

How about laughter?

Perceived naturalness of two laughing humanoid robots

Christian Becker-Asano Takayuki Kanda Carlos Ishi
Hiroshi Ishiguro

Advanced Telecommunications Research Institute International
Intelligent Robotics and Communications Laboratories
2-2-2 Hikaridai, Keihanna Science City, 619-0288 Kyoto, Japan

<http://www.irc.atr.jp/>

Abstract

As humanoid robots will have to behave socially adequate in a future society, we started to investigate laughter as an important para-verbal signal influencing relationships among humans quite easily. As a first step we investigate, how humanoid robots might laugh within a situation, which is suitable for laughter. Given the variety of human laughter, do people prefer a certain style for a robot's laughter? And if yes, how does a robot's outer appearance affect this preference, if at all?

Accordingly, we combined six recordings of female laughter with body movements for two different humanoid robots with the aim to evaluate their perceived naturalness using two types of video-based surveys. We not only found that people indeed prefer one type of laughter when being forced to choose, but the results also suggest significant differences in the perceived naturalness of laughter with regard to the participant's cultural background. The outer appearance seems to change the perceived naturalness of a humanoid robot's laughter only on a global level.

It is evident, however, that further research on this rather unexplored topic is needed as much as it promises to provide valuable means to support the development of social robots.

1. Motivation and related work

Laughter in humans has a socio-emotional function [10] and two major kinds of laughter can be distinguished, namely, aversive and friendly laughter [5]. Aversive laughter is also referred to as self-generated and emotionless "Non-Duchenne" laughter [2] and can be linguistically described as "laughing at so./sth.". Friendly laughter, on the contrary, (linguistically circumscribed as "laughing with so./sth.") is characterized as stimulus-driven and emotional

valenced "Duchenne" laughter [2]. Based on this distinction it is important to avoid a human's interpretation of a robot's laughter as negative, i.e. aversive, because we aim to establish positive human-robot relationships.

Laughter belongs to the more general class of "raw affect bursts" [11] and for a special type of raw affect burst, described as "contempt laughter", [12] reports an auditory recognition rate of 77%. In general terms, "raw affect bursts" are less conventionalized and less symbolic than "affect emblems" [11]. The latter consist of a certain verbal content and the authors of [13] could already show that the humanoid robot "Robovie-II" was allowed to have a slow response time, when it made use of "conversational fillers" such as the Japanese expression "etto" (resembling something similar to "well..." or "uh..." in English).

Furthermore, two of these robots successfully performed a Japanese kind of stand-up comedy enacting more laughter in human observers than a comparable performance by human actors [6]. Although some "affective sounds" have been used to improve affective interaction with a virtual agent [9], to the best of our knowledge, the use of laughter in robots or virtual agents as a powerful, para-verbal, social signal has not yet been systematically investigated.

Thus, the questions underlying the research presented subsequently can be stated as follows: How might a humanoid robot laugh that takes part in a situation, in which such laughter is natural to occur, e.g. that of responding to a joke? Given the variety of human laughter, do people prefer a certain style of a robot's laughter in such a situation? And if yes, how does a robot's outer appearance affect the formation of this preference?

The remainder of this paper is structured as follows. In Section 2 an online survey is described, its results are discussed and arguments for the second survey are given. Section 3 describes the second survey conducted with Japanese high school students and its results are presented. Finally,

in Section 4, we compare and discuss the results of both surveys, before in Section 5 conclusions are drawn.

2. The online survey

As motivated above, humans most often laugh within a social context [3], which in turn influences the style of their laughter [1]. Accordingly, we decided to establish a precise situational context in which our robots would start laughing, i.e. we told the observers that the robots would be laughing in response to a joke.

By choosing such a non-serious situational context we also tried to avoid that humans might interpret the robot’s laughter as negative, aversive laughter. Thus, we could focus on what kind of recorded human laughter would be judged as most natural when being produced by the two humanoid robots “Robovie-II” [8] and “Robovie-R2” [14] (cp. Figures 1(a) and 1(b) respectively). We used two different Robovie versions in order to check for a possible interaction effect between the robot’s outer appearance and the perceived naturalness of its laughter.

