

Coordinating the Generation of Signs in Multiple Modalities in an Affective Agent

Jean-Claude Martin, Laurence Devillers, Amaryllis Raouzaïou, George Caridakis, Zsófia Ruttkay, Catherine Pelachaud, Maurizio Mancini, Radek Niewiadomski, Hannes Pirker, Brigitte Krenn, Isabella Poggi, Emanuela Magno Caldognetto, Federica Cavicchio, Giorgio Merola, Alejandra García Rojas, Frédéric Vexo, Daniel Thalmann, Arjan Egges, and Nadia Magnenat-Thalmann

Abstract In order to be believable, embodied conversational agents (ECAs) must show expression of emotions in a consistent and natural looking way across modalities. The ECA has to be able to display coordinated signs of emotion during realistic emotional behaviour. Such a capability requires one to study and represent emotions and coordination of modalities during non-basic realistic human behaviour, to define languages for representing such behaviours to be displayed by the ECA, to have access to mono-modal representations such as gesture repositories. This chapter is concerned about coordinating the generation of signs in multiple modalities in such an affective agent. Designers of an affective agent need to know how it should coordinate its facial expression, speech, gestures and other modalities in view of showing emotion. This synchronisation of modalities is a main feature of emotions.

1 Introduction

As explained in [Sect. 6.1.1.1](#), an embodied conversational agent (ECA) is a multimodal interface in which an animated character displayed on a computer screen converses with the user with several human-like modalities such as speech, gesture and facial expressions (Cassell et al., 1999; Kopp et al., 2003). Using an affective ECA (affective agent for short) in an interface is expected to lead to an intuitive and friendly human–computer interaction, for example, via the display of emotional expressions that can be useful in games or education as a means of motivating students.

In order to be believable, ECAs must show expression of emotions in a consistent and natural looking way across modalities. The ECA has to be able to display

J.-C. Martin (✉)

Computer Sciences Laboratory for Mechanics and Engineering Sciences (LIMSI), Paris, France
e-mail: Martin@limsi.fr

coordinated signs of emotion during realistic emotional behaviour. Such a capability requires one to study and represent emotions and coordination of modalities during non-basic realistic human behaviour, to define languages for representing such behaviours to be displayed by the ECA, to have access to mono-modal representations such as gesture repositories.

This chapter is concerned about coordinating the generation of signs in multiple modalities in such an affective agent. Designers of an affective agent need to know how it should coordinate its facial expression, speech, gestures and other modalities in view of showing emotion. This synchronisation of modalities is a main feature of emotions (Scherer and Ellring, 2007; Scherer, 2000). Yet, generating multimodal expressions of emotions raises several questions to designers such as What are the key signs required for the expression of emotions in individual modalities? What are the different ways to combine and synchronise these signs across modalities? This also raises methodological questions such as How to be sure that the signs displayed by multiple modalities will be perceived and understood sign by sign, as well as a whole multimodal expression? How to collect knowledge on the coordination of signs in a specific situation? How can multimodal emotional corpora be useful and how to collect them? How to validate models of multimodal expressions of emotions?

This chapter has connections with the following other chapters of the handbook. In relation to Chapter “Emotion: concepts and definitions”, we are concerned with the multimodal expression of emergent emotions that are quite relevant for the design of affective agents and that might involve intensity, rapidity of change, brevity, event focus, appraisal elicitation and synchronisation. We will describe not only the display of basic emotions but also the display of several simultaneous emotions that we will call blends of emotions (cf. below). In relation to Chapter “Emotions in Social Interactions: Unfolding Emotional Experience”, this chapter also considers the perception of emotional expression and its impact on interaction.

While Chapter “Image and Video Processing for Affective Applications” focuses on the recognition of affective states from a user, for example, while interacting with an embodied agent or not), this chapter deals with the expression and generation of coordinated signs of affects displayed by an embodied agent. Thus, when considering video corpora or vision processing techniques in the current chapter, we do it only for the sake of collecting knowledge on how the ECA should be specified. We do not consider image processing of user’s behaviour.

As we will explain below, databases of multimodal emotional behaviours (Part III “Data and Databases”) can be useful for the design of affective agents. While the Part VII “Usability” is concerned with the design and evaluation of affective systems without a specific focus on affective agents, this chapter is of interest for designers of multimodal agents and provides detailed explanations on how the agent needs to coordinate the modalities. We do not consider how a user would perceive these signs when using a system as the field is not yet mature enough. Rather we focus on perceptual experiments studying how subjects perceive different signs and different ways to combine them. Such experiments are required for validating models of multimodal coordination.

Finally, as to the Part IV “Emotion in Interaction” itself, we explain how designers of affective agents can collect knowledge on how to coordinate several modalities during the display of affective messages. While Chapter “Fundamentals of Agent perception and Attention modelling” deals with perception, this chapter focuses on generation. While Chapter “Generating Listening Behaviour” deals with specific communicative functions related to listening behaviour, this chapter focuses on the expression of emotions. Chapter “Representing Emotions and Related States in Technological Systems” deals with internal representations of emotions; this chapter deals with their multimodal expressions. Finally, this chapter can be seen as input for Chapter “Embodied Conversational Characters: Representation Formats for Multimodal Communicative Behaviours” on behaviour representation languages, since here we will describe the approaches and methodology to collect knowledge on how to coordinate signs, which can then be described formally using behaviour representation languages.

2 Human Multimodal Communication

In this section, we provide some background information about multimodal communication by humans that needs to be considered when designing affective agents.

Human communicative modalities include speech, hand gestures (and other bodily movements), facial expressions, gaze (and pupil dilation), posture, spatial behaviour, bodily contact, clothes (and other aspects of appearance), non-verbal vocalisations, smell (Argyle, 2004). There is a wide range of literature on non-verbal communication (Argyle, 2004; Feldman and Rim, 1991; Knapp and Hall, 2006, Siegman and Feldstein, 1985). Different coding schemes for the different non-verbal modalities have been defined in psychology and can inspire the design of ECAs (Harrigan et al., 2005).

