The Virtual Storyteller: Enriching the Story by Expressing Emotions

- Master's Thesis -

Susanne Hendrickx
July, 2007

internal advisor
dr. Jennifer Spenader

external advisor
dr. Mariët Theune

Artificial Intelligence
Rijksuniversiteit Groningen

957

Abstract

This research project contributes to the development of the Virtual Storyteller of the University of Twente by studying emotional expressions. The Virtual Storyteller is a multi-agent framework that generates stories based on the emergent behaviour of character agents. The character agents in the storyteller contain an emotion model that makes them believable characters and the research strives for a natural and accurate translation of the emotions determined by the emotion model into natural language.

In this project the emotion model of the Virtual Storyteller is evaluated linguistically by studying a corpus of fairy tales. The corpus study shows that there is a mismatch between the emotions in the emotion model and the emotion lexicon in fairy tales. Some emotions of the model are only expressed implicitly in fairy tales and much variety exists in the frequencies of the expressed emotions, even within emotions that are regarded as a pair in the Virtual Storyteller. Based on our results we propose some adjustments to the emotion model of the storyteller and we identify lexical items, generally gradable adjectives, to express the emotions.

Gradable adjectives often give rise to conversational implicatures when combined with modifiers as in *I am not happy*. The interpretations of such expressions need to be clear if we want to use them in the Virtual Storyteller to express emotion intensities. We therefore test the theoretic assumptions of weak bidirectional Optimality Theory (OT) about the interpretations of such conversational implicatures by conducting empirical research. The experiment reveals what the interpretations of modified adjectives are, so that we can use these expressions in the storyteller. The results show that some interpretations cannot be systematically related to the weak bidirectional reasoning of OT. For some forms, there is no agreement about what the interpretation is and for other forms, weak bidirectional OT alone cannot explain the interpretations of the readers. We conclude how the theory of weak bidirectional OT gives the best fit to the experimental data.

Because we aim to relate emotions in the storyteller to other story aspects in a contrastive way contrast relations are studied theoretically. Just like Lakoff (1971), we distinguish between the semantic opposition and the denial of expectation. For each of these types of contrast, we discuss which relation between two propositions licenses the use of the contrastive connective *but* and how this restriction for the use of *but* affects the interpretation made by the hearer. This results in a theoretic claim about how the first proposition in the denial of expectation is weakened by the use of *but*. *But*-constructions for emotional expressions in the Virtual Storyteller are presented.

We propose how the language generation component of the Virtual Storyteller and the input to this component should be adjusted to construct emotional expressions using modifiers and how contrast relations can be used to relate the emotions to other aspects of the story.

Preface

This master thesis describes my final research project of the Master Human Machine Communication of the Rijksuniversiteit Groningen. This research was mainly conducted to contribute to the development of the Virtual Storyteller of the Universiteit Twente.

Due to excellent advice and help from my advisors, family and friends I have been able to work on this project successfully. I would like to thank a number of people in particular. First of all I am very grateful to my two advisors, Jennifer Spenader of the Artificial Intelligence department of the Rijksuniversiteit Groningen and Mariët Theune of the Human Media Interaction department of the Universiteit Twente. You have both been of great support to me by providing inspiring ideas and tips for the research and giving enthusiastic feedback on the thesis for improving the quality it. You stimulated me in a very positive way and I would like to thank you very much for that!

During the project I had lots of questions regarding the architecture of the Virtual Storyteller. Ivo Swartjes, who works on the Virtual Storyteller as a PhD-student, was always available for my questions, and I want to thank him very much for that.

For conducting my experimental research, I needed a large number of participants. To this end I sent an email to all AI students and to friends and family members and asked them to take part in the experiment. I was very positively surprised by the number of reactions to this email that resulted in a very large and varied group of participants. Thank you all very much for taking part!

Wilbert Heeringa helped me with the rather complicated statistical analysis of the experimental data. Thank you for helping me out.

For conducting a corpus study, Bart Cramer helped me using the Alpino parser that automatically selected all adjectives in the corpus. Thank you, Bart!

I also want to thank my parents for their patient confidence and support, not only during this research project but during the whole study. It was good to know that you were willing to help me at any time.

Finally, special thanks go to my boyfriend, Eric. We had fruitful conversations about my research topic and you read large parts of my thesis and helped me improving it. Thank you for your great support.

Susanne Hendrickx
Groningen, July 2007.

Contents

Chapter 1

Introduction

In our society stories have always been of great value. One may consider them to be a form of entertainment only, but they serve far more goals. Stories are particularly capable of communicating and preserving morals and knowledge as stories are easily remembered. In most fairy tales, for example, a moral message is hidden and this makes fairy tales very popular. If these stories are read in large numbers, they contribute to the standards of the culture.

Stories can also function as educative material. Stories often make matters much simpler by converting complex matter into imaginable scenarios. Beside that, stories can be applied in training pupils' problem solving skills. A scenario is then described and the pupil has to report what kinds of solutions exist, or the pupil can be asked to explain the situation.

Nowadays stories are more and more told via technical equipment like televisions (e.g. in films), and computers (e.g. in gaming). This equipment is able to present the story in a very lively manner, but the story creation is still left to humans. A successful story is not easily created. The story should be exciting, realistic (to a certain degree) and, most important, the reader or listener should be involved in the story. Automatically generating successful stories would be a great challenge and as described above, it could be applied in many fields.

The Computer Science department of the University of Twente (UT) has been working on the development of a virtual storyteller. The Virtual Storyteller is an agent-based system that can automatically generate fairy tales (Theune et al., 2002). The generated stories are based on events that take place in a simulated world with character agents. A formal story representation is constructed of the plot, which is then transformed into natural language.

The ultimate goal of the project is to generate stories that approach stories written by human authors. One element of the system that would make the story more natural is the use of realistic characters. Rensen (2004) has added an emotion model to the character agents of the storyteller in order to have the characters behave in a more realistic way. Currently, the emotions of the character agents are hardly mentioned in the story. If we expressed the emotions of the characters in the story, the reader would get more insight in the emotions and personalities of the characters and he or she would understand more about the agents' behaviour. The current research therefore strives to develop a way to express the character's emotions in the story.

1.1 Research goals

This research focuses on an extension of the Virtual Storyteller that can express the emotional states of the character agents. The emotional state of a character agent is represented by an intensity value for each emotion, which indicates to what extent the emotion is experienced by the agent. The main research question concentrates on how and when the emotional states should be converted into natural language in order to communicate the characters' emotions to the reader. It is formulated as follows:

In what way should the emotional states of the character agents be explicitly expressed in the story, in order to inform the reader about the characters' emotion intensities?

We aim to express the characters' emotions in a natural and accurate way in the story. To this end, we will study two linguistic fields that deal with the expression of emotions, namely Lexical Pragmatics and contrast relations. The field of Lexical Pragmatics is concerned with the effect of context on the interpretation of lexical items. Within this research field the pragmatics of adjectives is discussed. If we want to express emotions in an explicit and differentiated way, gradable adjectives are required in combination with modifiers. Complex expressions can then be constructed that refer to certain intensities on the emotional scale. An example of a complex expression is *not happy*, which can logically refer to many points on the ⟨happy, unhappy⟩ scale. Such an expression is called a conversational implicature because hearers generally have a preferred interpretation for it. This interpretation depends on the context of the utterance, in this case referring to alternative utterances that were not expressed by the speaker. Linguists have proposed how conversational implicatures like these are interpreted and how the interpretations are determined. We will discuss the literature and we will conduct empirical research for answering the following more specific research question:

a. *How do conversational implicatures of gradable adjectives correspond to the gradable spectrum of an emotion?*

Beside Lexical Pragmatics, we will study the field of contrast relations. When contrast relations are used for the linguistic expression of emotions, the reader will not only be informed about which emotions are experienced by an agent but also about how these emotions are related to other emotions or how they relate to an agent's behaviour or perception. Contrast relations are thereby very useful in giving the reader insight into the character's emotional state and the causes and consequences of this state, including the agent's personality. Markers of contrast, like *but*, are used to communicate that contrast holds between two sentence parts. Beside that, they indicate that the second conjunct is most important of the two (Spooren, 1989). When such markers are used, the contents of the first conjunct may be weakened or even be cancelled. We will discuss this hypothesis theoretically and propose how contrast relations can be used to express emotions. The following research question will be examined when studying contrast relations.

b. *How can contrast relations be used for expressing a character's emotional state and can contrast relations express emotion intensities?*

Having discussed the above two research questions and having gained insight in how linguistic expressions concerning emotions are interpreted by humans, we will propose how we can

expand the Virtual Storyteller with a module for the expression of emotions. The emotions should be expressed accurately; the intensities of the emotions in the emotion model should correspond to how the emotional expressions are interpreted by the reader. But the emotions should also be expressed in a natural way; the reader should not be distracted from the story by the emotional expressions and the storyteller should use expressions that are also used by human writers or speakers. These goals are expressed by the final specific research question:

c. *How can we expand the Virtual Storyteller so that the emotional states of the character agents are expressed in an accurate and natural way in the story?*

The three specific research questions a, b, and c will together answer the main research question, formulated above.

1.2 Approach

In the pursuit of the research goals described above, we will first study the architecture of the Virtual Storyteller and pay special attention to the emotion model and include a detailed analysis of the emotions in this model. Then we will investigate which lexical items can be used to refer to the emotions in the model by looking at how these emotions are expressed in existing fairy tales. This corpus study will also be used to verify whether the emotions that are included in the emotion model match the emotions that are expressed in existing fairy tales. Depending on the obtained frequencies of the expressed emotions in the corpus, we will add new emotions, or delete emotions from the emotion model of the Virtual Storyteller.

Once we have selected lexical items for the emotions in the model, we know how to refer to the two ends of an emotion scale, but not to the gradable spectrum in between. We will therefore construct complex expressions, using the lexical items in combination with modifiers, to refer to certain points on the scale. The field of Lexical Pragmatics will be explored to discuss the preferred interpretations of these complex expressions, called conversational implicatures. Because the theory about these conversational implicatures evoked by adjectives has not been tested empirically, we will conduct an experiment to check for the theoretical findings. From the experiment we will obtain what the preferred interpretations are for the complex expressions and how these fit into the theory of conversational implicatures.

Our next step is to conduct theoretical research on contrast relations. Contrast relations are capable of relating the expressed emotions to other aspects of the story. Using contrast relations, the experienced emotions are interweaved with the rest of the story. We will discuss what types of contrast relations exist and we will present how contrast relations can be used to express emotions in the Virtual Storyteller. We will also discuss whether contrast relations can be used to express emotion intensities.

Finally, we propose how our findings from the corpus study, the research on conversational implicatures and the research on contrast relations can be implemented in the Virtual Storyteller.

1.3 About this thesis

This Master's thesis presents the research that was conducted about the expression of emotions and how the Virtual Storyteller can be improved. Chapter two gives a description of the

Virtual Storyteller in its current state. Chapter three focuses on the emotion model of the storyteller and it explains the psychological OCC model that was used by Rensen (2004) to construct his emotion model. This chapter also presents the results of the corpus study that was conducted to study what emotional expressions human authors use. Chapter four discusses the field of Lexical Pragmatics and sheds a light on the interpretation of emotional expressions. Chapter five presents the results of an empirical study that tests several theoretical claims in the field of Lexical Pragmatics for emotion adjectives. Chapter six discusses how contrast relations can be used to express emotions. Chapter seven proposes how the Virtual Storyteller can be expanded so that emotions are expressed in the story. Chapter eight makes some final conclusions about incorporating emotional descriptions in story generation and proposes research goals for future work.

Chapter 2

The Virtual Storyteller

The Virtual Storyteller began as a research project of Sander Faas of the University of Twente (Theune et al., 2002). Since 2002, a team of students and their advisors have been working on the project and the Virtual Storyteller has evolved into a system with a well developed architecture. Although the elements of the architecture can still be improved and expanded, the architecture itself serves as a stable basis for the system.

The Virtual Storyteller can generate and present fairytale-like stories. The system has three subtasks: a plot has to be generated, the plot has to be converted into natural language (currently the Dutch language is used) and the story text should be converted into speech. So far, the focus of the project has been on the first two components of the storyteller. Some implementations of the text-to-speech components have been made, but these have not been integrated into the storyteller yet. As the current research strives to improve the Virtual Storyteller in its current state, the text-to-speech component is not studied in this chapter and it is left out in a description of the architecture below.

The plot generation task and the language generation task are two separate components in the Storyteller's architecture. The storyteller has a multi-agent framework and the system's tasks are executed by different agents. Three types of agents are responsible for plot generation and another agent is responsible for the language generation. The plot generation architecture contains the following agents:

- Character Agents - these are the actors in the plot and the characters of the story. They are associated with an emotion model that makes them believable characters;
- World Agent - keeps track of actions of the character agents and events of the plot agent and updates the world accordingly;
- Plot Agent - controls the plot by generating events in the world and constructs a formal description of the plot.

The plot is generated in a virtual story world, in which the Character Agents have to fulfil their goals and take actions that contribute to their goals. The World Agent keeps track of the story world by processing actions and events and checking their validity and effects on the world. The Plot Agent tries to manage the plot by influencing the world and controlling the characters. During the story, the Plot Agent constructs a formal description of the plot, called the fabula structure.

The language generation task is executed by the Narrator Agent, who receives the fabula structure from the Plot Agent and converts this structure into natural language. The Narrator's architecture is a pipeline of the following three components:

- Document Planner - converts the fabula structure into a global document plan. This is a tree structure in which all internal nodes have exactly one or two children;
- Microplanner - constructs a rhetorical dependency graph that contains dependency trees as plot elements;
- Surface Realizer - converts the dependency trees into text, performing syntactic aggregation, referring expression generation and surface form generation.

The narrator agent thus converts the fabula structure into natural language in several stages that have different tasks. In a future stage the output of the Narrator Agent will be sent to a Presentation Agent that performs a text-to-speech task.

The storyteller's architecture will now be described in more detail and the focus will be on the emotion model of the system and on how and where the emotion model is converted into natural language. During the description of the architecture, we refer to a translation of a story that was created by the Virtual Storyteller, called *Amalia and Brutus*. This story was presented in Hielkema (2005). A translation of the story is shown below. The original story in Dutch can be found in Appendix A.

Once upon a time there was a princess, called Amalia.
She lived in the small forest.
There was also a villain, called Brutus, who was in the bog.
Amalia went to the desert, because she wanted to pick cacti.
Brutus also went to the desert, to explore the environment.
She wanted to flee because she became scared when she saw Brutus.
Therefore she went to the barren field.
Brutus went to the barren field, because he wanted to imprison Amalia.
When Amalia saw Brutus, she became scared.
Brutus went to the mountains, because she went there to flee from Brutus.
There was a sword in the mountains.
She wanted to kill Brutus.
Because Brutus became scared, when he saw Amalia picking up the sword, he became aggressive.
He hit Amalia.
She lived happily ever after after stabbing Brutus.

We now present how the different components of the Virtual Storyteller produce a story like *Amalia and Brutus*. First we describe how the plot of the story is generated and after that, we focus on the language generation component of the Virtual Storyteller

2.1 Plot Generation

The architecture of the plot generation multi-agent system is displayed in figure 2.1.

Plot Agent The Plot Agent is the director of the plot and he tries to construct an exciting story. Most events that happen in the world are initialized by the Plot Agent. The rest of the events are consequences of other events or actions. For example, in *Brutus and Amalia*, the Plot Agent could have initialized the event that the sword was in the woods, and Amalia could have initialized the event of the sword being broken if she had dropped it in a canyon. Beside the initialization of events, the Plot Agent directs the plot by influencing the characters' behaviour indirectly in two ways. In the first place the Plot Agent can determine the characters' goals. A character's goal influences a character's behaviour in that the actions that are chosen by the character contribute to his goal. This will be explained in section 2.1.3. In the second place, all characters' perceptions are determined by the Plot Agent. Characters' perceptions have great influence on their behaviour, as the characters only choose actions that are possible in the world as they know it. For example, in *Amalia and Brutus*, Amalia would only have picked up the sword if she had seen the sword and believed that is was there. The Plot Agent sends all actions and events to the World Agent who maintains a world model and receives world updates from the World Agent. All information about the plot is stored in a fabula structure by the Plot Agent.

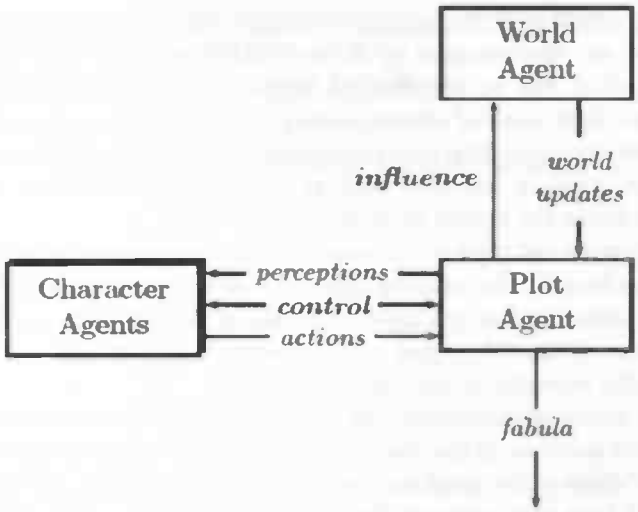


Figure 2.1: Architecture of plot generator taken from Swartjes (2006)

World Agent A representation of the world in which the story takes place is maintained by the World Agent. The Plot Agent makes sure the World Agent is informed about all actions and events. The World Agent processes these, checking their validity and their effects on the world. The world is then updated accordingly. The World Agent returns its world updates to the Plot Agent who keeps track of the plot. We can illustrate this with the story of Brutus and Amalia. Amalia went to the desert and Brutus went to the desert. This information was sent to the World Agent and the World Agent changed the world accordingly: Amalia and Brutus were both placed in the desert and were therefore at the same place.

Character Agents The Character Agents are semi-autonomous agents. They have their own internal states (emotions, beliefs and physical states), they have their goals and they can plan their own actions. This makes them believable characters. However, their behaviour is also partly controlled by the Plot Agent as described above. The Character Agents request world updates from the Plot Agent and they send their planned actions in combination with their internal states and their (sub)goals to the Plot Agent (via the control arrow in figure 2.1). In the story above, Amalia en Brutus are the only Character Agents. The story explicitly describes their emotions and actions. Their beliefs and physical states are not relevant enough to express in the story, because these directly reflect the events in the world. For example, because the sword was in the mountains, Amalia saw that the sword was in the mountains and this made her believe that the sword was in the mountains. When such perceptions and beliefs of characters are expressed, the story becomes very tedious.

We will now put some focus on the description of the emotion model of the character agents. The emotions in the emotion model are presented and we point out how the emotion model is connected to other aspects of the character agents.

2.1.1 Emotion model

The character agents have an advanced emotion model to represent their emotions. The emotion model was implemented by Rensen (2004) and based on the OCC model (Ortony et al., 1988), which will be described in section 3.2. The emotion model of the character agents contains eight pairs of corresponding positive and negative emotions. The emotion pairs with their corresponding interpretations in the storyteller are given in table 2.1. The five uppermost emotions represent feelings towards other character agents; the other three reflect feelings about the agents themselves.

An agent's emotional state is represented by an intensity value for each emotion pair in the agent's emotion model. The emotion pairs have an intensity value between -100 and +100. A positive value indicates that the agent experiences the positive emotion of the emotion pair, a negative value indicates the agent experiences the negative emotion. The numerical value is a measure for the intensity of the emotion.

Figure 2.2 shows schematically the input and output connections between the emotion model and other modules of the character agents. The intensity values of the emotion model depend on the beliefs of the agent and on the agent's personality. All beliefs that are possible in the story world have been assigned fixed emotion intensity values that determine the agents' emotional reactions to their beliefs. These assigned values are similar for all agents. For example, for every agent, the belief of some other character holding a sword and standing close results in an increase of the fear emotion. The personality was included in the input to the emotion model in order to realize that agents can react differently to the same event.

The events only indirectly influence an agent's emotion model; events are perceived by agents, which lead to agents' beliefs about the world, which can lead to changes in the agents' emotion model. Because of this, events can only change an agent's emotion if the agent perceives the event. Events that are not perceived do not influence the agent's emotion model. The *Amalia and Brutus* story does not describe the chain from event to emotion explicitly because this would make the story very boring. In fact the following happened in the plot: Brutus saw that Amalia was holding the sword which made him belief that Amalia was holding the sword which made him feel frightened.

The agents have eight fixed personality parameters, one for each emotion pair in the

Emotion pair	Interpretation
Happy for	Feeling happy for a beloved character
Pity	Having compassion for a beloved character
Gloating	Gloating over a hated character
Resentment	Feeling envious about a hated character
Admiration	Another agent executes an action that satisfies the own agent's norm
Reproach	Another agent executes an action that does not satisfy the own agent's norm
Love	Another agent or object is beloved
Hate	Another agent or object is not beloved
Hope (others)	Feeling hope for someone or something
Fear (others)	Feeling afraid of someone or something
Hope (self)	The agent's goal seems to succeed
Fear (self)	The agent's goal seems to fail
Pride	An executed action satisfies the agent's norm
Shame	An executed action does not satisfy the agent's norm
Joy	The agent's goal has been achieved
Distress	The agent's goal can hardly be achieved

Table 2.1: The emotions of the emotion model

Figure 2.2: Connections between the different modules of the Character Agent

character's emotion model. The parameters determine how the agent's personal emotions react to his beliefs. For stereotypical agents, the personality parameters are predefined. For example, all villain agents have similar parameters that represent an evil character. An agent's emotional reaction to his belief is represented by an update of the intensity values of the eight emotion pairs. The value of each emotion parameter is multiplied by the emotion value linked to the current belief to find the event's influence on the character's emotion intensity. This intensity number is then added to the current emotion value. The duration of the event's influence on the character's emotion equals the value of the emotion parameter of the agent. Thus, both the emotion duration and the emotion intensity depend on the emotion parameter of the agent. Consequently, the emotional reactions differ between agents. Figure 2.4 illustrates how an agent's personality and an agent's belief determine the agent's emotion. This figure is discussed in section 2.1.2.

2.1.2 Goal selection

The emotion model is used to determine the character's goal, which in its turn determines what actions the agent executes. Determination of a character's goal occurs via action tendencies, which are types of behaviour features. There are twelve action tendencies. Examples are behaving friendly towards another character and behaving aggressively. Each action tendency is assigned a value, based on an agent's emotional state and an agent's action tendency script which is part of the agent's personality. The translation of the emotion model of an agent into action tendencies of the agent can therefore differ between agents. For example, some agents may behave more friendly towards other agents when they get scared, other agents may behave more aggressively.

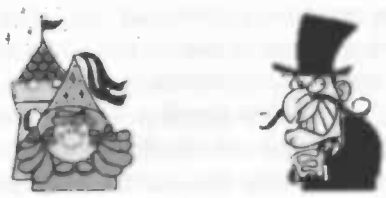
When all action tendencies have been assigned a value, the agent selects the goal that he wants to pursue. At first, goals are assigned *importance values*. There exist two types of goals: episodic goals that are chosen by the Plot Agent (or by the human author who initializes the system) and general goals, like attacking another agent, which depend on the agent's action tendencies. The action tendencies assign importance values to the general goals only; episodic goals do not have an importance value. The translation of the action tendencies into importance values for the general goals is similar for all agents. For example, the priority of attacking another agent is defined as the sum of the action tendencies behaving unfriendly towards the other agent and feeling superior with respect to the other agent.

out-of-control general goal > current goal (general or episodic) > episodic goal > 'in-control' general goal

Figure 2.3: priority list of goal types

When all general goals have been assigned an importance value (episodic goals do not have an importance value), the agent selects the goal that he will pursue. A character agent can only pursue one goal at a time and the agent's current goal determines his behaviour. Figure 2.3 shows the order of rank of an agent's goal selection. Goals that are *out of control* get the highest priority. These are general goals whose importance values exceed a certain threshold. The argumentation to give these types of goals the highest priority is that the emotions that form the basis for these goals are apparently too high for the agent to be in control. Therefore the corresponding goals must be pursued by the agent. If there are no out-of-control goals, the

agent will continue pursuing his current goal, which is either a general goal or an episodic goal. This prevents the character from switching goals very often, which could lead to unrealistic behaviour. When no goal is currently pursued (and there are no out-of-control goals) the agent will pursue an episodic goal, if this was instructed by the Plot Agent. At last, general goals that are 'in-control' are pursued. The agent selects one of the in-control goals semi-randomly. The chance that a goal is chosen is determined by its importance value.



Episodic goal	kill Brutus	imprison Amalia
Personality parameter	hope/fear = 5	hope/fear = 1
Action tendency script	passive = 0.8 * fear	passive = 0.3 * fear
Belief	villain Brutus is here	princess Amalia is here
with the fixed emotion value:	hope/fear = -20	hope/fear = 30
Results:		
Emotion	hope/fear (Brutus) = -100	hope/fear (Amalia) = 30
Action tendency	passive (Brutus) = -80	passive (Amalia) = 0
Goal	flee from Brutus	imprison Amalia

Figure 2.4: Fear emotion of Amalia and Brutus in the desert

Figure 2.4 illustrates how an agent's personality in combination with his belief influences his emotion and how this finally results in the selection of a goal. The figure shows the internal states of Amalia and Brutus when they are in the desert. For this figure, we focused on the hope/fear(others) emotion and the states that connect to this emotion.

When they enter the desert, Amalia and Brutus are both pursuing their episodic goals, 'killing Brutus' and 'imprisoning Amalia' respectively. For the hope/fear(others) emotion, their personality consists of a parameter for hope/fear and an action tendency script that describes how the 'passive' tendency depends on the hope/fear emotion. Amalia's hope/fear emotion parameter is higher, so her emotional reaction to an hope/fear event will be more intense. Furthermore, she will behave more passively than Brutus when fear is experienced. It is thus assumed that princesses behave in a more fearful way when they experience fear than villains do.

In the desert, Amalia and Brutus run into each other and this causes two beliefs: Amalia believes that the villain Brutus is there and Brutus believes that the princess Amalia is there. Amalia fears Brutus, but she also needs him to fulfil her goal. However, seeing Brutus increases her fear for Brutus, so her hope/fear emotion towards Brutus decreases. Brutus sees Amalia and this increases Brutus' hope for Amalia, since Brutus likes to capture princesses. Brutus' hope/fear intensity towards Amalia is therefore 30, and this does not change his goal. Brutus will therefore continue pursuing the episodic goal. Amalia, on the other hand, does experience fear of Brutus. This fear results in the action tendency 'passive' with an intensity of 80. The

importance value of the general goal of fleeing from Brutus then exceeds a certain threshold, which makes it an out-of-control general goal. Amalia thus pursues this general goal of fleeing from Brutus.

2.1.3 Action selection

Once a goal has been selected, the appropriate actions need to be chosen to fulfil the goal. The Virtual Storyteller contains an action ontology that specifies for each action what preconditions it requires and what effects it has on the world and the agents (Uijlings, 2006). The action ontology enables the character agents to construct an action plan that fulfills the agent's goal. The agent's world model that specifies where the objects and agents are located in the world tells the agent what actions are possible in the world.

The character agent uses the partial order planning (POP) algorithm to construct his action plan. The planning algorithm starts with the agent's initial state of the environment and the agent's goal state and it has a list of all available actions with their preconditions and effects on the world. From this situation, the algorithm searches for steps (actions) whose effects correspond to the preconditions of other actions and it thereby searches for an order of steps to connect the initial state to the goal state. The plan is complete if there are no more unsatisfied preconditions.

The actions from the action plan are executed by the agent. Each time an agent wants to execute an action, he sends his request to the Plot Agent who forwards it to the World Agent. The World Agent updates the world with the executed action. After each action, the agent constructs his action plan again in order to adapt it to new events in the world or to executed actions that did not succeed.

2.2 Narrator Agent

The Narrator Agent is responsible for converting the fabula structure, constructed by the Plot Agent, into natural language. It consists of a pipeline with three components: the Document Planner, the Microplanner and the Surface Realizer. This section roughly describes how each of the components contributes to the conversion of the fabula structure. The Narrator Agent was mainly developed by Slabbers (2006) who focused on the Document Planner (section 1.2.1) and the Microplanner (section 1.2.2) and Hielkema (2005) who worked on the Surface Realizer (section 1.2.3). For a description of the Narrator Agent the work of Slabbers was used.

2.2.1 Document Planner

The Document Planner receives the fabula structure from the Plot Agent. The fabula structure is a causal network that contains the most important plot elements from the whole story. The plot elements are represented by nodes and causal relations between the plot elements are represented by edges. Examples of plot elements are goals, actions and internal elements (emotions, physical states or beliefs). All plot elements have a time argument to specify when the element took place. The causal relations that hold between the elements can be e.g. psychological (an agent's perception leading to an agent's action) or physical (an event causing a perception). The Document Planner converts this fabula structure into an initial document

plan, which is a tree structure. After the initial document plan has been constructed, some operations are executed on the plan. The following operations are most relevant:

Background Information Supplier One of the tasks of the Document Planner is to specify what information should be present in the story text, because all information in the document plan will be converted into natural language. The Document Planner removes unimportant plot elements from the fabula structure and adds background information that was not present in the fabula structure. Perceptions and beliefs that an action succeeded are deleted as these make the story tedious. For example, when Amalia had killed Brutus, she believed that Brutus was dead. This belief was not expressed in the story. Other information that is not added to the document plan is an agent's supergoal if he has a more specific subgoal. For example, in *Amalia and Brutus*, Amalia's supergoal is to leave the desert and her subgoal is to go to the barren field. Only the subgoal will then be added to the document plan and finally be expressed in the story. Background information that can be added to the fabula structure concerns information from the story world and the character information modules. Character information can be a character's name or property and story world information can refer to locations or objects. Both types of background information can be used for generating referring expressions (see subsection 2.2.3).

State transformer Internal states that are connected to an agent's action are combined with the action. Internal states can be an agent's emotion, physical state (e.g. hungry) and belief. The internal state can be expressed directly by an adjective, e.g. *bang* (*scared*) and indirectly by a phrase, e.g. *met bonkend hart* (*with a pounding heart*) and the internal state can be combined with the action into one sentence or they can be expressed by two sentences. The choice of expression is currently made at random. The tree structure is changed according to the choice of combination between the action and internal state. If the internal state and the action are expressed by two sentences, the structure is left unchanged. If they are expressed by one combined sentence, the action plot element is expanded with a node that contains the internal state.

2.2.2 Microplanner

The Microplanner converts all plot elements in the document plan (the nodes) into rhetorical dependency trees. For this operation sentence templates and a lexicon are used by the sentence plan generator and the lexicalizer respectively.

Sentence plan generator To construct a dependency graph the sentence plan generator uses different kinds of templates for a plot element, mainly an action and event template, an internal state template and a perception and belief template. The templates contain fixed numbers of arguments that have to be filled in. Examples of arguments are a subject, an object and a modifier. The templates are used to determine how the arguments should appear in the dependency tree.

Of special interest for the current research is how the internal states, which can reflect emotions, are currently expressed. The template for an internal state is shown in 2.5. There are two types of internal states. If the object of the internal state is an adjective, the verb *to be* is assigned to the hd-node and if the object is a noun, the verb *to have* is assigned to the hd-node. With this template an internal state can be expressed explicitly, like in *Amalia is*

bang (*Amalia is afraid*) or *Brutus is ziek* (*Brutus is ill*). If an internal state is experienced intensely there are two other expressions that can be made to inform the reader about a state's high intensity: *Wat was ze blij* (*How happy she was!*) and *Ze was nog nooit zo blij geweest!* (*She has never been so happy before!*). The templates for these expressions do not differ much from the standard expressions for internal states.

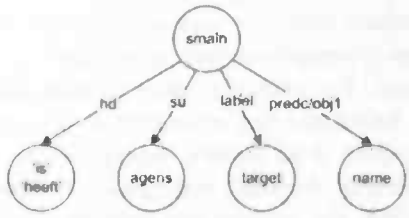


Figure 2.5: Template for internal state taken from Slabbers (2006)

Lexicaliser The lexicalizer maps concepts to Dutch words using a lexicon. The lexicalizer prevents word repetition, because the story would get monotonous if the same words were used very frequently. Concepts are looked up in the lexicon which returns the word for a concept and the word's corresponding properties, like the pos-tag, the determiner (for nouns), the gender (for nouns) and in case of a verb, some specific verb properties. The lexicalizer stores all used words in a list in order to prevent that the same word is used repeatedly. If a word was used recently the lexicalizer searches for synonyms that can be used.

2.2.3 Surface realiser

The surface realiser was designed and implemented by Hielkema (2005) and some improvements were made by Slabbers (2006). It is the final component of the narrator agent and it converts the rhetorical dependency graph into text, performing syntactic aggregation, referring expression generation and surface form generation.

Syntactic aggregator The rhetorical relations between dependency trees that are specified in the document plan are translated into cue words. Two related trees are then combined into one tree. The order of the trees is not changed here, which means that the plot elements are always told in chronological order. When two dependency tree are combined into one, the syntactic aggregator checks if the trees contain similar entities and syntactic aggregation is performed on these entities. In the story of Amalia and Brutus, syntactic aggregation was conducted on the second sentence among others: *The princess was called Amalia and lived in a small forest.* The cue word *and* was added to combine the two dependency trees and *the princess* was deleted from the second tree.

Referring expression generator Referring expressions are added to the rhetorical dependency graph. Three kinds of referring expressions can be used: a noun phrase (*the beautiful*

princess), the character's name, and a pronoun. Which expression is used depends on the fabula structure, the discourse history, the character information and the Story World. A detailed description can be found in Slabbers (2006). In Amalia and Brutus, we find many pronouns functioning as referring expressions.

Surface form generator The dependency graph of the story that contains combined trees and referring expressions is converted into text. A grammar is used to put the words in the dependency trees in the right order. The inflection of adjectives and orthography are applied and relative clauses are generated.

Chapter 3

Expressing the emotions

The first chapter has described the architecture of the Virtual Storyteller and where and how the emotion model of the character agents fits into this. The emotion model contains eight emotion antonym pairs that describe an agent's feelings towards another agent or towards himself. As the current research strives for a natural translation of the emotions in the storyteller, we need to study the emotions in the Virtual Storyteller more in depth. The emotion model of the storyteller was derived from the OCC model (Ortony et al., 1988) that is described in this chapter. We study what feelings the emotions in the OCC model represent and how well the antonym emotions fit together.

The structure of the OCC model was mainly based on psychological research but Rensen (2004) never tested whether the emotions in the OCC model are indeed experienced in stories. Rensen selected the eight most general emotion pairs from the OCC model for the Virtual Storyteller without clearly substantiating his choice for these eight emotions. In order to produce natural stories, the emotion model of the Virtual Storyteller should contain the emotions that are common in stories. To this end, we study existing fairy tales by conducting a corpus study and we discuss which emotions are expressed in the corpus and how they are expressed. One part of the corpus study is conducted automatically, the other part is conducted manually. For the manual study, we work through parts of the corpus by hand and search for expressions that refer to emotions. The results of the study show which emotions are common in fairy tales and what expressions are used to refer to these emotions.

Based on the corpus study, we give advice about which emotions should be included in the emotion model of the Virtual Storyteller and how the emotions should be expressed. Besides, the corpus study serves as an empirical evaluation of the OCC model. From the corpus, we obtain linguistic evidence for the emotions in the OCC model by studying the match between the emotions in the OCC model and the emotions expressed in the fairy tales. The corpus study shows if the emotions in the OCC model are linguistically grounded.

3.1 Models of emotion

The use of an emotion model in the Virtual Storyteller can contribute to the believability of the story by making the virtual characters more realistic. Constructing an exact model of human emotions in the Virtual Storyteller is impracticable, because of two main reasons, given by Magnenat-Thalmann et al. (2005). At first, psychological models of emotion and personality are not yet fully developed. How emotions and personality are related to perception, behaviour

and expression has yet to be determined. And, secondly, when modeling human emotions in a system like the Virtual Storyteller one is dependent on the computational properties of the system. The most suitable model for the system may not be the best representation for emotions.

In literature, the general view is that emotions are evoked from the perception of events, but there are different theories about how emotions result from events. The main question is whether physiological responses to perceptions of events play a role in the rise of emotions. For example, the Schachter-Singer theory of emotions (Schachter & Singer, 1962) claims that an event causes physiological responses and the emotion follows from the identification of these responses. Here, the perception of the physiological changes is crucial in the rise of the emotion. Another theory, the Lazarus theory of cognitive emotions (Lazarus, 1991), says that the thought following an event invokes both physiological changes and emotions and that the two are independent of each other.

