Avatar Body Language
Supporting Emotive Communication in Virtual Environments

Jakob Bandelin

Supervisor: Mattias Arvola
Examiner: Arne Jönsson

Cognitive Science Programme
Department of Computer and Information Science
Linköping University

LIU-IDA/KOGVET-A--10/012--SE
Abstract

This thesis tells the story of a design case creating an user interface for a MMORPG where the player are able control the body language of the avatar. By this the game can achieve a gameplay about drama and strong characterization. The thesis addresses considerations on what aspects of body language that can be important for computer games and other virtual environments. It also offers design considerations when designing interfaces for using gestures and other body signals to communicate emotions in virtual environments such as computer games. The main design consideration when creating the interface was to treat the player as an actor and the game world as a stage. The player needs to be in control of combinations of facial expressions, body posture and gestures as well as relative avatar positioning to other characters and objects. The interface was first tested as a paper prototype, re-designed, re-tested and then implemented into a computer prototype.
To complete this project has been a long journey, and I am indebted to many people. I would like offer my regards and blessings to all of those who supported me in any respect during the completion of this project.

Acknowledgments
# Table of Contents

1. **Introduction**
   - Thesis problem 7
   - Method 8
   - Contributions 8
   - Definitions 8

2. **Theory**
   - Face-to-face communication 11
   - Telling stories 17
   - Games 20
   - Computer mediated communication 23

3. **Method**
   - Case studies 27
   - Reflective studies 28
   - Prototyping 29
   - Procedure 30

4. **Design**
   - Design brief 33
   - Analysis of predecessors 34
   - Qualities-in-use 43
   - Avatar body language communication types 44
   - Designing for MMO 45
   - Design Considerations 46
   - Prototype building and testing 55
   - Specification 59
   - In summary 61

5. **Discussion**
   - Result 63
   - Reflections on game design process 66

6. **References**
   - Literature 67
   - Games 71
Table of Figures and Tables

Figure 1. Everquest 35
Figure 2. Neverwinter Nights 37
Figure 3. Anarchy Online 38
Figure 4. Resident Evil 40
Figure 5. Grim Fandango 41
Figure 6. The Sims 42
Figure 7. Prototype 1 57
Figure 8. Prototype 2 58
Table 1. Keys for controlling avatar body movement 59
Table 2. Keys for controlling the avatar body language interface 60
Figure 9. Ouroboros Pre Alpha partial implementation 61
1. Introduction

In a face-to-face situation spoken language is used to communicate thoughts, desires, emotions. However there is also the use of other signs and signals. There are smiles and laughs, and people shake their heads. People communicate when wearing different clothes and by inhibiting different bodies. These non-spoken signals can be called body language and are a major part of human communication.

In computer games the player is often in control of an avatar; a character or player entity. When the avatar is face-to-face with other characters in the virtual game environment the player often has little or no ability to control the avatar body language. Not being able to cry, smile, jump, dance, clap hands, nod, sneer, hug, kiss, slap, sit, point or spit limits how alive and believable the avatar and other characters in the game will seem.

If the avatar was capable of using body language for communicating there could be interesting new game experiences for players to immerse themselves in. The difficulty is how to design the interface for controlling avatar body language. There is also the need to determine what gestures are important for expressing emotions.

Thesis problem

There is little research on how the player can use the body language of the avatar to communicate with other players. The main question of this thesis is:

- What are important design considerations when building the interface for a computer game (especially a MMORPG) with player controlled avatar body language?
Method

The thesis problem was explored while creating the interface for the pre-alpha release of the research computer game Ouroboros. The game is a MMORPG where the gameplay is about strong social interaction and drama.

The development of the interface was treated as a case study. This thesis was in part written at the Zero Game Studio where the MMORPG Ouroboros where under construction. Explorative design investigated the different solutions to the design problem. Prototypes were built and cooperatively evaluated. All test users were from the game’s target group ranging from casual computer game players to hard-core role-playing gamers and computer gamers, ages ranging from 19–34.

An interface design specification was then produced to be implemented in the first hi-fi prototype. In that prototype two players could communicate over the Internet by chatting and each controlling an avatar and its body-language in a shared virtual environment.

Contributions

The thesis addresses considerations on what aspects of body language that can be important for computer games and other virtual environments. It also offers design considerations when building interfaces for controlling avatar gestures and other body signals to communicate emotions in virtual environments such as computer games.

Definitions

The term computer games is used in a wide sense including all computer mediated games independent from platform. Console games, video games, arcade games and games played using handheld devices (like mobile phones, Nintendo Game Boy, etc) are all considered computer games.

References to computer games are given in the form Game name (Publisher, Year) e.g. Animal Crossing (Nintendo, 2001).

Console games are games played using a computer gaming devices linked to a TV-set, e.g. Nintendo’s Gamecube and Wii, Sony’s Playstation and Microsoft’s Xbox.
In a **RPG**, Role Playing Game, participants assume the role of a fictional character and decide that characters actions. Early RPGs use pen and paper to record characters skills and inventory. Success of actions are often determined by dice.

**MMORPG**, Massive Multi-user Online Role Playing Game, is a game where hundreds or even thousands of players play at the same time in a shared, often persistent, online world. Some of the most influential MMORPGs are EverQuest, Asheron's Call, Anarchy Online, Neverwinter Nights and World of Warcraft. A game with less players on a shared online world is not considered Massive.

**LARP**, Live Action Role Playing, is a form of role playing game where the participants physically act out their characters' actions. The participants pursue goals within a fictional setting represented by the real world, while interacting with each other in character. The outcome of actions may be mediated by game rules, or determined by consensus among players.

An **Avatar** is a graphical representation of the player character controlled by the player.

The **Player** is the active user of a computer game.

A **NPC**, Non Player Character, is a character in a computer game not controlled directly by the player.
2. Theory

This chapter sets the theoretical framework for the thesis by introducing how face-to-face communication works and how identity and emotions are communicated in everyday encounters, movies, literature, storytelling and games. The chapter also presents theories of play and games.

Face-to-face communication

How face-to-face communication works especially how emotions and identity are communicated have design implications for an interface for virtual avatar face-to-face communication. Also of interest when working with games and drama is to understand how actor and audience can separate fiction from reality.

According to Clark (1996) the spoken face-to-face setting is the basic setting for language use because it is the universal setting in all human societies. All other forms of communication are deviations from this setting. The face-to-face setting have ten features (Clark 1996; Clark & Brennan 1991):

1. **Copresence.** The participants share the same physical environment.
2. **Visibility.** The participants can see each other.
3. **Audibility.** The participants can hear each other.
4. **Instantaneity.** The participants perceive each other’s actions at no perceptible delay.
5. **Evanescence.** The medium is evanescent—it fades quickly.
6. **Recordlessness.** The participants’ actions leave no record or artifact.
7. **Simultaneity.** The participants can produce and receive at once and simultaneously.
8. **Extemporaneity.** The participants formulate and execute their actions extemporaneously, in real time.
9. **Self-determination.** The participants determine themselves what actions to take when.

10. **Self-expression.** The participants take action as themselves.

All non-basic settings are deviations of the features of basic conversation (Clark, 1996). Communication in an on-line virtual environment share some of the features of the basic setting but not all. Regardless of setting communication requires some sorts of signs or signals transmitted between entities.

**Signs**

In both spoken language and body language signs are used for communication, i.e. signaling. When signaling signs are used deliberately and often consciously to communicate something. It is not considered signaling when something is communicated involuntarily or unconsciously. A sign can be classified as either an icon, an index or a symbol. An icon is something that resembles the object perceptually. An index has a spatial, or casual, connection between the object and the sign. A symbol is a conventional sign like a word, road sign, book or gesture. Many signs are combinations of icons, indices and symbols (Clark, 1996).

**Icons**

Icons can further classified as either demonstrations or illustrators. Demonstrations are selective depictions (Clark and Gerrig, 1990; in Clark, 1996). The point of demonstrating a thing is to enable the addressee to experience selective parts of what it would be like to perceive the thing directly. People are very creative using their bodies and faces as instruments to depict things. Many demonstrations are auditory, tactual and visual.

> "People can selectively depict all manner of speech characteristics — speed, gender, age, dialect, accent, drunkenness, lisping, anger, surprise, fear, stupidity, hesitancy, power." — (p.175, Clark, 1996)

Many gestures depict what is being talked about. These gestures are called iconic gestures (Clark, 1996) or illustrators (Ekman & Freissen, 1969). Although mostly used by the speaker the addressees may use iconic gestures, for example smile or grimace, in response to speakers. The iconic gestures are tightly connected with speech and anticipate the words they go with. The function of the iconic gesture is facilitative and communicative.

**Indices**

Another type of signal is indices, which are used to indicate an object, event, or state. A successful index must contain these parts (Clark, 1996):
1. **Attention.** The index is the participants joint focus.

2. **Location.** The index locates the object in space and in time.

3. **Physical connection.** The index locates by means of a physical connection with the object.

4. **Description.** The object is specified under a particular description.

5. **Computability.** The speaker presupposes that the addressees can work out parts 1 through 4 based on their current common ground.

Body parts are often used for indication for example by touching, pointing or gazing. Artificial instruments are also used; things like doorbells and starting pistols are extensions of the speaker’s body and voice (Clark, 1996).

**SYMBOLS**

The major parts of most languages are made up of symbols. Words and sentences are symbols. These sentences can be broken down into phrases, words and morphemes and be arranged in a tidy syntactic structure. However, for this thesis we are more interested in symbols as body language and therefore we leave out the linguistic world of syntactic analysis.

Gestures that are symbolic have been called emblems (Ekman & Freissen, 1969). Emblems can often be used instead of or replace words, for example a head nod for ‘yes’ or a greeting wave for ‘hello.’ A cultural common ground is usually required when using emblems since the emblems are based on conventional rules. The same body movement or gesture can have very different meaning in different cultures (Morris, 1977). There are also auditory emblems like handclaps, hissing, and whistling.

Another form of emblem is what Clark (1996) calls junctions. These are generally symmetrical and require behavior from two participants at once. Examples of junctions are hugs, kisses and holding hands. How junctions are performed and what the performers mean by performing them vary from culture to culture (Morris, 1977).

**Communicating identity and emotion**

What emotions to communicate, how and why depends on the identity and personality of the person—this is also true for fictional characters e.g. in a game. The signs and signals also are interpreted by observers.

