Towards an interactive Virtual Storyteller

Sybren Wijnia
Faculty of EEMCS, University of Twente
s.wijnia@student.utwente.nl

ABSTRACT
Good storytellers must respond to the audience's feedback to be more entertaining. The public's feedback can be given at the level of *what* is being told (story level) en *how* the story is being told (presentation level). These two levels of feedback must be incorporated in the digital storytelling system 'Virtual Storyteller' by including at least one type of interaction. As an onset for this implementation this article describes what the interaction possibilities are to create an interactive Virtual Storyteller.

Keywords
Digital storytelling, Interaction, Virtual Storyteller.

1. INTRODUCTION
A good human storyteller knows how to entertain his listeners. He can give them the experiences of joy, rapture and enlightenment [Hay98]. A storyteller not always tells stories in the same way. The public gives the narrator feedback during the process of storytelling. The storyteller responds to the reactions by adapting his story, which means that the public really has some influence on the development of the story [Sil03]. In my opinion the storyteller responds at two levels. First, on the level of *what* is being told, and second on *how* the story is being told.

Hayes-Roth [Hay98] gives in her article a nice example of the *course* the story of Alice in Wonderland takes during the telling process. She shows how Lewis Carroll, the author of Alice in Wonderland, presumably has developed his story through interaction with three of his students. He immersed his students in the story by representing each of them as a character in it. Carroll continuously acted upon the questions, suggestions, demands and commands of his students. Often he gave them choices, for example: Should Alice follow the white rabbit or not?; Does Alice need to drink the liquid from the bottle that says "Drink me"?; Is the Caterpillar really trustworthy and thus must Alice follow his advice? If the story drifted away from the main plot, Carroll let the white rabbit appear who brought the story back on track. Through all these interactions with his students he created a story aimed at their interests and wishes. He let them mainly influence the *what* of the storytelling process.

On the other hand, also the *how* of storytelling can be altered. For instance, a mother, telling a story to her children, is asked by her offspring to tell the story more scarier than she does initially. The content of the story can remain the same, while the facial expressions, gestures and tone of the voice of the mother may change. Or in another case, parents giving a reprimand to their child with a normal attitude or one with an angry point of view. These examples show that *what* is said remains the same, but *how* it is said differs. Even the impact of the reprimand on the children will differ. The latter style will probably have more influence.

The question is how to map the two levels of interaction to the digital environment. I have applied this to a specific system (the Virtual Storyteller) to give my research more focus. In the second section I will define some relevant terms, after which the need for this research is described in section 3, whereas section 4 gives an overview of the Virtual Storyteller. Section 5 shows how interactivity is shaped in existing storytelling systems on the story level and on the presentation level. After which section 6 deals with different modes of presentation. This paper concludes with a discussion and a conclusion.

2. DEFINITIONS
It is difficult to draw a line between what can be denoted as a story and what not. It is even so that in most articles about digital storytelling it is left open what is meant by a story. Since this paper is not about how to define a story I think it is sufficient to remark that a story must be seen as a succession of events organized by time sequence with some dramatic compression [Hay98, p18].

In the manner of this paper storytelling is seen as the reciting of a story by a human via voice or writings on paper. Digital storytelling or digital story creation is the autonomous creation of a story by a computer program with or without some influence of the audience (; the audience as a spectator vs. participant [Pre03]). When this digital storytelling is presented in a 2D or 3D virtual environment with virtual actors, then it is called virtual drama. If the audience has influence on the direction which the story takes then it can also be entitled as interactive drama.

3. RESEARCH NEED & PROBLEM STATEMENT
In the previous part I have shown that storytellers need feedback to be more entertaining and that this interaction process is possible at two levels, the presentation level and the story level.

Magerko [Mag02] proposes an interactive drama architecture in which there is a balance between writer flexibility and user flexibility; a balance between total plot specification by an author and emergent plot. Yet, the architecture is in its early phase. It needs to be implemented/developed any further to be of importance here.

Swartout *et al.* [Swa01] on the other hand, have built a more sophisticated first person prototype for military training in which some interactivity is possible through speech recognition. A participant can instruct its virtual soldiers to perform a

3rd Twente Student Conference on IT, Enschede June, 2005

specific action. In a future version the designers want to add a director agent to the system to have overall control of the simulation/story.

This kind of director agent is already part of the architecture of the Virtual Storyteller of Theune *et al.*, but is barely implemented [The04]. This storytelling system does not incorporate any interaction possibilities, and is therefore a fine subject of research. Adding interactivity to the system will boost its entertaining power. As an onset for this process I have researched what the interaction possibilities are to make the Virtual Storyteller interactive. The implementation of any interactivity will be part of future work.

Thus I have formulated the following problem statement:

"What are the interaction possibilities to create an interactive Virtual Storyteller?"

4. VIRTUAL STORYTELLER

Theune *et al.* [The04] have created a multi-agent framework for automatic story generation, which is aimed at creating short stories in the fairytale domain. The architecture of the Virtual Storyteller consists of a director agent (to guarantee well-structured plots and to guide the story characters), semi-autonomous character agents (for plot creation), a narrator (to offer natural text to the presenter), and a presenter (who presents the story to the user, see figure 1). This architecture is shown in figure 2 where it is combined with the idea of the two different levels of interaction.