Rensen (2004) decided to use the OCC model for the characters in the Virtual Storyteller. The OCC model is a cognitive appraisal model that describes how events, agents and objects elicit emotional responses. This model corresponds closely to Lazarus' theory of emotions, except that it leaves out the physiological responses to the events. In the Virtual Storyteller, physiological changes are not relevant to the story and their presence in the emotion model would serve no purpose. Leaving them out in the emotion model is therefore not a problem. However, we want to underline that the OCC model excludes a developed model of the physiological changes associated with the emotions and that it is thereby not a complete model of emotions.

3.2 The OCC model

The OCC model (Ortony et al., 1988) is a cognitive emotion model, based on psychological research, that assumes that emotions are a consequence of positive or negative reactions to events, actions and objects and it is therefore a so-called cognitive appraisal model. The emotions in the emotion model are categorized according to the eliciting conditions of emotions. All emotions are initiated by the thoughts that the agent has about events, actions of agents and aspects of objects and the model is therefore a representation of an agent's perception of the world. The actual events in the world only indirectly (via the agent's perception) influence the emotions of the agent.

Figure 3.1 shows the OCC model. Each emotion is represented by of a pair of emotions; one referring to the positive feeling, the other referring to the negative feeling. The emotions are grouped by the three eliciting conditions. The emotions (in italics) in the figure are merely labels for the type of emotion being considered. From top to bottom, the emotions in the figure become more specific.

An agent's emotional reaction to an event always stems from an agent's emotion of being pleased or unpleased about the consequences of that event. When an agent emotionally reacts to an action, the agent approves or disapproves his own action or the action of someone else. An agent liking or disliking something is initiated by an agent's reaction about aspects of an object or agent. Within each group of emotions, different aspects of the emotional feeling specify the emotion type.

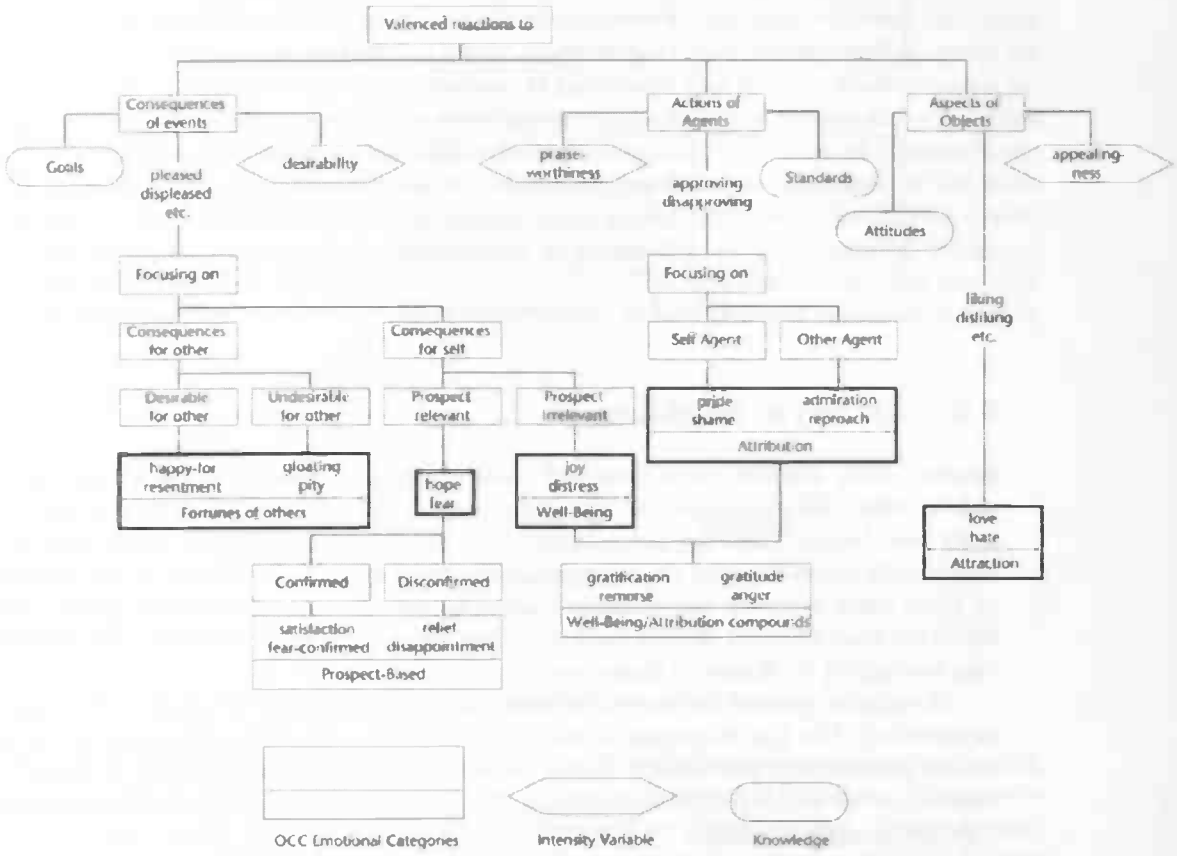


Figure 3.1: The OCC model, taken from Bartneck (2002)

Emotion label	Type specification
Happy for	Pleased about an event desirable for someone else
Pity	Displeased about an event undesirable for someone else
Gloating	Pleased about an event undesirable for someone else
Resentment	Displeased about an event desirable for someone else
Hope (self)	Pleased about the prospects of a desirable event
Fear (self)	Displeased about the prospects of an undesirable event
Hope (others)	Experience hope for someone or something
Fear (others)	Experience fear for someone or something
Joy	Pleased about a desirable outcome
Distress	Displeased about an undesirable outcome
Pride	Approving of ones own praiseworthy action
Shame	Disapproving of ones own blameworthy action
Admire	Approving of someone else's praiseworthy action
Reproach	Disapproving of someone else's blameworthy action
Love	Liking an appealing object
Hate	Disliking an unappealing object

Table 3.1: The emotion pairs of the storyteller

3.3 Antonym emotions

Rensen (2004) adopted seven emotions pairs of the OCC model for the Virtual Storyteller and he added one extra emotion. The emotions that Rensen used are marked in figure 3.1 by using bold boxes. These are the emotions that will be discussed below, as our aim is to find appropriate expressions for the emotions in the storyteller. A description of the emotions can be given when following the branches leading to the emotions in the OCC model. Table 3.2 shows the emotions that Rensen used with their corresponding descriptions. The emotion pair that was added by Rensen is hope/fear(others).

All emotion pairs of the storyteller have two dimensions that determine which emotion is experienced. The first dimension is the event, action or object that the agent reacts to. The second dimension is the agent's feeling (reaction) about this event, action or object. Both dimensions are either positive or negative. For most emotions, these two dimensions result in only two possible emotions:

- a positive event, action or object bringing about a positive feeling

- a negative event, action or object bringing about a negative feeling

Thus, for most emotions, the value of the emotion corresponds to the value of the eliciting event, action or object. For example, an agent experiences joy if and only if the event, he reacts to, is positive. If the event is negative, the agent experiences distress. The fortunes-of-others emotion is an exception to this generalization. With this emotion, an agent can react positively to an event undesirable for another agent and he can react negatively to an event desirable for another agent. These reactions are possible when an agent dislikes the other agent. So, for the fortunes-of-other emotion, the two dimensions result in four different emotion types. We will discuss these four types in more detail in section 3.3.1.

The emotion pairs in table 3.2 are treated as antonym emotions[1] in the Virtual Storyteller. For each emotion pair, the character agents experience either the positive emotion or the negative emotion to a certain degree. There is a gradual increase between the two antonym emotions and it is therefore important that the two emotions of an emotion pair are really opposite emotions. The emotions are now discussed in more detail. A description of each emotion type are given and we discuss how well the antonym emotions match together.

3.3.1 Fortunes-of-others emotions: happy-for, pity, gloating, resentment

These emotions arise when events have desirable or undesirable consequences for another agent and the agent himself is pleased or displeased about this. The implications of an event for another person's goal determine the desirability of the event. This requires that the agent experiencing the emotion knows another agent's plans and goals or that the agent knows which emotion the other agent experiences with respect to the event (joy/hope or distress/fear). These emotions are influenced by both the presumed desirability of the event for the other agent and by the desirability that the other agent experiences such an outcome. The type specifications for the fortunes-of-others emotions can be generalized as: Being pleased/displeased about an event desirable/undesirable for someone else. The specification contains two variables of which each can have two values. This results in four different emotions, shown in figure 3.2

Reaction \ Event	+	-
+	Happy-for	Gloating
-	Resentment	Pity

Figure 3.2: The fortunes-of-others emotions

Ortony et al. categorize the four emotions two by two, distinguishing emotions reflecting events desirable for others from emotions reflecting events undesirable for others. This results in two antonym pairs, each with a good-will emotion and an ill-will emotion. Happy-for and resentment are then antonyms and gloating and pity are antonyms. Rensen chose to deviate from the OCC model and he reorganized the antonym pairs, resulting in two other emotion pairs, one pair of two good-will emotions and one pair of two ill-will emotions. He thus claims that happy-for and pity are antonyms and that gloating and resentment are antonyms. This means that the attitude towards another person determines the emotion type and the desirability of the consequence for the other agent determines the intensity of the emotion.

Both distinctions are suitable. In fact, all four emotions exclude each other and we can form different antonym pairs. However, for the storyteller we need to construct an intensity scale for an emotion pair. So, the two emotions of an emotion pair should share a transition. We find emotions with opposite feelings more appropriate than emotions with opposite events, because the polarity of the emotion should change according to how the agent feels, and not according

[1] We use the term antonym to refer to the two opposite extremes on the same scale.

to what events take place. We thus agree with the OCC model and we will restructure the emotion model of the storyteller by forming emotions pairs happy-for/resentment and gloating/pity.

3.3.2 Prospect-based emotions: hope, fear

An event can also have a certain prospect as a consequence. That is, when an event occurs it may increase the likelihood of some other event (desirable or undesirable). If the future event is desirable the agent is pleased (hope), if the future event is undesirable the agent is displeased (fear). Figure 3.1 indicates that for this emotion the prospects of the event are relevant. Referring to the *Amalia and Brutus* story, used in chapter 1, an example of an event that initiated a prospect-based emotion for Brutus was Amalia holding the sword. The prospect of this event was the undesirable event of Amalia killing Brutus. The intensity of the emotion is influenced by both the likelihood of the prospect and the desirability of the prospect.

3.3.3 Well-being emotions: joy, distress

An agent can also emotionally react to the consequences of an event itself, if the consequences are desirable or undesirable for the agent. The prospects of the event are then irrelevant. The degree to which the consequence is desirable determines the intensity of the well-being emotion. The intensity can thus be positive (joy) and negative (distress) and the corresponding emotions constitute an antonym pair.

3.3.4 Attribution emotions: pride, shame, admiration, reproach

These emotions result from actions of an agent himself and from actions of other agents. When an agent judges something he himself has done as praiseworthy or blameworthy, the emotions pride and shame arise. And when an agent approves or disapproves of some other agent's action, the emotions admiration and reproach arise. Important here is that the emotion is a consequence of the agent's belief about the agent's own action, and it is not caused by the actual action. The praiseworthiness/blameworthiness of the action determines the intensity of the attribution emotion.

3.3.5 Attraction emotions: love, hate

The attraction emotions are reactions of dispositional liking or disliking another object or agent. The degree to which another object or agent is appealing affects the intensity of this emotion. Since the attraction emotion mostly comes about when certain aspects of an object are liked or disliked, a dislike of a certain category is a predisposition to dislike all instances of that category.

3.4 The language of emotions

Now that we have described the OCC model and the emotions that have been implemented in the emotion model of the Virtual Storyteller, we focus on how the emotions can be expressed. Introducing the OCC model, Ortony et al. suggested some lexical expressions for the emotions.

These expressions are shown in 3.2. The expressions were chosen theoretically and have not been established empirically.

Emotion label	Expressions
Happy for	delighted for, happy for, pleased for
Resentment	envy, jealousy, resentment
Gloating	gloating, Schadenfreude
Pity for	compassion, pity, sad-for, sorry-for, sympathy
Hope (self)	anticipation, excitement, expectancy, hope, hopeful, looking forward to
Fear (self)	apprehensive, anxious, cowering, dread, fear, fright, nervous, petrified scared, terrified, timid, worried
Hope (others)	not given
Fear (others)	not given
Joy	contented, cheerful, delighted, glad, happy, pleased, ecstatic, elated euphoric, feeling good, joyful, jubilant, pleasantly surprised
Distress	depressed, distressed, displeased, dissatisfied, distraught, feeling uncomfortable grief, homesick, lonely, lovesick, miserable, regret, shock, uneasy, unhappy, upset
Pride	pride
Shame	embarrassment, feeling guilty, mortified, self-blame, self-condemnation self-reproach, shame, uncomfortable, uneasy, etc.
Admiration	admiration, appreciation, awe, esteem, respect
Reproach	appalled, contempt, despise, disdain, indignation, reproach
Love	adore, affection, attracted-to, like, love
Hate	aversion, detest, disgust, dislike, hate, loathe, repelled-by, revulsion

Table 3.2: The emotions of the emotion model

For nearly all expressions in the table Dutch equivalents exist. The expressions that are suggested seem quite appropriate for the corresponding emotion types. The different tokens for an emotion pair differ in intensity (e.g. *adore* versus *attracted-to*), in specificity (e.g. *unhappy* expresses a general negative feeling whereas *homesick* refers to a specific negative feeling) and in which variable was responsible for the emotion (e.g. the token *hope* expresses an agent's desire for something and the token *hopeful* expresses an agent's expectation of something). Most of the expressions suggested by Ortony et al. are adjectives or nouns that can easily be converted to adjectives. The rest of the expressions are verbs that mostly represent feeling towards other agents or objects. In the current research, the task is to find the most appropriate expression regarding the emotions in the Virtual Storyteller.

3.4.1 The use of adjectives

Ortony et al. (1988) identified many adjectives to express the emotions in the OCC model. Adjectives seem appropriate lexical items to express emotions in a simple and straightforward way. Using adjectives, we could for example say: *y is nervous*, or *y is happy for x*. Much literature about the use of adjectives has been written (see Raskin & Nirenburg (1995) for a summary) and the conventional view is that adjectives modify nouns by denoting some property of the noun. The modifying function of an adjective is a formal syntactic property of this category. Many linguists believe that the lexical items in a syntactic category must also

share some semantic property (e.g. Lyons (1977) and Wierszbicka (1988)). Wierzbicka claimed that in general, adjectives can be semantically distinguished from nouns on two aspects. Nouns designate kinds of things whereas adjectives designate properties. And, nouns refer to a whole number of properties where adjectives refer to only a single property. This difference between nouns and adjectives illustrates the distinguishing property of adjectives: they designate some property.

Beside this semantic property, there is another semantic aspect related to the syntactic category of adjectives, which is inherent to the difference between adjectives, nouns and verbs. This aspect, time-stability, proposed by Givón (1984), refers to the dynamicity (the time-length) of the meaning. Verbs are regarded as very time-instable and nouns are regarded as very time-stable. The adjectives category falls in between these two ends of the time-stability scale. Note that this distinction is a very generalized one; some verbs are rather time-stable (e.g. *reign*) and some nouns are rather time-instable (e.g. *attack*).

Comparing the class of emotions to the semantic properties of adjectives, adjectives appear to be the most appropriate syntactic category for expressing the emotions. Emotions are interpreted as the properties of the characters in the story, each emotion referring to a single property of the character, and their time-stability is somewhere between that of nouns and verbs; emotions are momentary reactions that last longer than the events they react to, but they change over time for a certain character. The suggestion that adjectives are most appropriate to express emotions is confirmed by different taxonomies of adjectives. Dixon (1982) listed seven semantic categories that are usually (if the language provides them) expressed by adjectives. One of these is the category 'Human Propensity', referring to adjectives like *happy* and *hungry*. Frawley (1992) structured this Human Propensity category in three subcategories, namely mental states, physical states and behaviour. Emotions are categorized as mental states and they therefore belong to the taxonomy of adjectives. Thus, if we want to express emotions in an explicit and natural way, adjectives would be most suitable to use. This is illustrated by the fact that we can construct expressions like *I feel X*, when using an adjective for *X*. Because emotions refer to feelings, it is desirable that we can construct such forms when emotions are expressed.

The emotions of the character agents should be expressed in a differentiated way, so it is essential that we can put nuances in expressing the emotions. The emotional expressions should refer to certain intensities on the emotional scale. One group of adjectives, gradable adjectives, is very capable of referring to intensities on a scale. Gradable adjectives express qualities that can be measured on a scale and they can be modified by degree adverbs, such as *very* or *a bit* to put nuances in their expression. We will therefore most plausibly be able to represent emotions via gradable adjectives.

3.4.2 Linguistic evidence for the OCC model

The structure of the OCC model is based on psychological research and completely independent of the lexicon of emotions. The model was mainly supported by experimental studies in which subjects reported which emotions they would expect to arise under certain conditions. Ortony et al. believe that a theory of emotions should not be a theory about emotion words and they give two reasons for that. One is that the emotion lexicon simply has a different form than the structure of the emotions. The other is that a study about emotions should focus on the things to which emotion words refer and not on the emotion words themselves. Accordingly, they did not pay attention to the relation between the emotion types and the

emotion lexicon.

However, some scientists object the view of Ortony et al. and claim that the structure of words is isomorphic to the structure of the concepts they express (Fodor & Lepore, 1999). From studying the lexicon we can learn more about the structure of the concepts that the words refer to. Whether or not the lexicon is isomorphic to the structure of concepts, we believe that a study of the language of emotions (or other concepts) will give us insight in the structure of these. Although the words in the lexicon only refer to the concepts and are not the concepts themselves, they represent the conversational access that people have to the concepts.

Accordingly, we think that a study of the emotion lexicon can evaluate the OCC model. The accessibility to an explicit expression that refers to an emotion tells us how common the associated emotion is. As written above, adjectives are the most natural lexical items to express emotions. If an adjective is available to express an emotion, we can conclude that this emotion should be admitted to the OCC model. And, on the other hand, if no adjectives are available to express a certain emotion in the model, it should be discussed whether the associating emotion is correctly included in the model. The frequency of an emotional expression in a corpus tells us how often the emotion is experienced. Studying these frequencies we can recommend which emotions of the OCC model should be included in the emotion model of the Virtual Storyteller, assuming that frequently experienced emotions should be part of the stories that are produced by the storyteller in order to make the stories more natural.

Because we aim to evaluate the OCC model and the emotion model of the Virtual Storyteller we conduct a corpus study on emotional expressions. The corpus study surveys the match between the OCC model and the emotion lexicon and the linguistic grounds of the OCC model will be discussed. The corpus study also serves a more practical goal. We need to construct an emotion lexicon for the emotions in the Virtual Storyteller. We could choose to select the most appropriate expression from table 3.2 for each emotion type in the Virtual Storyteller. However, this method has some shortcomings. In the first place, we prefer to use an adjective for each emotion in the storyteller and the above table does not show an adjective for each emotion. We would then end up using different syntactic categories for different emotions and this would make the Narrator Agent more complex. Secondly, the above lexical items have only been obtained theoretically and we do not know how common the forms are. For some emotions in the table Ortony et al. suggest many lexical items and it is not indicated which ones are most frequently used in natural language. For other emotions, very few items are suggested and it would be useful to investigate what other expressions could be used for refer to these emotions. A corpus study will show how emotions are expressed in natural language and it will be very useful in constructing the emotion lexicon in the Virtual Storyteller.

3.5 Corpus study

We conducted a corpus study in the same domain as the stories produced by the Virtual Storyteller, the fairy-tale domain. The fairy-tale domain was studied, because the stories produced by the Virtual Storyteller should approach fairy tales written by human authors. A linguistic study of the fairy tales shows which emotions are common in fairy tales and should therefore be included in the emotion model of the Virtual Storyteller and it reveals

which lexical items are most suitable for the emotions in the stories produced by the Virtual Storyteller. Next to these two practical goals, the corpus study serves a theoretical goal; it determines whether the emotions in the OCC model are linguistically grounded.

Since adjectives are the most appropriate lexical items to express the emotions in a natural way, the study searches for adjectival emotional expressions in the corpus.

3.5.1 Method

The corpus was obtained from fairy tales written by H.C. Andersen (first published between 1835 and 1872 and translated by dr. Annelien van Hees, 1997) and the Grimm brothers (first published between 1812 and 1857 and translated by M.M. De Vries-Vogel, 1984), all in Dutch. The stories were available online in machine readable format.[2] From each author, forty fairy tales were selected randomly, which resulted in a corpus of eighty fairy tales. A list of the fairy tales that were in the corpus can be found in Appendix B.

The corpus was parsed by the Alpino dependency parser for Dutch.[3] All adjectives from the corpus were selected automatically. A random check of eighteen hundred words showed that the performance of the Alpino parser was as follows: 99.6 percent of the words were correctly classified as adjectives or non-adjectives, 0.2 percent of the words were adjectives that were missed by the parser and 0.2 percent of the words were incorrectly classified as adjectives.

Of all obtained adjectives, the adjectives that could represent emotions were selected manually. In order to discriminate between emotional expressions and general behavioural states or aspects of a character's personality, Ortony et al.'s definition for emotions was used: "Emotions are valenced reactions to events, agents, or objects, with their particular nature being determined by the way in which the eliciting situation is construed." Relevant aspects of the definition that we used to select emotional adjectives are that the adjectives should express reactions to something and they should be valenced (showing polarity). If necessary, the adjectives were looked up in a dictionary in order to determine whether they expressed emotions.

When all emotional adjectives were obtained, they were matched with the type descriptions in table 3.2 referring to the emotions of the storyteller. If the adjectives were ambiguous with respect to the emotion types, their usage was looked up in the corpus in order to disambiguate.

3.5.2 Results

13,400 adjective tokens were found, consisting of 2,035 adjective types. All 2,035 adjectives were worked through by hand to identify the adjectives that referred to emotions. Table 3.3 shows the emotional expressions that were obtained with their corresponding frequency in the corpus categorized according to the different emotion types. Adjectives that could not be categorized are shown at the bottom of the table.

The adjectives that we obtained from the corpus, are all gradable adjectives and therefore able to express emotions in a differentiated way. As seen in table 3.3, the emotional expressions were not completely compatible with the emotion types introduced by Ortony et al. Some

[2]Stichting Beleven has constructed an online database of 1092 folk stories, including fairy tales from Andersen and Grimm. These can be found at http://www.beleven.org/verhalen.

[3]The Alpino parser is presented in Bouma et al. (2001) and can be found at http://www.let.rug.nl/~vannoord/alp/Alpino

Emotion label	**Expressions**
Happy for	...
Resentment	jaloers op(5)
Pity for	medelijdend met(1)
Gloating	...
Hope (self)	...
Fear (self)	bang(51), angstig(6), bezorgd(1), bevreesd(1)
Hope (others)	...
Fear (others)	bang voor(11), angstig voor(1)
Joy	blij(60), gelukkig(55), vrolijk(35), verheugd(12), opgewekt(3)
Distress	bedroefd(42), verdrietig(8), treurig(7), droevig(6), weemoedig(2)
Pride	trots(26)
Shame	...
Admiration	...
Reproach	...
Love	verliefd op(3)
Hate	..
Other	
Anger	boos(83), kwaad(12), woedend(10), woest(4), driftig(2)
Gratitude	dankbaar(4)
Satisfaction	tevreden(18)
Disappointment	teleurgesteld(1)

Table 3.3: Obtained adjectives from the fairy-tale corpus

expressions could not be matched with an emotion type and some emotion types were not expressed by any adjective. We now first discuss the adjectives that we could match to some emotion.

Matching adjectives

For all adjectives that matched to emotions, only the adjectives that were general references to the emotions were selected. In order to vary in expressing an emotion, we selected more than one adjective if possible. From the obtained adjectives for an emotion, we selected all adjectives that were used in more than ten percent of the expressions. Below, we describe which items were selected.

Pity for Only one adjective with a very low frequency was found for the pity for emotion. We concluded that it is either not common to use an adjective for expressing this emotion or fairy-tale characters are seldom described as feeling happy/pity for some other object. We therefore preferred using a lexical item from a different syntactic category. When looking for Dutch equivalences for the expressions suggested in table 3.2, we found the verbal expression *medelijden hebben met (to be sorry for)* from which the adjective *medelijdend* could be derived to be the most common expression in the corpus with a frequency of seven. This expression will be used.

Resentment For the resentment emotion, the adjective *jaloers (op)* was obtained from the corpus with a rather low frequency. We also looked for other syntactic categories in the corpus to represent resentment, but none were found. We will therefore use the adjective *jaloers*.

Fear (self) We found four different expressions for the fear (self) emotion. The items differ much in their frequency in the corpus and the adjective *bang* is clearly the most common adjective to express fear. Using our selection criterion of frequency, we selected the adjectives *bang* and *angstig*.

Fear (others) The fear (others) emotion is a peculiar emotion that was added by Rensen. Humans can experience the emotion fear in general, but they can also have fear towards some other object or person. If one has fear towards another object, this does not imply that one does indeed experience fear at that very moment. For example, some people are afraid of spiders, but this does not influence their mood if there are no spiders nearby. Fear to others should therefore be interpreted as an attitude towards some other object or person and it only results in an emotional feeling of fear if the other object is nearby. Since we aim to express the emotions that are experienced by an agent, the fear (others) emotion should only be expressed if the feared object or the feared agent is nearby. For the fear (others) emotion, we found the adjective *bang* combined with the preposition *voor* (for). Its frequency in the corpus was nine; five of the occurrences represented the emotion fear, four of the occurrences represented the attitude fear. *Bang voor* will thus be used for expressing fear towards some other person or object.

Joy Five different adjectives were found whose frequencies varied only little for the three most frequent adjectives. The expressions *blij*, *gelukkig* and *vrolijk* were obtained in more than

ten percent of the cases. The adjective *blij* is most appropriate for expressing this emotion, as it represents a rather momentary feeling, mostly a reaction to a positive event. The adjective *gelukkig* expresses joy on a higher level; being in a fortunate situation. *Gelukkig* therefore often refers to a longer lasting situation and it is not as much directly changed by events. It will therefore not be used for expressing the joy emotion. The other frequently obtained adjective *vrolijk* mostly refers to someone's mood and it expresses that someone is in a good mood. This adjective thereby expresses a general feeling of happiness, which is not directly initiated by a positive event. We therefore chose to use *blij* only to express the joy emotion.

Distress We found five different expressions for the distress emotion. The expressions are synonyms, except for *weemoedig*, which is less intense than the other expressions. The other expressions differ much in their frequency in the corpus and *bedroefd* is obviously the most common adjective to express distress. Taking into account their frequency, we selected *bedroefd*, *verdrietig* and *treurig* for expressing distress. If we wanted to express a rather low intensity of distress, we could use the adjective *weemoedig*. We will come back to this in chapter 7.

Pride For pride we only found one adjective, *trots*. Its frequency is quite high so we selected the adjective *trots*.

Love This emotion type represents a complicated set of emotions, as there is a vague separation between liking and loving someone. An agent can like another agent, it can love another agent, and it can be in love with another agent. These three fulfillments of loving do not only differ in intensity, but also in the way an agent loves another agent. This makes it hard to define the differences between them. It might therefore be necessary to only express one of the fulfillments. The obtained expressions from the corpus should support our choice.

For expressing the love emotion, we found one adjective with a low frequency, *verliefd*. As this expression only represents a small subset of loving, namely being in love with another person, and since it can only refer to persons and not to objects, we have focused on other syntactic categories to express love more generally. *Houden van* (loving) was found with a rather high frequency (54). We had some trouble obtaining expressions for liking another object or person, as there is no straightforward translation of liking in Dutch. Most often liking is expressed as *leuk vinden* (finding it nice), where *leuk* can be replaced by some other aspect of the object. For example, an agent can find a painting pretty and he can find an ice cream delicious. We searched for instances of *vinden* in the corpus and obtained 21 tokens of *vinden*, differing in the aspects of the liked object. The most frequently obtained aspects were *leuk* (5) (nice) and *mooi* (6) (pretty). As these aspects are quite general, we selected them for the storyteller. We thus ended up with four expressions, referring to three different meanings of liking: *verliefd*, *houden van*, and *mooi/leuk vinden*. In section 7 we will discuss when each of these expressions should be used.

Emotions that were not expressed

We now discuss the mismatches between the emotion model and the adjectives from the corpus, first focusing on emotions for which we did not obtain adjectives. We can think of two explanations why we did not obtain any appropriate adjective for some emotions. One possibility is that the emotions are usually expressed by other lexical classes, like verbs or nouns. The other is that some emotions are simply not or seldom expressed in Dutch (or in

these fairy tales that were originally written in the nineteenth century). In order to find out which clarification is suitable for each unexpressed emotion type, we will look up the Dutch equivalences of the expressions suggested in table 3.2 in the corpus.

Happy for For this emotion Ortony et al. suggested adjectival expressions that we did not find in our corpus. This is remarkable. Either the emotion is not very common and the adjectival expressions were therefore not found in the corpus, or the Dutch language uses other constructions to express this emotion. Because we did find expressions for the antonym emotion, pity for, we insisted on selecting an expression for this emotion. We searched through the corpus manually to find out how the happy for emotion was expressed and we found rather implicit descriptions of this emotion type. Example 3.1 illustrates this.

(3.1) "*...de vogels wisten best dat de dode man nu daarboven in de hemel was, dat hij vleugels had veel mooier en groter dan de hunne, dat hij nu gelukkig was omdat hij goed was geweest hier op aarde, en daar waren ze blij om.*" (cited from De Reisgenoot, "Hans Christian Andersen - Sprookjes en vertellingen" uitgegeven door Van Holkema en Warendorf / Unieboek, Bussum, 1975.)

"*...the birds seemed as if they knew that the dead man was now in heaven, and that he had wings much larger and more beautiful than their own; and he was happy now, because he had been good here on earth, and they were glad of it.*"

In the above passage, the birds are glad that the dead man is happily in heaven. The event of the dead man being happily in heaven has no positive consequences for the birds. However, the fortune of the dead man does bring about a feeling of happiness to the birds. We can conclude that the birds feel happy for the dead man and that the happy-for emotion is only expressed implicitly in the above passage.

As we aim to express the emotions of the Virtual Storyteller in a natural way, we prefer using similar expressions to those obtained from the corpus. However, we are restricted to the current possibilities of the system. At this moment, the Virtual Storyteller is in a stage in which it is important that the natural language is generated completely automatically. Expressing emotions implicitly and automatically seems to be beyond this goal as it would require more linguistic insights from the storyteller than it currently has. We will therefore restrict ourselves to expressing emotions explicitly.

As we did not find any explicit expressions of the happy-for emotion in the corpus, we chose to simply translate the suggested happy-for expressions from table 3.2. We found the Dutch expression *blij voor*.

Gloating Ortony et al. suggested two non-adjectival expressions for gloating. We did not obtain these expressions from our corpus. Since we are not familiar with any other expressions, we again concluded that it is not very common to express this emotion explicitly. Just as for the happy for emotion, we conducted a corpus study manually to find implicit expressions for the gloating emotion. Example 3.2 shows how the gloating emotion can be expressed:

(3.2) "*Op de dag van het bal was Assepoester heel verdrietig. De gemene stiefzusters hadden dit door en lachten haar uit.*"(cited from Cinderella [4], The Grimm brothers)

"*On the day of the ball Cinderella felt very sad. The mean stepsisters noticed this and started to make fun of her.*"

As we see above, one way to express the gloating emotion is by laughing at someone who is feeling sad. Cinderella is not allowed to come to the ball, and her stepsisters feel happy about this misfortune of Cinderella. This is compatible with the type specification of gloating. The above construction can only be used if the characters are really having fun about the misfortune of someone else and it is therefore not a general description of gloating but just one possible description. Now that we have seen that gloating is indeed an emotion that is expressed in natural language, we need to find our own explicit expression for gloating. We selected the noun *leedvermaak* (schadenfreude), meaning being pleased about the misfortune of someone else. This corresponds to the type specification of gloating.

Hope (self) For the hope (self) emotion, many expressions were suggested. Again we did not find any of the suggested adjectives. As *bang* is a very common adjective for expressing fear (self), it is of importance that the antonym emotion hope (self) can be expressed as well. We should select a lexical item that expresses the opposite of *bang*. The suggested expressions differ in terms of the eliciting conditions. The token *hopeful* is used when someone expects a positive event to be happening, the token *hope* is used when someone has a great desire for some positive event (but his/her expectation that the event will indeed occur is not necessarily high). For example, an agent hopes that he will win the lottery, but he may not be very hopeful about this.

Which of the suggested token expresses the opposite of *bang* (afraid)? We use *bang* to refer to an emotion that is a reaction to the consequences of an event. The consequences of an event can bring about the *bang* emotion if they either enhance the likelihood of an undesirable event or enhance the undesirability of the event that is likely to occur. It is more likely that the consequences change the likelihood of an event than that they change the desirability. The desirability of an event is mostly fixed; it mainly depends on an agent's preferences. For example, most agents do not want to die. They will experience the fear emotion in risky situations, like when they stand on a big cliff. The agent's fear is then generally determined by the situation and not by the undesirability of the prospect.

We consider *bang* to refer to the fear emotion for which undesired prospects are likely to happen. The antonym emotion should then represent a situation in which desired prospects are likely to happen. From the suggested expressions, the meaning of the adjective hoopvol (hopeful) best represents this feeling.

Hope (others) No adjectives were obtained from the corpus for this emotion, and since this emotion was not part of the OCC model, no other expressions were suggested by Ortony et al. We could use the expression *hoop voor* (hope for), but this expression tends to refer to prospective events, rather than to hoping for some person or object. We therefore looked for synonyms of hope and came up with the verb *verlangen naar* (desire). This token expresses more concretely the hope for some object or person. The verb *verlangen naar* was obtained ten times from the corpus. This expression was selected for the Virtual Storyteller.

[4]This passage was taken from http://www.sprookjes.nu/sprookjes/assepoester

Shame Of all expressions suggested for shame by Ortony et al. the token *shame* itself was the most general one. We looked up the Dutch equivalence of shame, *zich schamen*, in the corpus and obtained eight hits. The other suggested expressions were not appropriate to express shame, because they were either too specific (e.g. feeling guilty) or their Dutch equivalences were not obtained from the corpus. We therefore selected the verb *zich schamen*.

Admiration Of the suggested expressions, we obtained *bewonderen* (admire), *bewondering* (admiration), and *respect* (respect) from the corpus with a frequency of two, one and one respectively. As the frequency of the tokens differed only little we made our choice taking into account the complexity of the tokens. The most straightforward way of expressing this emotion would be by using the verb *bewonderen*, rather than using the noun *bewondering/respect* in combination with another verb. We will therefore use *bewonderen*.

Reproach The results obtained for the reproach emotion were similar to those obtained for the admiration emotion. We obtained one noun, *minachting* (2) and two verbs, *minachten* (1) and *verachten* (1). For similar reasons, we preferred expressing this emotion by a verb. The verb *verachten* and *minachten* are synonyms, meaning *disdain*. Because variation in natural language is desirable, both verbs were selected for the Virtual Storyteller.

Disliking Of all suggested expressions for disliking, we obtained only few tokens in the corpus, namely *afkeer* (1) (aversion), *afkerig zijn van* (1) (being averse) and *vinden* (5) (finding) combined with negative properties of objects or agents. The properties that the token *vinden* referred to were all quite specific, e.g. *scary* and *ugly*. As none of the tokens were obtained frequently, we decided to select none of them directly. Instead, we focused on finding expressions that could function as antonyms for the obtained liking expressions.

Using Roget's Thesaurus[5], we obtained the antonym *afstotelijk vinden* (finding something repulsive) for *being in love*, the antonym *afkerig zijn van* for *liking* and the antonym *haten* (hating) for *loving*. In chapter 7 we will decide when and whether to use each of these expressions.

Expressions that could not be categorized

Some adjectives that were obtained from the corpus could not be categorized according to the emotion types that were selected by Rensen. The frequencies of these adjectives inform us about how common the adjectives are.

Gratitude/Anger The first two rows show synonyms of the emotions *anger* and *gratitude* which constitute an emotion pair and these are part of the well-being emotion in the OCC model. From all adjectives, the anger emotion *boos* was obtained most often from the entire corpus, and its frequency of 83 shows that it would be justified to add it to the emotion model of the Virtual Storyteller. Rensen left this emotion out because he thought it would be too specific in comparison with the other emotions. Although the OCC model indeed shows that this emotion is a sub-emotion of the well-being and attribution emotions, its frequency indicates that it should be part of an agent's emotion model. We therefore aim to add this emotion to the emotion model.

