

How Avatars in Care Context Should Show Affect

Sascha Fagel

Zoobe message entertainment GmbH
Schönhauser Allee 8
10119 Berlin, Germany
+49 30 28883888e
fagel@zoobe.com

Maher Ben Moussa

University of Geneva
Route de Drize 7
1227 Carouge, Switzerland
+41 22 379 00 73
maher.benmoussa@unige.ch

Donato Cereghetti

Maison de Retraite du Petit-Saconnex
Avenue Trembley 12
1209 Genève, Switzerland
+41 22 730 71 11
donato.cereghetti@mrps.ch

ABSTRACT

This paper presents an analysis on the role of affective expressions in the health care environment and particularly in the care of older adults, and proposes a computational model of affect for ECAs based on Roseman's emotion theory and adapted to the elderly care context. Further, this manuscript also describes how these expressions can be conveyed during spoken interaction.

CCS Concepts

• **Human-centered computing** → **Collaborative and social computing** • **Applied computing** → **Life and medical sciences**

Keywords

Expressive avatars; affective computing

1. INTRODUCTION

This paper focuses on the design and development of ECA (Embodies Conversational Agents) based solution attending to the daily activity and safety needs of elderly users while they go about their normal daily life. Acting as a virtual companion, the ECA should provide implicit support to the elderly, based on behavior and emotional understanding, and should exhibit distinctive human-like emotions, simulating in essence the interaction with a real life partner. Specifically, the virtual companion fuses together the user's facial expressions, voice intonation, gestures and other contextual information of the user's environment and provides appropriate responses.

The paper is structured as follows. Chapter 3 gives general recommendations for the creation of an ECA based on nursing science. Chapter 3 describes different models of emotion with respect to their applicability for an ECA in the care context and recommends replacements of expressions due to inapplicability in the given context. Chapter 4 outlines major properties of the combination of the visual and auditory representations of an ECA.

2. RECOMMENDATIONS FROM THE CARE LITERATURE

According to Orthonoy [1], consistency and coherence are essential requirements to make an emotional ECA believable. Importantly, believability is a context-dependent concept: "a character that is believable in an entertainment application might not be believable in an education or training application" [1]. In AAL (Ambient Assisted Living) systems, ECAs should not act as formal caregivers but rather as informal assistants characterized by a cheerful/friendly attitude [2]. Moreover, we believe that health psychology and nursing science studies should be considered, in addition to cognitive and affective models (see also chapter 3), to design a believable, caring ECA which provides support, assistance and safety to seniors in their daily life.

Consisting of facial expressions, gaze contact, prosody, body gestures, posture, and use of space; nonverbal communication plays an important role in care provision since it "qualifies verbal messages, regulates interaction, and communicates attitudes such as liking, warmth, supportiveness, agreement, and interest" [3]. Indeed, specific nonverbal patterns have been related to patient satisfaction [4, 5, 6], therapeutic efficacy [7, 8], and malpractice claim history [9]. For instance, patients' satisfaction is higher when physicians smile, nod, forward lean, have a patient-directed gaze, and have facial and vocal expressiveness [4, 5]. Physicians' patient-directed gaze has been associated with patients' share of their health problems [8]; while physicians' distancing behaviors (lack of patient-directed gaze and lack of smiling) has been linked with decreases in patients' cognitive and physical functioning [7]. Caris-Verhallen et al. [10] studied the relations between verbal and nonverbal communication in the nurse-senior relationship. They found that social communication (personal statements, banter, jokes and small talks) is associated with patient-directed gaze combined with smiling, while affective communication (verbal attentiveness, concern, empathy, and sympathy) is associated with patient-directed gaze combined with affirmative head nodding. According to Haskard et al. [6], affects are mostly communicated by nurses by vocal expressiveness, while instrumental content is stronger transmitted through visual nonverbal behavior (body movement, posture, interactional distance, and not-close facial expressions).

The following paragraphs discuss general guidelines that should be used by designers in order to model the nonverbal behavior of a believable ECA in AAL solutions:

(1) The ECA should maintain eye contact with the senior in order to express concern, empathy and to convey to the user the sense of being listened to [10, 11]. User-directed gaze can easily be ensured by the Mona Lisa Effect: "if the avatar is gazing towards the camera, the eyes seem to "follow" the beholder whatever vantage point he or she may assume" [12]. In wall-

mounted ECA solutions, vertical difference in eye level between the ECA and the user should be minimized in order to ensure eye contact between the agent and the user [see also 11]. The Gerontological Society of America (GSA) [13] also highlighted the importance of the symmetric eye contact between caregivers and seniors: “face older adults when you speak with them, with your lips at the same level as theirs”.

