A Framework to support the influence of Culture on Nonverbal Behavior generation in Embodied Conversational Agents

A Master's Thesis in
Computer Science
by

Joost van Oijen

Committee:

Dr. W. L. Johnson
Prof. Dr. Ir. A. Nijholt
Dr. D.K.J. Heylen
Dr. Z.M. Ruttkay

Information Science Institute (ISI)
Viterbi, School of Engineering
University of Southern California
The United States

Human Media Interaction (HMI)
Department of Electrical Engineering,
Mathematics and Computer Science
University of Twente
The Netherlands

August 2007

ABSTRACT

Embodied Conversational Agents (ECAs) are becoming more and more believable. One aspect of believability deals with the generation of realistic nonverbal behaviors when performing a communicative act. The choice of nonverbal behavior to express should depend on the ECA's identity as well as the context of the conversation.

The main goal of this research is to investigate the influence culture has on the choice of nonverbal behavior. An *information model* will be developed which structures cultural information about communicative behavior. This model will be integrated into a behavior generation framework for ECAs. This framework then makes it possible to generate communicative nonverbal behaviors adapted to an ECA's identity, cultural roots and contextual factors specifying the conversational setting.

The generation of nonverbal behaviors is determined by the communicative functions that need to be performed. The framework has two phases which determine the nonverbal behavior. First a choice will be made what behavior to express when performing a communicative function. This choice will be made using a mapping table converting functions to behaviors. Second, styles can be added to the generated behavior. Styles give additional information about the frequency and expressivity of certain behaviors.

PREFACE

This thesis is based on research I conducted from September 2006 to April 2007 at the Center for Advanced Research in Technology for Education (CARTE) located at the Information Science Institute in Marina del Rey, as part of the University of Southern California, Los Angeles, United States.

I finished the thesis at the Human Media Interaction (HMI) Research Group as part of the University of Twente, Enschede, the Netherlands. It marks the end of my graduation project and with it the end of my study Technical Computer Science.

First of all I would like to thank my main internal supervisor Prof. Dr. Ir. Anton Nijholt and main external supervisor Dr. W. Lewis Johnson for providing me the opportunity to conduct research at the Information Science Institute in California. It was a tremendous experience for me. What better way to study the importance of subtle communication differences between cultures than to spend the duration of the research living in a foreign culture.

I would like to thank Dr. Hannes Vilhjálmsson from the University of Iceland for his assistance in the search for an interesting research subject.

Furthermore I would like to thank the rest of my graduation commission Dr. Dirk Heylen and Dr. Zsófia Ruttkay for taking the time to evaluate and comment on my findings.

At last I would like to thank my family, friends and housemates who have supported me during these last months, during my foreign visits and throughout my whole study period.

Joost van Oijen, August 2007

CONTENTS

Part I

INTRODUCTION

1. INTRODUCTION

This chapter gives an introduction to several topics that are discussed in this document to get a better understanding of the research goal and the title of this document. The scope of this research encompasses the simulation of human communication using *Embodied Conversational Agents (ECAs)*. In short, ECAs are virtual characters whose purpose it is to simulate humans in face-to-face communication by expressing both verbal and nonverbal behavior.

Section 1.1 starts with a talk about human communication followed by a short summary of communication in virtual characters in section 1.2. Section 1.3 describes two levels of communication, namely communicative functions representing the *Mind* or the *meaning behind behavior* and communicative behaviors representing the *Body* or the *expression of behavior*. Section 1.4 discusses the factors which can cause different behavior expressions between different people performing the same communicative function. The chapter concludes with a description of the research goal in section 1.5 and the motivation behind this research in section 1.6.

1.1 Human Communication

Communication between humans is driven by goals or sub goals that are being pursuit. Take for example a person walking on the street on his way to a friend with whom he has an appointment. He is pretty hungry and is wondering if he still has time to buy a sandwich. In order to decide this, he needs to know the current time. He doesn't have a watch on him, so for him to accomplish this goal, he could communicate with a stranger passing by.

