Speaking and acting - interacting language and action for an expressive character

Sandy Louchart*
*Centre for Virtual Environments
University of Salford M5 4WT
S.Louchart@salford.ac.uk

Daniela Romano†
†Deparment of Computer Science
University of Sheffield S1 4DP
D.Romano@dcs.shef.ac.uk

Ruth Aylett*
‡Centre for Virtual Environments
University of Salford M5 4WT
R.S.Aylett@salford.ac.uk

Jonathan Pickering*
§Centre for Virtual Environments
University of Salford M5 4WT
J.H.Pickering@salford.ac.uk

Abstract

We discuss in this paper the FearNot! application demonstrator, currently being developed for the EU framework V project VICTEC. It details the language structure, content, interactions management, general architecture and design of the FearNot! Demonstrator, as well as presenting the VICTEC project and its motivations. This paper also focuses on the different sets of Speech Act inspired language action lists developed for the project and discusses their use for an interactive language and action system for the elaboration of expressive characters. The paper also presents early development and implementation work as well as system and speech evaluation planning.

1 Introduction

This paper discusses the language system being developed for the EU framework V project VICTEC1 - Virtual ICT (Information and Communication Technologies) with Empathic Agents. This seeks to use virtual dramas created by interaction between intelligent virtual agents as a means of dealing with education for children aged 8-12 in which attitudes and feelings are as important as knowledge. The project thus focuses on Personal and Social Education, which includes topics such as education against drugs, sex education, social behaviour and citizenship. The topic specifically addressed by VICTEC Victe (2004)is education against bullying. The project expects to contribute to an understanding of the role of empathy in creating social immersion, and to the evaluation of virtual environment ICT for child users. It also expects to contribute to a deeper understanding of empathy and its role in bullying, and to the relationship between Theory of Mind (TOM) Woods et al. (2003) and bullying behavior. The building of empathy between child and character is seen as a way of creating a novel educational experience.

An output of the project is the FearNot! 1 demonstrator, currently under construction. The overall interactional structure of this demonstrator alternates the enaction of virtual drama episodes in which victimisation may occur, and interaction between one of the characters in these dramas and the child user, who is asked to act as their 'invisible friend' and help them to deal with the problems observed in the dramatic episodes. The advice given by the child will modify the emotional state of the character and affect its behaviour in the next episode. The narrative approach undertaken by the VICTEC project is the one of Emergent Narrative Aylett (1999), Louchart and Aylett (2002), Aylett and Louchart (2004). The research on Emergent Narrative aims at finding and elaborating a narrative structure appropriate and suitable for an optimal use of Virtual Environments, combining the entertainment values of both storytelling and virtual experiencing.

The FearNot! Demonstrator represents an intuitive interface between the virtual world and the child user. The characters appearing in the demonstrator have been modelled to be believable rather than realistic, with the use of exaggerated cartoon-like facial expressions. Evaluation to date Woods et al. (2003) has shown that providing the narrative action is seen as believable, lack of naturalism is not perceived as a problem by prospective child users. FearNot! Draws upon feelings of immersion and suspension of disbelief, essential characteristics of Virtual Reality (VR) and Virtual Environments (VE), in order to build empathy between the child and the virtual character as the child explores different coping behaviours in bullying.

2 Integrating language and action

Unlike most dialogue systems or talking heads, VICTEC mixes language interaction with physical actions. Bullying can be categorised as verbal, physical, or relational (manipulating social relationships to victimise), so

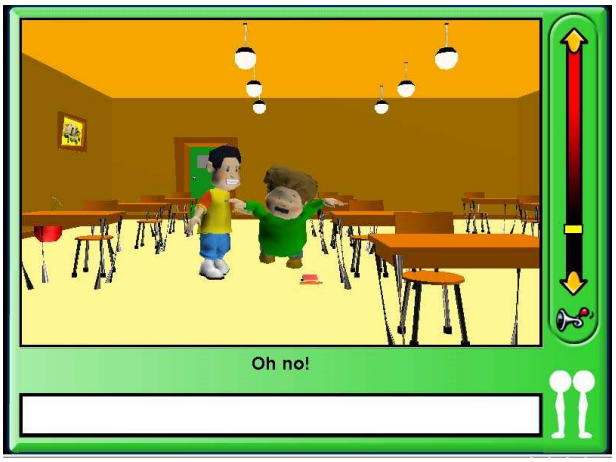

Figure 1: A screen shot of the FearNot! demonstator.