(2) The ECA should appropriately display affection, amiability and empathy to the user. Smiling, nodding and forward leaning should be designed in the nonverbal behavior of the ECA to convey warmth, friendliness, interest and the sense of being cared for [10].

(3) Similarly to formal caregivers [14], the ECA should be able to self-regulate and control the expression of his negative emotional states like anger, frustration, irritation or disgust.

(4) Facial events and acoustic cues should express the desire of the ECA to support and care. Validated tables matching facial events [15] and acoustic cues [16] with emotional events should be used to design a believable ECA emotional behavior. Note that combinations of action units (e.g. anger: AU4+AU5+AU7, disgust: AU9+AU10) and acoustic cues (e.g. irregularities in frequency, intensity and duration) associated to negatives emotions should be avoided or used sparingly.

(5) Importantly, in AAL systems, seniors should be able to easily modulate the ECA speech volume and speech rate according to their needs [17].

3. BUILDING EMOTIONAL BEHAVIOUR

3.1 Emotion Elicitation Theories

Affective science research on the elicitation of emotions distinguishes between several categories of theories where each category describes a different type of process for processing external or internal stimuli and eliciting emotions as a reaction. The most prominent of these theories are: (1) the basic emotion theories, which advocate the existence of a number of universal and prototypical emotions that are evolutionarily hard-wired in humans. (e.g. Ekman’s [18], Izard’s [19]); (2) the constructivist theories, which see emotions as subjective evaluation of the physiological signals (e.g. Russell’s [20], Barrett’s [21]); and (3) the appraisal theories, which focus on the evaluative process of emotions, where stimuli are evaluated against several appraisal variables (e.g. Scherer’s [22, 23], Roseman’s [24, 25, 26], OCC model [1, 27]). A more comprehensive comparison between the models can be found in [22] and [28].

For the simulation of realistic context-related affective behavior in ECAs, contrary to basic and constructivist theories, appraisal theories seem to be the most popular amount researchers and the most logical one as they define to a certain level the process of eliciting emotions based on an evaluation of the agent’s external and internal state against the so-called appraisal variables. For example, the emotion of joy would occur when a desirable event happens, while anger might happen when an undesirable event happens and someone is to blame for this event. The appraisal process and the employed appraisal variables may differ from one appraisal theory to another. Table 1 compares the five most well-known emotion appraisal models showing the similarities between them and the appraisal components they rely on to describe the human emotion appraisal process. In this table, we categorize the appraisal variables into *novelty*, *valence*, *goal relevance*, *goal consistency*, *certainty*, *agency*, *control*, and *value relevance*. Most appraisal models cover these appraisal variables in one way or

another where they might use the same variable or use more variables to describe the same affective component. However, the OCC (Ortony, Clore, & Collins) model, which is widely used in Computer Science, takes a different approach and does not cover goal relevance and consistency directly, but focuses more on distinguishing between event and action appraisal. An event can be either pleasant or unpleasant to self and deserved or undeserved to others. An action can be either approved or disapproved by the agent.

Table 1 Comparison of the emotion models (partly based on [29])

	<i>Frijda</i>	<i>Roseman</i>	<i>Scherer</i>	<i>Smith/Elsworth</i>	<i>OCC</i>
Novelty	Change Familiarity	Unexpectedness	Suddenness Familiarity Predictability	Attentional activity	
Valence	Valence		Intrinsic pleasantness	Pleasantness	Appealingness Desirability Praiseworthiness
Goals Relevance	Focality	Appetitive / Aversive motives	Concern relevance	Importance	
Goal Consistency	Open / closed	Motive consistency	Conduciveness	Perceived obstacle / Anticipated effort	Deservingness
Certainty	Certainty	Certainty	Outcome probability	Certainty	Likelihood
Agency	intent / Self-other	Agency	Cause: agent Cause: motive	Human agency	Self-agent / Other-agent
Control	Modifiability Controllability	Control potential	Control Power Adjustment	Situational Control	
Value relevance	Value relevance		Internal / external compatibility standards	Legitimacy	