Figure 1. Presenter of the Virtual Storyteller [The03]

The characters and the director are of course situated on the story level. The narrator is placed on the presentation level since it can tell the same thing in different ways, i.e. the narrator has influence on how the story is being told. However, the narrator lies outside the focus of this research, as I am not examining different narrative styles. Finally, it is obvious that the presenter can be found on the presentation level.

The first implemented version of the Virtual Storyteller is aimed at creating consistent and well-structured plots. Although these two requirements are necessary for achieving a well-formed plot, alone they do not make a story entertaining. An additional property like credible characters is added in the latest version. To be able to show later on where incorporation of interaction possibilities are conceivable I will give here some insight in the story creation process of this system.

Semi-autonomous character agents are responsible for the plot creation under the supervision of a director agent. The director agent uses knowledge about the virtual environment, the actions certain characters can take in it and general knowledge about what makes a 'good' plot to judge whether a character's intended action fits into the plot structure. The director has a few control methods at his disposal: He can introduce new objects and characters into the environment, he can give a character a goal to pursue or disallow a character's intended action. What he cannot do is force a character to perform a specific action [The02][The03].

In the Virtual Storyteller characters are made credible by providing them with emotions. This is done through representing an agent's emotional state by pairs of corresponding positive and negative emotions and their intensities. This state changes in reaction to events, actions and objects. The intensity and duration of emotional effects within an agent depend on the personality parameters of the agent, which can be changed by the user. Based on its emotional state, a character may develop certain action tendencies, which in their turn influence the importance a character attaches to certain goals [The04].

A story generated by the Virtual Storyteller consists of four episodes, which are selected from a database. Each episodic script contains information about the setting, goals and constraints of an episode. These episode properties define the boundaries within which the character agents are allowed to act while constructing the episode. The actual episode is created by the characters carrying out actions to pursue their individual goals (episodic goals or goals based on the character's emotions). Their actions can be shouting, walking, kicking, picking up or throwing away an object, or using an object. The global plot structure is predefined, yet there is variety possible. There are many different ways in which an episode can evolve before an episodic goal is reached. The 'emotional goals' take care of the plot diversity [The04].

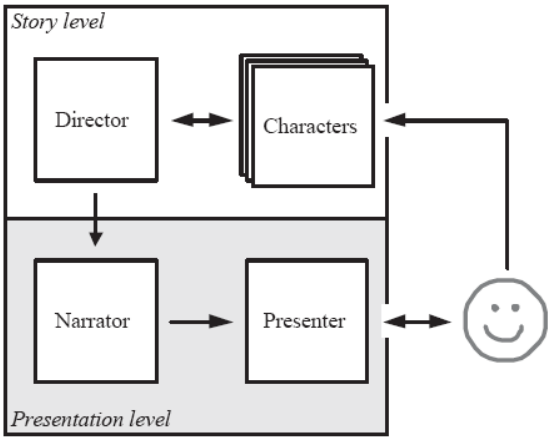

Figure 2. Architecture of the Virtual Storyteller combined with the two levels of interaction

5. INTERACTIVITY IN EXISTING SYSTEMS

This section shows how interactivity is shaped in existing storytelling systems. The first section deals with interactivity on the story level, the second on the presentation level. Each subsection gives a description of a particular storytelling system and its implications for the Virtual Storyteller.

5.1 Interactivity on Story Level

At the story-level I have examined systems (almost) similar to the Virtual Storyteller. One of these, Teatrix, also puts together a director agent and a character-based approach of story generation [Mac00]. Another, the Oz project uses 'drama manager' for well-structuredness and consistency [Mat99], and

Cavazza *et al.* [Cav02a] are using the character-based approach to develop a story.

5.1.1 Teatrix

Teatrix [Mac00] is a system developed to support collaborative drama and story creation by children. The children will use a set of predefined scenes and characters to create a story. The character agents are either system-controlled, and thus autonomous, or they are child-controlled.

Children can prepare the scenes, props and characters for every story in advance. The next step for each child is to choose a character, which they control. This character is then commanded during the entire process of story creation. The basic commands which can be given to the characters are: walk, get item, drop item, use item, activate item, interact, and talk.

The real acting in a collaborative 3D world can start after the characters have been chosen. In this virtual environment (see figure 3) children need to work together to create a story. This is done through interaction with other characters and the use of objects.

Figure 3. Interface of Teatrix [Mac00]

Just as in the Virtual Storyteller, the character agents are equipped with a range of possible actions, goals and emotions. These are all situated in the artificial mind of the agent. When a child controls a character agent, the mind has a more passive role, i.e. the child can control the actions a character takes. Another similarity is when it comes to the director agent. The director agent is responsible for narrative guidance of the story and is capable of inserting new characters and items, and gives other characters orders. He can even control them when needed. However, in order not to restrict the children's creativity, the director must not control their characters, but must play a guiding role in the story creation process.

The results of a user test revealed that children "should be given more control of the characters, in particular of their state of mind (and not only actions)"[Mac00, p.116].

