Using a Wizard of Oz study to inform the design of SenToy

Gerd Andersson, Kristina Höök
SICS[1]

Dário Mourão, Ana Paiva, Marco Costa
IST and INESC-ID[2]

ABSTRACT
We describe the design of an affective control interface, SenToy, a doll with sensors that allows users to control their avatars in an adventure game. A Wizard of Oz study was used early in the design process to find the best relationship between user movements of SenToy and the resulting affective expressions and movements of their avatar on the screen. The results from the study showed that there are gestures that most users will easily pick up to express emotions. It told us which dimensions of movements (distance to screen, movements of limbs, etc.) that work best. We describe in what way the results from the study have affected the design of the SenToy and the hardware requirements. Wizard of Oz studies have previously been used for natural language interface and intelligent agent design and we show that it can effectively be used also in the domain of affective input-device design.

Keywords
Affective interaction, Wizard of Oz study, User evaluation, Tangible user interface

INTRODUCTION
Gaming is a highly relevant and exciting area to explore for Interactive Characters and Affective Computing . Using an affective input device in the interaction with a game is a recent and fairly unexplored dimension and, to our knowledge, there is no other system like the SenToy. Designing such a novel interaction device is therefore quite difficult.

Our aim is to "pull the user into the game" through the use of SenToy. SenToy is a doll with sensors in its limbs, sensitive to movement and acceleration. The actions performed by players on SenToy are interpreted and the inferred emotion or movement action is used to make the players' avatars change facial expression or move. The avatar's emotional state, in turn, is crucial in advancing in the adventure game, FantasyA. Unless players can express the right emotional expression at the crucial point in the game, they will not be able to trade with characters in the game, gain magic gems, and thereby advance in the game.

SenToy is inspired by the work on tangible user interfaces where the goal is to "go beyond the current GUI (Graphic User Interface) and bring current objects to the interaction between users and computers" [6]. Following this vision, new types of interfaces have been built as ways of bridging these two worlds. In particular, for synthetic characters, the team at the MIT Media lab developed what is called a "sympathetic interface" [7], which can be seen as a "physical", "touchable" incarnation of the synthetic character. Such an interface allows the user to touch it and by its physical handling, influence the character's behaviour in the virtual world.

The second inspirational source for the SenToy development comes from synthetic characters. With the establishment of the area of affective computing, these characters are gaining emotional expressions and behaviours, which the user can controll or influence. In fact, some systems have already been developed where synthetic characters are "emotionally influenced" by the user. Examples of early commercial products are Active Worlds (www.activeworld.com) and ComicChat (www.microsoft.com) where the user can set the avatars facial expressions to reflect its emotional state. In ComicChat the avatar also automatically reflected some expressions derived from the text the user inputs. A question mark "?" in the text rendered a confused face. None of these systems allowed for tangible input.

In developing SenToy we had to tackle a major design issue: how can a user control the emotional state of a synthetic character using a physical interface, such as a doll? For example, what kind of gestures would be the best ones to express "anger" or "happiness"? Our first inspirational source is studies on gestures and emotions. These studies show that high arousal and high intensity is associated with arms [1]

[1] SICS (Swedish Institute of Computer Science)
Box 1263
SE-164 29 Kista, Sweden
{gerd, kia}@sics.se

[2] IST and INESC-ID (Instituto de Engenharia de Sistemas e Computadores)
Rua Alves Redol, 9, 1000 Lisboa, Portugal
{dario.mourao, ana.paiva}@inesc.pt Marco.Costa@taichung.inesc.pt

up, away from torso, whereas head down is associated with negative valence. However, how do users respond when controlling an avatar's emotional state through a doll? Do they follow a specific pattern of gestures? As the doll constitutes an intermediary between what the user wants to express and the resulting avatar expression, we cannot expect that users will mimic some kind of "natural" human bodily behaviour. Perhaps users are more influenced by how cartoon characters move, or preconceived ideas of emotions and bodily behaviours?

To get some input on these issues, we performed a Wizard of Oz study. In a Wizard of Oz study, the user is made to believe that they are interacting with a system, while in reality they are interacting with a human Wizard, sitting behind the screen pretending to be the system. The method can be useful in situations where the development of the system is expensive and where it is hard to know beforehand how users will behave as there are no "natural" behaviour experiences or given user tasks to build upon. The method has been used in the design of natural language systems [2] and intelligent agents [11]. One of the main advantages is that the method does not constrain users' possible expressions.