The reason for this communication is driven by a goal, in this case 'wanting to know the current time'. To successfully retrieve the desired answer, he has to follow certain interactional rules. One approach of achieving the goal is to first approach the stranger and get his attention, then to greet him and at last politely ask him for the time. These actions perform *communicative functions:* starting a conversation, greeting, asking a question. Other examples of communicative

functions are giving the turn, requesting something, expression emotions, giving advice and illustrating something.

To perform these *communicative functions* the person has to use his communication system to transmit his intentions to the other person. This transmission is done by expressing *communicative behavior*. This behavior can be verbal by using speech, or nonverbal by using gestures, facial expressions, gazing or other body movements. For example, he can ask the time in a verbal way by asking 'Could you tell me the time'. He could also accomplish this by making a gesture and tipping with his finger on his wrist, indicating he wants to know the time. Regularly, verbal and nonverbal behavior are both used to perform a communicative function. When people make an enquiry, they usually raise their eyebrows and slightly tilt their head while they ask a question.

So every behavior has a cause, in the form of a communicative function or intention. The behaviors can be conscious or unconscious to the person expressing the behavior. For example, the person asking for the time might also express some emotion even if he isn't aware of it. The emotion is then unintentionally leaked through his communication system to the other person.

To summarize things, the mind contains many planned goals and sub goals that can be achieved, each having their own priority. Some of these goals can be achieved by using communication. Executing such goals results in performing a sequence of communicative functions by using the person's communication system resulting in the expression of verbal and nonverbal behavior.

1.2 Embodied Conversational Agents

Embodied Conversational Agents (ECAs) are (autonomous) software entities representing a virtual human character demonstrating many of the same properties as humans in face-to-face conversation. This includes the ability to produce and respond to verbal and nonverbal communication, the ability to deal with conversational functions such as turn-taking, feedback, repair mechanisms and the ability to give signals that indicate the state of the conversation, as well as to contribute new propositions to the discourse [9]. The ECAs are presented as virtual humans that use the same modalities as humans do during conversation, such as speech, facial expressions, gestures and body stance.

Much research about ECAs involves increasing the believability of virtual characters. Believability is a term used to address how human-like a virtual character is. It can be used for many aspects of ECAs like natural language capabilities, personality, emotions, animation of gestures, lip-syncing, cognitive

and nonverbal behavior, interaction capabilities, i.e. all aspects contributing to the process of making ECAs more realistic, believable and human-like. An ECA is also more believable when he is able to express multimodal behavior. Multimodal behavior represents behavior across multiple modalities like face, gaze, gestures, head, etc. Usually the combination of behaviors across multiple modalities taken together gives meaning to an expression. Therefore an ECA must be able to express these behaviors in parallel.

ECAs can be created for many purposes. They can be used in human-computer interactions performing supportive or educational tasks (helping the user achieve some goal, providing information); they can interact with each other for simulation purposes; they can be used as characters in games for entertainment purposes, etc.

Architectures for ECAs usually consist of different components connected through interfaces each handling a different aspect of the agent's functionality. These components can include a model of the agent's mind, generating goals that require communication; a dialogue planner handling the state of the conversation and planning what to say next; a behavior planner determining what behavior to express when performing a communicative act; and a behavior realizer which is able to control the virtual agent to execute the plan generated by the behavior planner. Examples of ECA systems are the GRETA system [17, 64, 3] (see figure 1.1), BodyChat [75, 13], the REA system [11, 9] and its predecessor the Gandalf system [72].

1.3 Function versus Behavior

Currently many ECA systems adopt the strategy of using separate structures to specify an agent's communicative function and its communicative behavior, where the functional specification determines *what* the agent wants to communicate and the behavioral specification determines *how* the agent will communicate this using verbal and nonverbal behaviors.

This separation can be seen as two independent components where one component represents the *Mind* of an agent and the other component represents the *Body* [8]. During a conversation, the Agent's *Mind* decides what to communicate, while the *Body* reads what the Mind decides to communicate and renders it at the surface level, according to available communication channels [7].