2.1 Facial Expressions

Facial expressions have been studied by Ekman (1999, 2003; Ekman and Friesen 1975). He described how “the rapid signals (seconds or fractions of seconds) are produced by the contractions of the facial muscles, resulting in temporary changes in facial appearance, shifts in the location and shape of the facial features, and temporary wrinkles.” Facial signals not only send emotional signals but also send emblematic messages (meaning of which is very specific, the nonverbal equivalent of a common word or phrase such as a head nod for “yes” and “no”). Raising the brows and holding them while keeping the rest of the face blank is an example of emblem signalling questioning. If the brow raise is done together with an upward head movement, it is an exclamation. Some facial emblems called emotional emblems are a conventional reference to a feeling. Facial signals are also used as conversational punctuators (emphasise a particular word). Thus, any movement

of a given facial area might have several meanings. For example, the lowered, drawn-together brow is part of anger, but it is also an emblem (for determination, concentration and perplexity) and a punctuator (Ekman and Friesen, 1975).

2.2 Coding Scheme

The Facial Action Coding System (FACS) (Ekman and Friesen, 2002) is a physically based coding scheme and does not include behavioural interpretation. It explains how to classify facial movements as a function of the muscles which are involved. Action units (AU) are used to represent the muscular activity that produces momentary changes in facial appearance. An AU can represent the movements of one or several muscles. Symmetrically, for some facial muscles, the effects of different parts of the same muscles are detailed using several AUs. FACS also includes the description of gaze behaviours and head movements. Ekman described in detail the distinctive clues to each basic emotion (Ekman, 2003) and also depending on the intensity of the emotion and its meaning how many areas can/must be involved and how their signals are modified according to intensity. When considering three areas (brows, eyes, mouth), surprise can be shown in just two areas, with the third area uninvolved. The combination of these two-area surprise faces has a slightly different meaning:

- Eyes + brows = questioning surprise
- Eyes + mouth = astonished surprise
- Brows + mouth = dazed or less interested surprise

2.3 Complex Facial Expressions of Emotion

Real-life emotions are often complex and involve several simultaneous emotions (Ekman and Friesen, 1975; Scherer, 1998; Ekman, 2003). They may occur either as the quick succession of different emotions, the superposition of emotions, the masking of one emotion by another one, the suppression of an emotion or the over-acting of an emotion. These blends of emotions produce “multiple simultaneous facial expressions” (Richmond and Croskey, 1999). Depending on the type of blending, the resulting facial expressions are not identical. A masked emotion may leak over the displayed emotion (Ekman and Friesen, 1975), while superposition of two emotions will be shown by different facial features (one emotion being shown on the upper face while another one on the lower face) (Ekman and Friesen, 1975). Ekman described in a systematic manner the blending of facial expressions for all the pairs of six acted basic emotions (happiness, sadness, surprise, fear, anger, disgust) (Ekman and Friesen, 1975). Ekman photographed models who were instructed to move particular facial muscles. He photographed three areas which are capable of independent movement: the brow/forehead; the eyes/lids and the root of the nose;

the lower face including the cheeks, mouth, most of the nose and chin. He also illustrates the context which can lead to each of these blends (for example, a woman being both angry and sad after a driver has run over her dog). He explains the complexity of these blended behaviours with regard to the timing and the subtleties of individual emotions (the families of facial expressions). A blend does not require that different facial areas show the different emotions. It can also be achieved by blending the appearance of the two emotions within each area of the face. For example, the blend anger/disgust is described by lips pressed (anger), upper lip raised (disgust), nose wrinkled (disgust).

2.4 Bodily Expression of Emotion

Several researchers experimentally studied expression of emotion in movement. Wallbott conducted a study in which actors were elicited with the following emotions: elated joy, happiness, sadness, despair, fear, terror, cold anger, hot anger, disgust, contempt, shame, guilt, pride and boredom (Wallbott, 1998). The actors had to utter meaningless sentences. Twelve drama students were asked to code body movements and postures performed by these actors. Wallbott observed distinctive patterns of movement and postural behaviour associated with some of the emotions studied. Some of these distinctive features seem to be typical of certain emotions, such as “arms crossed in front of chest” for the emotion of pride or the number of “self-manipulators” for shame. A number of movements and posture categories merely distinguish between “active” emotions (like hot anger or elated joy) and more “passive” emotions. His results indicate that movements and postural behaviour are indicative of the intensity of different emotions. Movement and postural behaviour when encoding different emotions (at least as captured with the categories used here) are to some degree specific to at least some emotions. For example, expressions of hot anger were observed to have the following discriminative features of postural behaviour: shoulder up, arms stretched out frontal and lateralised movements, pointing, opening/closing, back of hands sideways, many illustrators. Expressions of hot anger were also observed to have discriminative features of movement quality: high movement activity, expansive movements, high movement dynamics.

2.5 Multimodal Emotional Corpora

Analog video has been used for a long time for observing and manually annotating non-verbal behaviours. Since the last 10 years, several studies using digital video and computer-based annotations have been conducted in a variety of context and data (laboratory, meeting, TV material, field studies). Both analog and digital video corpora are used in two main goals: experimental studies or computational models. Compared to the social sciences experiments, digital corpus-based studies

aim at producing computational models of multimodal behaviour including details of individual behaviours that can be useful for the design of affective ECAs with individual profiles. The ISLE project surveyed several corpora of human multimodal communication built before 2002 (Wegener Knudsen et al., 2002a, b). A key issue of multimodal and emotional corpora is the availability of so-called real-life data (Douglas-Cowie et al., 2003; Batliner et al., 2000). Different corpora enable to study different dimensions of emotional behaviours:

- location of recording (laboratory/TV on stage/TV in street/natural environment (e.g. at home));
- instructions staging/spontaneous/portrayed/posed;
- timescale (minutes (TV clip)/hours (meeting)/days);
- awareness of being filmed, awareness of emotion study;
- subjects (actors/experts on emotion/ordinary people);
- interactivity; monologue/multi-party/social constraints on interactivity;
- human–human or human–computer or mixed;
- number of recorded modalities, resolution and quality of collected data; intrusiveness of recording media

The reader can find more details about methodologies and tools for the collection and annotation of multimodal corpora in Martin et al. (2004, 2006), and Maybury and Martin (2002) and about emotional corpora in Devillers et al. (2006).

Videos of users interacting with the SAL system were manually annotated and used for the definition of a list of nonverbal behaviour generation rules (Lee and Marsella, 2006). The multimodal behaviour of a human storyteller was annotated in de Melo and Paiva (2006) and used for the design of an ECA in a narrative application. A parametric model for iconic gesture generation is defined by Tepper et al. (2004) following the analysis of a video corpus. Cartoon has been used as a video material for studying the role of irregularities and discontinuities modulations in expressivity (Ech et al., 2006). Recent emotional corpora aim at studying the synchronisation between the different modalities (Bänziger and Scherer, 2007).

