Evaluation of artificial mouths in social robots

Álvaro Castro-González, Jonatan Alcocer-Luna, Maria Malfaz, Fernando Alonso-Martín, and Miguel A. Salichs

***Abstract*—The external aspects of a robot affect how people behave and perceive it while interacting. In this paper, we study the importance of the mouth displayed by a social robot and explore how different designs of artificial LED-based mouths alter the participants' judgments of a robot's attributes and their attention to the robot's message. We evaluated participants' judgements of a speaking robot under 4 conditions: (1) without a mouth; (2) with a static smile; (3) with a vibrating, wave-shaped mouth; and (4) with a moving, human-like mouth. Seventy-nine participants evaluated their perceptions of an on-video robot showing one of the 4 conditions. The results show that the presence of a mouth, as well as its design, alters the perception of the robot. In particular, the presence of a mouth makes the robot to be perceived more lifelike and less sad. The human-like mouth was the one participants liked the most and, along with the smile, they were the friendliest ones. On the contrary, participants rated the mouthless robot and the one with the wave-like mouth as the most dangerous ones.**

***Index Terms*—social robot, mouth, human–robot interaction, HRI, LED-based mouth, robot mouth, evaluation, human factors**

I. INTRODUCTION

THE importance of social robotics is growing day by day. A large number of new applications have appeared where robots and people coexist and interact with each other. In these situations, it is crucial that social robots be accepted and not rejected by people. In order to ease this, robots need to be perceived as social beings able to communicate and interact with humans. This implies that (i) robots have to follow the behavioral norms to be socially accepted, and (ii) they have to be perceived as living entities rather than mere machines. For both purposes, the robot's appearance is a key aspect.

Many researchers have investigated how the external appearance of a robot alters the users' perceptions of it during an interaction, and hence the interaction itself [1]. For example, Kwak showed that human-like robots provide a stronger social presence and enriching interactions [2]. Kim et al. showed that anthropomorphic robots obtained more donations to a nonprofit fund-raising organization than functional ones (those with very limited social features) [3]. In many of these studies, the interaction capabilities of the robots were inspired by those existing in humans (e.g. movements of arms or utterances). However, robots can be endowed with new interaction modalities that humans do not have, such as screens or coloured lights. Another example of an extended interaction modality in robots is a matrix of LEDs to display an artificial mouth that is synchronized with the robot's speech. This artificial, LED-based mouth can be designed to resemble a human mouth or to display innovative forms, such as waves.

The authors are with the Department of System Engineering and Automation, Universidad Carlos III de Madrid, Robotics Lab, 28911-Madrid, Spain e-mail: acgonzal@ing.uc3m.es

Manuscript received January, 2017

Mouths are important for human–human interaction: the motion of the lips helps humans to understand the message and identify who is talking. We believe that human–robot interaction (HRI) can benefit from dynamic mouths or moving lips[1], too.

In the research presented in this paper, we studied the importance of having a mouth in a robot and explored how different forms of artificial, LED-based mouths alter the participants' judgments of a robot's attributes and their attention to the robot's message. In particular, we focused on analysing the effects of having a mouth (as opposed to not having one), its dynamics (a static mouth in contrast to a dynamic mouth), and its naturalness (i.e. similarity to mouths of existing living beings). Accordingly, we studied participants' judgements of a speaking robot under 4 conditions: (1) without a mouth; (2) with a static smile; (3) with a dynamic, wave-like mouth; and (4) with a dynamic, human-like mouth. In particular, we analysed the users' judgements in terms of the robot's anthropomorphism, animacy, safety, and likeability; in addition, we analysed how we can draw the user's attention to the robot's changing its mouth. Our ultimate goal is to investigate how, by means of the robot's mouth, to improve the user's engagement and experience while interacting with a social robot.

The rest of this paper is structured as follows. In the next section (Section II), we explain the relevant role of a mouth in human–human interaction, as well as previous research that has evaluated different mouths in robots. Next, in Section III, we present a review of the most typical types of mouths that have been implemented in social robots. Section IV gives details about the robot employed in the experiment and the system we have developed to generate different mouths using a matrix of LEDs. The experiment is described in Section V and the statistical results are included in Section VI. Then, we analyse the results and discuss their implications (Section VII). Finally, the paper is concluded in Section VIII.

II. PERCEPTION OF THE MOUTH IN COMMUNICATION

Researchers have shown that the perception of the mouth while humans communicate is a key element in understanding their speech. The well-known McGurk Effect proved this in the 1970s [4]. Mcgurk and Macdonald showed that the perception of speech changed depending on the visual input: adults reported hearing different utterances when the lip motions changed even though the actual sound did not. This experiment revealed clearly that there is a link between the utterance perceived and the visual information received, and both sources of information are closely related in the brain [5]. However, how they are combined or which one is the

[1]We will use the term *dynamic mouth* to describe those types of mouths that change their shape in sync with the robot's utterances.

most important is not clear and is outside the scope of this paper.

This important feature in human communication has been exploited in many fields. From 1926 [2] to the present day, film makers have provided animated characters with a mouth synchronized with their speech. These virtual characters open, close, and shape their mouths while talking [6, 7, 8, 9].

Researchers working on multimodal communication have benefited from the joint processing of multiple sources of information, both for perception and expressiveness, obtaining more reliable results [10]. In the area of social robotics, multimodal communication is paramount. Many works have focused on the importance of combining visual and auditory information. In this line, several authors have worked on synchronizing lip motion (visual) with speech (auditory), and on controlling the mouth movements [11, 12, 13]. Cid et al. stated that people's perception of the robot and understanding of the robot's utterances can be improved by combining two sources of information: (i) auditory cues (e.g. pitch, pauses, or emphasis) and (ii) visual clues (e.g. lip movements, facial expressions, or body poses)[14]. In the same study, researchers analysed how different mouths and sync algorithms affect the interaction. Cid et al. evaluated three types of mouths during an 'interaction experience': (i) the physical mouth of a 'teddy bear' robot; (ii) the mouth of an animated, virtual teddy bear; and (iii) a mouth formed by a matrix of LEDs operating as a VU-meter[3] in an animated, virtual robot head. The authors claimed that the physical mouth was more engaging and users were more attentive under this condition. However, the limited number of participants (15) and the high variability among the conditions (in some conditions, a real, animal-like robot was compared with an animated, virtual robot) make us believe that these results could be due to other reasons.

Hyung et al. evaluated the importance of lip-sync in an android robot where the mouth shape changed according to the vowel sounds played by the android [15]. They observed that although the shape of the mouth and the lip-sync timing were important components, the noise in the environment was the most relevant factor affecting the quality of the communication.

Summarizing, it seems that visual clues could help to improve human–robot interaction. Thus, it is likely that a proper dynamic mouth helps to foster the robot's attributes, to identify the subject speaking, to improve the understanding of its utterances, and to keep the users' attention on what the robot says. However, many questions have not been properly answered yet: is it really important to endow social robots with a mouth? What are the benefits of a dynamic mouth? And, finally, does the shape of the mouth matter? In this paper we aim to shed light on these questions.

[2]In 1926, Dave Fleischer presented his short animation film titled *My Old Kentucky Home* containing dialogs of animated characters that moved their lips according to the utterances.https://en.wikipedia.org/wiki/My_Old_Kentucky_Home_(film)

[3]A VU-meter is a device displaying a representation of the level of an audio signal, similar to an equalizer in audio equipment.

