For the human agent conditions, since we also wanted to evaluate how the employed laughter was perceived when used by a human, the participants also replied to

2. https://pub.uni-bielefeld.de/download/2999852/3000681.zip

two additional questions on the politeness and the empathy shown by the human agent. Lastly, the human agent conditions involving laughter (the original and the "more natural" laughter one) had two other questions asking the participants whether they perceived the agent to be laughing at the correct locations (from "always wrong" to "always right") and if the laughter sounded appropriate for that situation (from "not appropriate" to "appropriate"). All the questions used a 10-point scale (e.g., for professionalism, the value 1 of the scale was labelled as "unprofessional", while the value 10 corresponded to "professional"). The exact formulation of the questions can be found in Appendix A.

The study was reviewed and approved by the Ethical Committee of Bielefeld University (EUB No. 2019-085 and No. 2020-165). The majority of the participants were recruited through the Prolific crowdsourcing platform[3] and were paid according to the regular rate for participating in studies at Bielefeld University. The rest of them (about 10%) were either students at Bielefeld University (receiving course credit for their participation) or friends and family of the authors. All participants had to be residing in Germany and be fluent in German. We used a between-participants design, each participant taking part in only one condition. Of the 346 participants that completed the study, we had to exclude the data of 16 participants who did not listen to the entire conversation. The remaining participants[4] were split evenly (66 each) between the four conditions (human/virtual agent, no laughter/laughter). We aimed to have a balanced number of male and female participants in each condition (see Table 1 for exact numbers).

Agent type	Laughter present	# female	# male	# other	Total
Virtual	no	34	32	0	66
	yes	31	34	1	66
Human	no	33	31	2	66
	yes	33	33	0	66
	yes*	33	33	0	66

Table 1: Gender distribution (female/male/other) of the participants in the experiment, across all tested condition. The human agent yes* condition represents the case in which the human agent produces "more natural" laughter.

2.3 Analyses

All participant data from the non-laughter conditions, as well as from the laughter conditions using the same type of laughter (*yes*), were considered in the main analyses. Three participants, self-identified as neither female or male, were excluded from the analyses involving the predictor gender, since not enough data points were available for inclusion. In order to determine the effects of each of the investigated characteristics on the considered success measure, we employed generalized linear regression models. As the data best fitted a Poisson distribution, but it was under-dispersed, we used a generalized Poisson distribution, Conway-Maxwell Poisson, that handles both over-dispersed and under-dispersed data. The dependent variable of the models was either the apart-

3. https://www.prolific.com/
4. They include also a set of 66 participants, equally split between males and females, that was tested in the human agent, "more natural" laughter condition.

ment or the agent rating and, as predictors, the following: the agent type (human/virtual), the laughter presence (no/yes), the characteristics rated by the participants (communication, professionalism, pleasantness, formality and spontaneity), as well as the age and the gender of the participants. We also included all two- and three-way interactions between agent type, laughter presence, and each of the other predictors. This model was then reduced through step-wise elimination (Matuschek et al., 2017), by removing, at each step, the factor whose elimination would reduce the most the Akaike Information Criterion (AIC) value of the model. The process was stopped when the removal of any factor did not decrease the AIC of the model. All numeric predictors were scaled, by subtracting their mean value and dividing them by two standard deviations (Gelman, 2008). A sum to zero contrast was employed for the factor variables involved in the model. All analyses were performed using the appropriate functions of the R software (R Core Team, 2020, ver. 3.6.3), the models being fitted using the *glm.cmp* function of the COMPoissonReg package (Sellers et al., 2022).

3. Results

3.1 Overall results

The per-condition results obtained for the apartment rating, our task success measure, are illustrated in Figure 2. A visual inspection of the results reveals no effect of laughter in the human agent condition, but an increase in apartment rating when the virtual agent used laughter, compared to when they did not.