During the last decades, different emotion appraisal theories have been employed in the development of many computational models of emotions. Table 2 provides a comparison between a selection of these implementations. As can be noted, the OCC model is still the most popular appraisal model among computer scientists [30 – 40]. The simplicity and the well-defined structure of this model, makes it accessible to computer scientists as its properties can easily be computationally modelled in any programming

Table 2 Comparison between Computational Models of Emotions

Model	Appraisal theory
ALMA (Gebhard 2005)	OCC
Cathexis (Velásquez 1996)	Roseman
Coppélius’ Concoction (Hoorn, Pontier, and Siddiqui 2012)	Smith & Lazarus
Dastani’s model (Steunebrink, Dastani, and Meyer 2012)	OCC
EMA (Marsella and Gratch 2009)	Smith & Lazarus
Émile (Gratch 2000)	OCC
EVA (Kasap et al. 2009)	OCC
FAtiMA (Dias, Mascarenhas, and Paiva 2011)(Dias and Paiva 2005)	OCC
FLAME (El-Nasr, Yen, and Ioerger 2000)	OCC
GALAAD (Adam 2007)	OCC
GAMYGDALA (Popescu, Broekens, and van Someren 2013)	OCC
GRACE (Dang, Letellier-Zarshenas, and Duhaut 2008)	Scherer
Greta (de Rosi et al. 2003)	OCC
iGrace (Saint-Aim, Le-Pvdic, and Duhaut 2009)	Scherer
MAMID (Hudlicka 2008)	Lazarus
ParIE (Bui et al. 2002)	OCC
Soar-Emote (Marinier, Laird, and Lewis 2009)	Scherer
WASABI (Becker-Asano 2008)	Scherer

language. However, to many psychologists, the OCC model is considered to be outdated. This unpopularity of the OCC model in psychology has pressured several computer scientists to employ other models of emotion appraisal. Scherer’s model has been used by GRACE [41], iGrace [42], Soar-Emote [43], WASABI [44]; Roseman’s model has been used by Cathexis [45]; and Lazarus’s model has been by Coppélius’ Concoction [46], EMA [47], MAMID [48].

3.2 Our Implementation

In our approach, we employ Roseman’s model because of its computational simplicity and its acceptance in the affective science community. At each step of the interaction between the user and the ECA, we employ the following Roseman’s appraisal variables to determine the ECA’s emotional state

- *Appetitive/Aversive motives* – If the goal/motive is appetitive or aversive
- *Motive consistency* – If the event achieves or violates the goal/motive
- *Certainty* – If an expected event is certain to happen or not
- *Agency* – If the event is caused by the agent or by the user
- *Control potential* – If the agent has the potential to control the event

Roseman’s model defines appraisal schemes for 17 emotions. However, as pointed out in section 2, it is not advisable to let the ECA express negative emotions in the care context. Humans possess the ability to self-regulate their emotions based on current goals and context and in the care context the goal is to keep the elderly or patient in a positive emotional state. Emotional contagion should be avoided as negative emotions can easily be transferred unconsciously to the user [49]. We defined in Table 3 the Roseman appraisal schemes we use to determine the ECA’s internal emotion and the emotions we choose the avatar to express as a form of self-regulation in the care context. According to Roseman’s appraisal theory, a perceived event is evaluated based on the goal it achieves or violates, the type of the goal is important (Appetitive or Aversive), the certainty of the event and if the agent can cope with the results of the event. For events that violate the agent’s goals, we have chosen to express the emotion “compassion” instead of the negative internal emotions of sadness, distress, frustration and disgust. Compassion can be defined as “the emotion one experiences when feeling concern for another’s suffering and desiring to enhance that person’s welfare” [50]. This emotion is the nurse’s most precious asset [51] and research has shown that the lack of compassion in the care can affect the well-being and the satisfaction of the patient [52]. The ECA should also have this caring attitude, expressing compassion when the user is suffering and making the user feel ‘cared about’ and ‘cared for’ [53]. However, for negative events appraised as being ‘uncertain’ which elicit negative fear, we have chosen to keep the worried expression of the ECA for motivational reasons Table 3 Appraisal of event related emotions showing the user that the ECA is worried if she or he does not want to follow a healthy life style. Further, the emotion “hope” has been replaced by

