Room for one more? - Introducing Artificial Commensal Companions

Maurizio Mancini
School of Computer Science and Information Technology,
University College Cork
Cork, Ireland
m.mancini@cs.ucc.ie

Merijn Bruijnes
Interactive Intelligence Group
Delft University of Technology,
Delft, The Netherlands
m.bruijnes@tudelft.nl

Radoslaw Niewiadomski
Istituto Italiano di Tecnologia
Genoa, Italy
Radoslaw.Niewiadomski@iit.it

Conor Patrick Gallagher
School of Computer Science and Information Technology,
University College Cork
Cork, Ireland
cpg.gallagher@gmail.com

Gijs Huisman
Digital Society School,
Amsterdam University of Applied Sciences
Amsterdam, The Netherlands
g.huisman@hva.nl

CHI '20 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6819-3/20/04.
http://dx.doi.org/10.1145/3334480.3383027

Abstract
Commensality is defined as "a social group that eats together" and eating in a commensality setting has a number of positive effects on humans. In this paper, we discuss how HCI and technology in general can be exploited to replicate the benefits of commensality for people who choose or are forced to eat alone. We discuss research into and the design of Artificial Commensal Companions that can provide social interactions during food consumption. We present the design of a system, consisting of a toy robot, computer vision tracking, and a simple interaction model, that can show non-verbal social behaviors to influence a user's food choice. Finally, we discuss future studies and applications of this system, and provide suggestions for future research into Artificial Commensal Companions.

Author Keywords
food; eating; hci; robot; companion; interaction; non-verbal; commensality.

CCS Concepts
•**Human-centered computing** → *Human computer interaction (HCI);* Interaction paradigms;

Introduction

Food consumption is a highly social activity. Social psychology has shown that eating in the company of others, or being in a "commensality" setting, has a number of positive effects on humans, such as healthier food choices, reduced "over-eating", and increased enjoyment of food, as well as triggering a number of positive emotions [24]. In particular, "ambient effects on food intake" and the "social facilitation effect" of food consumption have been demonstrated [24]. For example, it has been shown that meal duration and group size are positively correlated; also, eating in a group increases the rating of food palatability, while eating with strangers decreases food intake [24].

Despite the benefits of social eating, current trends in Western society are resulting in reduced commensality [1, 2]: people consciously choose, or are forced, to eat alone. An example of the former is when people choose to eat in front of the television or while using their smartphone; an example of the latter is elderly people living alone, or those living with physical (e.g., living away from family due to work or study) and social (e.g., a lack of friends) barriers preventing them from experiencing commensality. Note that social isolation is expected to increase in the coming decades, especially in Western countries [23], so the decline in commensality is expected to continue in the near future. At the same time, according to a recent study carried out by the Oxford Economics and the National Centre for Social Research, eating meals alone is an important cause of unhappiness in UK (measured with by Sainsbury's Living Well Index in 2018[1]). It is therefore timely to investigate the potential role of HCI in enhancing social connectedness in general, and during food consumption in particular.

[1] https://www.about.sainsburys.co.uk/~/media/Files/S/Sainsburys/living-well-index/sainsburys-living-well-index-may-2018.pdf, visited 12 Feb 2020.

Commensality and Technology

Nowadays, technology is already widely present at the dining table. Think, for example, of people using their smartphone while eating to send text messages to friends, share photos of their food on social media, or watch videos [8]. While these technologies already have an impact on social dynamics during food intake [23], other more sophisticated technologies could be envisioned to better integrate into the social structures around food consumption. This is also one of the main interests in Human-Food Interaction (HFI) research, an emerging research area that looks at the intersection of technology, human interaction, and food practices [7]. Works within HFI have the goal of enriching the social dimension of a meal by investigating mealtime technologies, e.g., [5, 8, 18].