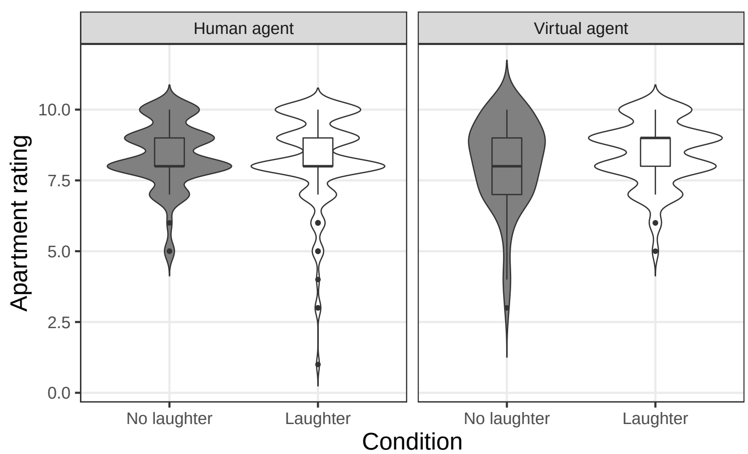

Figure 2: The apartment rating obtained in the four conditions considered in the study: human/virtual agent, no laughter/laughter. The bars inside each violin plot illustrate the median and the quartile values of the apartment rating.

The generalized linear regression model predicting the apartment rating showed that the presence of laughter had a significant effect directly, the apartment being rated higher if the agent used laughter ($\beta = 0.126, p = 0.029$), as well as through its interaction with the age of the participant ($\beta = 0.310, p = 0.017$). Significant effects were found also for professionalism ($\beta = 0.328, p =$

0.012), communication ($\beta = 0.365, p = 0.021$), and pleasantness ($\beta = 0.477, p = 0.001$), the apartment rating increasing with the increased rating of these characteristics, and for the age of the participant, with older participants rating the apartment lower ($\beta = -0.294, p = 0.023$). However, in the condition in which the agent laughed, the apartment was rated higher by older participants (see previously reported laughter-age interaction), suppressing the negative main effect seen by age.

The rating of the agent in the various conditions is displayed in Figure 3, while results related to the perceived characteristics of the agent are given in Table 2. Although laughter seems to have a more reduced effect for the agent rating, compared to the apartment rating, we verified the existing trends by means of a generalized linear regression model fitted with the agent rating.

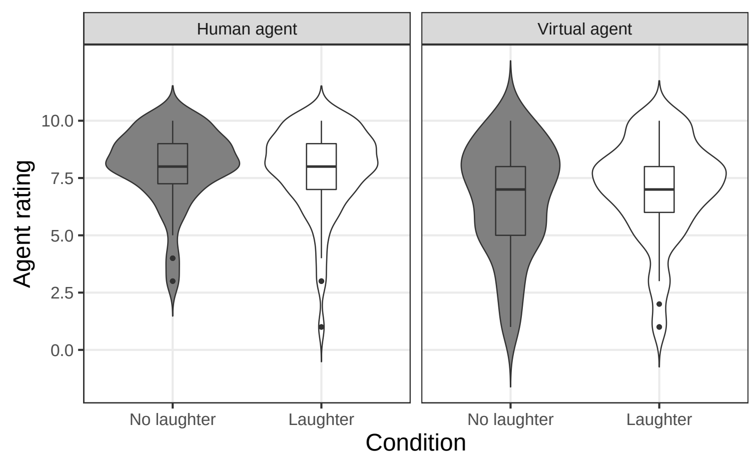

Figure 3: The agent rating obtained in the four conditions considered in the study: human/virtual agent, no laughter/laughter. The bars inside each violin plot illustrate the median and the quartile values of the rating.

The fitted model did not show a main effect of laughter, but significant interactions between agent type and laughter presence, the virtual agent being rated higher when they laughed ($\beta = 0.115, p = 0.016$), and between laughter and formality, the rating of an agent using laughter decreasing with its perceived increased formality ($\beta = -0.208, p = 0.039$). Further significant effects were found for professionalism ($\beta = 0.234, p = 0.049$), communication ($\beta = 0.491, p = 4.8e^{-4}$) and pleasantness ($\beta = 1.201, p = 4.7e^{-13}$) and for the interactions type-communication ($\beta = 0.384, p = 0.002$) and type-laughter-professionalism ($\beta = 0.261, p = 0.028$). Judging the agent to have better communication skills, to be more professional and to be more pleasant increased the rating of the agent. Moreover, rating the virtual agent higher on communication increased its rating in addition to the main effect seen by communication.