that actions such as pushing, taking possessions and hitting must be modelled. Each character displayed in the FearNot! Demonstrator is provided with its own autonomous action selection mechanism, and the overall architecture is shown in 2. An appraisal of events and the other characters is carried out, using the emotion-modelling system of Ortony, Clore and Collins Ortony et al. (1988) and the resulting emotional state is combined with the character's goals and motivations to select an appropriate action. Thus a common representation for both physical actions and language actions is needed so that both can be equally operated upon by the action-selection mechanism.

This representation is provided by the concept of a speech act Austin (1962), Searle (1969), defined as an action performed by means of language. Here, language is categorised by its illocutionary force, that is, the goal that the speaker is trying to achieve; the same view of action taken by an action-selection mechanism, and highly relevant to bullying scenarios. Speech Acts however work at a very high level of abstraction (e.g. assert, promise, threaten) and only a subset of those generally used are relevant to bullying scenarios. Moreover much of the subsequent work - such as that in Dialogue Acts Bunt (1981) - has taken place in language-only domains and does not address the close relationship between speech and actions required for the VICTEC project. It was therefore decided to define a set of language actions in the spirit of speech acts, using a corpus of bullying scenarios constructed by school children using a story-boarding tool Kar2ouche Education (2004).

Of course a speech act does not uniquely specify the utterance in which it is expressed - its locutionary form. Moreover it was created as an analytic tool, while the language system being created here must function in a generative capacity (see Szilas (2003) for other work with this aim). In addition, language and other actions must form coherent sequences, accepted as such by the child users. The approach must also take account of cross-cultural language practices such as the specific language used in schools in the UK, Portugal and Germany, the countries of the project partners.

Finally, there are two different contexts in which the language system must work. The first is within dramatic episodes in which characters interact with each other. The second is between episodes in which the character must interact with the child user.

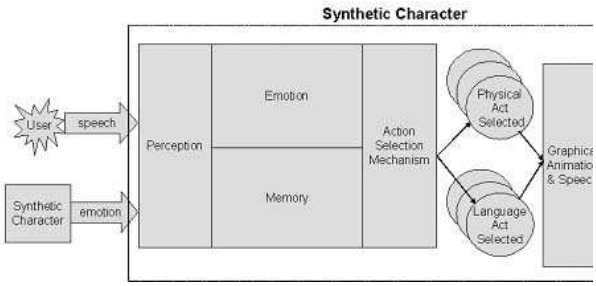

Figure 2: The architecture developed to support synthetic characters.

2.1 From action to utterance

An action can be described as a collection of instances of: an object on which the action can be performed (those being a object of the environment or another characters), the agent performing the action, the action priority used to order and deal conflicting actions, the context in which the action is performed (i.e. location, props, internal goal, history of previous actions, topics), the emotional status of the character at that time, and the utterance (relating to the language action) that should be played, and the animation of the part of the body of the character involved and the gesture. The emotional status of the character will determines whether the action to be performed is implemented via language action, physical activity or both.

Assuming that the next action selected is physical, from a current pose of the character a series of animations are possible, but to reach the current select one it might be necessary to introduce an intermediate pose that links the two (i.e. next action: walk to the door. Current pose: sitting. Intermediate pose necessary: stand up).

We can visualise this as a tree of behaviours where from a current state the next animation is possible only when the correct status of the character is reached and that action can began, requiring the introduction of an intermediate pose. See 3.