There are many attempts to classify personality. One of these is the five-factor model. According to this model each personality has five dimensions, extraversion, agreeableness, conscientiousness, emotionality and intelligence (Digman,
Other approaches of the personality are the psychodynamic (Freud 1900, 1901, 1915), the humanistic (Rogers 1947, Maslow 1935), the behaviorist (Skinner, 1935) and the cognitive approach (Kelly, 1955). While being very different all agree, more or less, that a person’s behavior is consistent, that each person has a unique pattern of behavior, thoughts, and that feelings and behavior arises from within. The personality theories agree in that emotions are an important part of human cognition. Attempts have been made to classify the emotional states in humans from observing behavior (Darwin 1872; Ekman & Friesen, 1971; Izard, 1992). Paul Ekman found six basic emotions in cross-cultural studies of facial expressions. Ekmans six basic emotions are anger, sadness, surprise, disgust, enjoyment, and happiness. Izards nine basic emotions are anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, surprise. Other researchers have found other numbers of basic emotions and used different basis of inclusion but it is enough for this thesis to conclude that there are emotions in humans and that these are apparent in the behavior and body language of a person. For instance has it been shown that people can judge emotions by only seeing a persons body movements (Walk & Homan, 1984).

People make judgments about others from appearance and manner to such extent that there is no such thing as not communicating in a face-to-face situation. The choice to try to not communicate is a communicative act in itself. If there are inconsistencies between appearance and manner it may confuse or upset the observer (Goffman, 1967). People use cues that are immediately apparent like talkativeness, grooming, smiling, gender and general appearance when determining the personality of a person they encounter (Pauonen, 1989, 1991). Much of these cues are gathered from kinematics information and people are remarkably accurate in determining things like intention, gender, identity and even deception from kinematics (Runeson & Frykholm, 1983). The body is a powerful communicative tool (Morris, 1977).

However it is important to remember that many signals are ambiguous. It is not always certain why a signal is being displayed. For example:

“[…] when we see a person with a disgust expression, we know that the person is responding to something offensive to taste or to smell, literally or metaphorically, […] and that the person is more likely to run away from the stimulation” (Ekman, 1999).

It is also important to remember that signals can be consciously produced or repressed. This can be done to display an emotion not currently being experienced.
“Not only can there be emotion without expression there can be expression without emotion.” (Ekman, 1999)

Some body parts and signals are easier to control than others. Blushing, crying and pupil dilation are very hard if not impossible for most people to control while hand gestures and poise can be controlled more easily (Morris, 1977). However there are subtle differences between involuntary smiles, social smiles and deliberately made false smiles (Ekman & Davidson, 1990).

**Separating fiction from reality using layering**

In a fictional setting an actor is only pretending to be expressing intentions. There is meta-communication that enables the participants to recognize that “this is play” or “this is fantasy”. This meta-communication is phrased by Gregory Bateson (1955) as:

> “These actions in which we now engage do not denote what those actions for which they stand would denote.”

This paradox-like aspect of language enables a wide range of expression for the speaker as well as addressees. This meta-communications is needed to appreciate novels, plays, jokes, irony, sarcasm, rhetorical questions, teasing, hyperbole and meiosis.

Building on what Bateson called frames Clark (1996) uses the term layering and has presented principles for how layering works:

- **Principle of layered meaning.** The speaker who means what is expressed in an utterance, and the addressee for whom it is meant, belong to the highest current layer of action.

- **Principle of imagination.** In layered actions, the primary participants are intended to imagine what is happening in the highest current layer of action.

- **Principle of appreciation.** In layered actions, the primary participants are intended to appreciate the instigator’s purposes and techniques in creating the highest current layer of action.

Clark (1996) has also gathered the properties of layering:

- **Relation.** Layering is a symmetric relation between joint actions in two domains.

- **Domains.** Each domain is specified, principally by its participants, roles, time, place, surroundings, and possible events.

- **Deixis.** The joint actions in the two layers have distinct deictic frames.
• **Simultaneity.** The two domains are present at the same time.

• **Recursion.** Layering is recursive.

The participants jointly map entities from one domain to another. For this mapping there is some properties of asymmetry (Clark, 1996):

• **Perspective.** The primary participants may construe any entity (an object, state, or event) one-way in domain 1 and, simultaneously, another way in domain 2.

• **Causality.** Many entities in domain 2 are caused by the occurrence of the corresponding entities in domain 1, but not vice versa.

• **Access.** The participants in layer 1 have informational access to entities in domain 2, but not vice versa.

• **Speaker’s meaning.** When there are two layers, the speaker who means what is expressed by a signal, and the addressee for whom it is meant belong to layer 2.

• **Imagination.** When there are two layers, the primary participants are to imagine the actions in layer 2, and appreciate the actions in layer 1.

Layering explain how e.g. an actress can act the role of a character in a theatre and how the audience appreciates that the actress is not always acting as herself and that she is deliberate in displaying signs trying to portray the character and at the same time empathize with the fictional character. The actress and the audience also understands that the scene and props are not real. The character does not know this and it would probably be considered a bad performance by the actress if the fictional character behaved otherwise e.g. bumped into backdrop on stage, she forgot the lines of the play or if when shooting a movie looked directly into the camera. However breaking the fourth wall can sometimes be used as a dramatic or comedic effect.

The use of layers is widespread in communication and gives it a spirit, edge and sense of imagination.
“It is not merely bad natural history that people might or should obey the Theory of Logical Types in their communications; their failure to do this is not due to mere carelessness or ignorance. Rather, we believe that paradoxes of abstraction must make their appearance in all communication more complex than that of mood-signals, and that without these paradoxes the evolution of communication would be at an end. Life would then be an endless interchange of stylized messages, a game with rigid rules, unrelieved by change or humor.” (Bateson, 1955)

Summary
The basic setting for communication is face-to-face. All other means of communication are deviations of this normal setting for everyday conversations present in all human cultures. When communicating signs in the form of icon, indices and symbols are used. Some of the signs communicate identity and emotion of the sender. Signs can be used deliberately to send messages that does not denote what the signs stands for. Emotions that are not felt by the sender can be displayed. The sender can for example be acting as someone else or exaggerate in order to upset or entertain. This has to be considered when designing a game that includes emotive communication. An on-line RPG is in itself also a form of layered communication where players can act not as themselves but as the characters of the game and the other players of the game are able to recognize this.

Telling stories
How stories in both literature, plays and cinema are told and how characters are presented may provide insights to factors that can be useful when designing a game with about drama with strong characterization.

Storytelling and literature
Certain plots and types of characters are more meaningful and entertaining than others. There are dramatic structures that are found more interesting than others. The ancient greek philosopher Aristotle state that artists imitate life, mimesis, when they produce literature. A character should have the following properties: goodness, appropriateness to social mores, truth to life (probability in small details), and consistency (not disturbingly divided in nature). Such a character will be most interesting for the audience. Further Aristotle argues that a good tragedy allows the audience to experience pity and fear. He states the story should be easily embraced by memory but long enough to move a character from calamity to good fortune or from good fortune to calamity. Effective struc-
tures in a plot are *peripeteia*, reversal of expectations, and *anagnorisis*, recognition of someone’s true identity; often one’s own.

Another later approach to storytelling is that of the Russian formalist Vladimir Propp. He found after a study of one hundred arbitrary chosen fairy tales 31 successive functions. The functions of the characters are based upon what effect they have for the story. These functions are *absence, interdiction, transgression*, etc. Such functions are basic units of action. Propp found that all the folk tales had the same structure. A tale always began with either an injury to a victim or the lack of some object. The end is therefore given; it will either include the retribution of the injury or the acquisition of the object. In the beginning of the story the hero is tested by a donor who after a test will supply the hero with some agent that will allow him to surpass the ordeals of the story. The stories always contained a number of the functions and they always appeared in a certain order (Propp, 1928, 1968).

**Movies**

The color design and range of tonalities are considered an important tool to evoke emotions. How master director Ingmar Bergman has described light can serve as an illustration for this:

“The gentle, dangerous, dreamlike, living, dead, clear, misty, hot, violet, bare, sudden, dark, spring-like, falling, straight, slanting, sensual, subdued, limited, poisonous, calming, pale light” (p. 38 in Ettedgui, 1998)

The composition of the frame, using space and time, can be used to direct the attention of the audience.

“The audience is only going to look at the most overriding thing in the frame. You must take charge of and direct their attention. It’s also the principle of magic: What is the single important thing? Make it easy for them to see it, and you’re doing your job” David Mamet, director (p. 176 in Bordwell & Thompson, 2001)

The speed of motion, perspective relations (focal length of lenses, depth of field and focus), aspect ratios, angle, level height and distance of frame all have implications for the interpretation of the shot. A moving camera can used for mobile framing where the camera can pan, tilt or track. Switching between shots can tell about rhythmic relations, spatial relations and temporal relations between the shots. The use of the techniques above can result in a certain film style (Bordwell & Thompson, 2001).
ANIMATION
A believable character supports suspension of disbelief. Animation artists are skilled in how to make images feel alive and have developed several techniques for animating a character.

“Believability. That is what we are striving for … belief in the life of the characters. That is after all the dictionary definition and meaning of the word ‘animation’: to invoke life” (Jones, 1989)

Many of these techniques are not based on producing realistic movements; instead animators focus on what is perceived as real or believable and accentuate certain elements and tone down others.

“Disney animation makes audiences really believe in … characters, whose adventures and misfortunes makes people laugh – and even cry. There is a special ingredient in our type of animation that produces drawings that appear to think and make decisions and act of their own volition; it is what creates the illusion of life.” (Tomas & Johnston, 1981)

One technique that can be used for realistic animation is motion capture. With this technique reflective markers is placed on an actor and by capturing these markers from many angels using an array of cameras the exact position of each marker can be calculated. From the marker data a virtual body can perform the same movements as the actor. For example was motion capture technology used when animating the Gollum character for the motion picture Lord of the Rings (New Line Cinema, 2001) and for the animation of the characters in the game Soul Calibur II (Namco, 2002).