3.2 Agent type analyses

In order to better understand the results, we performed separate analyses, for the human and for the virtual agent data, respectively. Each sub-analysis looked at both the agent and the apartment rating, employing them as dependent variables in regression models. The models were similar to

Agent type	Laughter present	Prof.	Comm.	Plea.	Form.	Spont.	Agent rating	Apartment rating
Human	no	8.58	8.71	8.64	5.03	5.52	8.03	8.39
	yes	8.30	8.59	8.39	4.29	6.17	7.71	8.48
Virtual	no	8.56	8.29	7.56	6.14	5.36	6.59	8.12
	yes	8.29	8.21	7.33	4.97	5.32	7.03	8.45

Table 2: Mean values of the scores given by the participants to the seven dimensions evaluated in the experiment (Professionalism - Prof., Communication - Comm., Pleasantness - Plea., Formality - Form., Spontaneity - Spon., Agent rating, and Apartment rating), for each type of agent (human/virtual) and each laughter condition (no laughter/laughter).

the ones previously described, except for the agent type no longer being a predictor. Moreover, the models fitted with the virtual agent data included two additional predictors: the self-reported speech technology exposure of the participants and its interaction with laughter use.

The first model fitted with the virtual agent data showed that the apartment rating depended on laughter use and on the perceived communication skill of the agent, as well as on the interaction between laughter and each of the following three factors: communication, speech technology exposure and age. A virtual agent using laughter ($\beta = 0.184, p = 0.044$) and having a higher communication skill ($\beta = 0.699, p = 6.9e^{-4}$) increased the rating of the apartment. The apartment presented by a virtual agent that laughed was also rated higher by participants reporting a higher speech technology exposure ($\beta = 0.455, p = 0.008$) and by older participants ($\beta = 0.479, p = 0.045$), and rated lower when the agent was judged to have higher communicative skills ($\beta = -0.382, p = 0.020$).

The second model for the virtual agent condition revealed that the agent rating increased when the agent laughed ($\beta = 0.211, p = 0.004$), while it decreased when the virtual agent used laughter and its perceived formality increased ($\beta = -0.304, p = 0.041$). Communication ($\beta = 0.640, p = 6.7e^{-4}$) and pleasantness ($\beta = 1.053, p = 1.6e^{-7}$) also had a significant effect on the agent rating, the latter being higher when the agent was perceived to have increased communication skills and a higher pleasantness.

For the models built on the human agent data (for the apartment and for the agent rating), the only predictors that had a significant effect on the dependent variables were professionalism ($\beta = 0.539, p = 0.007$ for the apartment and $\beta = 0.672, p = 1.4e^{-4}$ for the agent rating) and pleasantness ($\beta = 0.773, p = 0.002$ and $\beta = 1.620, p = 9.1e^{-9}$ for the apartment and the agent rating, respectively). An increase in the scores given to these characteristics resulted in an increase in both the apartment and the agent rating.

3.3 Detailed analysis of the human agent conditions

An analysis of the additional questions asked to the participants in the human agent conditions was performed. It examined the politeness and the empathy of the agent, as well as the perceived correctness of the place where the agent laughed and the perceived appropriateness of the employed laugh. We display these results in Table 3, along with all the scores given by participants in the "more natural" (yes*) laughter condition. For an easier comparison, we illustrate all the ratings obtained in the other two human agent conditions.

Laughter present	Prof.	Comm.	Plea.	Form.	Spont.	Agent rating	Apt. rating	Pol.	Emp.	Place	Appr.
no	8.58	8.71	8.64	5.03	5.52	8.03	8.39	9.15	7.73	-	-
yes	8.30	8.59	8.39	4.29	6.17	7.71	8.48	9.21	7.68	6.30	6.08
yes*	8.55	8.83	8.70	3.92	6.41	8.03	8.02	9.23	8.23	7.70	7.36

Table 3: Mean values of the scores given by the participants in the three human agent conditions: no laughter (no), laughter (yes) and "more natural" (yes*) laughter. Presented are the seven dimensions evaluated in the experiment (Professionalism - Prof., Communication - Comm., Pleasantness - Plea., Formality - Form., Spontaneity - Spon., Agent rating, and Apartment rating), as well as the perceived politeness (Pol.) and perceived empathy (Emp.) of the agent, and the appropriateness of the place where laughter was used (Place) and of the laughter itself (Appr.).