<i>Motive Consistency</i>	<i>Certainty</i>	<i>Control Potential</i>	<i>Motive Type</i>	<i>Internal Emotion</i>	<i>Expressed behaviour</i>
Achieves	Uncertain	Any	Any	Hope	Joy
Achieves	Certain	Any	Appetitive	Joy	Joy
Achieves	Certain	Any	Aversive	Relief	Relief
Violates	Uncertain	Low	Any	Fear	Fear(Worry)
Violates	Certain	Low	Appetitive	Sadness	Compassion
Violates	Certain	Low	Aversive	Distress	Compassion
Violates	Any	High	Appetitive	Frustration	Compassion
Violates	Any	High	Aversive	Disgust	Compassion

“happy” as “hope” does not have any prototypical facial expression and should cover a certain positive mood.

In Table 4 we have defined the appraisal schemes for events attributed to the user together with the regulated emotional expression of each scheme. The negative emotions of dislike, anger and contempt where an agent blames the user for a certain event are replaced by the directive behaviour. In a normal context, these blame emotions mainly serve to change or stop a situation caused by someone. In the care context, literature has shown that directive behaviour has achieved better result in changing the elderly’s behaviour in several situations [54]. We propose to design a “directive attitude” in the use cases in which the ECA motivates the non-compliant users to adopt a healthy behaviour (i.e., perform physical activity, taking medications).

Table 4 Appraisal of agency related emotions

<i>Motive Consistency</i>	<i>Certainty</i>	<i>Control Potential</i>	<i>Motive Type</i>	<i>Emotion</i>	<i>Expressed behaviour</i>
Achieves	Any	Any	Any	Love	Joy
Violates	Any	Low	Any	Dislike	Directive Behaviour
Violates	Any	High	Appetitive	Anger	Directive Behaviour
Violates	Any	High	Aversive	Contempt	Directive Behaviour

4. HOW TO CONVEY EXPRESSIONS

Spoken language is a highly developed and very efficient mean for the transmission of information between humans. Spoken language uses multiple modalities namely audition and vision to achieve high information density as well as robustness. Both modalities carry a vast amount of information along with the linguistic content when humans are communicating i.e., identity of the speaker, emotional content, personality, regional and social origin, and so on. Both modalities contribute in different ways to these properties and the differences should be taken into account when designing an affective ECA. The following technological constraints and audio-visual interactions have to be taken into account to create an efficient expressive avatar and to avoid a loss of quality and effectiveness.

4.1 Technology

There are two technical principles for generating an ECA that can ensure audio-visual coherence. First, as spoken communication is a physiological process that becomes visible and audible, rebuild the physical processes of communication with 3-D rendering and acoustic simulation (known as articulatory speech synthesis [55]) keeps all spacio-temporal linking between audio and video intact. However, this method is extremely complex and does not guarantee high-quality ECAs as an outcome. Secondly, segmenting and recombining recorded videos of a person to new communication sequences (known as audio-visual unit-selection) also keeps the audio-visual coherence but requires a huge amount of audio-visual data and their annotation to cover comprehensive speech capabilities and expressivities.

The usual way to create an expressive avatar is to generate the visual representation and the audible speech separately while using the best possible method quality-wise in each modality and to combine both modalities. Animated 3-D models allow for layering of facial animation mainly to convey emotions, body animation to include speech-accompanying, deictic and other gestures, and movements for lip-sync speech. Image-based ECAs are mainly used for video-realistic avatars and to represent real persons (see below) but require specific image databases for each expression.

Although there are several different approaches (see [56] for a review), all current high-quality audio speech synthesis systems use a unit-selection [57] or similar approach where comprehensive speech recordings (many systems require at least 4 hours of speech material) are annotated and re-combined to create new utterances. The larger the chunks are that can be taken from the original recording, the better the output quality will be. If the domain of the application is known (e.g. health or elderly care) special utterances can be included in the speech database during recording to improve the quality. Unless for each targeted vocal expression a separate speech database is created [58], the overall variability of this technology is limited to varying the process of selecting the units and post-processing of the speech output as the speech output is generated from previous recordings. Hence, spectral and temporal properties can only be slightly modified and a specific intonation cannot be generated.