This design strategy has several advantages. First of all it is now possible to use different bodies for the same mind. For example one agent can be designed to express a communicative function like referring to something by pointing at

Fig. 1.1: GRETA: Expressive Embodied Conversational Agent developed at the University of Paris 8. Figure taken from [54].

the reference target, while another agent may be designed to use his verbal communication channel and perform speech to refer to something. Second, the body now has the opportunity to behave differently when communicating the same communicative function. This makes it's possible to account for dynamic factors that influence someone's behavior (emotional state, natural environment, social environment). For example, if the agent from the example above is communicating in a very noisy environment like a bar or a music concert, than he will not use words to refer to something, because the addressee won't hear it. So he might consider pointing.

Recently a framework for ECA design has been proposed in which communicative functions and communicative behavior can be specified using two different markup languages, namely FML (Functional Markup Language) and BML (Behavior Markup Language). This framework is called the SAIBA (Situation, Agent, Intention, Behavior, Animation) framework [43]. In this framework, FML can be seen as a specification of what an agent's *Mind* wants to communicate. On the other hand, BML is a specification of how the agent's *Body* expresses the message that needs to be communicated. An example tag of FML is *emphasize* or *express joy*. An example tag of BML is *raise eyebrow* or *head nod*.

In this framework planning stages are defined. The first stage determines the ECA's communicative function which is presented in FML. The next stage

determines what behavior to express given a function. The resulting behavior is presented in BML. The last stage handles the realization of the behavior by interpreting the incoming BML and making sure the virtual character behaves accordingly. What is not specified by this framework is how the function is being mapped to behavior. This mapping is what makes one ECA behave differently from another ECA. A different mapping will result in a different style for an ECA.

1.4 Behavioral Style

People behave differently. Even though the message or intention that they communicate is the same, people still behave in a very different, distinct and personal way. For example the way people walk, the way people greet each other, the way people use their eyes during communication, the way people express their emotions, etc. One of many factors that causes this difference in behavior is *culture*. Studies have shown that the way people present information, describe events or talk about feelings are culturally dependent [4]. Culture is of course not the only factor. There are many factors involved in determining one's behavior in a given situation. These can be characteristics about the speaker (personality, culture, age, gender, emotional state), characteristics about the addressee (culture, status, relationship) or characteristics about the environment (noisy, indoor/outdoor). These factors can influence the communication process at different levels. Not only will these factors determine how a person behaves, they also influence a person's thoughts, desires, goals and therefore the content of a conversation. In this research the focus will be on the influences on behavior, the surface level of communication. As an example, consider the communicative function of greeting someone. A 45 year old Japanese man that wants to greet his wife when he comes home from work may do this in an entirely different way than an 16 year old Dutch girl that wants to greet her friends in a crowded, noisy bar on a saturday night.

These differences in behavioral style have not been explored that much in designing ECAs and most agents are very generic in their behavior. Studies have shown however that it is important to consider complex factors like culture [49], personality [36] and environmental setting [58] when designing an agent to increase its believability. Attempts have been made to model personality [67, 59]; personality, interlocutor characteristics and contextual factors to control emotional expressiveness [15]; culture, personality, and other factors by using a representation language based on several function-to-behavior dictionaries that

reflect an aspect of style [70]. Other research tries to influence behavior using expressivity dimensions to adjust space, time and fluency of behavior [47].

1.5 Research

From the title, two main tasks can be distinguished that need to be performed. First, the title talks about the *influence of culture on nonverbal behavior*. This is something that needs to be investigated. What is meant exactly by culture? What kind of nonverbal behaviors are there? What is the relationship between culture and nonverbal behavior? What causes nonverbal behavior and does culture play a role in this? What kind of behavioral differences exist between cultures and how can these be categorized? To answer these questions, a lot of information on cultural behavior will have to be investigated. This information will guide the design of a *Culture Information Model* which can be used to model cultural information about behavior. *Culture* in this domain will be defined as a *configuration of learned behavioral rules which are shared by members of a particular group*. For this research, groups have been defined on a national level, i.e. people from the same country. This is done because behaviors at this level unfold interesting differences.

