

## Article

# Creation and Staging of Android Theatre “Sayonara” towards Developing Highly Human-Like Robots

Takenobu Chikaraishi <sup>1,2,\*</sup> , Yuichiro Yoshikawa <sup>2</sup>, Kohei Ogawa <sup>2</sup>,  
Oriza Hirata <sup>1,2</sup> and Hiroshi Ishiguro <sup>2</sup>

<sup>1</sup> COI Site, Tokyo University of the Arts, Tokyo 110-8714, Japan; hhd00607@nifty.ne.jp

<sup>2</sup> Graduate School of Engineering Science, Osaka University, Suita 565-0871, Japan;  
yoshikawa@irl.sys.es.osaka-u.ac.jp (Y.Y.); ogawa@irl.sys.es.osaka-u.ac.jp (K.O.);  
ishiguro@irl.sys.es.osaka-u.ac.jp (H.I.)

\* Correspondence: chikaraishi.takenobu@pc.geidai.ac.jp

Received: 4 September 2017; Accepted: 18 October 2017; Published: 2 November 2017

**Abstract:** Even after long-term exposures, androids with a strikingly human-like appearance evoke unnatural feelings. The behavior that would induce human-like feelings after long exposures is difficult to determine, and it often depends on the cultural background of the observers. Therefore, in this study, we generate an acting performance system for the android, in which an android and a human interact in a stage play in the real world. We adopt the theatrical theory called Contemporary Colloquial Theatre Theory to give the android natural behaviors so that audiences can comfortably observe it even after long-minute exposure. A stage play is created and shown in various locations, and the audiences are requested to report their impressions of the stage and their cultural and psychological backgrounds in a self-evaluating questionnaire. Overall analysis indicates that the audience had positive feelings, in terms of attractiveness, towards the android on the stage even after 20 min of exposure. The singularly high acceptance of the android by Japanese audiences seems to be correlated with a high animism tendency, rather than to empathy. We also discuss how the stage play approach is limited and could be extended to contribute to realization of human–robot interaction in the real world.

**Keywords:** android theatre; robot theatre; contemporary colloquial theatre theory; social robots

---

## 1. Introduction

Humanoid robots with human body structures have been developed in recent years, and they are expected to emotionally engage with humans in the near future [1,2]. Previous studies of humanoid robots reveal the importance of nonverbal expressions during human–robot communications. In particular, android robots with the appearance of actual persons are expected to become familiar entities in future human society [3,4].

Humanoid robots can induce familiar feelings in humans because of their human-like appearance. However, humanoid appearances may induce a well-known effect called the “uncanny valley” [5]. The unnatural sensation of interacting with humanoid robots is caused by tiny differences between the robot and humans. In fact, 76% of subjects cannot distinguish an android robot from a human after watching the android for less than two seconds [6]. Therefore, the negative feeling is induced after long-term exposure to an android robot and remains a central barrier to comfortable human–android interactions. According to several studies, harmony between appearance and behavior of robots reduces the negative feeling in observers [7]. This consideration has led to some successfully implemented android behaviors that do not evoke the negative feeling [8,9]. However, these previous studies did not reveal the details of implementation or sufficiently evaluate its dependence

on the observers' backgrounds (such as nationality, age, and education). In other words, a methodology for building an android robot that is perceived as a human-like entity after long-term exposure has yet to be established. To tackle this issue, we have started an activity called "Robot theatre project," in which robots and humans act in a long-lasting stage production. The project integrates robot technology and knowledge with the art and entertainment arena [10].

The history of stage art is replete with implicit knowledge for directing actors on the stage. In other words, stage directors have useful knowledge about how to lead actors to portray realistic characters or emotions (pleasure, grief, and so on). Interestingly, these emotions can be performed independently of whether or not the actors actually feel them. Therefore, the practical knowledge of stage directors is expected to enable us to compose a stage play with a humanoid robot that is teleoperated and humanized. This will provide new insight for engineers designing robot behavior. Moreover, presenting such a robot at public performances in theatres enables a large-scale evaluation of audience impressions towards acting robots, which can then be fed back into the design process."

Stage plays have been performed all over the world from ancient times. A large number of art works have been produced based on the professional techniques of stage directors and actors. A director focuses on representing human behavior for improving his stage plays. Therefore, stage directors may have important knowledge for developing robots that represent human behavior.

In the 20th century, Stanislavski established a new method for training and rehearsing [11]. In this method, actors are supposed to build natural and real acting by focusing on the psychological aspects of the character to be played. This method has become mainstream in modern stage performances. Katie Mitchell kindly described a practical manner for the technique [12].

In contrast to Stanislavski's method, Gordon Craig suggested that artistic quality should be attributed only to the director, rather than the actors. According to his method, actors should be a type of puppet made of flesh and blood, faithfully representing the director's direction while suppressing their own psychologies and intentions [13]. This type of actor is known as an "über-marionette (super-marionette).

It is not yet easy for us to make a play with a humanoid robot based on Stanislavski's approach because intelligent robots capable of simulating the human mind (or conscious or unconscious behaviors stemming from it) have not yet been developed. In contrast, we could attempt to create scenes with our humanoid robot based on Craig's approach because the technology for teleoperating humanoid robots has become so sophisticated that we can treat it as an über-marionette of the director.

One of the authors of the present paper (Oriza Hirata) is widely esteemed as a playwright and stage director, and has advocated what he calls the "Contemporary Colloquial Theatre Theory (CCTT)" [14]. As CCTT replicates on stage the reality of everyday human activities, it is potentially applicable to designing robots with human behaviors. CCTT advocates for precise, rather than ambiguous, instructions for actors.

Actors are instructed when to alter their physical actions, such as utterances and body orientations. In other words, CCTT follows Craig's approach, which uses actors as über-marionettes. Such precise instructions are also expected to be directly applicable to android robots.

Therefore, by creating a CCTT-based stage play, we can hope to realize human-like robot behavior. Robots have occasionally been used in stage productions. For example, Duncan et al. played Shakespeare with micro-helicopter robots [15]. Demers made a stage performance of Kafka with a combined metal-bar shaped robot [16]. Lemaignan et al. made a pantomime performance with a two-arm robot [17].

Zeglin et al. created a human–robot drama using a butler robot with two hands and wheels [18].

Lu et al. ontologically classified the robot theatre on an axe of "automation level" and an axe of "control of the system" [19].

Jochum et al. successfully used live performance to explore creative solutions for design and control mechanisms enabling expressive robot motion [20] and showed that this framework is innovative in developing human care robots [21].