4.2 Speech

There is a strong link between speech production/synthesis and speech perception. If the audio and video representations match, the joint output is efficient and robust. In case of a mismatch between audio and video representation, a percept can be invoked that matches neither the audio nor the video representation (so-called McGurk effect [59]) leading to reduced intelligibility. Hence, synthesized multimodal speech of a newly created ECA should be assessed by human perceivers.

4.3 Age

Human perception in communication is susceptible to different properties of varying degrees. A person's age can be determined from speech [60] as well as from face [61]. If the age of the communicating party differs in audio and video representation, a weighted average of the ages perceived in audio and video separately is invoked with a higher impact of the visual modality [62]. Hence, the perceived age of an ECA is dominated by the visual representation and disturbed in case of severe mismatch with the audible speech.

4.4 Identity

The voice of a speaker can be used for identification by human perceivers as well as by machines [63]. The same is the case for the face [64] and can also be assumed for synthetically generated voices and faces. When combined, the face overrules the voice in case of limited mismatch. In a study of Morandell et al. [65] elderly users identified speaking photos of relatives although the voice was generated by a standard audio speech synthesizer without the impression of a mismatch between face and voice. Using a model of a real person as avatar in human-computer user interaction can lead to confusion and severe problems if the real person also interacts with the user without being fully aware of all details of the user's conversation with his/her ECA.

4.5 Emotion

Like with the linguistic content in audiovisual speech, the emotional content is significantly affected if the expressions in the audio and the visual modalities differ. An experiment by Fagel [66] discovered that when the video of an utterance spoken in one emotion is dubbed with the audio of the utterance spoken in another emotion the perceived emotion might be a third – neither present in the auditory nor in the visual modality. Hence, if the audio information with respect to the emotional content does not match the emotional content of the video, the recognition of the expression intended by the face can fail. The aspects of expressions conveyed by the two modalities differ where the face transports the positive/negative valence of the expression to a

great extent and the voice transports the level of arousal. Therefore, a 3-D animated avatar with different facial expressions added to the same mechanism of speech animation along with a unit-selection audio speech synthesis with modifications to match the level of arousal will be likely to convey the desired expression effectively.

5. CONCLUSIONS

In the chapters above we provided recommendations for the creation of an affective embodied conversational agent in a care context. These recommendations are grounded on accepted research. We proposed to use Roseman's model of emotion appraisal and map its emotions to the expressions Joy, Relief, Worry, Compassion, and Directive Behaviour. When creating or selecting the ECAs face and voice, we recommended to evaluate the intelligibility of the speech output, to use voice and face modules with minor mismatch between audible and visible age, and to consider using a fictional person as avatar. Furthermore, we recommended to display appropriate facial expressions in combination with audio speech with an audible level of arousal that matches the intended expressions.

6. ACKNOWLEDGMENTS

This work was funded by the European Commission through the FP7-ICT-Project Miraculous Life (grant agreement no. 611421).

7. REFERENCES

- [1] Ortony, A. 2003. On making believable emotional agents believable. In Trapp R. et al. (Ed.), *Emotions in Humans and Artifacts* (pp. 189--212). Cambridge, Massachusetts London, England: MIT Press
- [2] Cereghetti, D. M., Kleanthous, S., Christophorou, C., Tsiourti, C., Wings, C., Christodoulou, E. 2015. *Virtual Partners for Seniors: Analysis of the Users' Preferences and Expectations on Personality and Appearance*.
- [3] Stepanikova, I., Zhang, Q., Wieland, D., Releaser, G.P., & Stewart, T. 2012. Nonverbal communication between primary care physicians and older patients: How does race matter? *J Gen Int Med*, 27, 576-581.
- [4] Hall, J.A., Harrigan, J.A., Rosenthal, R. 1995. Nonverbal behavior in clinician-patient interaction. *Appl Prev Psychol*, 4, 21-37.
- [5] Griffith, C.H., Wilson, J.F., Langer, S., & Haist, S.A. 2003. House staff nonverbal communication skills and standardized patient satisfaction. *J Gen Intern Med*, 18, 170-4.
- [6] Haskard, K.B., DiMatteo, M.R., & Heritage, J. 2009. Affective and instrumental communication in primary care interactions: predicting the satisfaction of nursing staff and patients. *Health Communication*, 24(1), 21-32.
- [7] Ambady, N., Koo, J., Rosenthal, R., & Winograd, C.H. 2002. Physical therapists' nonverbal communication predicts geriatric patients' health outcomes. *Psychol Aging*, 17, 443-52.
- [8] Bensing J. M., Kerssens J. J., & van der Pasch M. 1995. Patient-directed gaze as a tool for discovering and handling psychosocial problems in general practice. *Journal of Nonverbal Behavior*, 19(4), 223-242.
- [9] Ambady, N., LaPlante, D., Nguyen, T., Rosenthal, Chaumeton, N., & Levinson, W. 2002. Surgeons' tone of voice: a clue to malpractice history. *Surgery*, 132, 5-9.