2.6 Coordination of Multiple Modalities

In this section, we explain the different means for coordinating modalities. One main feature of multimodal communication is that there are multiple mappings between meanings and multimodal signals (Poggi 1996, 2003). Several signals can be displayed in one or several modalities for a single meaning. For example, the meaning “emphasis” of a given word can be displayed via an eyebrow raise, a head nod or both. Symmetrically, a given signal can be used for several meanings: an eyebrow raise might mean surprise, emphasis or suggest.

Due to the ambiguity of meanings of behaviour in the different modalities, multimodal behaviours must be interpreted by analysing the different modalities

altogether. Several classifications have been proposed for describing how modalities may combine either in human communication or in multimodal interfaces (Knapp and Hall, 2006; Scherer, 1984; Buisine, 2005; Poggi, 2006; Martin et al., 2001; Martin and Béroule 1993; Coutaz et al., 1995; Kendon, 2004; Engle, 2000). The following notions are introduced in these classifications:

- Equivalence/substitution: one modality conveys a meaning not borne by the other modalities (while it could be conveyed by these other modalities)
- Redundancy/repetition: the same meaning is conveyed at the same time via several modalities
- Complementarity:
 - Amplification accentuation/moderation: one modality is used to amplify or attenuate the meaning provided by another modality
 - Additive: one modality adds congruent information to another modality
 - Illustration/clarification: one modality is used to illustrate/clarify the meaning conveyed by another modality
- Conflict/contradiction: the meaning transmitted on one modality is incompatible or contrasting with the one conveyed by the other modalities; this cooperation occurs when the meaning of the individual modalities seems conflicting but indeed the meaning of their combination is not and emerges from the conflicting combination of the meanings of the individual modalities.
- Independence: the meanings conveyed by different modalities are independent and should not be merged.

These combinations of modalities can be brought into play at different levels (Poggi, 2006). A signal (e.g. a pointing finger) may have a repetitive function with regard to its literal meaning, whereas its indirect meaning has an additive function (e.g. accusation).

3 Embodied Conversational Agents

In this section, we explain the importance of behaviour generation in ECAs. One makes a distinction between two parts in an ECA: its body and its mind. The body is in charge of the display of the behaviours. The mind is in charge of the reasoning, planning and management of the interaction with the user or the surrounding environment (Pelachaud et al., 2004).

3.1 Definitions and Examples

Several models have been proposed for the generation, selection and planning of multimodal behaviours in an ECA (Cassell et al., 2000; Prendinger and Ishizuka,

2004). André et al. (2000) describe the design of dialogue for agents. One approach consists in inserting gesture tags in the text to be spoken by the agent (Cassell et al., 2001). Automatic learning of these manual rules for gesture selection and timing from manual specifications is proposed in Kipp, (2006). With respect to the synchronisation of different modalities, the gestures of ECAs are often timed so that the stroke coincides with the stressed syllable of the co-occurring concept in speech (Cassell et al., 1994). The affective dimension is recognised of importance for pedagogical agents (Jacques et al., 2004). Jaques et al. (2004) describe some affective tactics that their mediating agent should bring into play to promote a positive mood in the student. Several experimental studies have evaluated the utility of such pedagogical agents for learning or to motivate students (Wonisch and Cooper, 2002; Lester et al., 1997; Moreno et al., 2001, Dehn and van Mulken, 2000).

Cosmo is a full body 3D cartoon like pedagogical agent in a learning environment for the domain of Internet packet routing (Lester et al., 2000). Given a communicative goal, the explanation planner determines the content and structure of an agent's explanations and then passes specifications of triplets (communicative act, topic, referent) to the deictic planner. The deictic planner exploits a world model, a user model, the current problem state, one focus history for speech and one focus history for gesture. Then the deictic behaviour planner selects and coordinates locomotive, gesture and speech behaviours. This deictic planning operates in three phases: (1) ambiguity appraisal, (2) gesture and locomotion planning, and (3) utterance planning and coordination (real-time composition of deictic gesture and gaze components). If the referent was mentioned in a recent utterance, only a spoken pronominal reference is generated (e.g. "it has low traffic" where "it" stands for a subnet that was pointed to by the agent in the previous utterance). The emotional planner that relies on a set of emotional behaviours has been designed so that it corresponds to variations of the most frequent pedagogical acts required in the application. The graphical design of the behaviours is based on knowledge from the literature of the animation industry.

The formation and motion characteristics of pointing by hand at objects of different sizes and locations have been investigated and used for embodied agents acting as weather reporters (Noma et al., 2000), cooperative assembly companion (Kopp et al., 2003) or explaining directions (Cassell et al., 2007). Pointing/reaching hand gesture should be preceded by gaze. The interplay of gaze, head and posture when looking at a direction has also been studied. Horizontal gaze shift greater than 25 degree, and vertical gaze shift greater than 10 degree are performed by the combination of head and gaze movement (Chopra-Khullar and Balder, 2001). These principles have been applied to visually explore the environment as well as follow moving targets by pointing or gaze. Gaze has been used to identify and follow "interesting", new and moving objects or people in a virtual environment (Hill et al., 2002), as well as in case of responsive faces reflecting emotions and attention according to the number of people appearing and moving around in a real world. See Sect. 6.2 (this handbook) for more details on such models.

3.2 Affective Agents

Greta incorporates conversational and emotional qualities (Pelachaud, Poggi et al., 2005). To determine speech accompanying nonverbal behaviours, the system relies on the taxonomy of communicative functions proposed by Isabella Poggi (2003). Communicative functions are defined by meaning, signal, expressivity. For example, looking away from the interlocutor (signal) conveys a take-turn meaning, within the communicative function “turn allocation”.

3.3 Multimodal Emotional Corpora for Informing the Design of Affective Agents

Several studies rely on computer-based annotation of video-taped human behaviours to inform the design of ECAs. Such an approach has two advantages: it is experimentally grounded and thus can be more relevant for the task at hand than general rules provided by the literature. The second advantage is that it enables to model individual use of nonverbal behaviours, which can be a means to achieve animated agents with an individual profile (Kipp, 2004).

