Shaping Robot Gestures: The Effect of Amplitude, Speed on Users' Perception

Amol Deshmukh, Bart Craenen, Alessandro Vinciarelli and Mary Ellen Foster
University of Glasgow, School of Computing Science, Glasgow, United Kingdom
Amol.deshmukh@glasgow.ac.uk

ABSTRACT

This study investigates the interplay between the way a robot displays different gestures with changes in amplitude and speed and the way these gestures are perceived by the users. The results with 30 participants show that, at least in some cases, there is an association between the speed and amplitude of a gesture—two parameters that account for energy and spatial extension—and the scores on the Godspeed questionnaire. It shows that the Godspeed scores tend to be different for different values of amplitude and speed. The main implication of such an observation is that, for a social robot, it is not sufficient to decide what gestures a robot should display during an interaction, but also how the gestures are performed in order to make them self-explainable.

CCS CONCEPTS

• **Human-centered computing → User models**; • **Computer systems organization → Robotic autonomy**; • **Computing methodologies → Computational control theory**;

KEYWORDS

Synthetic Gestures, Perception, Human-robot interaction

ACM Reference Format:
Amol Deshmukh, Bart Craenen, Alessandro Vinciarelli and Mary Ellen Foster . 2018. Shaping Robot Gestures: The Effect of Amplitude, Speed on Users' Perception. In *Proceedings of HRI (Workshop on Explainable Robotic Systems).* ACM, New York, NY, USA, Article 4, 3 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

Robots that operate in public spaces in noisy environments and crowded spaces, in which the level of acoustic noise tends to be high enough to make it difficult to hear and understand speech. Thus, the use of gestures and other bodily enacted cues play a critical role in conveying robot's intentions or behaviors for humans [3]. For this reason, the experiments presented in this study focus on isolated gestures that do not accompany or interact with spoken messages. Our approach proposed in this article does not only take into account what the gestures are that the robot displays, but also the way in which the robot displays these gestures. In particular, the experiments investigate the association between variations of *amplitude* and *speed*—two major parameters that characterize any natural and synthetic gesture—and variations of the users' perception as measured by the Godspeed questionnaire [1].

Workshop on Explainable Robotic Systems, March 2018, Chicago, USA

ACM ISBN 123-4567-24-567/08/06…$15.00
https://doi.org/10.475/123_4

2 SCENARIO

This work is being carried out in context of the MultiModal Mall Entertainment Robot (MuMMER) project, a four-year, EU-funded project[1], with the overall goal of developing a humanoid robot, *Pepper* (a robotic platform manufactured by Softbank Robotics), that can interact autonomously and naturally in the dynamic environments of a public shopping mall. The overall concept underlying MuMMER is that for a robot to be successful in such a situation, it must be *entertaining* and *engaging*: that is, they must possess the social intelligence to both understand the needs and interactive behaviour of the users, as well as to produce understandable behaviour in response [4]. From the default animations provided with the Pepper robot, we shortlisted a set of five which we anticipate to be useful for the shopping-mall scenarios addressed in the MuMMER project: animations for Engage-Gaining attention, Disengaging-Send-away, Pointing-Directions, Head-touching-Disappointment, Cheering-Success/Happy. The selection targeted gestures that, according to the criteria underlying the taxonomy proposed in [14], are relevant to the scenario addressed in this work, i.e., the interaction between people and robots in public spaces.

3 EXPERIMENTAL APPROACH

The goal of this study is to investigate the way the perception of the users changes depending on the properties of the gestures that a robot displays. During the experiments, 30 independent human observers selected randomly from a pool of subjects available at the university: 20 of them are female and 10 are male. Observers have been asked to watch 45 different gestures performed by *Pepper* and to complete, for each gesture, the Godspeed questionnaire [1]. The 45 gestures represent 9 variants of 5 animations selected from the standard library available with the robot, refer Table 1. The 9 variants of each core gesture have been obtained by manipulating two parameters, namely *speed* and *amplitude*.