III. Mouths in social robots

Although most social robots have a mouth, some still lack one. This is the case, for instance, with some functional robots (non-anthropomorphic) or the Keepon robot, a minimalist, little, yellow robot with just two eyes and a nose [16].

However, in the majority of the robots endowed with a head, robot designers have included a mouth. The simplest cases are those robots with a physical, rigid mouth which is part of the robot's shell, such as the robot Pepper, by Aldebaran Robotics and Softbank, with its static mouth [17, 18, 19].

In order to foster their expressiveness, other social robots have mechanical mouths or lips. This is the case with Flash [20] or Nexi, a robot with a highly articulated face including three degrees of freedom in the jaw [21]. Another well-known example is the robot iCat with its two motorized, rubber lips [22].

Android robots, that is, robots that clearly resemble humans, include realistic mouths with lips, teeth, and even a tongue [23]. Usually, an android's mouth moves, accompanying its speech. A particular case is the Furhat robot head which has a human-like back-projected face with computer animated mouth movements [24].

On the other hand, robots can be endowed with iconic mouths, i.e. mouths that are completely different from those present in nature but evoke them because of their location in the head or their operation. For example, the robot Robi [4] has a light that turns on and off in sync with the robot's utterances.

In this regard, many robots have LED-based mouths. These robots usually use a VU-meter similar to an equalizer, to turn on and off an array of LEDs. This can be observed in the robots Mini [25] and SnackBot [26], for example.

Another type of LED-based mouth has a matrix of LEDs. For example, Papero has a matrix of RGB LEDs to display a coloured mouth [27]. These mouths are easy to control and flexible enough to display not only the mouth but also other shapes, icons, or even text messages.

Lately, new robots have started to include screens to display appealing faces. For instance, the robots Buddy[5] [28] and DragonBot[6] [29] use a screen to display two eyes and a mouth. These devices are very flexible and allow displaying any figure or animation; however, the control is complex and requires dedicated software.

Although many robots are equipped with a mouth, or lips, the literature presents only a few studies of the relevance of the mouth in robotics. In particular, to the best of our knowledge, there is no thorough study of the importance of providing social robots with a mouth and how this mouth should be designed.

IV. The LED-based mouth in the MBot robot

In this paper we study how different robotic mouths designed in a matrix of LEDs affect some aspects of HRI. This research has been developed in the context of the European project MOnarCH (Multi-Robot Cognitive Systems Operating

[4]Robi TM, DeAgnostini, http://deagostini.jp/rbi/

[5]http://www.bluefrogrobotics.com/en/buddy/

[6]http://robotic.media.mit.edu/portfolio/dragonbot/

in Hospitals) [7]. In this project, a team of robots operates in the pediatric ward of an oncological hospital. In this scenario, the robots operate for long periods, interacting with the different people they encounter. The MOnarCH robots (MBots) (see Figure 1) have been designed to establish social interactions mainly with the hospitalized children.

Figure 1. Scenario of the MOnarCH project, where the robot mBot is interacting with a hospitalized child.

Considering the scope of the MOnarCH project, it is crucial that MBots are perceived as living entities with high expressive capabilities, able to communicate successfully. To achieve this, we take advantage of all the interaction modalities available to the robots, including the matrix of LEDs. We believe that during the operation of the robots, their mouth is a key modality to engage children in HRI, just as it is in human–human interaction (as just explained in Section II).

The matrix of LEDs in an MBot's head is a flexible mechanism that allows us to easily display different types of mouths according to the situation the robot is. For example, the robot can show a big smile when it is playing with a kid or an open mouth when it is surprised because of a kid's unexpected movement.

Besides, we can use these LEDs to display other figures that do not resemble the mouths of humans or animals. For example, we can use them to show a random point cloud, a geometrical shape (such as a wave), or even a text message.

In what follows, we will describe the mechanism used by the MBots to display different LED-based mouths.

A. Under the hood: Generating on-the-fly, LED-based mouths

On the front side of the robot's head, an MBot is endowed with a matrix comprising 283 red LEDs arranged in 8 rows, as shown in Figure 2 (note that not all the rows contain the same number of LEDs).

In our case, we are interested in achieving coherent configurations of the LEDs that provide mouth-like expressions intended to trigger a natural response from the users interacting

Figure 2. Detailed view of the hardware of an MBot's matrix of LEDs.

with the robot. Consequently, we have considered two different types of artificial mouths:

- *Static mouths* are considered as particular mouth expressions represented by a fixed configuration of the matrix of LEDs, for instance, a smile or a surprised mouth.
- *Dynamic mouths* change their form depending on the audio signal and give the impression of a moving-lip robot.

V. Evaluation

As already mentioned, we wanted, in this research, to investigate the effect of different types of mouths during human–robot interaction. In particular, we analysed the participants' evaluations in terms of anthropomorphism, animacy, likeability, and perceived safety, depending on the robot's mouth. In this section, we present our hypotheses, the experiment we have conducted, and how we evaluated the different LED-based mouths in the MBot.

A. Hypotheses

In order to study the importance of the mouth when social robots communicate and the effects of different designs, we have stated the following hypotheses:

H_1: A *mouthless* robot is perceived as less animated.
H_2: The perception of the robot can be altered by changing its mouth.
H_3: Participants will pay more attention to a speaking robot when it has a dynamic, human-like mouth.

The first hypothesis, H_1, refers to the importance of endowing robots with a mouth, regardless of its shape or movement. H_2 focuses on the type of mouth and how it changes the users' judgments of the robot. Finally, H_3 is related with a particular aspect during the interaction between the robot and an user: attention.

B. Conditions

In order to evaluate these hypotheses, we have considered 4 between-subject conditions. Initially, we wanted to study how important is the existence of a mouth in a robot. Hence, we had to consider the situation where the robot lacks a mouth: we designed the first condition as being when the robot does not show a mouth while speaking. We call this the *no mouth* condition (Figure 3a).

Besides, bearing in mind how humans communicate, we believed that a moving mouth would be more appealing than a static one. To study this, in the second condition, the robot

[7] http://www.monarch-fp7.eu

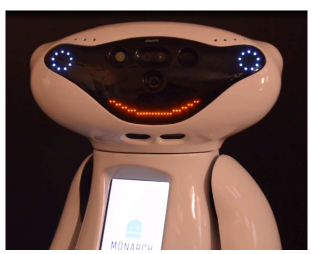
displayed a static smile while talking (Figure 3b). In this condition, when the robot is not talking, there is no mouth displayed.

The third and the fourth conditions correspond to dynamic mouths; that is, a robot whose ‘lips’ were synchronized with its speech. In this case, a mouth formed by a matrix of LEDs allows us to design many styles of mouths or lips. Moreover, we were interested in observing whether or not the design of a dynamic, moving mouth altered the perception of the robot. Consequently, for condition 3, we designed a dynamic mouth that resembled a human-like mouth composed of two moving lips that open and close (Figure 3c). In addition, we created a dynamic mouth that does not exist in nature and whose form is totally different from the previous ones: a wave-like mouth that uses a sinusoidal line changing its phase and amplitude (Figure 3d). This is applied in the fourth condition. As with condition 2, also in conditions 3 and 4, if the robot is not talking, there is no mouth displayed and all the LEDs are off.