- [10] Caris-Verhallen, W. M. C. M., Kerkstra, A. & Bensing, J. M. 1999. Non-verbal behaviour in nurse-elderly patient communication. *Journal of Advanced Nursing*, 29, 808–818.
- [11] Gorawara-Bhat R., & Cook, M.A. 2011. Eye contact in patient-centered communication. *Patient Education and Counseling*, 82, 442–7.
- [12] Beskow, J., & Al Moubayed, S. 2010. Perception of gaze direction in 2D and 3D facial projections. *Proceedings of the ACM/SSPNET 2nd International Symposium on Facial Analysis and Animation*, pp. 24–24.
- [13] The Gerontological Society of America 2012. *Communicating With Older Adults: An Evidence-Based Review of What Really Works*.
- [14] Sellman, D., & Snelling, P. 2009. *Becoming a Nurse: A Textbook for Professional Practice*. Harlow, UK: Pearson Education.
- [15] Ekman, P., Irwin, W., & Rosenberg, E. L. 1994. *The emotional facial action coding system (EMFACS)*. London, UK
- [16] Juslin, P. N, & Laukka, P 2003. Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code? *Psychological Bulletin*, 129(8/5), 770–814.
- [17] Christophorou, C., Georgiadis, D., Andreou, P., Kleanthous, S., Cereghetti, D. M., Meijers, J., Christodoulou, E., & Samaras, G. 2016. *ICT Systems and Services: Identification and Assessment of a Package of ICT Services for Ageing Well*. *Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON 2016)*, Paphos, Cyprus.
- [18] Ekman, Paul. 2009. “Darwin’s Contributions to Our Understanding of Emotional Expressions.” *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364(1535): 3449–51.
- [19] Izard, Carroll E. 2007. “Basic Emotions, Natural Kinds, Emotion Schemas, and a New Paradigm.” *Perspectives on Psychological Science* 2(3): 260–80.
- [20] Russell, James a. 2009. “Emotion, Core Affect, and Psychological Construction.” *Cognition & Emotion* 23(7): 1259–83.
- [21] Barrett, Lisa Feldman, Batja Mesquita, Kevin N. Ochsner, and James J. Gross. 2007. “The Experience of Emotion.” *Annual Review of Psychology* 58(1): 373–403.
- [22] Scherer, Klaus. 2009. “Emotions Are Emergent Processes: They Require a Dynamic Computational Architecture.” *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364(1535): 3459–74.
- [23] Ellsworth, P. C., & Scherer, K. R. (2003). 2003. “Appraisal Processes in Emotion.” In *Handbook of Affective Sciences*,.
- [24] Roseman, Ira J. 2013. “Appraisal in the Emotion System: Coherence in Strategies for Coping.” *Emotion Review* 5(2): 141–49.
- [25] Roseman, Ira J. 2011. “Emotional Behaviors, Emotivational Goals, Emotion Strategies: Multiple Levels of Organization Integrate Variable and Consistent Responses.” *Emotion Review* 3(4): 434–43.
- [26] Roseman, Ira J., Cynthia Wiest, and Tamara S. Swartz. 1994. “Phenomenology, Behaviors, and Goals Differentiate Discrete Emotions.” *Journal of Personality and Social Psychology* 67(2): 206–21.