In order to generate the utterance for a selected language action, it has been decided to use a shallow-processing approach, as originally used in ELIZA Weizenbaum (1966) and more recently in chat bots Mauldin (1994). The rationale for this approach is that it takes little processing resource compared to a deep approach based on parsing and semantics, thus allowing the graphics engine the resource it needs to run in real-time. In addition, such systems can show surprising resilience

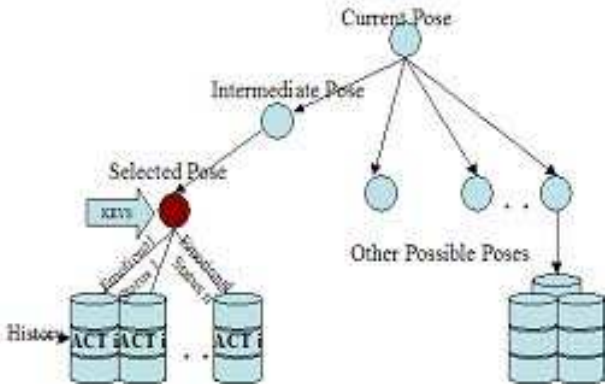

Figure 3: The search tree showing the space of possible behaviours.

in limited domains such as that of FearNot!, in which the language to be used is specific to the bullying scenarios. To prevent problems experienced with such systems in dealing with unexpected inputs, the FearNot! demonstrator will specifically drive the conversation in child-agent mode by using leading questions with a limited range of options for answer. Wizard of Oz studies are in progress to determine in more detail what language coverage will be required.

The FearNot! demonstrator 's natural language system is required to adopt and implement techniques and technologies inspired from research in conversational agents Braun (2002), Braun (2003), Rist et al. (2003), Prendinger and Ishizuka (2001), Prendinger and Ishizuka (2002). Similar approach as already successfully been implemented in Facade Mateas and Stern (2003), where the agent has the possibility of chosing between actions and language when interacting with another agent or a user. However, Facade's low level of abstraction approach would be difficilly manageable for VICTEC and would require more development than actual ressources would allow.

In agent-agent interaction, the language system starts with the language action generated by the action-selection system, which has the advantage of knowing exactly what action (language or otherwise) it is responding to. This indexes a group of utterance templates in which the previous utterance or physical action is used to fill in variable slots with an appropriate choice. For example, assume the utterance from the other agent was" I like flowers", the following group of utterance is selected: I like ... too, why do you like ...?, what do you find in ...?. The first unused utterance here is: "why do you like ...?" the dots are filled with the recognized object of the discourse in the user's input: flowers. The generated character utterance is "why do you like flowers?".

Child and character interaction is different. Here the action is not known, but must be inferred. The incoming text is matched against a set of language templates, and the language and action index is then taken as the starting point for the language action with which the agent must respond as discussed below. Since an objective is to retain control of this dialogue by keeping the conversational initiative with the character, the Finite State Machine structures discussed below can also be used to generate expectations about what language actions the child has produced.

3 The FearNot! Speech Act Knowledge-base

Since the FearNot! demonstrator developed for use in the VICTEC project includes in the same application both agent-to-agent and agent-to-user interactions, it is essential that such this particularity is taken into account when designing and developing the language actions' articulation and content. For this reason, the FearNot! demonstrator must combine the use of a bullying themed speech while operating on these two different and distinct levels.

In order for the FearNot! Demonstrator to successfully meet VICTEC's evaluation objectives, it is crucial that continuity and coherence is maintained during interactions (contextualisation) between agents while insuring that the communication is engaged and led by an agent when agents and users interact together. This not only fundamentally affects the design of the language system, it also requires the design of two distinct sets of actions,independent of each other as just discussed. For instance, in the case of an agent-to-agent communication, the process starts with the selection of a language action and ends with the selection of an utterance. The opposite occurs in the case of agent-to-user communication since the system needs to recognise an utterance via keywords and then select an appropriate language action or action in order to provide an answer to the user.

3.1 Action categorisation

A set of appropriate actions for bullying and victimization interactive scenarios has been identified. Those actions can be triggered and generate agent utterances according to their emotional states. As such a system is dealing with a number of actions and utterances, we have grouped the entire language content within three categories, Help, Confrontation and Socializing 1.

Each category includes a variable number of appropriate language and other actions. For instance, the confrontation category contains a considerably larger number of actions than the help section since there is a very limited number of coping behaviours available in dealing with bullying Woods et al. (2003).