To better understand the required behavioral nuances of fully humanoid robots, it is imperative to create a stage play using a robot that has human-like body structures such as android robots.

To this end, we have produced a stage play in our prior work, involving two humanoid robots, and we applied CCTT to realize co-acting of an android and a human actor. This trial revealed that CCTT is effective for engaging robots in a stage play and eliciting a positive response from audiences [10]. However, two issues remain untouched: first, whether human-mimicking robots that evoke the negative feeling are effectively directed by CCTT is debatable, second, as the trial was held only in Japan, we could not determine whether audience impressions would differ between nationalities. Therefore, in this study, we create a stage play using the Geminoid Fandroid, whose appearance exactly resembles that of an adult female, and evaluate it in different countries. Throughout the play, named “Sayonara”, the android’s actions are generated by tele-operation technology to convey the professional actor’s performance and emotional expressions directed by Hirata. In this paper, we describe not only the mechanisms of the proposed system to control the android, but also the process for making stage performance by using it under CCTT. It is considered to be worth precisely reporting them because such descriptions involve the important how-to information of this pioneering challenge of human–robot theatre.

“Sayonara” is a one-on-one conversation drama between an android whose appearance is elaborately modeled on an existing human model and a human actor. This is very advantageous to this type of cultural and social study because one-to-many conversations never occur in “Sayonara”. Audiences, therefore, can evaluate interactions with a robot, excluding the influence of the designed appearance of the robot.

We have presented two other dramatic works with androids: “Three Sisters, Android Version (2013)” and “La Métamorphose Version Androïde (2014)”. In “Three Sisters, Android Version”, an android with a human-like appearance, a robot with a mechanical design, and nine human actors appear on stage and interact with one another. In “La Métamorphose version androïde”, an android with a half-human-like and half-mechanical appearance performs with four human actors. Therefore, we chose “Sayonara” for this study.

The android interacts with a human actor on the stage for approximately 20 min. Through public performances, we demonstrate the feasibility of the proposed method for designing robots with long-term human-like behavior. To cover a wide cultural base, public performances were held in Japan, Austria, Germany and France. In questionnaires, audiences were invited to share their impressions of the robot and the stage play, and divulge their cultural and psychological backgrounds. The paper concludes with the implications of the current work and ideas for future work.

With growth of the internet, information or culture spreads immediately out into the world. In such an age, the study of the cultural and social reception of robots in everyday life among various countries is quite significant for designing future human–robot interaction. Therefore, investigating the reception of androids by different cultures via android–human theatre can contribute to the creation of robots that can harmoniously coexist with humans.

## 2. Creating a Stage Play with An Android

In this study, one of the two actors was “Geminoid F,” an android with an outwardly human appearance. This section, describes the system for generating Geminoid F’s acting performance via a tele-operating system and the theatrical instruction from Hirata.

### 2.1. Tele-Operated Android, Geminoid F

Figure 1 illustrates a scene from the stage play, acted by Geminoid F (left) and a human actress (right), who plays an opposite role. The degrees of freedom (DOF) of Geminoid F are provided by a pneumatic actuator. The robot is covered with a silicone rubber skin used in prosthetics. This skin is modeled to high accuracy from an existing human skin.

The upper torso of Geminoid F has 12 DOFs controlling the most essential expressions in human communications (raise/lower brow, scowl, open/close eyelid, vertically move eyeball, horizontally move eyeball, open/close mouth, lift corner of mouth, incline head, shake head from side to side, chest movement, and bowing). Each DOF is driven by pneumatic actuators, electromagnetic valves, and control boards installed inside the body. As the air compressor can be installed outside the android body and connected by a long air tube, the android's motions are less noisy than those of robots installed with electromagnetic motors, rendering the android suitable for theatrical use. In addition, because the working fluid (i.e., air) is flexible, when the android comes within collision reach of the human (e.g., the stage), its movements can be controlled within a safety margin, which is also a suitable feature for theatrical use.

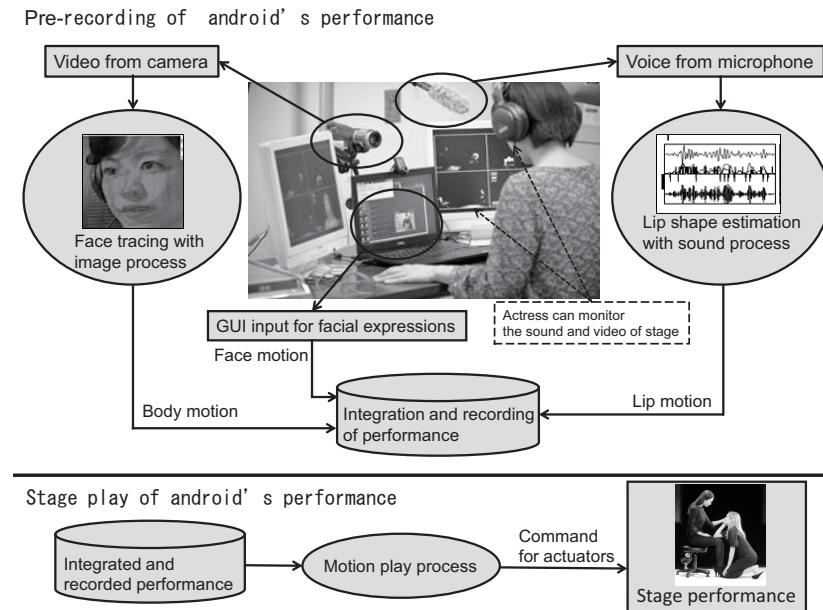


© Tatsuo Nambu

**Figure 1.** Android theatre “Sayonara”.

## 2.2. Acting Performance Generation System

The acting performance of the android is generated by a tele-operating system (Figure 2) [22]. Specifically, the android's part is first acted by a human actress. Her acting performance is captured by image processing and voice processing, and the performance is then transformed to android motion. Human actors interpret the director's intentions to properly express the emotions and bodily actions of their characters. As this skill is implicitly accumulated in the body language and experience of the actor, the corresponding actions of the android should be captured under the normal acting conditions of real actors.



**Figure 2.** System of android theatre.

Therefore, in this study, the head position of the actress is estimated by image processing with a face tracking program [23], and transformed into the head and coordinating back motions of the android. For opening and closing the mouth, face recognition by image processing is replaced by a voice processing technique developed by Ishii et al. [24]. As subtle unconscious behaviors such as blinking or breathing are important but subdued components of human-like presence, they are generated by a random process independent of the operator's acting performance, as implemented previously [8].