[5]Roget's Thesaurus can be found at http://thesaurus.reference.com/

According to the OCC model, the type specification of anger is: "Disapproving of someone else's blameworthy action and (being displeased about) the related undesirable event". This emotion is thus a reaction to two different eliciting conditions: the action of another agent and the consequence of the related event. Both conditions determine the intensity of this emotion. The antonym emotion, gratitude is defined as: "Approving of someone else's praiseworthy action and (being pleased about) the related desirable event". This emotion results from the same eliciting conditions as its antonym emotion for which the conditions have both positive values.

We determined how the emotion should be expressed, using our corpus. For the anger emotion, the expression *boos* (angry) was used most often in the corpus and applying our selection criterion we selected no other lexical items. For the gratitude emotion, Ortony et al. proposed the expressions *appreciation*, *gratitude*, *feeling indebted* and *thankful*. We only obtained *dankbaar* (thankful) from the corpus with a frequency of four so this emotion will be used to express gratitude.

Satisfaction Seventeen instances of *tevreden* (satisfied) were obtained from the corpus. The OCC model assumes that *satisfied* is an expression of the prospect-based emotion for which the desirable prospects are confirmed. For this emotion, the type specification is: "Being pleased about the confirmation of the prospect of a desirable event". We looked up the instances of *tevreden* in the corpus and found more general meanings of *tevreden*, like example (3.3).

(3.3) "*Er waren eens een koning en een koningin, die heel tevreden met elkaar leefden en twaalf kinderen hadden.*" (cited from De Twaalf Broers, "De Sprookjes van Grimm", translated by M.M de Vries-Vogel)

"*Once upon a time there were a king and a queen. They lived happily together and had twelve children, all boys.*"

We counted the instances of *tevreden* in the corpus that did correspond to the satisfaction emotion and found nine instances. Because this frequency was rather low we decided to stick with Rensen's choice and left this emotion out.

Disappointment This is also a specific emotion that is part of the OCC model but it was not included in the emotion model of the Virtual Storyteller. With a frequency of 1, it doesn't seem necessary to add it to the emotion model.

3.5.3 Conclusions

We have conducted a corpus study on fairy tales in order to evaluate both the OCC model and the emotion model of the Virtual Storyteller and in order to obtain lexical items for the emotions in the emotion model. The following conclusions can be drawn.

The OCC model There is a mismatch between the emotions in the OCC model and the emotion lexicon of fairy tales. We searched for emotion adjectives in the corpus and we found that not all emotions in the OCC model are expressed by adjectives. Some emotions in the OCC model were expressed by other parts of speech than adjectives (like nouns and verbs), others were only expressed implicitly in the corpus and it was difficult to identify these

emotions. The frequencies of the lexical items were often asymmetric for emotion pairs. For example, for the emotion pair hope/fear(self) we found 59 lexical items for the fear emotion, but the hope emotion was never expressed explicitly in the corpus. The question is then whether the hope emotion is just not experienced in the stories, or whether the emotion is not interesting or relevant enough to express in stories. Here, we are not sure what the reason is, but it is an interesting observation that tells us that we should pay attention to this aspect of an emotion.

If the hope emotion is indeed hardly experienced, the fear emotion can be regarded as an isolated emotion type, which is hardly related to its corresponding positive emotion, hope. Ortony et al. selected emotion pairs for all type descriptions, independently of the frequency/commonness of the positive or negative emotion of the pair. Because the commonness of the emotions is not taken into account for the construction of the OCC model, the relations in the OCC model appear to be quite simplified. For some emotion pairs, the two emotion types could be very much related and for others, it can be argued whether one of the two emotions types should actually be included in the model. The construction of the emotion types in the OCC model should rely on more than a theoretical type description.

What exactly can we conclude from the results of our corpus study? Although the frequencies of the emotions in the corpus question whether some emotions are correctly admitted to the OCC model, we are not justified to falsify the OCC model since the corpus only contains fairy tales and it is therefore not representative for all Dutch language. The fairy-tale world is a simplification of the real world and the language that is used in the storytelling is also rather simple. Taking into account that the fairy-tale domain is not complete, we can only derive conclusions from expressions that *are* in the corpus rather than expressions that are *not* in the corpus. We searched automatically for emotional expressions represented by the most natural lexical category, adjectives. The emotion adjectives we obtained from the corpus could all be categorized according to the emotion types in the OCC model. For selecting emotion adjectives, we used the definition of Ortony et al. We can therefore conclude that, according to Ortony et al.'s definition for emotions, the OCC model recognizes the emotions that are expressed by adjectives in fairy tales. We did not identify any missing emotions, based on our corpus study.

Concluding, the OCC model seems to be a complete model of emotions, given the fact that all obtained emotion adjectives could be categorized according to the emotion types in the OCC model. However, some emotion types in the OCC model appear redundant and lack empirical evidence from our corpus study. The commonness of the emotion types received no attention from Ortony et al., but we believe that it can represent the relation between the different emotion types. It should be questioned whether emotion types with very different frequencies can be admitted to the model as an emotion pair. We would therefore advise to conduct a corpus study on a more general domain than the fairy-tale domain and take the frequencies of the emotional expressions into account when evaluating the OCC model.

The emotion model of the Virtual Storyteller Rensen chose to select only the general emotions from the OCC model for the Virtual Storyteller. The data of our corpus study on fairy tales have shown that he left out one very common emotion, namely the *anger* emotion. The adjective *boos*, which refers to this emotion, was obtained very frequently from our corpus. We also obtained the corresponding positive emotion from the corpus, the *gratitude* emotion, expressed by *dankbaar*. The emotion pair gratitude/anger should be added to the emotion

model, in order to produce believable stories.

For each of the eight emotions that were already included in the emotion model of the Virtual Storyteller, we found enough linguistic evidence in the corpus. We will therefore keep using these emotions in the storyteller.

Emotion label	**Expressions**
Happy for	blij voor(0)
Resentment	afgunstig op(5)
Gloating	leedvermaak hebben over(0)
Pity for	medelijden hebben met(7)
Hope (self)	hoopvol(0)
Fear (self)	bang(51), angstig(6)
Hope (others)	verlangen naar(10)
Fear (others)	bang voor(11)
Joy	blij(60)
Distress	bedroefd(42), verdrietig(8), treurig(7)
Pride	trots(26)
Shame	zich schamen(8)
Admiration	bewonderen(1)
Reproach	minachten(1), verachten(1)
Liking	houden van(52), mooi vinden(6), lief vinden(5), verliefd op(3)
Disliking	haten(0), afstotelijk vinden(0), afkerig zijn van(1)
Gratitude	dankbaar(4)
Anger	boos op(83), kwaad op(12)

Table 3.4: Selected expressions for the Virtual Storyteller

Expressing the emotions Table 7.1 shows the expressions that we selected for the emotions of the Virtual Storyteller with their corresponding frequencies in the corpus. If possible, we selected adjectives to express the emotions. For some emotions, no lexical items were obtained from the corpus and we looked for Dutch translations of the lexical items suggested by Ortony et al. The lexical items in table 7.1 will be added to the lexicon of the Virtual Storyteller.

Chapter 4

Lexical Pragmatics

We have seen that each emotion in the emotion model of the Virtual Storyteller represents an intensity scale of an emotion pair. In chapter two we have identified lexical items, generally gradable adjectives, for the positive and negative ends of the emotion scales by conducting a corpus study of eighty fairy tales. The current chapter focuses on expressing the gradable spectrum between the two ends of the scale. Generally speaking, modifiers are used to express certain intensities of a gradable adjective. Gradable adjectives in combination with modifiers regularly give rise to conversational implicatures; the expressions mean more than their literal meaning. For example, the utterance *I am not happy* is generally interpreted as *I am unhappy*, whereas it literally only refers to any intensity other than happy due to the gradability of the adjective *happy*. If we want to express the emotions of the character agents in a differentiated way, we need to be sure about the meanings of conversational implicatures like this one.

Lexical Pragmatics provides us with a theoretical explanation for the rise of conversational implicatures. In natural language, the meaning of an utterance does not only depend on its formally defined meaning; it also depends on the meanings expressed by other, related words. The idea is that a speaker should try to use the simplest expression available for a certain meaning. If the speaker uses another more complex expression, he or she is considered to mean something else. The lexical alternatives thereby give rise to a conversational implicature. Here, we consider the lexical alternatives the context of the utterance, just like other linguistic aspects, such as previous discourse and non-linguistic aspects, such as social factors. We can say that the formal meaning of a word in combination with the context of the word determines how the word is interpreted by the reader. The field of Lexical Pragmatics is concerned with the effect of the context, mainly referring to the lexical alternatives, on the interpretation of an utterance. For a successful communication it is essential that both hearer and speaker know the alternative expressions.

We aim to express the emotions of the character agents in an accurate and natural way in the story. For the accuracy, we should put nuances in the emotional expressions so that the intensities of the emotions are communicated to the reader. Speakers often formulate conversational implicatures to this end, like *I am not happy* or *I am not very happy*. Such expressions were also found in the fairy-tale corpus. Expressing conversational implicatures like these in the story, would then contribute to a natural story text. However, the implications of the conversational implicatures must be clear to be able to model emotions in a natural way. It is absolutely necessary to capture the extra implicated meanings of such expressions if we want to use them in the story text. We need to be sure that what we intend to conversational

implicate by our expression is indeed interpreted by the reader. Therefore, we study the theoretical findings of Lexical Pragmatics in this chapter and highlight how conversational implicatures regarding gradable adjectives are interpreted according to the theory.

4.1 Grice's Maxims

Before focusing on the conversational implicatures mentioned above, we first present the foundations of the theory about conversational implicatures. Grice (1975) was the first to argue that it was possible to give a systematic account to meaning beyond just what was said. This account is based on the idea of the conversational implicature, and he introduced this term to refer to the meaning of an utterance that was conveyed by the context of the utterance, including related expressions that were not said. Conversational implicatures arise from the general assumption that the speaker complies with what Grice calls the Cooperative Principle. The Cooperative Principle presumes that a speaker tries to communicate information to the hearer in an efficient way that takes the least effort from the speaker and the hearer. The Cooperative Principle says:

Make your contribution as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged. (Grice, 1975: pp 307)

The Cooperative Principle can be translated into concrete rules which have to be obeyed by the speaker in order to comply. Grice defined four rules, which are called maxims, that underlie this co-operative use of language.

1. The maxim of Quality
   try to make your contribution one that is true, specifically:
   - do not say what you believe to be false
   - do not say that for which you lack adequate evidence
2. The maxim of Quantity:
   - make your contribution as informative as is required
   - do not make your contribution more informative than required
3. The maxim of Relevance:
   - make your contributions relevant
4. The maxim of Manner:
   - avoid obscurity of expression
   - avoid ambiguity
   - be brief
   - be orderly

The maxims have been constructed on the basis that spoken language strives for a maximally effective exchange of information. If one assumes that a speaker would normally obey the maxims, unless there is a valid reason to do not, conversational implicatures can be derived. The different types of conversational implicatures than can arise, are given in section 4.1.1 with corresponding examples. For now, we want to emphasize that none of the conversational implicatures are logically grounded. Conversational implicatures are very different from assertions. If a speaker contradicts the conversational implicature, by expressing the opposite, the implicature is cancelled and a logical contradiction is prevented. An important characteristic of conversational implicatures is thus that they can be cancelled by the speaker. We should also realize that a conversational implicature is not part of the conventional meaning of a word. Because of that, the same utterance can lead to different conversational implicatures in different contexts.

4.1.1 Interpretation of Grice's Maxims

we now give examples of conversational implicatures, grouped by the three categories that were distinguished by Grice.

Group A - No maxim is violated

I am out of petrol.
There is a garage around the corner. (Grice, 1977: pp 311)

In conversations, the first assumption is that a speaker does not violate any maxim. An utterance is therefore interpreted as obeying all maxims, unless it becomes apparent that this is not so. In the above sentence, the maxim of Relevance is only obeyed if the garage around the corner is open and has petrol to sell. Assuming that the speaker obeys the conversational maxims, the hearer assumes that the information about the garage is relevant in the current context. The conversational implicature is then that the garage is open and sells petrol.

Group B - A maxim is violated, but its violation can be explained by the supposition of a clash with another maxim

Where does Catherine live?
Somewhere in the South of France. (Grice, 1977: pp 311)

If a maxim is violated, the hearer has to find some reason why the speaker would violate a maxim. It is then likely that the speaker violates a maxim in order to obey another maxim. In the above sentence, the maxim of Quantity is violated; the speaker is less informative than the hearer required. Obviously, the speaker does not know the exact town and has to be vague not to violate the maxim of Quality ("do not say for which you lack evidence"). The implicature is thus that the speaker does not know where Catherine lives. From this example we can conclude that there must be an order of rank between the maxims. The maxim of Quality apparently gets more priority than the maxim of Quantity.

Group C - A maxim is flouted in order to getting in a conversational implicature

War is war (Grice, 1977: pp 311)

Here the maxim of Quantity is flouted. The utterance is a tautology whose truth value is always positive, so it is a totally non-informative remark. The hearer should then figure out why the speaker would say such a thing and should conclude that a conversational implicature is involved. The utterance could implicate: "This kind of drama is just what happens in war."

4.1.2 Generalized conversational implicatures

The above examples of conversational implicatures are all particularized ones; the implicatures are carried in special contextual cases and they would not arise from the utterances in an isolated context. Next to these particularized conversational implicatures, there exist generalized conversational implicatures. These are utterances that carry conversational implicatures even in the absence of any particular contextual features. The implicatures are normally conveyed, regardless of the context. Grice gives the following example of a generalized conversational implicature.

X is meeting a woman this evening. (Grice, 1977: pp 314)

This sentence normally implicates that the woman to be met is not X's wife. That is to say, the expression *an X* refers to someone who is only remotely related to the speaker. The implicature is derived because the speaker is assumed to obey the maxim of Quantity, which says that the speaker should be as specific as possible. Since the speaker does not denote his relation with the women to be met, the listener assumes that the speaker has no special relation with the woman and the conversational implicature is conveyed.

Just like the above example, the conversational implicature *I am not happy* that we started with in this chapter is also a generalized conversational implicature. The expression has obviously a preferred negative interpretation, even in an isolated context like it is presented here. We will explain the rise of this implicature in section 4.3

4.2 Principle of Quantity

Some scientists believe that Grice's maxims are more abundant than necessary to account for all conversational implicatures. Atlas & Levinson (1981) and Horn (1984) therefore reduced the Gricean maxims by proposing only two principles. The principles correspond to the two rules of the maxim of Quantity. The Q-principle says: "Say as much as you can" and it strives for informative communication. The I-principle says: "Say no more than you must" which focuses on communicating efficiently. The two principles keep the informativeness and the efficiency of a communication in balance. Acting on these principles the required cooperation from the speaker and the hearer are determined (Horn). The Q-principle minimizes the hearer's effort by requiring that the speaker is as informative as he can. The I-principle minimizes the speaker's effort by encouraging him to say no more than is necessary.

From the Q-principle, implicatures arise that interpret the utterance containing the maximum amount of information. Scalar implicatures belong to this group of implicatures. These implicatures apply when a speaker asserts that a point obtains on a scale. He then implicates that a higher point than the point that was expressed, does not obtain. This can be formulated as follows:

Given any scale of the form $\langle e_1, e_2, e_3, ...e_n \rangle$, *if a speaker asserts* $A(e_2)$ *then he implicates*

$\neg A(e_1)$, if he asserts $A(e_3)$, then he implicates $\neg A(e_2)$ and $\neg A(e_1)$, and in general, if he asserts $A(e_n)$, then he implicates $\neg A(e_{n-1})$, $\neg A(e_{n-2})$ and so on, up to $\neg A(e_1)$. (Levinson, 1983: pp 133)

Examples of scalar implicatures are the following two:

The show was good → The show was good but not excellent

You can buy fruit or vegetables → You can not buy both fruit and vegetables

The first implicature follows from the scale related to *good*. An example of such a scale is ⟨excellent, good, OK⟩. For the second scalar implicature the scale ⟨and, or⟩ applies.

As scalar implicatures arise when a speaker tries to indicate a point on a certain scale, gradable adjectives that express properties that can be measured on a scale can also raise this type of implicature. An example of a gradable adjective is the adjective *happy*, for which the antonym pair {happy, unhappy} refers to the intensity scale of happiness. Generally, the expressions *happy* and *unhappy* are interpreted as the two ends of the scale, but in fact intensity markers can be used to broaden the scale as in *very happy* and *very unhappy*. However this has no effect on the scalar implicatures. So, an expression like *I am rather happy* can be seen as a scalar implicature. It implicates that it is not the case that the speaker is *very happy* or *happy*, but that the speaker is only *rather happy*. A similar scalar implicature is derived from the expression *happy*, which implicates *happy, but not very happy*. Other, non-gradable adjectives like *male* cannot raise scalar implicatures since they can normally not be modified by intensity markers.

4.2.1 Formalizing the Q and I principle using Optimality Theory

The Q and I principles have been formalized by Blutner (2001), using the weak bidirectional version of Optimality Theory (OT) (Prince & Smolensky, 1993). Optimality Theory describes linguistic grammars in terms of ranked constraints. The idea is that a language can generate different candidate forms for a particular meaning and the optimal form, the grammatical form, is the one that violates the least important constraints. Weak bidirectional OT (Blutner, 2000) searches for optimal form-meaning pairs rather than for optimal forms only. Here, optimality is defined in terms of the complexity of the form or meaning. For example, the form *happy* is regarded as less complex (more optimal) than *rather happy*. And, for meanings, Blutner assumes that $\smile$ and $\frown$ are most optimal. [1] $\smile$ and $\frown$ are both more optimal meanings than any other meaning on the same scale. An optimal form-meaning pair is then (*happy*, $\smile$), because *happy* is the most optimal form to express $\smile$, and $\smile$ is the most optimal meaning that *happy* can refer to (while, literally, *happy* can refer to any intensity on the happy-scale). Weak bidirectional Optimality Theory thus says that *happy* refers to $\smile$, because the two are an optimal form-meaning pair.

Beside optimal form-meaning pairs, weak bidirectional OT also allows for super-optimal form-meaning pairs. An example of a super-optimal pair is the pair (*rather happy*, $\smile$). This pair is super-optimal because the alternative more optimal form for the meaning $\smile$, *happy*, is not available since *happy* is already used to express $\smile$. The same is true for its meaning. The

[1] For referring to the meanings on the {happy, unhappy} scale, we use the symbol $\smile$ for happy, $\smile$ for a bit happy, – for indifferent and $\frown$ for unhappy.

form *rather happy* can formally refer to the more optimal meaning ⌣, but this is impossible since another (more optimal) form, *happy*, already expresses this meaning. Thus, although *rather happy* and ⌣ have both more optimal alternatives, they are accepted as a super-optimal pair because the alternatives are already in use by another, optimal, form-meaning pair. With this analysis, we have explained the interpretation of the scalar implicature *rather happy* in a more formal way. Concluding, whether a form-meaning pair is grammatical (optimal or super-optimal) according to weak bidirectional OT depends on its alternative pairs.

Combining weak bidirectional Optimality Theory, which allows for the above described optimal and super-optimal pairs, with the Q and I principle (proposed by Atlas & Levinson and Horn), Blutner (2001) formalized the balance between efficiency and informativeness. The preferred meaning for a given utterance is determined by comparing different form-meaning pairs ((f, m) pairs), by means of their efficiency (>). A pair's efficiency depends on the complexity, called markedness in OT terms, of its form and meaning. A form or meaning is less marked than an alternative form or meaning if it is less complex. If the efficiency of a form-meaning pair is defined, it can be verified if the pair is optimal, using the following definition:

A form-meaning pair(f, m) is optimal if it can be generated and iff it satisfies both the Q-principle and the I-principle, where:

(Q) (f, m) satisfies the Q-principle iff there is no other pair (f', m) that can be generated such that (f', m) > (f, m)

(I) (f, m) satisfies the I-principle iff there is no other pair (f, m') that can be generated such that (f, m') > (f, m)

The Q-principle focuses on the speaker's position and determines which form the speaker should select for a certain meaning. The Q-principle requires that the speaker chooses the most unmarked expression available to represent a given meaning. To be more concrete, the form is required to be the most straightforward expression of what the speaker wants to communicate. Forms that are more marked than their alternative forms are dropped.

The I-principle focuses on the hearer's position and determines which interpretation would be correct for a certain utterance. The principle requires that a form's corresponding meaning is the most unmarked one of its alternatives.

Weak bidirectional OT also gives a formal definition of (super-)optimal pairs. Although super-optimal pairs do not contain the least marked meanings or forms, they are allowed because they are least marked among the forms and meanings not already associated in a pair. The definition of a (super-)optimal pair is an extension of the definition of an optimal pair because it adds a requirement to the alternative pair: the alternative pair should be a (super-)optimal in order to replace the current pair.

The definition of a (super-)optimal (f, m) pair is:

A form-meaning pair(f, m) is (super-)optimal if it can be generated and iff it satisfies both the Q-principle and the I-principle, where:

(Q) (f, m) satisfies the Q-principle iff there is no other pair (f', m) that can be generated that satisfies the I-principle such that (f', m) > (f, m)

(I) (f, m) satisfies the I-principle iff there is no other pair (f, m') that can be generated that satisfies the Q-principle such that (f, m') > (f, m)

According to the above definition, (f, m) pairs can be super-optimal even if both the form and the meaning have a more efficient counterpart. This is due to the extra requirement that a possible counterpart must satisfy both principles and therefore must be an (super-)optimal pair in order to replace the current pair. If alternative pairs with a more efficient meaning or a more efficient form are not (super-)optimal pairs, the current pair is super-optimal and thus accepted in the language.

With an OT-tableau, it can be illustrated how super-optimal forms arise. An example of an OT-tableau is given in figure 4.1. The hand marks the optimal candidate, the hand between brackets marks the super-optimal candidate. The arrows point to a pair that is more efficient than the current pair. In this example, form f1 is less marked than form f2, and meaning m1 is less marked than meaning m2. This makes pair (f_1, m_1) the optimal pair. Pair (f_2, m_2) is super-optimal because the counterparts (f_1, m_2) and (f_2, m_1) both have an optimal neighbour pair, namely pair (f_1, m_1). The existence of super-optimal pairs thus allows for marked meanings and forms.

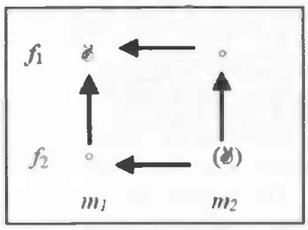


Figure 4.1: bidirectional OT tableau

Weak bidirectional OT differs from normal bidirectional OT in that it allows for super-optimal form-meaning pairs. Due to the existence of super-optimal pairs, this analysis has the tendency that less marked forms have less marked meanings and more marked forms have more marked meanings. With normal bidirectional OT, we would accept the same optimal form for different meanings. Consider the forms *knife* and *cutter*, assuming that *knife* is less marked than *cutter*. Normal bidirectional OT would assign the optimal form *knife* to any cutting object, including grinding machines. But, in English, the form *cutter* is used to refer to such objects. Weak bidirectional OT explains these denominations, due to the allowance of super-optimal pairs. Because *knife* already designates the object knife, the more marked form *cutter* is used to refer to other cutting objects than knives.

The same is true for ironic expressions like in example 4.1. Weak bidirectional OT can account for this implicature, because it concerns a super-optimal form-meaning pair. The form and meaning thus correspond to f2 and m2 (in figure 4.1) respectively. The optimal pair would be example 4.2 corresponding to pair (f_1, m_1) in figure 4.1. This pair is optimal, because its form and its meaning are least complex. Because the alternative form-meaning pairs for example 4.1 are replaced by their optimal neighbour, pair (f_1, m_1), example 4.1 is a super-optimal pair.

(4.1) Form: *Tom lifted the corners of his lips.*
Meaning: *Tom produced an artificial smile.*

(4.2) Form: *Tom smiled.*
Meaning: *Tom smiled in a regular way.*

4.3 Applying OT to antonym pairs

We have seen how weak bidirectional Optimality Theory can account for the effect of a marked form on its meaning. Marked forms are often used in expressing emotions with gradable adjectives. For instance, if a speaker wants to express how happy he is, he could say *I am not unhappy*, which is obviously more marked than the alternative utterance *I am happy*. Applying weak bidirectional OT, we can argue theoretically what the preferred interpretation of such an expression is.

4.3.1 Blutner's interpretation

Blutner (2001) applied weak bidirectional OT to the antonym pair {happy, unhappy} and formally described how the negations of these gradable adjectives should be interpreted. The OT tableau is shown in figure 4.2. The possible forms are *happy*, *not unhappy*, *not happy* and *unhappy* and there are three possible meanings, indicated by ⌣, – and ⌢. The pairs that cannot be generated (because the form literally excludes a given meaning, as in (*happy*, ⌢) are designated by grey areas. Applying the Q-principle on these forms their markedness needs to be specified. Blutner assumes that the number of negation morphemes is a measure for the markedness of the form. This means that the expressions *not happy* and *unhappy* have similar markedness because they contain one negation morpheme each. The form *happy* is then least marked and the form *not unhappy* is most marked. As a simplification, Blutner uses a discrete happiness scale of three intensities of happiness: ⌣, – and ⌢. Here Blutner's assumption is that the middle state is more marked than the two other states, so we can apply the I-principle that searches for the least marked meaning.

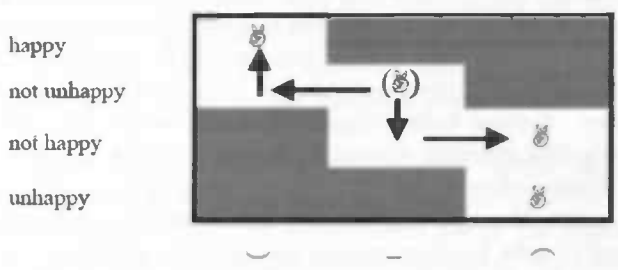


Figure 4.2: Bidirectional OT tableau for the antonym pair {happy, unhappy} taken from Blutner (2001)

We first have to find optimal (f, m) pairs for which the form and meaning are more efficient than their direct neighbours. The form *happy* with the meaning ⌣ constitutes an optimal

pair, because both its form and its meaning are least marked. On the other end of the scale the forms *unhappy* and *not happy* with the meaning ⌒ are also optimal pairs, as their complexities are similar and there are no alternative forms or meanings with lower complexities. Once the optimal pairs have been identified, we can look for super-optimal pairs. The form *not unhappy* combined with the meaning – is super-optimal because its direct neighbours are not (super-)optimal.

The OT tableau can explain two implicatures. The form *not happy* should be interpreted as ⌒, because the alternative form for ⌒ (namely *unhappy*) is not less marked than *not happy*. This phenomenon is called 'negative strengthening'; the negation of the adjective is interpreted as the opposite of the adjective, and it excludes the spectrum in between. The other implicature that the tableau can account for is the effect of a double negation, as in *not unhappy*. Logically, this utterance could refer to everything but unhappy, so – or ⌣, but the OT tableau suggests it should be interpreted as –, because the meaning ⌣ is already part of another optimal form-meaning pair. The informal explanation is that *not unhappy* should be interpreted as – because the other possible interpretation (⌣) could have been grasped more economically. Because the speaker used a more complex expression he must have meant something else. The term litotes is used to refer to these kinds of double negations.

Blutner's use of bidirectional OT seems to account for the effect of negative strengthening and litotes, but one remark has to be made. The choice of the antonym pair {happy, unhappy} is rather striking as the pair itself already contains a negation morpheme. We therefore replace the adjective *unhappy* by the adjective *sad* and apply weak bidirectional OT to the antonym pair {happy, sad} to study the effect of the lexical choice of the antonym on the interpretation of the conversational implicature.

4.3.2 Applying OT to non-morphologically complex antonym pairs

We use weak bidirectional OT to find out how marked forms of a non-morphologically complex antonym pair are interpreted. We use the pair antonym pair {happy, sad} and apply Blutner's criteria of markedness for forms and meanings. The corresponding OT tableau is shown in figure 4.3.

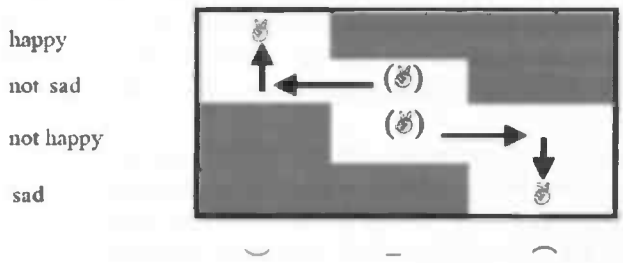


Figure 4.3: Bidirectional OT tableau for the antonym pair {happy, sad}

Now, there are only two optimal form-meaning pairs, namely (*happy*, ⌣) and (*sad*, ⌒). The pair (*not happy*, ⌒) is not optimal, because the form contains one negation morpheme and it is therefore more marked than the form *sad*. Now we have two forms left and only one

meaning. Both forms could refer to the meaning – by being super-optimal pairs. But, as it would be redundant if two very different forms would refer to the exact same meaning, it is plausible that the two forms have meanings that differ slightly from each other but do approach the meaning –. This indicates that it would be appropriate to use a smaller scale division in order to account for more subtle differences. Blutner's spectrum division of three intensities seems to be an inappropriate simplification for the current antonym pair. Unfortunately, he did not explain his choice for the spectrum division for the antonym pair {happy, unhappy}.

Although our OT tableau has a scale that may not be appropriate for the antonym pair, we can conclude that the results of applying bidirectional OT to the antonym pair {happy, sad} are different from the results of the antonym pair {happy, unhappy}. The interpretation of the form *not happy* is $\frown$ if we take *unhappy* as its alternative form, and it is close to – if we take *sad* as its alternative form. This is a consequence of the fact that the form *unhappy* is more marked than *sad*, which is the alternative form of *not happy*. The interpretation of a form thus depends on the antonym pair that is associated with the form. This corresponds to the assumption of weak bidirectional Optimality Theory that the interpretation of a word or utterance is relative to its alternative forms. We therefore have to be very cautious in selecting alternative forms for the OT tableau, as these indirectly determine the interpretation of a certain form.

Because alternative forms can determine the interpretation of a certain form, the context may play an important role in the interpretation. Levinson (2001) distinguishes general scalar implicatures from particularized scalar implicatures that are derived from the context. For general scalar implicatures, the lexicon contains alternates that are ordered in strength. For particularized scalar implicatures a temporary scale is constructed from the context. For example:

How far did you run?
I got to Assen

The implicature here is that the speaker did not get to Zwolle, considering that speaker and hearer both live in Groningen. Speaker and hearer construct a distance scale with Groningen as its starting point. Such a temporary scale is called a Hirschberg scale (by (Hirschberg, 1985)). This temporary scale thus determines how expressions are interpreted. So, in the case of the adjective *happy*, which can have two antonyms, *unhappy* and *sad*, the context may select one of the two as the antonym adjective. If the context does not refer to any scale, we should find out which antonym pairs humans generally prefer.

Weak bidirectional OT has provided a theoretical approach for the interpretation of complex forms. However, it also leaves us with a number of questions that Blutner simply answered with a number of assumptions. These are assumptions about the number of scale divisions that humans have for the meaning of an emotional expression, the complexity of forms and the complexity of meanings. Blutner used three scale divisions (meanings) for constructing the OT tableau, namely $\smile$, – and $\frown$. He assumed that the number of negation morphemes determines the complexity of a form. *Not happy* and *unhappy* are therefore similarly complex forms. And, about the complexity of meaning, he assumed that the meaning – is more complex than the other two.

The theoretical findings that Blutner obtained from applying weak bidirectional OT to the interpretation of complex expressions with gradable adjectives are not convincing as they

were based on the assumptions described above. Besides, Blutner's findings have never been tested empirically and have therefore never been confirmed. We assume that an approach of weak bidirectional OT to the domain of complex emotional expressions could be very fruitful. However, we believe it would be appropriate to first test how humans interpret complex expressions. Once we have more insight in the meanings of complex expressions, we can use weak bidirectional OT to account for these empirical findings, and replace Blutner's assumptions by founded facts. Chapter 5 describes the experiment that we set up.

Chapter 5

Experiment

In chapter 4 we have seen that weak bidirectional Optimality Theory (OT) can be applied to complex expressions with gradable adjectives to account for some conversational implicatures that are assumed to arise. Weak bidirectional OT assumes that the interpretation of a (negated) gradable adjective depends on the alternative forms and meanings of the adjective. A form-meaning pair is (super)optimal, and therefore accepted in a language, if its form and meaning are the most simple ones among their available alternatives. Weak bidirectional OT thereby requires that linguistic expressions are efficient and informative. The theory formally describes how complex forms are interpreted, taking into account the markedness of meanings and forms. However, markedness is ill defined making it difficult to apply the theory. Blutner (2001) made some assumptions about the markedness of meanings and forms in order to apply weak bidirectional OT to complex expressions with gradable adjectives (see section 4.3.1). We would like to adopt the vision of weak bidirectional OT to apply it to the field of gradable adjectives but we will not simply accept Blutner's assumptions about markedness. With the experiment described below, we aim to define empirically established criteria for markedness of forms and meanings. Our second research question of this chapter, formulated below, concentrates on this issue.

Additionally, the experiment tests if the conversational implicatures that are assumed about complex expressions with gradable adjectives are correct. While we may agree, when the example is presented to us, that e.g. *not happy* is interpreted as *unhappy*, it is not at all clear that speakers interpret these forms as is suggested. Without empirical grounding we will never know what the normal interpretations of such expressions really are. We will therefore study the interpretations of such expressions empirically. Our first research question below expresses this goal.

With the experiment we aim to answer the two research questions introduced above. The first question focuses on what the normal interpretations of complex expressions with gradable adjectives are, the second question studies whether weak bidirectional OT can explain these interpretations. Answering this second question we try to establish criteria for markedness of forms and meanings.

- *How are complex expressions with gradable adjectives, which can literally refer to a wide range of intensities, interpreted by human readers?*
- *Can weak bidirectional Optimality Theory account for these conversational implicatures?*
  - *What determines the markedness of forms?*

– *What determines the markedness of meanings?*

For the experiment, we will try to adopt the approach of weak bidirectional OT to explain the interpretations. We thus assume that the proposed optimal and super-optimal form-meaning pairs, as described in section 4.2.1, are accepted in language.

In order to answer the defined research questions, we study how the emotion pairs {gelukkig (happy), ongelukkig (unhappy)} and {blij (joy), bedroefd (sad)} in combination with the morpheme *niet* (*not*) and the degree modifiers *erg* (*very*) and *een beetje* (*a bit*) are interpreted, regarding their intensity. A comparison of the results obtained from both pairs gives us insight in the effect of the negation morphemes *on-* (*un-*) (as in *ongelukkig*) and *niet*, and the modifiers *erg* and *een beetje* on the interpretation of the adjectival expression.

The experiment also investigates the interpretations of the negation of the lexical items that we identified for the emotions of the Virtual Storyteller, in order to apply our results to the language generation component of the storyteller. None of the lexical items that were selected for the storyteller contain a negation morpheme and they may therefore behave similarly to the antonym pair {blij, bedroefd}. If the results obtained from the detailed test of the {blij, bedroefd} emotion are compatible with the results of the other lexical items selected for the storyteller, the results of the {blij, bedroefd} pair will be applied to the linguistic expressions of the other items.

5.1 Method

The experiment consisted of an online test in which subjects were asked to indicate how they interpreted certain complex expressions with adjectives, verbs and nouns. For each trial, subjects had to compare two expressions on the same emotion scale and they had to indicate which of the two was more positive or negative.

5.1.1 Experimental Design

The test was conducted online in order to reach a wide range of participants. The test consisted of 61 trials, in which subjects had to compare two expressions. In each trial, a context sentence was presented and two named faces were displayed that expressed their emotions about the context sentence. The two expressions differed from each other, but they focused on the same emotion scale. The expressions were followed by a question that asked which of the two persons had the most positive (or negative) feelings. Three options were given: *person A*, *person B* or *both equally much*. Subjects had to select one of the options. An example of this is shown in figure 5.1. In the figure the English translations of the Dutch sentences are put between brackets.