The main differences of Teatrix, compared with the Virtual Storyteller, are that in the first "the character and director agents function as aids in the children's story creation process, rather than creating the story by themselves" [The04]. There is also interactivity, which is missing in the Virtual Storyteller.

The results of applying the interactivity mechanisms of Teatrix to the Virtual Storyteller are that the user can be involved in:

- Choice of the (four) necessary episodes to construct the story. Story coherence can be maintained by showing only relevant options to the user during the choice process.
- Preparation of props.
- Selection or generation of the (main) characters. Selection means choosing from a predefined set of possible characters, generation means that the user can adapt the appearances and clothes of the characters, or even choose the emotions and their intensities of the actors. Via this means it even should be possible for the user to create a character resembling himself, just like Carroll Lewis did by representing three of his students as characters in his famous story 'Alice in Wonderland'.
- Controlling a protagonist or other kind of character. This can vary from choosing where to go, what interactions to be involved in, what actions to take, to how to react to a specific character e.g. angry, irritated, or lovable.
- Inserting new characters or items.

5.1.2 OZ Project

In the OZ-project [Mat99], a drama manager is used for well-structuredness and consistency. It ensures that some essential points of the plot occur during the story creation process. The drama manager takes the history of a story into account and then calculates at a plot point every possible future. The direction of the highest ranked future is then taken. Unfortunately Mateas [Mat99] describes in his article mainly believable agents. He doesn't show how these agents create the story together and in what kind of environment, he only mentions they are under supervision of the drama manager and that they have to meet the following set of requirements: personality, emotion, self-motivation, change, social relationships, and illusion of life.

This implies for the Virtual Storyteller that the user could be able to:

- Define how a character talks (e.g. dirty or polite), with what voice he talks (e.g. masculine or feminine), how he moves (e.g. striding or running), and how intelligent he is. This idea is based upon the personality requirement introduced by the OZ-project, and is most likely not part of OZ.
- Define relationships between certain characters before the story starts. The type of these relationships can be for instance friends, lovers or foes.
- See the characters change during time and in reaction to certain actions charged by the audience. For example, if a character gets punched in the eye, then it must turn black. Others can then ask the character what has happened with his eye. (Actually this point should be the result of choices of the user, i.e. the result of interactivity.)

5.1.3 Friends System

Cavazza *et al.* [Cav02a][Cav02b] have created a 3D virtual environment inspired by the popular TV sitcom "Friends". Therefore I will call their system the 'Friends system'. This system aims at generating comic situations. The stories start with a basic scenario which dynamically evolves through the interaction between the characters, whose behaviours are implemented in the system by using 'AI planning techniques'. Instead of using a control agent who takes care of a consistent plot, Cavazza *et al.* believe their solution of describing characters' behaviour in terms of roles is adequately for narrative control. This can be right because the behaviours are laid down in hierarchical task networks (an HTN is a sort of goal tree with a main goal, which is decomposed in sub-goals,

sub-sub-goals, etcetera) which are predefined. Thus, in the end the designers take care of consistent character behaviour. The freedom of the character agents lies in the fact that they can choose which subordinate goals they pursue to achieve their main goal.

Figure 4. A story instantiation generated by the Friends system [Cav02b]

The approach taken by Cavazza *et al.* [Cav02b] is based on the user-as-spectator paradigm. The user is shaped as an invisible avatar and is able to wander around and see the story unfolding (see figure 4). Users cannot physically interfere with the characters, for instance by denying them access to a room. Conversely, the audience is capable of removing or changing the location of objects and is able to provide the characters with information.

In the given example in [Cav02a] the character Ross wants to date Rachel. To realize this goal he first wants to acquire some information about Rachel. Therefore he wants to steal her diary. A friend of her may be in Rachel's room to prevent Ross from getting it. Then Ross needs to re-plan a solution for the information problem. The real life audience has some influence on this. They should be able to decide where to locate all the characters in de virtual environment and they can, for example, place the diary somewhere else than in Rachel's room or just keep it for themselves.

The other way of interaction is the use of natural language. Through the use of speech recognition the audience can provide true or false information to a certain character, instruct or warn him and give generic advice to behave somewhat differently. The audience can for example say "Ross, don't let Rachel see you with Phoebe or "Ross, don't be rude". The system can more or less understand it as long as the name of the instructed character is mentioned first.

From the description given above, the following interactivity options for the Virtual Storyteller can be elicited in which the user can:

- Determine the role a character takes, for instance the role of problem solver, quarrelsome person, or disciple.
- Place the characters and usable objects in the story world before the beginning of the real story creation process.
- Walk around in the environment as an invisible avatar.
- Use speech recognition or keyboard input to provide information to a certain character, instruct or warn him, or to give him generic advice.
- Directly manipulate virtual objects, which could play a role in the story.

5.2 Interactivity on Presentation Level

As far as the presentation level is concerned I have inquired two systems: A system with a head only presenter and paper books. The first system has thus the same presentation approach as the Virtual Storyteller, while the paper books can roughly be seen as the total collection of the balloons of the presenter.