The expression of facial emotions (such as smiling or sadness) requires delicate timing. Therefore, we preprogrammed the actuators with a sequence of commands for proper facial expressions. By selecting an arbitrary timing, the actress controls the android's facial expression. In addition, during her acting performance (i.e., while she is tele-operating the android), the actress can observe the motions and hear the voice of the opposite actress and know the scene being enacted on stage through on-stage cameras and microphones. Thus, the acting performance of the android is recorded as a sequence of commands input to each actuator. This sequence drives each actuator at the same sampling rate, thereby realizing the same acting performance and proper interaction with the human actress on the stage.

### 2.3. Acting and Directing

The motions and utterances of the android during the stage performance are pre-recorded by the above described system. In other words, the acting performance is regenerated from pre-recorded motions. The director adjusts the acting performance via the following procedures:

- (1) **Read-through:** The script is read aloud. The main purpose of this process is to remember the words and timing.
- (2) **Initial direction:** The actress practices the script with a real android. She receives instructions on utterance timing, standing and sitting positions, and the body motions of the android and opposite actress from the director. When the stage director assesses that the performance has reached a threshold level of proficiency, the voices of the opposite actress and the actress playing the android are recorded.

- (3) **Voice recording:** During this process, the voice of the android is recorded in the recording studio. The actress playing the android's role utters the script while listening to the voice recorded in (2). In this way, the actors can reconstruct the acting performance with the correct timing and speed of their utterances.
- (4) **Recording motion:** During this process, the actress exactly times her bodily motions with the voices recorded in phase (3). Using image processing, the measured motions are transformed and recorded as sequences of command values for each actuator.
- (5) **Complete performance:** Now, the human actor synchronizes the timing of her acting performances with those of the android governed by the data recorded during phases (3) and (4).

The director repeats directions about the interpersonal timing of turn-taking as well as intrapersonal timing between motion and emotion, and the actor playing the android's role practices until the latency is acceptable. In the actual stage performance, an operator replays the motion data obtained in the manner described above.

As reported in a previous analysis of various CCTT-produced stage plays [25], the instructions for the present play pertain to the onset timing and speed of utterance, the direction of gaze, and the standing (or sitting) positions of the actors. Interestingly, the android and the actress with whom scenes are enacted received instructions for the same points.

### 3. The Android Drama “Sayonara”

Based on the system and procedure proposed in Section 2, we created the stage play “Sayonara”. The CCTT-based play was written and directed by Oriza Hirata, a world famous stage director and coauthor of this study. It was produced by Hirata’s theatre company “Seinendan”, which also employs a technician, lighting designer, acoustic designer, and scenographer. These specialists supplement the actors’ performance with their techniques to complete the stage play.

#### 3.1. Script

In “Sayonara”, a human actress and Geminoid F walk on to the stage. During the 20 min performance, the android reads the poetry to a dying girl.

The following five poems are cited:

- “Le bateau Ivre” by Arthur Rimbaud (France),
- “Tooku” by Shuntaro Tanikawa (Japan),
- “Sayonara” by Shuntaro Tanikawa (Japan),
- “über den Bergen” by Karl Busse (Germany), and
- “No title (in short tanka style)” by Bokusui Wakayama (Japan).

Through a calm dialogic play, Hirata encourages the audience to question what constitutes human life and death.

The stage play is dialogued in Japanese, English, German, and French. The participating actress can speak all of these languages, and the appropriate language was selected for each theatre. At international festivals, however, the spoken language was supplemented with English subtitles, or presented in English subtitled with the language spoken by the theatre audience.

Although the nominal performance time was 20 min, the actual performance time depended on the language in which the play was performed. In the English version, the android spoke 49 times (718 words), while the opposite actress spoke 48 times (363 words). As the android’s role was to read poems to her owner, the android uttered more words than the human actress.

#### 3.2. Stage Effects and Scenic Art

The stage set is simply constructed, comprising a rocking chair placed on a white rug (Figure 3). The actress on the rocking chair and the android faced each other two meters apart on the stage.

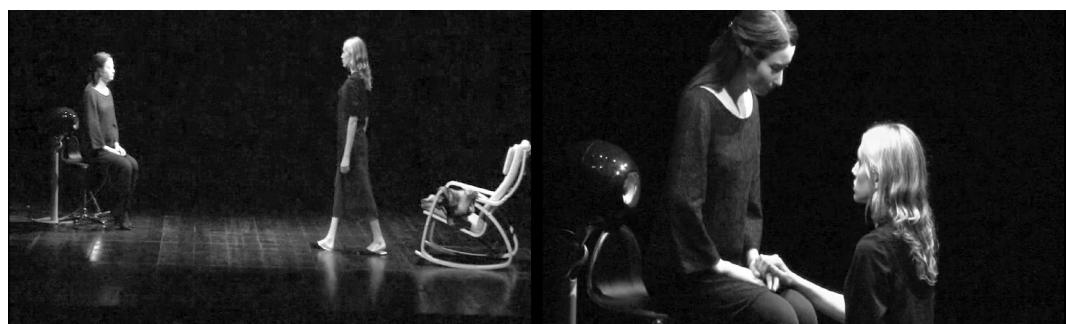


**Figure 3.** Stage set.

Although the details depended on the individual theatre, the lighting could be broadly divided into two categories: whole stage lighting, and spot lights illuminating each actor. As the audience entered the auditorium, the lighting was dimmed, and the actress was seen as sleeping in the rocking chair under the android's gaze. In this opening scene, the audience could immediately empathize with the characters on the stage. Once the performance had started, the stage brightened and remained bright throughout the performance. The performance proceeded in the absence of background music and sound effects. However, calm music was played while the audience was entering the auditorium. The android was heard through a speaker installed nearby. The speaker volume was adjusted to resemble the human voice that was copied in the android.

### 3.3. Acting Performance

Figure 4 depicts a scene from the stage performance. The story progresses by the conversation between the android and the actress. In this scene, the android remains seated and is approached by the actress, who takes the android's hand and converses with her. In this performance, the android expresses moderate emotion. As the performance contains no exaggerated actions, it can be assumed to reflect the distinctive feature of CCTT, namely, the natural expressions of people undertaking everyday activities.



**Figure 4.** One scene of “Sayonara”; The actress silently comes close to the Android (left), and the Android looks at the actress's face (right). Then, the Android looks away and recites the poems.