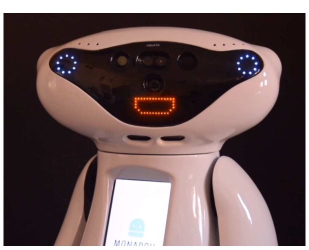
In short, we ended up with 4 conditions:

- No mouth: all LEDs are off.
- Static smile: the matrix of LEDs is configured to show a fixed smile when the robot is talking.
- Human-like mouth: two-lip mouth that opens and closes.
- Wave-like mouth: a sinusoidal mouth with varying phases and amplitudes.

(a) No mouth



(b) Smile



(c) Human-like

(d) Wave-like

Figure 3. The 4 types of mouths considered in the conditions of the experiment: no mouth (3a), smile (3b), human-like mouth (3c), and wave-like mouth (3d).

C. Experimental procedure

In order to obtain the highest number of participants, we decided to conduct an experiment using videos of the MBot robot speaking in Spanish and on-line questionnaires to collect the participants’ impressions.

The videos and the questionnaires were integrated in one web-based application that was accessed via an URL provided to the participant. Figure 4 presents the experimental procedure for each participant. Once the participant accessed the web-based questionnaire, the first page contained the instructions for taking part in the experiment. Here, the participants were informed about the length of the experiment (15 minutes approximately) and that once they started the experiment, they could not pause and resume it. We decided to limit the viewing of the two videos used in the experiment to once per participant to avoid multiple repetitions that could affect the results, in particular those related to the subjects’ attention.

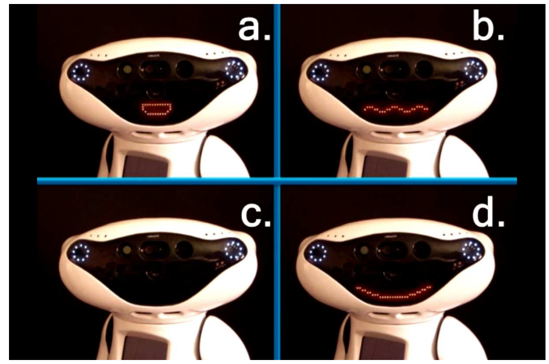


Figure 5. A frame of the video shown in the second part of the experiment. The full video can be watched at https://www.youtube.com/watch?v=5IQtK1Hmfz0

The experiment was divided into two main parts. In the first one (Part I), the participant watched a 6-minute movie where MBot was telling its ‘personal’ story (its functionalities and features, its role in the MOnarCH project, and a description of the laboratory and other robots). In this part of the experiment, depending on the condition the participant was assigned to, the robot displayed the corresponding mouth [8] . The other types of mouths remained unknown to the subjects. After watching the video of the robot speaking, the participants completed a questionnaire to evaluate their impressions of the robot. After that, some questions related to the content of the robot’s speech were included to measure the subject’s attention to MBot.

In the second part (Part II), participants watched a one-minute movie containing 4 videos (one per condition) that were played simultaneously and synchronized with the audio (Figure 5 shows the layout of the movie). After the video, the participants completed another questionnaire with several multiple choice questions, where the participant selected one of the robots.

After that, a few demographic and personal questions had to be answered before the experiment was over. At this point, the questionnaire closed automatically after thanking the participant. Only the data from the participants who completed the experiment were stored and included in the results.

Because the effects of the different mouths might be obscured by a floor effect in the participants’ ratings, the robot

[8] In order to clearly understand the differences between the conditions, the videos used during the first part of the experiment are available online: no mouth (https://www.youtube.com/watch?v=SKzR9mS4epo), smile (https://www.youtube.com/watch?v=mRzriVuoRMA), human-like mouth (https://www.youtube.com/watch?v=2_CPq-H5TZI), and wave-like mouth (https://www.youtube.com/watch?v=KYuMPqKSfwU)

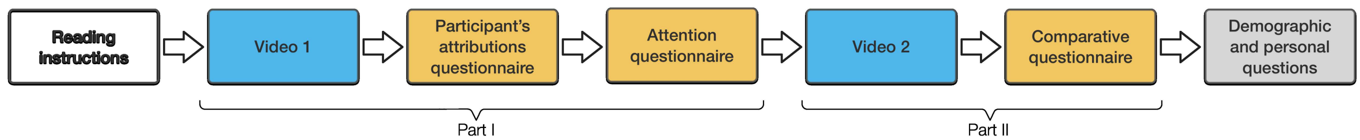


Figure 4. Steps followed by each participant

made slight movements of its head and arms, to promote the robot's lifelikeness, independently of the condition.

D. The participants

In this experiment, 85 voluntary subjects participated in the experiment, but 6 of them were discarded since they did not fully complete the on-line experiment. Thus, data from 79 participants were employed for the statistical analysis. They were recruited using social networks and mailing lists. There were 38 women (44.7%) and 47 men(55.3%) with an average age of 34 years ($SD = 10.3$). Since the call for participants was done in general-interest channels, their backgrounds were very diverse. All participants spoke fluent Spanish. It was crucial that the participants had not met the MBot before or heard about the MOnarCH project, so members from the research group and the MOnarCH project were avoided.

The participants were divided into 4 groups, corresponding to the 4 conditions presented in Section V-B. Each participant was randomly assigned to one of the groups. The number of participants in each group was balanced automatically.

E. The questionnaire

As said before, the participants' ratings were obtained by means of online questionnaires. During Part I, after watching the first video (Video 1 in Figure 4), participants answered (i) questions related to their perception of the robot (*Participant's attributions questionnaire* in Figure 4), and (ii) questions to measure their attention (*Attention questionnaire* in Figure 4). It is worth recalling that the questionnaire was the same for all conditions, but the displayed movie in this Part I changed depending on the condition.

The evaluation of the participants' impressions of the MBot was inspired by the Godspeed Questionnaire [30]. Bartneck et al. designed this questionnaire to measure users' perceptions of robots and it has been extensively used in robotics. In that questionnaire, the items are composed of adjectives related to the robot's anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety. Considering that participants completing this questionnaire were unattended and in order to avoid dropouts due to a long participation, in this paper we consider 12 of these adjectives (mainly from anthropomorphism, animacy, and likeability) and the participants have rated each one using a 7-point Likert scale, where 1 means 'little' and 7 means 'a lot'. The adjectives used in this research are: *natural*, *human-like*, *mechanical*, *inert*, *artificial*, *lively*, *friendly*, *nervous*, *dangerous*, *angry*, *sad*, and *happy*.

In order to evaluate the user's attention, we defined several questions related to the content of the robot's speech (Attention questionnaire). We considered that the more correct answers there were, the more attention the user paid to the robot. We used 10 questions: 8 *yes/no/don't know* questions, and 2 multiple choice questions with 4 options. The full list of these questions can be seen in Table I. All questions have been accurately translated from Spanish.

In the second part of the experiment, where the 4 conditions were shown simultaneously (Video 2 in Figure 4), participants made a direct comparison within all the conditions. Here, 6 comparative questions were presented (*Comparative* questionnaire) and the possible answers were the four conditions (labeled as *a*, *b*, *c*, and *d*). These questions were related to the robots' animacy, safety, and likeability. In the Comparative questionnaire, participants selected (i) the most lively, (ii) the most inert, (iii) the most dangerous, (iv) the most friendly, (v) the most likeable robot, and (vi) the robot they were looking forward to meet.

Finally, demographic and personal questions about age, sex, and background were included.