Elisabeth André addresses several issues related to the use of corpus-based approaches to behaviour modelling for embodied agents via observation, modelling, generation and perception (André, 2006). She reviews different approaches such as “understanding via learning to generate” (e.g. Wachsmuth), “analysis by observation and synthesis” (e.g. Kraher), “Study, model, build, test cycle” (e.g. Cassell). The author also recommends to start from a model when collecting data.

3.4 Coordination of Multiple Modalities in ECAs

Redundancy was observed to provide better results than complementarity in a study investigating recall and preference after technical presentations made by an ECA (Buisine and Martin, 2007a). Yet, individual differences were observed with respect to gender and extraversion of the subjects (Buisine and Martin, 2007b). Redundancy might also seem unnatural if too frequent and involving too many modalities in a systematic way during a long presentation (Cassell et al., 1994).

Synchrony between speech and facial expression occurs at all levels of speech (Condon, 1986): phonemic segment, word, phrase or long utterance (Cassell et al., 1994). Different facial motions are characteristic of these different groups. Some of them are more adapted to the phoneme level, like an eye blink, while others act at the word level, like a frown. In the illustrative example “Do you have a blank check?”, a raising eyebrow might start and end on the accented syllables “check”, while a blink might start and end on the pause marking the end of the utterance. Facial expression of emphasis can match the emphasised segment, showing synchronisation at this level (a sequence of head nods can punctuate the emphasis). Moreover, some

movements reflect encoding–decoding difficulties and therefore coincide with hesitations and pauses inside clauses. Many hesitation pauses are produced at the beginning of speech and correlate with avoidance of gaze (the head of the speaker turns away from the listener) as if to help the speaker to concentrate on what she is going to say (Duncan and Friske, 1977).

One example of coordination between multiple modalities occurs in the case of response to external stimuli. An embodied agent should react to the visual and acoustic stimuli from the surrounding virtual and/or real world, in order to be perceived as life like and believable. Reacting to signals from the real world contributes to “sharing the real world with the human interlocutor” experience. Such signals play a role in human nonverbal behaviour in two different ways:

- (a) Invoke some behaviour, often on a biological reflex level. For example, a sudden noise triggers a blink and attracts attention; the pupil size adjusts itself to the lighting intensity; sudden intense light evokes blinking.
- (b) Modify some behaviour, concerning its timing and/or formation. For example, when looking at a moving object, the eyes, the head, the upper torso or even the entire body may be turned, depending on the relative location and speed of the object being looked at. Or when approaching a train to get on, one not only generates a shortest path with respect to the other passengers and obstacles, but may speed up his locomotion when hearing the whistle of the conductor.

Coordination/intersynchrony to accomplish a common task: In case of interpersonal gestures involving physical touch, the behaviours of two embodied agents need to be simultaneously coordinated, like in the case of a handshake, a hug or kiss. Both physical and social factors play a role in the coordination strategy. Though among humans the physical contact is established by getting feedback via peripheral vision, the protocol of such gestures too involves synchrony with mutual gaze: eye contact is to be established during the handshake. The duration of the handshake, the proximity and eventual head nod or bowing move of the upper body are to not only be subtly coordinated, but also be used according to the protocol reflecting the gender, ethnicity and social relationship of the two parties.

3.5 Facial Expressions of Emotions in Affective Agents

In this section, we focus on the signs of emotions in facial expressions.

3.5.1 MPEG-4

MPEG-4 is an ISO/IEC international standard defined in 1998. It aims at overcoming all the divisions in the world of facial animation by defining a standard way to deal with synthetic faces using a model-based approach. It led to the definition of a set of animation parameters and semantic rules which can be used to drive any synthetic face compliant with MPEG-4 (Pandzic and Forchheimer, 2002). The

six archetypal expressions (joy, sadness, anger, fear, disgust and surprise) can be described using MPEG-4 (Malatesta et al., 2007).

Emotions that belong to the same category can be rendered by animating the same FAPs using different intensities. For example, the emotion group fear also contains worry and terror; these two emotions can be synthesised by reducing or increasing the intensities of the employed facial animation parameter (FAP). Creating profiles for emotions that do not clearly belong to a universal category is not straightforward. Apart from estimating the range of variations for FAPs, one should first define the FAPs which are involved in the particular emotion.

Very few models of blended emotions have been developed so far for ECAs. The interpolation between facial parameters of given expressions is commonly used to compute the new expression (Pandzic and Forchheimer, 2002; Albrecht et al., 2005; Ruttkay, 2003; Duy Bui, 2004; Malatesta et al., 2007).

3.5.2 Complex Expression of Emotions

Distinguishing various types of blends of emotions in ECA systems is relevant as perceptual studies have shown that people are able to recognise facial expression of felt emotion (Ekman, 1982; Wiggers, 1982) as well as fake emotion (Ekman, 1982) from real life as well as on ECAs (Pandzic and Forchheimer, 2002). Moreover, in a study on deceiving agent, Rehm and André (2005) found that the users were able to differentiate when the agent was displaying expressions of felt emotion or expression of fake emotion. Aiming at understanding whether facial features or regions play identical roles in emotion recognition, researchers performed various perceptual tasks or studied psychological facial activity (Bassili, 1979; Gouta and Miyamoto, 2000; Constantini et al., 2005; Cacioppo et al., 1986). They found that positive emotions are mainly perceived from the expression of the lower face (e.g. smile) while negative emotion from the upper face (e.g. frown). One can conclude that reliable features for positive emotion, that is, features that convey the strongest characteristics of a positive emotion, are in the lower face.

On the other hand, the most reliable features for negative emotion are in the upper face.

Based on such findings computational models for facial expressions of blend of emotions can be defined (Ochs et al., 2005; Niewiadomski, 2007). Such models compose facial expressions from those of single emotions using fuzzy logic-based rules, for example, for assessing the similarity between facial expressions (Niewiadomski and Pelachaud, 2007).

3.6 Bodily Expression of Emotions in Affective Agents

Studies from the social sciences observed that discriminative features of emotions are found (1) in postures and (2) in movement quality (Wallbott, 1998). Thus, both should be considered when designing an affective agent. In posture, three aspects are required in posture design: achieving a set of world space constraints, finding a body shape that reflects the character's inner state and personality and making

adjustments to balance that act to strengthen the pose and also maintain realism (Neff and Fiume, 2005). Experimental studies are being conducted to collect more knowledge on the posture features which can be used with a discrete or a dimensional model (Kleinsmith and Bianchi-Berthouze, 2007). Few studies have explored the combination of postures with other modalities.