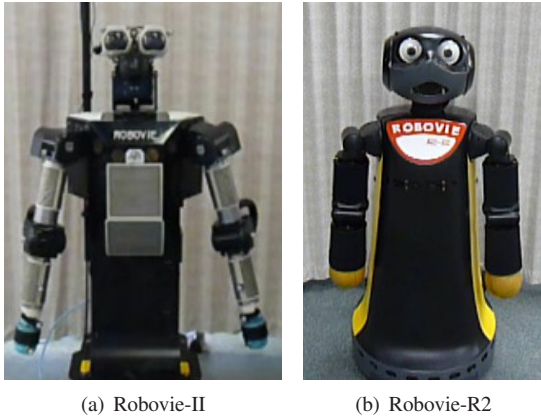


Figure 1. The two humanoid robots in their rest postures adopted before and after laughing

2.1. Design

The participants, first, had to choose one of the languages German, English, or Japanese, in which the survey was then presented. Next, the situational context was described in that the participants could read the complete joke and were told that the robot would start laughing in response to that joke. Further written explanations together with a screenshot of the interface (cp. Figure 2) were given as well and the participants could listen to the last sentence of the joke, which afterward was played in the beginning of each video.

Next, the participants were requested to provide some personal data, such as their gender, their age, their nationality, and their email address. The latter was only used to confirm each participant’s identity by email and to prevent multiple participations.

Finally, each robot’s laughter had to be judged based on the pairwise sequential presentation of a total of six short video clips per robot (cp. Figure 2). The content of these video clips will be described in the following.

Each robot’s laughter was followed by the Japanese exclamation “Arieheh!” (meaning “unbelievable”), which was rendered by a speech synthesizer. Five laughter sounds were manually chose out of a total of 402 Japanese laughters that originate from dyadic smalltalk recordings [7]. The restriction to female laughter for both robots was motivated, first, by the robot’s speech synthesis being based on a female voice as well and, second, by our belief that there were still enough variations for laughter realization. Because child-like laughter seemed to fit to our humanoid robots as well, we pitched one sample up by 25% (keeping its duration constant using the GoldWave software [4]) to produce an artificial, more child-like laughter.

The length of the laughter ranged from 0.9 seconds for laughter number six to a maximum of 1.74 seconds in case of laughter number five (cp. Table 1). The following characteristics describe the different laughters (labeled L_x):

- L_1 : Very high pitch; artificial, child-like laughter; rather constant pitch contour; six pulses; gender ambiguous
- L_2 : Same as L_1 , but with mid-height pitch; female
- L_3 : Starting with higher pitch and continuously decreasing to mid-height pitch; seven pulses; female
- L_4 : Rather low pitch; “smoky” voice quality; eight pulses; female
- L_5 : Mid-height pitch; quickly alternating in- and exhaling; six pulses; female
- L_6 : High pitch; three pulses; rather short; female

Graphical representations of each laughter’s amplitude can be found in Table 1.

Table 1. Amplitudes, lengths, and number of pulses for each of the six female laughter bouts

Laughter 1	Laughter 2	Laughter 3
1.25 seconds, 6 pulses	1.25 seconds, 6 pulses	1.47 seconds, 7 pulses
Laughter 4	Laughter 5	Laughter 6
1.48 seconds, 8 pulses	1.74 seconds, 8 pulses	0.9 seconds, 4 pulses

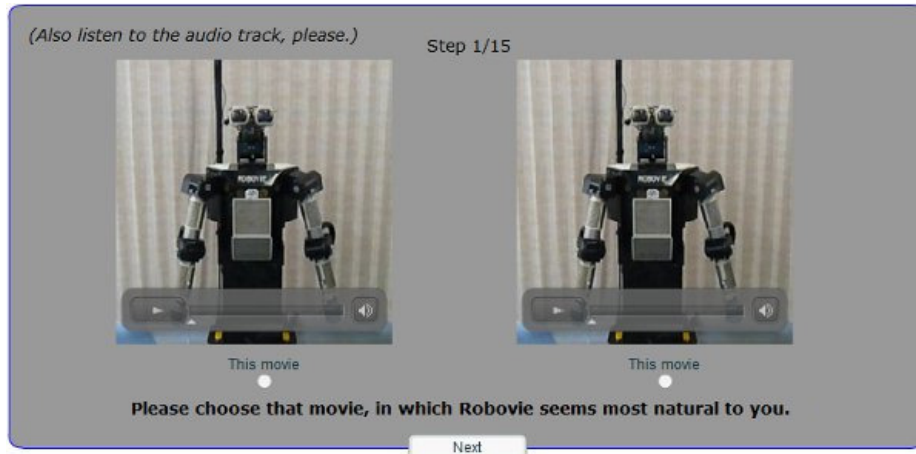


Figure 2. A screen-shot of the online interface: The presented movies were always different to each other and participants were forced to chose one of them, before they could proceed to the next pairing by pressing “Next”.