VI. Results

A. Data analysis for Part I

Here we present the statistical analysis run on the data obtained from the questionnaire for the first part. As already explained, these questions were answered after participants watched the video of the MBot's personal history under their corresponding condition.

To ensure discriminability in the items, we considered questions that were sufficiently sensitive. To that end, we considered items with a standard deviation larger than 1.0. Only one item, *dangerous*, obtained a smaller standard deviation. It was excluded from the subsequent statistical analysis.

We conducted bivariate correlations between all items with pairwise deletion as the missing data treatment. None of the items showed coefficients larger than 0.9, indicating that the questions were not redundant.

We tested for the normality of the ratings of all items considered in the analysis. Since normality was violated in all items, we used non-parametric methods for the statistical analysis. Each item has been considered as a dependent variable (DV) with ordinal values. Besides, since we designed the experiment to use independent groups and independent observations, we ran Kruskal–Wallis H tests to compare the mean ranks for each DV between each condition (i.e. independent variable, IV).

Therefore, 12 Kruskal–Wallis H tests were conducted. They showed that there were statistically significant differences between conditions in the ratings of 4 items: *inert* ($\chi^2(3) = 7.949, p = .047$), *nervous* ($\chi^2(3) = 7.991, p = .046$), *sad* ($\chi^2(3) = 8.639, p = .034$), and *happy* ($\chi^2(3) = 8.734, p =$

Table I
ATTENTION QUESTIONNAIRE

Questions	**Possible answers**
The robot is from a Spanish project	yes - no - don't know
Carlos III University of Madrid is in charge of the MBot's expressivity	yes - no - don't know
The robot is named Monarch	yes - no - don't know
The robot operates in a hospital with children with pneumonia	yes - no - don't know
Maggie and Mini are other robots in the laboratory	yes - no - don't know
The robot plays with hospitalized children in their dormitories	yes - no - don't know
Children and the robot play ball	yes - no - don't know
Carlos III University of Madrid is in charge of the safety navigation of the robot	yes - no - don't know
Where are the educational activities between the robot and children conducted?	infirmary - canteen - corridor - dormitory - don't know
What are the other robots from the lab intended for?	elderly care - car manufacturing - teaching English - playing with children - don't know

.033). The mean ranks of the scores of these 4 items are presented in Table II.

Table II
MEAN RANK VALUES OBTAINED FROM THE 4 KRUSKAL–WALLIS H TESTS FOR *inert*, *nervous*, *sad*, AND *happy*.

	Condition	N	Mean Rank
Inert	no mouth	20	48,80
	smile	22	36,07
	human-like	20	30,98
	wave-like	17	45,35
	Total	79	
Nervous	no mouth	20	35,28
	smile	22	33,82
	human-like	20	41,45
	wave-like	17	51,85
	Total	78	
Sad	no mouth	20	50,53
	smile	22	31,95
	human-like	20	39,45
	wave-like	17	38,68
	Total	79	
Happy	no mouth	20	27,95
	smile	22	45,86
	human-like	20	41,55
	wave-like	17	44,78
	Total	79	

Specifically, post-hoc analyses were performed on the 4 items that presented statistically significant differences using Dunn-Bonferroni tests and considering the adjusted p-values to control the familywise type I error. We found the following results.

In the case of the items *inert* and *nervous*, we did not find significant differences and consequently we cannot state where the differences were.

The post-hoc analysis on the item *sad* came up with higher ratings when no mouth was exhibited by the robot ($Mdn = 2.50; mean\ rank = 50.53$) than it did when a static smile was displayed ($Mdn = 1.00; mean\ rank = 31.95$), and the differences were statistically significant ($p = .021$). No other significant differences were found for the *sad* item.

For *happy*, the *no-mouth* group scored significantly lower than the *smile* ($Mdn_{no-mouth} = 3.00$, $mean\ rank_{no-mouth} = 27.95$; $Mdn_{smile} = 5.00$, $mean\ rank_{smile} = 45.86; p = .044$) . No other significant differences were found for the *happy* item.

Figure 6 shows all the statistically significant differences found by the *Participant's attributions questionnaire*.

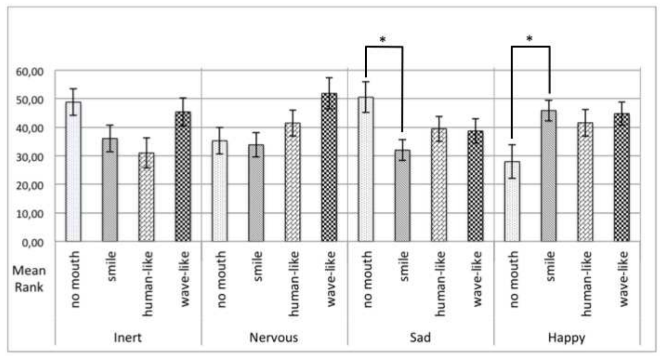


Figure 6. Mean ranks obtained from the Kruskal–Wallis H tests for the items *inert*, *nervous*, *sad*, and *happy*; * indicates significant differences ($p < .05$) considering the adjusted p-values.

In relation with the 10 questions about the MBot's story included in the *Attention questionnaire*, we considered the number of right answers as the DV. After checking that it was approximately normally distributed for each condition, we confirmed that there were no outliers in our data by following the Outlier Labeling Rule [31]. There were no statistically significant differences between the category means in the number of right answers, as determined by one-way ANOVA ($F(3, 75) = 2.007, p = .120$). However, if we look at its descriptive statistics (see Figure 7), we observe that the *no-mouth* condition presented the lowest median ($Mdn = 5, 50$), and the *human-like* group the highest one ($Mdn = 7, 50$). The scores in the conditions *smile* and *wave-like* were rather higher too ($Mdn_{smile} = Mdn_{wave-like} = 7, 00$).

B. Data analysis for Part II

As already detailed in Section V-C, in the second part of the experiment, participants watched a video where the 4 conditions were shown simultaneously. They had to answer 6 questions by choosing one of the 4 different configurations of the robot's mouth.

Considering that, depending on which condition they were assigned to in Part I, they had watched one of the robot's configurations before, first we tested if their selections were related with the condition they were assigned to. We employed

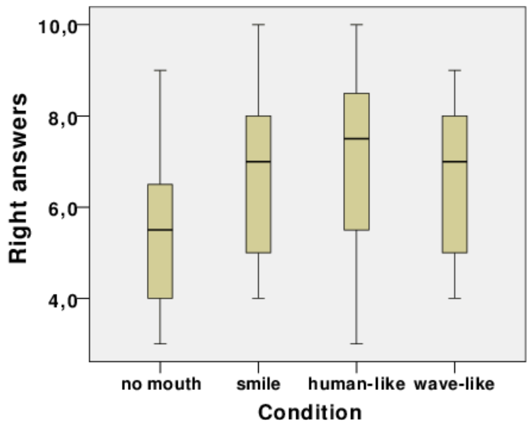


Figure 7. Boxplot of right answers depending on the condition. Each box is delimited by the 1st and 3rd quartiles and the median is represented by the inside horizontal line. The upper and lower whiskers mark the maximum and minimum scores, respectively.

Table III
PEARSON'S CHI-SQUARE TEST RESULTS. STATISTICALLY SIGNIFICANT ASSOCIATIONS BETWEEN THE PARTICIPANTS' RESPONSES AND THEIR CONDITION HAVE BEEN HIGHLIGHTED (p-VALUE $< .05$).