## 4. Evaluation of Performance

To analyze the impression of our performance on the audience, we publicly presented the performance on stage and distributed questionnaires. In this study, the questionnaires were delivered to audiences in Austria, Germany, Italy, and France during the autumn of 2011, and in Japan during the summer of 2012. We distributed the questionnaire to all audience members and requested

them to voluntarily complete the forms. The questionnaire contained questions on personal background (e.g., age and gender), psychological background, and evaluation of the android in the performance. The questionnaire data (numbers collected and analyzed) are summarized in Table 1. In addition, 867 questionnaires were collected, but some were not considered because the participants experienced difficulty seeing the stage (108 cases) or hearing the voices of the actresses (244 cases), had a nationality that differed from that of the host country (292 cases), or the participants were under 15 years of age (54 cases). Some questionnaires were rejected for multiple reasons. The reason why we wanted the participants to be nationals of the host country was to explore the cultural influence. We also removed participants under 15 years of age because the psychological scale used to assess the audiences' psychological background was inappropriate for children. Having removed the invalid data, we retrieved no questionnaire data from Italy.

**Table 1.** Number of questionnaires collected and analyzed in each location.

Country	City	Date	Collected	Analyzed
Austria (AUT)	Linz	11 October	346	74
Germany (GER)	Berlin	11 October	110	76
Italy (ITA)	Palermo	11 November	23	0
France (FRA)	Paris	11 November	29	22
	Toulouse	11 November	43	24
Japan (JPN)	Kumamoto	12 June	163	140
	Osaka	12 August	153	113
total			867	449

#### 4.1. Evaluation of the Android

As we based the acting performance of the android on CCTT, we expected that the human-like motions imparted to the android were more refined than those of previous humanoid robots. To determine the extent to which the audience could discriminate the human and the robot performing human motions, we requested participants to state which actor they regarded as the android. We found that 7.4% and 3.6% of the audience could not distinguish the human from the robot at the beginning and end of the performance, respectively. Although our robot was not sufficiently human-like to confuse most of the audience, it certainly confused a portion of participants.

We were also interested in the degree to which the audience perceived the android as human, when they correctly distinguished robot from human. The tendency of attributing human-specific features to androids is expected to be increased. In this study, we requested the audience to evaluate the android in terms of mind (M: two questions), attractiveness (A: three questions), and weight of words (W: two questions). These three aspects collectively indicate perceived human-likeness. Specifically, we asked the audience whether they agreed with the following seven statements on a 7-point Likert scale. We did not adopt verified scales to evaluate difference of robot behavior among conditions in laboratory experiments in the previous study (e.g., [26]). The reason why is that we do not intend to find small differences between conditions in this study because it is difficult to prepare different conditions in public performances. Instead, we adopt exaggerated questions to investigate the accomplishment level of the robot's behavior based on CCTT by checking whether the audience positively grades it against for several baselines of judgment.

#### 1. Mind

- (M1) You could feel emotions from the android.
- (M2) You occasionally felt that the android has a purity that all humans lack.

## 2. Attractiveness

- (A1) The appearance of the android was extremely beautiful.
- (A2) The voice of the android was supremely impressive.
- (A3) You occasionally felt that the android was more attractive than any human.

## 3. Weight of words

- (W1) Did you sense that the android uttered poems from her heart?
- (W2) The android could better enhance the original impression of poem than any human.

M-questions are chosen to evaluate whether audiences perceived a human-like mind inside the android (M1), and which aspects of such a mind were perceived (M2). Purity of mind was chosen as an aspect to be evaluated because the dialogue prepared for the android in the script sounded like the dialogue for a person with a very pure mind. It was considered to be difficult for ordinary persons to have.

A-questions are chosen to evaluate whether the audiences found human-like attractiveness in the android in terms of its appearance (A1) and voice (A2). Additionally, questions assessed whether the observed attractiveness might be felt to exceed human attractiveness (A3). W-questions are chosen to evaluate whether the audiences can find human-like weight of words in the utterances of the android (W1), and whether the found weight of words might be felt to exceed that of human utterances (W2). In the last questions in the above categories (M2, A3, and W2), strong representations were chosen to avoid problems associated with the potential floor effect in evaluations.

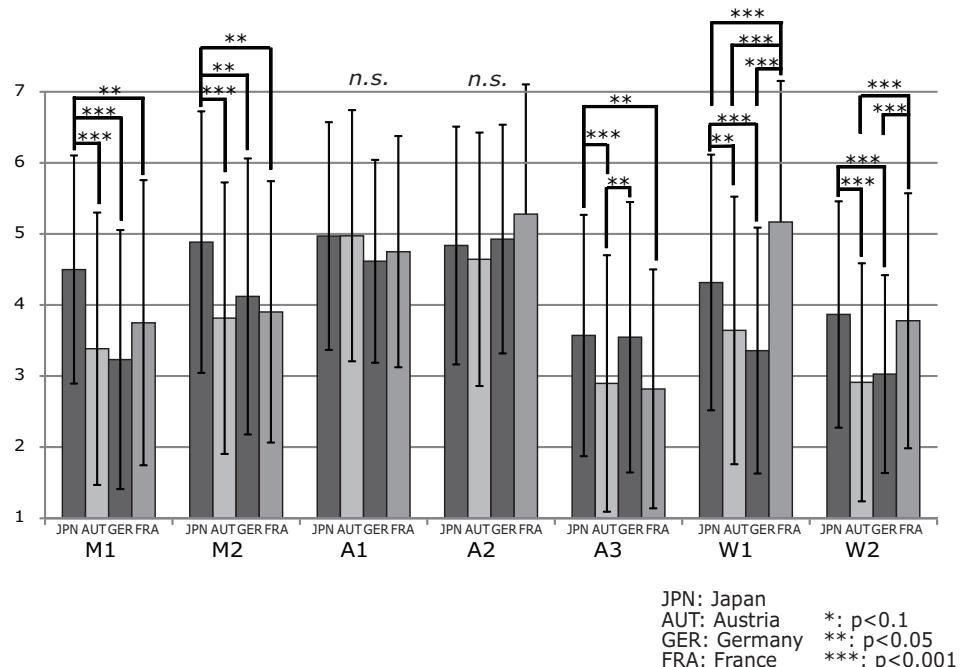
Figure 5 shows the average and standard deviation of the seven questions for each country. The ANOVA (Analysis of variance) results and multiple comparisons with Tukey HSD (honest significant difference), including information on significant differences, are summarized in Table 2. We confirmed that Japanese audiences more strongly recognize human-like emotion (M1) and purity of mind (M2) in the android than audiences of other countries. The effect size is moderate (Table 3).