Three variants were generated by adopting three different values of the speed λ per core stimulus: 15, 25 and 35 *frames per second* (*fps*), where 25 *fps* is the original speed of the core stimuli. For each of the 15 resulting gestures, another three stimuli can be obtained by modifying the differences $\Delta_i(t) = \theta_i(t) - \theta_i(t-1)$, where $\theta_i(t)$ is the angle between the two mechanical elements connected by joint i at frame t. In particular, the values of the $\Delta_i(t)$ were multiplied, for all values of i and t, by a factor α—the *amplitude* hereafter. Three different values of α were adopted, namely 0.50, 0.75 and 1.00.

[1] www.mummer-project.eu/

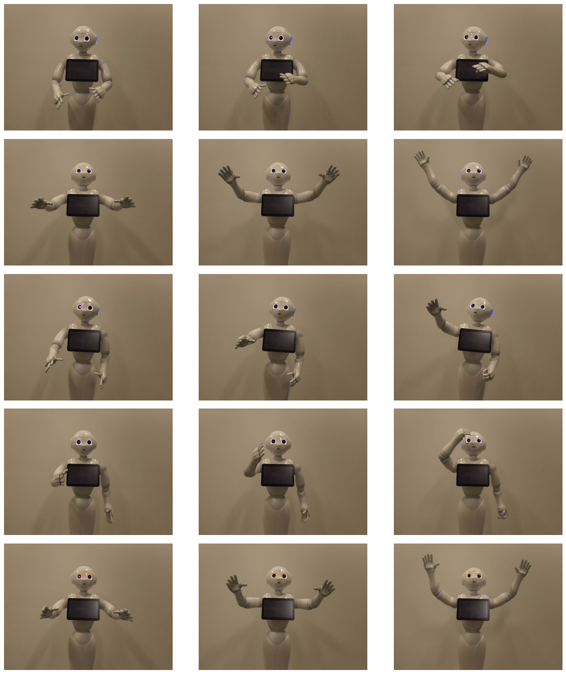

Table 1: The figures show, for each of the five core stimuli, the effect of the parameter α. The rightmost column ($\alpha = 1.00$) contains the core stimuli.

In this way, it is possible to investigate whether there is an association between the way a gesture is performed and the perception of the users. The motivation behind the choice of speed and amplitude is that they are related to energy and spatial extension, respectively, two characteristics that have been shown to play a crucial role in the expressiveness of artificial agents [8].

4 GESTURES AND PERCEPTION

The first question addressed in this study is whether users differently perceive robots that display different gestures and, if yes, how the perception changes based on the characteristics of the gestures. During the experiments, $N = 30$ human observers filled in the Godspeed questionnaire [1] after watching each of the 45 stimuli (all observers observed and rated all stimuli). The Godspeed questionnaire is widely accepted as a standard measurement tool for Human Robot Interaction and aims to quantify the following tendencies underlying users' perception:

(1) *Anthropomorphism*: tendency of human users to attribute human characteristics to a robot;
(2) *Animacy*: tendency of human users to consider the robot alive and to attribute intentions to it;
(3) *Likeability*: tendency of human users to attribute desirable characteristics to a robot;
(4) *Perceived Intelligence*: tendency of human users to consider intelligent the behaviour of a robot;
(5) *Perceived Safety*: tendency of human users to consider safe the interaction with a robot.

Completing the questionnaire results in five scores that measure the tendencies above: the higher the score, the more pronounced the tendency (see [1] for full details).

Table 2 shows the cases in which the distribution of the Godspeed scores across the multiple variants of the same core stimulus deviates, to a statistically significant extent, from the uniform distribution. Furthermore, when the deviation is statistically significant, the table shows whether increasing amplitude and speed of a gesture corresponds to higher or lower Godspeed scores. A deviation from the uniform distribution is considered statistically significant when a χ^2 test results into a p-value lower than 0.05. The *False Discovery Rate* (FDR) correction [2] has been applied to tackle the multiple comparisons problem.