Question	X2	df	Asymp. Sig. (2-side)
Which is the most lively robot?	16.817	9	.052
Which is the most inert robot?	5.069	9	.828
Which robot is the most dangerous?	17.950	9	**.036**
Which is the friendliest robot	29.244	9	**.001**
Which robot did you like the most?	16.281	9	.061
Which robot would you like to meet?	16.968	9	**.049**

Pearson's chi-square test for each question to test if the selected robot's mouth was due to the condition they were assigned to. In other words, we checked whether the first impression of the robot in Part I affected the participants' answers in the questionnaire of Part II. Considering the Pearson's chi-square test, the null hypothesis for each question was

> *There are no differences in the selected robot's mouth between conditions.*

Accordingly, the alternative hypothesis would be

> *The selected robot's mouth will be different depending on the condition the participant was assigned to.*

The questions and the results of the 6 Pearson's Chi-square tests are presented in Table III.

For the items *Which is the most lively robot?* ($\chi = 16.817$, $df = 9$; $p > .05$), and *What robot did you like the most?* ($\chi = 16.281$, $df = 9$; $p > .05$), the null hypothesis should be accepted: there is not enough evidence to suggest that any differences between conditions is for any reason other than chance (e.g. sampling error).

In the case of the item *Which is the most inert robot?*, Pearson's chi-square test suggested that the null hypothesis should be accepted ($\chi = 5.069$, $df = 9$; $p > .05$) too.

However, in the case of the item *Which robot is the most dangerous?*, the alternative hypothesis should be accepted ($\chi = 17.950$, $df = 9$, $p = .036$; $p < .05$). This means that the users' responses were affected by the condition they were assigned to. Similarly, the item *Which is the friendliest robot* presented a probability of the chi-square test statistic ($\chi = 29.244$, $df = 9$) $p = .001$, less than the alpha level of significance of 0.05.

When asking *Which robot would you like to meet?*, we found a marginal effect on the users' responses based on their condition ($\chi = 16.968$, $df = 9$; $p = .049$).

In order to ease the understanding of the results, Figure 8 shows the responses of the participants depending on their condition.

VII. DISCUSSION

A. Part I: Participants' perceptions of the robot

The Kruskal-Wallis tests conducted on the items *inert* and *nervous* found significant differences but the post-hoc analysis could not determine where those differences were. Nevertheless, it is worth mentioning the ratings for these items shown in Table II and Figure 6. In the case of the *mouthless* configuration, it obtained the highest ratings in the item *inert*, just followed by the *wave-like* mouth robot. These two configurations, *no-mouth* and *wave-like*, presented the biggest differences compared to the mouths existing in nature: in the one case, the robot did not show a mouth at all, while the other configuration showed a moving sinusoidal wave. Therefore, considering that most animals in nature emit sounds by opening and closing a mouth, it might be very difficult for a person to link these configurations with a living being causing higher ratings of the adjective *inert*.

On the contrary, the *human-like* and *smile* configurations presented the lowest ratings of *inert*. Both configurations might resemble to human mouths and this can get lower ratings.

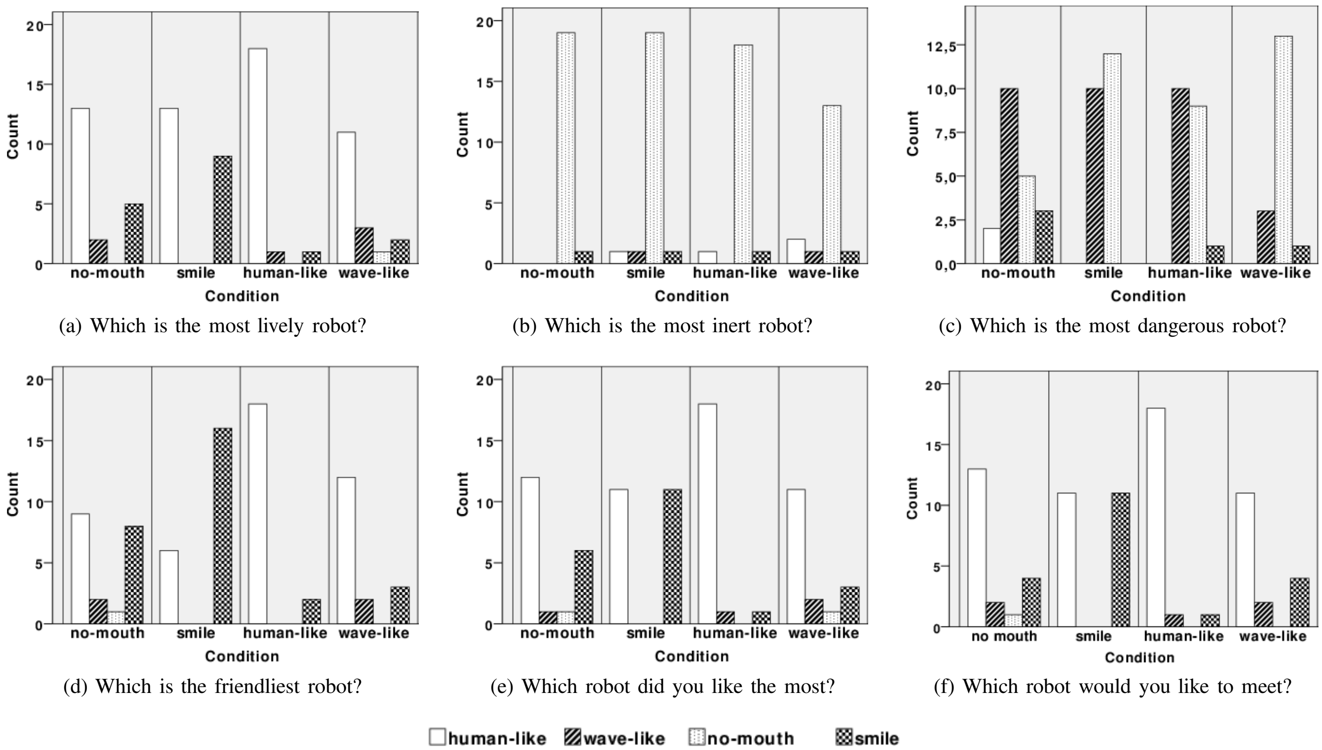
These results could lead to the idea that the presence of a mouth that is similar to those existing in animals is important for improving the perceived liveliness of a robot. However, more experiments are needed to confirm this idea.

In the case of the adjective *nervous*, the two dynamic mouth configurations scored the highest: the *wave-like* mouth and the *human-like* mouth. In contrast, the smiley mouth, which was static, obtained the lowest ratings. We believe that a moving mouth helped produce the perception of our MBot as a jittery robot and therefore they obtained higher ratings of nervousness.

In addition, the animation industry has used a wave-shaped mouth to show characters' anxiety or fear. The authors believe that this definitely contributed to increase the ratings of the adjective *nervous* over the other configurations. Participants who took part in the *wave-like* condition could link the robot's mouth with the memories they had about cartoons using this technique.

Again, according to the statistical results, we cannot point out where the differences of nervousness are and additional research is required to corroborate these ideas.