As previously shown people are very skilled in determining identity and gender as well as emotional states from kinematics (Ibid). This is also what skilled professional motion capture artist experience when working with motion capturing for film and computer games. Motion capture technician Richard Widgery\(^1\) emphasize the importance of using the right actor when motion capturing. If one captures data from a female it is very hard to apply that data to a male character without this looking weird. A persons anatomy and personality influence the captured data so much that for capturing a small character one have to use a small actor for best result. It is also his experience that it is not a good idea to use the same actor for capturing data for different characters in a project as people will easy pick up when two characters are moving in the same manner.

\(^1\) Sharing his knowledge in discussions during a training course in Motion Capture.
However there are ways to synthesize and distort motions to give them specific biological and psychological attributes (Troje, 2002). A walk-cycle animation can for example be tweaked into looking more or less happy.

**Summary**
Certain plots makes stories more interesting. Depending on how a character is portrayed it can engage the audience, viewer or reader. When working with animation there are also techniques that can be used to make the characters feel more alive. Motion capturing can be used to record movement of actors and apply them to characters in a computer game.

**Games**
This sub-chapter examines play and games and how activities can be designed to best support playing.

**Play and games**
Playing is acting freely. Play is the opposite of seriousness, however play can be serious. Games and playing are performed mainly for the play itself. There can be secondary interests but the game is a means and an end in itself.

“Play is older than culture, for culture, however inadequately defined, always presupposed human society, and animals have not waited for man to teach them their playing. We can safely assert, even, that human civilization has added no essential feature to the general idea of play. Animals play just like men. We have only to watch young dogs to see that all the essentials of human play are present in their merry gambols. They invite one another to play by certain ceremoniousness of attitude and gesture. They keep the rule that you shall not bite, or not bite hard, your brother’s ear. They pretend to terribly angry. And—what is most important—in all these doings they plainly experience tremendous fun and enjoyment. [...] There are other much more highly developed forms: regular contests and beautiful performances before an admiring public.” (p. 1, Huizinga, 1955)

A game has a playground or a special setting apart from the mundane reality and creates a temporary world within the real world. During the game there is a time-limited perfection—games has a tendency of beauty.
“The play mood is labile in its very nature. At any moment ‘ordinary life’ may reassert its rights either by an impact from without which interrupts the game, or by an offense against the rules, or else from within, by a collapse of the play spirit, a sobering, a disenchantment.” (p.21, Huizinga, 1955)

An essential part of playing is antithetic. There is usually a division between two parties. The tension enables the participants to test their skills and abilities. The winner of the game does not only win the game but also honor and creditability. The competitive instinct has nothing to do with lust for power; the primary is to outdo the others, to be the best and to be regarded a such.

“From the life of childhood right up to the highest achievements of civilization one of the strongest incentives to perfection, both individual and social, is the desire to be praised and honoured for one’s excellence. In praising another each praises himself. We want to be honoured for our virtues. We want the satisfaction of having done something well. Doing something well means doing it better than others. In order to excel one must prove one’s excellence; in order to merit recognition, merit must be made manifest. Competition serves to give proof of superiority. This is particular true of archaic society.” (p. 66, Huizinga, 1955)

Roger Callios (1961) has classified human games and play into these categories:

- **Agon** – competition like sports and athletics
- **Alea** – games of chance
- **Ilinx**, or vertigo – alters consciences or scrambles perception
- **Mimicry** – creates alternative realities, dance, theatre, and arts in general

Callios also uses the terms paidia and ludus. Paidia presents itself in spontaneous play and the play instinct. Paidia is “letting oneself go” and can be found in activities of tumult, agitation, improvisation and joy. On the other side of the scale is ludus which is present in games that involve skill and is the “taste for the gratuitous difficult”. In ludus type of games the participants are conscious of the performance. All agon, alea, ilinx and mimicry activities can all be placed somewhere on the paidia – ludus scale. For agon would the paidia activities of racing and wrestling are contrasted with football and chess on the ludus end. For alea, counting out rhymes are paidia and betting and roulette are ludus. For ilinx paidia can be horseback riding or waltzing and ludus can be skiing or tightrope walking. In mimicry, children’s tag and the use of masks represent the paidia end of the scale while theatre and spectacles are typical ludus activities.
Flow

The term flow is defined by Mihaly Csikszentmihaly (1990) as “joy, creativity and the process of total involvement in life”. The experience of flow is not something that happens by chance but rather something that is made happen. This optimal experience happens by having “realistic goals, and when skills match the opportunities for action. The pursuit of a goal brings order in awareness because the person must concentrate attention on the task at hand and momentarily forget everything else.” By achieving or experiencing flow a person becomes an “increasingly extraordinary individual” (Csikszentmihaly, 1990).

Csikszentmihaly makes a distinction between pleasure and enjoyment. Pleasure is the feeling one achieves when biological and social goals are met. Pleasure is achieved from food when we are hungry, and passively absorbing information and dulling a mind with drugs and alcohol is pleasant. Pleasure does not add complexity to the person it just maintains order in consciousness. Enjoyment on the other hand is something that adds or increases complexity of the person. Activities can be designed to conduct flow and thus being enjoyable optimal experiences. In order to be enjoyable an activity should have:

• Rules that require the learning of skills
• Clear goals
• Provide feedback
• Make control possible

Computer games

The principles of Aristotle’s Poetica can be used to make the stories interesting in computer games. Even so in one aspect computer games are different from dramas, books and movies. In computer games the user becomes the agent and this changes the nature of the experience. Not being limited to being a passive reader or viewer but an active participant or player opens up new possibilities for dramatic experiences (Laurel, 1991).

Summary

There are different types of games and play activities. One fundamental aspect of play and games is the strive for perfection and to out do the other participants; e.g in acting in character better than the others.

Depending on how an activity is structured it can be more or less interesting and rewarding to be engaged in. If an activity have rules that require the learning of skills, provide clear goals and feedback, and make control possible it will be more enjoyable to be engaged in.
The principles of literature and storytelling can be applied to computer games which enables new kinds of experiences.

**Computer mediated communication**

This subchapters investigates how communication is affected when computers are used and how computers can be interpret and display emotive communications.

**Chat and virtual worlds**

In chat systems there is a higher cost of producing statements, which can make the user, produce less words than in a spoken conversation. The cost of producing messages and in chat can also make the conversation seem less polite (Brennan & Oharei, 1999).

However users of chat systems are very innovative in coming up with new forms of communication. The use of emotes are well spread. In Multi User Dungeons and in chat rooms users, or players, provide cues about personality with their avatars appearance and try to compensate for the lack of nonverbal behavior and visual social cues with typed information about actions, feelings and appearance. Computer mediated communication is highly social and participants are often engaged in highly emotional behavior; playing, flaming, making friends and falling in love (Chenault, 1998, Danet, 1991).

One of the attractive factors of chats and virtual worlds seems to be the anonymity and freedom from the physical world.

>“The absence of non-verbal and other social or material cues to identity frees participants to be other than ‘themselves’ or more of themselves than they normally express.” (Danet, 1997).

**Natural Language Processing**

Can computers be programmed to understand and respond to the everyday language of humans?

The biggest problem in natural language processing and for agents that try to communicate like humans is that most utterances are ambiguous. There is lexical ambiguity, where one word has more than one meaning, syntactic ambiguity, which can cause semantic ambiguity, referential ambiguity, caused by anaphoric expressions, pragmatic ambiguity and local ambiguity. Another difficulty when
processing natural language is that it is often vague. Problems also arise when dealing with metonymy and metaphors (Russell & Norvig, 1995).

There are some different approaches on how to extract affective information from text. The most accessible and economic approach is to link certain keywords with emotions. The keyword spotting approach weakness is detection when negation is involved and relies on the presence of the keywords. A more sophisticated approach is lexical affinity, which assigns probability of affect content to words. While often out-performing keyword spotting lexical affinity still performs poor on negations and alternative word meanings. Another approach is statistical language processing which gives a huge corpus to a machine-learning algorithm of affectively annotated texts. While performing fairly well on page and paragraph level the statistical method is often semantically weak if trying to affectively value sentences. Another approach is “hand-crafted models of affect based on psychological theories about human, needs, goals, and desires”. Yet another approach is the an affect sensing engine which can classify written text into Ekman’s six emotion categories in a more robust way than previous approaches (Liu et al., 2003).

**Believable agents**

How is it to interact with an agent, e.g. an avatar controlled by a computer? It has been found that avatars having gestures and facial expressions produce a significantly higher level of co-presence when compared to static avatars. Emotion in agents makes them and the experience of interacting more interesting (Bates, 1994).

Having a behavior-based autonomous cinematography system, a camera agent, for presenting the agents or actors can make interacting in virtual worlds more interesting. Knowledge from film and television can be used; an actor is only as good as the cinematographer. The cinematography system needs to work closely with the interface so that the user is connected to the system (Tomlinson et al., 2000).

Emotion is one of the primary means to achieve believability. To achieve this there sometimes is a need to “fore-shadow, exaggerate and tone down simultaneous action to get the main point across.” Full out realism is not useful for creating a subjective feeling of realism. The arts in general and animation in particular are skilled in making characters believable (Bates, 1994).

Agents can also interactively generate stories, e.g. there are systems that can generate short fictional texts from a line provided by a user (Liu & Singh, 2002).
Input devices
A keyboard is good because it has many keys and it is very easy to type text messages fast – up to 150 word per minute. Many keyboards have functions keys; F1–F12. Function keys positions are sometimes matched on screen to support mapping (Shneiderman, 1998).

The use of a pointer saves the user from learning commands and keeps their attention on the screen. The most common pointing device is the mouse pointer (Shneiderman, 1998).

Another form of input device are joysticks are good for precise analog control (Shneiderman, 1998). The contemporary joypad often features two analog joysticks and an array of buttons.

There are also other forms of input devices with accelerometers that sense the positioning and movement of the device; e.g. the Apple iPhone and Nintendo’s Wii. There are also solutions that use video or audio input; e.g. eyetracking or speech recognition (Shneiderman, 1998).

Signs can be drawn on a trackpad or other touch sensitive surfaces or devices and be translated into text (Kristensen, 2004).

Switching input devices or hand positions on the keyboard during a task is disruptive. For tasks that mix typing and pointing solutions is generally performed faster by using only the keyboard than by keyboard and mouse (Shneiderman, 1998).

Summary
In a virtual environment there are deviations from the basic face-to-face setting of language use. Users of virtual environments and chat rooms often make up ways to compensate for these limitation and can engage in highly social and emotional interactions.