5.1.2 Stimuli

A stimulus consisted of a context sentence, followed by two expressions and a question about these expressions. There were two types of questions. One asked which of the two expressions was most positive, the other asked which of the two expressions was most negative. For each stimulus the choice of the question was made randomly beforehand. Because we wanted to exclude that the choice of the question would influence the results, a second test was created. This test was similar to the first test, but it contained the opposite question for each stimulus.

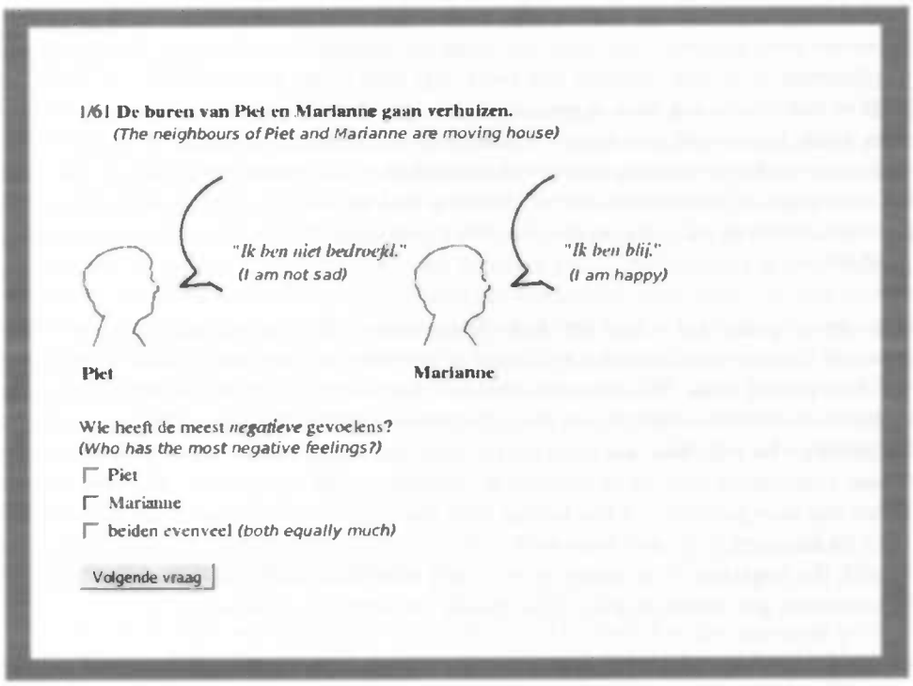


Figure 5.1: Example of a trial

Both tests had a fifty percent chance of being selected, so that for each stimulus, we would approximate an equal division over the corresponding question types.

There were three categories of stimuli. One category consisted of expressions with the antonym pairs {gelukkig, ongelukkig} and {blij, bedroefd}. For each of these antonym pairs twelve comparisons were made between different expressions. Another category consisted of expressions with the antonym pairs selected for the Virtual Storyteller. Of each antonym pair, three comparisons were made. The third category focused on the synonyms that were found for some emotion types. Six comparisons were made between obtained synonyms. All comparisons were preceded by a context sentence that was compatible with the expressions. The names of the two persons that expressed their emotions were chosen randomly.

{gelukkig, ongelukkig} and {blij, bedroefd} Of these two antonym pairs complex expressions were formed using only the negation morpheme *niet*, using the morpheme *niet* in combination with the modifier *erg* (*niet erg*) and using the modifier *een beetje*. Thus, for each of the adjectives, four expressions were possible. For each antonym pair, comparisons were made between expressions for which the difference between them was not obvious. For instance, *een beetje gelukkig* was not compared to *gelukkig* as their difference was obvious. But we did compare the expressions *niet gelukkig* and *ongelukkig*. Twelve comparisons were made for each antonym pair. Appendix C1 lists the stimuli for the {gelukkig, ongelukkig} and the {blij, bedroefd} pair.

Antonym pairs selected for the Storyteller For the eleven antonym pairs that we selected for the emotions of the Virtual Storyteller we constructed one complex expression for each lexical item. We chose to construct only one complex expression per lexical item in order to restrict the duration of the experiment. The complex expressions were constructed combining the negation morpheme *niet* with the lexical item, as in *niet trots* (*not proud*). These expressions were of most interest, because, of all morphemes, the negation *niet* would affect the interpretation of the lexical item the most. We compared the negated lexical item with its antonym (e.g. *niet trots* with *schamen* and *trots* with *niet schamen*) and we compared it with the negation of its antonym (e.g. *niet trots* with *niet schamen*). This resulted in three comparisons per antonym pair. The stimuli are listed in Appendix C2.

Synonyms The test was expanded with comparisons between obtained synonyms for the emotions of the Virtual Storyteller. The results on these comparisons would show whether the synonyms are indeed interpreted equally positive or negative. Six items that could function as a synonym were tested. This resulted in six comparisons. These stimuli can be found in Appendix C3.

5.1.3 Data Collection

The data collection during the test was automated and all selected answers of each participant were stored. Additionally, some personal information about the participants and their linguistic background was stored, namely their name, age, native language, the country where they grew up, the country and area where they lived the most of their adult life, the country where they lived at the time of participation, and the highest educational institution they attended. The form that the subjects had to fill in is shown in Appendix C4. The linguistic background was asked for in order to determine how familiar the participant was with the Dutch language.

5.1.4 Subjects

Via e-mail, the URL of the online test was sent to all AI students of the Rijksuniversiteit Groningen and to a wide range of friends and family members who were naïve with respect to the goal of the experiment. Receivers of the e-mail were asked to forward the URL to others who might be interested in taking part.

83 subjects filled in the test, among which were 35 AI students (or graduates). All subjects spoke Dutch fluently; their native language was Dutch or they had grown up in The Netherlands. The subjects had a mean age of 26 (with $\sigma = 10$), varying between 17 and 60 years of age. Nearly all subjects (81 out of 83) had attended a higher educational institution (high vocational education or university).

5.1.5 Procedure

Subjects first received a written introduction to the experiment with a description of the task. Next they had to fill in the form with the background questions. (The introduction text and the form are shown in Appendix C4). Once the form was completed, subjects started the test by pressing the start-button. The test consisted of 61 multiple choice questions, asking which of the two presented expressions was most positive or negative or whether they were equally positive or negative. The subject had to select the preferred answer by checking the corresponding checkbox with the mouse. When they clicked on 'next question', the succeeding question was displayed. Subjects were instructed not to go back to the previous question by using the back-button in their browser and revise their answer. We wanted the subjects to follow their first intuition rather than revise and reason about their answers. After the 61st question, subjects were given the opportunity to send comments about the experiment. They could write their feedback in an input field and press the submit-button to send their comments.

5.2 Results and Discussion

Three categories of data were obtained from the test, i.e. data for the antonym pairs {blij, bedroefd} and {gelukkig, ongelukkig}, data for the antonym pairs selected for the Virtual Storyteller and data for the synonyms that we obtained for the emotions in the Virtual Storyteller. We present the results and discuss them according to these three categories.

5.2.1 {blij, bedroefd} and {gelukkig, ongelukkig}

Subjects compared complex expressions to simple expressions on the same scale. Their interpretations of the complex expressions (versus the simple expressions) are shown in figure 5.2. This chart was constructed to gain insight in the overall trends of the data. The x-axis shows the simple expressions and it is thereby an intensity scale of the emotion pair (from positive to negative). The subjects' answers were converted to values (1 for 'more positive than', 0 for 'equally positive' and -1 for 'more negative than') and for each comparison, the mean value was computed. These values are projected on the y-axis. The datapoints are a measure for the difference between the complex expressions and the simple expressions on the x-axis. However, they do not reflect the diversity of the interpretations per comparison. For

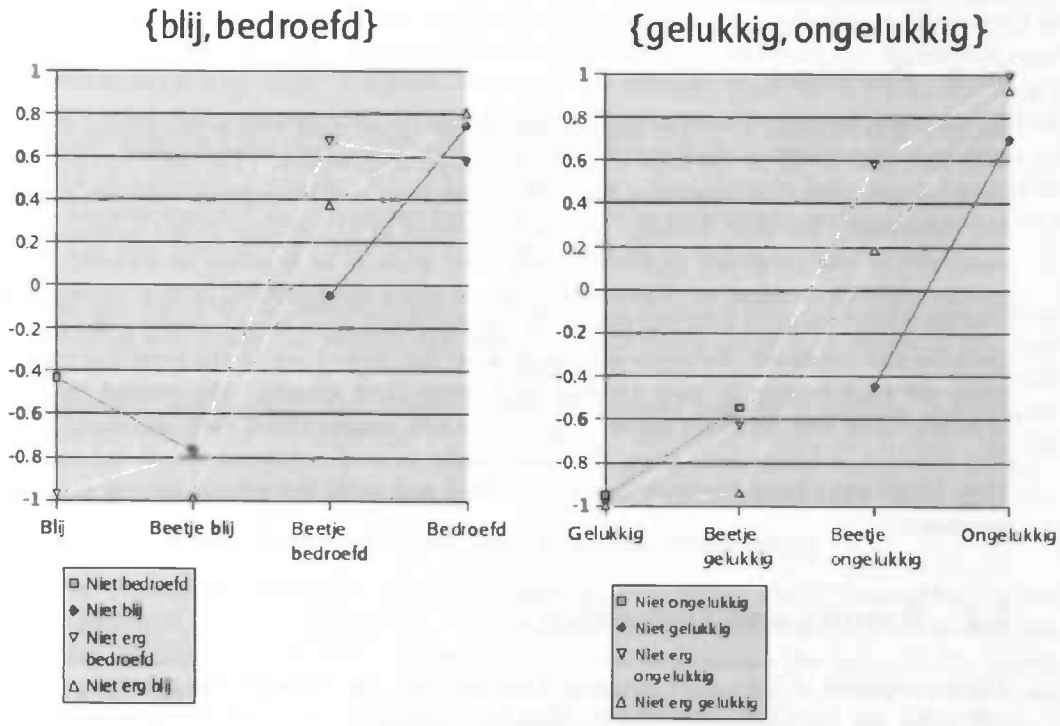


Figure 5.2: Interpretations of complex expressions versus simple expressions on a scale of -1 to 1

instance, the expression *niet erg gelukkig* was overall interpreted as more positive than *een beetje ongelukkig*, but the exact proportions of answers are not visible in the chart.

The lines in the charts in figure 5.2 should increase from left to right due to the scale that they were put on. We identified two odd datapoints. The first one represents the interpretation of *niet bedroefd* versus *blij*. The context sentence of the test clarified this odd point: it said that the neighbours of the two persons expressing their feelings were moving house (this trial was actually given as an example in figure 5.1). The subjects probably interpreted positive feelings about this event as negative (towards the neighbours) and vice versa. Hence, they may have interpreted *niet bedroefd* as more positive than *blij*. In a statistical analysis of the data we will try to replace this datapoint.

The other odd point was the *niet erg bedroefd* versus *bedroefd* interpretation, which was relatively negative. The context sentence here also referred to a negative event for some other person, namely the death of the grandmother. Therefore, subjects may have interpreted negative feelings as positive and vice versa. We will take this odd interpretation into account when analysing the data.

Furthermore, one datapoint is missing, due to an error in the experimental design. This datapoint, 'niet erg blij versus blij' is not very crucial, we assume that it would approach the -1 value.

From figure 5.2 we can roughly derive the preferred interpretations of the complex expressions. These will now be discussed.

Preferred interpretations

Results Focusing on our first research question of this chapter, we can derive some conclusions about how complex expressions are interpreted. In figure 5.2 the mean interpretations of the complex expressions were given relative to the simple expressions on the x-axis in the chart. For each complex expression we can derive its corresponding interpretation on the x-axis. If the value of the datapoint is zero, the interpretations of the associating complex expression and the associating simple expression correspond. For example, the interpretation of *niet blij* corresponds to that of *een beetje bedroefd*, because the mean interpretation of 'niet blij versus een beetje bedroefd' was zero. Table 5.1 sums up the interpretations of all complex expressions in figure 5.2.

Complex expression	Mean interpretation
niet bedroefd	more negative than *een beetje blij*
niet blij	*een beetje bedroefd*
niet erg bedroefd	between *een beetje blij* and *een beetje bedroefd*
niet erg blij	between *een beetje blij* and *een beetje bedroefd*
niet ongelukkig	more negative than *een beetje gelukkig*
niet gelukkig	between *een beetje ongelukkig* and *ongelukkig*
niet erg ongelukkig	between *een beetje gelukkig* and *een beetje ongelukkig*
niet erg gelukkig	between *een beetje gelukkig* and *een beetje ongelukkig*

Table 5.1: Mean interpretations of the complex expressions

As already mentioned above, the data in figure 5.2 represent the mean interpretations per compared expressions, but the variation of the interpretations per comparison is not visible.

Therefore, table 5.2 was constructed to show the variety of interpretations per complex expression. For each comparison, the frequencies of the three given answers and their corresponding proportions are given in the figure. We see that for some complex expressions, the subjects do not agree on what the interpretation should be. For example, the frequencies for the 'niet erg gelukkig versus een beetje ongelukkig' comparison show much diversity which indicates that it is not clear to everyone how to interpret the form *niet erg gelukkig*. The same is true for the comparisons 'niet blij versus een beetje bedroefd' and 'niet erg blij versus een beetje bedroefd'.

Discussion The interpretations of two complex forms in our test were theoretically predicted by Blutner (2001). The mean interpretations of the forms *niet gelukkig* and *niet ongelukkig* approach those assumed by Blutner. Blutner suggests that *niet gelukkig* is interpreted as *ongelukkig* and that *niet ongelukkig* is interpreted as *indifferent*. However, the diversity of some interpretations is very large. Subjects obviously have different preferences for interpreting a complex form. OT assumes fixed criteria for the markedness of forms and meanings, which indirectly determine the interpretation of a form. We would then also expect agreement on the interpretation of a complex form. Our results show that there is not always agreement. This indicates that not every complex form can have its interpretation predicted by the theory alone and that humans may have personal preferences for a certain interpretation.

We will now focus on the second research question and discuss which criteria for markedness of meanings and forms can explain the data. To that end a number of trends in the data will be discussed focusing on the difference between the two antonym pairs {blij, bedroefd} and {gelukkig, ongelukkig}, the difference between negated negative adjectives and negated positive adjectives and the difference between negated expressions with and without the use of *erg*.

Difference between {blij, bedroefd} and {gelukkig, ongelukkig}

Results For investigating the effect of a negation morpheme, like in *ongelukkig*, on the interpretation of complex expressions, we compared the data between {blij, bedroefd} and {gelukkig, ongelukkig}. We first used a matched t-test to compare the means in the charts in figure 5.2. These obtained means for the interpretations were not very exact, because we assigned the value 1 to any expression interpreted more positively than the other (and -1 to any expression interpreted more negatively than the other), whereas we were not certain about the degree of positivity or negativity. Beside that, the diversity of the interpretations was not reflected by the data in the chart. The matched t-test showed that seven out of eleven datapoints differed significantly, but we aimed to identify only very deviating datapoints. Due to this undesirable outcome and the shortcomings of the method, described above, we used a different statistical test.

We used table 5.2 to compare the proportions between the expressions of {blij, bedroefd} in the leftmost column with the expressions of {gelukkig, ongelukkig} in the rightmost column. We only compared the positive and the negative proportions for every two items. The neutral answers ('two expressions interpreted equally positive') were not compared, but they were included in total number of answers for each item. A two-tailed test for the difference between independent proportions at the 0.05 level of significance was used to compare the proportions.[1]

[1] The two-tailed test for the difference between independent proportions requires that the proportions lie

Expression	Ans	Frequency	Frequency	Ans	Expression
Niet blij versus bedroefd (N = 83)	+	65 (78 %)	64 (77%)	+	Niet gelukkig versus ongelukkig (N = 83)
	0	15 (18 %)	13 (16%)	0	
	-	3 (4 %)	6 (7%)	-	
Niet blij versus een beetje bedroefd (N = 83)	+	35 (42%) *	14 (17%) *	+	Niet gelukkig versus een beetje ongelukkig (N = 83)
	0	9 (11%)	18 (22%)	0	
	-	39 (47%)	51 (61%)	-	
Niet erg blij versus bedroefd (N = 46)	+	39 (85%)	78 (94%)	+	Niet erg gelukkig versus ongelukkig (N = 83)
	0	5 (11%)	4 (5%)	0	
	-	2 (4%)	1 (1%)	-	
Niet erg blij versus een beetje bedroefd (N = 83)	+	48 (58%)	36 (43%)	+	Niet erg gelukkig versus een beetje ongelukkig (N = 83)
	0	18 (22%)	26 (31%)	0	
	-	17 (20%)	21 (25%)	-	
Niet erg blij versus een beetje blij (N = 166)	+	2 (1%)	0 (0%)	+	Niet erg gelukkig versus een beetje gelukkig (N = 83)
	0	6 (4%)	5 (6%)	0	
	-	158 (95%)	78 (94%)	-	
Niet erg blij versus blij (N = 0)	+		0 (0%)	+	Niet erg gelukkig versus gelukkig (N = 83)
	0		0 (0%)	0	
	-		83 (100%)	-	
Niet bedroefd versus blij (N = 83)	+	22 (27%) *	0 (0%) *	+	Niet ongelukkig versus gelukkig (N = 83)
	0	3 (4 %)	4 (5%)	0	
	-	58 (70%) *	79 (95%) *	-	
Niet bedroefd versus een beetje blij (N = 83)	+	5 (6%)	12 (14%)	+	Niet ongelukkig versus een beetje gelukkig (N = 83)
	0	9 (11%)	14 (17%)	0	
	-	69 (83%) *	57 (69%) *	-	
Niet erg bedroefd versus bedroefd (N = 83)	+	64 (77%) *	82 (99%) *	+	Niet erg ongelukkig versus ongelukkig (N = 83)
	0	3 (4 %)	1 (1%)	0	
	-	16 (19%) *	0 (0%) *	-	
Niet erg bedroefd versus een beetje bedroefd (N = 83)	+	61 (73%)	60 (72%)	+	Niet erg ongelukkig versus een beetje ongelukkig (N = 83
	0	17 (20%)	15 (18%)	0	
	-	5 (6%)	8 (10%)	-	
Niet erg bedroefd versus een beetje blij (N = 83)	+	4 (5%)	8 (10%)	+	Niet erg ongelukkig versus een beetje gelukkig (N = 83)
	0	10 (12%)	15 (18%)	0	
	-	69 (83%)	60 (72%)	-	
Niet erg bedroefd versus blij (N = 83)	+	0 (0%)	0 (0%)	+	Niet erg ongelukkig versus gelukkig (N = 83)
	0	2 (2%)	1 (1%)	0	
	-	81 (98%)	82 (99%)	-	

* *significant difference*

Table 5.2: Frequencies (and proportions) of subjects' responses on comparisons between complex expressions and simple expressions

The proportions that differed significantly are indicated by an '*' in table 5.2. Six statistically differing proportions were identified, distributed over four compared expressions.

Niet blij differed from *niet gelukkig* relative to *een beetje bedroefd* and *een beetje ongelukkig* respectively (P = 0.0004). *Niet blij* was interpreted more positively than *niet gelukkig*.

The next three differing proportions concern the comparison of *niet bedroefd* and *niet ongelukkig*. The first two concern the already identified incorrect data for the item 'niet bedroefd versus blij'.[2] *Niet bedroefd* was interpreted rather positively and this is reflected by the proportions. The correct data for this item would approximate the data for the 'niet erg bedroefd versus blij' item, because the data for 'niet bedroefd versus een beetje blij' and 'niet erg bedroefd versus een beetje blij' also nearly correspond. If we substitute the 'niet erg bedroefd versus blij' data for 'niet bedroefd versus blij', then there is no significant difference between 'niet bedroefd versus blij' and 'niet ongelukkig versus gelukkig'.

For the negative proportions for 'niet bedroefd versus een beetje blij' and 'niet ongelukkig versus een beetje gelukkig' a difference was identified (P = 0.0294). *Niet bedroefd* was interpreted more negatively than *niet ongelukkig*.

Finally, two significant differences were found between the proportions for 'niet erg bedroefd versus bedroefd' and 'niet erg ongelukkig versus ongelukkig' ($P < 0.002$ for both the positive proportions and the negative proportions). These proportions concern the other odd datapoint in figure 5.2.[3] Because the proportions for the other three items do not differ significantly, we assume that the difference for these proportions are due to the context sentence of the item. The interpretations for *niet erg bedroefd* and *niet erg ongelukkig* are therefore assumed to be similar.

Discussion From a comparison of proportions, we found that *niet gelukkig* was interpreted more negatively than *niet blij* and *niet ongelukkig* was interpreted more positively than *niet bedroefd*. Both differences stem from only one significant difference in proportions, but the other proportions for these expressions support these data although not significantly.

The difference between the interpretations of *niet blij* and *niet gelukkig* can be ascribed to the complexity of the antonym forms: *bedroefd* contains no negation morpheme, but *ongelukkig* contains the negation morpheme *on-*. Because these forms function as alternatives for *niet blij* and *niet gelukkig*, they may contribute to their interpretation. *Niet gelukkig* and *ongelukkig* both contain one negation morpheme. If both forms are equally complex, their interpretations could correspond. The right chart in figure 5.2 illustrates that *ongelukkig* is interpreted as more negative than *niet gelukkig*. Speakers probably prefer using *ongelukkig*

between 5% and 95%. If they don't, the risk is that two distributions are not identified as significantly different, while in fact they would be different in different circumstances. We realize that some of our compared proportions do not meet the requirement of the two-tailed test, but this is not a problem because our results focus on the proportions that are identified as significantly different rather than on those that are not. For the same complex expressions, we compare different proportions so that a difference in interpretation (if there is one) will probably be obtained on at least one of the comparisons.

[2] As already mentioned in this section, this datapoint was quite positive in relation to the other datapoint concerning the form *niet bedroefd*. The context sentence said that the neighbours of the two persons expressing their feelings were moving house. The subjects probably interpreted positive feelings about this event as negative (towards the neighbours) and vice versa. This may have had an effect on their responses.

[3] As already mentioned in this section, this datapoint was quite negative in relation to the other datapoints concerning the form *niet erg bedroefd*. The context sentence refers to a negative event for some other person, namely the death of the grandmother. Therefore, subjects may have interpreted negative feelings about the event as positive and vice versa, which may have had an effect on their responses.

rather than *niet gelukkig*. When *niet gelukkig* is used, the speaker means something else than *ongelukkig*. However, the difference in interpretations between *niet gelukkig* and *ongelukkig* is rather small. The interpretation for *niet gelukkig* approaches that of *ongelukkig* just like its complexity approaches that of *ongelukkig*. For the antonym pair {blij, bedroefd}, *niet blij* is obviously more complex than *bedroefd* and its interpretation may therefore deviate more from the interpretation of *bedroefd*.

Why was *niet ongelukkig* interpreted as more positive than *niet bedroefd*? *Niet ongelukkig* is a rather complex form, as it contains two negation morphemes, so we would expect its interpretation to deviate quite much from the interpretation of *gelukkig*. Maybe, the combination of two negation morphemes in *niet ongelukkig* triggers the reader to interpret the expression as rather positive. The two negation morphemes may result in a double negative, so that readers have the tendency to interpret this form as quite positive. Due to the form's complexity its interpretation is less positive than that of *gelukkig*.

Difference between negated negatives (litotes) and negated positives

Results The proportions in table 5.2 were used to analyse the effect of *niet* on the interpretation of the adjective. A two-tailed t-test with the 0.05 level of significance was used to test whether the effect of *niet* on a positive adjective was different from the effect of *niet* on a negative adjective. To that end, we compared the positive proportion of a negated positive form with the negative proportion of the corresponding negated negative form (or vice versa, depending on which proportion was highest). Table 5.3 shows the results.

Negated positive adjectives	Negated negative adjectives	Positive % – Negative %	Negative % - Positive %	P-value
Niet blij versus bedroefd	Niet bedroefd versus blij	8%**		.215
Niet blij versus een beetje bedroefd	Niet bedroefd versus een beetje blij	-41%*		<0.0002
Niet erg blij versus blij	Niet erg bedroefd versus bedroefd		*missing*	
Niet erg blij versus een beetje blij	Niet erg bedroefd versus een beetje bedroefd		22%*	<0.0002
Niet erg blij versus een beetje bedroefd	Niet erg bedroefd versus een beetje blij	-25%*		0.0004
Niet erg blij versus bedroefd	Niet erg bedroefd versus blij	-13%**		0.0062
Niet gelukkig versus ongelukkig	Niet ongelukkig versus gelukkig	-18%*		0.0008
Niet gelukkig versus een beetje ongelukkig	Niet ongelukkig versus een beetje gelukkig	-52%*		<0.0002
Niet erg gelukkig versus gelukkig	Niet erg ongelukkig versus ongelukkig		1%	0.32
Niet erg gelukkig versus een beetje gelukkig	Niet erg ongelukkig versus een beetje ongelukkig		25%*	<0.0002
Niet erg gelukkig versus een beetje ongelukkig	Niet erg ongelukkig versus een beetje gelukkig	-29%*		<0.0002
Niet erg gelukkig versus ongelukkig	Niet erg ongelukkig versus gelukkig	-5%		0.1

* *significant difference*
** *odd datapoint*

Table 5.3: Difference in proportions for negated positives and negated negatives

Leaving out the two odd datapoints, we see that for all compared expressions, the negated positives were interpreted as more negative than the corresponding negated negatives were interpreted as positive. Seven out of nine proportions differed at the 0.05 level of significance.

Discussion The interpretations of a negated positive adjective and a negated negative adjective were not symmetric. The interpretation of a negated positive was more negative than the interpretation of a negated negative was positive. Negated positives thereby differed more

from the original positive forms. The negation *niet* thus appeared to have more effect on a positive form than on a negative form. This effect was identified for both antonym pairs {blij, bedroefd} and {gelukkig, ongelukkig}.

Changing the polarity of a negative adjective (by adding *niet*) is harder than changing the polarity of a positive adjective. The use of a negative adjective seems to bring about a negative interpretation which is hard to cancel (or deny). When a positive adjective is combined with the negation morpheme *niet*, the reader has no problems interpreting this as negative. The meaning or polarity of the adjective obviously plays a role in the interpretation of the expression as a whole whereas logically speaking, the interpretations of the two negated forms would be symmetric.

If the interpretations of the expressions cannot be explained by their logical meaning only, Optimality Theory uses the complexity of meanings and forms to account for the interpretations. In this case, our intuition is that the complexity of the forms plays no role, because the effect was observed for both the antonym pairs {blij, bedroefd} and {gelukkig, ongelukkig} and seems to be independent of the form of the adjective. The complexity of the meaning, in terms of the intensity of the meaning, would then explain the difference in interpretation. With regard to these observations, we would conclude that readers are tempted to interpret a negated form as negative. The complexity of a meaning is then determined by its polarity, negative meanings are less complex than positive meanings.[4] In the Integrated Discussion we try to fit this provisional conclusion into the other assumptions about complexity of forms and meanings.

However, we think that there is also an alternative explanation for the effect of *niet* on the interpretation of an adjective that could work. This explanation lies in the complexity of the forms. For both antonym pairs, the negative adjective can be regarded as more complex than the positive adjective. So, *ongelukkig* is more complex than *gelukkig* in terms of the number of the negation morphemes and *bedroefd* is more complex than *blij* in terms of the number of morphological elements. When a negation is added to the positive adjective, the difference in complexity with the antonym form is much smaller (as in *niet blij* compared to *bedroefd*) than when a negation is added to the negative adjective (as in *niet bedroefd* compared to *blij*). According to weak bidirectional OT, the difference in complexity of the forms should be reflected by the difference in the meanings. So, we would expect a relatively small difference in meaning between the forms *niet blij* and *bedroefd* and a relatively large difference in meaning between the forms *niet bedroefd* and *blij*. This was indeed observed.

Beside our intuition that the first explanation, which takes into account the complexity of meanings, is a more plausible explanation for the effect of *niet*, there is an additional reason to suggest that this explanation is the right one. The effect of *niet* was observed for all complex forms, including the adjectives preceded by *niet erg*. We think that for these complex forms the effect of the complexity of the adjective should be rather low. When comparing the forms *niet erg blij* and *niet erg bedroefd*, we think that the contribution of the extra morphological

[4]This preference for a negative meaning can be caused by a speaker's preference to use a positive form. Positive forms may generally be less marked than negative forms. As a consequence, if a speaker does use a negative form combined with *niet* the hearer is tempted to interpret the expression as quite negative to explain the speaker's choice of a negative form. On the other hand, if a speaker uses a positive form combined with *niet*, the meaning may approach the meaning of the negative form because the complexity of the expression also approaches the complexity of the negative form. The general preference for the use of positive forms can then be seen as a way to nuance negative expressions. An expression like *I am unhappy* sounds very depressed and a speaker probably does not prefer such negative expressions. The hearer then accounts for this preference of the speaker and tries to find a negative interpretation where possible.

element in *bedroefd* to the complexity of the complete form *niet erg bedroefd* is relatively low. We consider it unlikely that this extra element would result in such different interpretations of the complex forms *niet erg blij* and *niet erg bedroefd*. Therefore, the complexity of the meanings is more likely to result in different interpretations.

Differences between negated expressions with and without *erg*

Results The proportions in table 5.2 were used to compute the difference between expressions in which *erg* was used and in which *erg* was not used. A two-tailed t-test computed the difference between the positive proportions (for negative adjectives) and negative proportions (for positive adjectives). The differences are displayed in table 5.4.

Expressions with "erg"	Expressions without "erg"	Difference in positive %	Difference in negative %	P-value
Niet erg blij versus bedroefd	Niet blij versus bedroefd	6% *		0.37
Niet erg blij versus een beetje bedroefd	Niet blij versus een beetje bedroefd	16% *		0.0436
Niet erg bedroefd versus blij	Niet bedroefd versus blij		28% **	<0.0002
Niet erg bedroefd versus een beetje blij	Niet bedroefd versus een beetje blij		0%	1
Niet erg gelukkig versus ongelukkig	Niet gelukkig versus ongelukkig	17% *		0.002
Niet erg gelukkig versus een beetje ongelukkig	Niet gelukkig versus een beetje ongelukkig	27% *		0.0002
Niet erg ongelukkig versus gelukkig	Niet ongelukkig versus gelukkig		4%	0.1732
Niet erg ongelukkig versus een beetje gelukkig	Niet ongelukkig versus een beetje gelukkig		4%	0.6101

* *significant difference*
** *odd datapoint*

Table 5.4: Difference in proportions of forms with and without *erg*

For negated positive adjectives, *erg* had a positive effect on the interpretation and for negated negative adjectives, *erg* had a negative effect on the interpretation. Leaving out the odd datapoint, we found that for the positive adjectives three out of four differences were significant (with the 0.05 level of significance), and one of the differences for negative adjectives were significant.

Discussion The use of *erg* in a negated adjective nuanced the negation; negated positive adjectives became less negative and negated negative adjectives became less positive. However, for the positive adjectives only, this effect was significant.

The nuancing effect of *erg* can be explained logically: where e.g. *niet blij* logically negates the whole positive range of the adjective *blij*, *niet erg blij* only negates *erg blij* and leaves everything between *blij* and *erg bedroefd* as potential interpretations. Thus, by adding *erg* to a negated adjective, the negated range is narrowed and a wider range remains for being interpreted by the reader.

For positive adjectives only, the observed differences were significant. The explanation can be found in the difference between negated negatives and negated positives. For instance, *niet blij* was interpreted as rather negative. When *erg* is added to this negated expression, its interpretation is nuanced and becomes significantly more positive. On the other hand, *niet bedroefd* was already interpreted as rather negative. Nuancing this expression by adding *erg* makes no significant difference. *Niet bedroefd* leaves little room for a nuancing effect by adding *erg*.

Integrated Discussion

We have described a number of features in the data of the {blij, bedroefd} and {gelukkig, ongelukkig} antonym pairs. We will now try to derive some more general explanations for the interpretation of complex expressions with gradable adjectives. We realise that these conclusions are based only on the interpretations of the two antonym pairs given above and that the context sentences used in the experiment may have influenced the data for these antonym pairs. It would therefore be wise to conduct an experiment like this again with more variation in the antonym pairs and context sentences. However, for now, we propose some provisional conclusions that are supported by the data of this experiment.

One general observation was that at least for some complex forms there was no consensus about what the interpretation was. Blutner's theory suggests that at least for many of the comparisons we should expect agreement on the difference in interpretation if the interpretation really is systematically related to the markedness of forms and meanings and the weak bidirectional reasoning of OT. The fact that there are some large and unexpected variations in the subjects' interpretations shows that the theory is only a rough estimate of what the real interpretations of complex forms are.

The fact that *niet blij* was interpreted less negatively than *niet gelukkig* indicates that the alternative forms *bedroefd* and *ongelukkig* influence the interpretation. This is in correspondence to Blutner's weak bidirectional acceptation of OT. Blutner assumes that the number of negation morphemes determines the complexity of the form, and he would expect the forms *niet gelukkig* and *ongelukkig* to be interpreted equally negative. Our results do not support his assumption; *niet gelukkig* was interpreted less negatively than *ongelukkig*. We would conclude that the form *niet blij* is more complex than *bedroefd* and that *niet gelukkig* is more complex than *ongelukkig*, but that the difference between the latter two is rather small. This difference can be explained if both the number of negation morphemes *and* the number of words determine the complexity of the expression. Then *niet blij* is definitely more complex than *bedroefd* in that it contains an extra word and an extra negation morpheme. The difference between *niet gelukkig* and *ongelukkig* is smaller; the first expression contains two words and a negation morpheme and the second expression contains one word and a negation morpheme.

We can also say something about the complexity of meanings. The complex forms above (all containing an adjective preceded by the negation *niet*) could logically refer to any intensity between neutral and the opposite end of the scale. The fact that the expressions were all interpreted as negatively as possible (taking into account alternative form-meaning pairs) indicates that there was a preference to try to get to the negative end of the scale.[5] Blutner would explain these data by determining that the complexity of meanings increases from the negative end of the scale to the positive end of the scale. We realise that these preferences for meanings may be attached to these specific data in which negations are used in expressions. The use of a negation may trigger a negative interpretation by the reader. However, because we want to verify whether Optimality Theory can explain the data, we try to derive general assumptions about forms and meanings and stick with the assumption that the complexity of meanings increases from negative to positive.

We saw that *niet erg gelukkig* and *niet erg blij* were interpreted towards neutral. Logically, these expressions can refer to any intensity on the scale between *gelukkig/blij* and *erg ongelukkig / erg bedroefd*. The assumed complexity for meanings says that they should be

[5] As discussed, negated positives were interpreted rather negatively, negated negatives were interpreted close to neutral

interpreted as negative as possible. Because the negative meanings can be grasped more economically by alternative forms, the neutral meanings appear to be most negative. The same argument can be given for the interpretations of *niet erg ongelukkig* and *niet erg blij*, which were also interpreted close to neutral.

With the above assumptions about complexity of meanings and forms there is one trend in the data, concerning double negatives, that cannot be explained. There was a difference in interpretation between *niet ongelukkig* and *niet bedroefd* with the first expression being interpreted less negatively. Using the assumptions for complexity of forms, the difference between *niet ongelukkig* and *gelukkig* is bigger than the difference between *niet bedroefd* and *blij*. We would therefore expect the meaning of *niet ongelukkig* to be more distinct from *gelukkig* and therefore rather negative. The experimental data showed that the opposite was true. We proposed that a double negative, like *niet ongelukkig* may have triggered some positive interpretation, but an explanation like this cannot be given by Optimality Theory, only taking into account the complexity of forms and meanings. Although OT cannot account for the small difference in interpretation between *niet ongelukkig* and *niet bedroefd*, we have seen that it can at least explain why both expressions are interpreted as close to neutral.

5.2.2 Antonym pairs selected for the Virtual Storyteller

Results The interpretations of the comparisons made for the antonym pairs selected for the Virtual Storyteller are given in table 5.5 (with N = 83, unless stated otherwise). The odd datapoints for the {blij, bedroefd} pair were replaced (indicated by an '*') in order to compare the {blij, bedroefd} pair to the other antonym pairs. The frequencies for 'niet bedroefd versus blij' were substituted by the frequencies for 'niet erg bedroefd versus blij'. This substitution was already justified because the frequencies for 'niet bedroefd versus een beetje blij' nearly correspond to the frequencies for 'niet erg bedroefd versus een beetje blij'.