Digital Commensality and *Computational Commensality* have been defined in [19, 23] to refer to, respectively, "eating and drinking in the company of technology" and "computational models for social food preparation and consumption in HCI". Computational Commensality, more specifically, refers to the computational models and techniques that may underlie digital technology at the dining table and help guide HCI researchers and designers to create scenarios in which technology plays a key role. The idea behind these concepts is the following: by properly designing and implementing commensality interfaces, that is, computer interfaces creating the social interactions around eating activities, we could provide users with the beneficial effects characterizing human commensality in contexts that, otherwise, would force them to dine alone. However, there are challenges in implementing digital commensality to allow human users to eat "in company of technology" [23]. Discussing artificial dining assistants, authors of [23] highlight that there is a "limited extent of research investigating the extent to which these approaches deliver the same health/well-being benefits that are associated with physically dining together with another person/other people". Others, like [19], propose a *computational commensality scenario*, in which "human(s) interact with an artificial companion, such as a social robot during meal time. The companion uses sensors and computational models of commensality to guide its behavior toward the human interlocutor". However, as stated in [19], the existing "research in AI and HCI and technologies dedicated to Food- & Eating-related Activities often focus on food (or eating) itself (e.g., food recognition and sensory augmentation) rather than on its social dimension". Here, we take a first step in designing a companion and its underlying computational model that are focused on social eating and drinking.

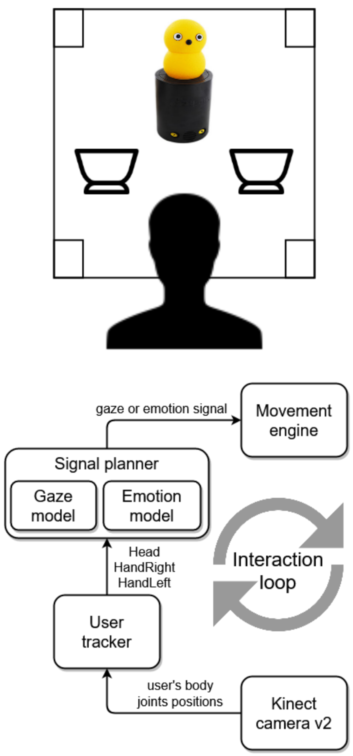

Figure 1: Artificial Commensal Companion setup (upper part) and system overview (lower part).

Artificial Commensal Companions

An Artificial Commensal Companion (ACC) can be thought of as a virtual agent or social robot designed to interact with humans during meal time, while taking into account, and being able to act upon, all the social intricacies that occur during a meal. A few examples of ACCs already exist, such as the robotic companion called FoBo [10]. Its aim is to create playful and entertaining interactions around a meal with no clear "real-world" goal. It may "consume" batteries, perform sounds related to eating (e.g., burping and purring) and mimics some of the human behaviors. Liu and colleagues [15] proposed a virtual eating companion being at the same time an active listener. The companion role is to support the generation of new ideas. Takahashi and colleagues [25] developed a virtual co-eating system allowing enjoyable conversations related to the meal, as well as typical daily conversation. The results of a preliminary evaluation with 5 participants showed positive outcomes of using this technology. Other works within HFI [7] have the goal of enriching the social dimension of a meal by investigating mealtime technologies, e.g., [8, 18]. Though they are not much focused on co-located social interactions, they could inform and inspire our work on designing and implementing ACCs, as they investigate mealtime social interactions and experiences.

Building an ACC can be an interesting challenge for HCI and the examples above demonstrate more playful implementations, build around simple social scenarios. In the design of the behavior of artificial companions to "replace" humans in more complex interactive scenarios, the most common approach is to imitate human behavior as much as possible [6]. Usually, a large corpus of human-human interactions is collected, then machine learning or other algorithms are used to model the companion's behavior. Examples of such an approach exist in literature, for example laughing interactive virtual agents [16, 20] or the SEMAINE project's listening characters displaying backchannels [4].

The same approach cannot work for ACCs: in the case of an eating-related interaction, users will not expect that ACCs consume food or drinks. Let us imagine two humans dining together: they continuously switch their attention between their meal and their interlocutor, so their gaze continuously moves from the plate to the other person, and vice-versa. For obvious reasons, ACCs do not need to gaze at their own plate and they cannot continuously focus their attention on the (human) interlocutor's face (looking at someone eating for too long will make them feel uncomfortable). As a consequence, the gaze behavior of an ACC will have to be carefully designed in order to allow it to seamlessly integrate in eating-related interactions. Thus, we need to design and study other forms of interaction that will be "accepted" by humans while eating, that will not be perceived as inappropriate, unusual, or disturbing during