When interacting with computer controlled characters, also called AI agents, the interaction are found to be more interesting if the characters are believable. However it is very difficult getting a computer to understand and respond to natural language.

An interface for a game should be designed to limit or completely avoid switching of input devices.
3. Method

This chapter introduces the methodological framework of the thesis and the methods used during the design process.

Case studies

To be a relevant case study the design problem has to be explored and defined in relation to all contextual factors. A case study reports about a specific case and makes it possible for others to transfer insights and conclusions from one specific case to other specific cases. The question for the case study researcher is: “What can be learned from this specific case?” In a qualitative case study the researcher is spending much time on the site and is personally involved in what is going on.

The researcher interprets the theories and applies them on the specific case. From analysis and reflections on how the theories work with the specific case the researcher can report on how the ideas of the theories were valid and how they could be refined. A case study can be seen as small steps of abstraction towards grand generalizations.

Each case is unique and exists only in a unique context. As everything about the case cannot be told or reported it is up to the researcher to decide what and how the case is told. The case study is subjectively told by the researcher with ideas connected and embedded with illustrations and contexts. The reader of the case study will have to add, subtract and reshape ideas to determine how the insights in one specific case is relevant for another specific case (Stake, 1994).
Reflective studies
In a design situation there is much uncertainty and complexity. The means and ends are unknown. In such a situation the designer has to find the consequences and implications of their design. The designer should reflect-in-action while exploring the design space. Using a conversation composed of “what if?”-questions the designer can find a frame to work with. Using a repertoire, composed of previously encountered design problems, read about cases and stylistic preferences the designer will use reflection as a means of seeing this as that. By reflecting-in-action the designer see similarities with this design and other. (Schön, 1983)

Explorative design
The solution to a design problem has to be searched for and to find a good solution it is desirable to explore many solutions to the problem. Often the objectives of the design process are unstable and tentative and problem boundary is undefined. In order to come up with a solution the designer has to understand the design problem. The design process can be divided into three different stages, divergence, transformation and convergence (Jones, 1992). In the first stage of the design process, divergence, the designer aims to:

“destructure, or to destroy, the original brief while identifying these features of the design solution that will permit a valuable and feasible degree of change. To search divergently is also to provide, as cheaply and quickly as possible, sufficient new experience to counteract any false assumptions that the design team members, and the sponsors held at the start”. (p. 66, Jones, 1992).

Next is transformation from which:

“the general character, or pattern, of what is being designed, a pattern that is perceived as appropriate but cannot be proved right.” (p. 66, Jones, 1992).

In the transformation stage the boundaries of the problem that earlier where undefined are being fixed and divergent work has enabled understanding of the design brief and the design problem. It is important the have the freedom of changing sub-goals to find a solution without compromises. There will probably be many different transformations that can lead to a good result but these should not be mixed. Two good solutions to one problem does not combined necessarily form a great solution.

The last stage is convergence which is to:
“reduce a range of options to a single chosen design as quickly and cheaply as can be managed and without the need for unforeseen retreats” (p. 69, Jones, 1992).

Design space analysis
Design space analysis (DSA) is a way of arguing for a move from vision to design specification (MacLean, 1989). It represent choices and why they where made in the design process and on what basis decisions are made. DSA is a tool for switching views when solving the design problem. DSA helps the designer working divergently and to challenge so called pet solutions.

Qualities-in-use
Social, practical, aesthetickal considerations and material limitations as well as ethical considerations for the design can be formulated as qualities-in-use. Qualities-in-use has also been called abstract usability objectives (Howard, 2002). The qualities-in-use tend to be short sentences or words that can be used to guide the everyday decisions in the design process. These word can be seen as the keywords of the design and is well suited to communicate a shared view on how the design should behave among different people involved in the design process (Arvola, 2003).

Prototyping
Prototypes can represent many different states of evolving design and help the designer to explore options. The terms lo-fi (low fidelity) and hi-fi (high fidelity) are sometimes used for prototypes to indicate how close to the final system the prototype is. Often, but not always, a lo-fi prototype is a paper prototype of the system and the hi-fi is run on a computer. Effective prototypes produce answers to their designers’ most important questions in the least amount of time. A prototype is any representation of a design idea. There are four kinds: role, look and feel, implementation, and integration prototypes. Role-prototypes investigate what an artifact could do for a user. Look & feel prototypes investigate the options for the experience of the artifact. Implementation prototypes are built to answer technical questions. Integration prototypes are a combination of the three types mentioned above (Houde & Hill, 1997).

Provatotyping
A prototype can be used to provoke the current activity and create new activities based on current ones. This method, called provotyping, is intended to be used between the activities of initial analysis of the current and the design of the new.
Cooperative evaluation
A useful way of gathering information about a prototype is cooperative evaluation. Cooperative evaluation aims to identify problems with the prototype and how these can be solved. The method involves users by encouraging them to explain their behavior when they interact with the prototype. The users can be a good source of constructive criticism, which helps designers overcome possible tendencies to design for themselves. Doing a cooperative evaluation with five users is enough. The users should be given specific tasks and the observations made should be recorded in some way. At the end of the observation two categories should be examined: unexpected behavior and user comments (Monk et al., 1993).

Procedure
The case in this thesis is from building an interface for a pre-alpha release of a research project MMORPG called Ouroboros. Much of the time was spent at the Zero Game Studio working as an interaction designer. The specifications of Ouroboros was not firmly set and there were no specifications on how the interface for the game nor the gameplay. There was a general idea that gestures should be used to enhance characterization in the game.

The first step in the design process was to better understand the design problem. This was done by exploring and reflecting on the design brief and how similar problems had been solved earlier. From the exploration qualities-in-use was found. As the design problem started to become clearer a possible solution was sought for.

During the convergence phase two paper prototype was constructed and cooperatively evaluated. The paper prototypes were constructed to find out if players understood the interface and to see if players found it meaningful and fun to be in control of the avatar body language. As such the prototype can be viewed as a provotype as it inspired the players to solve communication problems in a new way and got them to think and respond to how it could be improved.

The final product was a specification for a hi-fi prototype that also served as a provotype. It would further test how players will respond to being in control of
Method

the avatar’s body language and how this will change their action and game experiences. However this was not within the scope for this thesis.
4. Design

This chapter tells the story of the design process for this case.

The process was not as linear as it appears in written form. Many of the activities were more often than not in progress in parallel; e.g. the analysis of predecesors was not a task that was completed and left but during the whole design phase it revisited and expanded upon.

Design brief

The design problem was to create an interface for the pre-alpha version of the game Ouroboros. The game was intended to be a MMORPG about drama, strong characterization and social interaction. For the game there was a general idea that more than 200 gestures per character was to be used. The cast in the game was gods and spirits from different pantheons. The brief stated that the game was to be in first person view. The input to the system to communicate with other characters was either spoken, using voice over IP or written input using the keyboard. The game was primarily aimed for the PC and Mac. There was no exact specifications on what the game at large was going to be about. There was not any exact definitions on game play elements and technicalities. However the intention was to have a magic and metaphysics system, a fighting system–although the gameplay was not to be about fighting–and an inventory system. The game was going to be run as a persistent on-line world.

The first characters designed and specified for the game was Bast an Egyptian cat goddess, Toth, Egyptian god of the moon, magic and writing. Other characters for the game where Thor, the Norse god of thunder, and Loki, half-giant half-
god, notorious rascal of Norse mythology. During the design process scenarios with these and sometimes other characters where used as examples to illustrate ideas and exploring the pros and cons of different solutions. These were also the characters that was talked about in discussions in the studio about the game.

During the design process the brief was changing back and forth. There were discussions in the team if the game was going to be massive or not, if the game play should be session based or persistent. Also there were big gaps in the brief e.g. no clear picture of how magic should work, no fighting system, and no clear idea on how to generate quests or similar structures to provide the players with momentum to start playing and exploring the game world. Because of this the design work for the interface could be concentrated on specifying an alpha version prototype for avatar control supporting emotive communication using verbal and nonverbal communication. The expandability of the prototype for inclusion of inventory, fighting, and magic systems was also important during the design.

The goal of the design process was to provide an interface for controlling the avatars in the pre-alpha release of the game. This version was only going to have two player-controlled characters, or avatars, played over a network connection in a small virtual world. From an evaluation of this first computer prototype the design process could continue with a more solid specification of the game.

Analysis of predecessors

To better understand the design problem and what had been constructed earlier when facing similar design problems the major MMORPG on the market was analyzed. Further some other games were analyzed as they were thought to have interesting elements for certain features of the game interaction.

The analysis had somewhat of an open end and was not finished until the design was finished as it was useful during the whole design phase to go back and expand on certain features of games and to find arguments for and against certain paths in the design space. But for practical layout and readability reasons the major part of the analysis of the predecessors is presented here.

In the descriptions below there is little or no information on what the games are about. The text focus on the qualities has been used in the design as arguments against something or as inspiration towards something.
**Everquest**

One of the first MMORPGs was *Everquest* (Sony, 1999), figure 1. The interface is visually unattractive with text, numbers and meters taking up much of the screen estate. The first person view makes it difficult to determine distances to other characters and objects. It also feels like the avatar is floating or is hovering about in the game world when moved. Consequently Everquest provided strong arguments against first person view during the design of the Ouroboros interface.

Interaction with non-player characters only uses keywords which easily felt like interacting by a computer command prompt; even more so as non-keywords were ignored by the game. The game was controlled by keyboard and mouse and the player frequently has to switch between input devices and this can interrupt the game experience.

![Everquest interface](image)

*Figure 1. Everquest.*

**Neverwinter Nights**

The game *Neverwinter Nights* (Atari, 2002), figure 2, is not massive in the sense that many players inhabited the same world at the same time; it could perhaps be classified as semi-massive, semi-multi on-line RPG. The game consist of a portal
in which the player can decide which game world to enter. Other players often create the game worlds and the game has a strong link to the classic role-playing game *Dungeons & Dragons* (created by Gygax and Arneson and first published by TSR in 1974). This was indeed the game that had the most role-playing going on. That there were players who actually engaged in role-playing in an online computer RPG suggested that there could be an interest in the type of game presented in the design brief.

Parts of the interface is very obtrusive. Especially when a pie-menu is opened and navigated to find gestures or other options disrupts the game experience. The pie menu opens on top of the game world and sometimes block the view. The icons are hard to interpret and sometimes the gestures selected were not gestures but pre-recorded speech.