- [27] Ortony, Andrew, Gerald L. Clore, and Allan Collins. 1994. *54 The Cognitive Structure of Emotions*. New York, NY, US: Cambridge University Press.
- [28] Moors, Agnes. 2009. “Theories of Emotion Causation: A Review.” *Cognition & Emotion* 23(4): 625–62.
- [29] Scherer, Klaus R. 1999. “Appraisal Theory.” In *Handbook of Cognition and Emotion*, eds. T Dalglish and M Power. Chichester, U.K.: Wiley, 637–63.
- [30] Gebhard, Patrick. 2005. “ALMA: A Layered Model of Affect.” In *International Joint Conference on Autonomous Agents and Multiagent Systems - AAMAS ’05*, New York, NY, USA: ACM Press, 29–36.
- [31] Steunebrink, Bas R., Mehdi Dastani, and John-Jules Ch. Meyer. 2012. “A Formal Model of Emotion Triggers: An Approach for BDI Agents.” *Synthese* 185(S1): 83–129.
- [32] Gratch, Jonathan. 2000. “Émile: Marshalling Passions in Training and Education.” In *International Conference on Autonomous Agents - AGENTS ’00*, New York, New York, USA: ACM Press, 325–32.
- [33] Kasap, Zerrin, Maher Ben Moussa, Parag Chaudhuri, and Nadia Magnenat-Thalmann. 2009. “Making Them Remember—Emotional Virtual Characters with Memory.” *IEEE Computer Graphics and Applications* 29(2): 20–29.
- [34] Dias, J, Samuel Mascarenhas, and Ana Paiva. 2011. “Fatima Modular: Towards an Agent Architecture with a Generic Appraisal Framework.” In *International Workshop on Standards for Emotion Modeling 2011*, , 1–8.
- [35] Dias, João, and Ana Paiva. 2005. “Feeling and Reasoning: A Computational Model for Emotional Characters.” In *Portuguese Conference on Artificial Intelligence, EPIA 2005, Lecture Notes in Computer Science*, eds. Carlos Bento, Amílcar Cardoso, and Gaél Dias. Covilhã, Portugal: Springer Berlin Heidelberg, 127–40.
- [36] El-Nasr, Magy Seif, John Yen, and Thomas. R. Ioerger. 2000. “FLAME — Fuzzy Logic Adaptive Model of Emotions.” In *Autonomous Agents and Multi-Agent Systems (AAMAS 2000)*, Berlin / Heidelberg, Germany: Springer-Verlag, 219–57.
- [37] Adam, Carole. 2007. “Emotions: From Psychological Theories to Logical Formalization and Implementation in a BDI Agent.” *INP de Toulouse*.
- [38] Popescu, Alexandru, Joost Broekens, and Maarten van Someren. 2013. “GAMYGDALA: An Emotion Engine for Games.” *IEEE Transactions on Affective Computing X*: 1–1.
- [39] de Rosis, Fiorella, Catherine Pelachaud, Isabella Poggi, Valeria Carofiglio, and Berardina De Carolis. 2003. “From Greta’s Mind to Her Face: Modelling the Dynamics of Affective States in a Conversational Embodied Agent.” *International Journal of Human-Computer Studies* 59(1-2): 81–118.
- [40] Bui, The Duy, Dirk Heylen, Mannes Poel, and Anton Nijholt. 2002. “ParLE: An Adaptive Plan Based Event Appraisal Model of Emotions.” In *Annual German Conference on AI - KI 2005, Lecture Notes in Computer Science*, eds. Matthias Jarke, Gerhard Lakemeyer, and Jana Koehler. Berlin / Heidelberg, Germany: Springer-Verlag, 129–43.