The Help set articulates the actions needed to generate offering-help interactions between agents. It covers the interactions needed for the generation of enquiries from agent-to-agent with respect to emotional states and related goals. In addition, this function also generates advice and

Table 1: Actions catagories and example listings

Catagorie	Listings
HELP	Ask for help / Offer help / Help question / Help advice / Help introduce to friend / Help talk to someone / Help invitation / Offer protection / Non assistance confirmation
CONFRONTATION	Order / Aggressive questioning / Do / Forbid / Defiance / Tease / accusations / Insult / Threat / Aggressive answer / Apology / Abandon action / Action / Hit / Lie / Steal / Obey / Deny / Ask why / Beg / Claim back / Leave / Struggle
SOCIALISING	Greeting start / Topic introduction / Exclusion topic introduction / Information topic / Information exclusion topic / Questions topic 2 / Question topic 3 / Exclusion question 2 / Exclusion question 3 / Exclusion invite / Invitation / Greeting end

offers such as help, protection or assistance. As with the other categories, the Help language and action set category has been designed according to a potential sequential structure. This can be triggered either by an agent asking for the help of another or in response to an aggressive action carried out on a particular agent.

The Confrontation language and action set provides the necessary content for an altercation between two different agents. This category covers most of the physical bullying expressions and involves threats, insults, orders, aggressive behaviour that leads to aggressive actions and violent behaviour. Finally, the Socialising category includes language and actions tha can be used in social discussion by pupils in schools (sports, homeworks, music, video games) and language and actions tha can be used in generating relational bullying. Relational bullying is different from physical bullying, depending on socia exclusion and should therefore be integrated into socialinteraction, as opposed to help or confrontational actions. Although the structure is simple in theory, its implementation requires a large number of utterances and topics.

3.2 Actions Finite State Machine (FSM)

Each action category also possesses its own organisation and consequently requires the design of its own Finite State Machine (FSM). A language action is coherent to both the system and the user if organised into structured speech sequences. While this has to be taken into account it is also essential that the speech system focuses on organising the possible sequences of utterances and ensure the transfer and communication of content without interfering with the agent action selection mechanism. Since, as with all speech system, there are issues of contextualisation, the utterances that constitute the content of the system are formed of templates that can be filled appropriately by the speech system, based on keyword recognition.

Table 2: An example sequence of speech act utterances

Speech act	Utterance
DO	You, [order] now!
If speech act = DENY	You must be joking, [rejection] [insult]
If speechact = OBEY	Ok, but please don't hurt me!

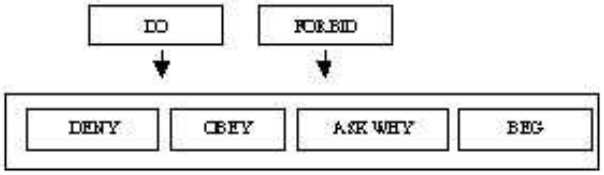

Figure 4: An example of a sequence of speech acts.

Each FSM integrates the language actions relative to the category itself but also potential elements of answers for discussion or interaction. For instance, the actions 'DO' or 'FORBID' in a confrontational situation will be followed by the actions 'DENY', 'OBEY', 'ASK WHY' or 'BEG' Figure 4, to retain conversational coherence.

The VICTEC language actions and utterances have been elaborated according to sequence of actions observed in the scenarios developed by school children mentioned above.

Speech acts are materialised on the FearNot! Demonstrator by utterances. The situation presented in Figure 4 would produce, in case of denial or obedience from the victim the following exchange shown in table 2.

3.3 User-to-agent language action design

Since, the language generated by the user is highly ambiguous and there are no means for the system to understand the meaning of a sentence, the user-to-agent interaction, as we mentioned previously, needs a different approach. As a sentence can only be "understood" by the keywords included in it, it seems sensible to leave the initiative to the agent rather than the user. The fact that the system leads the conversation with the user presents an advantage in terms of believability for the speech system

in the sense that, the system can be expectation driven and can expect a certain type of answer from the user and adjust and compare the answer to a set of pre-defined templates. Although the system could not understand its human interlocutor, it could generate a high level of believability and interact with its user by asking simple and adequate questions.

In order for the agent to keep the upper hand in terms of interaction with the child user, it must be the one asking for advice and the one who generally ask questions to which the child user is expected to answer.

It has been decided, due to the high possibility of misspelling from the children who are going to use the system, that the language system includes a keyword recognition feature that should allow it to recognize the intention of the user and make the association with existing categories of actions.