We employed the same type of generalized linear regression model as in the previous analyses to determine whether there was an effect of laughter condition on several scores. The following scores were considered: apartment rating, agent rating, perceived agent politeness, perceived agent empathy, appropriateness of the place where laughter was used, and appropriateness of the type of employed laughter. We performed pairwise comparisons between the conditions: no laughter (no) / laughter (yes) and laughter (yes) / "more natural" laughter (yes*).

There was no effect of the presence of laughter (no/yes conditions) on any of the investigated measures: on the apartment rating ($\beta = 0.023, p = 0.760$), on the agent rating ($\beta = -0.0521, p = 0.280$), on the perceived agent politeness ($\beta = 0.008, p = 0.912$), or on the perceived agent empathy ($\beta = -0.010, p = 0.851$).

Also the type of laughter produced (similar to the one employed by the virtual agent vs. the "more natural" laughter) did not have any effect on the apartment rating ($\beta = 0.086, p = 0.116$), on the agent rating ($\beta = -0.041, p = 0.356$), on the perceived agent politeness ($\beta = -0.006, p = 0.938$), or on the perceived agent empathy ($\beta = -0.108, p = 0.055$). There were differences, however, between the two conditions on the remaining examined measures: the appropriateness of the place where the laughter was used ($\beta = -0.183, p = 2.5e^{-4}$) and the appropriateness of the type of employed laughter ($\beta = -0.126, p = 0.002$). In both cases, the laughter produced similarly to how the virtual agent laughed was perceived less appropriate. However, the place of the laughter events did not vary between the two conditions and the significant difference in reported laughter place must be due to the participants not being able to separate the two appropriateness measures. This comparison shows that, while the participants did perceive the laughs produced by the human agent in the "more natural" condition as being more appropriate, these laughs did not have an effect on the overall perception of the agent or on the task success measure.

4. Discussion and conclusions

Investigating the use of laughter in conversation by a human or a virtual agent, we observed an effect of laughter on the success of the task, defined as the likelihood of recommending the apartment presented by the agent, with a higher task success when the agent employed laughter. This effect was mainly driven by the ratings given in the virtual agent case, where also the likelihood of recommending the agent increased with the use of laughter. These findings add to the body of evidence that laughter helps the interaction with human interlocutors. Our study took one step further, though, by providing confirmation that the use of laughter by a virtual agent increases not only how the agent is perceived socially, but also the success rate of the performed task, when compared to an agent that does not laugh. Although, due to the experimental setting employed here, there was no explicit embodiment given to the virtual agent, we believe our findings may be easily extended to embodied agents, taking into account that the physical embodiment of an agent actually increases their evaluation by humans (Lee et al., 2006), and may also have an indirect effect on task success. Moreover, similar findings have been previously reported concerning the use of laughter, regardless whether a conversational agent (Niewiadomski et al., 2013; Pecune et al., 2015) or a robot (Becker-Asano and Ishiguro, 2009; Türker et al., 2017; Inoue et al., 2022) was employed in the experiment.

We compared the effect of laughter use also by a human agent, on the same task, revealing that the increase in task success can be seen only in the virtual agent case. Our results did not corroborate those previously observed in human-human interaction (Kangasharju and Nikko, 2009; Brosy et al., 2021) and this lack of an increase in task success may be explained by the reduced naturalness of the laughter stimuli used in our study, probably caused by how they were elicited. While no differences, neither in the apartment or agent rating, nor in the perceived politeness or empathy of the agent, have been observed between the human agent laughter and "more natural" laughter conditions, the fact that the participants indicated differences in the appropriateness of laughs used (including their place, although that was not the case) might be an indicator of this limited naturalness. In order for us to be able to directly compare the human with the virtual agent conditions we had to use the same types of laughter and thus, we were limited by the selection available to use in the virtual agent conditions. However, despite the reduced naturalness of the employed laughs, in both the human and the virtual agent conditions, it seems that this was more strongly perceived in the former, rather than in the latter conditions. These results are encouraging for using laughter in human-machine interaction, seeing how humans are more forgiving towards reduced laughter naturalness in a machine than towards a human. Furthermore, our results provide evidence that it may be overly simplistic to expect that any behavior that is either helping or impeding interactions between human interlocutors, may have the same or, at least similar, effects on communication between humans and machines (but see Krämer et al. 2012 for a different view). Our findings may also indicate that humans may have differentiated expectations from virtual agents, which influence their assessment and processing of the virtual agent's signal, in line with previous work in the literature (Horstmann and Krämer, 2020).