**Table 2.** Summary of scores for impressions about the android. HSD: honest significant difference.

Q.	F(3,345)	p	f	Post Hoc Tests (Tukey HSD)
M1	22.03	<0.001	0.39	JPN > AUT, GER, FRA
M2	12.3	<0.001	0.29	JPN > AUT, GER, FRA
A1	2.29	<0.1	0.13	n.s
A2	1.84	n.s	0.11	n.s
A3	7.39	<0.001	0.23	JPN > FRA, AUT GER > AUT
W1	18.55	<0.001	0.36	JPN > AUT, GER FRA > JPN, AUT, GER
W2	13.5	<0.001	0.31	JPN > AUT, GER FRA > AUT, GER

**Table 3.** Effect size of Tukey HSD with JPN.

Q.	JPN	AUT	GER	FRA
M1	$M = 4.50$ $SD = 1.60$	$M = 3.38$	$M = 3.23$	$M = 3.75$
		$SD = 1.92$	$SD = 1.82$	$SD = 2.00$
M2	$M = 4.88$ $SD = 1.83$	$r = 0.30$	$r = 0.32$	$r = 0.17$
		$M = 3.81$	$M = 4.12$	$M = 3.90$
		$SD = 1.91$	$SD = 1.94$	$SD = 1.71$
		$r = 0.27$	$r = 0.18$	$r = 0.21$



**Figure 5.** Scores for impressions about the android (ANOVA(Analysis of variance)).

Japanese and German audiences tended to rate the attractiveness of the android higher relative to humans (A3) than audiences of other countries, but the effect size is small. Scores for the attractiveness of appearance and voice did not significantly differ between countries. On the other hand, even when exaggerated options were provided, the audience rated the robot as physically attractive with a medium to large effect size, (A1:  $M = 5.0, SD = 1.5, t(439) = 14.5, p < 0.01, d = 0.69$ ). Beauty of voice also scored highly (A2:  $M = 5.1, SD = 1.5, t(439) = 15.0, p < 0.01, d = 0.71$ ) relative to Point 4 on the Likert scale (i.e., Point 4 means ‘undecided’). The weight of words was more highly rated in Japan and France than in Austria and Germany. The differences were significant, and the effect size was medium.

#### 4.2. Psychological Background

In the former section, we identified potential differences in android evaluations between countries. This finding suggests that an evaluation may be affected by the psychological background of the audiences, itself dependent on cultural background. Therefore, when introducing humanoid robots designed to emotionally engage with human minds into society, we must consider the personalities of users and the cultural background of the society. The type of users who will accept robots with human-likeness is an important practical consideration. To argue this point, we questioned the participants on their personal background.

In evaluating the psychological background, we consider the propensity of humans to animate inanimate objects [27]. To this end, we adopted an animism scale for adults (ASA) [28] in this study. We also used the multidimensional empathy scale (MES) [29] as a measure of understanding toward others. In the present study, where participation is entirely voluntary, raters are frequently reluctant to answer an excessive number of questions. Therefore, we extracted pertinent questions from the above-mentioned psychological scale. Note that the audiences knew that the robot performed on the stage before the stage performance because of poster and flyer advertisements.

From the ASA scale, we selected the following items:

- (Q-1) I can accept that God lives in a big stone or a big tree in the wild.
- (Q-2) I can accept that a sea God lives in the sea and a mountain God lives in the mountain.
- (Q-3) I can feel the ex-owner’s soul in a memento of him or her.

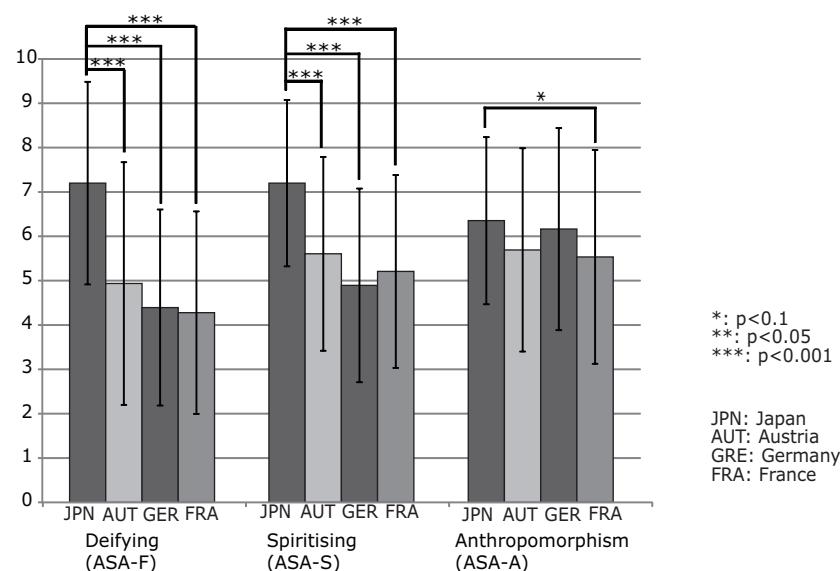
- (Q-4) I sometimes feel the creator's soul in an item made by his or her hand.
- (Q-5) I sometimes feel that I love a possession as if it were human.
- (Q-6) I sometimes name my possessions.

The raters evaluated the extent to which audience members agreed with these six items on a 5-point Likert scale. Six items were extracted from prior research [29], concerning "Deifying (ASA-D)", "Spiritualizing (ASA-S)", and "Anthropomorphizing (ASA-A)" as factors of animism tendency.

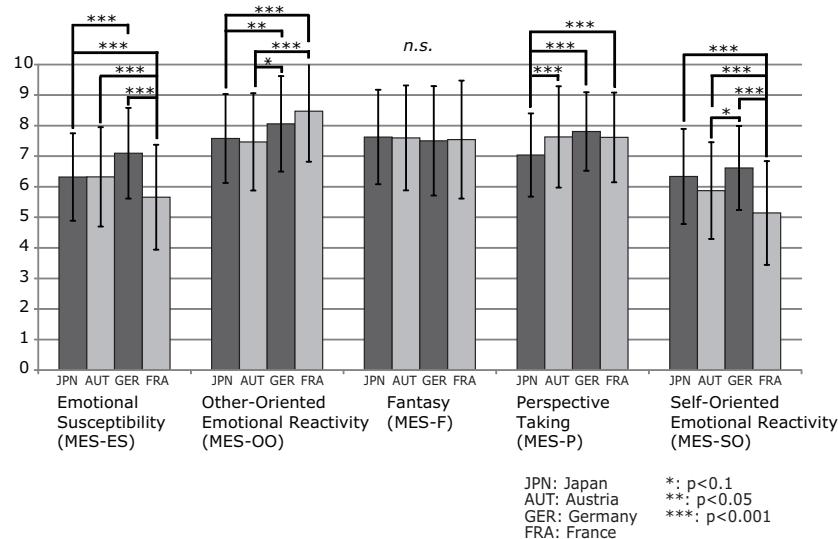
From the MES scale, we selected the following items.