These six laughers were systematically combined with videos of the two robots, in which they both performed the same movements: after they listened to the last sentence of the joke and while the laughers were being played, they moved their heads backward to the left and lifted their arms resembling an “open-hand” gesture (cp. Figure 3). With finishing their laughter they moved back into their initial positions looking straight into the camera with their arms next to their bodies and finally they said “Ariehen!” without moving at all.

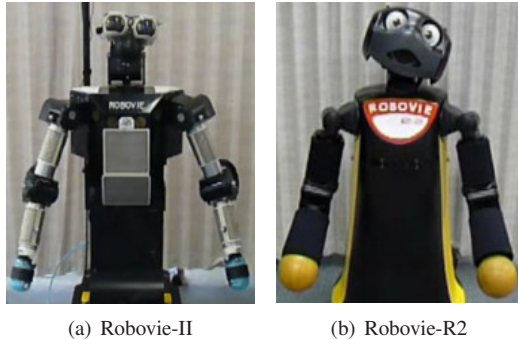


Figure 3. Head and arm movements of the robots during laughter

2.2. Procedure

The two robots were presented in two independent online surveys, which are subsequently labeled “survey A” for Robovie-II and “survey B” for Robovie-R2. In each survey the videos were presented pairwise in random order, such that each of the corresponding six videos was presented in combination with each other video. Accordingly, both surveys followed the within-subject, forced-choice design, because in the resulting total of 15 pairs of videos per survey the participants were forced to decide for that video, in

which the robot seemed to behave most naturally.

Table 2. Distribution of gender per survey

	male	female	Σ
survey A	30	20	50
survey B	25	8	33
Σ	55	28	83

*Note that one participant did not reveal his or her gender.

Fifty participants took part in survey A (30 male, 20 female) and 34 participants in survey B (25 male, 8 female; cp. Table 2). Four participants joined both surveys. Of these 84 participants 24 (12 survey A, 12 survey B) originate from Asia, 22 (17 survey A, 5 survey B) from America, and 38 (21 survey A, 17 survey B) from Europe (cp. Table 3).

Table 3. Distribution of participants’ origin per survey

	Asian	European	American	Σ
survey A	12	21	17	50
survey B	12	17	5	34
Σ	24	38	22	84

Comparing the age distributions between surveys with respect to the participants’ cultural backgrounds (cf. Figure 4) no significant differences are shown with all mean values between 27 and 32.4 years.

2.3. Results

Not taking the robots’ different appearances into account by pooling the data of both surveys, laughter number two containing six laugh pulses was chosen most often (cp. Figure 5, mean 3.6, standard deviation (STD) 1.13) and dominates all other laughers (i.e., the second best laughter is judged significantly lower (mean 3.07, STD 1.07,

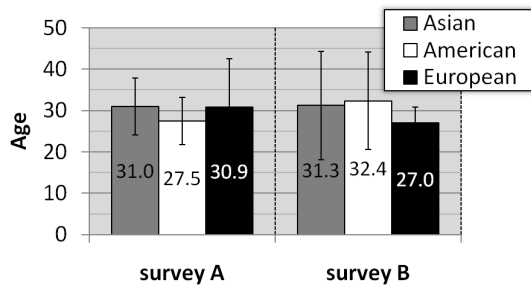


Figure 4. Mean values and standard deviations of the age distributions per survey depending on cultural background

$p = 0.006^1$). Interestingly, when this laughter is pitched up to artificially produce child-like laughter (i.e. laughter number one), it belongs to the group of three laughters evaluated as most unnatural (mean 2.04, STD 1.51).