Although the speech system, in the case of a user-to-agent interaction mainly requires language actions expressed through utterances from the agent rather than the user, it is however important to predict and anticipate answers in regards to the different possibilities that are been offered to the user. Since the VICTEC project is mainly being tested by children aged between 8 and 12 years old, it has been thought that such approach would also have the advantage of helping them in formulating their answers to the agent. The speech system is, in the particular case of the VICTEC project divided into two distinct sections, the agent language actions and utterances and the user's answers.

Table 3: User and Agent Language actions lists

Catagorie	Listings
AGENT	Ask for advice / Ask again /
	Prompt / Cannot understand statement /
	Ask for reason / justification /
	Thank user for advise /
	Confirm advice with user /
	Express reproach to user /
	advice rejection /
	Express disappointment towards user /
	Report result of interaction /
	Beg for help
USER	Give advice /
	Refuse to give advice /
	Ignore the agent / No answer /
	No helpful comments /
	Advice confirmation / Justification

4 Implementing Language Acts

In this section we address the problem of implementing speech acts within the Victec system. The approach will be to take the small set of speech acts defined above in 3 and find a structural similarities in the sentences used to represent them.

4.1 Speech and Dialog Acts

Speech acts are descriptions of utterances in terms of the function they perform Searl (1975). The simplest example is the sentence, "I name this ship Lady Day.", which can be classified as the act of naming.

More recently *dialog acts* have been proposed Bunt (1994), which extend the concept of speech acts to include analysis of the conversational use of an utterance. For example the utterance "I'm sorry, do you mean the first or second letter", would be classified as a repair action, as is clears up failure to communicate properly.

Formally speech act theory may be defined as a branch of pragmatics that classifies utterances by the role or use that they serve in a communication. The role played by an utterance is the function that it provides as distinct from the semantic meaning of the utterance. This is best clarified by an example, consider the two following sentences.

1. Give me the keys.

2. Please may I have the keys.

The meaning of both sentences is the same, the utterer is asking for the keys. However the pragmatics is very different, in the first case the utterer is giving an order while in the second a request is being made. It is important to note that speech acts are not unique classifiers of utterances. It is equally possible for instance, to classify both the previous utterances as communications of a desire. Some examples of speech acts follow.

Question to gain information the questioner needs some information, for example "Where is the milk?".

Exam question the questioner knows the answer but is testing the candidate, for example "Who is the president of Mexico?".

Continuer during a long monologue a speaker will pause, giving chance for the lister to indicate that they are still following the speaker, for example "Yes go on".

Each of the above examples can be satisfied by many utterances and it is impossible to tell from the language act alone how to construct or choose and utterance. The problems are: speech acts alone contain no semantic information; speech acts are not unique; and speech acts cannot in general be mapped to syntax. It is claimed Jurafsky and Martin (2000), that classifying utterances into speech acts is an AI complete problem, meaning that a human being, or a computer system equivalent to a human being, would be required to correctly classify them.

4.2 Microgrammes

Although the general problem of classifying speech acts is currently unformalisable. It is possible to produce automatic classifiers that give partial coverage of common acts. The method for doing this exploits the fact that many speech acts correlate to structural features in a conversation. These structures, introduced by Goodwin (1996) have been called *microgrammes*. They comprise set of features which are classified into three different types.

Words and collocation certain words and particularly combinations of words (collocation) indicate some speech acts. For example the words 'who, when, where', indicate questions.

Prosody the tone of voice used in an utterance may indicate its intended act. In English questions,for example, can be indicated by a rising intonation at the end of a sentence.

Conversational Structure the current context and the immediate predecessor statements may give an indication of the speech act. A simple example of which is that the utterance after a question is probably a reply or a request for clarification.

In the case of a textural system such as Victec prosody will have no part to play. The burden of the work will have to be achieved using pattern matching to identify words and collocation. Although hopefully, some support can be provided through the use of context.

4.3 Language Acts

Because the Victec project is centred on the development of autonomous agents that interact in a virtual environment by the use of actions, it was natural to use speech acts to define the agent's speech system. This would allow the agent to remain in an action reception, action appraisal, action selection loop.