THE SENTOY CONCEPT

The SenToy is an affective physical interface to control/influence synthetic characters that inhabit a 3D role playing game named FantasyA. In FantasyA, characters are magi that live in a fantasy world. The game develops around the use of gemstones that influence the emotional state and personality of the magus. Thus, users must control the emotional states of their magus in order to develop in the game. The characters in FantasyA can act in the world by moving, manipulating magic gems, and interacting with others (be it trading or duelling). Each interaction depends on the emotions of the magi.

Users can perform two kinds of actions through the SenToy: *movement actions* and *emotion related actions*. Movement actions include walking, picking up items, and stopping. Emotion related actions concern making the avatar's emotional state change to exhibit one of six different emotions: happiness, sadness, anger, disgust, surprise and fear. We chose the 6 basic emotions (according to Ekman [5]) for two main reasons. First, people independent of culture recognize these emotions from facial expressions. Second, this set of emotions was considered as the minimum for the game to have enough variety so that the player can see the influence of different emotions in the development of the game. The initial identification of the gestures was inspired by research on emotion theories (in particular Ortony et al [12]; Lazarus [9] and Ekman [5]), work in body expression, such as the Laban theory [4] and on how cartoons move and express emotions [16].

The System SenToy

The architecture of the SenToy system consists of two main components: (1) the physical interface (the doll) and its acquisition module and (2) the affective user model component. The physical interface (the doll) transmits a set of signals, generated by its sensors, resulting from the interaction with the user. The acquisition module transforms these signals, so they can be easily interpreted and consequently used to infer the underlying action or emotion.

Emotion	Gestures	Reference
FEAR	Put SenToy's hands in front of its eyes or move it backwards vigorously.	According to Lazarus [9] fear is associated with avoidance.
ANGER	Place SenToy's arms crosswise or shake it vigorously.	According to Lazarus [9] anger is associated with the "tendency to attack".
SURPRISE	Open SenToy's arms backwards inclining its torso slightly backwards.	According to Laban [4] surprise is associated with attention and with a sudden event and inclination of torso backwards [3].
SADNESS	Bend down SenToy's neck or bend down the entire torso.	According to Scherer [14] sadness is expressed through slow movement inwards and head down.
HAPPINESS	Swing SenToy (make it dance) and/or play with its arms.	Joy is portrayed with open arms, movements such as clapping or rhythmic movement according to Darwin[3]
DISGUST	Move SenToy slightly backwards (squeezing it slightly).	According to Lazarus [9] action tendencies for disgust include "move away", nausea and even vomiting.

Table 1 Initial mapping between emotions and gestures used as a first approach for the acquisition rules

RELATED WORK ON AFFECTIVE INPUT

Affective computing [13] is a growing research area where recently there have been some developments in the area of affective input. There are basically two lines of work on affective input: inferring users' "real" emotional states, and allowing users to manipulate the emotional state of their representation (be it an avatar or user model). It is mainly the latter that interests us here.

One problem with controlling the emotional state of an avatar is the amount of work it requires of the user (which often comes in addition to other user tasks such as communicating with others or playing a game). Antonini [1] therefore propose solutions attempting to find a

Movements	Gestures
WALK	Swing the legs forward and backward alternately or moving the SenToy forward with small jumps.
STOP	A bouncing jump vertically.
PICK	To bend down and move the arm like it was picking something or moving the SenToy like it was diving.

Table 2 Initial mapping between movements and gestures

balance between character autonomy and user control: "semi autonomous avatars". Neither of their solutions solves the problem at hand as the goal of the SenToy is to be part of the game's challenge, thus mastering it and taking control over the avatar is crucial for the player..

A few systems have explored affective input devices similar to SenToy. In Swamped! [7] the user takes on the role of a chicken that is trying to protect its eggs from a hungry raccoon. The chicken has a set of behaviours such as squawking to get the racoon's attention and make him angry, scratching his head, or kicking him. The user partially controls the synthetic character in the Swamped! virtual world through a set of gestures with a plush toy. The control of the character is done at a behavioural level where sets of predefined behaviours (such as walk) are associated with patterns of gestures and handling of the doll. With SenToy we also want to express emotional states and not only behaviours.

Martinho and Paiva [10] allow users to influence the emotional state of a synthetic dolphin (Isolda) through touching a set of sensors placed in a porcelain dolphin. Also, in a recent art installation created by Tosa [17], users can influence the behaviour of two synthetic mermaids through their hand gestures and physiological signals such as the heart rate. Each mermaid will move in sync with the user's heart rate (which is captured through an electrode attached to the user's collarbone). Finally, the Affective Tigger [8] is a toy that reacts emotionally to the emotions of the user, though the Tigger does not control a synthetic character in a virtual world.