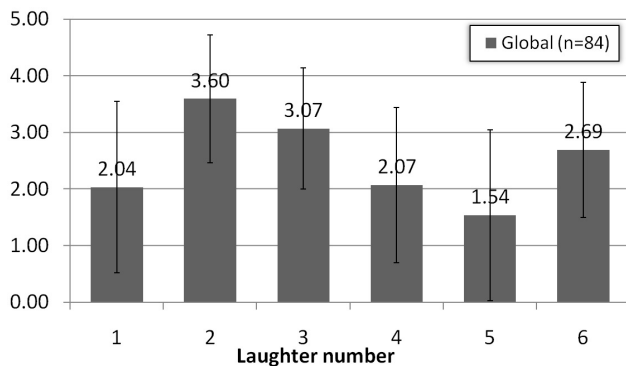


Figure 5. Global mean values with standard deviations per laughter

The different outer appearances of the two robots seem to have no effect (cp. Figure 6). Only laughter number five, which contains a lot of breathing, was judged significantly less natural ($p = 0.02$) for Robovie R2 (mean 1.12, STD 1.2) than for Robovie II (mean 1.82, STD 1.64). Neither any other statistically significant differences between robots nor any global gender effects were found. Although a difference between robots seems to exist with regard to the naturalness of the child-like laughter number one (survey A: mean 1.78, STD 1.45; survey B: mean 2.47, STD 1.55, cp. Figure 6), this result is not statistically significant ($p = 0.054$).

Taking a global look at intercultural differences by pooling the data of both surveys again, we found that the participants originating from the American continent ($n=22$, mean 1.45, STD 1.41) judged the child-like laughter number one as significantly less natural than did the Asian participants ($n=24$, mean 2.42, STD 1.47, $p = 0.03$, cp. Figure 7). The respective means of either of these groups, however, do not differ significantly from the mean rating of the European

¹Our analysis is based on a two-tailed t-test assuming unequal variances with a 5% level for statistical significance.

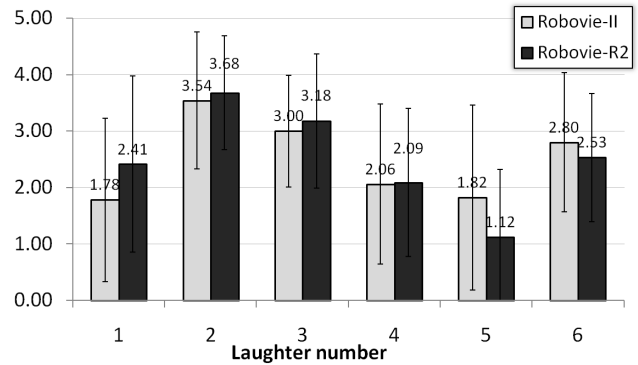


Figure 6. Mean values with standard deviations to compare the results for Robovie-II with those for Robovie-R2

participants (mean 2.04, STD 1.52) for this child-like laughter. Furthermore, participants of each of the three groups evaluated laughter number two as most natural.

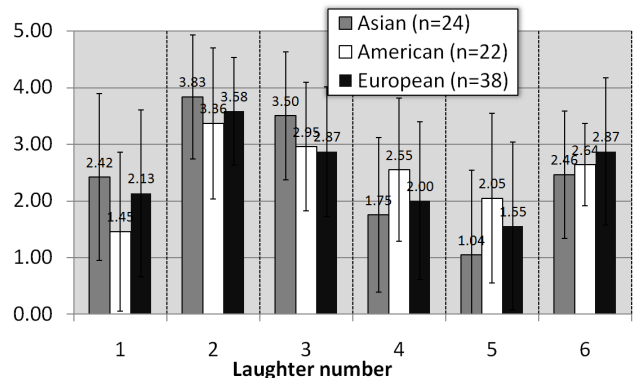


Figure 7. Mean ratings with standard deviations split up according to participants' cultural background distinguishing Asian, American, and European origin

2.4. Discussion

In summary, the robot's outer appearance seems to have a much smaller effect on the perceived naturalness of laughter than we expected. Probably any real differences are dominated by the judged naturalness of the different laughters themselves, but some participants also reported that they thought any kind of male laughter would fit better to our humanoid robots than any of the female laughters that we presented. Furthermore, the forced-choice design of our study might have overshadowed any inter-robot difference in the perceived naturalness of laughter as well. Aiming to investigate these open questions, we decided to acquire additional data by conducting a second survey.