The problem is the lack of semantics and multiple definitions of speech acts. To allow for the first some semantic information has had to be added to the agents actions. We have called the combination of speech act plus semantic information *language acts*.

We intend to solve the second problem of how to identify sentences with speech acts by applying microgrammers to the very small set of sentences that have been classified in the knowledge base.

A microgrammer can be written for each speech act. When parsed in conjunction with the semantic information and contextual knowledge of the source and sender of the speech act the microgramme will generate a sentence.

For example consider the act of greeting a person. The set of possible sentences is very small, consisting of a greeting word, possibly the name of the person being greeted, and possibly a greeting question.

Hello
Hello Sue
Hi
Hi Tom
Hi Jo, how are you?

We can immediately see a general form to these sentences and written down in Backus Naur Form (BNF) it is.

⟨ Greeting word ⟩ ⟨ ToName? ⟩ ⟨ status_question? ⟩
⟨ Greeting word ⟩ = ⟨ Hello ⟩
= ⟨ Hi ⟩

⟨ ToName? ⟩ = ⟨ `receiver` ⟩

⟨ status_question? ⟩ = ⟨ how are you? ⟩
= ⟨ are you all right? ⟩

The term `receiver` is a context variable that is set by the semantic information in the language act.

4.4 Implementing Agent to Agent Language Acts

The database of language acts will be specified in XML, which has the expressive power of a context free grammar and so can express BNF statements. Each language act would be implemented as a template consisting of rules. The rules are implemented separately and may be recursive.

A language act generated by an agent will contain the name of the sending and receiving agents, the type of act and some semantic information, that is act specific. When received the template for the requested act is found, context variables are set and then the rules are repeatedly applied until a response has been formed.

For example the following XML specifies a request for a greeting from agent Tom to agent Sue, with an optional question about Sue's current status.

⟨ Type ⟩ Greeting ⟨ /Type ⟩
⟨ Sender ⟩ Tom ⟨ /Sender ⟩
⟨ Receiver ⟩ Sue ⟨ /Receiver ⟩
⟨ SemanticValue name="true" statusQuestion="random" /⟩

Fist the variables `sender` and `receiver` would be instantiated to the names Tom and Sue, then the template for the language act looked up. Using a greeting language act as specified in 4.3 the template requires a greeting word that can be 'Hello' or 'Hi', as there is no other information a random choice would be made.

Next the because the name attribute of the semantic information was set to `true` the 'ToName' must be added. The rule for this evaluates to the context variable `receiver`, so the value of Sue would be added to the reply.

Finally the status question attribute of the semantic information was set to `random` so the the language system

will chose with equal probability between the adding and not adding a status question. If a question is to be added the rule is evaluated which gives a random choice of two possible questions.

This results in one of the six following possible greetings being generated.

Hi Sue
Hello Sue
Hi Sue how are you?
Hello Sue how are you?
Hi Sue are you all right?
Hello Sue are you all right?

The rules can be recursive, allowing a rule to contain other rules, which will allow the case and gender agreements of German and Portuguese to be applied.

4.5 Implementing User Agent Dialog

In the case of user agent dialog the problem of classifying the user's speech acts and extracting the semantic information for appraisal by an agent must be addressed. As was stated in 4.1 this is in general a very difficult problem.

The problem may be simplified by noting three features that will apply in the case of FearNot.

- The dialog will be very short and focused only on the previous bullying episode.
- The users will be children of age 10 and so will only type simple sentences.
- In order to help the children buttons providing part formed sentences will be provided.

It is hoped that these features will so constrain the input domain as to allow the identification of speech acts using pattern matching to look for words and collocation supported by some conversational structure information.

5 Conclusion

This paper has described the interactional structure and articulation of the language system being developed for the VICTEC project and reported on progress made so far. It also detailed the different language actions and their categorisation in relation to the specific theme of bullying.