The adjective *sad* was rated significantly higher in the *no-mouth* condition than in the *smile* one. This could be attributed to the fact that a happy face is clearly identified with a U-shaped mouth. In contrast, when a face does not have a mouth,



(a) Which is the most lively robot?
(b) Which is the most inert robot?
(c) Which is the most dangerous robot?
(d) Which is the friendliest robot?
(e) Which robot did you like the most?
(f) Which robot would you like to meet?

Figure 8. Frequency of counts of the participants' responses to the comparative questions. The color of the bar corresponds to the different responses: option 'a', human-like mouth, uses the blue color; option 'b', wave-like mouth, is shown in green; the yellow bar is reserved for the option 'c', no-mouth; and the smile mouth, option 'd', is represented in purple. For each question, the responses are divided by the condition they were assigned to.

participants have perceived this as a sad trait and their ratings were higher.

Moreover, when the robot shows a neutral mouth, either the *human-like* or the *wave-like* mouths, the *sad* ratings were higher than the smiling robot but lower than the *mouthless* one. In these cases, no significant differences were found, but it might indicate the importance of showing an emotional mouth, i.e. a mouth that clearly recalls an emotion such as sadness or happiness, when we wish to endow our robots with the ability to express emotions.

In relation to happiness, in the *no-mouth* condition, the item *happy* was rated significantly lower than in the *smile* one (the highest one). Although no significant differences were found for the other two conditions, their mean ranks are clearly higher than the *no-mouth* condition. These results are inline with the results for the item sad.

In view of these results, we can strengthen the importance of a mouth because its absence can lead to false impressions; in our experiment, we did not intend to attribute emotional states to the robot but the lack of a mouth made Mbot perceived as less *happy* and more *sad*. Considering the above mentioned results, we can support hypothesis H_2: different impressions have been observed between conditions, the only difference being the type of mouth.

B. Part I: Attentiveness to the robot

According to the results above, in the questions related to the content of the robot's speech, we did not find significant differences in the number of right answers between the different conditions. Therefore, we cannot confirm hypothesis H_3. However, looking at Figure 7, we have observed that the *no-mouth* condition presents the lowest number of right answers, one and a half points away from the next condition. Again, this important difference brings up the effect of having a mouth on focusing one's attention. In this case, the varied conditions presented mouths, either dynamic or static, that synchronized with the robot's speech. Despite the apparent simplicity of the mouths, we have observed more correct answers than when it is a *mouthless* robot.

It is also important to mention that the highest number of right answers was obtained under the *human-like* condition. This could indicate that the more 'human' the mouth is, the more focused a user is on the robot, and, therefore, the more correct answers. This would be a very significant result that can be relevant to many applications of social robots. For example, social robots are extensively being applied to educational tasks [32, 33]. In these scenarios, the attention of the users is crucial, and a human-like mouth would help to improve the learning rates. To confirm this hypothesis, more experiments are required.

C. Part II: Comparing the robots' mouths

In the Comparative Questionnaire (Part II), the participants contrasted the 4 conditions. Figure 8 shows important differences in the preferences of the participants, depending on the robot's mouth. This confirms hypothesis H_2, which assumed that the type of mouth changes the participants' perception of the robot.

Considering the statistical analysis presented in Section VI-B, the differences found for the questions 'Which is the most lively robot?', 'Which is the most inert robot?', and 'Which robot did you like the most?' were independent of the condition under which the participants took part.

In the case of the item where the participants selected the most lively robot, clearly the robot with a *human-like* mouth is the favourite one under all conditions, followed at a considerable distance by the robot with a smile. Only under the *wave-like* condition did more participants prefer this configuration over the smile, but far from the *human-like* mouth. In this case, it seems that when users initially interact with a robot endowed with a dynamic mouth, they perceive it as being more lively than a robot with a static mouth, even if the mouth is as affective and positive as a smile. The *no-mouth* robot was rarely selected: participants seldom perceived a *mouthless* robot as lively.

In this same line, when participants picked the most inert robot, they undoubtedly chose the *mouthless* MBot. Other configurations of a robot's mouth were rarely selected.

These results confirm hypothesis H_1; this points to the importance of the presence of a mouth in robots that want to be perceived as a living entity rather than a mere machine. In contrast, when there is no mouth, the robot is seen as a lifeless object.

The question about the robot participants like the most reveals that the *human-like* configuration is clearly the favourite one. Even participants from the *smile* condition approximately equally rated the *smile* and the *human-like* configurations. In view of these results, we can say that the type of mouth affects the robot's likability, confirming H_2.

The results obtained from the items 'Which robot is the most dangerous?', 'Which is the friendliest robot?', and 'Which robot would you like to meet?' were different, depending on the condition to which the participant was assigned.

The question related to the most dangerous robot shows two mouths that stand out: *no-mouth* and *wave-like* (Figure 8c). Participants who observed the *mouthless* robot speaking in the video of Part I selected the *wave-like* mouth as the most dangerous. In contrast, those who watched MBot with a *wave-like* mouth while speaking, rated the *no-mouth* configuration as the most dangerous. In the case of participants under the conditions *smile* and *human-like*, they selected, approximately equally, the *wave-like* and *no-mouth* robots as the most dangerous. We can derive two conclusions here: (i) the robot's first impression matters - if people have already observed a robot with some traits that might be perceived as dangerous, they will not consider it as dangerous as others, probably because they have gotten acquainted with it; (ii) the lack of a mouth and a waveform mouth contribute to having the robot perceived as more dangerous. Consequently, we can say that the type of mouth can influence the perceived safety of a robot, which enforces H_2 again.

The ratings of the friendliest robot present two configurations over the rest: the *human-like* and the *smile* mouths (Figure 8d). Both are easily associated with human mouths, and this could favour their friendliness. In particular, participants under the *no-mouth* condition rated both configurations roughly equally friendly. Considering that, under this condition, participants selected the *wave-like* mouth as the most dangerous one, it makes sense that the *human-like* and the *smile* mouths were perceived as friendlier. When seeing the *wave-like* condition, participants rated the *human-like* mouth as the friendliest one. This could be due to the fact that they have already observed a dynamic mouth in sync with the robot's utterances and expected a dynamic mouth; having a static mouth could be perceive as weird, and thus damage the robot's friendliness. In the case of the *smile* and the *human-like* conditions, participants rated more highly those robots that had initially told their story. Again, this proves the importance of the first impression, in robots too.

Finally, the last comparative question reveals that most of the participants would like to meet the robot with a *human-like* mouth (Figure 8f). Only participants from the *smile* condition presented different results: MBots with the *human-like* and the *smile* mouths were chosen evenly. Once more, the fact that these participants had already watched the smiling robot resulted in more participants wanting to meet that robot.

These last two questions, about the robot's friendliness and being willing to meet it, are related to its likeability. Again, in this category, the participants had clear preferences, which sustains H_2.

D. Limitations

The results presented in this paper should be considered with caution due to the following limitations. First, the conducted experiment was performed by volunteers watching videos of a robot telling its personal story. These results need to be confirmed for real, physical, human–robot interactions. However, in order to recruit a large number of heterogeneous participants to evaluate the robots' appearance, online videos represented an interesting setup.

In addition, in this experiment we have employed the MBot robot with different mouths. MBot presents a head with a mouth, a torso, and two arms, which can resemble a human or animal body. In case we consider robots with other external appearances, the results could differ. The study needs to be extended to more heterogeneous robots before we generalize the conclusions.