3. High school survey

During a lecture at a Japanese high school given by the first author, 36 students were asked to provide their opinions

about how well each of the six laughters fits to each of the two robots. The laughters as well as the robots were the same as in the online survey described in Section 2.



Figure 8. The mode of presentation during the lecture at a Japanese high school

3.1. Design and procedure

Although the same six laughters in combination with the same movements of the robots had to be judged, the way of presentation was different from the one applied in the online survey. In the lecture room a projector was used (cp. Figure 8) to present 12 video clips of the laughing Robovies without giving any further context information. The videos were grouped by type of laughter, i.e. six video pairs were shown each of the pairs consisting of one video for Robovie-II and the according one for Robovie-R2. After each pair the students were asked to independently rate the degree to which the respective laughter fits to each of the two robots. Their rating was based on a five-point scale ranging from minus two (labeled with “did not fit”) to plus two (labeled with “did fit”). We also gathered each student’s age and gender and invited them to write down comments.

A total of 36 students took part in this survey (26 male and 10 female). Nine of them were 16 years old (all of them being male students) and the remaining 27 students were 17 years old.

3.2. Results

Most surprisingly, laughter number two was rated to not fit to either of the two Robovies (mean -1.56 and STD 0.81 for Robovie-II, mean -0.36 and STD 1,29 for Robovie-R2, cp. Figure 9). This result contradicts the results of the previous online survey (cp. Figure 5), in which the same laughter was judged as most natural. Moreover, for Robovie-II the high school students judged this laughter as most unnatural.

Only laughter number six in combination with Robovie-R2 gained a positive evaluation by the students with its mean being significantly different from the mean of the second best laughter number one for Robovie-R2 (mean 0.08, STD 1.4, $p < 0.01$). No significant gender effects occurred.

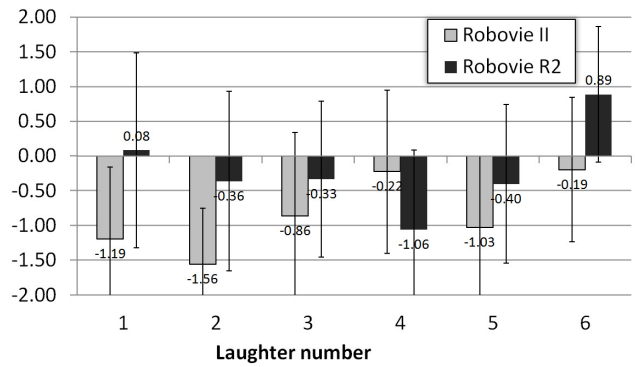


Figure 9. Mean ratings with standard deviations of the laughter ratings comparing Robovie-II and Robovie-R2

4. Discussion

The results of our surveys provide us with valuable insights into the question of how a robot might laugh. We are well aware that due to the design of our surveys the possible conclusions are limited. Nevertheless, we derive the following ideas from our data, which inform our future research:

- **Laughter & robot’s gender:** Especially the comments given by the high school students reveal that they expected Robovie-II to produce male laughter, whereas this is not the case for Robovie-R2. Accordingly, we might compare the effect of male vs. female laughter for both Robovies in future research.
- **Laughter & situational context:** A possible reason for the difference in the results of the online and the high school survey might be that for the latter no concrete situational context, in which the robots would start laughing, was explained to the students. Furthermore, by allowing participants of future studies to directly interact with Robovie instead of only presenting video clips, they could be much more involved in the situation probably leading to a significant change in their impression of a laughing Robovie.
- **Laughter & intercultural differences:** The results of the online survey led to the assumption that there might exist intercultural differences in perceived naturalness of laughing humanoids. With the participants of the high school survey being all Japanese and of nearly the same age, we currently plan to conduct a very controlled, similar study at a German high school in the aim to explicitly check for intercultural differences.
- **Laughter & its degree of artificiality:** Although the mean evaluation of the more artificial, child-like laughter number one only gained a slightly positive evaluation in case of Robovie-R2 in the high school survey, the question arises, whether completely artifi-

cial laughter might be judged even more natural than recorded, human laughter.

One major point of skepticism can be brought up when comparing the results of both types of surveys:

Why was laughter number two judged as most natural for both Robovies in the online surveys (cp. Figure 6), but as least natural for Robovie-II and as unnatural for Robovie-R2 by the high school students (cp. Figure 9)?