5.2.1 Granddad system

Silva *et al.* [Sil03] show by what means children can influence *how* the story is being told. This in turn also affects the story itself. In their system a 3D granddad (figure 5) is equipped with voice, gestures and facial expressions to tell stories.

Figure 5. Granddad [Sil03]

Darcy *et al.* [Dar03] subscribe to the importance of congruence of gestures and body posture if the granddad appears not only as a head but also with his torso and arms. If there is no congruence, the audience's feeling of realism is reduced extremely.

The audience of the granddad system [Sil03] can show the system that he or she wants the story to be told e.g. less frightening by inserting a special card into it. The system can adapt itself to this by letting the granddad express himself somewhat different. Unfortunately, this adapting can only be done by the system at the moment it picks a new StoryBit at a higher level to continue (or start) a story. To understand this I need to explain how a story is being created.

A human author creates story files in advance with a specific tool. At several levels he or she writes one or more pieces of a story (which are the StoryBits). A story told by the granddad starts with a StoryBit at the bottom level and evolves continuously to a StoryBit at the highest defined level. At each level one or more StoryBits are present. During the storytelling process the system decides which StoryBit to choose at the next level based on the omitting or occurrence of user input.

It is left a little vague what the difference is between StoryBits at the same level: Do the different StoryBits comprise really different story contents or are they needed just to represent the moods of the presenter in order to let the granddad express himself differently? I have assumed the latter option.

It is notable that this storytelling system is only equipped with one type of narrator, the granddad. It's a fact that people rather mingle with persons similar to them [Pru01]. When users of the system can choose between several storytellers, they may like the generated stories more because of the heightened affection with the narrator. Nevertheless, a granddad is an applicable choice for a raconteur.

Possibilities for an interactive Virtual Storyteller are:

- The user can choose between different predefined narrators, or can even choose their gender, looks and voice.
- The audience has an influence on how the story is performed by letting the narrator express himself somewhat different (e.g. happy, frightening, or brusque).

5.2.2 Books

Books can also be seen as a storytelling system, only they are not digital. In her article about narrative for children in a historical perspective [Mad03] Madej shows that a certain kind of interactivity is also possible in books. Pop-up, open-the-flap, and peep-through-the-hole books give children the opportunity to read a story somewhat different a couple of times. Actions taken by the children can be incidental or integral. The first case means that the interaction does not affect the outcome of the story, the latter denotes the contrary. "Anecdotal evidence shows that, whether incidental or integral to the story, children, young children in particular, dote on all of these interactive devices" [Mad03, p.12].

Up to here I have mainly described integral interactivity, but also the notion of incidental action can be very useful for the Virtual Storyteller. The generated stories can become more fun for young children and it doesn't take many extra resources to develop a story since incidental action is not very important for the story progress.

Implication for the Virtual Storyteller:

- The use of incidental actions. Recall that incidental actions are actions taken by the audience, which have an influence on what can be seen, but do not have any influence on the story outcome. Children can for example point at an object to see an animation in which that object plays a part.

5.3 Summary of the interactivity options

Every interaction possibility mentioned in the previous two subsections has its own properties. Some options are only feasible before the real storytelling begins, others only after the start, while some options can be executed at both moments in time. Also the level on which an option could have some impact may differ. The audience can via these options have influence on the story level (*what*), the presentation level (*how*), or both. The results are shown in table 1.

Table 1. Interactivity options related to time of execution and to the level which they have most impact on.

Choices/Inter-activity options	**Execution**		**User influence on**	
	Before	*During*	*What*	*How*
Episode choice	**x**		**x**	
Props	**x**	**x**[1]	**x**	
Generation of (main) characters	**x**	**x**[1]	**x**	**x**
Control a character		**x**	**x**	**x**
Personality	**x**	**x**	**x**	**x**
Relationships	**x**[2]		**x**	**x**
Role of characters	**x**		**x**	**x**
Initial positions of characters and objects	**x**		**x**	
Provide information		**x**	**x**	**x**
Manipulate virtual objects		**x**	**x**	
Choice of narrator	**x**	**x**		**x**
Influence on narrator	**x**	**x**		**x**
Incidental actions		**x**		

6. PRESENTATION MODES

I believe the interaction possibilities on the story level with the system of subject are constrained by its presentation mode. The current presenter of the Virtual Storyteller is only telling you the story, he doesn't show any objects.

So, direct manipulation of virtual objects which is possible with the 'Friends system' [Cav02], mentioned before, doesn't seem appropriate for the Virtual Storyteller right now. The audience of the Virtual Storyteller can't, for example, hide Rachel's diary from Ross. Yet, it's maybe possible to create the map of the story world and place the main objects somewhere on it. In the future the presentation mode might be changed. As a result there are more interaction styles possible.

The different presentation styles can be grouped in four top categories: audio, textual, presenter and visual.