We believe that overall more work is needed on the integration of laughter in human-machine interaction systems. Currently, the integration of social laughter (such as here) is mainly done by means of natural laughter instances extracted from speech corpora (Inoue et al., 2022). This seems to be due to limited capabilities of the current synthesis systems to generate naturally sounding laughter. State-of-the-art laughter synthesis systems based on deep-learning approaches (Mori et al., 2019; Tits et al., 2020), reach mean opinion scores of around three (on a scale up to five, on which natural laughter is evaluated between four and five). Even if the quality of laughter synthesis reaches a good level, one needs to take into account that the human laughter inventory does not only contain laughs. The acoustic production of laughter varies considerably (Bachorowski et al., 2001) and speech-laughs are often employed in spontaneous communication (Ludusan and Wagner 2019b reported between 20%-40% of all laughter instances to be speech-laughs, across materials in three languages). We observed this also in our data, with the human "more natural" laughter condition containing three instances of speech-laughs (out of a total of eight laughter events) and an inspection of the stimuli revealed that the use of speech-laughs was appropriate in those cases (maybe even more so than that of laughs). Unfortunately, there is less work done on the synthesis of speech-laughs and the existing studies suggest it being even more difficult than laughter synthesis, with lower mean opinion scores (El Haddad et al., 2015). Finally, as noted also by Türker et al. (2017), it is difficult to incorporate naturalistic laughter (be it laughs or speech-laughs) to synthetic speech. This is made even more challenging by the fact that speech preceding laughter is subject to co-articulation effects (Ludusan and Wagner, 2019a). For instance, in the "more natural" laughter condition, the produced laughter instances can be predicted before their actual occurrence, due to the acoustic changes in vocal tract in anticipation to the following laughter. Thus, more advances in speech technology are needed to seamlessly integrate laughter in synthetic speech, preferably through joint synthesis of speech and laughter (see also more recent attempts to generate conversational phenomena, such as laughter, directly in the synthesized speech signal, using deep learning models (Nguyen et al., 2023)).

Some observations on the laughter employed in the conversation are also in order. The improvements previously reported in the literature (on how the agent is perceived) were found when using shared laughter. In our case, out of the eight laughter events included in the stimuli, three of them were shared. While shared laughter seems to play an important role in human-machine interaction studies (e.g., Niewiadomski et al., 2013; Türker et al., 2017) and also in human-human communication (e.g., Ludusan and Wagner, 2019b), our findings show that not all laughs produced by the virtual agent need to be synchronised to those of the human interlocutor. Further analyses of the functions played by laughter will be required to better understand its roles in conversation and to better model laughter in human-machine interaction.

Future work could also include an investigation into whether an increase in task success may be seen for other type of settings (e.g., social settings – personal assistants). Another factor worth exploring is the effect of the perspective on the interaction with the virtual agent: either a first-person one, such as when the user actively participates in the interaction, or a third-person one, when they are an observer to the study (similar to here). An increase in task success in the case of laughter use by a virtual agent, in a first-person experiment, would be of particular applied importance. Lastly, independent of the experimental setting and the task involved, the human-machine interaction system will need to ensure an appropriate integration of laughter within the synthesized speech, one that takes into account also the functions that laughter plays, similar to the proposals put forward by Trouvain and Weiss (2022) for smiled speech.

Acknowledgments

The authors would like to thank Paul Kaminski, Leonie Schade and Marin Schröer for their help with setting up the dialogue and Annett Jorschick for helpful discussions regarding the experiment and the statistical analyses. Part of the results included here have been previously presented at ESSV 2021. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) project number 461442180. It was also supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement no. 799022.