- (Q-7) I usually change my mind based on another's opinion.
- (Q-8) I am not influenced by the feelings of others.
- (Q-9) I feel no sympathy when people make mistakes.
- (Q-11) I regularly daydream and fantasize about things that might happen to me.
- (Q-12) When I watch a movie or play, I do not become emotionally engaged with the actor's or player's situation.
- (Q-13) I sometimes try to understand my friends better by imagining situations from their perspective.
- (Q-14) I sometimes find it difficult to see things from another's viewpoint.
- (Q-15) When I see someone make mistakes, I usually feel repelled by their actions.
- (Q-16) I am easily upset by the success stories of others.

The raters agreed or disagreed with these ten statements on a 5-point Likert scale. Ten items were extracted from a previous study [29], concerning "Emotional Susceptibility (MES-ES)", "Other-Oriented Emotional Reactivity (MES-OO)", "Fantasy (MES-F)", "Perspective Taking (MES-P)", and "Self-Oriented Emotional Reactivity (MES-SO)" as factors of empathy. Figures 6 and 7 display the ASA and MES scores, respectively, of audiences in different countries. The ANOVA results and multiple comparisons with Tukey HSD are summarized in Table 4. Significant differences are found between the scores.



**Figure 6.** Scores for animism scale for adults (ASA).



**Figure 7.** Scores for multidimensional empathy scale (MES).

Although the dependence of these scores on the nationalities of raters is interesting, we focus here on the relationship between the ASA and MES scores and the robot evaluations.

As evidenced by the ASA-D and ASA-S scores, Japanese audiences demonstrate a stronger animism tendency than audiences in the other three countries. The effect size is large. Relating these scores to the above android evaluation scores, we found significant correlation between ASA-D and M1 ( $r = 0.25, p < 0.01, N = 433$ ), ASA-D and M2 ( $r = 0.28, p < 0.01, N = 433$ ), ASA-S and M1 ( $r = 0.25, p < 0.01, N = 431$ ), and ASA-S and M2 ( $r = 0.33, p < 0.01, N = 431$ ). Other correlations were found, but are not reported because their effect sizes were small ( $r < 0.2$ ).

Although the MES scores significantly differ between countries (Table 4), they were not significantly correlated with the android evaluations. Although correlations were found between some pairs, the effect sizes were very small ( $r < 0.2$ ).

**Table 4.** Summary of ASA and MES scores. ASA: animism scale for adults; MES: multidimensional empathy scale. (\*:  $p < 0.1$ , \*\*:  $p < 0.05$ , \*\*\*:  $p < 0.001$ )

	<b>F-Value</b>	<b>Post Hoc Tests (Tukey HSD)</b>
ASA-D	$F(4704) = 48.30, p < 0.01$	JPN > AUT***, GER***, FRA***
ASA-S	$F(4790) = 36.41, p < 0.01$	JPN > AUT***, GER***, FRA***
ASA-A	$F(4761) = 4.35, p < 0.05$	JPN > FRA*
MES-ES	$F(4,781) = 11.47, p < 0.01$	GER > JAN***, FRA*** AUT > FRA***, JPN > FRA***
MES-OO	$F(4791) = 8.20, p < 0.01$	FRA > AUT***, JAN*** GER > JAN**, AUT*
MES-F	$F(4779) = 1.20, n.s.$	n.s.
MES-P	$F(4787) = 9.38, p < 0.01$	JPN > GER***, FRA***, AUT***
MES-SO	$F(4776) = 12.85, p < 0.01$	FRA < GER***, JPN***, AUT*** GER > AUT*

## 5. Discussion

### 5.1. Conquering Negative Feeling and Cultural Gap

Regardless of nationality, all audiences highly evaluated the android relative to Point 4 on the Likert scale (undecided) in terms of beauty (A1, A2). Even though the audience viewed the android for 20 min on the stage and exaggerated options were included in the questionnaire, we may indicate that the android behavior was widely accepted as attractive. The fact that an android with a beautiful appearance is perceived as being beautiful demonstrates that the play-acting android avoids eliciting a negative response. The play was performed in 35 cities (including 16 overseas cities) from September 2010. Since then, we have received invitations from around the world. In the realm of stage art, it is popularly said that the number of invitations is a barometer of a play's quality. Therefore, our stage play might gain more widespread appreciation. In addition, our android was perceived as an attractive entity rather than an negative entity.

Furthermore, we confirmed that Japanese audiences more strongly attributed psychological characteristics to the android (M1, M2) than European audiences. This implies that the android was mentally regarded as a human-like entity. Such characteristics, which are important for the acceptance of fully humanoid robots in society, appear to be culturally dependent, and are unique to Japan [30]. A correlation analysis between psychological background and mental aspects might provide the key to understanding such differences. The uniquely high acceptance of androids by Japanese audiences appears to be related to a high animism tendency (correlation with ASA scores), rather than to the empathy scale (small or no correlation with MES scores). Furthermore, in addition to Japanese audiences, German and French audiences evaluated the android highly in terms of attractiveness (A3) and weight of words (W1 and W2). Although there appeared to be a unique tendency in the MES scores for these audiences, it is not easy to clearly interpret any causal structure between the evaluation of the android and the empathic tendencies of the audiences: how can emotional susceptibility or other-oriented emotional reactivity be the only or dominant factor in a bias evaluation of the android (in terms of its attractiveness or weight of words)? Unfortunately, because the questionnaire volume was limited by practical considerations, we could not evaluate the full impact of psychological background in this study, and further verification and discussion is required to determine whether or not there are other causalities or cultural dependencies in the evaluation of the android.

### 5.2. Future Issues

By creating a stage play partially enacted by robots, we can contribute knowledge to the construction of highly human-like robots. Our approach is very different from the conventional approach adopted in previous robotics studies. The most important advance of our current study is that the android is regarded as attractive and mentally engaging, even after long-term observation. Future work must identify the factors or rules that yield attractive behaviors, and generalize them into a methodology for social robot development.

CCTT is ideally suited for designing robot behavior because it provides definitive instructions. Thus, CCTT will form the focus of our future study. In the current study, we could find that the CCTT instructions were very physical and similar to those provided to human actors, such as "move its (your) standing point 5 cm to the right". To utilize these instructions in humanoid robot design, we must integrate a robotic system with sensors through which the stage director can mediate his or her instruction. The extent to which robots can be rendered physically and emotionally attractive by CCTT-based physical instruction is an interesting question.