6.1 Audio

Audio only presentation means that the users can only hear a voice telling the story to them. They do not see any visual images related to the story or the narrator. An advantage of this approach is that the user can choose just to sit back and listen, without doing anything else. On the contrary, interaction is more restricted than with other presentation mechanisms, because there is nothing to see. The choices the listener has must be explained to him. This will take a lot of time. To communicate with the system a good solution will be speech recognition. This supports communication within the same channel as is used for storytelling (rule of matched modality [Ree96]); it makes interaction easier. This actually only yields when the recognition performance is almost or really perfect. Bilici *et al.* [Bil00] state that when users work in a multimodal environment with speech recognition and typed input, they often prefer speech as input modality. If the performance of ASR (automatic speech recognition) drops, then the users will switch to input via keyboard because of higher efficiency.

When it comes to audio only presentation Mullenix *et al.* [Mul03, p.419] note that "It appears that, when listening to a persuasive appeal, female human speech is preferable to female synthetic speech and male synthetic speech is preferable to female synthetic speech." Thus system designers could try to incorporate a female human speech in their system. However, pre-recorded utterances are far from flexible as text-to-speech (TTS) speech synthesis systems [Swa01]. Current TTS lacks the emotional content that a human voice has, but its flexibility is

[1] Generation of props or characters that are inserted during the unfolding of the story.

[2] Relationships can change *during* the development of the story, but they are the result of actions of one or more characters determined by the audience or otherwise.

really needed in interactive digital storytelling. Therefore at least a male synthetic voice should be included in the system. This will make interacting with a digital storytelling system more attractive.

6.2 Textual

Another possible output is text only. In fact this is just the generation of stories which are published as text. During the story generation process the user can have influence.

When the complete text is bound together it can be called an e-book. Only the e-book cannot be interactive. An e-book combined with a read aloud program results in an audio book. An audio book is almost the same thing as described in the previous section, but without any interactivity.

What has to be kept in mind is that a textual representation of a story, just like audio only, has other demands than a visual one. Much more has to be described, because it is not visible to the audience. They have to visualize the surroundings, characters, and objects for themselves on the basis of the description.

6.3 Presenter

The narrators of the Virtual Storyteller [The03] and of the Granddad system [Sil03] fall under this category. The presenter of the first system is really basic; it is a standard Microsoft agent without any emotions with respect to its voice, facial expressions and gestures. The granddad is more mature, but is not capable of taking the audience's emotional reactions (i.e. the facial expressions or gestures) into account. Braun & Rieger's approach [bra04] is closely related to Silva's granddad. They have also created a human-like virtual narrator that uses facial expression and gestures to portray suspenseful situations. From their system evaluation results they concluded that the audience had more fun when the narrator could express himself when telling a story. Yet, they also have stressed the need for congruence between facial emotion and gestures.

The ideas of Waters *et al.* [Wat98] can be of use for taking the audience's reactions into account. They have made a prototype of a public kiosk interface, which uses simple color and motion stereo tracking to provide visual information for the kiosk and which is equipped with a synthetic agent like the granddad to engage in an interaction with multiple humans. The system is able to detect and attract people by recognizing them as humans and calling them. Through eye gaze the agent can communicate his focus of attention to the audience. The agent is further capable of recognizing arrivals and departures of users, and of identifying the position of current users. The agent must be able to allocate its resources to different people fair enough and to handle conflicting demands. In next versions speech understanding is added to the system in such a way that users can interact with the kiosk in a direct way.

In my opinion, determining (potential) users of a storytelling system can make the stories more entertaining. The emotions of the audience can then be taken into account, even as their position in the real world. The presenter agent can, for example, get annoyed when the audience is not paying attention to him. He may speak with a different tone, ask the users what actions a certain character should take, or ask them what else is so important to them. Also more active user involvement is possible by asking for instance the users to move to the left if they agree with a certain statement of the protagonist of the story, or to move to the right if they disagree. Facial recognition can be used to store user preferences too. Every time the narrator tells a story he learns more about the user and is thus probably better capable of increasing the user satisfaction.

6.4 Visual

The notion visual presentation represents pictures (like in a comic), or a 2D or 3D virtual environment inhabited by characters, who are able to use several virtual objects. Examples of the type last mentioned are the Friends system (see figure 4) and Teatrix (see figure 3).

In digital storytelling two main paradigms have emerged [Cha04]: the immersed user paradigm and the interactive TV approach. In the immersed user approach, or the "Holodeck™" approach [Swa01], the user is immersed in a virtual environment and has a real actor role to carry out. Together with the virtual actors on the screen in front of him he has to achieve certain goals. In the other paradigm, the user is a spectator with influence on the story being generated. The systems mentioned in section 5.1 made use of this paradigm. The first approach is actually an onset in creating the holodeck from the world famous series 'Star Trek'. Details about this advanced type of storytelling can be read in Murray's 'Hamlet on the Holodeck: The Future of Narrative in Cyberspace' [Mur97]. Here, I will just state that the only means of interaction in Swarthout *et al.*'s system [Swa01] is speech. The user as an actor can instruct or ask its virtual allies to do something.

The focus of Charles *et al.* [Cha04] is a combination of both approaches mentioned above based on the magic mirror model of Marichal & Umeda [Mar03]. Charles *et al.* have used Mixed Reality (MR) in which the user (as an actor) has a virtual alter ego. This means that the actions of the user in the real world are captured with a video camera and microphone, after which they are translated to the screen as actions of the virtual alter ego. By this is meant that the user is able to make e.g. a real pushing movement which is then translated to his virtual alter ego pushing another virtual character. Or the user can, via his virtual character, have a conversation with some characters. Users do also have direct impact on virtual objects, but there is a limitation: the transfer of objects from the real world to the virtual one and vice versa.