Appendix A

This appendix contains the questions asked to the participants after watching the video containing the interaction between the agent and the client.

The following questions were asked for all conditions:

- Wie bewerten Sie die Interaktion des Maklers in den folgenden Dimensionen:
  How would you rate the broker's interaction along the following dimensions:
  - Professionalität: unprofessionell (1) - professionell (10)
    Professionalism: unprofessional (1) - professional (10)
  - Kommunikation: schlecht (1) - sehr gut (10)
    Communication: poor (1) - very good (10)
  - Annehmlichkeit: unangenehm (1) - angenehm (10)
    Pleasantness: unpleasant (1) - pleasant (10)
  - Förmlichkeit: locker (1) - förmlich (10)
    Formality: informal (1) - formal (10)
  - Spontanität: gestellt (1) - spontan (10)
    Spontaneity: acted (1) - spontaneous (10)
- Wie wahrscheinlich ist es, dass Sie Ihren Bekannten die Wohnung empfehlen würden? sehr unwahrscheinlich (1) - sehr wahrscheinlich (10)
  How likely is it that you would recommend the apartment to your acquaintances? very unlikely (1) - very likely (10)
- Wie wahrscheinlich ist es, dass Sie Ihren Bekannten den Makler empfehlen würden? sehr unwahrscheinlich (1) - sehr wahrscheinlich (10)
  How likely is it that you would recommend the broker to your acquaintances? very unlikely (1) - very likely (10)

An additional question was asked in the three conditions involving a human agent:

- Wie haben Sie das Verhalten des Maklers gegenüber der Interessentin wahrgenommen?
  How did you perceive the broker's behavior towards the client?
  - unhöflich (1) - höflich (10)
    rude (1) - polite (10)
  - abweisend (1) - mitfühlend (10)
    dismissive (1) - compassionate (10)

Two additional questions were asked in the two conditions involving a human agent that used laughter:

- Hatten Sie den Eindruck dass der Makler immer an den richtigen Stellen gelacht hat? immer falsch (1) - immer richtig (10)
  Did you have the impression that the broker always laughed in the right places? always wrong (1) - always right (10)
- Wenn der Makler lachte, wie klingte das für Sie? unangemessen (1) - angemessen (10)
  When the broker laughed, how did it sound to you? inappropriate (1) - appropriate (10)

References

Jo-Anne Bachorowski, Moria Smoski, and Michael Owren. The acoustic features of human laughter. *The Journal of the Acoustical Society of America*, 110(3):1581–1597, 2001. doi: 10.1121/1.4778613.

Christian Becker-Asano and Hiroshi Ishiguro. Laughter in social robotics-no laughing matter. In *Proceedings of the International Workshop on Social Intelligence Design*, pages 287–300, 2009. URL `https://gki.informatik.uni-freiburg.de/papers/becker-asano-ishiguro-sid2009.pdf`.

Julie Brosy, Adrian Bangerter, and Joaquim Sieber. Laughter in the selection interview: impression management or honest signal? *European Journal of Work and Organizational Psychology*, 30 (2):319–328, 2021. doi: 10.1080/1359432X.2020.1794953.

Justine Cassell, Timothy Bickmore, Mark Billinghurst, Lee Campbell, Kenny Chang, Hannes Vilhjálmsson, and Hao Yan. Embodiment in conversational interfaces: Rea. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 520–527, 1999. doi: 10.1145/302979.303150.

Kevin El Haddad, Stéphane Dupont, Jérôme Urbain, and Thierry Dutoit. Speech-laughs: an hmm-based approach for amused speech synthesis. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4939–4943, 2015. doi: 10.1109/ICASSP.2015.7178910.

Kevin El Haddad, Hüseyin Çakmak, Emer Gilmartin, Stéphane Dupont, and Thierry Dutoit. Towards a listening agent: a system generating audiovisual laughs and smiles to show interest. In *Proceedings of the ACM International Conference on Multimodal Interaction*, pages 248–255, 2016. doi: 10.1145/2993148.2993182.

Andrew Gelman. Scaling regression inputs by dividing by two standard deviations. *Statistics in medicine*, 27(15):2865–2873, 2008. doi: 10.1002/sim.3107.