INFORMING DESIGN THROUGH WIZARD OF OZ

We needed to know exactly which movements and behaviours of the doll users will pick up "naturally". Building the whole system with sensor recognition and affective interpretation before having any input on what will most probably work best, is a waste of time and energy. This is why we picked the Wizard of Oz method.

Expectations for the Wizard study

We anticipated two reasons why the rules from Table 1 and Table 2 above would not necessarily work. First, users will not behave in the same way when expressing emotions through a doll rather than through their own bodily behaviours. There are numerous reasons for this, among them the cultural notions of how dolls and cartoon characters behave when expressing emotions. Secondly, we needed to put users in a loop where they are given feedback from the system through how the avatar reacts. Users will *learn* how to create the right behaviour through watching the face of the avatar on the screen when they perform actions on the SenToy. Thus there is room for "unnatural" learnt behaviours, but we have to make sure that the expected behaviour does not defer too much from the movements of the doll that the subjects use.

Concerning the doll, we wanted to know which kind of doll and properties of the doll would be best suited to expressing emotion-related related and motion actions. Three dolls were selected varying in size, shape, softness, and in being a puppet or just an ordinary doll.

STUDY

We used eight subjects (from Sweden) in this study, four males and four females, between 14 and 30 years old.

We used three different dolls, a puppet, a Barbie doll and a teddy bear. All subjects tested two different dolls. They were told that the doll contained sensors, whose signals were picked up by the video camera, placed in front of them and that the sensors reacted upon movements and distance from the video camera. They had a computer screen placed to the right, where they could see the avatar and they were told that they controlled the avatar through moving the doll. The subjects were asked to express six different emotions and perform three actions, one at a time, through moving the doll. The Wizard had a set of control buttons that she could push to make the avatar express a particular emotion or movement action, in the case that the subjects had moved the doll in the correct way. When the subjects moved the doll in another way, she had to make the avatar express the wrong emotion. If the movement performed was not on the list, the avatar was made to express a neutral facial expression.

Subjects finally filled in a form, providing feedback of which doll they preferred to use.

The Setting

The subjects sat in a chair with the video camera placed at a distance of 2 meters from the chair. The screen with Papous was placed to the right (see Figure 1). The chair was placed so that the subjects could see both the screen and the camera if turning their heads. Thus, subjects would turn to the camera to do the doll movement properly, and then they would turn towards Papous. The Wizard sat behind a desk beside the camera.

Software Description: SenToy Alternations for Wizard of Oz

In order to perform the experiment we developed software that allows the Wizard to control the avatar. We adapted a system that contains a synthetic character that conveys emotions through gestures, facial expressions, voice and text named *Papous*. The character used is a granddad that tells stories (taken from another project)[15].

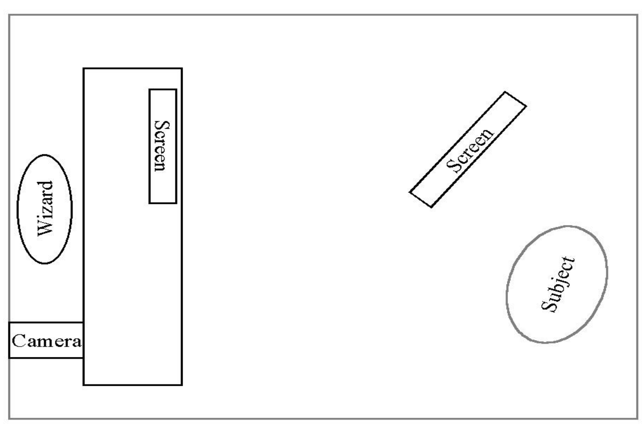

Figure 1 Setting of the room for experiment

Figure 2 Facial expressions of Papous.

The animations of the emotional states are expressed using both vocal and facial expressions of Papous. The vocal expressions are also part of the animations of the emotional states. The speech is generated through an affective speech component that receives sentences and the current emotional state and synthesizes the sentences using the voice to express the current emotions. We used a TTS system from Eloquent (www.eloq.com).

The Wizard Task

The Wizard had a set of control buttons that she could push to make the avatar express a particular emotion or perform a movement action. The Wizard's task was quite difficult, as she had to recognize whether, unintentionally, the subject had expressed another emotion than the one asked for. In such a case, she had to make the avatar express the wrong emotion. When the subjects moved the doll in a way that was not on the list of the suggested gestures, the Wizard pressed a neutral button, which gave the avatar a neutral facial expression (see Figure 2).