Second, the title talks about a *framework* which supports the influence of culture during behavior generation in ECAs. This is something that needs to be designed. *Behavior Generation* is a process which determines the behavior an agent will express during communication. A framework will be designed in which cultural factors can influence this process of behavior generation for ECAs. This framework will make it possible to specify and design multiple culturally adapted ECAs each behaving according to their identity and cultural roots.

To give an example of culturally adapted behavior, take the following piece of information: *"In Japan people bow deeper and longer when greeting someone of a higher status"*. This piece of information actually consists of two different types of information. The first piece says that people from Japan bow when greeting each other, while the second piece says that they bow deeper and longer when the person they greet is of a higher status. Now the *Culture Framework* should make it possible to specify an ECA which is adapted to the Japanese culture and is able to generate the correct behavior in his own communication encounters.

The *Culture Framework* that will be designed creates the ability to more easily design believable ECAs with complex behavioral rules creating a personal

style based on the utilization of personality, social and cultural factors. The framework must be able to be used in a broad range of domains, from ECAs in the form of a talking head interacting with the user to ECAs being used in multi-agent based systems that can simulate human communication. Figure 1.2 shows an example of an existing system which also tackles the problem of implementing cultural differences in communication in multi-agent systems.

Fig. 1.2: American soldiers adapting to Iraqi culture. Figure taken from the Tactical Iraqi Language and Culture Training System [48].

1.6 Motivation

The motivation for this research is driven by the desire to make ECAs more believable and variable in expressing their behavior. One way to make ECAs more believable, i.e. making them more human-like, is to give them an *identity*. This identity can consist of assigning the ECA to have a certain age, gender, personality or cultural background. These identity factors can have a great influence on human communicative behavior. The focus of this research is mainly on culturally determined behavior. The reason for this is that culture play a large role in human communication and this has to be taken into account when

designing ECA systems.

To illustrate these cultural differences, take for example an Arab starting a conversation with an American. During the conversation, the American expects the Arab to maintain a certain distance. This is the distance the American feels comfortable with to talk to a stranger. He expects the Arab to respect his personal space. If the Arab comes too close he may seem aggressive or intrusive in the eyes of the American. If he stays too far he may seem anxious or not interested. The Arab person on the other hand has other expectations and standards with which he feels comfortable. Some behavioral differences that may seem insignificant could result in cross-cultural conflicts.

From this example follows that the believability of an ECA is dependent on cultural differences. People belonging to the same culture have a shared understanding of behavioral rules. When designing a believable or culturally adapted ECA, behavioral rules can be specified for one specific culture. But if it is desirable to have multiple agents representing different cultures, this process can become very time consuming. It would be more efficient to identify exactly what the cultural differences are and try to adapt the behavior to follow the norms of some culture.

There are many basic communication scenarios that occur all over the world where some communicative goal is achieved. Take for example the act of asking for directions. The communicative functions that perform this act could consist of greeting someone, taking the turn and asking a question. When the communicative functions are the same, the resulting behavior that is being expressed might still be different between different cultures. Therefore it would be efficient to make a distinction between the specification of communicative functions and behaviors. This can make it more easy to influence an ECA's behavior when its functions stay unmodified. A behavior generation framework for ECAs that has been developed, called the *SAIBA Framework*, makes this distinction between function and behavior and therefore will be the starting point of the framework to design.

2. RELATED WORK

There has been a lot of research in the development of Embodied Conversational Agents. They constitute of a multimodal interface with modalities like speech, gesture, facial expressions; a software agent which represents the computer or a human user depending on the purpose of the agent; and a dialogue system where verbal and nonverbal communication regulates the conversation. Each of these areas have received a lot of attention recently [12]. For instance, the desire to make ECAs more believable has led to more and more research on nonverbal communication skills in ECAs. Believable agents need the ability to express their intentions through a richness of used modalities and gestures.