At least three reasons can be speculated about, which might explain this difference:

1. Differences in the design of the two surveys:
 - (a) For the online survey we provided the participants with a concrete situational context (i.e. laughing in response to a joke) and the two Robovies said “Arieheh!” in the end of each movie. In the high school survey we only presented laughing robots without any context.
 - (b) The online survey followed a forced choice, randomized design and participants rated the naturalness of only one of the two robots at a time. The high school students, in contrast, could rate the performances of both Robovies in comparison based on a five point scale and the order of presentation was not randomized.
2. Differences in age and cultural background of the participants between surveys: The age as well as the cultural background of the participants of the online survey naturally varied much more than in case of the quite uniform group of high school students.

Reason 1(a) might be worth further investigation, because it is critical to gain an understanding of how changes in situational contexts might influence the perceived naturalness of laughter. This is also related to the critical question of how to avoid a negative effect of a robot’s laughter as motivated in Section 1.

Because reason 1(b) might also be responsible for the differences in the results, we have to take care, however, not to over-interpret these differences. In order to investigate possible effects of age and cultural background (i.e. reason 2) we plan to conduct another survey at a German high school following the same design as reported in Section 3.

5. Conclusions

In summary, people in a forced choice design indeed prefer one type of laughter. Outer appearance, however, seems to have only a global effect on the perceived naturalness of laughter in that female laughter appears to fit better to

Robovie-R2 than to Robovie-II, because the latter is perceived to be male. Furthermore, interesting intercultural differences appeared that inform our future research.

Thus, we feel encouraged to keep investigating the applicability and effects of laughter in Human-Robot Interaction and we believe that many interesting as well as challenging new questions will arise along the way.

6. Acknowledgments

The first author is supported by a post-doctoral fellowship of the Japan Society for the Promotion of Science.

References

- [1] N. Campbell. Whom we laugh with affects how we laugh. In *Interdisciplinary Workshop on The Phonetics of Laughter*, 2007.
- [2] M. Gervais and D. S. Wilson. The evolution and functions of laughter and humor: A synthetic approach. *The Quarterly Review of Biology*, 80:395–430, 2005.
- [3] P. Glenn. *Laughter in Interaction*. Cambridge University Press, 2003.
- [4] GoldWave Incorporated. GoldWave Audio Editor. <http://www.goldwave.com>, 2009.
- [5] K. Grammer and I. Eibl-Eibesfeldt. The ritualisation of laughter. In *Natürlichkeit der Sprache und der Kultur*, chapter 10, pages 192–214. Brockmeyer, 1990.
- [6] K. Hayashi, T. Kanda, T. Miyashita, H. Ishiguro, and N. Hagita. Robot manzai: Robot conversation as a passive-social medium. *International Journal of Humanoid Robotics*, 5:67–86, 2008.
- [7] C. Ishi, H. Ishiguro, and N. Hagita. Analysis of inter- and intra-speaker variability of head motions during spoken dialogue. In *Proceedings of the International Conference on Auditory-Visual Speech Processing*, pages 37–42, 2008.
- [8] T. Kanda, H. Ishiguro, T. Ono, M. Imai, and K. Mase. Multi-robot cooperation for human-robot communication. In *International Workshop on Robot and Human Interactive Communication*, pages 271–276, 2002.
- [9] H. Prendinger, C. Becker, and M. Ishizuka. A study in users’ physiological response to an empathic interface agent. *International Journal of Humanoid Robotics*, 3(3):371–391, 2006.
- [10] R. R. Provine. Laughing, tickling, and the evolution of speech and self. *Current Directions in Psychological Science*, 13:215–218, 2005.
- [11] K. Scherer. Affect bursts. In *Emotions: Essays on emotion theory*, pages 161–196. Erlbaum, 1994.
- [12] M. Schröder. Experimental study of affect bursts. *Speech Communication*, 40:99–116, 2003.
- [13] T. Shiwa, T. Kanda, M. Imai, H. Ishiguro, and N. Hagita. How quickly should communication robots respond? In *HRI’08*, 2008.
- [14] Y. Yoshikawa, K. Shinozawa, H. Ishiguro, N. Hagita, and T. Miyamoto. Responsive robot gaze to interaction partner. In *Proceedings of robotics: Science and systems*, 2006.