- [41] Dang, Thi-Hai-Ha, Sabine Letellier-Zarshenas, and Dominique Duhaut. 2008. "Grace – Generic Robotic Architecture To Create Emotions." In *International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines*, Singapore: World Scientific Publishing Co. Pte Ltd, 174–81.
- [42] Saint-Aim, Sbastien, Brigitte Le-Pvdic, and Dominique Duhaut. 2009. "iGrace – Emotional Computational Model for EmI Companion Robot." In *Advances in Human-Robot Interaction*, ed. Vladimir A. Kulyukin. InTech.
- [43] Marinier, Robert P., John E. Laird, and Richard L. Lewis. 2009. "A Computational Unification of Cognitive Behavior and Emotion." *Cognitive Systems Research* 10(1): 48–69.
- [44] Becker-Asano, Christian. 2008. "WASABI: Affect Simulation for Agents with Believable Interactivity." University of Bielefeld.
- [45] Velásquez, Juan David. 1996. "Cathexis : A Computational Model for the Generation of Emotions and Their Influence in the Behavior of Autonomous Agents by." Massachusetts Institute of Technology.
- [46] Hoorn, Johan F., Matthijs Pontier, and Ghazanfar F. Siddiqui. 2012. "Coppélius ' Concoction : Similarity and Complementarity Among Three Affect-Related Agent Models." *Cognitive Systems Research* 15-16: 33–49.
- [47] Marsella, Stacy C., and Jonathan Gratch. 2009. "EMA: A Process Model of Appraisal Dynamics." *Cognitive Systems Research* 10(1): 70–90.
- [48] Hudlicka, Eva. 2008. "Modeling the Mechanisms of Emotion Effects on Cognition." In *AAAI Fall Symposium on Biologically Inspired Cognitive Architectures 2008*, Menlo Park, CA: AAAI Press, 82–86.
- [49] Parkinson, B., & Simons, G. (2009). Affecting others: social appraisal and emotion contagion in everyday decision making. *Personality and Social Psychology Bulletin*, 35(8), 1071–1084. doi:10.1177/0146167209336611
- [50] Leiber, S., Klimecki, O., & Singer, T. (2011). Short-term compassion training increases prosocial behavior in a newly developed prosocial game. *PLoS One*, 6.
- [51] Schantz, M. (2007). Compassion: a concept analysis. *Nursing Forum*, 42(2), 48-55.
- [52] Halifax, J. (2014). G.R.A.C.E. for nurses: Cultivating compassion in nurse/patient interactions. *Journal of Nursing Education & Practice*, 4, 121–128.
- [53] Henderson A, Van Eps MA, Pearson K, James C, Henderson P. 2007. 'Caring for' behaviours that indicate to patients that nurses 'care about' them. *Journal of Advanced Nursing*, 60(2), 146-153.
- [54] Christenson, A. M., Buchanan, J. A., Houlihan, D., & Wanzek, M. 2011. Command use and compliance in staff communication with elderly residents of long- term care facilities. *Behavior Therapy*, 42, 47-58.
- [55] Birkholz, P., Jackel, D. 2003. A three-dimensional model of the vocal tract for speech synthesis. *Proc. ICPhS*, 2597–2600.
- [56] Balyan, A., Agrawal, S.S., Dev, A. 2013. Speech Synthesis: A Review. *International Journal of Engineering Research & Technology (IJERT)* ISSN: 2278-0181, Vol. 2 Issue 6, 57-75.
- [57] Hunt, A.J., Black, A.W. 1996. Unit selection in a concatenative speech synthesis system using a large speech database. *Proc. ICASSP*, 373-376.
- [58] Schröder, M., Burkhardt, F. and Krstulovic, S. 2010. Synthesis of emotional speech. In Scherer, K. R., Bänziger, T. and Roesch, E. (Eds.). *Blueprint for Affective Computing*, pp. 222-231. Oxford, UK: Oxford University Press.
- [59] McGurk, H., MacDonald, J. 1976. Hearing lips and seeing voices. *Nature* 264, 23-30.
- [60] Winkler, R., Brückl, M., Sendlmeier, W. 2003. The aging voice: an acoustic, electroglottographic and perceptive analysis of male and female voices. *Proc. of ICPhS*, 2869-2872.
- [61] Burt, D.M., Perrett, D.I. 1995. Perception of Age in Adult Caucasian Male Faces: Computer Graphic Manipulation of Shape and Colour Information. *Proc. Royal Society London* 259, 137-143.
- [62] Fagel, S. 2007. Auditory-visual integration in the perception of age in speech. *Proc. ICPhS*, 725-728.
- [63] Steeneken, H.J.M., Van Leeuwen, D. 1997. Speaker Recognition by Humans and Machines. *Proc. of Eurospeech*, 2319-2322.
- [64] Chellappa, R., Wilson, C.L., Sirohey, S. 1995. Human and machine recognition of faces: A survey. *Proc IEEE* 83(5), 705-741.
- [65] Morandell, M., Hochgatterer, A., Fagel, S., Wassertheurer, S. 2008. Avatars in assistive homes for the elderly. In: Holzinger, A. (ed.) *USAB 2008*. LNCS 5298, 391–402.
- [66] Fagel, S. 2006. Emotional McGurk effect. *Proc. Speech Prosody*, 229-232.