To move the avatar around the player mouse-click and the avatar moves to that location. When the avatar is moving by itself like that it has great potential to break the illusion of a link between the player and the avatar. In that aspect of the game the player act more as the director or commander of the avatar and not as the actor.

The player-to-player communication is adapted to the gameplay experience and players can send messages that are heard by all players in a game world or talk to members of a group despite distances between avatars. The game puts much emphasis on acquisition of objects and experience points in the same way that its paper based parent does. The input devices for the game is keyboard and mouse. Often the avatar is moved around by mouse and the camera viewpoint is controlled by keyboard. As with Everquest this requires switching between input devices although not as much.
Anarchy Online

The MMORPG *Anarchy Online* (Funcom, 2001), figure 3, has one annoying feature. If the player is to send a message to another character the player is first required to press the spacebar on the keyboard to enter chat mode. This can often and easily be forgotten and hard to get used to with the result that typed messages is lost. The feedback that something is wrong comes when the avatar jumps when the player presses enter; the enter key is used for sending messages in chat-mode but for jumping when not in chat mode. Not only does this require the message to be typed again the inappropriate jumping is visible to other characters and players of the game.

The interface for the game is somewhat similar to a typical Linux GUI where there is many semitransparent windows with figures and text. This can change the player focus from the game world and make the game take place in the semitransparent windows filled with text, numbers and meters rather than in the virtual environment. It is apparent that how the game is presented to the player influence where the attention is directed.
Anarchy Online is a 3rd person view which made it easy to see where the avatar was in relation to other characters in the game.

Mouse and keyboard is used as input devices requiring frequent switching between them.

In *The Legend of Zelda: The Wind Waker* (Nintendo, 2002) there is strong emphasis on the facial expression of the characters, which enhance the players concern for the characters of the game. The visual style of the game is cel-shading which brings about associations of cartoons and comic books.

The view is 3rd person and the camera is following the avatar most of the time intelligently providing good shots to tell the story. The camera can easily be controlled by the player if not behaving as desired. The avatar also looks at objects that it is possible to interact with thus communicating to the player that there might be something of interest. This is helpful for things the player has not noticed and for things that is not visible on screen or hard to see.
The game uses context sensitive avatar control and representations of the buttons of the controller are visible on-screen e.g. when walking by a boulder the choice of lifting the stone will be presented to the player by the A-button label reading ‘lift’ and if the player presses the A-button the boulder is lifted. After lifting the boulder the A-button label reads ‘throw’ and the R-button label reads ‘drop’.

Interaction with other characters is very simple and often they repeat the same phrase over and over.

When the player character was low on energy it was apparent not only by meters and sound alerts but also in the body language of the avatar.

As with almost all console games the game use only one input device. It is so easy and natural to use and the player can focus completely on playing the game watching the screen and the characters presented.

**Resident Evil**

The horror adventure game *Resident Evil* (Capcom, 2002), figure 4, has a very minimalist interface. In the game there are no meters or numbers visible there are only the avatar and the game world. The game uses fixed camera positions and by sometimes providing a visually less informative view of the avatar the gameplay is actually improved. The game gestalt successfully resembles that of horror movies.

The game used avatar body language to communicate about the health of the player character and gave information about the surroundings, e.g. the avatars moved differently near a shot zombie about to get up again than close to a dead, or rather a destroyed, zombie which is not a threat any longer. This makes the player focus even more on what was going on on-screen.

The player only has one input device which quickly becomes transparent to the player.
Grim Fandango
In *Grim Fandango* (LucasArts, 1998), figure 5, the avatar looks at objects that can be interacted with by moving his head and eyes. This game serves as another great example on how the avatar can communicate with the player using body language and how it can communicate things about the game environment to the player.

The only input device needed is keyboard, but the game can be played using a joystick and keyboard.
The animation of the characters in ICO (Sony Computer Entertainment, 2001) communicate their feelings in a subtle way and this influence how the game as a whole was perceived. This game is one of the few examples of games where the emotions of the characters are a major part of the gameplay. For the design case ICO serves as proof that there are computer games with strong characterization and drama.

The control and on screen interface was minimal and consequently does not interfere with the strong emotional gameplay.

**Animal Crossing**

In the game Animal Crossing (Nintendo, 2001) symbols are used for communicating the emotions of the NPCs as well as facial expression. A thinking NPC will have cogwheels turning, new ideas is signaled by a lightbulb, happiness by flowers, love by a floating heart. The effect is somewhat lessened by the fact that the NPCs has a very repetitive dialog and often say the same thing as another character had said before. However the notion that body language in games does not have to limit itself to the bodies of the characters is useful.
**The Sims**
The Sims (Maxis, 2000) and the expansion kit The Sims: Hot date (Electronic Arts, 2001), figure 6, use of body language and symbols floating over the head of the characters to communicate to the player what is going on. A part of the game is to select the appropriate social interaction for the avatar, especially in Hot Date where the player has to decide on things like whether the avatar was going to ‘Shake Hands’, ‘Kiss’ or ‘Hug’ when greeting another character. The Sims serves as a good example on how social interaction can make new and interesting kind s of games.

The game is mainly played with the mouse sometimes using one hand at the keyboard mainly for command shortcuts. Complete switching between input devices were seldom necessary while playing.

![Figure 6. The Sims.](image)

**Phantasy Star Online**
Phantasy Star Online, Ep. 1 & 2 (Sega, 2002) has an icon based chat and universal translation system so that players could communicate the most common messages to each other quickly and without worrying about language barriers. The game has no pausing; when a player wishes to change weapon or do something else that in other games often require entering a sub-screens this game is still playing while the player is browsing the inventory of the avatar. This adds much
to the gameplay especially as the avatar still can move around while browsing. E.g. trying to activate a healing potion for a fellow player while trying to keep clear of attacking beasts is not an easy task but it adds much to the excitement and in a sense to the realism of the action sequences of the game.

Soul Calibur II
Button mashing fighting game *Soul Calibur II* (Namco, 2003) shows how input frequency and controller type could support the gameplay as well as communicate the interaction gestalt of the game.

The game is also of interest since the characters are animated using motion capture technology.

Unreal Tournament
First person shooter Unreal Tournament (GT Interactive, 1999) shows how input devices can become transparent to users of the game. The first person view works well in this kind of game.

Qualities-in-use
From parts of the analysis of predecessors and the theories described in chapter two qualities-in-use was extracted to be used as objectives to orient towards when navigating the design space.

- **Immersive.** Clashes with reality and the interface should be avoided. Flow experience both in the interaction with the interface and in the game experience.

- **Entertaining.** The game has to be fun and meaningful to use.

- **Role-playing and acting.** To achieve the desired game play experience specified an important quality has to be acting. The player should be able to act out what the player character is feeling.

The three qualities-in-use are interlinked and overlapping however they can individually be expanded upon to better show what was thought to be important for each quality.

Immersion
The interface should support the game in being immersive. The interaction should be elegant and seductive in a sense; it should take the player to the side and direct attention away from the normal world and immerse the player in the game experience. To achieve this the appropriate input frequency has to be im-
The controls and the interface has to feel right when it is used and there should to be no clashes where the player has to waste cognitive energy trying to figure out how the interface is working. The visual representation is also important for keeping the player immersed. The player should be kept in the layer 2, playing the game, as much as possible.

**Entertainment**

To get the player immersed the game also has to be entertaining. The story and the characters must be interesting and the interactions meaningful to get the player playing and to keep the player playing. The game is good if it provides pleasure but it is best if it is enjoyable. The game should provide the player with challenges of the appropriate difficulty. Playing the game has to be rewarding for the player.

**Role-playing and acting**

In order to have the player playing a game with strong social interaction the interface and gameplay has to provide the right feedback for the player and enable the right kinds of interaction. The interface affords a certain play-style that the player will pick up on. The interface also should support empathy for characters in the game world and they have to be projected as individuals. Interaction for using body language has to be economic to get the players to use it. The gaming aspects of agon and mimicry should be maintained where the players try to out do each other (agon) in being the best actor (mimicry).

**Avatar body language communication types**

The avatars body language can serve several purposes. One basic function of the avatar body language is to provide feedback that input is working e.g. tilting the control stick or pressing a key causes the avatar to move. Temporary loss of control of the avatar is very frustrating and can be used to create a more exiting game experience (temporary loss of control is also a common feature for player punishment in games) but can also be misused and break the link between the player and the avatar and thus destroying immersion and ruin entertainment and role-playing aspects.

Instead of having meters and numbers for injuries the avatar can move as if injured (as in Resident Evil) or tired, dizzy (The Legend of Zelda: Wind Waker). The avatars body language can also communicate to the player things about the game world. In Resident Evil the avatar moves differently near a destroyed zombie than around one that is about to get up. In The Legend of Zelda: The Wind
Waker and in Grim Fandango the avatar looks at objects that can be interacted with which sometimes the player cannot see. The only way of detecting some objects is by watching the body language of the avatar.

The nonverbal communication of other characters encountered is used by the player to make assumptions about these characters. The size, appearance, gender, weapons, manner and poise of other characters are used for making assumptions about the personality and function of the characters. In multiplayer games the visual cues are even more important especially if the signals are explicitly communicated like gestures and body positioning and the player recognizes that the signs the other characters are displaying are sent to communicate something.

Designing for MMO
When designing for a massive multi-user online game there are some things that have to be dealt with differently than in single player or multi-player games.

**No pausing**
In massive games it is not possible, or appropriate, to enable pauses in the game as it is in single player or small (2-4 players) multiplayer games—the more players the more pausing and more frustration. It will be too disruptive for all the players to be put on hold repeatedly. Thus it can be no pausing in the same way as in single player games where the game usually stops when the player enters e.g. their inventory sub screen. Phantasy Star Online has solved this in a good way; the inability to pause actually adds excitement of the game experience.

**Entering and leaving**
Multiplayer online games have to be able to handle players joining and leaving and even being disconnected. Therefore the game has to provide some sort of entering mechanism so that the avatar enters the world in a manner that does not interrupt the game play and when a player leave the same thing applies; it also has to be dealt with what to do with a player avatar when the player for some reason is disconnected from the game. Again Phantasy Star Online (PSO) is a good model on how to enter a game. In PSO the player first enters a portal where the player can chat and find companions to go out on a mission with.