	Ant		Ani		Lik		Int		Saf	
Core Stimulus	α	λ	α	λ	α	λ	α	λ	α	λ
Engaging	↑	↑	↑	↑	↑	↑				
Disengaging					↓	↓			↓	↓
Pointing										
Head-Touching			↑	↑						
Cheering			↑	↑						

Table 2: The symbols "↑" and "↓" account for statistically significant effects. The symbol "↑" means that increasing amplitude or speed corresponds to observing higher Godspeed scores. The symbol "↓" means that decreasing amplitude or speed corresponds to observing lower Godspeed scores. Empty cells correspond to cases in which no statistically significant effects have been observed.

For the Disengaging gesture, a significant effect was found for Likeability and Perceived Safety. In the former case, the scores tend to decrease when α and λ increase, while in the latter case the scores tend to increase when α and λ decrease, respectively. The possible explanation behind the Likeability effects is that this gesture aims to increase the physical distance between the robot and its users. Given that physical and social distances have been shown to be equivalent (the longer the former, the longer the latter) [10], increasing the energy of the gesture might appear to be an attempt by the robot to push people towards distances that, according to proxemic theories [7], correspond to less friendly and more formal relationships. As far as the relationship with Perceived Safety, the probable explanation is that slower movements (lower λ) that do not extend far from the robot's body (lower α) appear less likely to harm the users.

In the case of the Engaging gesture, statistically significant effects have been observed for Anthropomorphism, Animacy and Likeability. In all three cases, increasing the amplitude and speed corresponds to higher Godspeed scores. In the case of Anthropomorphism, one possible explanation is that the human brain has been shown to be more anthropomorphic—meaning that it is more prone to process artificial agents like it processes human ones—when synthetic movements are more similar to those displayed by humans [6]. Lowering α and λ actually produces gestures that, at least in the case of the Engage core stimulus, are less similar to those performed by humans.

The possible explanation for the Animacy effects is that higher speed and amplitude result into higher energy and motor activation, two factors that play a crucial role when it comes to consider an agent alive or lively [1]. Finally, the increase in Likeability scores is likely to depend on the correlation between Anthropomorphism and positive judgements about the robots that have been observed earlier in the literature [16]. Overall, the three effects observed for the Engaging gesture are an advantage in those scenarios in which the robot is expected to pro-actively initiate an interaction with the users: in particular, the effects indicate a possible mechanism for making the perception of the users more positive—a prerequisite towards successful interactions with machine that display human-like behaviour (see, e.g., [13])—at the very moment they enter in contact with the users.

No statistically significant effect was observed for the Pointing gesture. A possible explanation is that deictic gestures such as pointing are meant to convey information about spatial knowledge [9]—in particular when it comes to the position of an object of interest in the environment—and not about the social and psychological phenomena underlying the items of the Godspeed questionnaire [1].

Finally, both the Head-Touching and Cheering gestures show significant effects on the Animacy scale. The main probable reason is that both gestures, when displayed by people, tend to convey information about one's inner state. In particular, Head-Touching is typically associated with a situation of confusion [12, 15] while Cheering tends to be displayed as a sign of success, satisfaction, and accomplishment [5]. This means that a robot displaying the two gestures above can elicit the attribution of the same inner states and, ultimately, of Animacy, defined as the very property of being alive [1].

For both Head-Touching and Cheering, the Animacy scores tend to increase when both α and λ increase. In the case of α, the probable reason is that lowering this parameter leads to gestures that have a morphology different from the core stimulus and, hence, fail in conveying the same impression. In the case of λ, the probable reason is that movements have been shown to play a crucial role in the attribution of Animacy, the very difference between animate beings and inanimate objects [1]. Thus, increasing the movement's energy (proportional to speed) tends to attract higher Animacy scores.