The Dolls Tested

We used three different dolls in this experiment: one teddy bear, one Barbie doll, and one puppet. Three dolls were selected varying in size, shape, softness, and if it could be used with only one hand.

The teddy bear was quit soft and the subject could move his arms and legs. It was about 30 cm high and had a neutral facial expression. The Barbie doll – Super Model Kenneth™ – was a traditional male Barbie doll made of plastic. It was possible to move his arms and legs. Since the Barbie doll's hands and feet were to small to contain a sensor inside, we had taped "sensors" on his hands and feet so that the subjects would get the impression that there were actually sensors on the doll.

The Puppet was about 15 cm high. It was possible to put ones fingers inside the dolls arms and then move his arms around. There were no holes in the doll's legs, so the subjects could not move the legs. Also this doll had "sensors" placed on its hands and feet to give the subjects the impression that there were sensors on the doll.

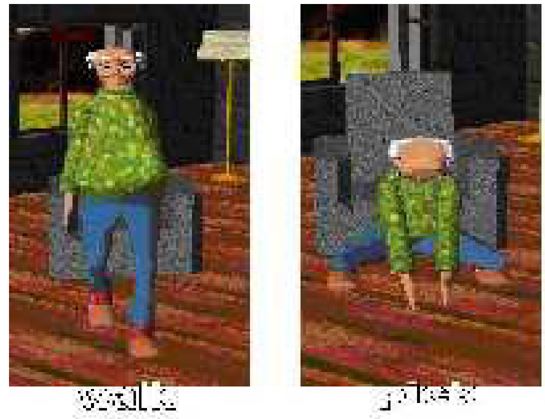

Figure 3 Papous walking and picking up an item

RESULTS

We divide the results obtained in the experiment into the following categories: the dolls tested and portraying emotions through gestures.

Dolls Tested

Subjects preferred dolls with a neutral facial expression, as it was easier to portray emotions other than those expressed by the doll's facial expression. This was foremost a problem in the case of the puppet, which had a happy face (see Figure 4). The Barbie doll and the teddy bear had fairly neutral expressions.

It was also important that the doll was soft enough to be bent in various directions. This was a crucial problem with Kenneth who was made of hard plastic. Even if it was possible to bend his limbs, it was hard to portray continuous movements back and forth. When using the Kenneth doll, the subjects were quite absorbed with moving his arms and legs, and therefore had difficulties focusing on Papous and his appearance on the screen.

The size of the doll was also deemed important. If the doll is too small, the movements become hard to perform.

Portraying Emotions

The Wizard knew which emotion the subjects were trying to express, but in many cases the movements were hard to distinguish. For example "happiness" and "anger", make the doll's limbs move in a similar way, and for a human wizard it is very hard to recognize the

Figure 4 The teddy bear, puppet and Barbie doll.

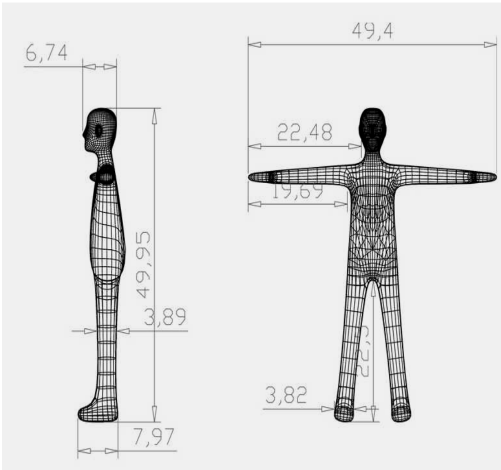

Figure 5 Model and dimensions of SenToy

difference in speed. In this experiment this didn't really matter, because we only wanted to know how the user would naturally move the doll in order to express an emotion and if the way they moved the doll corresponded with the initial mapping. The Wizard was helped by the fact that she knew which emotion they were trying to express, so that she could initiate the correct avatar expression.

The results showed that all movement actions and three of the emotion related actions (happiness, surprise and sadness) were expressed according to the rules suggested in Table 1 by all of our subjects. When expressing anger most of the subjects where boxing with the doll's arms instead of shaking the doll as suggested in the table, but this movement made the doll's limbs move as it where shaken. To express fear half the subjects put the doll's hand in front of its eyes and the rest turned the doll away from the camera. All subjects experienced the emotion "disgust" very hard to express. None of them moved the doll backwards as suggested, instead most of them stroke the doll's hands in front of its eyes as wiping something away and the rest made a "vomiting" gesture.