Much of the excitement, horror or possibly frustration when playing Resident Evil is that the game can only be saved (entered and left) at certain locations. Having to make it back alive to these safe rooms adds much to the gameplay. In
Animal Crossing the player is punished for not leaving the game following the right procedure—thus preventing players resetting the console without saving unwanted outcomes in the game.

**Communicating**

In games spanning over vast distances of virtual space and with a large number of players it has to be decided upon some sort of screening process to limit the number of received messages from other players and NPCs. Message overload has to be prevented while the messaging system still is being supportive enough to enable the player to receive enough communication to enjoy the game. If it is difficult communicating within a quest group the game will not be enjoyable. For example can members of the same team in Unreal Tournament hear or see each other’s messages, in Neverwinter Nights players can communicate within a party unaffected by distance in the game world.

Another challenge for making an MMO-game is that of language and language skills. Role-playing games requires good communications skills from the players and those unskilled in the game language of choice will be left out or will not have an as fulfilling experience as language skilled players. Bad actors can possibly ruin, interrupt, or disturb the positive experience of other players. It is difficult to monitor player behavior and restricting communication abilities. If this is done it also limits the acting aspects of the game. Especially in a game where the gameplay might require the avatar to use offensive language shouted in capital letters.

Perhaps even worse than bad acting are conversations taking place in the game world that is on a lower layer than the game world. Talk about real world matters should be kept outside the game world. The need for off-stage communication should be handled outside the game world and it could be of benefit to provide a backstage or off-stage area for the players maybe combined with the entering and leaving mechanism similar to that of Neverwinter Nights. With a construction like that players can hang out with each other before and after each game session much like the off-stage areas in a LARP.

**Design Considerations**

Since the design problem was presented rather loose and was changing some important design decisions had to be made to narrowing down the design space and to close in on a solution. These choices are represented one at the time but in reality they where not treated in this manner. The design choices represent de-
sign work in divergence, transformation and convergence. Even so it is more practical to collect the thoughts about different aspects and explain work on them one at the time.

**Platform**

**CONSOLE vs. PC**

Designing for both consoles and computers is difficult since they have somewhat different players, market, input devices and network technology. Although they are converging it was thought as they was not close enough to make one general enough solution for both platforms. Console technology still requires solutions of its own; probably every console has to be treated as platform of its own as well.

The design choice was to use PC as platform but to keep in mind portability to console so that the design solution did not include something that was impossible or very difficult to port to console.

**Input**

**SPEECH: KEYBOARD vs. VOICE**

Using PC technology there is a choice of communication avatar speech by either entering text using a keyboard, or using voice-over-IP technology. For using voice-over-IP speech distortion was a requirement, or there will be limitations on what characters a player can choose. This being an interesting choice for communication it still required special technology and the project team did not know if there was time or resources to ensue that research. So the choice for input was made to use keyboard input similar to most chat systems.

**MOVEMENT: CHOOSING INPUT DEVICE**

For the movement of the avatar doing both gestures and general movements around the world, e.g. walking and running, the input device of choice had to be elegant and support the gameplay. To focus the players’ attention to the game the input device had to be easy to use and feel natural. Since the input of speech was using the keyboard it was thought to be better to use the keyboard as the only source of input for the player—eliminating mouse, joystick, trackball, and what have you that would force the player to switch between input devices.

When looking at other computer games it was always disruptive when one hand had to move to the keyboard for typing and then to the mouse for clicking a button or doing some other movement and then back to the keyboard again. The games that can have the hands in fixed locations are much nicer to use (all console games, Sims, Unreal Tournament, Grim Fandango). It was judged as important to in the same time to be typing a message and reading the screen to be able
to move one step closer, have the avatar smile in response to what another player character was doing without having to switch input device and have the player searching for the mouse.

Gesture recognition and other input methods that use devices with accelerometers, video cameras, trackpads or other touch sensitive surfaces might provide some interesting solutions. Such devices were not available at the Zero Game Studio and consequently they were not explored further.

**View**

**FIRST PERSON VS. THIRD PERSON**

The initial design brief stated that the game was to be a FPA, first person actor, game. Despite the catchy phrase (compare with the common genre of FPS, first person shooter) thinking about it a game with acting is much better if the player can see the avatar. Looking at the predecessors first person view is suitable for some action games but for a drama or role-playing game where the avatar is performing gestures the player will need much feedback about how the gestures are being performed. This feedback is not available in first person view. It is also hard to determine distance in first person view. The player needs to be informed how a gesture was looking when executed and where and how the avatar is positioned.

The choice for point of view had to be third person and sometimes where it is suitable going into first; e.g. when examining something up close.

Another really good thing with having a third person view is that it enables camera positioning.

**CAMERA CONTROL VS. CAMERA AGENT**

The player view of the avatar and the game world is determined by what can be called a camera. There is a choice between giving the player full control over the camera or no direct control over the camera. There might be indirect control as the camera has to keep the avatar in view i.e. moving the avatar will often affect the camera. If the game world is very limited and similar to a stage in a theatre the camera can be fixed.

For a game about social interaction and drama it is better to use a camera agent. This provides the game with the opportunity to use cinematography principles to add to the style of the game; which is done with excellence in Resident Evil.

Having a camera agent also frees the player from having to control the camera. The inability to directly controlling the camera also reinforces the link between the player and the avatar; as the only choice the player has to get to see what is
going on somewhere is to move the avatar to that location instead of perhaps rotating or tilting the camera.

The negative aspect of having a camera agent is that the camera agent has to be smart enough or it will have a great chance of interrupting game experience for the player.

There is also the option to sometimes let the player control the camera to override the camera agent that is good for games with puzzle or maze elements like Ico and The Legend of Zelda: Wind Waker. If the gameplay were to expand and include such elements it should be considered if the player should get to control the camera agent when appropriate.

**Avatar control**

**AUTONOMY VS. AUTOMATION, PART 1**

One issue to resolve was how much ability to control the avatar and the game world the player should have.

The player could be given control over everything from eye blinks, pupil dilation and heart rate of the character. This would lead to a very complicated control system that will take too much energy to use or even learn. Another option is that the game engine could use an AI procedural attachment to determine what the avatar does and when. With too much of this the game soon cease to be a game and starts becoming a virtual theatre where the player is no more than a spectator. This might be interesting but it would no longer be a game. The best solution was found somewhere in-between the two extremes. The game engine should help the player with the petty parts of avatar controls while still allowing the player great freedom of expression.

But the one question had to be investigated further: How much can the system understand—especially about affective interaction between two characters?

**PARSING FOR EMOTION?**

One interesting aspect of the design problem is that of parsing game context for emotional content. If the system could determine what mood the avatar should be in it would be a helpful for the player and interesting for the gameplay. However it is extremely hard to parse a context with player characters for emotional content. The more freedom of expression the players have the harder the parsing would be as natural language, especially in combination with gestures and non verbal cues, is full of ambiguities and is used for sarcasm, irony and jokes, etc.
The system could also parse for the changes of relationships between the characters and for example make it impossible to use certain gesture towards characters to which the avatar has certain kinds of relationships. The problem with such a solution is that the system might not be able to keep up with the jargon and gestures of the players and therefore interrupt the game experience.

Another interesting aspect is the personality of the avatar changes from what interaction choices the player makes. Experience might enable some new gestures and alter the avatars walk cycle. Here it is easier to parse the interaction style of the player and apply a fail-soft approach as it will not be visible to the player when the system is making the “wrong” decisions. However the player might enjoy the game more if the control of avatar development is more in the control of the player. Even if it is an interesting idea that the style of the player changes the avatar in the long term it has no effect on the design of the interface, especially if the mechanism is hidden from the player. Thus long term character development was left out from the design of the interface for the time being anyway.

Apart from emotional parsing is the game physics engine that also will be doing a parsing of some sort. The avatar might slide if running and stopping quickly on low friction terrain. If the avatar is affected by external forces or violence, and collision detection should be handled by the game and not the player. If the avatar is hit in the head by a rock, shot by an arrow or runs too close or perhaps over the edge of a cliff the system not the player should make the avatar perform the appropriate body movements.

To sum up either the system was going to use parsing for emotional content and making the choices, or provide a small set of choices, for the player; or the player was to be given much freedom and there would be no parsing except from that of the physics engine. Before moving to a decision some thoughts about games and acting has to be further investigated.

THE PLAYER AS THE ACTOR

If the game is about strong social interaction and drama then the game can be classified as a sort of mimicry game, which is very different from most other computer games, of which most has to be classified as competitive games (agon) or games of chance (alea). The object of the game for the player is to experience the other world and in doing so providing others with a similar experience.

There is a passage by Monti (1969, pp. 9–15; in Csikszentmihaly, 1990, p. 256) concerning the uses of West African ritual masks that is useful for explaining the experience of mimicry; coincidentally relevant for a game which in which the players is acting the roles of gods and spirits:
Design

“From a psychological point of view the origin of the mask can also be explained by the more atavistic aspiration of the human being to escape from himself in order to be enriched by the experience of different existences—a desire which obviously cannot be fulfilled on the physical level—and in order to increase its own power by identifying with universal, divine, or demonic forces, whichever they may be. It is a desire to break out of the human constriction of individuals shaped in a specific an immutable mould and closed in a birth-death cycle which leaves no possibility of consciously chosen existential adventures”

In one aspect there is no winner in the same way as there is no winners among actors in a theatre or singers in a choir. But in an indirect sense the agon quality of games is evident in that everybody wants to be the best actor or the best singer. Respect and admiration is paid to the person who is making the best performance. Consider the awards given to actors; it is not always the actor who has been playing the most heroic character, nor does the actor portraying the character who has killed most enemies and found all the lost treasures that gets all the awards and respect. The awards is given to the best actors; who often gets to play the heroes. This aspect is essential for making a game about drama and interaction because there has to be stories and plots as frames and foundations to build the drama on. In a good plot or drama not all characters can be the classical heroes.

AUTONOMY VS. AUTOMATION, PART 2

The player should be provided with the best tool for being able to be the greatest actor. Not only is extensive parsing for affective content in a game context difficult it would also lessen the opportunity for the players to make impressive performances. The controls should be elegant and provide the right amount of support for the player. The interpretation of the character and acting should be left to the player as the major part of the gameplay experience. This also provides for a game design that will enable the player to enter a state of flow and makes the player less passive. It also makes the game more demanding, which is both good and bad. The demands have to be small at first and then escalate. New players should be given easy performances and their roles can increase with time and rising skills of the player.