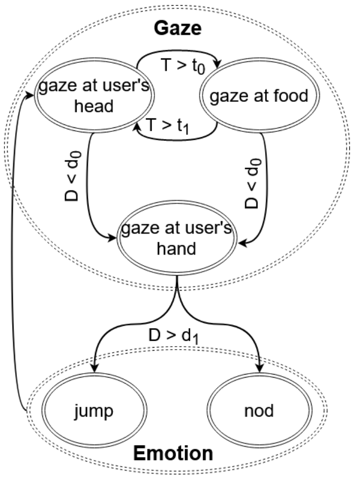

Figure 2: Non-verbal signals state machine. The robot's gaze continuously shifts between the user's head and the food depending on 2 time thresholds t_0 and t_1, unless the distance D between the user's hand and the food is less than threshold d_0: in that case, the robot gazes at the user's hand and, when D becomes higher than threshold d_1, it produces a random emotional feedback.

food consumption. While it might be relatively easy to create a humanoid robot, or a virtual agent, carrying on a conversation with humans about food quality, or providing dietary suggestions (e.g., [17, 21]), it is still unclear what kind of nonverbal behavior ACCs should display during meals [23]. Therefore, in our research we focus on nonverbal behaviors.

Our long-term aims are: 1) to build ACCs that are able to create a meaningful interaction with humans while eating and 2) to measure and replicate with ACCs at least some of the benefits of human-human commensality [24]. For this purpose, our first step will be to measure whether or not there is an impact of ACC technology on humans while eating. So, initially, we aim to investigate two basic research questions: RQ1) can an ACC influence food choice through only its nonverbal behaviors? RQ2) is the presence of an interactive and social ACC preferred over a) eating alone or b) eating with a toy robot not showing any social behavior?

Answering the above questions is a fundamental step before trying to create more complex interactive ACCs. By answering to the first question, whether or not an ACC is able to influence a human's food choices, we can confirm that the ACC nonverbal behavior has a concrete impact on its human co-diner. Regarding the second question, according to our knowledge, while some interesting exploratory work exist [10, 15], there is no previous study evaluating the social acceptance of an artificial companion interacting with a human while eating.

System Description

For the purpose of this work we decided to use a myKeepon, a simple toy robot, which already has been exploited for research purposes [13]. It has 3 degrees of freedom (left/right rotation, front/back and left/right leaning) and it can "jump" up and down. We preferred to use a robot instead of a virtual agent, because in our study we want to explore gaze behavior, which is a particularly important nonverbal behavior in human-human interaction in general [11] and, as a consequence, in commensality settings. Usually the impact of gaze behavior is easier to measure in human-robot interaction than human-virtual agent interaction, because of the robot's physical presence [22]. The gaze of a robot (i.e., "what it is looking at") is easier to predict in the real world, compared to the gaze of a virtual agent. A virtual character can look at the camera (and through the "Mona Lisa effect", all users will perceive the agent as looking at him/her) or not look at the camera, but it is less clear what it is looking at [22].

Despite its simple design, myKeepon can be programmed to mimic proper gaze behavior. At the same time, we expect that by using a toy robot we will reduce user's expectations about what the robot can do and the interaction complexity in general. Consequently, even a limited repertoire of nonverbal behaviors might be acceptable to a human user [26]. Last but not least, thanks to its pet-like design, it may elicit positive attitudes in human participants.

Figure 1 illustrates the system we developed to address our research questions. The user is sitting on a chair and the robot is placed on a dining table together with 2 bowls containing food. The robot produces non-verbal signals to communicate to the user its food preference and emotional feedback (e.g., by producing gaze behavior).

User Tracker

We exploit a Kinect version 2 sensor with a Python wrapper for the Microsoft Kinect SDK to track the user's position. Since the user is sitting in front of the robot, we expect that tracking of the user's head and hands is stable (i.e., there is no occlusion). From the list of body joint names, locations

Gaze signals

GazeAt(x,y) - (x, y) is the target location, in the robot's view space, that the robot has to gaze at. Assuming that a camera is placed in correspondence with the robot's head, the point $(0, 0)$ is the bottom-left corner of the camera field of view, while $(1, 1)$ is the top-right corner. The x coordinate of the target location is reached by rotating the robot on the horizontal plane and by converting rotation values to linear values with the equation:

$$rot_H = \arcsin{(x - 0.5)}$$

The same idea is applied to the vertical rotation:

$$rot_V = \arcsin{(y - 0.5)} * 3.33$$

and confidence score (a value in [0,1] corresponding to the level of confidence the sensor has in respect to that tracked body joint location) in the 3D camera space returned by the wrapper, we pick 3 joints: Head, HandRight, HandLeft.