To end this section, I would like to note that the choice for a visual presentation style for the Virtual Storyteller has a serious implication in a sense that there is an extra problem. How to visualize a description of something? If the presenter tells the audience that a young boy is walking on a winding forest path from the castle to the lake in the centre of the forest, the user can visualize it himself, but what if the system has to show the scene to the user via a virtual environment? Where are the castle, the lake and the forest exactly located, how does the path go, what does the young boy look like and how does he walk? This kind of questions need extensive further research.

6.5 Add-on: A Physical Interface

Holmquist, Helander, and Dixon [Hol00] are using a slightly different definition of digital storytelling than me. All the story parts used in their system are pre-recorded and thus fixed; autonomous story generation is not possible. The only variety in their storytelling is the choice the user has to take of which story component to uncover. Yet, their concept of memory objects might be useful.

A memory object is a real physical object that plays a part in the story. The audience can show the object to the system to uncover more information about it. The user can for instance show a (bar coded) spoon to the system, as a result the system shows a scene from the perspective of the virtual spoon or shows what has happened with the spoon in the past [Hol00].

Another possible utilisation of memory objects would be to show a virtual character what to look for, or which object he really needs to use right now. This may heighten the user's involvement in the story. A serious drawback can be that much different objects are needed to create variety; do system designers really want to supply a whole bunch of story objects with their software (and hardware)? This disadvantage can be diminished by using a camera, not for facial recognition, but for object recognition. The audience is then able to use objects from their own environment.

6.6 Presentation mode conclusion

I have shown four different modes of presentation with their own possibilities and limitations. These different modes can also be mixed together in one system, so that the user can choose which kind of mode he wants or in a manner that the presentation styles are truly mixed. A useful example of the latter is the presenter who is able to let the audience hear an audio sample, or see a letter, picture or movie scene.

It seems that most possibilities of interaction, distilled from section 5, are suitable for each presentation mode, especially when deployed in advance. After all, a graphical user interface is generally used beforehand. Only the audio mode may not have that. Exceptions of the mentioned suitability for each presentation mode are:

- Direct manipulation of virtual objects. This manipulation is reserved for the visual and the textual presentation mode.
- Choice of a narrator, and influence on him. These are the only means of interaction which are not suitable for the visual style.

Regarding the contemporary developments in our society the presenter and the visual style are probably the most appealing presentation modes. Audio books do still exist, but are generally seen as a bit outdated and command interfaces, like the ones used in the first adventure games, are old fashioned. Children for example, choose in most of the times a computer game instead of a book when given a choice between them [Mad03]. The presenter and visual mode have common grounds with computer games and are thus attractive.

The previous subsections have also demonstrated the usability of techniques such as speech recognition, motion tracking, facial/object recognition, and memory objects. They (presumably) enable more entertaining storytelling.

7. DISCUSSION

From the previous sections appears that there exist many options to make the Virtual Storyteller interactive at two levels, at several times during the story creation process and for a number of presentation modes. This leads to various questions which need to be kept in mind by the designers of the Virtual Storyteller:

- What is our target group of audience? This determines what needs to be told, how things should be brought, the extent of details which should be presented and the vocabulary in which the story ought to be narrated. This should all be decided based upon user characteristics like age, gender, intelligence, and personality type.

  The current version of the Virtual Storyteller is aimed at generating fairy tales. These are initially meant for children, aside from the popularity of e.g. Harry Potter and Lord of the Rings among adults. Madej [Mad03] states that there is a need for more complex stories when children grow older. During this natural process they can understand more, and read and speak better. Toddlers who are even not capable of speaking should not use a system with speech recognition. Young children who cannot read are a plausible audience when using the presenter or audio presentation mode. Youths and adults will probably like the visual style or the presenter mode more as concluded in section 6.6.

- How much influence do we want the user to have on the story and when (in advance or during the story)? For example, only choice of episodes will have a huge influence on the main plot, but during the process of storytelling the influence of the user is reduced to nothing. A user in control of a character on the other hand, may have serious impact on the story progress, but not on the main plot. Users should have influence on the story both in advance and during the story. Influence even before the real beginning of the story will absorb the audience in it, as the influence during the story will maintain their interest.

- Which options of interactivity do we really want to incorporate in our system and can they coexist?

  The interactivity options should be limited when the audience is young. They should for instance only be able to pick a character, and choose yes or no as answers to certain questions. If stories become more complex (as the target users are older) more interactivity options, perhaps more complex, are possible. From my personal experience I would recommend the following options: Generate and control a character and provide information. When extending the current presenter of the Virtual Storyteller I would like to be able to choose the narrator and his mood and I would like him to take my expression and gestures into account.

- What mode of presentation do we want to use? This question is related to the previous ones and is already elaborated in the last part of section 6.