For each comparison and each antonym pair the distribution over the proportions was compared to the distribution of the {blij, bedroefd} pair over the proportions, using a two-sided Pearson Chi-Square test with weighted frequencies. Table 5.6 shows the obtained P-values. The distributions with P-values below 0.05 significantly differed from the distributions for the {blij, bedroefd} pair. This concerns seven out of ten antonym pairs. We will discuss the distinct distributions one by one.

The antonym pair {blij voor, jaloers op} was interpreted more negatively than {blij, bedroefd} on the first comparison, 'niet blij voor versus jaloers op'. For 'niet blij voor versus jaloers op' the number of negative interpretations was even higher than the number of positive interpretations. This indicates that *jaloers op* is not interpreted as a scalar opposite of *blij voor*, but more positively than that.

The antonym pair {leedvermaak, medelijden} (with *leedvermaak* as the positive emotion and *medelijden* as the negative emotion) differed on two distributions from the antonym pair {blij, bedroefd}. *Geen leedvermaak* was interpreted rather negatively in relation to *medelijden* and *geen medelijden* was interpreted rather positively in relation to *leedvermaak*. These results could suggest that the difference between *leedvermaak* and *medelijden* is quite small. However, the third distribution does not significantly differ from the {blij, bedroefd} distribution and it does not support this hypothesis.

The results for {hoopvol, bang} are similar to those of {leedvermaak, medelijden}. Here, the third distribution also deviates from the {blij, bedroefd} distribution. All three distri-

		comparison		
		niet + versus -	+ versus niet -	niet - versus niet +
antonym pair	interpretation	*e.g. niet blij versus bedroefd*	*e.g. blij versus niet bedroefd*	*e.g. niet bedroefd versus niet blij*
blij – bedroefd	positive	65	81 *	70
	equal	15	2 *	8
	negative	3	0 *	5
blij voor – jaloers op	positive	25	78	26 (N = 30)
	equal	4	4	4 (N = 30)
	negative	54	1	0 (N = 30)
leedvermaak over – medelijden met	positive	42	68	63
	equal	14	11	12
	negative	27	4	8
hoopvol – bang	positive	58	57	57
	equal	9	13	15
	negative	16	13	11
verlangen naar – bang voor	positive	79	83	59
	equal	2	0	14
	negative	2	0	10
trots – schamen	positive	63	75	66
	equal	18	7	9
	negative	2	1	8
bewonderen – minachten	positive	79	81	40
	equal	2	2	15
	negative	2	0	28
houden van – haten	positive	75	81	58
	equal	5	2	12
	negative	3	0	13
mooi vinden – afstotelijk vinden	positive	81	78	51
	equal	1	3	9
	negative	1	2	23
verliefd – afkerig	positive	80	76	39
	equal	1	6	14
	negative	2	1	30
dankbaar – boos	positive	76	80	67
	equal	6	3	9
	negative	1	0	7

* *substituted datapoint*

Table 5.5: Frequencies of the possible interpretations for the three comparisons made for each antonym pair

antonym pair	comparison		
	niet + versus -	+ versus *niet* -	*niet* - versus *niet* +
{blij voor, jaloers op}	P = 0 *	P = 0.422	P = 0.349
{leedvermaak over, medelijden met}	P = 0 *	P 0.003 *	P = 0.394
{hoopvol, bang}	P = 0.005 *	P = 0 *	P = 0.058
{verlangen naar, bang voor}	P = 0.003 *	P = 0.155	P = 0.12
{trots, schamen}	P = 0.777	P = 0.135	P = 0.648
{bewonderen, minachten}	P = 0.003 *	P = 1	P = 0 *
{houden van, haten}	P = 0.057	P = 1	P = 0.065
{mooi vinden, afstotelijk vinden}	P = 0.001 *	P = 0.324	P = 0.001 *
{verliefd, afkerig}	P = 0.001 *	P = 0.206	P = 0 *
{dankbaar, boos}	P = 0.057	P = 0.65	P = 0.795

* *significant difference*

Table 5.6: P-values for distributions of antonym pairs compared to distributions of {blij, bedroefd}

butions indicate that the difference between *hoopvol* and *bang* is smaller than the difference between *blij* and *bedroefd*.

The other distinct distributions concern the four antonym pairs {verlangen naar, bang voor}, {bewonderen, minachten}, {mooi vinden, afstotelijk vinden} and {verliefd op, afkerig van}. The deviating distributions of these three antonym pairs correspond (except for the {verlangen, bang voor} pair for which the third comparison did not differ statistically). The distributions indicate that the difference between the two antonym expressions is very large. The subjects agreed that the negation of the one expression still differs from the antonym expression. And when both negated expressions were compared, the difference between them was less obvious than the difference between *niet blij* and *niet bedroefd*. Both distributions support the hypothesis that the two antonym expressions refer to the two ends of a scale and that the difference between them is large.

Using Multidimensional Scaling, the proximities between the antonym pairs are geometrically displayed in figure 5.3, based on the positive frequencies of the three comparisons made for each antonym pair. The proximities in the space represent normalized Euclidean distances. Each antonym pair is labelled by the positive expression only in the figure due to space restrictions. The antonym pairs for which the distributions do not differ from the {blij, bedroefd} pair are surrounded by a circle (including the {blij, bedroefd} pair).

We see that the pairs in the bottom left corner are the pairs for which the difference between the antonyms is supposed to be quite large. Furthermore, we see that the difference between {blij, bedroefd} and the other deviating pairs is very large. This is in correspondence with the deviating distributions.

Discussion The adjective *jaloers* was obviously interpreted as more positive than the opposite of *blij voor*. In Dutch *jaloers* can express a negative feeling about the fortunes of others, but it can also express a desire to own what others own. In that sense, *jaloers* can have two meanings, the first reflecting a negative feeling and the second reflecting a more neutral feeling. The negative feeling is the antonym emotion of *blij voor*. The data showed that the second,

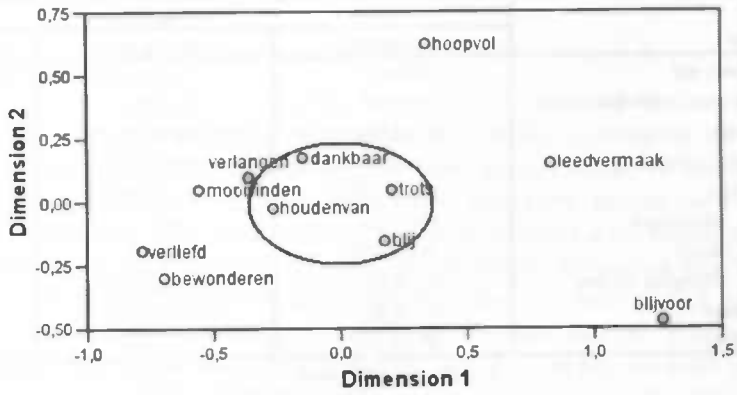


Figure 5.3: Proximities between antonym pairs, based on their three positive frequencies

neutral meaning, was the preferred interpretation for *jaloers* so the use of the adjective *jaloers* would not be appropriate for the Virtual Storyteller, unless it will be clear from the context what the interpretation of *jaloers* should be. In order to exclude any misinterpretations of the emotional expressions, we will use a different, less frequent, adjective, *afgunstig op* with the literal meaning 'having negative feelings about the fortunes of someone else'.

The other six deviating antonym pairs all showed that the conversational implicatures predicted by the {blij, bedroefd} pair may not arise to them due to different intensity scales. Figure 5.4 shows the difference between the antonym pairs {verliefd, afkerig} and {blij, bedroefd}. Just like for the antonym pairs {bewonderen, verachten}, {mooi vinden, afstotelijk vinden} and {verlangen naar, bang voor}, the difference between the two emotions is very large. The negation of one of the two emotions leaves a large neutral space as possible interpretation and this affects the interpretation of the reader. For the emotion pairs {leedvermaak, medelijden} and {hoopvol, bang} the difference between the emotions may be smaller or the emotions may even overlap, resulting in different conversational implicatures.

For these six deviating antonym pairs we will stick to the choice of the selected adjectives in the Virtual Storyteller, but we will not construct complex expressions like we do for the {blij, bedroefd} emotion. These complex expressions could be interpreted differently by the reader.

5.2.3 Obtained synonyms

Results We have tested whether the obtained synonyms were interpreted as equally positive, so that we could use them as synonyms for the Virtual Storyteller. The results of comparisons between the synonyms are displayed in table 5.7, with N = 83 for all synonym pairs. For each compared synonym pair the distribution over the answers is displayed. The last column shows the sum of the given answers (with 1 taken for each positive answer, 0 taken for each equal answer and -1 taken for each negative answer. The sum gives a measure for the difference between the synonyms. For each synonym pair, we tested whether a significant majority of the subjects agreed on the difference between the adjectives (being either positive, equal or negative), using a two-tailed test for the difference between two independent proportions

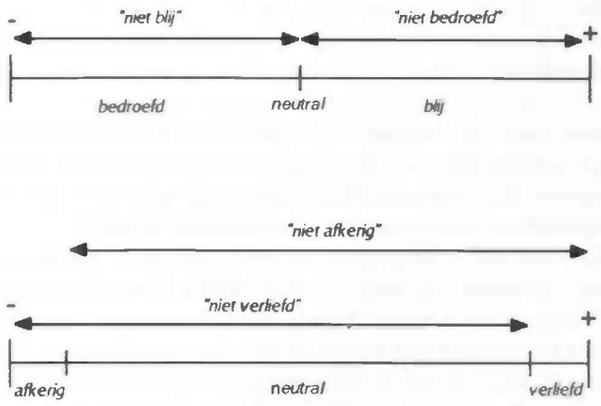


Figure 5.4: Difference between {blij, bedroefd} and {verliefd, afkerig}

at the 0.05 level of significance. Subjects agreed on the interpretation if one of the three answers had obtained a significantly higher proportion than the sum of the other two answer's proportions. We found that for only one synonym pair, {boos, kwaad}, the subjects agreed on the difference in interpretation, with *kwaad* being more negative than *boos* ($P < 0.002$).

	frequencies			
compared synonyms	positive	equal	negative	sum
bang versus angstig	23	24	36	-13
bedroefd versus verdrietig	37	33	13	24
bedroefd versus treurig	17	40	26	-9
treurig versus weemoedig	12	26	45	-33
boos versus kwaad	65	16	2	63
minachten versus verachten	39	28	16	23

Table 5.7: Difference between obtained synonyms

Discussion There was little consensus about the difference between the obtained synonyms. Not a single synonym pair was interpreted by a majority of the subjects as equal. Nevertheless, we would like to add the synonyms to the lexicon of the Virtual Storyteller as to allow more variation in the expression of emotions. The right-most column is used to derive an overall measure for the difference between the antonyms. We decide to use *angstig* and *treurig* as synonyms for *bang* and *bedroefd* respectively, because both corresponding sums were close to zero. The other synonyms will be used to express intensities that are in correspondence with their obtained sums. Chapter 7 describes in more detail how the synonyms are used in the

Narrator Agent.

5.3 Conclusion

From comparisons between complex expressions and simple expressions on the same scale ({blij, bedroefd} and {gelukkig, ongelukkig}), we have gained insight in the interpretation of complex expressions. We have seen that, for negated adjectives, people have a preference for a negative interpretation. This preferred interpretation is called a conversational implicature. For negated adjectives, the polarity of the adjective probably determines which conversational implicature arises: negated negatives are interpreted close to neutral and negated positives are interpreted between the antonym (negative) adjective and neutral. Furthermore, we have seen that complex expressions like *niet erg bedroefd* and *niet erg blij* are both interpreted close to neutral, with *niet erg bedroefd* being more positive than *niet erg blij*. From a comparison of the data for {blij, bedroefd} and {gelukkig, ongelukkig}, we found that *niet gelukkig* is interpreted rather negative, and *niet ongelukkig* is interpreted rather positive.

The antonym pairs that were selected for the Virtual Storyteller show differing scales. The interpretation of the complex expressions with the pairs {trots, schamen}, {houden van, haten} and {dankbaar, boos} correspond to those of the antonym pair {blij, bedroefd}. For these antonym pairs, the Virtual Storyteller can construct complex expressions for referring to the points on the scale that correspond to the conversational implicatures described above. For the deviating antonym pairs, the Virtual Storyteller will not construct complex expressions, because these may be interpreted differently. The scales of these deviating pairs can differ in two ways: the difference between the two lexical items is too big or too small. In the first case, the two lexical items only refer to the two ends of the scale and not to the whole positive or negative spectrum, which leaves a big neutral area between them, like in {verliefd, afkerig}, {bewonderen, verachten}, {mooi vinden, afstotelijk vinden} and {verlangen naar, bang voor}. In the other case, the positive and negative parts of the scale are very close to each other and may even overlap, like in {hoopvol, bang} and {leedvermaak, medelijden}.

Weak bidirectional Optimality Theory describes which form-meaning pairs are (super-)optimal and thereby accepted in a language, taking into account the complexity of the form, the complexity of the meaning and the logical meaning of the form. For complex expressions with gradable adjectives, the complexity of the form differs between antonym pairs (depending on the number of negation morphemes), the complexity of the meaning is similar to all antonym pairs (depending on the polarity of the meaning) and the logical meaning may differ between antonym pairs (depending on what scale is referred to by the antonym pair). The obtained differences between the pairs {blij, bedroefd} and {gelukkig, ongelukkig} can be ascribed to the complexity of the forms. The obtained differences between {blij, bedroefd} and other antonym pairs selected for the Virtual Storyteller can be ascribed to the logical meaning of the complex expressions, due to distinct emotion scales.

The current experiment contained an expanded test for only the antonym pairs {blij, bedroefd} and {gelukkig, ongelukkig}. We have seen that weak bidirectional Optimality Theory can mainly account for the conversational implicatures that arise from complex expressions with these antonym pairs. The complexity of meanings then increases from negative to positive and the complexity of forms is determined by the number of words and the number of negation morphemes. Applying these criteria to the complex expressions used in the experiment, we obtain interpretations that are close to the interpretations given by the subjects. Besides,

the difference between *niet blij* and *niet gelukkig* can be explained, because the alternative forms have different complexities. However, explaining the obtained difference between *niet ongelukkig* and *niet bedroefd* is not possible with this OT approach. Summarizing, the overall trends in the data can be explained, but not all individual datapoints.

Now that we have fitted the criteria for the complexity of meanings and forms to the data obtained for the antonym pairs {blij, bedroefd} and {gelukkig, ongelukkig}, we want to verify whether the interpretations of complex expressions with other antonym pairs can also be explained. Thus, is weak bidirectional Optimality Theory appropriate for explaining conversational implicatures concerning gradable adjectives in general? Can all differences in interpretations be ascribed to the complexity of forms, meanings and the logical meanings? In order to verify the appropriateness of Optimality Theory to explain conversational implications, and the correctness of the selected criteria for markedness of forms and meanings, we would recommend to conduct an experiment like this again, with an expanded test for varying antonym pairs.

Chapter 6

Contrast Relations

In natural language, produced stories consist of a series of coherent text elements. The text elements are somehow related to each other and all related text elements together create the discourse structure of the story. Using related sentences is what distinguishes a story from an arbitrary collection of separate sentences. In order to produce one coherent text passage, the speaker or writer must make sure that relations hold between the parts of the text and that the total of relations is consistent. Moreover, the reader or listener should be able to recognize these relations.

There are a number of coherence relations that can link together clauses of a text passage. Many researches have attempted to list these relations, but they have not reached agreement on the number of relations that exist. Their lists vary in length, containing between seven and thirty relations (e.g. Mann & Thompson (1988), Hobbs (1979)). Examples of relations are temporal relations, causal relations and contrast relations. Sentences 6.1, 6.2 and 6.3 illustrate these kinds of relations.

(6.1) Peter bought a book, while Mary bought a shirt. (*temporal*)

(6.2) Kate stayed at home, because her daughter was ill. (*causal*)

(6.3) Brian passed his exam, but Bill failed his exam. (*contrast*)

The stories that are created by the Virtual Storyteller are also coherent stories that contain different rhetoric relations. The emotions of the characters are interweaved in this discourse structure, as the characters' actions result from their emotions to a large extent. For instance, an agent that experiences much fear will flee from his enemy. The relation between an agent's emotion and an agent's behaviour is thereby mostly causal.

In some situations, a contrast relation may hold between an agent's behaviour and his emotion. An agent that experiences fear may nevertheless decide to stay near his enemy. The agent may behave this way because of certain circumstances or because his fear emotion is not intense enough to have him flee right away. In this situation, the agent's fear emotion does not result in a corresponding action. Nonetheless is it important to communicate this emotion to the reader and to express the contrast between the emotion and the action. The reader is then informed about the agent's emotion with a contrastive marker like *but*. By signalling to the reader that this behaviour might normally be considered odd but is in this instance correct, the story becomes more natural. The effect of a contrastive marker becomes clear

from the difference between sentence 6.4 and 6.5, in which Alexia's emotion and behaviour are in contrast. The use of *but* in sentence 6.5 helps the reader establishing the contrast.

(6.4) ♯ Alexia was afraid. She didn't flee from Brutus.

(6.5) Alexia was afraid but she didn't flee from Brutus.

We discuss the use of contrast relations in natural language and show how such relations are expressed. We concentrate specifically on sentence parts that contain gradable adjectives, which are generally used to express the emotions in the Virtual Storyteller, and we present how contrast relations can be used to express emotions.

6.1 The connective *but*

The relation between two text elements can be communicated to the listener or the reader by the use of operative markers that can conjoin two sentences or connect two text elements. Examples of operative markers were given in sentences 6.1, 6.2 and 6.3: *while*, *because* and *but*. For contrast relations, the markers *although*, *but*, *yet* and *nevertheless* are the most common expressions in English for which the Dutch equivalences are *alhoewel*, *maar*, *toch* and *desalniettemin* respectively. Dutch adds the connective *echter* to this frequency list. *But* (*maar*) is the prototypical marker of contrast. This connective can conjoin two sentence parts into one sentence. The meaning of *but* slightly differs from the meaning of the other markers, as the others are more specific than *but* and cannot be used in all contrastive uses of *but*. We will therefore focus on the use of *but* only and propose how *but*-coordinations can be constructed in the Virtual Storyteller.

Generally speaking, the connective *but* can join any two sentences that are somehow opposed but not contradictory to each other. *But* can therefore be used by the speaker to communicate that a contrast relation holds between the sentence parts. The general vision is that markers of contrast contribute mainly to the discourse structure and minimally to the meaning of the propositional content of the sentence parts. In some circumstances though, we believe that *but* can have an effect on the interpretation of the propositional contents of the sentence. Consider the sentence *The princess was scared, but she did not flee*. The connective *but* marks a contrast between the emotion and the behaviour of the princess. Additionally, *but* could indicate that the princess was not extremely scared, because this would have made her flee right away. This additional interpretation changes the interpretation of the first proposition; the intensity of the emotion is changed.

We cannot just assume that such interpretations of *but*-coordinations are derived; we first have to understand the *but*-coordination in more detail. To that end, we first study which types of contrast *but* requires. When we have gained more insight in what semantic relationship between two sentence parts licenses the use of a contrastive connective like *but* we can argue how this restriction on the use of *but* affects the interpretation made by the hearer.

6.1.1 Semantic approach

In 1971, Lakoff was one of the first to study the use of the connective *but* in natural language. She stated that there are two requirements for the contrastive use of *but* to conjoin two sentences. First, there must be some semantic relation between the two sentences. This

relation is mostly realized by some common property of the two objects or persons. For instance, a sentence that expresses the colour of one car can be conjoined with a sentence that expresses the colour of another car (see example 6.6). Beside such a similarity, the two sentences must differ in some way, so that *but* marks a contrast between the sentences. In 6.6 the colours of the cars differ. This second restriction of dissimilarity is not connected to the neutral connective *and*, which only requires some similarity between the two sentence parts, but may in some circumstances also conjoin things that are dissimilar. The second restriction thereby points out the difference between *and* and *but*. *But* is more specific than *and* and it marks a contrast. The combination of similarity and dissimilarity allows for the use of *but*.

(6.6) My sister has a black car, but my brother has a green car.

Semantic distinction

When the two rules described above are obeyed, there are two main types of *but*-coordinations that can be distinguished semantically (Lakoff, 1971). 6.7 and 6.8 are examples of these two types.

(6.7) The prince loves Snow White but the stepmother hates her.

(6.8) John is small but he's good at basketball. (Lakoff, 1971)

The first use of *but*, is the **semantic opposition** *but*. This type refers to conjoined sentences in which the sentence parts are directly opposed to each other in some semantic aspect. 6.7 shows an example of a semantic opposition use of *but*. This sentence contains a similarity and a dissimilarity. The similarity between the sentence parts is that the two subjects have emotional feelings towards Snow White that are described. The dissimilarity is the type of emotional feelings: love is opposed to hate. Lakoff says that for these sentences, "the subjects of the sentence parts are directly opposed to each other in a particular property" (Lakoff, 1971: pp 133). In this example, the prince and the stepmother are directly opposed in their feelings towards Snow White.

Lakoff calls the other use of *but* the **denial of expectation**. The two sentence parts in example 6.8 are not explicitly similar and dissimilar like in example 6.7. There is no pair of words opposed in the conjoined sentence. However, an implicit contrast between the two sentence parts is present. This contrast stems from a defeasible rule which is presupposed about small people: *Someone who is small is not good at basketball.* The first conjunct in 6.8 says that John is small and it matches the antecedent of the defeasible rule, which makes some general statement about small people. Applying the defeasible rule to this expression about John results in the expectation that John is not good at basketball. This defeasible inference is contradicted by the second conjunct. The sentence thus contains both an implicit similarity, John's affinity with basketball, and an implicit dissimilarity, bad versus good. Because the second conjunct denies the expectation that John is not good at basketball, this use of *but* is called denial of expectation.

The embedded defeasible inference in sentences with denial of expectation mostly reflects knowledge of the world or context knowledge. In the above example, the presupposed inference represents a generalized assumption about small people. As the example shows, this knowledge can be denied for some specific case. We must thus notice that the world or context knowledge that is associated with a denial of expectation is defeasible and that corresponding conclusions can therefore be cancelled.

Semantic restrictions

Sentences 6.7 and 6.8 are prototypical examples of the two types of *but*-coordinations and for these sentences it is easy to decide which *but* is involved. However, with the above descriptions for semantic opposition and denial of expectation the difference between these two types of *but* is not completely specified. Lakoff described four distinguishing properties or restrictions of semantic opposition usage of *but* that denial of expectation lacks. These properties thereby specify the difference between semantic opposition and denial of expectation. They were summarized in Spooren (1989: pp 12-13):

1. The two conjuncts express propositions that differ from each other in the value of one variable (e.g. *the prince* versus *the stepmother*) and in the fact that the predicate of the one proposition is negated in the other (e.g. *love* versus *hate*).

2. There is no logical relation between the two propositions (e.g. no entailment).

3. The *but*-coordination can be paraphrased with a *while*-construction or an *and*-construction.

4. The order of the conjuncts can be exchanged freely (symmetry).

The first restriction actually gives a definition of semantic opposition. This description is not applicable to denial of expectation. The other restrictions focus on characteristics of semantic opposition, which are not valid on denial of expectation. We have already described why the second restriction is not applicable to denial of expectation. The conjuncts of denial of expectation are logically related; the first conjunct defeasibly negates (a defeasible implication of) the second conjunct. A consequence of the logical relation in denial of expectation is that the *but*-coordination is asymmetric; the two conjuncts cannot be exchanged without changing the meaning of the sentence (as indicated by the fourth restriction). The marker *but* can therefore not be replaced by the symmetric *and* in denial of expectation (which corresponds to the third restriction).

Lakoff's proposal to distinguish two *but*-coordinations semantically on the above properties has gained much criticism. First, Spooren argues that Lakoff's analysis does not account for the effect of context on the interpretation of the *but*-coordination. The first property requires that the predicates in the two conjuncts are contrastive, but it does not mention the role that the context may play. Secondly, there is another type of *but*-coordination that Lakoff does not distinguish. This is the concessive opposition (Spooren, 1989), an argumentative interpretation of contrast, first observed in Anscombre & Ducrot (1983) and Dascal & Katriel (1977), and accepted by many researchers. In these *but*-coordinations, the first conjuncts expresses an argument in favour of a certain statement and the second conjunct expresses an argument against this statement. Lakoff's definition of semantic opposition does not clearly distinguish it from the concessive opposition and it seems to be incomplete. We will not discuss the concessive opposition in this chapter, because the Virtual Storyteller, in its current stage, does not require an implementation of this *but*-coordination for its language production. The last kind of criticism that was expressed in literature concerns the need for a semantic distinction at all and it marks Lakoff's theory as superfluous. Some linguists believe that there is only one type of *but*-coordination, and that a semantic distinction is not appropriate. For example, Lang (1984), Foolen (1991), Spenader & Stulp (2007) and Winter & Rimon (1994) have tried to explain all *but*-coordinations in terms of one analysis. We will come back to Winter & Rimon's analysis in section 6.3.1.

For now, we adopt Lakoff's distinction between semantic opposition and denial of expectation. Both types of *but*-coordinations would be useful to implement in the Virtual Storyteller. The semantic opposition can express explicit contrast, e.g. between emotions, properties of agents or actions of agents. The denial of expectation can express contrast between an agent's emotion and his behaviour. We aim to define formal criteria for these *but*-coordinations, so that their construction can be implemented in the Narrator Agent of the Virtual Storyteller. We therefore study these two types of *but*-coordination in detail and in chapter 7, we propose how they can be used to express emotions in the Virtual Storyteller.

6.2 Semantic Opposition

Lakoff (1971) has given an analysis of semantic opposition in terms of distinguishing properties and we have described some shortcomings of her analysis in the form of criticism. The first shortcoming concerned the absence of the context in the first defined property of semantic opposition. The semantic opposition seems not always identifiable without access to the context but Lakoff presents no theory about this. The other two objections doubted whether Lakoff's distinction between two types of *but*-coordinations was justified at all. One concentrated on the incompleteness of her theory, the other on the superfluousness. Because we aim to distinguish between semantic opposition and denial of expectation for the Virtual Storyteller, we only focus on solving the first shortcoming.

Spooren (1989) signalled the absence of the role of context in Lakoff's analysis. 6.9 gives an example of a *but*-coordination that cannot be classified unambiguously using Lakoff's distinguishing properties.

(6.9) John is tall but Charlie is fat.

This sentence violates the first restriction; the predicate *tall* is not negated by the predicate *fat* and vice versa. But 6.9 does have all the other properties of semantic opposition: there is no logical implication related between the conjuncts, *but* can be replaced by *and* or *while* and the conjuncts are symmetric. It is therefore desirable to categorize 6.9 as a semantic opposition. Lakoff's first restriction, in which the contrast between the predicates is defined as semantic, seems then overspecified. We need to generalize this restriction so that a sentence like 6.9 is accepted as a semantic opposition. In this sentence, *tall* and *fat* are not semantically contrastive, but due to the context they may be interpreted as such.

Spooren reformulated Lakoff's first property in a definition for semantic opposition, and he added the involvement of the context, resulting in the following definition:

Semantic opposition: A relation between two conjuncts each having different subjects[1], to which properties are attributed that are mutually exclusive in the given context. (Spooren, 1989: pp 31)

We see that 6.9 satisfies this definition. John and Charlie are different subjects with properties tall and fat. In the given context, tall and fat must be mutually exclusive. So, this sentence is accepted as a semantic opposition if there is some plausible reason to assume that the two predicates mutually exclude each other. An example of context knowledge that justifies this

[1]The term 'subjects' refers to the elements that are focused in the two conjuncts, rather than to the elements that function as syntactic subjects. For a focus-based analysis of *but*-coordinations, we refer to Umbach (2005).

assumption is that both persons have only one characteristic shape, in this example either tall or fat (or e.g. small or broad etc.). If we want to construct a semantic opposition we thus have to make sure that the compared properties are mutually exclusive in the context.

If a semantic opposition is introduced by incompatible predicates the subjects in the two conjuncts have to differ, as was stated in Spooren's definition. The definition contains no restriction on how the subjects are related, but Spooren claims that the two subjects have to have a comparable semantic content. He explains this claim with sentences 6.10 and 6.11.

(6.10) In France the people are Roman Catholic, but the Brits are Anglican.

(6.11) ♯ In France the president is Roman Catholic, but the Brits are Anglican.

In 6.10 properties of people with a certain nationality are compared. Although the subjects differ syntactically, their semantic content is comparable. In the first conjunct, the subject is the people of the country France, in the second conjunct the subject *the Brits* refers to the people of the country Britain. This sentence is therefore a correct semantic opposition.

In 6.11 the first conjunct contains an uniquely identifiable referent as subject, *the president of France*, whereas the subject in the second conjunct refers to a whole group of people. These two subjects cannot be compared in a semantic opposition. Spooren thus concludes that the subjects of the two conjuncts must be part of the same class.

In order to compare two elements they must at least be part of the same class. But this claim is still vague because classes are defined on different levels. For example, an ontology can have a class 'organisms', which subsumes the classes 'animals' and 'plants'. This structure is shown schematically in figure 6.1. Semantic opposition can compare the concept animal to the concept plant (like in sentence 6.12), because they are both part of the class organisms. But, without specifying more constraints on the relation between the compared subjects, the concept elephant could also be comparable to the concept pine tree, as they are also both organisms. However, sentence 6.13 shows that this comparison results in an incorrect *but*-coordination. So it does seem that only elements belonging to the same level in the ontology can be compared.

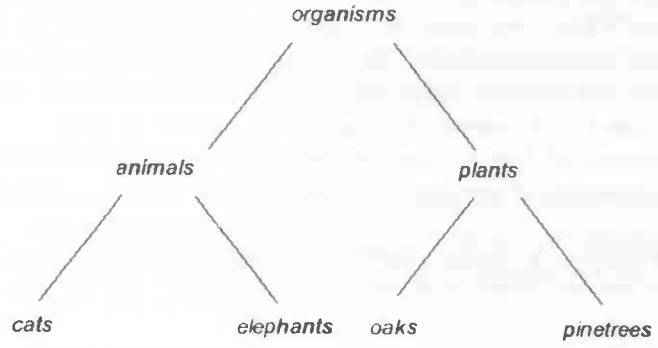


Figure 6.1: An ontology of organisms

(6.12) An animal can move freely from place to place but a plant cannot.

(6.13) ♯ Elephants have trunks, but pinetrees don't have trunks.

Lang (1984) has claimed that coordinated elements have to be semantically independent, neither of them subsuming the other. This means, that according to our ontology an animal cannot be compared to a cat, although they are both part of the class 'organisms'. The difference between the relations plant - animal, elephant - pine tree and animal - cat is that in the first relation the two concepts are both directly subsumed by the same class (organisms), whereas in the other two relations, the subsuming classes are different. We could then specify our restriction by claiming that the two compared concepts must be directly subsumed by the same class.

Notice that this restriction is still not very precise. For example, context can play a great role in defining (temporary) classes. Cats can be compared to human individuals when talking about a class of all living things that were present in an apartment, like in the well formed semantic opposition in sentence 6.14. And if we expanded our ontology in figure 6.1, by distinguishing domestic animals from wild animals, we would not be allowed to compare cats to elephants according to our claimed restriction.

(6.14) The child survived the fire in the apartment, but the cat died.

As these examples show, the context and the nature of the class largely determine which elements can be compared in a semantic opposition. There is obviously a very complex rule associated with the compared elements and applying such a rule requires agreement on the structure of the ontology. It it is beyond the scope of this research to construct such a rule. Instead, we can adopt the view of Umbach (2004): An alternative has to be added in the second conjunct to what was discussed in the first conjunct and *but* requires the alternative in the second conjunct to be excluded. Umbach thereby refers to the meaning of *but* as anti-additive; it excludes an alternative. So, in sentence 6.14 the alternative subject for *the child* is *the cat*. The *but*-coordination excludes *the cat* from the category of organisms that survived the fire (which is introduced by the first conjunct). Thus, the coordinated element in the second conjunct has to be an alternative for the subject in the first conjunct. Whether or not an element can function as alternative, depends of course on both the world knowledge (an ontology) and the context.

Formulating one general rule that exactly defines the required relation between the subjects in semantic opposition is too ambitious at this point. However, we need some direction when constructing semantic opposition in the Virtual Storyteller. We will therefore hold on to the restriction that two compared elements in semantic opposition must be directly subsumed by the same class in the storyteller's ontology. This restriction is not a necessary condition for semantic opposition, but at least a sufficient one. In other words, if we know that two concepts are directly subsumed by the same class, it is safe to construct a semantic opposition that has these concept as subjects. In chapter 7 we propose how semantic opposition can be constructed in the Virtual Storyteller.

6.3 Denial of expectation

Denial of expectation is semantically distinguished from semantic opposition in that it expresses implicit contrast between the two conjuncts. Like we did for semantic opposition, we

specify the restrictions for denial of expectation in this chapter and in chapter 7 we propose how denial of expectation can be used for expressing emotions.

The difference between semantic opposition and denial of expectation was made clear in the previous section, but we have not formulated a definition for denial of expectation. Spooren (1989) defined it as follows:

Denial of expectation: On the basis of the first conjunct one expects a certain state of affairs to be true, which is denied by the second conjunct. (Spooren, 1989: pp 31)

The use of *but* in denial of expectation can in fact be viewed as indicating that the second proposition obeys the maxim of Relevance (Grice, 1975), discussed in chapter 4. Blakemore (1989) suggests that a speaker uses *but* to indicate how the second conjunct is consistent with the principle of Relevance. The relevance of the second conjunct depends on the interpretation of the first conjunct; the first conjunct may have yielded a defeasible inference that the speaker wishes to deny. By using *but*, the speaker constrains the interpretation of the first conjunct. Consider example 6.15, taken from Lakoff (1971) and 6.16, where *but* is replaced by a full stop. In 6.16 the hearer will not establish the relevance of the second proposition which is indicated by the use of *but* in 6.15. The interpretation is that the two propositions are causally related. The need for the marker *but* in denial of expectation distinguishes this *but*-coordination from semantic opposition. In semantic opposition, the marker but can be replaced by a full stop without changing the interpretation, as examples 6.17 and 6.18 show.

(6.15) John is a republican, but he is honest.

(6.16) John is a republican. He is honest.

(6.17) John is small but Peter is tall.

(6.18) John is small. Peter is tall.

So far, we have only given an informal, linguistic analysis of the use of *but*. The distinction between semantic opposition and denial of expectation is based purely on some informal properties of the conjoined sentence. In order to obtain more insight in the restrictions of denial of expectation (and the effect of this *but*-coordination on the interpretation of the propositional contents), we present a more formal approach of denial of expectation.

6.3.1 Formal approach

Winter & Rimon (1994) have formalized the informal analysis of denial of expectation proposed by Lakoff (1971). They defined the logical relation between the two conjuncts in a *but*-relation more formally. Let us reconsider example 6.8, which is repeated here for the ease of reading:

"John is small, but he's good at basketball"

If we use the propositions p to refer to *John is small* and q to refer to *John is good at basketball* we have a sentence of the form 'p but q'. The inference that can be derived from the marker *but* can be roughly formulated as:

(A) p defeasibly implies not q

The general inference connected to this particular sentence is that all small people are not

good at basketball. John is an instance of a small person, thus we can infer the particular implication that says that *John is small* implies *John is not good at basketball.* Winter & Rimon emphasize that the imply-relation between p and q defeasible. If it was a normal implication, the sentence would contain two contradictory propositions. It should therefore be interpreted as a defeasible implication, like in 'normally, p implies q', so that its consequent, q, may not always be derived from its antecedent, p, like in denial of expectation.

What happens if the defeasible implication that is derivable from the first proposition of a *but*-sentence is not valid according to the world knowledge or context knowledge? Consider example 6.19, taken from Winter & Rimon. Given the marker *but*, the inference evoked from this sentence would be that p (= it was cloudy) implies *not* q (= it was not raining). Because this implication does not hold according to our world knowledge, this use of *but* is inappropriate (unless the context would suggest something else). The two sentences should therefore be conjoined by some other marker, like *because*, or by the neutral conjoint *and*.

(6.19) ♯ ^{p}[It was cloudy], but ^{q}[it was raining].

Winter & Rimon use the term *infelicitous* to refer to a sentence like 6.19, in which the use of *but* raises an invalid implicature. We must notice that logically speaking, the use of *but* in this sentence does not change the truth values of the sentence in comparison with the use of an appropriate connective. The truth conditions of any contrastive conjunction 'p but q' are exactly the same as those of 'p and q'.