3.6.1 Expressivity

The Greta ECA (Pelachaud) features a model of gesture expressivity that acts on the production of communicative gestures. The model of expressivity is based on studies of nonverbal behaviour (Wallbott, 1998; Wallbott and Scherer, 1986; Gallaher, 1992). Expressivity is described as a set of six dimensions (Hartmann et al., 2002, 2005). Each dimension acts on a characteristic of communicative gestures. Spatial extent describes how large the gesture is in space. Temporal extent describes how fast the gesture is executed. Power describes how strong the performance of the gesture is. Fluidity describes how two consecutive gestures are co-articulated one merging with the other. Repetition describes how often a gesture is repeated. Overall activity describes how many behaviour activities there are over a time span. For the face, the expressivity dimensions act mainly on the intensity of the muscular contraction and its temporal course. In order to consider individual agents, an individual expressive profile can be assigned to the agent specifying in which modalities the agent is the most expressive. The expressivity attributes can act over a whole animation, on gesture phases or on every frame (Ech et al., 2006). A corpus-based approach with acted data collected in laboratory was used to provide values for such parameters using image processing techniques (Caridakis et al., 2006). Image processing was also used for analysing bodily movement of a dancer (Volpe, 2005). Emotional reaction of affective agents: An example of a body gesture that depends on the emotional context is the reaction movement. Reactions are unconscious behaviours that improve the realism of affective agents. Based on observation experiments of real people reacting, different types of reactions were identified: avoid, face and protect. These types of reactions are associated with personality traits (Garcia-Rojas et al., 2007). To synthesise these reactions into embodied agents, a semantic model that represents the animation synthesis, agent's geometry and a description of individual agents through personality, emotion, age, gender, etc. was created. The advantage of such a semantic model is that it enables scalability for implementation of concepts.

3.7 *Coordinating the Generation of Signs in Multiple Modalities in an Affective Agent*

When expressing superposition of two emotions in several audio and visual modalities, how should the signs of the different emotions be split into different modalities? Some researchers study how to generate mixed emotions in ECAs (Carofiglio

et al., in press, Karunaratne and Yan, 2006; Lee et al., 2006). Arfib (2006) uses time warping to align video of one emotion and vocal tempo of another emotion using a time-warping algorithm, hence artificially creating multimodal blends of emotions.

Video corpora of TV interviews enable to explore how people behave during such blended emotions (Devillers et al., 2006; Martin et al., 2005) not only by their facial expression but also by their gestures and their speech (Douglas-Cowie et al., 2005). Yet, these corpora call for means of validating subjective manual annotations of emotion.

Annotation is composed of two steps: Step 1 aims at the automatic annotation of the video with data that can be useful either for the manual annotation of the video or the specification of the agent's behaviour: pitch, intensity, etc. Step 2 involves manual annotations of the video. The word by word transcription including punctuation is achieved following the LDC norms for hesitations, breath, etc. The video is then annotated at several temporal levels (whole video, segments of the video and behaviours observed at specific moments) and at several levels of abstraction. The global behaviour observed during the whole video is annotated with communicative act, emotions and multimodal cues. The segments are annotated with emotion labels and the modalities perceived as relevant with regard to emotion. We have grounded this coding scheme in requirements collected from the parameters known as perceptually relevant for the study of emotional behaviour and the features of our emotionally rich TV interviews. Movement expressivity is annotated for torso, head, shoulders and hand gestures.

Manual steps can be followed for defining an animation of an affective agent replaying annotated multimodal emotional behaviours observed in an annotated video. Such a copy-synthesis approach enables to identify the levels of representation that are required (Martin et al., 2006) via several steps: annotation, extraction and generation. Perceptual experiments are recommended to compare the perception of the original videos and the animations of the corresponding behaviour replayed by an affective agent (Buisine et al., 2006).

Similarly, automatic mimicry approach lies in the fact that both facial and gestural aspects of the user's behaviour are analysed and processed. The mimicry consists of perception, interpretation, planning and animation of the expressions shown by the human, resulting not in an exact duplicate rather than an expressive model of the user's original behaviour (Caridakis et al., 2007).

4 Conclusions

In this section, we explained the complexity of expressions of emotions in individual modalities. Some discriminative features of specific emotions have been pointed out in facial expressions, in posture and in expressive movements. Several possible ways to combine different modalities were presented for multimodal coordination and synchronisation.

From a methodological point of view, we also described how to apply corpus-based approaches for studying how different modalities are coordinated in human communication. We explained how to specify such expressions of complex emotions in an affective agent using facial expressions and gestures.

Where the display of signs in individual modalities by an affective ECA can be seen as quite advanced (e.g. displaying expression of basic or intermediate emotions using MPEG-4), the coordination of different modalities during the expression of complex emotions was only recently brought on the forefront and deserves dedicated experimental studies.

Researchers are also paying attention to individual differences in the perception of multimodal behaviours displayed by affective agents. For example, extrovert and introvert versions of static text and static graphical display of posture were combined by Isbister and Nass (2000). It was found that users identify the intended personality trait. Furthermore, users preferred consistent verbal and nonverbal, as well as agents with a complementary extroversion compared to their own.

References

- Albrecht I, Schröder M, Haber J, Seidel H-P (2005) Mixed feelings: Expression of non-basic emotions in a muscle-based talking head. *J Virtual Reality Lang Speech Gesture* 8(4 Special issue):201–212
- André E (2006) Corpus-based approaches to behavior modeling for virtual humans: a critical review. Modeling communication with robots and virtual humans. In: Workshop of the ZiF: research group 2005/2006 “Embodied communication in humans and machines”. Scientific organization: Ipke Wachsmuth (Bielefeld), Günther Knoblich (Newark)
- André E, Rist T, van S, Mulken, Klesen M, Baldes S (2000) The automated design of believable dialogues for animated presentation teams. In: Cassell JSJ, Prevost S, Churchill E (eds) *Embodied conversational agents*. MIT Press, Cambridge, MA, pp 220–255
- Arbib D (2006) Time warping of a piano (and other) video sequences following different emotions. In: Workshop on “subsystem synchronization and multimodal behavioral organization” held during Humaine summer school. Genova
- Argyle M (2004) *Bodily communication*, 2nd edn. Routledge, London and Taylor and Francis, New York, NY
- Bänziger T, Scherer K (2007) Using actor portrayals to systematically study multi-modal emotion expression: the GEMEP corpus. In: 2nd international conference on affective computing and intelligent interaction (ACII 2007) Lisbon, Portugal, pp 476–487
- Bassili JN (1979) Emotion recognition: the role of facial movement and the relative importance of upper and lower areas of the face. *J Pers Soc Psychol* 37(11):2049–2058
- Batliner A, Fisher K, Huber R, Spilker J, Noth E (2000) Desperately seeking emotions or: Actors, wizards, and human beings. ISCA Workshop on speech and emotion: a conceptual framework for research Newcastle, Northern Ireland, pp 195–200
- Buisine S, Abrilian S, Niewiadomski R, Martin J-C, Devillers L, Pelachaud C (2006) Perception of blended emotions: from video corpus to expressive agent. 6th International Conference on Intelligent Virtual Agents (IVA’2006). Best paper award. Springer, Marina del Rey, CA, pp 93–106
- Buisine S (2005) Conception et évaluation d’Agents conversationnels multimodaux bidirectionnels. PhD Thesis. Doctorat de Psychologie Cognitive – Ergonomie, Paris V. 8 avril 2005. Direction J.-C. Martin et J.-C. Sperandio. 2005. URL <http://stephanie.buisine.free.fr/>. Accessed on 4 November 2010