Even though I still believe there are two separate levels of interaction in storytelling, they are very much intertwined. The next two examples show (again) that the two levels can be independent from each other. A teacher is not satisfied with an assignment made by his student. He can tell him this by speaking with a normal voice or with a more angry one. In both cases the *what* will be the same, but the *how* differs ánd determines the impact his words have. On the contrary, it is easy to understand that the content (the *what)* of a story can be different, while speaking monotone (the *how* remains the same). However, regarding the outcomes of table 1, I think the influence of the user on the story should preferably be on both levels at the same time. Then there is a match between *what* is being told and *how* it is being told. I think this is more truthful. For example, in Teatrix it is possible to prepare the characters. Suppose we can choose the looks and clothes of the main character, which can be seen as user influence on the presentation level.

Now if our character looks like a rough, fat, angry drifter, then in most cases this character would be avoided in the real world. If this happened in our virtual world, this can be seen as user influence on the story level. Thus, the users choice of the looks and clothes of a characters influences the presentation and the story level at the same time. This seems realistic to me. In the afore mentioned example, the *what* is automatically adapted to the *how*. The reverse should also be possible. If a character is for instance really happy, then he should also express himself like that.

Although it has been claimed that digital storytellers should be interactive, there is still no solid evidence that the audience really wants this. Maybe they just want to 'sit back and relax'. As far as I know designers of storytellers have not evaluated their interactive systems with their previous versions without interactivity. However, there were no previous versions in most cases. (Braun & Rieger [Bra04] have only evaluated their system with an emotional narrator and their old system with a more monotone one, in both cases there was no sign of interactivity. And Callaway & Lester [Cal01] have evaluated the effects of natural language generation techniques on reader satisfaction, also no interactivity was present.) I believe system designers should not only evaluate the new features of their systems stand alone, but also with respect to the older version which had more or less interactive options. Through these kinds of user tests, the user satisfaction with more or less interactivity can be measured.

Designers should not only underpin possible variety in created stories through the use of interactivity with theoretical arguments. They should also do this with an evaluation of all the story variants which are of importance that a certain system is capable of generating. Just like Ong and Leggett [Ong04] are planning to do.

8. CONCLUSION

In this paper I have presented several possibilities for implementing interactivity into the Virtual Storyteller varying from choosing the episodes of a story, to controlling characters and manipulating virtual or even real objects. Choice of a narrator, or determining a character's personality are probably easy to implement, because these are already more or less part of the system. Also taking over some roles of the director should not be that difficult. Trying to encompass a virtual environment and its related interactivity options in the system will take more effort (, though the presenter agent will not be necessary then).

It may be clear by now that implementation of certain interactivity possibilities will be part of future work, where designers also cannot refrain from user tests. Finally, other types of systems need to be investigated, i.e. systems that are not similar to the Virtual Storyteller and systems that have little to do with digital storytelling in general. They may give rise to innovative ideas concerning possibilities to make the Virtual Storyteller interactive.

ACKNOWLEDGEMENTS

I would like to thank Mariet Theune and the students of track C (Intelligent Interaction) of the third "Twente Student Conference on IT" for their helpful comments and suggestions. I also would especially like to thank Lieke Wijnia who volunteered to review this paper.

REFERENCES

[Bil00] Bilici, V., Krahmer, E., Riele, S. t. & Veldhuis, R. (2000). Prefered Modalities in Dialogue Systems. In: *Proceedings ICSLP 2000*, Beijing, Vol. II, pp. 727-730.

[Bra04] Braun, N. & Rieger, T. (2004). Evaluation of a Virtual Narrator's Expressiveness in Terms of Suspense Signaling. In: *TIDSE 2004*, pp. 89-94.

[Cal01] Callaway, C. & Lester, J. (2001). Evaluating the effects of natural language generation on reader satisfaction. In: *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society*, Edinburgh, UK, pp. 164-169.

[Cav02a] Cavazza, M., Charles, F. & Mead, S. J. (2002). Emergent Situations in Interactive Storytelling. In: *Proceedings of SAC2002*, pp. 1080-1085.

[Cav02b] Cavazza, M., Charles, F. & Mead, S. J. (2002). Interacting with virtual characters in interactive storytelling. In: *Proceedings of the first international joint conference on Autonomous agents and multi agent systems: part 1*, Bologna, Italy, ACM Press, pp. 318 - 325.

[Cha04] Charles, F., Cavazza, M., Mead, S. J., Martin, O., Nandi, A. & Xavier, M. (2004). Compelling Experiences in Mixed Reality Interactive Storytelling. In: *IEEE Multimedia Magazine*, pp. 30-39.

[Dar03] Darcy, S., Dugdale, J., El Jed, M., Pallamin, N. & Pavard, B. (2003). Virtual StoryTelling: A Methodology for Developing Believable Communication Skills in Virtual Actors. In: *ICVS 2003*, Vol. LNCS 2897 (Eds. Balet, O., *et al.*) Springer-Verlag Berlin Heidelberg, pp. 78-81.

[Hay98] Hayes-Roth, B., Stern, A. & Murray, J. H. (1998). Interactive Fiction. *IEEE Intelligent Systems*, Vol.13, Iss.6, pp.12-15.