The language system and its content have been developed based on actual language currently in use amongst school children, however it requires iterative refinement and testing of both its efficiency and the coherence as well as evaluation of its capacity to suspend or limit the initial disbelief commonly generated by this type of system. A series of Wizard of Oz experiments Maulsby et al. (1993) along with psychological and usability evaluations Woods et al. (2003) are therefore planned. Further evaluation of the whole FearNot! Demonstrator is also planned: for example, a large sample of children (N: 400) will take part in a psychological evaluation at the University of Hertfordshire in June 2004. However, while the agent architecture of the system and systems integration is still under development, language graphical content has already been produced for preliminary evaluation and the VICTEC team is working with the aim to present a first prototype of the system by April 2004.

References

J L Austin. *How to Do Things with Words*. Harvard University Press, 1962.

R Aylett. Narrative in virtual environments - towards emergent narrative, 1999. URL `citeseer.nj.nec.com/mateas98narrative.html`.

R Aylett and S Louchart. Towards a narrative theory of vr. *Virtual Reality Journal, special issue on storytelling*, 7, January 2004.

N Braun. Automated narration - the path to interactive storytelling. In *Workshop on Narrative and Interactive Learning Environments, Edinburgh, Scotland*, 2002.

N Braun. *Storytelling & conversation to improve the fun factor in software applications*, chapter 19. Kluwer Academic, 2003.

H Bunt. Rules for the interpretation, evaluation and generation of dialogue acts. IPO annual progress report 16, Tilburg University, 1981.

Harry C. Bunt. Context and dialogue control. *THINK Quarterly*, 3:19–31, 1994. URL `citeseer.nj.nec.com/bunt94context.html`.

Immersive Education. Kart2ouche. www.kar2ouche.com, 2004.

C Goodwin. Transparent vision. In E Ochs, E Schlegloff, and S Thompson, editors, *Interaction and Grammer*. Cambridge University Press, 1996.

D Jurafsky and J Martin. *Speech and Langugage Processing*. Prentice Hall, 2000.

S Louchart and R Aylett. Narrative theories and emergent interactive narrative. In *Proceedings Narrative and Learning Environments Conference NILE02 Edinburgh, Scotland*, pages 1–8, 2002.

M Mateas and Andrew Stern. Integrating plot, character and natural language processing in the interactive drama faade. In *Proceedings Technologies for Interactive Digital Storytelling and Entertainment (TIDSE) conference, Darmstadt, Germany*, pages 139–152, March 2003.

M Mauldin. Chatterbots, tinymuds, and the turing test: Entering the loebner prize competition. In *Proceedings AAAI 94*, 1994.

D Maulsby, S Greenberg, and R Mander. Prototyping an intelligent agent through wizard of oz. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, pages 277–284. ACM Press, 1993.

A Ortony, G L Clore, and A Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.

H Prendinger and M Ishizuka. Let's talk! socially intelligent agents for language conversation training. *IEEE Transactions on Systems, Man,and Cybernetics - Part A: Systems and Humans*, 31(5):465–471, 2001. (Special Issue on Socially Intelligent Agents - The Human in the Loop, ed. K Dautenhahn).

H Prendinger and M Ishizuka. Evolving social relationships with animate characters. In *Proceedings Symposium of the AISB-02 Convention on Animating Expressive Characters for Social Interactions, London, UK*, pages 73–78, 2002.

T Rist, M Schmitt, C Pelachaud, and M Bilvi. Towards a simulation of conversations with expressive embodied speakers and listeners. In *16th International Conference on Computer Animation and Social Agents*, pages 5–10. IEEE Computer Society, 2003.

J Searl. A taxonomy of illocutionary acts. In P Cole and J Morgan, editors, *Speech Acts: Syntax and Semantics Volume 3*, pages 59–82. Academic Press, New York, 1975.

J Searle. *Speech Acts*. Cambridge University Press, 1969.

N Szilas. Idtension: A narrative engine for interactive drama. In *1st International Conference on Technologies for Interactive Digital Storytelling and Entertainment (TIDSE 2003)*, pages 24–26, 2003.

Victe. Victec home page. http://www.VICTEC.org, 2004.

J Weizenbaum. Eliza - a computer program for the study of natural language communication between man and machine. *Communications of the Association for Computing Machinery*, 9:36–45, 1966.

S Woods, L Hall, D Sobral, K Dautenhahn, and D Wolke. Animated characters in bullying intervention. In *IVA 2003*, number 2972, pages 310–314. Springer-Verlag LNAI, 2003.