Developing a robot system that can modify its behavior according to an instruction database is an important but nontrivial problem. One promising approach is to integrate a sensing system, such as a network of laser range finders, which identifies actors' stage positions. The data so collected could then be associated with instructions dictating the stage situation. A generalized robot mechanism

that behaviorally adapts to situational changes recognized by sensor networks should be realized in the future.

As we are interested in factors on the audience's evaluation, we adopted the excerpted version of scales established in the previous work. However, identifying essential predispositions on the perception of robots from the data obtained in the experiments in public performance could be another interesting topic of discussion. For this direction, a large amount of answers for a larger number of questions would be required, rather than those used in this study. As collecting such data is difficult in our setting of evaluation, public performances, we should consider more effective ways of collecting answers from one public performance and keep holding public performance.

We have already created other drama works with android. In one, an android with a human-like appearance, a robot with a mechanical design, and nine human actors appear on stage and interact with one another. In another, an android with a half-human-like and half-mechanical appearance performs with four human actors. In other words, we have already created different or more complex experimental situations. For example, we can design other conditions such as one-to-many conversations or different appearances of robots on stage.

### 5.3. Limitations

Although we show that an android directed by CCTT can have such a natural human robot interaction that observers' positive feelings are induced, the behavior is more or less limited to the situation presented in the script. Generalizing knowledge of naturalistic human–robot interaction involved in successful examples such as ones evaluated in this study is considered to be necessary for an android to work in the real daily environment beyond the stage. It is, therefore, a very important next step, and we have already started preliminary trials exploring this direction [31].

The questionnaires are translated into each language from Japanese. Therefore, questions of questionnaire might be recognized as having different meaning in different countries because of cultural gaps. It is the difficulty of starting cross-cultural studies in a new domain. We need to keep developing appropriate scales for investigations in multi-country theatrical settings, based on the obtained data, as one of the important future works.

This study has a limitation that is related to the duration of the interaction. Noma et al. reported that after 2 s, many humans can find an android to be unnatural [6]. On the other hand, we conjectured that almost all interactions in our daily life are less than 20 min. We therefore assume that the 20 min of natural interaction are sufficiently long enough for improving the current android. However, interactions can sometimes be over 20 min, especially between people who share close relationships, such as family members, intimate friends, and co-workers. Therefore, building androids that humans can interact with for over 20 min is part of our future work.

We found that androids did not evoke negative responses when used in theatrical plays. From this, we can discuss the possible impressions a human might have of a robot when interacting with it in a daily situation. There is a possibility, however, that the android would evoke a negative response during a face-to-face interaction. A control experiment is therefore needed to more precisely understand the differences in the evaluation of androids between daily situations and theatrical uses.

Several poems were read out on stage as listed in the Section 3.1. The different scores for "weight of word" among countries might be ascribed to the different familiarity of the audiences with each poem. However, because one of the poems was German, it was difficult for us to argue that the different evaluations of W-questions by Japanese and French audiences were caused by the fact that they were the only groups with a chance to listen to a national and familiar poem. However, it is also difficult for us to completely reject this possibility because we did not collect data on audiences' familiarity with each poem read on stage. For potential future meta-analyses, the full list of poems used on the stage is included in this paper.

In the experiment, we had to limit ourselves to a small number of questions, in order to decrease the burden on the audience as much as possible. This was done because we had to ask for volunteers

to answer our questions. Because we prioritized collecting as much data as possible, we selected a limited number of direct and specific questions of interest. When subjective reports are used to evaluate subjective impressions, researchers typically must ask several questions about each topic of interest. Repeatedly asking participants the same or similar questions can produce a more reliable investigation. Therefore, further investigations, including extensive surveys of audience impressions, should be conducted to reach more reliable conclusions.

In this study, we are not certain of the extent to which the obtained evaluations were influenced by the stage context. The treatment of subjective evaluation independently of the interaction context is an open problem of HRI (Human-Robot Interaction), and one that should be addressed in our work. To move beyond a report on a single experiment focusing on a specific context, and meaningfully generalize it to real-world situations involving live (non-scripted) interactions, we need to gather and accumulate many evaluations of realistic scenarios of human–robot interaction in changing contexts. We believe that android theatre studies can effectively contribute to this approach from the viewpoint of preparing necessary, very realistic pieces of interaction to be evaluated.

In this study, it is not very clear how much the obtained evaluations were influenced by the context of the stage. The treatment of subjective evaluation independently of the context of interaction is a kind of open problem of HRI and should be also concerned in our work. To go beyond the report of a single experiment focusing on a specific context to meaningfully generalize it, we need to pile up many evaluations of realistic scenes of human–robot interaction by changing contexts. We believe that android theatre studies can effectively contribute to this approach in terms of creating necessary, very realistic pieces of interaction to be evaluated.

## 6. Conclusions

This study investigated a novel challenge in the interdisciplinary field that lies between robotics and art, namely, creating “Android–Human Theatre”. Adopting the stage art methodology of CCTT, we developed a natural human–robot conversation. The play was publicly performed in many theatres throughout several countries. The feasibility of the proposed methodology was indicated by the high physical and emotional attractiveness scores awarded to the android by audiences of different cultures. Notably, audiences awarded these high scores after watching 20 min of human–android interaction on the stage, different from previous studies. This suggests that robot behavior can be rendered acceptable through public performances, although human acceptance of androids is culturally dependent to some extent. It is an important future direction to establish a more firm, CCTT-based methodology for designing highly human-like robots through successive creation and further analysis of androids’ behavior.

**Acknowledgments:** This research is partially supported by the Center of Innovation Program from Japan Science and Technology Agency, JST. The author would like to thank the members of the Seinendan Theatre Company who supported this study incredibly.

**Author Contributions:** Hiroshi Ishiguro conceived and designed the android robot; Oriza Hirata wrote the manuscript and directed the stage play; Kohei Ogawa contributed the robot material and robot software, and contributed to the analysis of the data; Yuichiro Yoshikawa designed the experiments and contributed to the analysis of the data; Takenobu Chikaraishi performed the stage play and the experiments, analyzed the data, and wrote the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Hirai, K.; Hirose, M.; Haikawa, Y.; Takenaka, T. The development of Honda humanoid robot. In Proceedings of the 1998 IEEE International Conference on Robotics and Automation, Leuven, Belgium, 20 May 1998; Volume 2, pp. 1321–1326.
2. Kanda, T.; Ishiguro, H.; Ono, T.; Imai, M.; Maeda, T.; Nakatsu, R. Development of Robovie as a Platform for Everyday-Robot Research. *Electron. Commun. Jpn.* **2004**, *87*, 55–65.