[Hol00] Holmquist, L. E., Helander, M. & Dixon, S. (2000). Every Object Tells a Story: Physical Interfaces for Digital Storytelling. In: *Proceedings of NordiCHI 2000*.

[Mac00] Machado, I., Prada, R. & Paiva , A. (2000). Bringing drama into a virtual stage. In: *Proceedings of the third international conference on Collaborative virtual environments (CVE 2000)*, San Francisco, California, United States, ACM Press, pp. 111-117.

[Mad03] Madej, K. (2003). Towards Digital Narrative for Children: From education to Entertainment: A Historical Perspective. *ACM Computers in Entertainment*, Vol.1, Iss.1.

[Mag02] Magerko, B. (2002). A Proposal for an Interactive Drama Architecture. In: *Artificial Intelligence and Interactive Entertainment: Papers from the 2002 AAAI Spring Symposium*, Menlo Park, CA.

[Mat99] Mateas, M. (1999). An Oz-Centric Review of Interactive Drama and Believable Agents. In: *AI Today: Recent Trends and Developments*. Lecture Notes in AI 1600, (Eds. Wooldridge, M. and Veloso, M.) Berlin, NY: Springer.

[Mar03] Marichal, X. & Umeda, T. (2003). Real-time segmentation of video objects for mixed-reality interactive applications. In: *VCIP 2003 - SPIE Visual Communication and Image Processing Intl*

Conference, Lugano, Switzerland, Vol. 5150, pp. 41-50.

[Mul03] Mullennix, J. W., Stern, S. E., Wilson, S. J. & Dyson, C.-l. (2003). Social perception of male and female computer synthesized speech. *Computers in Human Behavior*, Vol.19, Iss.4, Elsevier, pp.407-424.

[Mur97] Murray, J.H. (1997) *Hamlet on the Holodeck: The Future of Narrative in Cyberspace*, The Free Press, New York.

[Ong94] Ong, T. & Leggett, J. J. (2004). A Genetic Algorithm Approach To Interactive Narrative Generation. In: *Proceedings of the fifteenth ACM conference on Hypertext & hypermedia*, Santa Cruz, CA, ACM Press, pp. 181-182.

[Pre03] Prendinger, H. & Ishizuka, M. (2003). The Storification of Chances: Corporate Training with Life-like Characters in Virtual Social Environments. In: *Chance Discovery*, Advances in Information Processing Series (Eds. Ohsawa, Y. and McBurney, P.) Springer Verlag Berlin Heidelberg, pp. 189-207.

[Pru01] Pruyn, A. & Wilke, H. (2001) *Sociale psychologie voor managers*, Bohn Stafleu Van Loghum, Houten / Diegem.

[Sil03] Silva, A., Raimundo, G. & Paiva, A. (2003). Tell Me That Bit Again…Bringing Interactivity to a Virtual Storyteller. In: *Proceedings ICVS 2003*, LNCS 2897 (Eds, Balet, O., *et al.*) Springer-Verlag Berlin Heidelberg, pp. 146-154.

[Ree96] Reeves, B. & Nass, C. (1996) *The Media Equation: How people treat computers, television, and new media like real people and places*, CSLI Publications/Cambridge University Press, Stanford, CA.

[Swa01] Swartout, W., Hill, R., Morie, J., Rickel, J., Thiébaux, M., Tuch, L., Whitney, R., Douglas, J., Gratch, J., Johnson, W. L., Kyriakakis, C., LaBore, C., Lindheim, R., Marsella, S., Miraglia, D. & Moore, B. (2001). Toward the holodeck: integrating graphics, sound, character and story. In: *Proceedings of the fifth international conference on Autonomous agents*, Montreal, Quebec, Canada, ACM Press, New York, NY, USA, pp. 409 - 416.

[The02] Theune, M., Faas, S., Nijholt, A., Heylen, D. & ACM, P. (2002). The Virtual Storyteller. *ACM SIGGROUP Bulletin*, Vol.23, Iss.2, ACM Press, pp.20-21.

[The03] Theune, M., Faas, S., Heylen, D. & Nijholt, A. (2003). The virtual storyteller: Story creation by intelligent agents. In: *Proceedings TIDSE 03: Technologies for Interactive Digital Storytelling and Entertainment*, (Eds. Göbel, S., *et al.*) Fraunhofer IRB Verlag, pp. 204-215.

[The04] Theune, M., Rensen, S., Akker, o. d., R., Heylen, D. & Nijholt, A. (2004). Emotional characters for automatic plot creation. In: *Technologies for Interactive Digital Storytelling and Entertainment. Second International Conference, TIDSE 2004*, Darmstadt, Germany, Vol. Lecture Notes in Computer Science 3105 (Eds. Göbel, S., *et al.*) Springer-Verlag, Berlin Heidelberg New York, pp. 95-100.

[Wat98] Waters, K., Regh, J., Louhlin, M., Kang, S. B. & Terzopoulos, D. (1998). Visual Sensing of Humans for Active Public Interfaces. In: *Computer Vision for Human-Machine Interaction*, (Eds. Cipolla, R. and Pentland, A.) Cambridge University Press, pp. 83-96.