- Buisine S, Martin JC (2007a) The effects of speech-gesture cooperation in animated agents' behavior in multimedia presentations. *Interact Comput* 19:484–493
- Buisine S, Martin J-C (2007b) The influence of personality on the perception of embodied agents' multimodal behavior. In: 3rd conference of the international society for gesture studies
- Cacioppo JT, Petty RP, Losch ME, Kim HS (1986) Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *J Pers Soc Psychol* 50:260–268
- Caridakis G, Raouzaoui A, Karpouzis K, Kollias S (2006) Synthesizing gesture expressivity based on real sequences. In: Workshop "multimodal corpora. From multimodal behaviour theories to usable models". 5th international conference on language resources and evaluation (LREC'2006), Genova, Italy, pp 19–23
- Caridakis G, Raouzaoui A, Bevacqua E, Mancini M, Karpouzis K, Malatesta L, Pelachaud C (2007) Virtual agent multimodal mimicry of humans. *J Lang Res Eval* (Special issue) "Multimodal Corpora". *J Lang Res Eval* (41):367–388
- Carofiglio V, de Rosis F, Grassano R (2008) Dynamic models of multiple emotion activation. In: Canamero L, Aylett R (eds) *Animating expressive characters for social interactions*. John Benjamins, Amsterdam, pp 123–141
- Cassell J, Pelachaud C, Badler N, Steedman M, Achorn B, Becket T, Douville B, Prevost S, Stone M (1994) Animated conversation: rule-based generation of facial expression, gesture and spoken intonation for multiple conversational agents. In: *ACM SIGGRAPH'94*, pp 413–420
- Cassell J, Bickmore T, Billinghurst M, Campbell L, Chang K, Vilhjálmsón HH, Yan H (1999) Embodiment in conversational interfaces: Rea. CHI'99 (SIGCHI conference on Human factors in computing systems: the CHI is the limit) Pittsburgh, PA, USA, pp 520–527
- Cassell J, Bickmore T, Campbell L, Vilhjálmsón H, Yan H (2000) Human conversation as a system framework: designing embodied conversational agents. In: Cassell J, Sullivan J, Prevost S, Churchill E (eds) *Embodied conversational agents*. MIT Press, Cambridge, MA, pp 29–63
- Cassell J, Vilhjálmsón H, Bickmore T (2001) BEAT: the Behavior Expression Animation Toolkit. In: 28th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '01) Los Angeles, CA, pp 477–486
- Cassell J, Kopp S, Tepper P, Ferriman K, Striegnitz K (2007) Trading spaces: how humans and humanoids use speech and gesture to give directions. In: Nishida T (ed) *Conversational informatics*. Wiley, New York, NY, pp 133–160
- Chopra-Khullar S, Badler N (2001) Where to look? Automating attending behaviours of virtual human characters. In: 4th Conference on AAMAS, pp 9–23
- Condon WS (1986) *Communication: rhythm and structure. rhythm in psychological, linguistic and musical processes*. Charles C Thomas Publisher
- Constantini E, Pianesi F, Prete M (2005) Recognizing emotions in human and synthetic faces: the role of the upper and lower parts of the face. In: *Intelligent User Interfaces (IUT'05)* San Diego, CA, USA, pp 20–27
- Coutaz J, Nigay L, Salber D, Blandford AE, May J, Young RMY (1995) Four easy pieces for assessing the usability of multimodal interaction. In: *Interact'95* pp 115–120
- de Melo C, Paiva A (2006) A Story about Gesticulation Expression. In: 6th International Conference on Intelligent Virtual Agents (IVA'06) Marina del Rey, CA, pp 270–281
- Dehn DM, van Mulken S (2000) The impact of animated interface agents: a review of empirical research. *Int J Hum Comput Stud* 52:1–22
- Devillers L, Cowie R, Martin J-C, Douglas-Cowie E, Abrilian S, McRorie M (2006) Real life emotions in French and English TV video clips: an integrated annotation protocol combining continuous and discrete approaches. In: 5th in-ternational conference on Language Resources and Evaluation (LREC 2006), Genoa, Italy
- Devillers L, Martin J-C, Cowie R, Douglas-Cowie E, Batliner A (2006) Workshop "Corpora for research on emotion and affect". In: 5th international conference on language resources and evaluation (LREC'2006). Genova, Italy