Winter & Rimon (1994) argue that all uses of *but* can be explained in logical terms and they make no distinction between semantic opposition and denial of expectation, as Lakoff did. We will now give another example of denial of expectation and describe the logical framework that Winter & Rimon developed to explain the *but*-sentences.

(6.20) ^{p}[It was raining], but ^{q}[I took an umbrella].

In 6.20 the *but*-conjunction is appropriate, although 'p implies not q' does not hold. The use of *but* in 6.20 indicates some other contrast between the sentence parts. This contrast is present by means of an implicit statement that is defeasibly denied by p and defeasibly implied by q. Winter & Rimon use the notation r to refer to this implicit statement. In this context, r can refer to *I will not get wet*, because *It was raining* defeasibly denies that the speaker will not get wet (or defeasibly implies that the speaker will get wet) and I *took an umbrella* defeasibly implies that the speaker will not get wet. Due to the need of an implicit statement r, the contrast in 6.20 is indirect.

Winter & Rimon combined formulation A, given above, with the description above, in which r refers to an implicit statement, and formulated a generalized condition, assumed to represent the underlying semantics of all contrastive *but*-sentences:

(C1) p implies *not*(r) and q implies r

This general condition can also account for the use of *but* in example 6.8. Applying C1 to this sentence, we get the implicatures p, *John is small* implies r, *John is not good at basketball* and q, *John is good at basketball* implies r, *John is good at basketball.* In this case, r is identical to q (and q thus implies r). Because here r is identical to q, we speak of direct contrast.

As was already suggested by Lakoff, the second implication, 'q implies r', is stronger than the first implication, 'p implies not r'. The conclusion *not*(r) that can be derived from the

first implication is therefore cancelled by the second implication. Winter & Rimon formulate this as follows:

(C2) q's implication of r is 'stronger' than / 'cancels' p's implication of *not*(r).[2]

Winter & Rimon claim that the two conditions formulated above can also account for Lakoff's semantic opposition use of *but*. Consider example 6.21, in which the heights of the prince and the princess are compared. How does this sentence satisfy both C1 and C2? C1 says that *the prince is tall* should deny some r and that *the princess is small* should imply some r. C2 says that the second implication is stronger than the first one. Winter & Rimon would choose r to refer to *not all characters are tall*, so that both C1 and C2 are satisfied. *The prince is tall* then defeasibly denies *not all characters are tall*. This conclusion is cancelled by *the princess is small* which implies *not all characters are tall*.

(6.21) The prince is tall but the princess is short.

Although Winter & Rimon's logical framework can also account for semantic opposition *but*, we find their explanation rather implausible. In order to apply their framework to semantic opposition, the two sentence parts must be related to some far-fetched implicit statement r in cases where the context does not supply the implicature. In the above example, r could refer to *not all characters are tall*, but we believe that this statement is not related to what the speaker tries to communicate with the sentence. We think that a *but*-coordination like 6.21 is interpretable without identifying an implicit inference and selecting some r. In a semantic opposition the speaker makes a comparison between the two subjects in the conjuncts and tries to communicate the contrast between the two. For semantic oppositions, the internal structure of the sentences is characteristic and it distinguishes these sentences from denial of expectations. The properties of the subjects in the conjuncts are explicitly in contrast since they refer to the two opposite ends on the same scale. Winter & Rimon ignore the internal relationships that are obviously present in semantic oppositions. For example, in sentence 6.21 they ignore the relationship between 'tall' and 'short'.

Beside the fact that Winter & Rimon's approach is implausible for semantic opposition it is also impracticable for automatically generating semantic oppositions. In a semantic opposition, there are numerous r's that can be selected to satisfy both C1 and C2 and it is unclear which one is most plausible (in cases where the context does not supply any). For automatically identifying semantic opposition relations Winter & Rimon's restriction is too vague to us and there is a simpler way to do it. We will instead follow the approach of Spooren (1989) and search for semantic contrast when constructing semantic opposition.

Of course, there are contextual circumstances in which a semantic opposition can indeed be explained by Winter and Rimon's framework. If someone asked: "Are all characters in the Virtual Storyteller tall?", then the contrastive statement 6.21 could indeed be interpreted as Winter & and Rimon suggest. However, in the absence of such a context, Winter & Rimon's explanation of semantic opposition is very unlikely. Nevertheless, we continue presenting their framework and give their definition for contrast sentences, as we believe their logical framework is appropriate for denial of expectation use of *but*.

[2]The difference between the imply relations is intuitive here. Winter & Rimon also give a formal approach to these relations. For the interested reader we refer to section 3.1, Restatement of (C1) and (C2) in possible world semantics, in Winter & Rimon (1994).

Definition

Winter & Rimon combined C1 and C2 and drew the following definition for the use of *but* in contrast relations:

(D) The contrast relation: A proposition r *establishes* contrast between two (ordered) propositions p and q iff $\Diamond(p \rightarrow \neg r) \wedge (q \rightarrow r)$ is true. [3]

From the above definition that distinguishes between the types of the implications that are related to p and q (p's implication is weak, q's implication is strong), we can conclude the following. If we assume that p and q are both true, so that the speaker is not lying, it must be concluded that $(p \rightarrow \neg r)$ is false for this specific case. This step is necessary in order to prevent an inconsistency between the two conjuncts. Because the conclusion r that is derived from the second conjunct cannot be cancelled (the second implication is strong), we must cancel the conclusion that is derived from the first conjunct.

We have seen that Winter & Rimon's definition for contrast relations assigns *but* the same interpretation in all situations. As stated above, we wish to distinguish between the different semantic uses of *but* that were defined by Lakoff. We therefore use the above definition to analyse the use of *but* in denial of expectation only. For semantic opposition, we stick to the approach in section 6.2.

6.3.2 Effect on sentence content

In this section, we study what effect denial of expectation can have on the interpretation of the propositions of the conjuncts. An example of denial of expectation contrast is given in 6.22.

(6.22) Sam is a nice guy, but he hits his sister.

This example contains a contrast relation, given that our world knowledge says that nice guys do not hit their sisters. The property 'nice' entails a number of qualities, like 'behaving politely', 'not teasing anyone' and 'not hitting anyone'. 'Not hitting' is thereby an assumed characteristic of nice people. If we apply this world knowledge to the above sentence we obtain two conflicting conclusions: according to the first conjunct Sam does not hit his sister, and according to the second conjunct Sam hits his sister. Because the second conjunct cannot be denied (assuming that the speaker does not give any false information), the conclusion linked to the first conjunct must be cancelled. This was also stated by definition D.

The general view about the interpretation of a sentence like 6.22 is that the implication associated with the first conjunct ('p implies not q') is false for this specific case. Winter & Rimon used the modal operator of possibility to indicate the violability of this implication. So, the nice guy Sam does not have the quality of not hitting his sister. This shortcoming of Sam does not change his degree of niceness.

Weakening the first conjunct

We think that another explanation could be given to cancel the false conclusion from the first conjunct in 6.22. The world knowledge that corresponds to the implicature embedded in this

[3] $\Diamond(p \rightarrow \neg r)$ should be interpreted as a weak implication. This implication is not necessarily true. In the context of contrast relations, a weak implication refers to an expected or default implication.

sentence states that nice guys do not hit their sisters. This formulation is not very precise, but we can assume that it only focuses on guys that are nice to a certain degree. For example, very nice guys do not hit their sisters, but guys that are only little nice may hit their sisters. The first conjunct of sentence 6.22 that contains the gradable adjective *nice* does not specify how nice Sam is. The exact intensity of the expressed adjective is not clear. Because of the vagueness of this expression, the antecedent of the inference may not match the first conjunct, so that the consequent that says that Sam does not hit his sister cannot be derived. In other words, Sam may be less nice than the inference about nice guys presumes. The marker *but* thereby nuances how nice Sam is.

The contrast in sentence 6.22 comes from the typical interpretation of the gradable adjective *nice* when it is not modified by any degree marker. It is then typically interpreted as quite strong, referring to the positive end of the nice scale. The use of *but* nuances this typical interpretation. If the first proposition had already contained a weaker adjectival expression like in *Sam is not entirely nice*, we would not be able to derive contrast by adding a proposition like *but he hits his sister*. The first proposition then already shows that Sam's niceness has some shortcomings and the hitting behaviour of Sam would not come as a surprise to the hearer and the relation between the two propositions would not be contrastive.

Our proposed interpretation of the *but*-coordination is obviously different from the general interpretation that was described in the literature. It supposes that the marker *but* has an effect on the interpretation of the scalar adjective of the first conjunct; the interpretation is put to a weaker point on the scale.[4] It thereby focuses on the interpretation of the first expression rather than on the validity of the implication, as was done by e.g. Lakoff (1971) and Winter & Rimon (1994).

Cancellation of information

Spooren (1989) also considered the marker *but* to influence the interpretation of the first conjunct. In his work, he studied the interpretation of different perspectives that were expressed by one speaker. An example is given in 6.23.

(6.23) Smith informs me that Burns employs thousands of workers, but according to Robert Burns employs 100 men.

In this example, the information related to the two conjuncts is directly in conflict. The first conjunct defeasibly implies that Burns employs thousands of workers and the second conjunct defeasibly implies that Burns employs a hundred workers. We could say that the second implication is stronger than the first implication (as in definition D) and solve the conflict by claiming that the conclusion from the first defeasible implication must be cancelled. Spooren argued that when the speaker expresses different perspectives the second conjunct is often interpreted as the speaker's opinion and he formulated the Asymmetry Hypothesis:

> If a *but*-coordination is used, the information related to the second conjunct stands a better chance of percolating to a higher perspective than the information related to the first conjunct. In this respect *but*-coordinations differ from *and*-

[4] In the current work, we do not formalize what a weakened interpretation is. As we have seen in the previous chapters, (emotion) scales are regarded as numerical spectra. We thus suggest that a weakened interpretation is interpreted as a lower numerical value than the value of the prototypical interpretation of the scalar adjective.

coordinations, in which neither one of the conflicting informational units percolate (Spooren, 1989: pp. 41).

The Asymmetry Hypothesis says that in *but*-coordinations, the second conjunct is generally interpreted as more important than the first. This means that a hearer has the tendency to interpret the information related to the second conjunct as true (or as the speaker's opinion), although in some circumstances the hearer may also just not be sure about which of the two perspectives is true. Spooren supported his hypothesis with empirical research (Spooren, 1989). For example, subjects were presented argumentative texts in which two contrastive claims were related by means of the marker *but*. The subjects had to judge how much the writer of the text agreed with the claim and they generally regarded the second claim as the speaker's opinion.

If we apply the Asymmetry Hypothesis to example 6.23 the information related to the second conjunct, "Burns employs a hundred men", is likely to be interpreted as the speaker's opinion. Does Spooren suggest that the *but*-coordination affects the interpretation of the sentence parts themselves? In sentence 6.23, by using the marker *but*, the hearer is more likely to interpret the second conjunct as the speaker's opinion than when the marker *but* was not used. If the speaker had solely expressed the first conjunct, the principle of Relevance (Grice, 1975) ("make your contributions relevant") would have suggested that the hearer interpreted the speaker to believe Smith's words and that Smith's words are true. So, the marker *but* affects how the hearer interprets the first conjunct. That *but* has an effect on the contents of the sentence parts becomes also clear when we exchange the two conjuncts. We then see that the hearer is likely to interpret Smith's opinion to be true and Robert's opinion to be false.

6.3.3 A new approach: *but* as a weakener

As said before, the general view is that the use of *but* in denial of expectation cancels the conclusion that can be derived from the first conjunct. However, we assume that in some *but*-coordinations, the use of *but* may be more appropriately interpreted as simply weakening the interpretation of a scalar element in the first conjunct. We will explain this effect combining Winter & Rimon's and Spooren's approach.

Selection of r

Winter & Rimon's definition says that for a sentence 'p but q', p defeasibly denies some r and q implies some r. We have seen that r can refer to an implicit statement and that it can correspond to the second conjunct q. If r refers to an implicit statement, there is indirect contrast in the *but*-coordination, in other cases the contrast is direct. This section shows how the selection of an implicit statement r in sentence 6.22 relates our approach of *but* as a weakener to the approach of Winter & Rimon and Spooren. We now reconsider sentence 6.22, which is reformulated in 6.24 for the ease of reading.

(6.24) Sam is a nice guy but he hits his sister.

In this sentence, p refers to *Sam is nice* and q refers to *Sam hits his sister*. We have seen that r can refer to *Sam hits his sister* in order to satisfy definition D. Then r matches q, so q directly implies r. The use of *but* in this sentence only expresses that some expectation about Sam is negated. The negation of the expectation does not contribute to the interpretation of

the first sentence part; Sam is still a nice guy. However, he misses the quality of not hitting his sister.

Could we select some other statement for r that still satisfies definition D? Spooren's Asymmetry Hypothesis claimed that if a sentence expresses two conflicting perspectives, the second perspective is generally regarded as true. We could apply this hypothesis to sentence 6.24 and assume that the speaker himself expresses two perspectives about Sam:

perspective 1: Sam is an ordinary nice guy
perspective 2: Sam is not an ordinary nice guy

The first perspective stems from the first conjunct, *Sam is a nice guy.* The adjective *nice* is not preceded by a degree modifier so we are not completely certain about how nice Sam is; he may be a little nice and he may be nice in an ordinary way. Because the speaker did not add any modifier, the hearer can assume that the speaker meant that Sam is just ordinary nice (which is in accordance with Grice's principle of quantity (Grice, 1975), described in chapter 4).

The second perspective can be derived from the expression *Sam hits his sister.* Our world knowledge says that guys who are ordinary nice, do not hit their sister. This world knowledge can be treated as a hard implicature rather than as a defeasible rule (as was true for denial of expectations discussed so far). Because Sam does hit his sister, we must conclude that he is not ordinary nice. From this denial, we cannot infer how nice Sam exactly is. However, we know that he is at least less nice than ordinary nice.

Sentence 6.24 can thus be explained as containing two conflicting perspectives, both expressed by the same speaker. Spooren (1989) suggests that the information related to the second conjunct corresponds to the speaker's opinion and thereby cancels the first perspective. So, the first sentence part remains true, but the perspective that was expressed by the first sentence part is cancelled. If we adopt this view, we obtain the truth values in table 6.1.

	First conjunct	**Second conjunct**
Expression	Sam is a nice guy [true]	Sam hits his sister [true]
Perspective	Sam is ordinary nice [false]	Sam is not ordinary nice [true]

Table 6.1: Truth values of the expressions and perspectives

If we interpret sentence 6.24 this way, can we then speak of a contrast relation? Definition D say that if we can denote some r that p defeasibly denies and q implies we can. In the above interpretation, r refers to *Sam is not an ordinary nice guy.* This statement is defeasibly denied by the first conjunct and it is defeasibly implied by the second conjunct. *Sam is a nice guy* defeasibly denies *Sam is not ordinary nice* (or defeasibly implies *Sam is ordinary nice*) and *Sam hits his sister* defeasibly implies *Sam is not ordinary nice.* This assignment of r thus satisfies definition D. We will now discuss how this selection of r affects the interpretation of the sentence.

Effect on interpretation

Although the use of *but* rather than *and* in sentence 6.24 does not directly change the truth values of the sentence parts it does contribute to the interpretation of the sentence as a whole.

The implicit statement r (*Sam is not ordinary nice*) is assumed to be true in the contrast relation, as we have seen in Winter & Rimon's approach. Besides, both expressions *Sam is a nice guy* and *Sam hits his sister* are regarded as true. The first of these is related to the implicit statement *Sam is not nice*, so we should combine them in order to derive some general judgement about Sam's niceness.

The adjective *nice* is gradable and the denial of some degree of nice can logically refer to many points on the scale. The implicit statement r says that Sam is less nice than ordinary nice. This can vary from not nice to nearly ordinary nice. The expression p says that Sam's degree of niceness is somewhere on the nice side of the scale, which varies from only little nice to very nice. These two statements about Sam's niceness are illustrated in figure 6.2. The thick line indicates the overlap between the two statements, which determines how the combination of the two should be interpreted. Combining these two statements about Sam's niceness thus results in a degree of niceness that lies somewhere between neutral and ordinary nice, preferably in the middle of the two ends. We call this degree *little nice* for the ease of writing.

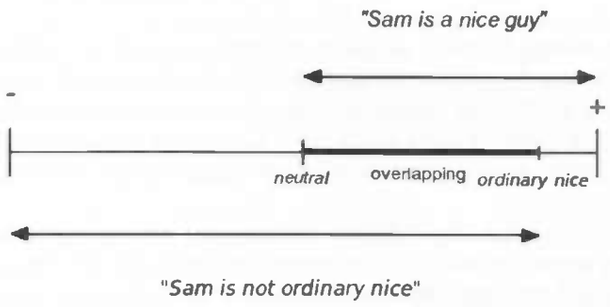


Figure 6.2: Resulting interpretation about Sam

Concluding, how does the use of the marker *but* contribute to the interpretation of sentence 6.24? It weakens the adjective *nice* that was used in the first conjunct. We could analyse its meaning as follows: Sam is a nice guy, but not so nice that he doesn't hit his sister. He is therefore only little nice. The degree of niceness that was expected from the first expression about Sam is decreased by the second expression about Sam. So, by denying what we would expect from the first conjunct, the first conjunct is weakened or nuanced.

A crucial difference with the interpretation of denial of expectation according to Winter & Rimon's approach is that here, a hard implication (and not a defeasible one) is associated with the first proposition. Since the second proposition denies what was expected from the first proposition, the interpretation of the first proposition must be changed. This is realized by interpreting the scalar adjective below a certain threshold on the scale. Thus, rather than denying the associated implication as a whole, the antecedent of the implication is weakened so that the implication does not go through. Of course, this analysis only applies to *but*-coordinations in which the first proposition contains a scalar expression.

Cases of weakening

We have seen how we can explain the weakening function of the marker *but* for sentence 6.24. We now discuss in what cases we can derive a weakening effect from *but*, so that we know when we can use a *but*-relation in the Virtual Storyteller to weaken emotional expressions. In this section we present some prototypical examples of *but*-coordinations in which the interpretation of the first proposition is obviously weakened or not. The regularity that we obtain in the weakened expressions will then be applied to expressions in the Virtual Storyteller that we present in chapter 7, so that we know when we can use this weakening method.

Generally speaking, a weakening effect occurs only if p contains some gradable expression that is related to q. This means that the first conjunct of the *but*-coordination must contain a scalable proposition. The question then is to which subset of these *but*-coordinations a weakening effect is derived. The main idea is that the second conjunct q denies some degree of the first conjunct p.[5] We illustrate this rule with sentences 6.25 and 6.26 that are clear examples of the weakening effect of *but*.

(6.25) It was a cold winter day, but the pond was not frozen.

(6.26) It was early in the morning, but the sun had already risen.

In 6.25 the adjective *cold* is scalable and some degrees of coldness are denied by the second conjunct, namely below zero degrees Celsius. The second conjunct thereby nuances the interpretation of the first conjunct. The temperature is thus higher than zero degrees Celsius. In 6.26 the adjective *early* is scalable. The second conjunct states that is it later than sunrise and it thereby denies all hours before sunrise. So, the second conjunct nuances what time it is.

In sentences 6.25 and 6.26 the second conjunct necessarily denies some degree of the first conjunct. There are no normal circumstances that could account for a very early sunrise or a very cold day without frozen ponds. We see that the first proposition raises a prototypical interpretation of the adjective that is not compatible with the second proposition and it is therefore weakened. In order to specify when such an interpretation is appropriate we first give some examples in which the weakening effect of *but* is not present.

(6.27) John is small, but he is good at basketball.

(6.28) Anne is intelligent, but her exam results were bad.

In 6.27 John's height is not weakened by the second conjunct. The second conjunct does not deny some length of John. The implicit statement that relates a person's height to his/her ability in basketball says that *most* good basketball players are tall. Thereby, the statement allows for exceptions. John is an exception to this rule; he is small and he is good at basketball. The fact that John is good at basketball does not nuance John's height but it only indicates that John's height is in contrast with his affinity with basketball. In other words, *in spite of the fact that John is small, he is good at basketball.*

[5] From now on, we treat the weakened denial of expectation as a form of direct contrast, where the first conjunct p directly denies q. In the previous section we proposed the selection of some implicit r and this may have suggested that the *but*-coordination contains indirect contrast. However, this implicit r was only used to relate our theory to the theory of Spooren and the theory of Winter & Rimon and we will not pay attention

Sentence 6.28 should be interpreted in a similar way. From intelligent people we expect good exam results. However, it is very well possible that some intelligent person gets bad exam results. We can come up with many reasons to clarify why Anne's results were bad. The implicit statement that relates someone's intelligence with his/her exam results also allows for exceptions. So, this example can also be reformulated as: *in spite of the fact that Anne is intelligent, her exam results were bad.*

The main difference between the above two sets of examples is the imply relation between the second conjunct and the first conjunct. In the weakened constructions the implication is hard. That is, the first conjunct contains some predicate that has a set of properties. One of these properties is then denied in the second conjunct. So, in example 6.25, a temperature less than zero has the property of frozen water under normal circumstances. The fact that the water is not frozen implies that the temperature is above zero.

For the non-weakening effect of *but*, we could say that the second conjunct denies some *expectation* about the first conjunct and that the imply-relation between the second and the first conjunct is therefore defeasible. So, like in example 6.27, as John is small we expect that he is not good at basketball. However, this is not a property of small people, it only applies to a large number of small people. If the expectation that stems from the first conjunction is denied, we can simply remove the defeasible implication.

Summarizing, in cases of weakening the second conjunct denies the prototypical interpretation of the first conjunct and the hearer has to interpret the position of the scalar proposition as lower. In other cases the second conjunct defeasibly denies the prototypical interpretation of the first conjunct and the conflict is solved by removing the defeasible implication.

Using these criteria we can explain the interpretation of sentence 6.24. Sam is said to be nice and he is also said to hit his sister. The character type nice includes some behavioural properties like 'not hitting'. This makes the implication $p \rightarrow \neg q$ a strong one. Someone who is ordinary nice does not hit. We assume that hearers have this detailed information about properties of predicates. When the second conjunct expresses that Sam does hit, the hearer weakens Sam's niceness on the scale. This interpretation could be changed if some reason for Sam's odd behaviour was expressed. For instance, the speaker could have continued saying that Sam's sister was a very mean woman and that she deserved to be hit. This extension would have explained why Sam, who is a nice guy, hits his sister and it would not weaken the interpretation of the first conjunct. If so, the sentence could be reformulated as: *in spite of the fact that Sam is nice, he hits his sister.*

Although our established criterion for a weakening effect of *but* is a bit vague and we have not tested this assumption empirically it helps us reasoning about how specific examples of contrast relations could be generated in the Virtual Storyteller to help the reader get the intended interpretation. We think that speakers use *but*-coordinations to specify certain scalar propositions and it would also be desirable to use such constructions for the storyteller. Because scalar expressions on their own are highly unspecified, one way speakers have available to make their actual value more specific is to add information about a part of the scale that is not relevant. Thus, contrast is an implementation tool for speakers to be more specific about scales and we wish to also use contrast in the Virtual Storyteller to this end.

Our weakening approach to a subset of denial of expectation is very different from the analysis of Lakoff and Winter & Rimon but it has a big advantage when generating denial of expectation in the Virtual Storyteller. Our *but* contributes to the propositional meaning of the sentence and it can thus be used to change the scalar position of the proposition.

This makes it possible to derive a plausible interpretation of the *but*-coordination without having access to a database of world knowledge implicatures. We simply use *but* to signal weakening when the position on a scale for a given emotion is associated with actions that are unexpected given a normal position on the scale. Emotions are predicates with certain behavioural properties and we expect that the interpretations of emotional expressions (which are scalar) are weakened if one of their properties is denied in a *but*-coordination. For this set of *but*-coordinations our approach removes the need for referring to defeasible implicatures because all implicatures related to emotions are assumed to be hard. If the first and second proposition of a *but*-coordination contradict this is solved by weakening the first proposition.

However, our approach does not work for all *but*-coordinations in the Virtual Storyteller. It will not work for antecedents that are not scalar, so we need to clearly distinguish a weakening interpretation form a non-weakening interpretation in the Virtual Storyteller. And, even if the first proposition is scalar, the weakening interpretation may not be derived. We have seen some examples of these above (e.g. sentence 6.27 and we admit that the border between hard implications and defeasible implications is vague. But, we think that our theory does at least work for a very large subset of the scalar expressions in the Virtual Storyteller since we believe that emotions are scalar predicates with a number of fixed properties. This means that emotions on certain points on the scale really imply certain properties. In chapter seven we will discuss the interpretation of these expressions in more detail.

6.4 Conclusion

We have discussed two types of contrast relations: semantic opposition *but* and denial of expectation *but*. In semantic opposition, two sentence parts are directly opposed to each other in their semantic contrast or their contextual contrast. The denial of expectation expresses implicit contrast, which, according to Lakoff (1971) and Winter & Rimon (1994), results from a defeasible implication related to the first conjunct. The second conjunct denies what was defeasibly implied from the first conjunct and the implication is cancelled for these *but*-coordinations.

For the implementation of the two contrast relations in the Virtual Storyteller, we hold on to the criteria discussed for each type of contrast. The semantic opposition requires the compared subjects to be directly subsumed by the same class in the Virtual Storyteller and the contrastive properties in the two conjuncts need to exclude each other in the context. Constructing semantic opposition in the storyteller is useful when describing a number of related, yet different, things (like emotions or agents)

The denial of expectation can be constructed if the second conjunct denies what was (defeasibly) implied by the first conjunct. If the first conjunct contains some gradable lexical item which is not modified explicitly and if the second conjunct excludes some property of the prototypical interpretation of the first conjunct, we assume the interpretation is weakened. For these *but*-coordinations, the imply-relation between the prototypical interpretation of the first proposition and the second proposition is hard and not defeasible. This means that, in case of a contradiction between these two, the implication cannot be canceled and instead, the interpretation of the first proposition is changed. The scalar proposition is placed on a lower point on the scale. This is prevented if some explanation for the denial is expressed explicitly, which therefore cancels the imply relation by making it a defeasible one. We think that, from a generation perspective, our weakening approach is more practical than the approach of Lakoff

and Winter & Rimon, since we do not have to deal with defeasible implications and the scalar positions of unmodified emotional expressions can be specified.

The next chapter proposes how semantic opposition and denial of expectation can be generated in the Virtual Storyteller.

Chapter 7

Design Proposal

In the previous chapters we have studied how emotions are expressed in natural language, by conducting a corpus study, experimental research and a theoretical study. We have identified lexical items for the emotions in the emotion model, we have accessed complex expressions with their preferred interpretations using the lexical items in combination with modifiers and we have found ways to use contrast relations in expressing emotions. The current chapter proposes how the Narrator Agent of the Virtual Storyteller can be improved with the results of these studies.

Before presenting the main changes of the Narrator Agent, we briefly describe the current architecture of the Narrator Agent. Next we discuss how and when the emotions of the character agents with their corresponding intensity values can be expressed in an explicit way, and after that we outline how contrast relations can be implemented in the storyteller.

The Narrator Agent For expressing the emotions of the character agents the language generation module of the Virtual Storyteller needs some adjustments. The Narrator Agent of the Virtual Storyteller performs the language generation task and it transforms the fabula structure (the plot) generated by the Plot Agent into natural language. We have described the architecture of the Narrator Agent in chapter 2 and we have seen that it consists of a pipeline of three components, the Document Planner, the Microplanner and the Surface Realizer. The architecture is shown in figure 7.1 with the input and output of each component. The first component, the Document Planner converts the fabula structure and some background information structures into a document plan (which contains all story-information). The document plan is then transformed into a rhetorical dependency graph by the Microplanner and the Surface Realizer produces the story text from the rhetorical dependency graph. The main operations that are executed in each module are given in figure 7.1.

The main input to the Narrator Agent is the fabula structure. If we want to adapt the Narrator Agent we have to start focusing on the fabula structure which contains the emotional states. The fabula consists of plot elements (e.g. events, beliefs, perceptions) that are connected by different types of causal relations. The emotions of the character agents are included as a plot element if they have a causal relation to any other plot element. They can roughly be connected to two other kinds of plot elements: events, when events lead to certain emotions via the agent's perceptions, and actions, when emotions lead to pursuing corresponding goals. The emotions in the fabula structure are therefore relevant elements of the story and we aim to express all emotions that are included in the fabula structure.

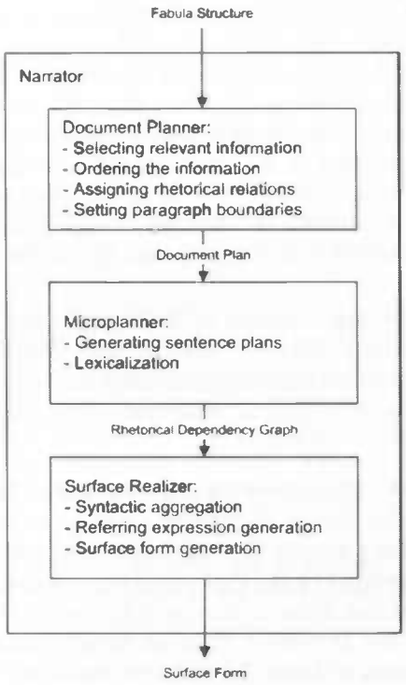


Figure 7.1: The Narrator Agent

Now that we have revisited the Narrator Agent and its main input, we propose how the three components of the Narrator Agent should be changed in order to express the emotions in the plot. Section 7.1 describes how the emotions in the fabula structure can be expressed in an explicit way using modifiers. Section 7.2 discusses how contrast relations can be constructed for expressing emotions that were included in the fabula structure and emotions that were not.

7.1 Expressing emotions using modifiers

Before presenting the main changes on the three different components of the Narrator Agent, we discuss how the fabula structure should be adjusted. Currently, the emotions in the fabula structure are not marked by their intensity values. Because the emotions are to be expressed in a differentiated way in the story, the Narrator Agent needs the intensity values of the emotions. This requires some adjustments to the fabula structure.

7.1.1 The fabula structure

Emotions are a subclass of internal states, just like physical states and beliefs. These relations between concepts are defined in the story world ontology. In the fabula structure, there is a fixed element structure for internal states. This structure is the internal element, which has the following arguments: a *time* argument that specifies at what time the internal state is experienced, a *character* argument for the character that experiences the internal state and a *contents* argument to specify what the internal state entails. The internal states are represented in OWL (the Ontology Web Language).[1] An example of an emotional state represented in OWL, is the following:

```
rdf:type emotion12 Fear
fabula:time 7
fabula:character emotion12 princess2
```

The *contents* argument is empty for emotional states. For belief elements, the argument refers to other states like actions or goals and it thereby specifies the contents of the belief. This *contents* argument is redundant for emotional states, because their contents are the emotions themselves. Beside this redundancy, the current structure misses some arguments for emotions. The emotions that are directed to other agents (like love and admire) require some object. And, for all emotional states, the intensities of the emotions should be added to the structure, so that the Narrator Agent can access these. Because of these differences between the emotional states and the belief elements, we propose splitting the internal state in two different structures.

The structure for belief elements remains as it was. The structure for emotional states has to be expanded with an argument for the object and the intensity. The object argument is called *target*, and it specifies to what agent or object the emotion is directed. The intensity of the emotion is saved in an *intensity* argument, which contains a value between -100 and 100. Because the value of the intensity already expresses whether the experienced emotion is positive or negative, the emotion type argument refers to one of the nine emotion pairs, rather

[1] A description can be found at http://www.w3.org/TR/owl-features/

than to one of the eighteen emotion types. We then get the following structure for emotions that represents here *The princess is afraid of the villain*:

```
rdf:type emotion13 hope/fear
fabula:time 8
fabula:character emotion13 princess2
fabula:target emotion13 villain1
fabula:intensity emotion13 -80
```

The *intensity* indicates which of the two emotions of the emotion pair is experienced. In this example, the emotion fear is experienced with an intensity of 80. The *target* is optional, because self-reflecting emotions do not have an object. The other arguments all get a value for each emotional state. We now discuss how the Document Planner processes the fabula structure.

7.1.2 Document Planner

The fabula structure is first converted into a document plan by the Document Planner. This document plan contains all plot elements that will be expressed in the story. The plot elements in the fabula structure are represented by plot element structures in the document plan. All emotions in the fabula structure are included in this document plan. Currently the plot element structure for internal states is used to represent emotions and it contains five Strings: the type of element, the name, the agens, the patiens and the target. In order to express the intensity of the emotion, we propose adding an Integer to this structure that copies the intensity value of the emotion in the fabula structure. This Integer, also called *intensity*, is left empty for the internal states other than emotions. Figure 7.2 shows an example of the new plot element structure which now contains six arguments.

"De prinses was bang voor de schurk."
("The princess was afraid of the villain.")

type: state
name: hope/fear
agens: princess
patiens: null
target: villain
intensity: -80

Figure 7.2: Example of an internal state

In order to bring about some variation in how the internal states are expressed, the Document Planner contains a state transformer, which combines internal states with connected actions into one or two sentences. There are five ways to combine for example the internal state *the princess was scared* with the action *the princess went to the forest*:

1 *The princess was scared. She went to the forest.*

2 *The princess, who was scared, went to the forest.*
3 *The scared princess went to the forest.*
4 *The princess went to the forest with a pounding heart.*
5 *The princess went to the forest. Her heart was pounding in her throat.*

The state transformer selects one of these five descriptions at random. The first and the second description leave the two plot elements unchanged, where the second only changes the relation between the plot elements into an elaboration relation. For these two descriptions, we can express intensities of the adjective by using modifiers. Examples of modifiers are *een beetje* (*a little*), *erg* (*very*) and *not erg* (*not very*). How modifiers are used is explained in section 7.1.3. The other three ways are rather unnuanced descriptions of the emotion. Using modifiers in the third description is unnatural, since we cannot easily modify the adjective in this sentence and the fourth and fifth descriptions use complete sentences (that are stored as canned text in an additional database) to express the emotions. We prefer using one of these undifferentiated descriptions for default intensities where we would normally not use any modifier. Section 7.1.3 shows that these are intensities between -60 and -90 or between 60 and 90.

So, the random selection mechanism of one of the five possibilities is replaced as follows. For intensities between -60 and -90 and between 60 and 90, one of the five ways is chosen at random. For the other intensities, a random choice is made between the first and the second description, to which modifiers can be added to express the intensities. The additional database is expanded with a phrase for each emotion in the emotion model that expresses a high intensity of that emotion. By adding these phrases to the database we can vary more in the expression of emotions.

7.1.3 Microplanner

The Microplanner replaces all plot elements in the document plan by dependency trees. This is where the main part of the lexicalization takes place. For the construction of emotional expressions, this part of the Narrator Agent has to be adjusted.

For the straightforward expression of the emotions of the character agents, we selected lexical items to refer to the emotions and we investigated how modifiers can be used to express the intensities of the emotions. Table 7.1 shows the lexical items that have to be added to the lexicon in the Narrator Agent.[2] The first mentioned lexical items are the main items to refer to the emotion. The others are used as synonyms. A '≈' is used to indicate that the meaning of the synonym slightly differs from the meaning of the main item. The paragraph below describes how these items are used in the storyteller. The lexicalizer prevents word repetition by selecting (non-differing) synonyms if necessary. This mechanism was already implemented by Slabbers (2006).

We see that the lexical items in table 7.1 concern adjectives, nouns and verbs (indicated between brackets). The required prepositions are shown in *italics*. The Microplanner of the Narrator Agent uses templates to convert plot elements into dependency trees. The template

[2]We deleted the antonym pairs {mooi vinden, afstotelijk vinden} and {verliefd, afkerig} from this list, because the pairs are too specific to refer to the liking/disliking emotion pair and we are not sure when they apply to a situation. The lexical items *houden van* and *haten* could in fact be used to refer to any liking and disliking emotions respectively so we prefer using these lexical items. Furthermore, we deleted the token *weemoedig*, because we find its interpretation to distinctive from *bedroefd*.

Emotion label	**Expressions**
Happy for	blij *voor* (adjective)
Resentment	afgunstig *op* (adjective)
Gloating	leedvermaak *over* (noun)
Pity for	medelijden *met* (noun)
Hope (self)	hoopvol (adjective)
Fear (self)	bang (adjective), angstig (adjective)
Hope (others)	verlangen *naar* (verb)
Fear (others)	bang *voor* (adjective)
Joy	blij (adjective)
Distress	bedroefd (adjective), treurig (adjective), ≈ verdrietig (adjective)
Pride	trots (adjective)
Shame	zich schamen (verb)
Admiration	bewonderen (verb)
Reproach	minachten (verb), ≈ verachten (verb)
Liking	houden *van* (verb)
Disliking	haten (verb)
Gratitude	dankbaar *aan* (adjective)
Anger	boos *op* (adjective), ≈ kwaad *op* (adjective)

Table 7.1: lexical items selected for the Virtual Storyteller

for emotions, the internal state template, only accepts plot elements with an adjective or noun as label and it does not allow for plot elements with objects, like emotions that are directed to other agents. The state template therefore has to be expanded.