Signal Planner

We implemented a software module to plan the non-verbal signals to be produced by the robot depending on the user's tracked skeleton. The signal planner is based on two components: Gaze Model and Emotion Model.

Gaze Model. Gaze is a non-verbal cue having several functions: it can show interest, emotional states and contribute to regulate conversation turns [22]. In our system, we focus on the first function, showing interest, that has been described as referring to regions, features and values of interest in the visual field [27]. In particular, the robot can gaze at 3 regions of interest: the bowls containing food, the user's head and the user's hands. The gaze signals are generated by the state machine depicted in the upper part of Figure 2. The goal of the robot's gaze behavior is to establish and maintain gaze contact with the user and to signal to the user that the robot is interested in the food that is contained in the bowls placed on the table.
Emotion Model. Emotional feedback can be successfully expressed by robots, even by those with very limited degrees of freedom, as demonstrated by [3]. It corresponds to the non-verbal signals produced by the robot to communicate its emotional state in response to the non-verbal actions of the user (e.g., picking some food from a bowl). Given the limited amount of available non-verbal signals of myKeepon, we focus on two types of emotion feedback: showing approval and joy. By producing an emotional feedback, the robot aims to "celebrate" and "cheer" the user's choice of eating from one of the bowls (the robot's preferred one, see the last Section of the paper to know how we plan to exploit this type of non-verbal signal). We base this behavior on the theoretical finding about the emotion regulation function of food [9].

Emotion feedback is produced whenever the user picked food that the robot prefers. We can detect this event by checking the distance D between the user's hand and the food bowl: if D becomes less than a threshold d_0 (i.e., the user is picking food) and, subsequently, it becomes higher than another threshold d_1 (the user is bringing food to mouth), then robot chooses a random emotional feedback (see Figure 2, bottom part).

Movement Engine

The movement engine receives the signals to be produced and sends commands to the robot to execute them. The myKeepon toy robot can be program-controlled by following the robot makers' instructions on GitHub[2]. The signals and the corresponding commands that are sent to the robot are described in the margin paragraphs ("Gaze signals" and "Emotion signals").

Future Work

The system we described in the previous section will soon be exploited to test research questions mentioned in the previous sections. We will compare the effects of two versions of the system. In the first version (V1) the robot is a reactive/active dining companion, demonstrating the social behavior described in the previous section. In the second version (V2) the robot is a passive "dining companion", intended to entertain the participant, but it will not demonstrate social behaviors or attempt to influence the participant's food choices. Instead, it will perform pre-defined movement sequences. We hypothesize that the gaze behavior of the robot in V1 will influence the

[2]https://github.com/beatbots/mykeepon

Emotion signals
HeadNod - The robot can produce a head nod by performing two tilt movements in a row (starting from the centered position), the first one of 25 degrees forward, the second one of 25 degrees backward (going back to the neutral starting position).
Jump - The myKeepon robot can "jump" by compressing and extending its body in the vertical direction, a movement called "pon" by the makers of the robot. The jump signal is implemented by a sequence of "pon up" and "pon down" commands (repeated 3 times).

participant's food choice and perception in three measurable ways: (1) the participants will prefer the food in the bowl gazed at by the robot and for which the robot produces cheerful movements when the user eats from it, (2) the food intake from that bowl will higher in the V1 condition compared to the V2 condition, and (3) the user's perception of social presence will be higher in the V1 condition.

To address the research questions mentioned in previous sections we will apply a mix of subjective (post-experiment questionnaires) and objective measures (e.g., food intake). In particular, the questionnaires will evaluate the user experience in terms of food preference, as well as in terms of robot's (social) skills (e.g., [14]), and finally the effect of the robot's behavior on the user's perceived sense of being in a commensality setting.