- Douglas-Cowie E, Campbell N, Cowie R, Roach P (2003) Emotional speech; Towards a new generation of databases. *Speech Commun* 40
- Douglas-Cowie E, Devillers L, Martin J-C, Cowie R, Savvidou S, Abrilian S, Cox C (2005) Multimodal databases of everyday emotion: facing up to complexity. In: 9th European Conference on Speech communication and technology (Interspeech'2005), Lisbon, Portugal, pp 813–816
- Duncan S, Fiske D (1977) *Face-to-face interaction: research, methods and theory*. Lawrence Erlbaum, Hillsdale, NJ
- Duy Bui T (2004) *Creating emotions and facial expressions for embodied agents*. PhD Thesis. University of Twente
- Ech Chafai N, Pelachaud C, Pelé D, Breton G (2006) Gesture Expressivity Modulations in an ECA Application. In: 6th international conference on intelligent virtual agents (IVA'06) Marina del Rey, CA, pp 181–192
- Ekman P (1982a) *Emotion in the human face*. Cambridge University Press
- Ekman P, Friesen W (1982) Felt, false, miserable smiles. *J Nonverb Behav* 6:4
- Ekman P (1999) Basic emotions. In: Dalglish T, Power MJ (eds) *Handbook of cognition & emotion*. Wiley, New York, NY, pp 301–320
- Ekman P (2003a) *Emotions revealed. Understanding faces and feelings*. Weidenfeld and Nicolson, London
- Ekman P (2003b) *The face revealed*. Weidenfeld and Nicolson, London
- Ekman P, Friesen WV (1975) *Unmasking the face. A guide to recognizing emotions from facial clues*. Prentice-Hall, Englewood Cliffs, NJ.
- Ekman P, Friesen WC, Hager JC (2002) *Facial action coding system. The Manual on CD ROM*
- Engle RA (2000) *Toward a theory of multimodal communication: combining speech, gestures, diagrams and demonstrations in instructional explanations*. PhD Thesis, Stanford University
- Feldman RS, Rim B (1991) *Fundamentals of nonverbal behavior. Studies in emotion and social interaction*. Cambridge University Press, Cambridge
- Gallaher P (1992) Individual differences in nonverbal behavior: Dimensions of style. *J Pers Soc Psychol* 63:133–145
- Garcia-Rojas A, Vexo F, Thalmann D (2007) Semantic representation of individualized reaction movements for virtual humans. *J Virtual Reality* 6(1):25–32
- Gouta K, Miyamoto M (2000) Emotion recognition, facial components associated with various emotions. *Shinrigaku Kenkyu* 71(3):211–218
- Harrigan JA, Rosenthal R, Scherer K (2005) *The new handbook of methods in nonverbal behavior research. Series in Affective Science*. Oxford University Press, Oxford
- Hartmann B, Mancini M, Pelachaud C (2002) Formational parameters and adaptive prototype instantiation for MPEG-4 compliant gesture synthesis. In: *Computer animation (CA'2002)* Geneva, Switzerland, pp 111–119
- Hartmann B, Mancini M, Pelachaud C (2005) Implementing expressive gesture synthesis for embodied conversational agents. In: *Gesture Workshop (GW'2005)*, Vannes, France
- Hill R, Han C, van Lent M (2002) Perceptually driven cognitive mapping of urban environments. In: *First international joint conference on autonomous agents and multiagent systems*, Bologna, Italy
- Isbister K, Nass C (2000) Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *Int J Hum Comput Stud* 53:251–267
- Jacques PA, Vicari RM, Pesty S, Bonneville J-F (2004) Applying affective tactics for a better learning. In: 16th European Conference on Artificial Intelligence (ECAI 2004). Valencia, Spain, IOS, Amsterdam, pp 109–113
- Karunaratne S, Yan H (2006) Modelling and combining emotions, visual speech and gestures in virtual head models. *Signal Process Image Commun* 21:429–449
- Kendon A (2004) *Gesture : visible action as utterance*. Cambridge University Press, Cambridge
- Kipp M (2004) *Gesture generation by imitation. From human behavior to computer character animation*. Boca Raton, Dissertation.com Florida

- Kipp M (2006) Creativity meets automation: combining nonverbal action authoring with rules and machine learning. In: 6th international conference on intelligent virtual agents (IVA'06) Marina del Rey, CA, pp 230–242
- Kleinsmith A, Bianchi-Berthouze N (2007) Recognizing affective dimensions from body posture. In: 2nd international conference on affective computing and intelligent interaction (ACII 2007) Lisbon, Portugal, pp 48–58
- Knapp ML, Hall JA (2006) Nonverbal communication in human interaction, 16th edition. Thomson and Wadsworth, Belmont, CA
- Kopp S, Jung B, Lessmann N, Wachsmuth I (2003) Max - A multimodal assistant in virtual reality construction. *KI-Künstliche Intelligenz*. Vol. 4/03, pp 11–17
- Lee J, Marsella S (2006) Nonverbal behavior generator for embodied conversational agents. In: 6th international conference on intelligent virtual agents (IVA'06) Marina del Rey, CA, pp 243–255
- Lee B, Kao E, Soo V (2006) Feeling ambivalent: a model of mixed emotions for virtual agents. In: 6th international conference on intelligent virtual agents (IVA'06) Marina del Rey, CA, pp 329–342
- Lester J, Converse S, Kahler S, Barlow T, Stone B, Bhogal R (1997) The Persona effect: affective impact of animated pedagogical Agents. In: CHI '97 Atlanta, pp 359–366
- Lester JC, Towns SG, Callaway CB, Voerman JL, P, F (2000) Deictic and emotive communication in animated pedagogical agents. *Embodied conversational agents*. The MIT Press, Cambridge, MA
- Malatesta L, Raouzaïou A, Karpouzis K, Kollias S (2007) MPEG-4 facial expression synthesis. *J Pers Ubiquitous Comput 'Emerg Multimodal Interfaces'* (Special issue) following the special session of the AIAI 2006 Conference. Springer, 13(1):77–83
- Martin JC, Bérroule D (1993) Types et buts de coopérations entre modalités. *Cinquièmes Journées sur l'Ingénierie des Interfaces Homme-Machine* Lyon, France, pp 17–22
- Martin JC, Grimard S, Alexandri K (2001) On the annotation of the multimodal behavior and computation of cooperation between modalities. In: Workshop on “Representing, Annotating, and Evaluating Non-Verbal and Verbal Communicative Acts to Achieve Contextual Embodied Agents” in conjunction with the 5th International Conference on Autonomous Agents (AAMAS'2001) Montreal, Canada, pp 1–7
- Martin J-C, den Os E, Kuhnlein P, Boves L, Paggio P, Catizone R (2004) Workshop “Multimodal corpora: models of human behaviour for the specification and evaluation of multimodal input and output interfaces”. In: Association with the 4th international conference on language resources and evaluation LREC2004 URL <http://multimodal-corpora.org/>. Accessed on 4 November 2010. Centro Cultural de Belem, LISBON, Portugal
- Martin J-C, Abrilian S, Devillers L (2005) Annotating multimodal behaviors occurring during non basic emotions. In: 1st International Conference on affective computing and intelligent interaction (ACII'2005), Beijing, China, pp 550–557
- Martin J-C, Kuhnlein P, Paggio P, Stiefelhagen R, Pianesi F (2006) Workshop “Multimodal Corpora: from Multimodal Behaviour Theories to Usable Models”. In: Association with the 5th international conference on language resources and evaluation (LREC2006), Genoa, Italy
- Martin J-C, Niewiadomski R, Devillers L, Buisine S, Pelachaud C (2006) Multimodal complex emotions: gesture expressivity and blended facial expressions. *J Humanoid Rob (Special issue)*. In: Pelachaud C, Canamero L (eds). *Achieving human-like qualities in interactive virtual and physical humanoids*. 3(3):269–291
- Maybury M, Martin J-C (2002) Workshop on “Multimodal Resources and Multi-modal Systems Evaluation”. In: Conference on language resources and evaluation (LREC'2002), Las Palmas, Canary Islands, Spain
- Moreno R, Mayer RE, Spires HA, Lester JC (2001) The case for social agency in computer-based teaching: do students learn more deeply when they interact with animated pedagogical agents? *Cognition Instr* 19:177–213
- Neff M, Fiume E (2005) Methods for exploring expressive stance. *Graph Models*. Special issue on SCA 2004. 68(2):133–157