In the same line, during the video, the robot was synthesizing utterances. Many roboticists have started to endow robots with non-verbal sounds. What kind of mouth should be applied with this type of sounds, as well as their pros and cons, is not clear yet and needs further study.

The experiment was conducted in Spanish. Some parts of the questionnaires were carefully translated from their original languages (for example, items from the Godspeed questionnaire). Likewise, some of the content of the present paper has been accurately transcribed from Spanish into English (e.g. the comparative questions). Despite the effort, some inaccuracies could be present.

VIII. Conclusions

In this paper we have observed that the presence of a mouth, in sync with the robot's utterances, improves the robot's impressions on users watching the robot. In particular, we have evaluated 4 conditions with the MBot robot: two synchronized dynamic mouths (human-like and wave-shape), a static one (a smile), and no mouth at all.

In terms of happiness, the presence of a mouth makes a difference: the *no-mouth* condition was rated as the saddest, and its scores for the *happy* item were significantly lower than those of the *smile* condition.

When comparing the 4 conditions, independently of the participant's background, the MBot showing the human-like mouth was selected as the most lively and the one they liked the most. Conversely, the "mouthless" robot was perceived as the most inert robot. In the case of the most dangerous robots, participants perceived the robot with the wave-shaped mouth and the one without a mouth as the most dangerous. In line with this, the friendliest robots were those with a smile or with a human-like mouth, and consequently these were the robots participants wanted to meet.

Considering the previous results, we can confirm that including a mouth in social robots increases its perception as life-like (H_1). Besides, we have observed that the shape and movement of the mouth affects the perceptions of the robot (H_2). As a consequence, roboticists should put more effort in the robots' mouths. Finally, we cannot confirm that participants will pay more attention to a robot speaking when it has a human-shaped, synchronized mouth (H_3). However, the results obtained point in this direction.

After these findings, we can conclude that our dynamic mouths can benefit human–robot interaction. They help to communicate the robot's state as well as intentions. Then, depending on the context, the robot could show the most appropriate type of mouth. For example, if the robot is performing a task that may be dangerous for the people around it, it can display a wave-shape mouth to prevent users from coming close. To give another example, in case the robot needs to engage people, it could display a human-like mouth, which was found the most appealing one.

Acknowledgment

The research leading to these results has received funding from the projects: Development of social robots to help seniors with cognitive impairment (ROBSEN), funded by the Ministerio de Economia y Competitividad (DPI2014-57684-R); MOnarCH, funded by the European Commission (grant agreement No 601033); and RoboCity2030-III-CM, funded by Comunidad de Madrid and cofunded by Structural Funds of the EU (S2013/MIT-2748).

References

[1] A. Castro-González, H. Admoni, and B. Scassellati, "Effects of form and motion on judgments of social robots' animacy, likability, trustworthiness and unpleasantness," *International Journal of Human-Computer Studies*, vol. 90, pp. 27–38, Jun. 2016. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S107158191600032X

[2] S. S. Kwak, "The Impact of the Robot Appearance Types on Social Interaction with a Robot and Service Evaluation of a Robot," *Archives of Design Research*, May 2014. [Online]. Available: http://www.aodr.org/_common/do.php?a=full&b=12&bidx=155&aidx=1817

[3] R. H. Kim, Y. Moon, J. J. Choi, and S. S. Kwak, "The effect of robot appearance types on motivating donation." ACM Press, 2014, pp. 210–211. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2559636.2563685

[4] H. Mcgurk and J. Macdonald, "Hearing lips and seeing voices," *Nature*, vol. 264, no. 5588, pp. 746–748, Dec. 1976. [Online]. Available: http://www.nature.com/doifinder/10.1038/264746a0

[5] Jana M. Iverson and Esther Thelen, "Hand, Mouth and Brain. The Dynamic Emergence of Speech and Gesture." *Journal of Consciousness Studies*, vol. 6, no. 11-12, pp. 19–40, Nov. 1999.

[6] David F. McAllister, Robert D. Rodman, Donald L. Bitzer, and Andrew S. Freeman, "Lip Synchronization for Animation," in *ACM SIGGRAPH 97 Visual Proceedings: The Art and Interdisciplinary Programs of SIGGRAPH '97*, ser. SIGGRAPH '97. Los Angeles, California, USA: ACM, 1997, oCLC: 833286155. [Online]. Available: http://doi.acm.org/10.1145/259081.259312

[7] Keith Waters and Thomas M. Levergood, "DECface: An Automatic Lip-Synchronization Algorithm for Synthetic Faces," Digital Equipment Corporation, Cambridge Research Laboratory, Technical report CRL 93/4, 1993. [Online]. Available: http://www.hpl.hp.com/techreports/Compaq-DEC/CRL-93-4.pdf

[8] J. P. Lewis and F. I. Parke, "Automated lip-synch and speech synthesis for character animation," *ACM SIGCHI Bulletin*, vol. 17, no. SI, pp. 143–147, May 1986. [Online]. Available: http://portal.acm.org/citation.cfm?doid=30851.30874

[9] T. Frank, M. Hoch, and G. Trogemann, "Automated Lip-Sync for 3d-Character Animation," Berlin, Aug. 1997.

[10] Tsuhan Chen and R. Rao, "Audio-visual integration in multimodal communication," *Proceedings of the IEEE*, vol. 86, no. 5, pp. 837–852, May 1998. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=664274

[11] I.-W. Park, J.-Y. Kim, B.-K. Cho, and J.-H. Oh, "Control hardware integration of a biped humanoid robot with an android head," *Robotics and Autonomous Systems*, vol. 56, no. 1, pp. 95–103, Jan. 2008. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0921889007001376

[12] S. Nak, F. Kanehiro, K. Miura, M. Morisawa, K. Fujiwara, K. Kaneko, S. Kajita, and H. Hirukawa, "Creating facial motions of Cybernetic Human HRP-4c." IEEE, Dec. 2009, pp. 561–567. [Online]. Available: http://ieeexplore.ieee.org/document/5379516/

[13] K.-G. Oh, C.-Y. Jung, Y.-G. Lee, and S.-J. Kim, "Real-time lip synchronization between text-to-speech

(TTS) system and robot mouth." IEEE, Sep. 2010, pp. 620–625. [Online]. Available: http://ieeexplore.ieee.org/document/5598656/

[14] F. A. Cid, L. J. Manso, L. V. Calderita, A. Sánchez, and P. Nuñez, "Engaging human-to-robot attention using conversational gestures and lip-synchronization," *Journal of Physical Agents*, vol. 6, no. 1, pp. 3–10, 2012. [Online]. Available: http://193.146.101.101/index.php/jopha/article/viewArticle/101

[15] H.-J. Hyung, B.-K. Ahn, B. Cruz, and D.-W. Lee, "Analysis of Android Robot Lip-Sync Factors Affecting Communication," in *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. IEEE Press, 2016, pp. 441–442. [Online]. Available: http://dl.acm.org/citation.cfm?id=2906917

[16] H. Kozima, M. P. Michalowski, and C. Nakagawa, "Keepon: A Playful Robot for Research, Therapy, and Entertainment," *International Journal of Social Robotics*, vol. 1, no. 1, pp. 3–18, Jan. 2009. [Online]. Available: http://link.springer.com/10.1007/s12369-008-0009-8

[17] Doug Gross, "Meet Pepper, the emotional robot," Jun. 2014. [Online]. Available: http://edition.cnn.com/2014/06/06/tech/innovation/pepper-robot-emotions/