For the expression of emotions with adjectives and nouns, the internal state template is used, but the predicate node contains an ap (or an np for nouns), rather than just the emotion name. Figure 7.3 shows the template for these emotions and other internal states. The ap consists of a modifier-node (the intensity of the emotion), a hd-node with a name (referring to the emotion pair) and a pc-node with a pp. The pc-node is optional and it is used to express the person or object to which the emotion was directed. The preposition of the pp is determined by the adjective and it is found in the lexicon. So, for example, the preposition for *boos* is *op*. Emotions that are expressed by nouns can use the internal state template as well, with the hd-node *heeft* (*has*) and the np as predicate-node. Although the pos-tags for these expressions are different, the construction of these expressions can be conducted in the same way. With this template, emotional expressions like *Amalia heeft medelijden met de prins* (*Amalia feels pity for the prins*) and *Amalia is trots* (*Amalia is proud*) are constructed. The state template can also be used to construct other internal states than emotions, such as *Amalia heeft honger* (*Amalia is hungry*). The modifier-node and the pc-node are then left empty.

For the emotions with a verb as part of speech tag, a new template is designed which is very similar to the action/event template. Figure 7.4 shows this template. The head-node contains the emotion pair (verbs), the subject-node refers to the character experiencing the emotion, the modifier contains the intensity of the emotion and the pc-node consists of a preposition (determined by the emotion) and the target. An example of an emotional

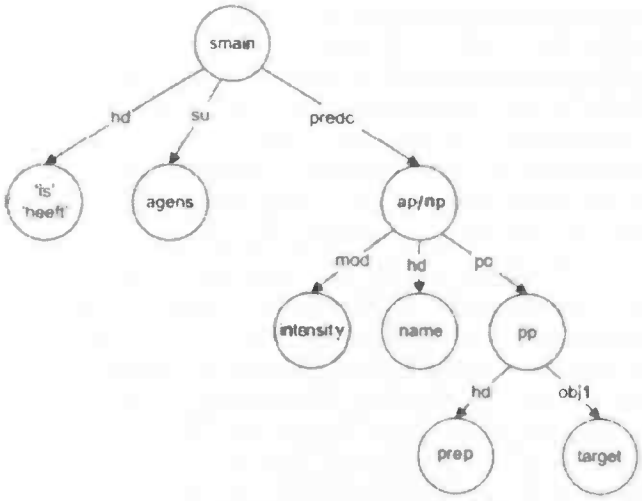


Figure 7.3: plot element for internal states with an adjective or a noun

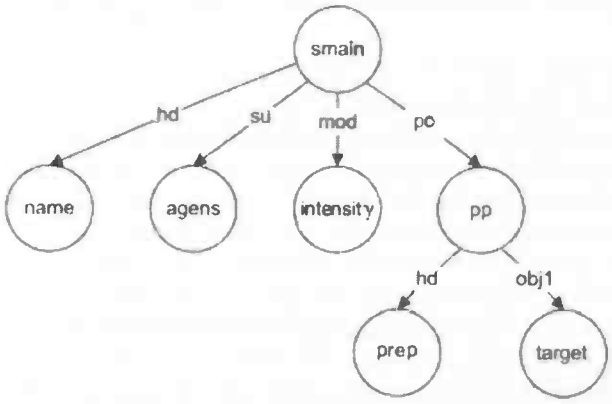


Figure 7.4: plot element for emotional states with a verb

expression, constructed by this template is *Amalia houdt erg van de prins* (*Amalia loves the prins very much*).

Selection of emotional expressions

Once the templates have been filled in with concepts, the lexicalization is performed which maps each concept in the dependency graph onto a Dutch word using the lexicon. This word-for-word lexicalization is not appropriate for the emotional expressions, because the modifiers and adjectives are interdependent. Figure 7.5 shows the emotional expressions for the ⟨blij, bedroefd⟩ scale, based on the results of our empirical study (see chapter 5). The upper expressions are suitable to all emotion types, since their intensities depend on the used modifiers only. The lower expressions are conversational implicatures and they are therefore only used by the antonym pairs whose scales correspond to the scale of the {blij, bedroefd} pair. The intensity boundaries on the scale were chosen intuitively.

We see that the choice of the adjective (positive or negative) depends on the used modifier. So, for example, an intensity of -50 can be expressed by either *een beetje bedroefd* or by *niet blij*. In the first case, the Microplanner should select the negative adjective, in the second case the positive adjective should be selected. The lexicalization of the dependency tree should therefore be preceded by a function for composing the emotional expression, by selecting a modifier and an emotion name (the positive or the negative adjective of the emotion pair). This function takes as input the emotion pair and the adverb represented by an intensity value. For the emotion pairs hope/fear(self), hope/fear(others), happy-for/resentment, gloating/pity and admiration/reproach, the function selects one of the upper expressions, depending on the intensity value of the adverb. For the other emotion pairs, the lower expressions can also be selected (with a fifty percent chance of being selected) if they correspond to the intensity value of the emotion. If the intensity value lies between -15 and 30, the emotion will not be expressed at all and the plot element will be removed from the tree. When the expression is constructed, the lexicalizer is called to map the selected expression onto Dutch words. Both the lexical items and the modifiers in figure 7.5 are to be added to the lexicon.

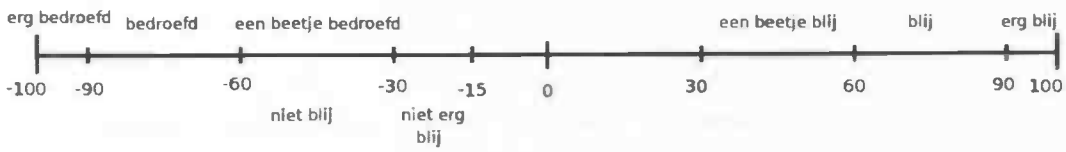


Figure 7.5: Scale of expressions

Table 7.1 contains some synonyms that were interpreted differently (more positively or more negatively) than the main lexical item for the same emotion. Based on the results of our empirical study, described in chapter 5, we calculated intensity values for these synonyms. The difference between the synonym and the main word was scaled to the intensity scale of -100 to +100, and we defined an intensity range for each synonym of 10. We obtained ranges [-77, -86], [-90, -99] and [-78, -87] for respectively the synonyms *verdrietig*, *kwaad* and *verachten*. These words have a fifty percent chance of being used, if the emotion intensity falls in the

specified range.

Example
A plot element contains the subject 'princess', the predicate 'gratitude/anger' and the modifier '-95'. This can be expressed in two ways: *Amalia is erg boos* (50% chance) and *Amalia is kwaad*, with 50% chance for each expression. Figure 7.6 shows the output of the Microplanner for the first expression.

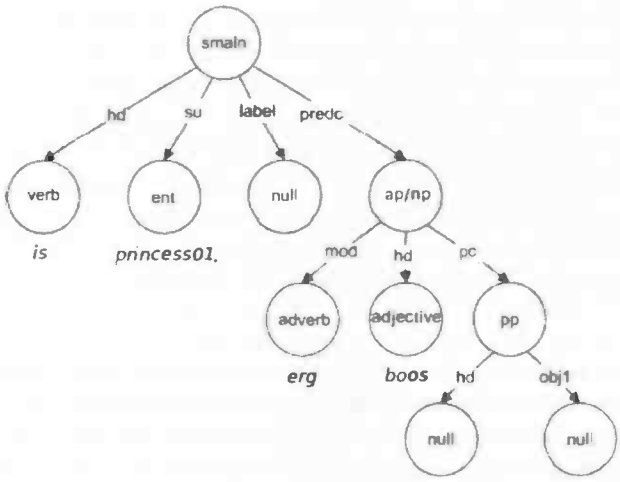


Figure 7.6: plot element for the expression *Amalia is niet blij*

7.1.4 The Surface Realizer

The dependency graph created by the Microplanner is converted into final text by the Surface Realizer. The first module, the Syntactic Aggregator, transforms the rhetorical relations in the dependency trees into text and combined trees. For cause relations, two trees can be combined into one by adding the cue word *omdat* (*because*). But, the two trees can also be generated separately by adding the cue word *daarom* (*therefore*) as an adjunct to one of the two trees or by simply leaving the cue word out. If emotions in the fabula structure have a causal relation to an action, the Surface Realizer currently uses the cue words *zo...*, *dat* to express this relation like in *Amalia was zo bang, dat ze schreeuwde* (*Amalia was so scared that she screamed*). These cue words will still be used if the intensity of the emotion is higher than 80 or lower than -80. The modifier label (which is empty or contains the adverb *erg*) will then be replaced by the modifier *zo* and the two dependency trees are combined by the cue word *dat*.

The second module of the Surface Realizer, the referring expression generator executes general transformations which are independent of the type of plot element. This module need not be adjusted. The third and last module is the surface form generator who decides in which order the nodes in the dependency trees should appear in the text. To this end, the grammar libraries are used that contain the necessary rules to order the nodes. For processing the new

plot elements, the following rules have to be added to the grammar:

```
SMAIN -> SU HD MOD PC
AP -> MOD HD PC
NP -> MOD HD PC
```

In each of these rules, PC stands for prepositional complement. The first rule is used to construct expressions with emotion verbs. The hd-node contains the emotion. This rule already existed in the library files without the modifier argument. The modifier argument is added to express the intensity of the emotion and it directly succeeds the emotion. The second rule describes how the ap-node that contains an emotion adjective is processed. This rule was also already in the library file, but here the modifier argument is present to express the intensity. For adjectival phrases the modifier directly precedes the emotion. Finally, the third rule is for noun emotions. This rule is similar to the rule for adjectival phrases, because these phrases have the same word order. Comparing, for example, *Amalia is trots op de prins* (*Amalia is proud of the prince*) to *Amalia heeft medelijden met de prins* (*Amalia feels pity for the prince*), we see that the head verb *is* (*is*) is replaced by the verb *heeft* (*has*) and that the adjective *trots* (*proud*) is replaced by the noun *medelijden* (*pity*), but the order of words is similar.

With the current approach, all emotions in the fabula structure, with an intensity lower than -15 or higher than 30 are expressed in the story in a differentiated way. Modifiers are used to communicate the intensity of the emotions to the readers. Due to the use of complex expressions and simple expressions, and due to the synonyms in the lexicon and the complete phrases that are stored in an additional database for very high emotion intensities, the Narrator Agent produces varied emotional expressions. The expression of emotions contributes to a lively and realistic story and we expect that the reader gets more involved. The reader understands more about how the character agents feel and why they behave in certain ways.

7.2 Constructing Contrast Relations

The proposal to implement contrast relations in the Narrator Agent is split up in the two types of contrast relations: semantic opposition and denial of expectation. We give a description of both implementations.

7.2.1 Semantic Opposition

In chapter 6 we have defined some restrictions for semantic opposition and we now propose how we can use semantic opposition to express emotions in the Virtual Storyteller. Our main goal of this thesis is to express the emotions of the character agents of the Virtual Storyteller in a natural and accurate way during the story. The semantic opposition can contribute to this goal, by emphasizing that two emotions are contrastive. As described in chapter 6, two properties can be coordinated by the semantic opposition if they are mutually exclusive in the given context. The two elements therefore need not be semantically contrastive, as long as they exclude each other in the given context.

Before explaining how semantic opposition can be implemented in the Virtual Storyteller we give an example of semantic contrast that would be appropriate in a story produced by the storyteller.

(7.1) Amalia is sad, but Alexia is happy.

This sentence directly compares the emotions of Amalia and Alexia which are semantically contrastive. Amalia experiences the negative emotion of the joy/distress emotion pair and Alexia experiences the positive emotion. The subjects of the conjuncts of this *but*-coordination are *Amalia* and *Alexia*, two individuals that are subsumed by the class 'persons'. The semantic opposition requires that the two subjects are related, which is in fact true for all character agents since they are subsumed by the class 'persons'. However, constructing a *but*-coordination like 7.1 is only useful if it contributes to an understanding of the setting. We therefore require that, in order to construct a *but*-coordination, the subjects must have something in common that makes them comparable subjects. The two subjects must have a relation with the current context. Which exact relation is required is hard to determine at this moment and this is beyond the scope of this research. However, it is an important issue because the use of contrast should really contribute to an understanding of the story and one should take care that it does not lead to an isolated irrelevant expression.

How can semantic contrast between emotions be detected by the Narrator Agent? We propose to first look at the emotions in the fabula structure as these emotions are causally related to other aspects of the plot and are therefore relevant emotions. Only emotions with high intensities (above 60 or below -60) are selected, because we want to compare between emotions on either the positive or the negative end of the scale. The Narrator Agent then has to start searching for an emotion that is contrastive to the emotion in the fabula structure. Currently the Narrator Agent can only access the emotions that are included in the fabula structure. In order to search for contrastive emotions, we propose logging all emotions that are experienced during the plot and provide access to this log-file to the Narrator Agent. The Narrator Agent can then send a query to the log-file that specifies the emotion that the agent searches for. For example, the Narrator Agent can specify the time and the agens argument of the emotion. The Narrator Agent then receives the emotions that meet the demands of his query and can construct a *but*-coordination, combining the emotion from the log-file and the emotion from the fabula structure.

Now that we have seen an example of a semantic opposition we propose how the semantic opposition can be constructed by the Narrator Agent. We discuss the two types of contrast: contrast on the same emotional scale in which the two emotions refer to the opposite ends and contrast on different (emotional) scales.

Contrast on the same scale

The emotion model of the character agents contains nine emotion pairs. Six of these reflect an agent's feeling towards another agent or object, the other three reflect feelings towards the agent himself (or towards his goals). Two emotions are contrastive if the one regards the positive emotion of an emotion pair and the other regards the negative emotion of the same emotion pair. The emotions that are experienced within one agent can only be contrastive if they reflect feelings towards some other agent or object. For example, an agent can experience both the contrasting emotions love and hate, as long as they are directed to different persons. But an agent cannot experience joy and distress at the same time, as he experiences either the positive or the negative emotion of this emotion pair.

For the three self-reflecting emotion types joy/distress, hope/fear and pride/shame we must compare between feelings of different agents in order to construct a semantic opposition

with semantic contrast. Sentence 7.1 already gave an example. For the other six emotion types, which reflect feelings towards other agents, *but*-coordinations can be constructed in different ways. Assuming, as an example, *John loves Mary*, we can construct three different *but*-coordinations with the antonym emotion *hate*.

(7.2) Amalia loves Alexia, but Amalia hates Brutus.

(7.3) Amalia loves Alexia, but Brutus hates Alexia.

(7.4) ? Amalia loves Brutus, but Alexia hates William.

(7.5) Amalia loves Alexia, but Alexia hates Amalia.

In 7.2, *Alexia* and *Brutus* are the stressed 'subjects' (in the sense of Spooren's analysis) that are compared on their relation with Amalia. They are subsumed by the same class ('persons') and the properties *Amalia loves* and *Amalia hates* are semantically contrastive for these elements, due to the antonyms *love* and *hate*. 7.2 is therefore a correct *but*-coordination. If one of the emotions is included in the fabula structure, the Narrator Agent constructs this *but*-coordination if it searches in the log-file for an emotion with the same characters, but with the opposite emotion and with a different target. Besides, it is required that Amalia, Alexia and Brutus all play a role in the context.

In 7.3, *Amalia* and *Brutus* are the stressed subjects that are compared on their relation with *Alexia*. They are subsumed by the same class ('persons') and the properties *loves Alexia* and *hates Alexia* are again semantically contrastive. 7.3 is therefore also accepted as a *but*-coordination. For this *but*-coordination the Narrator Agent needs to search for an antonym emotion with a different target, and the same agens. Here, it is required that Amalia and Brutus are related to the current context as they are the compared subjects.

7.4 is a bit different from the other two constructions. Here, both *Amalia* and *Brutus* are replaced by different persons in the second conjunct. Either of them can be interpreted as the stressed subject. Assuming Amalia and Alexia are the subjects, the comparing properties are their relation towards some other man. *Loving Brutus* and *hating William* are not semantically contrastive. However, we can imagine some context (e.g. Brutus being Amalia's husband and William being Alexia's husband) in which they are contrastive (e.g. *Amalia loves her husband* and *Alexia hates her husband*). This type of contrast is rather difficult to construct in the Virtual Storyteller, because it depends on the person-person and person-object relations in the story. These relations can change during the story. So, in order to detect that a *but*-coordination of this kind is possible, the Narrator Agent should check repeatedly for relations during language generation. This function is too complex for the current architecture of the Virtual Storyteller.

Although 7.5 is much like 7.4, it is allowed as a semantic opposition. This is because the subject and the object in the first proposition are swapped in the second proposition. As a result, the relation between the two is similar in the two propositions. Thus, the compared subjects are *Amalia* and *Alexia* and they are compared on their relation with each other. For this semantic opposition, the Narrator Agent only has to search for an antonym emotion for which the character is the target and for which the target is the character and for which the emotion is the opposite. There is no requirement on how the character and the target are related to the context, since the first proposition, which was included in the fabula structure, already points out that there is a relevant relation between the two agents. This is indicated by the fact that one agent has an emotional feeling towards the other agent.

Concluding, for the six emotions that reflect feelings towards other agents, *but*-coordinations are constructed like 7.2 and 7.3 and 7.5 in which either the functional objects differ, the functional subjects differ or in which they are swapped. Figure 7.7 summarizes the required queries to construct a *but*-coordination with corresponding examples. The state element on top represents a self-reflecting emotion, the element below represents an emotion directed to some other agent. Only the variables that differ from the state element are mentioned in the query. So, for all queries, the time variable equals the time variable in the state element. This is correct as we always search for an emotion that was experienced simultaneously to the emotion in the fabula structure. The intensity of the emotion in the query should be opposite to the emotion in the internal state and it should be a number between 60 and 100. This is abbreviated by $-x$.

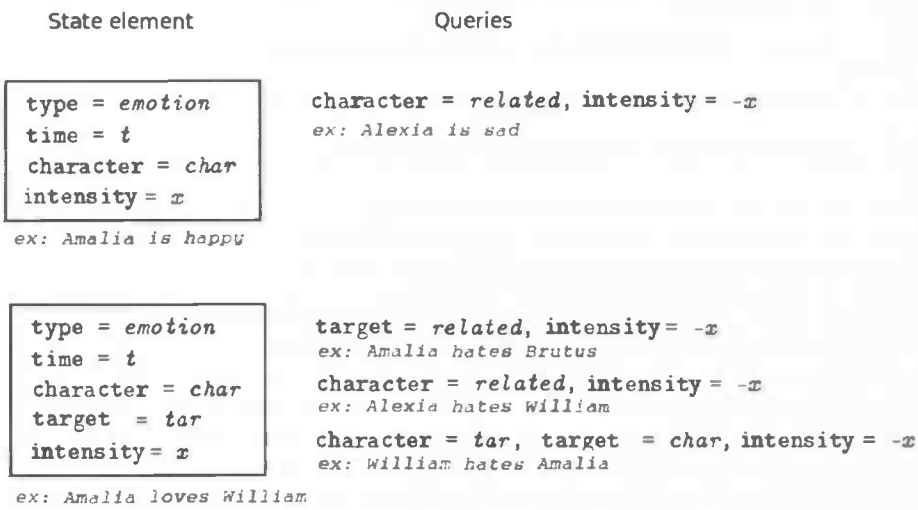


Figure 7.7: state elements with corresponding queries for the log-file

Contrast on different scales

This type of semantic opposition concerns *but*-coordinations in which the two properties of the conjuncts are mutually exclusive in the given context, but they are not opposite values on the same scale. An example is *John is tall but George is fat.* These two subjects have semantically quite different properties but they can be related by the use of *but.* The properties *tall* and *fat* must mutually exclude each other if this is a correct *but*-coordination. The reader will derive context information, namely that John is not fat and that Charlie is not tall. This conclusion can easily be made, because the two properties have a semantic relation; they are both exterior characteristics. We can thereby imagine that the one excludes the other.

When can we make such a *but*-coordination with emotional expressions? For the self-reflecting emotions of the Virtual Storyteller, we can make constructions like 7.6. Here, the hearer interprets that the emotions are mutually exclusive in the context. Following Gricean Quantity maxim, the speaker would not have expressed 7.6 solely if he or she knew that

Amalia was sad or that Alexia was scared. Thus, for such type of contrast, the hearer can assume mutual exclusivity.

(7.6) Amalia is scared, but Alexia is sad.

What are the restrictions for this type of *but*-coordination? The emotions (properties) need to differ, but there are no other restrictions on their relation. As we see, their polarities can be similar and need not be contrastive. The contrast arises, because the reader concludes that both persons experience only one emotion. If the latter is correct, a *but*-coordination like 7.6 can be constructed.

For the other emotions which are directed towards other persons or objects, we can vary more (7.7 to 7.10).

(7.7) Amalia is afraid of Brutus, but Amalia hates William.

(7.8) Amalia is afraid of Brutus, but Alexia hates Brutus.

(7.9) Amalia is afraid of Brutus, but Brutus hates Amalia.

(7.10) Amalia is afraid of Brutus, but Alexia is afraid of William.

7.7, 7.8 and 7.9 correspond to the constructions 7.2 7.3 and 7.5 respectively, but here we did not use antonyms as properties. These constructions are accepted as *but*-coordinations, as long as they exclude each other in the context, thus e.g. Amalia does not hate Brutus and Alexia is not afraid of Brutus in 7.7. In 7.10 the contrasting properties are *being afraid of Brutus* and *being afraid of William*. These properties are not values on the same scale (unless the context provides such a scale) but they do differ in the value of one element namely the person to which the fear emotion is experienced. Such properties can therefore be connected by the marker *but*, as long as their mutual exclusivity is correct in the context. So, e.g. when Amalia is not afraid of William and Alexia is not afraid of Brutus.

A construction like 7.11 in which all variables are substituted in the second conjunct will not be used in the Virtual Storyteller (just like 7.4). This construction only becomes felicitious if we have an explanation for the constrast between *being afraid of Brutus* and *hating William*. Furthermore, 7.12 cannot be a semantic opposition, because the properties *William* and *ice cream* are not subsumed by the same class and also because *likes* is different in meaning in the two conjuncts.

(7.11) ? Amalia is afraid of Brutus, but Alexia hates William.

(7.12) ♯ Amalia likes William, but Alexia likes ice cream.

Constructions like 7.7, 7.8, 7.9 and 7.10 can be used in the Virtual Storyteller, taking into account the restriction that the compared properties are mutually exclusive. The queries for the Narrator Agent correspond to the queries that were given in figure 7.7, except for the type of contrast. Figure 7.8 presents the queries with corresponding examples. The emotion types in the queries are different from the emotion in the state element. When an emotion element is retrieved from the database, the Narrator Agent must check that the database contains no element similar to the retrieved element but with the same emotion as in the fabula structure. So, for example, when the state element in the fabula represents *Amalia is scared* and a retrieved contrasting emotion from the database represents *Alexia is happy*, the Narrator Agent must make sure that the database contains no element like *Alexia is scared*.

State element | Queries

type = *emotion*
time = *t*
character = *char*
intensity = *x*

ex: Amalia is happy

type = other than *emotion*, character = *related*, intensity = [-100, -60] or [60, 100]
ex: Alexia is scared

type = *emotion*
time = *t*
character = *char*
target = *tar*
intensity = *x*

ex: Amalia loves William

type = other than *emotion*, target = *related*, intensity = [-100, -60] or [60, 100]
ex: Amalia fears Brutus

type = other than *emotion*, character = *related*, intensity = [-100, -60] or [60, 100]
ex: Alexia fears William

type = other than *emotion*, character = *tar*, target = *char*, intensity = [-100, -60] or [60, 100]
ex: William fears Amalia

character = *related*, patiens = *related*, intensity = *x*
ex: Alexia loves Brutus

Figure 7.8: state elements with corresponding queries for the log-file

Connecting the two state elements

If the query to the log-file results in the retrieval of a contrastive state element, the Narrator Agent has to add this state element to the document plan so that it will be expressed in the story. In the document plan, the causal relations in the fabula structure are replaced by different kinds of rhetorical relations. So, the Document Planner has to connect the new internal state element to the contrasting emotion element from the fabula structure. The two plot elements are then connected by a contrast relation. In the Surface Realizer such a rhetorical relation can be expressed explicitly by the cue words *maar* (*but*) and *echter* (*however*) (example 7.13 and 7.14). The Surface Realizer can also choose not to express the relation explicitly, as in example 7.15.

(7.13) Amalia is bang, maar Alexia is blij. (Amalia is scared but Alexia is happy)

(7.14) Amalia is bang. Echter, Alexia is blij. (Amalia is scared. However, Alexia is happy)

(7.15) Amalia is bang. Alexia is blij. (Amalia is scared. Alexia is happy)

7.2.2 Denial of expectation

We now explore how a denial of expectation relation fits into the expression of emotions of the Virtual Storyteller. As discussed in chapter 6, there must be some presupposition associated with the emotions and we discuss whether the associated presuppositions raise a weakening effect of the first conjunct.

In the Virtual Storyteller, emotions result from beliefs and personalities and they indirectly determine an agent's behaviour. In chapter 2 we described how emotions are triggered by events and how an agent's behaviour is influenced by his emotions. The chain from events to behaviour (or actions) exists of some default implications, which can be adjusted by an agent's personality. For example, when a character agent sees a dragon, the agent would normally flee. However, some very brave agents will not flee when they get scared. These unusual

behavioural consequences lend themselves to being expressed with a denial of expectation use of *but*, because a defeasible inference is associated with the emotion, which is denied by the agent's behaviour.

Focussing on denial of expectation in expressing emotions, we can discern two kinds of *but*-coordinations: the emotion can be expressed in the first conjunct, and it can be expressed in the second conjunct. If the emotion is expressed in the first conjunct, the second conjunct expresses an unexpected action. If the emotion is expressed in the second conjunct, the emotion itself is unexpected given a certain belief or event. We now discuss how both coordinations can be constructed in the Virtual Storyteller. For both coordinations we restrict ourselves to denial of expectations with direct contrast. Thus, the second conjunct (defeasibly) denies some expectation of the first conjunct.

Emotional expression as first conjunct

As discussed at length in chapter 6, generating a complex sentence with *but* can evoke a denial of expectation interpretation where the first conjunct defeasibly implies some proposition which is in conflict with the second conjunct. If this first conjunct contains an emotional expression, we have to find out what sorts of conclusions can be defeasibly implied. Taking into account the event-action chain, in which the emotion model was inserted, emotions lead to action tendencies which lead to the selection of a goal which leads to certain actions. Action tendencies are not expressed in the story, but the pursued goals and the executed actions are. So, we can construct a denial of expectation with an emotional expression as first conjunct and an unexpected pursued goal or executed action as second conjunct.

In what circumstances is a pursued goal or executed action unexpected? Emotions can result in strange actions or goals if the emotion-action chain deviates from default transitions. There are two possible deviations, shown in figure 7.9.

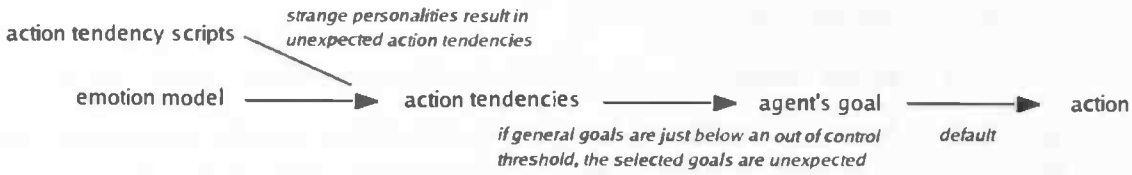


Figure 7.9: The emotion-action chain, illustrated with potentially unexpected consequences

The action tendency script is combined with the emotion model to derive action tendencies (behavioural features). An action tendency script is part of an agent's personality. It determines how an agent's emotions are mapped to action tendencies. There are more or less default action tendency scripts and an agent's personal script may deviate from these. For example, usually the joy emotion influences the action tendency 'friendly' positively. Yet, a villain may behave more unfriendly as he experiences joy. Deviating action tendency scripts result in unexpected action tendencies. If these action tendencies finally determine the agent's actions, these actions can be expressed by the denial of expectation, because they are also unexpected with respect to the emotions of the agent.

Action tendencies only determine an agent's actions when they result in an out-of-control goal. Out-of-control goals are general goals with such high importance values that they are pursued by the agent and replace the episodic goals. The importance values of general goals are determined by the intensities of the action tendencies. If none of the general goals exceed the out-of-control threshold, the agent continues pursuing his episodic goal with the associated actions. Such actions reflect none of the agent's action tendencies. Thus, the denial of expectation, resulting from unexpected action tendencies, can only be expressed if the unexpected action tendencies are so intense that they result in out-of-control goals. If so, the denial of expectation will express that an agent's action or goal is in conflict with his emotion. 7.16 gives an example.

(7.16) The villain was scared, but he attacked the princess.

The villain experiences the emotion fear. As distinct from what one would expect, the villain decides to attack the princess. This occurs, because for villains, being scared results in behaving aggressively. 'Attacking' is then the out-of-control goal that the villain pursues, which stems from the action tendency 'aggressive'. This goal (or the action that the agent executes to accomplish his goal) can be expressed using the denial of expectation. The *but*-coordination then illustrates the personality of the villain.

The other transition in figure 7.9 where unexpected conclusions may be derived is where action tendencies are mapped to goals. Goals only listen to action tendencies if they are out-of-control. When the importance value of a goal is just below threshold, the action tendencies are not intense enough to determine the goal. When the action tendencies correspond to the agent's emotions (due to default action tendency scripts), we can conclude that the intensities of the emotions are just too low to result in certain behaviour. For example, the fear emotion of an agent results in the action tendency 'passive'. This action tendency determines the importance of the general goal 'flee'. The fear emotion may be intense but not intense enough to result in fleeing behaviour. We can then speak of the denial of expectation. 7.17 expresses this.

(7.17) Amalia was scared, but she did not flee.

Amalia's behaviour of not fleeing form Brutus is unexpected and can be expressed using the denial of expectation. So, the denial of expectation can be used here if the agent experiences some emotion to a rather high degree, but the agent keeps pursuing his episodic goal.

Does the use of *but* in the constructions given above weaken the first conjunct? In 7.16 the emotion is high enough to bring about an out-of-control goal, so a weakening effect of the emotion is not desirable. In 7.17 the emotion is not high enough to bring about an out-of-control goal. So, here it would be appropriate if the interpretation of the emotional expression was weakened by the use of *but*. But does our approach of a weakening effect account for these different interpretations?

In 7.16 the context must explain the unexpected action of the agent in order to prevent a weakening effect of the expressed emotion since such an interpretation is not intended. Although the reader may know that villains behave aggressively in general, the reader may still be confused about the villain's action and derive wrong conclusions about the villain's emotions. It is therefore desirable to explicitly explain why the villain attacked the princess.

So, we should add some explanation to the text, like *The villain's fear made him behave aggressively.*

In 7.17 the reader has no information about why Amalia chooses to stay near her enemy. In this case, generating a sentence with *but* aids the reader in interpreting Amalia's position on the hope/fear scale as less fearfull than is normally appropriate. The reader should then interpret that Amalia is not scared enough to flee from her enemy.

Concluding, we can construct the denial of expectation for expressing agent's odd reactions (out-of-control goals) to their emotions. Here, we need to specify why the agent behaves in an unexpected way in order to prevent the hearer form interpreting *but* as weakening the interpretation of the emotion in the first conjunct since such an interpretation is not intended. For other examples we can use *but* to express that an agent's emotion is not intense enough to bring about a corresponding out-of-control goal. Specifying that the emotion was not intense enough is not necessary as the reader probably interprets the *but*-coordination as a weakener and this interpretation is desirable.

Emotional expression as second conjunct

We now need to discuss when an agent may have unexpected emotions. In the event-action chain we have to focus on the transitions left of the emotion model. These are specified in figure 7.10 and in the figure we have indicated where the denial of expectation can be applied to express an emotion.

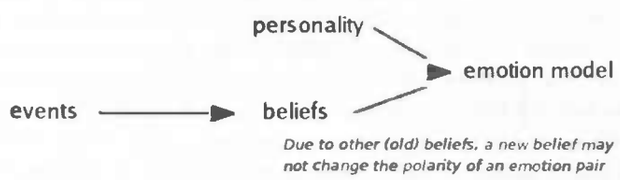


Figure 7.10: The chain from event to emotion

The intensity values in the emotion model are continuously updated on the basis of an agent's beliefs and an agent's personality. Each possible belief in the Virtual Storyteller has been assigned fixed emotion values that describe which emotions are evoked by the belief. The personality parameters of an agent determine for each emotion to what extent it reacts to a belief. The emotional reaction to a belief, which is a numerical value, is added to the old emotion intensity to update an agent's emotion.

The personality parameters cannot cause unexpected emotions to react to beliefs, as the parameters only determine *to what extent* emotions react to beliefs and not *which* emotions react to beliefs. An emotion's reaction to a belief is therefore always as the reader would expect (due to fixed emotion values for beliefs). However, the updated emotion model can deviate from an agent's belief as the old emotions play a role here. For instance, an agent can be happy before a sad event takes place. After this event he may still be happy. This can be expressed by a denial of expectation like 7.18.

(7.18) Amalia saw Brutus with a sword, but she was (still) hopeful.

By using a *but*-coordination, we emphasize that Amalia's emotions do not correspond to her belief. Amalia's belief is that a villain with a sword is in front of her and the reader would expect her to become frightened. The explanation for Amalia's odd emotional reaction is that she still has other reasons to be hopeful. A *but*-coordination can therefore be constructed when the reaction of an emotion is not intense enough to change the polarity of the experienced emotion. We can therefore add adverbs such as *still* as to clarify the contrasting emotions.

A denial of expectation in which an event or belief is contrasted with an emotional reaction does not bring about a weakening of the first conjunct. The first conjunct contains no gradable expression, so weakening is not possible.

Constructing the denial of expectation

For each of the three ways in which the denial of expectation can be used to express emotions (thus for unexpected reactions to emotions, for no reactions to emotions and for unexpected emotional reactions to beliefs) we now propose how they can be constructed in the Virtual Storyteller.

Unexpected reactions to emotions An example of this kind of denial of expectation was given in 7.16: *The villain was scared but he attacked the princess.* In order to detect contrast between an agent's emotion and his action, there must be some knowledge base for default reactions to emotions. Reactions to emotions are determined by an agent's action tendency script, so the knowledge base has to contain the default action tendency scripts. For each of the eighteen emotions, the knowledge base then indicates which out-of-control goal is expected. For example, it says that from the emotion fear the out-of-control goal fleeing is expected. If an emotion then results in a different out-of-control goal, a differing action tendency script is held responsible for this.

The Document Planner of the Narrator Agent converts the fabula structure into a document plan and replaces the causal relations in the fabula structure by different kinds of rhetorical relations. If emotions result in certain actions the two elements are causally connected in the fabula structure. Emotions only have a causal relation to actions, when they result in out-of-control goals. For these causal relations, the Document Planner needs to verify whether the relation between the emotion and the action is indeed causal. In other words, whether the relation between the emotion and the action is as the reader would expect. It does this by searching for the emotion in the knowledge base and comparing the out-of-control goal (with corresponding actions) to the action or goal executed by the agent. If the executed action deviates from the goal in the knowledge base, the agent's deviating action tendency scripts have resulted in an unexpected action. The cause relation between the emotion and the action is then replaced by a contrast relation in the Document Plan.

As stated above, the unexpected action must be explained to prevent a weakening effect of the emotional expression. This can be done by a sentence like *Zijn angst maakte hem aggressief* (*His fear made him behave aggressively*). Thus, the emotion of the agent is causally connected to the action tendency of the agent. This can be done by adding two plot elements to the fabula structure. One element is an internal state element with the agent's emotion (again), the other element is also a state element with the agent's action tendency. The two are connected by a causal relation and they connect causally to the out-of-control action. This is displayed schematically in figure 7.11. Here we see that the causal relation between the agent's emotion and his action tendency cause the contrast relation between his emotion and his goal.