5 CONCLUSION

This study presented experiments on the interplay between the way a gesture is performed and the perception of the users. The results show that, at least in some cases, there is an association between the speed and amplitude of a gesture—two parameters that account for energy and spatial extension—and the scores on the Godspeed questionnaire [1]. Overall, the coherent picture that emerges is that, for gestures expected to achieve a social goal—Engaging and Disengaging—effects were found on the Godspeed dimensions that better account for social aspects of Human Robot Interaction, namely Anthropomorphism (the tendency to attribute human characteristics to the robot) and Likeability (the tendency to attribute desirable characteristics to the robot). Similarly, gestures designed to simulate an "inner state"—Head-Touching and Cheering—show effects on Animacy, the Godspeed dimension that captures the tendency to consider the robot alive and, hence, capable of experiencing the world. Finally, there are no effects for Pointing—which, unlike the other stimuli used in the experiments, aims to share knowledge about the environment more than to convey information about the robot properties embodied in the Godspeed questionnaire.

The above suggests that the stimuli have been designed correctly and, most importantly, it shows that the Godspeed scores tend to be different for different values of amplitude and speed. The main implication of such an observation is that, for a social robot, it is not sufficient to decide what gestures a robot should display to convey its intention, but also how the gestures are performed in order to make it more understandable and self-explainable to humans. In particular, the same gesture should be displayed with different amplitude and speed depending the target impression to convey on the tendencies underlying the Godspeed scores. However, future work will aim at investigating how the findings of this study can possibly change when the gestures are accompanied by speech, as in the most frequent case in everyday interactions [11].

6 ACKNOWLEDGEMENTS

This research has been partially funded by the European Union's Horizon 2020 research and innovation programme under grant agreement no. 688147 MuMMER project.

REFERENCES

[1] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics* 1, 1 (2009), 71–81.
[2] Y. Benjamini and Y. Hochberg. 1995. Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B* (1995), 289–300.
[3] C. Breazeal, C.D. Kidd, A.L. Thomaz, G. Hoffman, and M. Berlin. 2005. Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on.* IEEE, 708–713.
[4] Maartje MA de Graaf and Bertram F Malle. 2017. How People Explain Action (and Autonomous Intelligent Systems Should Too). *AAAI Technical Report* (2017).
[5] C. Gardair. 2013. *Assembling Audiences.* Ph.D. Dissertation. Queen Mary University of London.
[6] V. Gazzola, G. Rizzolatti, B. Wicker, and C. Keysers. 2007. The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. *Neuroimage* 35, 4 (2007), 1674–1684.
[7] E. Hall. 1959. *The silent language.* Doubleday.
[8] B. Hartmann, M. Mancini, and C. Pelachaud. 2005. Implementing expressive gesture synthesis for embodied conversational agents. In *Proceedings of International Gesture Workshop.* 188–199.
[9] J. Haviland. 2000. Pointing, gesture spaces, and mental maps. In *Language and Gesture*, D. McNeill (Ed.). Cambridge University Press, 13–46.
[10] A. Kendon. 1990. *Conducting Interaction.* Cambridge University Press.
[11] A. Kendon. 2000. Language and gesture: unity or duality? In *Language and Gesture*, D. McNeill (Ed.). Cambridge University Press, 47–63.
[12] M.L. Knapp and J.A. Hall. 1972. *Nonverbal Communication in Human Interaction.* Harcourt Brace College Publishers.
[13] C. Nass and S. Brave. 2005. *Wired for speech: How voice activates and advances the human-computer relationship.* MIT Press.
[14] C.L. Nehaniv, K. Dautenhahn, J. Kubacki, M. Haegele, C. Parlitz, and R. Alami. 2005. A methodological approach relating the classification of gesture to identification of human intent in the context of human-robot interaction. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on.* IEEE, 371–377.
[15] V.P. Richmond, J.C. McCroskey, and S.K. Payne. 1991. *Nonverbal behavior in interpersonal relations.* Prentice Hall.
[16] M. Salem, F. Eyssel, K. Rohlfing, S. Kopp, and F. Joublin. 2013. To err is human (-like): Effects of robot gesture on perceived anthropomorphism and likability. *International Journal of Social Robotics* 5, 3 (2013), 313–323.