[18] Time video, "Meet Pepper, the Robot Who Can Read Your Emotions," *Time*, Jun. 2014. [Online]. Available: http://time.com/2845040/robot-emotions-pepper-softbank/

[19] Erico Guizzo, "How Aldebaran Robotics Built Its Friendly Humanoid Robot, Pepper," Dec. 2014. [Online]. Available: https://goo.gl/h0MXlQ

[20] J. Kedzierski, R. Muszynski, C. Zoll, A. Oleksy, and M. Frontkiewicz, "EMYS-Emotive Head of a Social Robot," *International Journal of Social Robotics*, vol. 5, no. 2, pp. 237–249, Apr. 2013. [Online]. Available: http://link.springer.com/10.1007/s12369-013-0183-1

[21] J. K. Lee and C. Breazeal, "Human social response toward humanoid robot's head and facial features." ACM Press, 2010, p. 4237. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1753846.1754132

[22] A. van Breemen, "Animation engine for believable interactive user-interface robots," vol. 3. IEEE, 2004, pp. 2873–2878. [Online]. Available: http://ieeexplore.ieee.org/document/1389845/

[23] Minato Takashi, Michihiro Shimada, Hiroshi Ishiguro, and Shoji Itakura, "Development of an android robot for studying human-robot interaction," ser. IEA/AIE'2004, vol. 3029. Ottawa, Canada: Springer Springer Verlag Inc, 2004, pp. 424–434. [Online]. Available: http://link.springer.com/10.1007/b97304

[24] S. Al Moubayed, J. Beskow, G. Skantze, and B. Granström, *Furhat: A Back-Projected Human-Like Robot Head for Multiparty Human-Machine Interaction*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 114–130. [Online]. Available: https://doi.org/10.1007/978-3-642-34584-5_9

[25] R. Perula-Martinez, E. Salichs, I. P. Encinar, A. Castro-Gonzalez, and M. A. Salichs, "Improving the Expressiveness of a Social Robot through Luminous Devices." ACM Press, 2015, pp. 5–6. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2701973.2702021

[26] M. K. Lee, J. Forlizzi, P. E. Rybski, F. Crabbe, W. Chung, J. Finkle, E. Glaser, and S. Kiesler, "The snackbot: documenting the design of a robot for long-term human-robot interaction," in *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction*. La Jolla, California, USA: ACM Press, Mar. 2009, p. 7. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1514095.1514100

[27] J. Osada, S. Ohnaka, and M. Sato, "The scenario and design process of childcare robot, PaPeRo." ACM Press, 2006, p. 80. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1178823.1178917

[28] "Buddy robot snubs creator," Aug. 2015. [Online]. Available: http://www.bbc.com/news/technology-33739290

[29] J. Kory Westlund, "Telling Stories with Green the DragonBot: A Showcase of Children's Interactions Over Two Months." ACM Press, 2015, pp. 263–263. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2701973.2702089

[30] C. Bartneck, D. Kulic, E. Croft, and S. Zoghbi, "Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots," *International Journal of Social Robotics*, vol. 1, no. 1, pp. 71–81, Jan. 2009. [Online]. Available: http://link.springer.com/10.1007/s12369-008-0001-3

[31] D. C. Hoaglin, B. Iglewicz, and J. W. Tukey, "Performance of Some Resistant Rules for Outlier Labeling," *Journal of the American Statistical Association*, vol. 81, no. 396, p. 991, Dec. 1986. [Online]. Available: http://www.jstor.org/stable/2289073?origin=crossref

[32] E. Short, K. Swift-Spong, J. Greczek, A. Ramachandran, A. Litoiu, E. C. Grigore, D. Feil-Seifer, S. Shuster, J. J. Lee, S. Huang, Svetlana Levonisova, Sarah Litz, Jamy Li, Gisele Ragusa, Donna Spruijt-Metz, Maja Mataric, and Brian Scassellati, "How to train your dragonbot: Socially assistive robots for teaching children about nutrition through play," in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 2014, pp. 924–929. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6926371

[33] D. Leyzberg, S. Spaulding, and B. Scassellati, "Personalizing robot tutors to individuals' learning differences." ACM Press, 2014, pp. 423–430. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2559636.2559671

Álvaro Castro-González received the B.Sc. degree in Computer Engineering from the University of León, León, Spain, in 2005, and the M.Sc. and Ph.D. degrees in Robotics and Automation from the Carlos III University of Madrid, Madrid, Spain, in 2008 and 2012, respectively. He is a member of the Robotics Lab Research Group and Assistant Professor at the Department of Systems Engineering and Automation of the Carlos III University of Madrid. His current research focuses on personal and social robots, human–robot interaction, decision making systems, autonomy, and artificial emotions and motivations.

Jonatan Alcocer-Luna obtained his Bachelor's Degree in Computer Science from the Rey Juan Carlos University of Madrid, in 2013. He obtained a Master's Degree in Robotics and Automation from the University Carlos III of Madrid, in 2015. From December 2014, he has been working as software systems administrator and researcher at the RoboticsLab. His line of research focuses on the automation of human–robot interaction. His research interests include artificial intelligence, bioengineering and nanorobotics.

María Malfaz is an assistant professor of the Systems Engineering and Automation Department at the Carlos III University of Madrid. María Malfaz received her degree in Physics at La Laguna University in 1999. In October 2001, she received a MSc in Control Systems from Imperial College of London. She received the Ph.D. degree in Industrial Engineering in 2007; the thesis was 'Decision Making System for Autonomous Social Agents Based on Emotions and Self-learning'. Her research area follows the line carried out in her thesis and, more recently, she has been working on multimodal human–robot interaction systems. She belongs to several international scientific associations: the IEEE-RAS (IEEE Robotics and Automation Society), the IFAC (International Association of Automatic Control), and the CEA (Comité Español de Automática). Moreover, she is also a member of research networks such as EURobotics (European Robotics Coordination Action), and HispaRob (Plataforma Tecnológica Española de Robótica).

Fernando Alonso-Martín works as a researcher and teacher at the Systems Engineering and Automation Department, Carlos III University of Madrid. His research fields are personal robots, human–robot interaction, dialogues, and other related issues. Currently he is mainly involved in two projects: RobAlz is a project born from the collaboration between the Spanish Alzheimer Foundation and the RoboticsLab. The aim is to develop robots that assist in the daily tasks of caregivers to Alzheimer's sufferers. In the case of the MOnarCH (Multi-Robot Cognitive Systems Operating in Hospitals) project, it is a European Union FP7 project that aims at the development of a network of heterogeneous robots and sensors in the pediatric area of an oncological hospital.

Miguel A. Salichs received the electrical engineering and Ph.D. degrees from Polytechnic University of Madrid, Madrid, Spain. He is currently a Full Professor of the Systems Engineering and Automation Department at the Carlos III University, Madrid, Spain. His research interests include autonomous social robots, human–robot interaction, bio-inspired robotics, and cognitive architectures. Prof. Salichs was member of the Policy Committee of the International Federation of Automatic Control (IFAC), chairman of the Technical Committee on Intelligent Autonomous Vehicles of the IFAC, head of the Spanish National Research Program on Industrial Design and Production, the President of the Spanish Society on Automation and Control (CEA), and the Spanish representative to the European Robotics Research Network (EURON).