The problem with this document plan is that the nodes will be expressed in the wrong order: *The villain's fear resulted in aggressive behaviour. Because of that, the villain was scared but he attacked Amalia.* The contrast relation should be expressed first, and then an explicit explanation should be given, but this is not possible with this document plan. Therefore, the cause relation on top is replaced by an additive relation, and the two branches are swapped like in figure 7.12. The additive relation should not be expressed explicitly. The contrast relation will then be expressed first, followed by an explanation of the agent's behaviour. We then obtain the following text passage: *The villain was scared but he attacked Amalia. The villain's fear resulted in aggressive behaviour.*

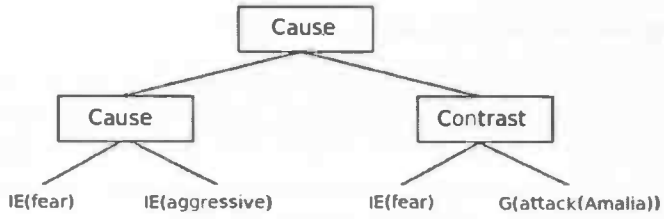


Figure 7.11: Document Plan with rhetorical relations

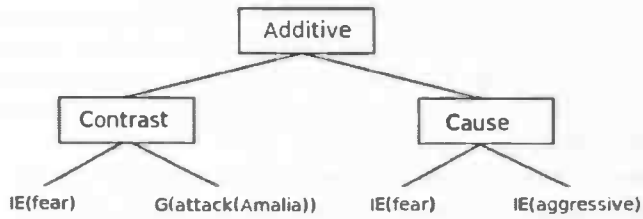


Figure 7.12: Document Plan with rhetorical relations

The Microplanner and the Surface Realizer translate the Document Plan into natural language. Contrast relations can be explicitly expressed by cue words. For the denial of expectation interpretation, the cue word *maar* (*but*) is required (see chapter 6). Therefore, these kinds of contrast relations need to be marked with *maar*, so that the Surface Realizer knows that the cue word *maar* has to be selected.

No reaction to emotions These concern emotions whose intensity values are just too low to result in out-of-control goals. Such contrast relations can be detected easily by the Document Planner. If an emotion with a high intensity is included in the fabula structure, but it does not cause any actions, then contrast can be established between the agent's emotion and his behaviour. To that end, the Document Planner needs to add a new plot element to the Document Plan, which is a negation of the expected out-of-control goal. This out-of-control goal can be found in the knowledge base, presented above, that describes the default reactions to emotions. The action template is used for the new plot element and it is connected to the emotion by a contrast relation.

As written above, the first conjunct may be weakened by the use of *but*. Therefore, emotions with a very high intensity should not be conjoined with *but* and an unexpected action. The reader may then have wrong ideas about the intensity of the emotion. In the current architecture, there can be other reasons for why an emotion does not results in an out-of-control goal (e.g. a low action tendency script value). We propose to only express emotions with an intensity between 60 and 85 (or between -60 and -85) using the denial of expectation. This means that for each emotion in the fabula structure, the Document Planner checks whether the intensity is between 60 and 85 and, if so, whether the emotion is or is not causally connected to any action. If the Document Planner finds an emotion (with an intensity between 60 and 85) that is not causally connected to any action, the Document Planner searches for the expected goal in the knowledge base and it adds a negation of this goal to the Document Plan and connects it to the emotion by a contrast relation. This type of contrast is marked with *maar*, so that the cue word *maar* will be selected by the Surface Realizer to express the relation.

Unexpected emotional reactions to beliefs A denial of expectation interpretation is also desirable if an agent's belief (or perception) is not strong enough to change an agent's emotion. The fabula structure only contains causal relations between plot elements. So, if a belief does not result in a corresponding emotion, the emotion is not added to the fabula structure. The Document Planner can easily detect such beliefs, as they do not result in any other plot element in the fabula structure. However they are included in the fabula structure since they are a reaction to an agent's perception and they therefore connect to the perception plot element.

We wish to only construct contrast relations for beliefs from which the reader would expect a typical emotional reaction. All beliefs in the story world have fixed corresponding emotion values and the beliefs for which certain emotion values are really high can be used in the denial of expectation, because these believes usually evoke typical emotional reactions. If the Plot Agent marks such beliefs as *emotional beliefs* in the fabula structure, the Document Planner can decide whether to construct a denial of expectation. This algorithm works as follows. If a marked belief does not connect to any other element, the Document Planner first consults the knowledge base that contains the default emotional reactions to beliefs. Once the Document Planner knows which emotion is expected to arise, it queries the log-file. The Document Planner sends a query with the same time, same character and the expected emotion type and it will retrieve the intensity of the emotion type that corresponds to the query. If the intensity is opposite to the expected intensity, the emotion is added to the Document Plan with a contrast relation to the belief. This contrast is of course marked with *maar*, so that *maar* will be used to express the contrast. Additionally, the emotion will be modified by *nog steeds* (*still*), so that the reader understands what is going on. An example is the following.

Example The belief of seeing a witch usually strongly decreases the intensity of the hope/fear emotion. It is therefore a so called emotional belief. If this belief is experienced by Amalia at time t, the Plot Agent adds this belief to the fabula structure and the belief is marked as emotional. When the Document Planner converts the fabula structure into a document plan, it detects the marked belief and it checks whether this belief indeed causes some emotion in the fabula structure. In other words, it checks whether Amalia was indeed fearful due to her perception of the witch. If not, the Document Planner queries the database with the hope/fear emotion, the time t+1 and the character Amalia so that it knows the intensity of

the hope/fear emotion just after Amalia believed that she saw a witch. When the intensity is positive, which is opposite to what the reader would expect, the emotion is added to the Document Plan and contrastively connected to the belief. The modifier of the state element will represent the adverb *still*. The contrast will then be expressed as follows: *Amalia saw a witch but she was still hopeful.*

7.3 Conclusion

We have proposed how the Narrator Agent can be expanded in order to express intensities of emotions. Generally, modifiers are used to do this. Straightforward expressions like *I am a bit sad* can be constructed and other more complex expressions can be constructed like *I am not happy*. This last type of expression only applies to antonym pairs for which the interpretations of complex expressions were identical to those of the {blij, bedroefd} pair. In order to express the intensities, the emotions in the fabula structure need to be marked by their intensity values.

We have discussed how the two types of contrast relations, semantic opposition and denial of expectation, can be constructed in the Virtual Storyteller. Both types of contrast relations are useful in expressing emotions. For emphasizing general contrast between the emotions of two characters or between the emotions within one character, a semantic opposition interpretation can be evoked using a number of diffrent contrast markers. The reader is then pointed to the fact that two emotions are contrary and this helps the reader realize that agents feel different or that one agent has different feelings about different characters.

A denial of expectation interpretation can be evoked in situations where the reader has expectations about a character's emotions or behaviour, but these expectations are not realized. There are three cases. The reader expects an agent's emotional reaction to an event, but the agent reacts differently, because the event had only little effect on the agent's emotions. Or, the reader expects that an emotion results in some action, but this action is not executed by the agent. This occurs when the agent's emotion is not high enough to influence its behaviour. By using the contrastive marker *but* to conjoin the emotion and the unexpected behaviour the reader is encouraged to interpret the emotional expression in the first conjunct as weaker than normal. A last possibility is that the agent's peculiar personality results in unexpected actions given its emotions. The use of *but* then illustrates the agent's personality.

The types of contrast relations have their own use in expressing emotions and they would contribute to a natural translation of the emotions in the story into text. We have therefore proposed a way to implement these contrast relations in the Virtual Storyteller.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

In this project we have strived for a natural and correct translation of the emotional states of the character agents into natural language. To this end we have conducted a corpus study on fairy tales, empirical research on conversational implicatures and theoretical research on contrast relations. The most important conclusions are now presented.

Corpus study In a corpus of fairy tales we searched for emotional expressions. The study showed that there is a mismatch between the fairy-tale lexicon of emotions and the OCC model. Some emotions of the OCC model are not explicitly expressed in fairy tales and there is much variety in the frequency of the emotional expressions in fairy tales even within an emotion pair. When constructing an emotion model for a storyteller it is important to add emotions to the model that are relevant emotions in stories and we think that the OCC model is therefore not a good starting point. The OCC model is a complete model of emotions and some of its emotions are uncommon in fairy tales. Existing fairy tales show which emotions of the OCC model should be included in an emotion model of a storyteller. Based on the results of our corpus study, we recommend that the emotion model of the storyteller is expanded with the emotion pair gratitude/anger and we suggest different lexical items, generally gradable adjectives, to refer to the emotions in the model.

Conversational Implicatures For a natural translation of the emotions in the emotion model we suggest the use of complex expressions with gradable adjectives that give rise to conversational implicatures. The interpretations of these expressions have been established by conducting empirical research. We compared the results of our study to the theoretical findings of Blutner (2000) who applied weak bidirectional OT to expressions with antonym pairs. Our conclusions are that weak bidirectional OT can only roughly estimate what the interpretations of conversational implicatures are. Blutner's assumptions about the complexity of meanings and forms need to be adjusted in order to account for our empirical findings. We suggest that the complexity of forms depends on both the number of negation morphemes and the number of words. The complexity of meanings should then increase from negative to positive, because readers interpret complex forms as negative as the alternative forms allow. For the interpretations of complex expressions, we have seen how the intensity scale that the antonym pair refers to influences the interpretation.

Contrast relations When contrast relations are used in the Virtual Storyteller, the emotions of the character agents can be related to other aspects of the story in a contrastive way. Two types of contrast relations can be constructed in the Virtual Storyteller: semantic opposition and denial of expectation. A semantic opposition interpretation is evoked by joining two emotional expressions that are opposed to each other. The emotional expressions can refer to emotions on the same scale or to emotions on different scales. If two emotions constitute an emotion pair in the Virtual Storyteller there is contrast on the same scale. Emotions on different scales can also evoke a semantic oposition interpretation but then it is important that the two emotions that are attributed to the two subjects in the conjuncts are mutually exclusive in the given context. There are no clear restrictions proposed in the literature about the way in which the compared subjects are related. We therefore propose to define some restriction that is suitable to the expression of emotions in storyteller. Our idea is that the two subjects' emotions must be reactions to the same event, action or object.

Denial of expectation can be used to express an unexpected emotional reaction to an event, or to express an unexpected action of an agent given its emotion. Combining Winter & Rimon's definition of denial of expectation with the Asymmetry Hypothesis of Spooren, we predict that in some cases a denial of expectation weakens the interpretation of the first conjunct if this contains a gradable expression. In the storyteller, such an interpretation would be realized if the first conjunct of a *but*-coordination denotes some emotion to an agent in an unnuanced way and the second conjunct expresses that the agent does not behave as would be typically expected from his emotion. A weakening of the first conjunct is then desirable as, basically, the agent's emotion is too weak to result in some corresponding goal.

The results of our three studies can be applied to the Virtual Storyteller. The Narrator Agent, which is responsible for translating the fabula into natural language, needs to be adjusted in order to express the emotions in a nuanced way by using expressions in which the emotions are modified, and in order to construct contrast relations that contain emotional expressions. Generally this means that the emotions in the fabula structure are to be marked by their intensity values and that the emotional states of the character agents during the story should be logged in a database that the Narrator Agent can query. These adjustments provide the Narrator Agent with the required information to construct nuanced emotional expressions and contrast relations.

8.2 Suggestions for future work

In this section we give some suggestions for future research. The research that we suggest elaborates on the research that we conducted in this master's project.

Testing our ideas about weak bidirectional OT

Our empirical study about conversational implicatures was conducted because we wanted to test whether weak bidirectional OT can explain the interpretations and because the interpretations of conversational implicatures needed to be clear in order to use them in the Virtual Storyteller. Due to the existence of these two goals, we had to restrict the number of trials in the experiment for each goal. We admit that our conclusions about weak bidirectional OT are based on a small number of data. That is to say, only the antonym pairs {gelukkig,

ongelukkig} and {blij, bedroefd} were tested in depth and each item (comparison) was presented only once in the test, so that the context sentence of the trial could influence the results to a certain extent.

The question now is whether our conclusions about weak bidirectional OT and the criteria for the markedness of forms and meanings are correct or whether they are an overfit to the data. We therefore suggest to set up a new experiment with the same design as our experiment but in which the expressions with different antonym pairs are tested in depth. A number of these antonym pairs should contain negation morphemes so that general conclusions can be drawn on the effect of a negation morpheme on the complexity of a form.

Another suggestion for empirical research on this topic is to test where the differences in interpretation come from. Do subjects really have personal preferences (possibly related to their linguistic background) for certain interpretations or are they just not sure what the interpretations of certain complex expressions are and does this cause differing data between the subjects? We did not find any correlation between a reader's linguistic background and his or her preferred interpretations, but this could be investigated in more detail.

Empirical research on contrast relations

Our idea is that in contrast relations the interpretation of the first conjunct can differ from the typical interpretation of this conjunct. If the first conjunct contains a gradable expression, the interpretation is weakened in certain circumstances. In chapter 6 we have explained the theoretical foundations of this idea, but we have no empirical evidence for our suggestion. If we want to implement contrast relations in the Virtual Storyteller it is absolutely essential that the interpretations of these expressions are clear, especially in *but*-coordinations where a weakening effect is desirable in the storyteller. Therefore, human readers should be tested on their interpretation of contrast relations and it should be investigated when a weakening effect of *but* is derived. Next to that, it is important to know how weakening of the first conjunct is prevented so that we can apply this in the storyteller in text passages where we wish to do so.

Examining the relevance of emotional contrast in the storyteller

In chapter 7 we have presented how the Narrator Agent can construct contrast relations to express the emotions in the Virtual Storyteller and we have given an overview of the possible realizations of emotions in contrast relations. The question is now when we want to express contrast in the Virtual Storyteller. If each form of contrast in the plot is expressed in the story the reader may be distracted from the real story topic. Currently, mainly causal relations (which are included in the fabula structure) are expressed in the story and these contain the most relevant story aspects. Story aspects that are only contrastively related to the rest of the story can be seen as the story background. It should be examined when such aspects are relevant story aspects and when it is desirable to express these story aspects using contrast relations.

Semantic opposition in particular also requires future research. For the semantic opposition the emotions of two character agents are compared. It must then be specified what the relation between the two compared agents must be. We have already given an example of a relation that would make sense: the two contrastive emotions of the agents should be a reaction to the same event, object, or action. But this requirement seems rather specific. For instance, two agents could emotionally react to each other, in which case an expression of contrastive

emotions would also add useful information to the story. Future research should determine which relation is required between the subjects to express the semantic opposition in the storyteller. By starting to implement contrast relations in the Virtual Storyteller insight may be gained in this issue.

Expanding the emotion model and the Narrator Agent

First, the emotion model needs some adjustments. The pair gratitude/anger has to be added to the model with corresponding connections to events and action tendencies. Next to that, the pairs happy-for/resentment and gloating/pity-for need to be restructured as explained in chapter 3. The adjusted emotion model is a better representation of the emotions that are important in fairy tales.

For expressing emotion intensities and contrast relations the Narrator Agent needs to be expanded and the emotions of the character agents during the plot need to be logged in a database. In chapter 7 we have proposed how this can be done, but this proposal has not been implemented. An implementation of our proposal should be tested and evaluated.

Bibliography

Anscombre, J. & Ducrot, O. (1983), 'Deux mais en français?', *Lingua* **43**, 23–40.

Atlas, J. & Levinson, S. (1981), It-clefts, informativeness and logical form, *in* P. Cole, ed., 'Radical Pragmatics', Vol. 1-61, Academic Press, New York.

Bartneck, C. (2002), 'Integrating the occ model of emotions in embodied characters'. Workshop on Virtual Conversational Characters: Applications, Methods and Research Challenges, Melbourne.

Blakemore, D. (1989), 'Denial and contrast: a relevance theoretic analysis of but', *Linguistics and Philosophy* **12**(1), 15–37.

Blutner, R. (2000), 'Some aspects of optimality in natural language interpretation', *Journal of Semantics* **17**, 189–216.

Blutner, R. (2001), Two case studies in lexical pragmatics, *in* E. Németh & K. Bibok, eds, 'Pragmatics and the Flexibility of Word Meaning', pp. 11–28.

Bouma, G., Noord, G. V. & Malouf, R. (2001), Alpino: wide coverage computational analysis of dutch, *in* W. Daelemans, K. Sima'an, J. Veenstra & J. Zavrel, eds, 'Computational Linguistics in the Netherlands', Rodolphi, Amsterdam, pp. 45–59.

Dascal, M. & Katriel, T. (1977), 'Between semantics and pragmatics: the two types of 'but' - hebrew 'ava' and 'ela", *Theoretical Linguistics* **4**, 143–172.

Dixon, R. (1982), *Where have all the adjectives gone? And other essays in semantics and syntacs*, Mouton, Berlin, chapter Where have all the adjectives gone?

Fodor, J. & Lepore, J. (1999), 'Impossible words?', *Linguistic Inquiry* **30**, 445–453.

Foolen, A. (1991), 'Polyfunctionality and the semantics of adversative conjunctions', *Multilingua* **10**, 79–92.

Frawley, W. (1992), *Linguistic Semantics*, Erlbaum, Hilsdale.

Givón, T. (1984), *Syntax: a functional typological introduction*, John Benjamins, Amsterdam.

Grice, H. (1975), Logic and conversation, *in* P. Cole & J. Morgen, eds, 'Syntax and Semantics', Vol. 3, Academic Press, pp. 41–58.

Hielkema, F. (2005), Performing syntactic aggregation using discourse structures, Master's thesis, University of Groningen, The Netherlands.

Hirschberg, J. (1985), A theory of scalar impliature, PhD thesis, University of Pennsylvania.

Hobbs, J. (1979), 'Coherence and coreference', *Cognitive Science* **3**, 67–90.

Horn, L. (1984), Toward a new taxonomy for pragmatic inference: Q-based and r-based implicatures, *in* D. Schiffrin, ed., 'Meaning form and use in context', Vol. 11-42, Georgetown University Press, Washington.

Lakoff, R. (1971), If's and's and but's about conjunction, *in* C. Fillmore & D. Langendoen, eds, 'Studies in linguistic semantics', New York, pp. 114–149.

Lang, E. (1984), *The semantics of coordination*, Amsterdam: Philadephila: John Benjamins.

Lazarus, R. (1991), *Emotion and adaptation*, Oxford University Press, New York.

Levinson, S. (1983), *Pragmatics*, Cambridge University Press.

Levinson, S. C. (2001), *Presumptive meanings*, The MIT Press, chapter The phenomena.

Lyons, J. (1977), *Semantics*, Cambridge University Press.

Magnenat-Thalmann, N., Kim, H., Egges, A. & Garchery, S. (2005), 'Believability and interaction in virtual worlds', *Proceedings of the 11th IEEE International Multimedia Modelling Conference, 2005*. pp. 2–9.

Mann, W. & Thompson, S. (1988), 'Rhetorical structure theory: Toward a functional theory of text organization', *Text 8* **3**, 243–281.

Ortony, A., Clore, G. & Collins, A. (1988), *The cognitive structure of emotions*, Cambridge University Press.

Prince, A. & Smolensky, P. (1993), Optimality theory: Constraint interaction in generative grammar, Technical report, MS., Rutgers University, New Brunswick and University of Colorado, Boulder.

Raskin, V. & Nirenburg, S. (1995), Lexical semantics of adjectives: A microtheory of adjectival meaning, Technical report, Memoranda in Computer and Cognitive Science, MCCS-95-288, New Mexico State University.

Rensen, S. (2004), De virtuele verhalenverteller: agent-gebaseerde generatie van interessante plots, Master's thesis, University of Twente, The Netherlands.

Schachter, S. & Singer, J. (1962), 'Cognitive, social and physiological determinants of emotional state', *Psychological Review* **69**, 379–399.

Slabbers, N. (2006), Narration for virtual storytelling, Master's thesis, University of Twente, The Netherlands.

Spenader, J. & Stulp, G. (2007), 'Antonymy in contrast relations', Proceedings of the Seventh International Workshop on Computational Semantics, Tilburg, 10-12 January, 2007.

Spooren, W. (1989), Some aspects of the form and interpretation of global contrastive coherence relations, PhD thesis, University of Nijmegen: Sneldruk Enschede.

Swartjes, I. (2006), The plot thickens - bringing structure and meaning into automated story generation, Master's thesis, University of Twente, The Netherlands.

Theune, M., Faas, S., Nijholt, A. & Heylen, D. (2002), 'The virtual storyteller', *ACM SIGGROUP Bulletin* **23**(2), 20–21.

Umbach, C. (2004), 'On the notion of contrast in information structure and discourse structure', *Journal of Semantics* **21**, 155–157.

Umbach, C. (2005), 'Contrast and information structure: a focus-based analysis of *but*', *Linguistics* **34**(2), 207–232.

Wierszbicka, A. (1988), *The Semantics of Grammar*, John Benjamins, Amsterdam, chapter What's in a noun? (or: How do nouns differ from meaning from adjectives?).

Winter, Y. & Rimon, M. (1994), 'Contrast and implication in natural language', *Journal of Semantics* **11**, 365–406.

Appendix A

A produced story

Er was eens een prinses.
De prinses heette Amalia en woonde in het klein bos.
Er was een schurk.
De schurk heette Brutus en bevond zich in het moeras.
Amalia ging naar de woestijn, want zij wilde cactussen plukken.
Brutus ging, om de omgeving te verkennen, tegelijkertijd erheen.
Zij wilde vluchten, want, doordat zij Brutus zag, was zij bang.
Zij ging dus naar de kaal vlakte.
Brutus ging naar de kaal vlakte, want hij wilde Amalia gevangen nemen.
Doordat Amalia Brutus zag, was zij bang.
Brutus ging naar de bergen, want zij ging, om te vluchten, erheen.
Er lag een zwaard daar.
Zij wilde Brutus doden.
Omdat Brutus, doordat Amalia daarom het zwaard oppakte, bang werd, werd hij aggressief.
Hij sloeg Amalia.
Nog lang en gelukkig leefde zij, nadat zij Brutus had neergestoken.

Appendix B

B.1 Fairy-tales of Grimm brothers used in the corpus

1. De Kikkerkoning of IJzeren Hendrik
2. Kat en muis samen thuis
3. Het kind van Maria
4. Sprookje van iemand die erop uittrok om te leren griezelen
5. De wolf en de zeven geitjes
6. De trouwe Johannes
7. De goede ruil
8. De wonderlijke speelman
9. De twaalf broers
10. Het gespuis
11. Broertje en zusje
12. Raponsje (Rapunzel)
13. De drie mannetjes in het bos
14. De drie spinsters
15. Hans en Grietje
16. De drie slangenbladeren
17. De witte slang
18. Strohalm, boontje en kooltje vuur
19. Van de visser en zijn vrouw
20. Het dappere snijdertje
21. Assepoester

22. Het raadsel
23. Van het muisje, het vogeltje en de braadworst
24. Vrouw Holle
25. De zeven raven
26. Roodkapje
27. De Bremer Stadsmuzikanten
28. Het zingende botje
29. De duivel met de drie gouden haren
30. Luisje en vlootje
31. Het meisje zonder handen
32. Slimme Hans
33. De drie talen
34. Knappe Elsje
35. De kleermaker in de hemel
36. Tafeltje dek je, ezeltje strek je en knuppel uit de zak
37. Klein Duimpje, Duimedik
38. De bruiloft van vrouw vos
39. De kabouters
40. De roverbruidegom

B.2 Fairy-tales of Andersen used in the corpus

1. Het meisje met de zwavelstokjes
2. Het lelijke jonge eendje
3. De kleine zeemeermin
4. De nieuwe kleren van de keizer
5. De prinses op de erwt
6. Duimelijntje
7. De Chinese nachtegaal
8. De wilde zwanen

9. De laatste droom van de oude eik
10. De sneeuwkoningin
11. De vogel Feniks
12. De mooiste roos van de wereld
13. De engel
14. De rode schoentjes (naverteld)
15. De tondeldoos
16. Het zwanennest
17. Ib en de kleine Christine
18. Wat vader doet is altijd goed
19. De kranige springers
20. De stopnaald
21. Boekweit
22. Holger De Deen
23. Het is echt waar
24. De sneeuwman
25. Het kaboutertje bij de winkelier
26. De Poppenspeler
27. Alles op zijn plaats!
28. Domme Hans
29. Het herderinnertje en de schoorsteenveger
30. Het geluk kan in een stukje hout liggen
31. De hardloper
32. In de kinderkamer
33. Kleine Klaas en grote Klaas
34. Het madeliefje
35. Het meisje dat op het brood ging staan
36. De oude kerkklok
37. De reisgenoot

38. De rozenelf
39. Het spaarvarken
40. De varkenshoeder

Appendix C

Experiment

C.1 Stimuli

Antonym pair {gelukkig, ongelukkig} (happy/unhappy)

1. *Marie en Sophie zijn beiden onlangs gescheiden*
   Marie: "Ik ben gelukkig." Sophie: "Ik ben niet ongelukkig."

2. *Het huis van Kees en Merel is verkocht.*
   Kees: "Ik ben niet ongelukkig." Merel: "Ik ben een beetje gelukkig."

3. *De contracten van Frans en Kim zijn niet verlengd.*
   Frans: "Ik ben een beetje ongelukkig." Kim: "Ik ben niet gelukkig."

4. *Er komt voorlopig geen einde aan de oorlog.*
   Paula: "Ik ben ongelukkig." Johan: "Ik ben niet gelukkig."

5. *Kees en Merel hebben een meisje gekregen.*
   Kees: "Ik ben niet erg gelukkig." Merel: "Ik ben gelukkig."

6. *Remco en Nicole werden tweede op de Olympische Spelen.*
   Remco: "Ik ben een beetje gelukkig." Nicole: "Ik ben niet erg gelukkig."

7. *De regering zal het wegennet in Nederland gaan uitbreiden*
   Laura: "Ik ben niet erg gelukkig." Mirjam: "Ik ben een beetje ongelukkig."

8. *Hans en Peter horen dat koopzondagen in Nederland worden verboden.*
   Hans: "Ik ben ongelukkig." Peter: "Ik ben niet erg gelukkig."

9. *Laura en Inge zijn beiden geslaagd met een zes.*
   Laura: "Ik ben ongelukkig." Inge: "Ik ben niet erg ongelukkig."

10. *Linda en Pieter moeten over een jaar met pensioen.*
    Linda: "Ik ben niet erg ongelukkig." Pieter: "Ik ben een beetje ongelukkig."

11. *De zoon van Els en Cas gaat voor een jaar naar Australie.*
    Els: "Ik ben een beetje gelukkig." Cas: "Ik ben niet erg ongelukkig."

12. *De regering wil het aantal dierproeven in Nederland gaan terugdringen.*
Karel: "Ik ben niet erg ongelukkig." Jan: "Ik ben gelukkig."

Antonym pair {blij, bedroefd} (happy/sad)

1. *De buren van Piet en Marianne gaan verhuizen.*
Piet: "Ik ben niet bedroefd." Marianne: "Ik ben blij."

2. *Volgende week zal het flink koud gaan worden.*
Nadien: "Ik ben niet bedroefd." Frits: "Ik ben een beetje blij."

3. *Wegens ziekte vervalt de muziekles van Peter en Daan.*
Peter: "Ik ben een beetje bedroefd." Daan: "Ik ben niet blij. "

4. *Het voorstel van Laurens en Anne is afgewezen.*
Laurens: "Ik ben niet blij." Anne: "Ik ben bedroefd."

5. *De moeder van Tamara en Chantal heeft een nieuwe vriend.*
Tamara: "Ik ben niet erg blij." Chantal: "Ik ben blij."

6. *Het college dat aanvankelijk zou vervallen, gaat toch door.*
Erik: "Ik ben een beetje blij." Loes: "Ik ben niet erg blij."

7. *De bioscoop naast het huis van Joop en Joanne zal sluiten.*
Joop: "Ik ben niet erg blij." Joanne: "Ik ben een beetje bedroefd."

8. *De huurtoeslag in Nederland zal gaan stijgen.*
Jos: "Ik ben niet erg blij." Moniek: "Ik ben bedroefd."

9. *De grootmoeder van Tanja en Wies is gisteren overleden.*
Tanja: "Ik ben niet erg bedroefd." Wies: "Ik ben bedroefd."

10. *Er komt geen vervolg op het boek dat Wim en Victor gelezen hebben.*
Wim: "Ik ben niet erg bedroefd." Victor: "Ik ben een beetje bedroefd."

11. *Het bedrijf waar Ria en Simon werken heeft dit jaar een redelijke omzet gehaald.*
Ria: "Ik ben een beetje blij." Simon: "Ik ben niet erg bedroefd."

12. *De hittegolf in Nederland is voorlopig nog niet voorbij.*
Annemiek: "Ik ben niet erg bedroefd." Tjeerd: "Ik ben blij."

C.2 Stimuli

Antonym pair {blij, bedroefd} (happy/sad) (overlaps with appendix C1)

1. *De buren van Piet en Marianne gaan verhuizen.*
Piet: "Ik ben niet bedroefd." Marianne: "Ik ben blij."

2. *Het voorstel van Laurens en Anne is afgewezen.*
Laurens: "Ik ben niet blij." Anne: "Ik ben bedroefd."

3. *De vakantieplannen van Harold en Katja gaan niet door.*
Harold: "Ik ben niet bedroefd." Katja: "Ik ben niet blij."

Antonym pair {blij voor, jaloers op} (happy-for/resentment)

1. *Fred heeft veel waardering gekregen voor zijn werk.*
   Jaap: "Ik ben niet jaloers op Fred." Koos: "Ik ben blij voor Fred."

2. *De verwende nicht van Anne en Iris is kampioen zwemmen geworden.*
   Anne: "Ik ben niet blij voor onze nicht." Iris: "Ik ben jaloers op onze nicht."

3. *De kruidenier heeft dit jaar goede zaken gedaan.*
   De bakker: "Ik ben niet blij voor de kruidenier." De slager: "Ik ben niet jaloers op de kruidenier."

Antonym pair {leedvermaak, medelijden} (gloating/pity)

1. *Voetbalclub FC Groningen is uitgeschakeld in de competitie.*
   Paul: "Ik heb medelijden met FC Groningen." Joost: "Ik heb geen leedvermaak over de uitschakeling van FC Groningen."

2. *Een neef van Bart en Jos heeft zijn been gebroken.*
   Bart: "Ik heb geen medelijden met hem." Jos: "Ik heb leedvermaak."

3. *De tegenstander van Jesse en Paul is door zijn enkel gegaan.*
   Peter: "Ik heb geen leedvermaak over onze tegenstander." Daan: "Ik heb geen medelijden met onze tegenstander."

Antonym pair {hoopvol, bang} (hope/fear)

1. *Het weerstation voorspelt dat de naderende orkaan in kracht zal afnemen.*
   Joris: "Ik ben niet bang." Bas: "Ik ben hoopvol."

2. *De werkgever van Alex en Stijn gaat binnenkort veel mensen ontslaan.*
   Alex: "Ik ben bang." Stijn: "Ik ben niet hoopvol."

3. *De uitslag van het onderzoek van opa wordt straks bekend gemaakt.*
   Henk: "Ik ben niet hoopvol." Arie: "Ik ben niet bang. "

Antonym pair {verlangen naar, bang voor} (hope/fear(for))

1. *Marlies en Ilse gaan vandaag naar het strand.*
   Marlies: "Ik ben niet bang voor de zee." Ilse: "Ik verlang naar de zee."

2. *Marie en Fleur zijn op safari in Afrika.*
   Marie: "Ik ben bang voor tijgers." Fleur: "Ik verlang niet naar tijgers."

3. *De docent Nederlands is de strengste van de school.*
   Fleur: "Ik verlang niet naar de docent." Merel: "Ik ben niet bang voor de docent."

Antonym pair {trots, schamen} (pride/shame)

1. *Daan en Eelco hebben een brief geschreven naar de minister.*
   Daan: "Ik schaam me niet." Eelco: "Ik ben trots."

2. *Doris en Evelien hebben een erg dure televisie gekocht.*
   Doris: "Ik schaam me." Evelien: "Ik ben niet trots."

3. *De presentatie van Julia en Karlijn ging redelijk.*
   Julia: "Ik schaam me niet." Karlijn: "Ik ben niet trots."

Antonym pair {bewonderen, minachten} (admiration/reproach)

1. *Een vriend van Floris en Gerard zet zich in voor het milieu.*
   Floris: "Ik bewonder onze vriend." Gerard: "Ik minacht onze vriend niet."

2. *De minister heeft besloten te bezuinigen op de zorg.*
   Hella: "Ik bewonder onze minister niet." Willem: "Ik minacht onze minister."

3. *De buurvrouw van Pim en Janneke heeft haar zieke hond laten inslapen.*
   Pim: "Ik bewonder haar niet." Janneke: "Ik minacht haar niet."

Antonym pair {houden van, haten} (love/hate)

1. *De hond van onze buren is erg gehoorzaam.*
   Alex: "Ik hou van de hond." Stijn: "Ik haat de hond niet."

2. *De vader van Karolien en Rik heeft zijn gezin bedrogen.*
   Karolien: "Ik hou niet van onze vader." Rik: "Ik haat onze vader."

3. *Francien en Pieter hebben hun broer al lange tijd niet gezien.*
   Francien: "Ik hou niet van onze broer." Pieter: "Ik haat onze broer niet."

Antonym pair {mooi vinden, afstotelijk vinden} (like/dislike)

1. *Een vriendin van Henk en Klara heeft haar haar geverfd.*
   Henk: "Ik vind het haar mooi." Klara: "Ik vind het haar niet afstotelijk."

2. *De werkgever van Ruud en Marc heeft een nieuwe auto.*
   Ruud: "Ik vind de auto niet mooi." Marc: "Ik vind de auto afstotelijk."

3. *De ouders van Xandra en Karel hebben een nieuw huis gekocht.*
   Xandra: "Ik vind het huis niet mooi." Karel: "Ik vind het huis niet afstotelijk."

C.3 Stimuli

Synonyms

1. *Karin en Victoria horen voetstappen.*
   Karin: "Ik ben bang." Victoria: "Ik ben angstig."

2. *De hond van Laurien en Niek is ziek.*
   Laurien: "Ik ben bedroefd." Niek: "Ik ben verdrietig."

3. *Dirk en Mieke zijn gezakt voor hun rijexamen.*
   Dirk: "Ik ben treurig." Mieke: "Ik ben bedroefd."

4. *De kat van de buren is overreden.*
   De buurmam: "Ik ben treurig." De buurvrouw: "Ik ben weemoedig."

5. *Het kabinet heeft besloten de belastingen te verhogen.*
   Paul: "Ik ben boos." Pieter: "Ik ben kwaad."

6. *De winkelbediende is altijd erg onaardig tegen haar klanten.*
   Maurice: "Ik minacht de winkelbediende." Ralph: "Ik veracht de winkelbediende."

C.4 Introduction form

Algemene informatie

Afstudeerder/Onderzoeker: Susanne Hendrickx <s.hendrickx@ai.rug.nl>

In dit experiment wordt onderzocht hoe mensen bepaalde uitspraken interpreteren. Er zullen straks 61 situaties beschreven worden waarin telkens twee personen een uitspraak doen. Vervolgens wordt een meerkeuzevraag gesteld over deze twee uitspraken. Het beste antwoord dient daarbij aangevinkt te worden. Het gegeven antwoord wordt dan automatisch opgeslagen. Denk eraan: er bestaan geen goede en foute antwoorden. Het gaat erom dat je goed weergeeft hoe jij de uitspraken interpreteert. Je mag daar rustig de tijd voor nemen.

Nadat je een vraag hebt beantwoord en op "Volgende vraag" hebt geklikt kun je niet meer je vorige antwoord veranderen. Probeer dus niet terug te gaan naar het vorige scherm dmv de "Back" knop in je browser, omdat dit de resultaten kan aantasten. De test duurt ongeveer 15-20 minuten. Als je opmerkingen hebt over de test, kun je die plaatsen in het invulveld aan het eind van de test.

Voordat je deelneemt zou ik graag persoonlijke informatie over jezelf en je taalkundige achtergrond willen krijgen, omdat dit de resultaten kan beinvloeden. Deze informatie zal alleen gebruikt worden voor de verwerking van dit experiment.

Wat is je voornaam?

Wat is je achternaam?

Wat is je leeftijd?

Wat is je moedertaal?

In welk land ben je opgegroeid?

In welk land of landdeel heb je heb je grootste deel van je volwassen leven gewoond?

In welk land woon je momenteel?

Wat is de hoogste opleiding die je hebt gevolgd?
(selecteer)

Wat is je email adres?

Start de test