Phillip Glenn. Towards a social interactional approach to laughter. In *Laughter in Interaction*, chapter 1, pages 7–34. Cambridge University Press, 2003. doi: 10.1017/CBO9780511519888.003.

Aike C Horstmann and Nicole C Krämer. Expectations vs. actual behavior of a social robot: An experimental investigation of the effects of a social robot's interaction skill level and its expected future role on people's evaluations. *PloS one*, 15(8):e0238133, 2020. doi: 10.1371/journal.pone.0238133.

Julian Hough, Ye Tian, Laura de Ruiter, Simon Betz, Spyros Kousidis, David Schlangen, and Jonathan Ginzburg. DUEL: A multi-lingual multimodal dialogue corpus for disfluency, exclamations and laughter. In *Proceedings of the Language Resources and Evaluation Conference*, pages 1784–1788, 2016. URL `https://aclanthology.org/L16-1281`.

Koji Inoue, Divesh Lala, and Tatsuya Kawahara. Can a robot laugh with you?: Shared laughter generation for empathetic spoken dialogue. *Frontiers in Robotics and AI*, page 234, 2022. doi: 10.3389/frobt.2022.933261.

Helena Kangasharju and Tuija Nikko. Emotions in organizations: Joint laughter in workplace meetings. *The Journal of Business Communication*, 46(1):100–119, 2009. doi: 10.1177/0021943608325750.

Nicole C Krämer, Astrid von der Pütten, and Sabrina Eimler. Human-agent and human-robot interaction theory: Similarities to and differences from human-human interaction. *Human-computer interaction: The agency perspective*, pages 215–240, 2012. doi: 10.1007/978-3-642-25691-2_9.

Laura E Kurtz and Sara B Algoe. When sharing a laugh means sharing more: Testing the role of shared laughter on short-term interpersonal consequences. *Journal of Nonverbal Behavior*, 41: 45–65, 2017. doi: 10.1007/s10919-016-0245-9.

Kwan Min Lee, Younbo Jung, Jaywoo Kim, and Sang Ryong Kim. Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction. *International Journal of Human-Computer Studies*, 64(10):962–973, 2006. doi: 10.1016/j.ijhcs.2006.05.002.

Bogdan Ludusan and Petra Wagner. No laughing matter: An investigation into the acoustic cues marking the use of laughter. In *Proceedings of the 19th International Congress of Phonetic Sciences*, pages 2179–2182, 2019a. URL `http://www.assta.org/proceedings/ICPhS2019/papers/ICPhS_2228.pdf`.

Bogdan Ludusan and Petra Wagner. Laughter dynamics in dyadic conversations. In *Proceedings of INTERSPEECH*, pages 524–528, 2019b. doi: 10.21437/Interspeech.2019-1733.

Bogdan Ludusan and Petra Wagner. Knock-knock! Who's there? The laughter-enhanced virtual real-estate agent. In *Elektronische Sprachsignalverarbeitung 2021. Tagungsband der 32. Konferenz*, volume 99, pages 281–288, 2021. URL `https://www.essv.de/essv2021/pdfs/16_ludusan.pdf`.

Maurizio Mancini, Beatrice Biancardi, Florian Pecune, Giovanna Varni, Yu Ding, Catherine Pelachaud, Gualtiero Volpe, and Antonio Camurri. Implementing and evaluating a laughing virtual character. *ACM Transactions on Internet Technology*, 17(1):1–22, 2017. doi: 10.1145/2998571.

Vladislav Maraev, Chiara Mazzocconi, Christine Howes, and Jonathan Ginzburg. Integrating laughter into spoken dialogue systems: preliminary analysis and suggested programme. In *Proceedings of the Workshop on Artificial Intelligence for Multimodal Human Robot Interaction*, pages 9–14, 2018. doi: 10.21437/AI-MHRI.2018-3.

Hannes Matuschek, Reinhold Kliegl, Shravan Vasishth, Harald Baayen, and Douglas Bates. Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, 94: 305–315, 2017. doi: 10.1016/j.jml.2017.01.001.