Discussion and Conclusion

The preliminary work presented in this paper inspires follow-up questions relating to commensality and ACCs. For example, the current implementation uses a toy robot, but it is unknown what is the most appropriate embodiment for an ACC. Does an ACC need to have a human-like appearance or not? Similarly, the question whether a robot or virtual agent provides a more appropriate ACC is relevant. The same setup as the one discussed here could be applied to a virtual agent (e.g., in an augmented reality setup [12]), or a more advanced humanoid robot. The advantage of the interaction model proposed in 2 is that it can be easily implemented on different platforms, permitting cross-platform comparison. Other possible extensions of the current setup regard the interaction with food. For instance, previous work using virtual agents introduced an ACC capable of "eating" virtual food [15]. It is not clear, however, what the human expectations are regarding such behavior. The advantage of using robots is that a robot can interact with actual food and drink; it can pass the food along, or participate in collaborative food preparation. Using this approach, there is great potential for ACC technology to provide health and well-being benefits. For example, in the near future, systems could be implemented that enhance a user's liking of healthy or sustainable food, such as plant- or insect-based food, or systems that could use social interactions with an ACC to make repetitive meals (e.g., in a hospital setting or, in the future, on a space ship traveling to Mars [23]) more palatable.

Importantly, these speculations are currently untested and need to be investigated in-depth before any real benefits of ACCs can be claimed. Indeed, there may also be risks associated with introducing ACCs at the dining table. Like other media, ACCs might be distracting during food consumption, which could contribute to mindless eating and increase unwanted calorie intake [23]. ACCs, especially when poorly designed, could disrupt social interactions that already exists during commensal settings, which could reduce any inherent positive effects in such settings [24].

We see the work presented in this paper as a first step towards ACCs. Our work provides a platform to build commensal interaction scenarios on, provides suggestions on how to test the efficacy of the ACC, and a first example of a computational interaction model focusing on social eating. It is our hope that it will foster further research into the potential benefits and risks of ACCs. Indeed, as they say "the proof of the pudding is in the eating".

ACKNOWLEDGEMENTS

Authors would like to thank F. Rea and D. Pasquali (Istituto Italiano di Tecnologia, Genoa, Italy) for their technical advice on the myKeepon robot.

REFERENCES

[1] 2018a. Number of single-person households in the U.S. from 1960 to 2017 (in millions). Statista - The Statistics Portal. `https://www.statista.com/statistics/242022/number-of-single-person-households-in-the-us/`. (2018). Accessed September 26, 2019.

[2] 2018b. U.S. Census Bureau Releases 2018 Families and Living Arrangements Tables. `https://www.census.gov/newsroom/press-releases/2018/families.html`. (2018). Accessed September 26, 2019.

[3] A. Beck, B. Stevens, K. A. Bard, and L. Cañamero. 2012. Emotional body language displayed by artificial agents. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 2, 1 (2012), 2.

[4] E. Bevacqua, S. Pammi, S. J. Hyniewska, M. Schröder, and C. Pelachaud. 2010. Multimodal Backchannels for Embodied Conversational Agents. In *Intelligent Virtual Agents*, J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud, and A. Safonova (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 194–200.

[5] M. Bruijnes, G. Huisman, and D. Heylen. 2016. Tasty tech: human-food interaction and multimodal interfaces. In *Proceedings of the 1st Workshop on Multi-sensorial Approaches to Human-Food Interaction*. 1–6.

[6] G. Castellano, M. Mancini, C. Peters, and P. W. McOwan. 2011. Expressive copying behavior for social agents: A perceptual analysis. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 42, 3 (2011), 776–783.

[7] Altarriba B. F., S. Jhaveri, R. Lutz, K. Isbister, and D. Wilde. 2019. Making sense of human-food interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[8] H. S. Ferdous, B. Ploderer, H. Davis, F. Vetere, and K. O'hara. 2016. Commensality and the social use of technology during family mealtime. *ACM Transactions on Computer-Human Interaction (TOCHI)* 23, 6 (2016), 1–26.

[9] M. E. Hamburg, C. Finkenauer, and C. Schuengel. 2014. Food for love: the role of food offering in empathic emotion regulation. *Frontiers in psychology* 5 (2014), 32.