The designer had envisioned that the subjects would use both gestures and distance from the camera to express an emotion. Even though subjects were told that the sensors reacted to movement and distance from the camera, they did not make use of the distance dimension. This might be a consequence of the experiment setup, where subjects were sitting in front of a computer, rather than standing up. It was therefore not really natural for them to move the doll backwards.

When expressing the emotion related actions the subjects often performed two types of movements at the same time or tried a set of movements before the system reacted. This could cause trouble using sensor recognition. The expression for sadness, for example, might get mixed up with the expression for fear in the final SenToy system, since subjects in the study put the doll's hands in front of its eyes in both cases. The design envisioned was that they would only do this when expressing fear. In the same way happiness could get mixed up with surprise, even if all of the subjects expressed both emotions as suggested. In both cases, subjects would open the arms backwards. To express surprise, the movement would stop in a frozen position when the arms where extended, while happiness would be expressed by continuously moving the whole doll back and forth or its arms back and forth. These two emotion expressions would more easily be distinguishable if the sensors were able to recognize when the movement stops in a frozen position.

NEW SENTOY DESIGN

The results of the study influenced significantly the design of SenToy. The impact can be examined in several different categories: the type of the doll, its "look and feel"; the gestures to be identified as ways of expressing emotions; the hardware needed for the capture of gestures; and the impact that this affective control will give in the game and its dynamics. Here we will describe the impact the results had on the hardware chosen and the look and feel of SenToy.

The hardware of SenToy and gesture interpretation

Given the gestures chosen, the hardware of SenToy was designed with a set of sensors connected to a micro controller through a digital or an analogical interface. The sensors used include two magnetic switches to detect the movement of putting the hands in front of the eyes. These sensors work by the proximity of a magnetic field (created by a magnet placed in the hands of the toy) closing a switch and establishing a connection between the two terminals. This means that whenever the user places the hands of the toy in front of the eyes that gesture is recognised. To detect walk, Piezoelectric Force Sensing Resistors are used in the arms and legs together with a plastic structure that allows for detecting a bending and movement with the legs.

Finally, to detect all movements with the torso of the doll, we used accelerometers, which provide a measure of the acceleration in one or two axes. The accelerometers also allow for obtaining a measure of the vibration of the doll. This will allow for distinguishing movements from one another, bending, turning away from the screen, etc. It will also allow for the distinction of rapid and slow movements, which will be extremely important to distinguish between certain emotions. In the cases of confusion, like in the gestures Fear

Emotion	Most common action	No	Second most common action	No
Anger	Boxing with its arms	12	Shake the doll	6
Fear	Hands in front of the eyes	8	Turn the doll away from the camera	7
Disgust	Arm in front of face as if wiping something away	10	"Vomiting"	4
Happiness	Dancing/jumping, continous movement	16	Arms in the air, waving them back and forth	11
Sadness	Bending down its trunk	16	Hands in front of the eyes	8
Surprise	Arms in the air, frozen position	16	Lifting the doll upwards into a frozen position	4

Table 3 Gestures used to express emotions 16 is maximum, 8 subjects, two dolls each.

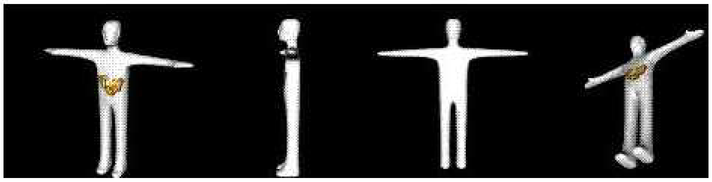

Figure 6 SenToy 3D model

versus Sadness (where subjects have put the hands in the eyes) the timing of the movement is also detected, and according to the results, this hardware can make a distinction between these gestures. Happiness versus Anger is the other case where confusion may occur (see Table 3), but again we can detect the bouncing of the body versus the boxing of the arms solely (without dancing) with the accelerometers.