- Niewiadomski RA (2009) model of complex facial expressions in interpersonal relations for animated agents. PhD Thesis. PhD. dissertation, University of Perugia
- Niewiadomski R, Pelachaud C (2007) Fuzzy similarity of facial expressions of embodied agents. In: 7th international conference on intelligent virtual agents (IVA'2007) Paris, France, pp 86–98
- Noma T, Zhao L, Badler N (2000) Design of a Virtual Human Presenter. *IEEE J Comput Graph Appl* 20(4):79–85
- Ochs M, Niewiadomski R, Pelachaud C, Sadek D (2005) Intelligent expressions of emotions. In: 1st International Conference on Affective Computing and Intelligent Interaction (ACII'2005), Springer-Verlag, Beijing, China, pp 707–714
- Pandzic IS, Forchheimer R (2002) MPEG-4 facial animation. The standard, implementation and applications. Wiley and Sons, LTD
- Pelachaud C (2005) Multimodal expressive embodied conversational agent. ACM Multimedia, Brave New Topics session, Singapore, 6–11 November, ACM
- Pelachaud C, Braffort A, Breton G, Ech Chadai N, Gibet S, Martin J-C, Maubert S, Ochs M, Pelé D, Perrin A, Raynal M, Réveret L, Sadek D (2004) AGENTS CONVERSATIONNELS : Systèmes d'animation Modélisation des comportements multimodaux applications : agents pédagogiques et agents signeurs. Action Spécifique du CNRS Humain Virtuel
- Poggi I (1996) Mind markers. In: 5th International pragmatics conference, Mexico City
- Poggi I (2003) Mind markers. Gestures. Meaning and use. University Fernando Pessoa Press
- Poggi I (2006) Social influence through face, hands, and body. In: Second Nordic Conference on Multimodality Goteborg, Sweden, pp 5–29
- Poggi I, Pelachaud C, de Rosiis F, Carofiglio V, De Carolis B (2005) GRETA. A Believable Embodied Conversational Agent. In: Stock O, Zancarano M (eds) Multimodal intelligent information presentation. Kluwer, Dordrecht, pp 3–26
- Prendinger H, Ishizuka M (2004) Life-like characters. Tools, affective functions and applications. Springer, Berlin
- Rehm M, André E (2005) Catch Me If You Can – Exploring Lying Agents in Social Settings. In: International Conference on Autonomous agents and multiagent systems (AAMAS'2005) Utrecht, the Netherlands, pp 937–944
- Richmond VP, Croskey JC (1999) Non Verbal Behavior in Interpersonal relations. Allyn and Bacon
- Ruttkay Z, Noot H, ten Hagen P (2003) Emotion Disc and Emotion Squares: tools to explore the facial expression face. *Comput Graph Forum* 22(1):49–53
- Scherer KR (1984) Les fonctions des signes non verbaux dans la communication. In: Cosnier J, Brossard A (eds) La communication non verbale. Delachaux & Niestl, Paris, pp 71–100
- Scherer KR (1998) Analyzing Emotion Blends. In: Proceedings of the 10th conference of the international society for research on emotions Wrzburg, Germany, pp 142–148
- Scherer KR (2000) Emotion. Introduction to social psychology: a European perspective. Blackwell, Oxford
- Scherer KR, Ellgring H (2007) Multimodal expression of emotion: affect programs or componential appraisal patterns? *Emotion* 7:1
- Siegmán AW, Feldstein S (1985) Multichannel integrations of nonverbal behavior. LEA–Routledge, New York, NY
- Tepper P, Kopp S, Cassell J (2004) Content in context: generating language and iconic gesture without a gestionary. Workshop on balanced perception and action in ECAs at autoumous agents and multiagent systems (AAMAS), New York, NY, USA
- Volpe G (2005) Special issue on expressive gesture in performing arts and new media. *J New Music Rese Taylor and Francis* 34:1
- Wallbott HG (1998) Bodily expression of emotion. *Eur J Soc Psychol* 28:879–896
- Wallbott HG, Scherer KR (1986) Cues and channels in emotion recognition. *J Pers Soc Psychol* 51(4):690–699
- Wegener Knudsen M, Martin J-C, Dybkjr L, Berman S, Bernsen NO, Choukri K, Heid U, Kita S, Mapelli V, Pelachaud C, Poggi I, van Elswijk G, Wittenburg P (2002a) Survey of NIMM data

- resources, current and future user profiles, markets and user needs for NIMm resources. ISLE natural interactivity and multimodality. Working Group Deliverable D8.1
- Wegener Knudsen M, Martin J-C, Dybkjr L, Machuca Ayuso M-J, Bernsen NO, Carletta J, Heid U, Kita S, Llisterri J, Pelachaud C, Poggi I, Reithinger N, van Elswijk G, Wittenburg P (2002b) Survey of multimodal annotation schemes and best practice. ISLE natural interactivity and multimodality. Working Group Deliverable D9.1
- Wiggers M (1982) Jugments of facial expressions of emotion predicted from facial behavior. *J Nonverb Behav* 7(2):101–116
- Wonisch D, Cooper G (2002) Interface agents: preferred appearance characteristics based upon context. In: *Virtual conversational characters: applications, methods, and research challenges in conjunction with HF2002 and OZCHI2002* Melbourne, Australia