[10] R. A. Khot, E. S. Arza, H. Kurra, and Y. Wang. 2019. FoBo: Towards Designing a Robotic Companion for Solo Dining. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems (CHI EA '19)*. ACM, New York, NY, USA, LBW1617:1–LBW1617:6.

[11] C. L. Kleinke. 1986. Gaze and eye contact: a research review. *Psychological bulletin* 100, 1 (1986), 78.

[12] J. Kolkmeier, M. Bruijnes, D. Reidsma, and D. Heylen. 2017. An asap realizer-unity3d bridge for virtual and mixed reality applications. In *International Conference on Intelligent Virtual Agents*. Springer, 227–230.

[13] H. Kozima, M. P. Michalowski, and C. Nakagawa. 2009. Keepon. *International Journal of Social Robotics* 1, 1 (2009), 3–18.

[14] C. U. Krägeloh, J. Bharatharaj, S. K. Sasthan Kutty, P. R. Nirmala, and L. Huang. 2019. Questionnaires to Measure Acceptability of Social Robots: A Critical Review. *Robotics* 8, 4 (2019).

[15] R. Liu and T. Inoue. 2014. Application of an Anthropomorphic Dining Agent to Idea Generation. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct)*. ACM, New York, NY, USA, 607–612.

[16] M. Mancini, B. Biancardi, F. Pecune, G. Varni, Y. Ding, C. Pelachaud, G. Volpe, and A. Camurri. 2017. Implementing and evaluating a laughing virtual character. *ACM Transactions on Internet Technology (TOIT)* 17, 1 (2017), 3.

[17] D. McColl and G. Nejat. 2013. Meal-time with a Socially Assistive Robot and Older Adults at a Long-term Care Facility. *J. Hum.-Robot Interact.* 2, 1 (Feb. 2013), 152–171.

[18] R. Mitchell, A. Papadimitriou, Y. You, and L. Boer. 2015. Really eating together: a kinetic table to synchronise social dining experiences. In *Proceedings of the 6th Augmented Human International Conference*. 173–174.

[19] R. Niewiadomski, E. Ceccaldi, G. Huisman, G. Volpe, and M. Mancini. 2019. Computational Commensality: from theories to computational models for social food preparation and consumption in HCI. *Frontiers in Robotics and AI* 6 (2019), 1–19.

[20] R. Niewiadomski, M. Mancini, G. Varni, G. Volpe, and A. Camurri. 2016. Automated Laughter Detection From Full-Body Movements. *IEEE Transactions on Human-Machine Systems* 46, 1 (2016), 113–123.

[21] N. Randall, S. Joshi, and X. Liu. 2018. Health-e-Eater: Dinnertime Companion Robot and Magic Plate for Improving Eating Habits in Children from Low-Income Families. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. Association for Computing Machinery, New York, NY, USA, 361–362.

[22] K. Ruhland, C. E. Peters, S. Andrist, J. B. Badler, N. I. Badler, M. Gleicher, B. Mutlu, and R. McDonnell. 2015. A review of eye gaze in virtual agents, social robotics and HCI: Behaviour generation, user interaction and perception. In *Computer graphics forum*, Vol. 34. Wiley Online Library, 299–326.

[23] C. Spence, M. Mancini, and G. Huisman. 2019. Digital commensality: Eating and drinking in the company of technology. *Frontiers in psychology* 10 (2019), 2252.

[24] N. Stroebele and J. M. De Castro. 2004. Effect of ambience on food intake and food choice. *Nutrition* 20, 9 (2004), 821–838.

[25] M. Takahashi, H. Tanaka, H. Yamana, and T. Nakajima. 2017. Virtual Co-Eating: Making Solitary Eating Experience More Enjoyable. In *Entertainment Computing – ICEC 2017*, N. Munekata, I. Kunita, and J. Hoshino (Eds.). Springer International Publishing, Cham, 460–464.

[26] K. R. Thórisson and J. Cassell. 1996. Why put an agent in a human body: The importance of communicative feedback in human-humanoid dialogue. *Proceedings of Lifelike Computer Characters' 96* (1996), 44–45.

[27] J. K. Tsotsos, S. M. Culhane, W. Y. K. Wai, Y. Lai, N. Davis, and F. Nuflo. 1995. Modeling visual attention via selective tuning. *Artificial intelligence* 78, 1-2 (1995), 507–545.