If the players are the ones in charge of the performance it is in agreement with Clarks principles for layers. The players will be able to appreciate each other’s skill in performance on one level and use their imagination in experiencing the higher layer.
The design choice was that the player is in control to a large extent of what is going on. Physical things like injuries, breathing and such is to be handled by the system. There will be automatic reactions by some characters to some events is also to be dealt with by the system. Perhaps some characters will jump when there is a loud bang and some of the cooler ones will not. To determine exactly what behaviors is automated and what behavior that are left out altogether the game has to be specified in greater detail. A compromise is also possible in that newbies can have access to the whole array of communicative signals but the system can highlight the choices calculated as appropriate. If it is possible to make such a system it is important that it can be deactivated as to not disturb experienced players. Some actions should not be available in acting-help-mode to increase the desire of the player to play in expert-mode to impress the other players.

Without parsing it is an art to act as a liar and get caught. It can be part of the fun of the game to see which player can act the best liar. In a game with extensive parsing and little player control that element of gameplay would be lost; and play-acting is at the very core of emotive communication.

Something that has to in part be parsed to detect is joint gestures, like kissing, hugging and holding hands. In such gestures one character has to request participation to perform the gesture or perform another gesture, which is quite different by force. A hug that is answered is different from an unanswered hug. One solution is to have the players see the invitation or request made by the character as it “moves in” for the joint gesture. Backing away or standing still will not answer the gesture and it will either be stopped or forced. Meeting, by for example moving the avatar slightly forward will answer the joint gesture. This enables a delicate sense of control and increases the focus that has to be kept on screen. To miss an invitation to a joint gesture can have serious effects. Getting hugged by an enemy or missing the opportunity to kiss a love interest is not preferable.

**Visual representation**

**SPEECH**

After it was decided that what is said by the avatar is represented by text typed on the keyboard by the player there was an issue on how to represent this text.

In most games with text representations of speech the dialog appears in a window and is visually very similar to an Internet Relay Chat (IRC). The text is preceded by the name of the speaker and there is often the possibility to scroll back in the chat window to see what has been said previously. Another option is to have speech bubbles, like in Phantasy Star Online. The speech bubbles visually
link what is said to the character speaking and the player does not have to link name and text to the image representation of the characters and their gestures. Having speech bubbles however take up much of the visual space and the player can have a hard time reading all the conversations going on in a room crowded with avatars. This is not necessary a bad thing. If trying to immerse the player in the game the attention of the player should be focused on the game. The game requires attention and the attention is directed at what is going on on-screen, on-stage, rather than watching a chat window as well as the stage.

Although a windowed IRC-style chat has benefits the benefits of a more evanescent representation of speech bubbles was decided to enhance the gameplay in part by their “bad” properties. Speech bubbles are closer to the basic setting of language use in the game. Perhaps the log could be available for discussion off-play, off-stage.

For the design problem it was more elegant to show onscreen who said what and in what style. Not knowing the name of who said it and letting the statement fade out rather quickly makes it hard to follow many conversations going on in one room and not seeing what is going on behind the character and not hearing much of it either. Speech bubbles are good for doing this. However they are a very distinct style element that might not be suitable for different designs. The design of the game was moving towards an art-nouveau style and from talks with the art director it was concluded that not only could speech bubbles work but they could contribute to the desired graphical appearance and atmosphere of the game.

Another thing about speech bubbles is that they can add information about how something is said and even what is not said if that would be the case with think bubbles or similar constructions used for telepathy.

The time a text is visual has to be worked out but something close to the time of real time speech seems to be a good starting point. This is also good if the text is to be phonetically parsed to properly lip-sync the avatar.

**Body Language**

In animation the laws of physics can be broken. What it meant for representing emotional states of the avatar was that not only could naturalistic body language be used. There where choices to have the avatar bend to have anime-like facial expressions changes in color and size and even physical representation all together. An angry avatar could morph into a dragon for example. Using symbols along with body language also works just as it does in comic books. An avatar thinking greedy thoughts might have dollar signs for eyes and a depressed avatar
a black cloud over her head. Another interesting effect for communicating emotions is that of using auras and filters; e.g. could an angry avatar could be glowing of red, a sad avatar could have the world around it turn grey. These special effects were left out for the pre-alpha version and do not really influence or add to the design solution. Controlling a gesture is controlling a gesture whether it is glowing lights or waving arms the input system is fairly the same. What signals that is available for the player to use is to a large extent based on the desired game style. Resolution is also a factor. Subtleties of smiles for example have little value if they are too hard or impossible to discern; unless the desired game play includes some aspect in which is necessary to differentiate deliberately made false smile and real smiles.

**Gesture set selection**

Body language is too big and complex to include control of all aspects of it in a game. It would be difficult for the user to learn and control such an interface as well as impossible to implement, animate and build an interface for such control. Of course there has to be found a range where the interface allows enough control to serve the purpose of the game and the abilities and skills of the player. Therefore the interface should enable control for the gestures most important to the game play and other parts of body language has to be left out.

For the Ouroboros prototypes this meant that the body language was such body language that was useful for social and emotive communication. Important body language in face-to-face social interactions are body positioning. Where one stands and what is looked at is important. Also body language that is used in response to others leaning back or forward, nodding, folding arms, etc. Social gestures such as greetings (friendly, neutral and hostile), goodbyes and expressions of ones emotions smiles, frowns, thumbs up, hugs, kisses, yawns are also important. There should probably also be some gestures for playing and goofing around such as dancing, burping and pretending to falling asleep. Most, or probably all, gestures can be used in many different situations and for many different meanings. Context will have a strong impact on how body language will be interpreted. To be easier to access body language was broken down into categories: appearance, body positioning, facial expressions, body posture and gestures.

**Appearance**

Gender, clothing, race and so on will have affect how the avatar is perceived. These signal are not something that player can control, at least not very often after character generation or selection. However appearance is a very important
signal in the avatar face-to-face communication and it will influence how everything else is perceived and interpreted.

**BODY POSITIONING**
Most basic is the need to move the avatar around by walking and running. It is also important to be able to control exactly where the avatar is standing and to be able to turn on the spot to show where attention is directed.

**FACIAL EXPRESSION**
The most communicative part of the body is the face. It is especially good for conveying emotive information. The player need to control if the avatar looks happy, disgusted, surprised, angry, etc.

**BODY POSTURE**
Some gestures that are being maintained for a longer time and that can serve as a starting point for different gestures, predominantly hand gestures, are different body postures. If the avatar is standing with arms folded, is sitting on the ground, assumes a fighting pose or is dancing sends out different messages.

**GESTURES**
Using the postures as a starting point the player needs to be provided with some choices on what gestures the avatar can preform. Based on the posture there will be different gestures. If avatar is dancing it will greet differently than when sitting on the ground or standing with arms folded. But there is the need for similar types of gestures from the different poses.

**COMBINATIONS**
Using combinations of the five types of body language control the player will have great possibility to for controlling the avatars body language when communicating in the avatar face-to-face situations that will arise in the game.

Prototype building and testing
During the design space analysis a pattern of what was felt useful was taking form. The next step was constructing and testing the pattern with prototypes.

**Prototype 1**
**CONSTRUCTION**
The first prototype was built for testing how players would react to and use an interface where it was possible to control the body language of an avatar. A test was also needed to guide speculations on what gestures are desirable for the players. Gestures selected for inclusion was mainly of two kinds emblems, gestures that can be used instead of words, and certain kinds of iconic gestures. The
interface was constructed to be able to handle more than 200 specific gestures controlled from a keyboard.

The gestures where sorted into groups and subgroups to form a hierarchical structure. The paper prototype, figure 7, was on purpose built to look shabby so that the users did not avoid criticism out of politeness. Also little effort was spent with the paper avatar as to avoid focus being put on how the gestures were implemented and how they looked. The prototype was built to test the interface and interaction using the keyboard to control the GUI and the avatar.

**Testing**
The test was made to see what the player did with the availability of gesture control.

The test environment was the interaction design studio at Linköpings University. It was tested on two persons both fairly experienced computer game players.

Starting out the test persons were instructed to make the paper prototype avatar perform some simple gestures that was explicitly specified. This was done so that they got a feel on how the interface worked and what gestures was available. The test persons was asked to make the avatar smile, jump, cry, etc.

Next came scenarios. The test persons was read a scenario and another character was introduced on the screen. E.g. “you meet your friend” or “you meet a stranger who is rude to you”.

After the scenarios the test persons where asked some questions about what they felt about the test and the prototype. They where encouraged to come up with ideas and to provide criticism of the prototype.

During the test and the discussion short notes where taken. Immediately after the test person left these notes and the test situation where written down and reviewed.
RESULT

Sorting the gestures into subgroups and having a different sets of gestures each emotion made it hard for the test users to learn where to find a desired gesture with the right emotional category.

However it was apparent that the test users found it fun to use gestures and said that they would like to play a game or use a virtual chat where they could use gestures to communicate. To categorize gestures according to emotional state did not work as well as hoped.

Prototype 2

REDESIGN AND CONSTRUCTION

The gestures and buttons were regrouped. The facial emotional expressions no longer determined the available subgroup of gestures as in previous prototype. Instead facial expression could be selected independently from posture. For each posture there was a set of gestures that used the posture as a starting and ending point. In this prototype, figure 8, the player could change the facial expression without changing the posture of the avatar or the player could change the posture without changing the facial expression.
For most of the postures the gesture set maintained the same meaning for the gestures. For many postures there was a yes-, no-, maybe- and greeting-gesture available at the same location in the interface.

For the new prototype some new gestures were also included.

The tab-key switched the active menu between control for face, posture and gesture. F1-F11 were used for selecting facial expression, posture or gesture depending on what menu was active. F12 presented more choices for the current menu; i.e. more choices for face, posture or gesture.

**TESTING**

The prototype was tested on five persons. Age ranging from 19 to 35 and gaming experience from little to extensive. Some of them had role-playing or acting experience. The test environment was in a small conference room at the Zero Game Studio in Visby.

The second test followed the procedure as first. First the test user performed simple gestures with the avatar, then acted in small scenarios and the test was concluded with an interview. Notes taken during test and interview and immediately reviewed after each interview.