3. Ishiguro, H. Scientific Issues Concerning Androids. *Int. J. Robot. Res.* **2007**, *26*, 105–117.
4. Hashimoto, T.; Hiramatsu, S.; Kobayashi, H. Dynamic display of facial expressions on the face robot made by using a life mask. In Proceedings of the 2008 8th IEEE-RAS International Conference on Humanoid Robots, Humanoids 2008, Daejeon, Korea, 1–3 December 2008; pp. 521–526.
5. Mori, M. The uncanny valley. *Energy* **1970**, *7*, 33–35.
6. Noma, M.; Saiwaki, N.; Itakura, S.; Ishiguro, H. Composition and evaluation of the humanlike motions of an android. In Proceedings of the 2006 6th IEEE-RAS International Conference on Humanoid Robots, Genova, Italy, 4–6 December 2006; pp. 163–168.
7. MacDorman, K.F.; Ishiguro, H. Opening Pandora’s uncanny Box: Reply to commentaries on “The uncanny advantage of using androids in social and cognitive science research”. *Interact. Stud.* **2006**, *7*, 361–368.
8. Chikaraishi, T.; Minato, T.; Ishiguro, H. Development of an Android System Integrated with Sensor Networks. In Proceedings of the 2008 IEEE International Conference on Intelligent Robots and Systems (IROS 2008), Nice, France, 22–26 September 2008; pp. 326–333.
9. Ogawa, K.; Taura, K.; Ishiguro, H. Possibilities of Androids as poetry-reciting agent. In Proceedings of the 2012 IEEE RO-MAN, Paris, France, 9–13 September 2012; pp. 565–570.
10. Ishiguro, H.; Hirata, O. Robot theater. *Robot. Soc. Jpn.* **2011**, *29*, 35–38. (In Japanese)
11. Stanislavski, K. *An Actor’s Work*; Routledge Classics; Taylor & Francis: Abingdon, UK, 2016.
12. Mitchell, K. *The Director’s Craft: A Handbook for the Theatre*; Taylor & Francis: Abingdon, UK, 2008.
13. Craig, E. *The Actor and the Über-Marionette; The Mask*; Florence, Italy, 1908.
14. Hirata, O. *Gendai Kogo Engeki No Tamemi (for Contemporary Colloquial Theater)*; Banseisha: Tokyo, Japan, 1995. (In Japanese)
15. Duncan, B.; Murphy, R.; Shell, D.; Hopper, A. A Midsummer Night’s Dream: Social proof in HRI. In Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction, Osaka, Japan, 2–5 March 2010; pp. 91–92.
16. Demers, L. Machine Performers: Neither Agentic nor Automatic. In Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction, Osaka, Japan, 2–5 March 2010.
17. Lemaignan, S.; Gharbi, M.; Mainprice, J.; Herrb, M.; Alami, R. Roboscopie: A theatre performance for a human and a robot. In Proceedings of the 2012 7th ACM/IEEE International Conference on Human-Robot Interaction, Boston, MA, USA, 5–8 March 2012; pp. 427–427.
18. Zeglin, G.; Walsman, A.; Herlant, L.; Zheng, Z.; Guo, Y.; Koval, M.C.; Lenzo, K.; Tay, H.J.; Velagapudi, P.; Correll, K.; et al. HERB’s Sure Thing: A rapid drama system for rehearsing and performing live robot theater. In Proceedings of the 2014 IEEE International Workshop on Advanced Robotics and its Social Impacts, Evanston, IL, USA, 11–13 September 2014; pp. 129–136.
19. Lu, D. Ontology of Robot Theatre. In Proceedings of the ICRA 2012 Workshop Robotics and Performing Arts: Reciprocal Influences, St. Paul, MN, USA, 14 May 2012.
20. Jochum, E.; Millar, P.; Nuñez, D. Sequence and Chance: Design and control methods for entertainment robots. *Robot. Auton. Syst.* **2017**, *87*, 372–380.
21. Jochum, E.; Vlachos, E.; Christoffersen, A.; Nielsen, S.G.; Hameed, I.A.; Tan, Z.H. Using Theatre to Study Interaction with Care Robots. *Int. J. Soc. Robot.* **2016**, *8*, 457–470.
22. Sakamoto, D.; Kanda, T.; Ono, T.; Ishiguro, H.; Hagita, N. Android as a telecommunication medium with a human-like presence. In Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction, Arlington, VA, USA, 9–11 March 2007; pp. 193–200.
23. Seeing Machines Inc. faceAPI. Available online: <https://www.seeingmachines.com/download/new-faceapi-product-release/sm10907-1-faceapi-release-announcement/> (accessed on 1 November 2017).
24. Ishii, C.; Liu, C.; Ishiguro, H.; Hagita, N. Speech-driven lip motion generation for tele-operated humanoid robots. In Proceedings of the International Conference on Auditory-Visual Speech Processing, Volterra, Italy, 1–2 September 2011; pp. 131–135.
25. Goan, M.; Fukaya, T.; Tsujita, K. Two Succeeding Stages in Acquisition Process of a Rehearsed Drama: Applying System Dynamics to Human Collaborative Behavior. In Proceedings of the 12th World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2008), Orlando, FL, USA, 29 June–2 July 2008; Volume 7, pp. 126–131.
26. Bartneck, C.; Croft, E.; Kulic, D.; Zoghbi, S. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *Int. J. Soc. Robot.* **2009**, *1*, 71–81.

27. Kitano, N. Animism, Rinri, Modernization; the Base of Japanese Robotics. In Proceedings of the IEEE International Conference on Robotics and Automation, Roma, Italy, 10–14 April 2007; pp. 10–14.
28. Ikeuchi, H. Animistic thinking in adults : The memorial service for dolls as a voluntary loss. *Jpn. J. Soc. Psychol.* **2010**, *25*, 167–177. (In Japanese with English abstract)
29. Suzuki, Y.; Kino, K. Development of the Multidimensional Empathy Scale (MES) : Focusing on the Distinction Between Self- and Other-Orientation. *Jpn. J. Educ. Psychol.* **2008**, *56*, 487–497. (In Japanese with English abstract)
30. Kaplan, F. Who is afraid of the humanoid? Investigating cultural differences in the acceptance of robots. *Int. J. Humanoid Robot.* **2004**, *1*, 465–480.
31. Ogawa, K.; Chikaraishi, T.; Yoshikawa, Y.; Nishiguchi, S.; Hirata, O.; Ishiguro, H. Designing Robot Behavior in Conversations based on Contemporary Colloquial Theatre Theory. In Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication, Edinburgh, UK, 25–29 August 2014.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).