Hiroki Mori, Tomohiro Nagata, and Yoshiko Arimoto. Conversational and social laughter synthesis with wavenet. In *Proceedings of INTERSPEECH*, pages 520–523, 2019. doi: 10.21437/Interspeech.2019-2131.

Tu Anh Nguyen, Eugene Kharitonov, Jade Copet, Yossi Adi, Wei-Ning Hsu, Ali Elkahky, Paden Tomasello, Robin Algayres, Benoit Sagot, Abdelrahman Mohamed, et al. Generative spoken dialogue language modeling. *Transactions of the Association for Computational Linguistics*, 11: 250–266, 2023. doi: 10.1162/tacl_a_00545.

Radoslaw Niewiadomski, Jennifer Hofmann, Jérôme Urbain, Tracey Platt, Johannes Wagner, Bilal Piot, Huseyin Cakmak, Sathish Pammi, Tobias Baur, Stephane Dupont, et al. Laugh-aware virtual agent and its impact on user amusement. In *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*, pages 619–626, 2013. doi: 10.5555/2484920.2485018.

Eva E Nwokah, Hui-Chin Hsu, Patricia Davies, and Alan Fogel. The integration of laughter and speech in vocal communication: A dynamic systems perspective. *Journal of Speech, Language, and Hearing Research*, 42(4):880–894, 1999. doi: 10.1044/jslhr.4204.880.

Florian Pecune, Maurizio Mancini, Beatrice Biancardi, Giovanna Varni, Yu Ding, Catherine Pelachaud, Gualtiero Volpe, and Antonio Camurri. Laughing with a virtual agent. In *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*, pages 1817–1818, 2015. URL `https://dl.acm.org/doi/pdf/10.5555/2772879.2773452`.

Robert R Provine. Laughter. *American Scientist*, 84(1):38–45, 1996. URL `http://www.jstor.org/stable/29775596`.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020. URL `https://www.R-project.org/`.

Hannes Ritschel, Ilhan Aslan, Silvan Mertes, Andreas Seiderer, and Elisabeth André. Personalized synthesis of intentional and emotional non-verbal sounds for social robots. In *Proceedings of the 8th International Conference on Affective Computing and Intelligent Interaction*, pages 1–7, 2019. doi: 10.1109/ACII.2019.8925487.

Shane Saunderson and Goldie Nejat. How robots influence humans: A survey of nonverbal communication in social human–robot interaction. *International Journal of Social Robotics*, 11:575–608, 2019. doi: 10.1007/s12369-019-00523-0.

Kimberly Sellers, Thomas Lotze, and Andrew Raim. *Package 'COMPoissonReg'*, 2022. URL `https://cran.r-project.org/web/packages/COMPoissonReg/index.html`. R package version 0.8.0.

Gijsbert Stoet. PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4):1096–1104, 2010. doi: 10.3758/BRM.42.4.1096.

Gijsbert Stoet. PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1):24–31, 2017. doi: 10.1177/0098628316677643.

Noé Tits, Kevin El Haddad, and Thierry Dutoit. Laughter synthesis: Combining seq2seq modeling with transfer learning. *Proceedings of INTERSPEECH*, pages 3401–3405, 2020. doi: 10.21437/Interspeech.2020-1423.

Jürgen Trouvain and Marc Schröder. How (not) to add laughter to synthetic speech. In *Proceedings of the Tutorial and Research Workshop on Affective Dialogue Systems*, pages 229–232, 2004. doi: 10.1007/978-3-540-24842-2_23.

Jürgen Trouvain and Benjamin Weiss. Thoughts on the usage of audible smiling in speech synthesis applications. *Frontiers in Computer Science*, 4, 2022. doi: 10.3389/fcomp.2022.885657.

Bekir Berker Türker, Zana Buçinca, Engin Erzin, Yücel Yemez, and T Metin Sezgin. Analysis of engagement and user experience with a laughter responsive social robot. In *Proceedings of INTERSPEECH*, pages 844–848, 2017. doi: 10.21437/Interspeech.2017-1395.

Brian J. Zhang and Naomi T. Fitter. Nonverbal sound in human-robot interaction: A systematic review. *Journal of Human-Robot Interaction*, 2023. doi: 10.1145/3583743.