---

*Figure 8. Prototype 2.*
RESULT
The test users did not object to the button placements as was the case with the first prototype. The interface was also very learnable. Test users often did not bother to look at the interface and after a few scenarios knew the key combinations for their favorite gestures. Instead of waiting for the interface being updated (which can slow when working with paper prototypes) they just quickly tapped the appropriate keys. Categories and gesture set felt good for the scenarios and for general acting. All of the test users reported that they found it fun to use gestures to communicate and would like to play a game where they could use body language to communicate. The test users seemed to enjoy using avatar body language as a mean of communicating. It was remarkable how creative they were with coming up with responses in the scenarios and how different they used gestures in combination written input.

Specification
After the second prototype test specifications for a computer hi-fi prototype was handed over to the technical team for implementation in the pre-alpha release of Ouroboros along. The gestures to be used where determined and the animations where directed and recorded using motion capture. The look-and-feel of the interface was decided upon in collaboration with the art director and drawn by the lead artist.

The following is a short version of the specification for how the interface should work:

• Avatar speech appears in speech bubbles.

• The interface was built for using keyboard only. The avatar is moved about the game world by using the keyboard numpad. See table 1.

• Modes for gesture interface are: face, posture and gesture. The posture selected determines what gestures can be browsed and executed in the gesture mode. The face mode is for controlling facial gestures and expressions. This enables eleven gestures, or postures, or facial expressions, are available in each mode. Pressing F12 displays the next set of eleven gestures that are available for the mode. See table 2.

• Body movements and gestures can be combined so that an avatar can lean forward and smile or turn away and frown.
**Table 1. Keys for controlling avatar body movement**

<table>
<thead>
<tr>
<th>Key(s)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tab</td>
<td>Change mode (Face, Posture, Gesture)</td>
</tr>
<tr>
<td>F1 through F11</td>
<td>Execute animation depending on set and mode.</td>
</tr>
<tr>
<td>F12</td>
<td>Change set</td>
</tr>
</tbody>
</table>

**Table 2. Keys for controlling the avatar body language interface**

With 1 set of 11 facial expressions, 1 set of 11 body postures each with 2 sets of 11 different gestures for each posture the interface handles about 2,500 different combinations of body language expression based on combinations of 262 animations. With 2 sets of 11 facial expressions, 2 sets of 11 body postures each with 3 sets of 11 different gestures for each posture the player has access to more than 12,000 means of expressions based on combinations of 748 animations. The sum of available means of body language communication greatly increases when combining it with more general body movements—moving backwards, leaning in, turning—and positioning—standing close, side by side, face to face.

Some things where left out due to time constraints and lack of technical skills within the team. How the camera agent works had to be left unfinished. For the pre alpha the camera position was fixed behind the avatar. The specifications did
not require gestures to blend with emotions and there where no happy or sad
gestures for example. Personality changes and changes to the gesture set depend-
ing on play style and mechanisms for entering and leaving was left out. These
thing has to be worked out in parallel with the specifications of the gameplay.
Since there was only going to be two characters in a small world there was no
need to bother with broadcasting. Both players could see each others messages if
they where close enough and within view.

Other things that were not specified was how things such as magic and inventory
should be accessed or used in the game. However a system similar to the gesture
input system worked should probably work for other aspects of the game as well
– depending on the gameplay.

A screenshot of the prototype is presented in figure 9.

![Figure 9. Ouroboros Pre Alpha, partial implementation.](image)

In summary

The specifications for the game were changing or undefined, during the whole
design process. The gameplay at large was not decided upon which prevented
work on large aspects of was could be an interface for the whole game. This en-
abled a focus on how to create an interface which afforded emotive communication. Such an interface where thought to lead to a game style that is more about drama and social interaction between the characters. Ability to communicate using body language were thought to deepen characterization.

Analysis of computer games was used as to provide positive and negative examples as well as inspiration for the interface design and the game experience desired. Not only were predecessors in the MMORPG genre analyzed, games with other interesting aspects were also looked at.

The analysis of games and a literature analysis of game and communication literature resulted in a list of qualities-in-use. The qualities for the game were: immersion, entertaining and role-playing. During the exploration of the design space and the design problem different aspects of interfaces was found. The input of speech or text and control of avatar had to be elegant to enable the player to focus attention on screen. The view of the character should be a third person perspective which would allow better feedback to the player on how avatar gestures where performed. Having a third person view also would enable the use of an intelligent camera system which could use cinematic principles to improve to the game experience. The control system for the avatar should enable players to get skilled in its use. It would have to be easy to use but its complexity must enable players to learn increasingly sophisticated methods of communicating. Parsing for content could help newbies but must not intervene with the interface learnability and the possibilities of using ambiguities in language and emotive gestures when acting. The parsing will also have difficulties with handling complex emotional situations that the characters will come across. The gestures most supportive for emotive communication are iconic gestures and emblems. Speech bubbles are good for communicating not only who said what but also how. The evanescence of speech bubbles keeps the player focused and is more in accordance with basic setting of language use.

Prototype tests hint that players like to use gestures when communicating. It seems than an interface for more than 200 gestures can be constructed to be both usable, learnable and more important fun to use. It is also possible to implement the interface for a game.
5. Discussion

In this chapter the research question is answered and the results are discussed, criticized and reflected upon.

Result

*What are important design considerations when designing an interface for an MMORPG that allows the player to control the avatar body language?*

- Think of the player as an actor and the game world the stage.
- Using a 3rd person view enables the player to see the avatar and its relative positioning more accurately. This supports the acting and enables the avatar to communicate about the world to the player by its body language.
- 3rd person view enables the use of a camera agent that can enhance the game experience by using appropriate cinematographic principles.
- Let the player can choose from large number of emotive expressions and allow freedom in how to use the avatar body language together with spoken, or typed, language.
- Facial expressions and emblems are probably the most useful gestures for emotive communication. Joint gestures are also important.
- Relative positioning can also be a form of communication. Standing close to, turning away, etc communicate something.
- The player must be able to out perform other players and characters of the game. The competitive element of the game have to be focused on acting.
• The game also needs some sort of story to support the player and let them know what characters they are aiming to portray.

Thus the gameplay experience is more like LARP and in the realm of mimicry in that the game is about the experience of acting and interacting with other players and that the best actor is the best player. All the considerations for the interface build upon this intention. It is important to remember that much of these are specific for controlling the gestures of an avatar in a MMORPG that stresses characterization and drama. Designing an interface for an action adventure or something else might need different solutions but the considerations and choices made for in this thesis might be of help when designing other avatar control interfaces.

The game needs to set the stage for acting to emerge—without game mechanics, incentives, hints, bells and whistles the acting probably will not happen. For each avatar some sort of story or plot needs a to be stated or supported. This is also supported by the presentation of the avatar in the virtual world which enhance characterization.

**Significance of result for research area and practice**

The linguistic and semiotic fields of knowledge are highly applicable to the design of game interfaces for MMORPGs and understanding how (avatar) body language can be used. Theories of the nature of and strive to play explain why, and provide arguments for constructing, a game where players enjoy acting out stories with, or against, each other.

It seems that switching between input devices should be avoided. Switching devices is similar to an actor bumping into the decor on stage ruining the illusion for the audience.

Working with lo-fi prototypes and provotypes is a effective tool that provides useful information during game interface construction.

**Critique of method**

**CASE STUDIES & REFLECTIVE STUDIES**

The nature of case and reflective studies can be criticized for being (too) subjective. This has been a reflection on and the story of one case. Of how much value generalizations from this case have to other similar cases can be debated.

This case has been about the creation of the interface for a game in a research studio. As such the priorities and setting has differed from that of a real, i.e. commercial, game studio and this case might not be applicable to cases in the
computer game industry. The lack of specifications on the game design at large adds to this.

**Design**

During the navigations thru the design space the choices made might not have been the best. Another set of choices can lead to another different but equally or more satisfying solution to the design brief.

**Prototypes**

Testing prototypes there is always a risk that the researcher influence the users of the prototype during the test. To minimize the risk of users being too kind in their opinions about the interface it was made to look sloppy.

**Consequence of result for further research**

One problem with a large set of gestures and body language is how they are going to be animated. Motion capturing more than 200 gestures for each avatar in a massive game is not possible. Techniques for combining and tweaking data are available and should be useful.

Alternatives for input devices are of high interest. Face analysis that can detect emotions from the users webcam and other motion or image detection input devices can be of interest for creating games with players acting.

There are currently not many games that enables acting, complex social interaction and drama. That the users of the prototypes enjoyed using body language as well as the large interest in LARP suggests that mimicry style games could have a market.

It is also of interest how gestures and text can be parsed for emotional content which could make the NPC of a game more interesting to interact with.

**Analysis of antecedents**

A brief look at some of the games that has come after the work on the prototypes for this thesis show that some of the design choices has been implemented in other games.

**World of Warcraft**

For this MMORPG what is interesting is that uses a third person view with a camera agent. The gesture-control in the game is limited and the gestures and body language are with a few exceptions the same for all avatars. The gameplay is not so much about playing a role as it is about solving quests, gaining levels, items and reputation. All in line with what the bells and whistles of the game tell the player. Although the game is very social and features a great deal of commu-
nication it is not on a player character basis. The players are not using the avatars for acting out characters—the social activities are about players cooperation in solving quest and organizing guild activities.

**Sims 2 and Sims 3**
For the sequels of the Sims features such as goals, interests and aspirations of the characters has been added to provide incentives for the players on what to do next in the game. The gestures and body language are still large part of how the characters in the game communicate with each other and also to the player.

**Reflections on game design process**
It was difficult to work very loose specification for the game. It is probably easier for an interaction designer to work with either a rather firm game concept or to start working with the game development team at an early stage of the game development. In interaction design there is much knowledge that can be used for creating not only better game interfaces but also better games altogether. The designer should help design the game and the game specifications and constructing the interface—in that order. Coming up with a game is much about coming up with an interface. There are no good games with useless interfaces and interactions.
6. References

Literature

Aristotle. (350 BCE). *Poetics.*
(From http://classics.mit.edu/Aristotle/poetics.html)


Morris, D. 1977. *The pocket guide to manwatching*


References


Games

Given in the format: Game name. Developer. Publisher. Release year. System.

Anarchy Online. Funcom. Funcom. 2001. PC.


Everquest. Sony Online Entertainment. Sony Online Entertainment. 1999. PC.