The type of doll

According to the results, users preferred the soft and cuddly toy prior the hard plastic one. Our design is to create the doll's exterior from soft fabric, filled with latex and covered with a kind of skin made with synthetic nape. Its material and construction must take into account: production (allow for opening the toy) and also for haptic feedback from the toy – important areas are marked with a softer material. As for the anthropomorphism of the doll, it should only be partially anthropomorphic (to allow for legs, face and arms) with reference points to the avatar appearance in the virtual world. As suggested by the subjects, we decided that it should not have a defined facial expression, character or personality so that it will not constrain the manipulation and its adequacy to different synthetic characters and different emotions. Its minimalist appearance will allow for the same toy to be used for controlling many different types of characters (if needed, a symbol like a different logo in the trunk of each produced doll) can be added for each character). Given that the arms and legs of the toy were quite important for the gestures, arms, legs and torso should allow for the amplitude of movements needed so are long enough for an adult hand to hold (see Figure 5). Also, they should not be constrained by the sensors' position and the support structure that should not affect too much the flexibility of the movements. Given that FantasyA is like an adventure game with emotions, the image of the toy should be somewhere in-between "a toy" and "an electronic gadget". Figur 6 shows the final design of SenToy.

The set of emotions and associated gestures

Given the results from the study and taking into account the requirements or the game, some of the emotions considered (disgust in particular) need to be rethought in the context of the game. Further, although the tests were performed using the six emotional states, other emotions, more social ones such as pride, shame or envy are also quite important in the context of FantasyA. We expect now to rethink the set of emotions to be captured and evaluate the possibility of including such type of emotions. Note that in such cases, the gestures may have to be mainly "learned" and perhaps less natural than the gestures we have discussed so far.

CONCLUSIONS

The Wizard of Oz method provided valuable design input on: which doll to choose; the orientation of the player, game and sensors; which movements and dimensions of movements that work best; and, finally, which emotions that are most easily expressed. The extra work of implementing the Wizard of Oz setup for the experiment, perform the study and analyze the results took approximately two weeks. We consider this to be a cost-efficient method for testing the system.

ACKNOWLEDGEMENTS

Thanks to André Silva and Marco Vala for the development of Papous. Thanks to all our partners in the Safira project for their comments and criticisms on the SenToy interface.

REFERENCES

1. Antonini, R Implementing an Avatar Gesture Development Tool as a Creative Evolutionary Collaborative System, *In Proceedings of AISB'99*, Edinburgh, 1999
2. Dahlbäck, N., Jönsson, A., and Ahrenberg, L. Wizard of Oz studies—Why and how. In *Proceedings of. ACM International Workshop on Intelligent User Interfaces* ACM, 1993.
3. Darwin, C, 3:rd ed. by Paul Ekman, *The expression of emotions in man and animals*: Oxford: Oxford University Press. 1872/1998.
4. Davies, E. Beyond Dance: Laban's Legacy of Movement Analysis, Eden Davies; Seven Locks Press, 2001.
5. Ekman, P. *Emotion in the Face*, New York, Cambridge University Press, 1982
6. Ishii, H. and Ullmer, B. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms, *In Proc of CHI'97*, ACM Press, 1997.
7. Johnson, M., Wilson, A., Kline, C., Blumberg, B., Bobick, A. Sympathetic Interfaces: Using an Plush Toy to Direct Synthetic Characters, In Proc of CHI'99, ACM Press, 1999.
8. Kirsch, D. The Affective Tigger: A study on the construction of an emotionally reactive toy, TR-496, MIT Media Lab, 1999.
9. Lazarus, R. *Emotion and Adaptation*, Oxford University Press. 1991.
10. Martinho, C. and Paiva, A. Pathematic Agents, *Proceedings of the Autonomous Agents'1999*, ACM Press.
11. Maulsby, D., Greenberg, S. and Mander, R. Prototyping an intelligent agent through Wizard of Oz. *In Proceedings of the CHI'93,* The Netherlands, ACM Press, 1993.
12. Ortony, A., Clore, G. and Collins, A. *The Cognitive Structure of Emotions*, Cambridge Univ. Press, 1988.
13. Picard, R. Affective Computing, *MIT Press*, 1997.
14. Scherer, K. Emotion effects on voice and speech: Paradigms and approaches to evaluation, presented at the ISCA Workshop on Speech and Emotion, Belfast 2000. Available at: http://www.qub.ac.uk/en/isca/proceedings/pdfs/scherer.pdf
15. Silva, A., Vala, M. and Paiva, A. "Papous: The Virtual Storyteller", In *Intelligent Virtual Agents,* Eds António, A., Aylett, R. and Ballin, D. , Springer, 2001.
16. Thomas, F. and Johnston, O. The Illusion of Life: Disney Animation, *Walt Disney Productions*, 1981.
17. Tosa, N. Unconscious Flow *in SIGGRAPH'99 Art Show*, ACM Press.