Many markup languages have been proposed throughout the years to specify multimodal behavior in ECAs. Section 2.1 describes the most important markup languages that have been developed.

2.1 Multimodal Behavior Languages

The conversational content an ECA chooses to deliver is highly domain specific and is determined by its dialog system. The way an agents delivers this content though is a reusable skill. The ECA plans communicative functions which results in communicative behaviors. The different functions and behaviors an ECA can perform are limited and are reused in conversations. Therefore many function and behavior representation languages have been developed to specify these components. Most of these languages are XML based.

ECAs are usually developed for a very specific purpose and domain. The complexity and the richness of multimodal behavior differs from agent to agent. The abstraction level of the described behavior is also domain dependent. This is the reason why most ECA system designers developed their own behavior and functional language specifically designed for their agent's needs. Depending on their complexity, some architectures made clear distinctions between behavioral and functional elements, others did not.

MURML

Multimodal Utterance Representation Markup Language [44] is a language designed for anthropomorphic agents to describe the results of behavior planning. Speech can be annotated with tags. The focus of the language is on specifying very detailed face, body or hand movements and less on the function causing the movements. For instance, a hand movement description can contain information about the hand shape, the palm orientation, finger locations and movement directions.

VHML

Virtual Human Markup Language [34] is designed to accommodate various aspects of human-computer interaction. It is a collection of partly previously existing markup languages describing gestures, emotions, speech, facial animation, body animations and dialogues. There is no clear difference between function and behavior and both types are used intertwined. For example, there is an *emphasis* tag in the emotion markup language which is a communicative function and can be used in combination with tags from the body animation markup language, which specifies the communicative behavior.

BEAT/Spark

Behavior Expression Animation Toolkit [14] is a system which generates synchronized nonverbal behavior based on linguistic and contextual analysis of some input text. It includes different generators for beat gestures, iconic gestures, gaze and eyebrow behavior, etc. The Spark system modified BEAT for an avatar-based chat system. In Spark two separate XML tag sets were defined to make a distinction between communicative function (intent) and behavior. Here the input text is first automatically annotated with function tags. Then the behavior generators transform the function tags into behavioral tags which define the resulting behavior.

APML

The Affective Presentation Markup Language [8] specifies an agent's behavior at the meaning level. In the architecture, the agent is made up of two components: the mind and the body. The language assures the independence of both components. It is based on Poggi's taxonomy of communicative functions defined as a meaning-signal pair [63]. A meaning can be expressed by different signals.

This allows us to create different types of agents in which different aspects can be represented like personality, age or culture. An agent is thus defined by a dictionary of meaning-signal pairs.

SAIBA/FML/BML

At the AAMAS 2002 workshop "Embodied Conversational Agents - let's specify and evaluate them!", it became clear that many behavior and functional languages contained a lot of similarities. Therefore a push was initiated to standardize the behavioral and functional specifications of ECAs. The resulting languages are FML (Functional Markup Language) and BML (Behavior Markup Language). Their development is an ongoing process and tags will be proposed based on current research and experience. Besides the two markup languages, a unifying model of representations for multimodal generation was created: the SAIBA framework. This framework defines the planning stages involved in multimodal generations and the knowledge structures (FML/BML) that mediate between them. The SAIBA framework is described extensively in chapter 4 and will be used as the basis for the behavior generation framework in the design of the *Culture Framework*.

2.2 ECAs with Style

Although there are many ECA behavior specification languages out there, there hasn't been a lot of extensive research on the behavioral *style* of ECAs. An agent having a personal style which is influenced through factors like the gender, age or culture he must represent makes him more believable and usable for a broader set of applications. Different styles are also desired in different conversational settings. It may also be desirable for an ECA to behave in accordance with his physical and mental state. A few research projects which attempt to address this notion of style for ECAs are described in this section.

2.2.1 GESTYLE

GESTYLE [60, 68, 69] is a markup language to annotate text to be spoken by an ECA. The annotation tags can be low-level behavior tags specifying nonverbal behaviors like gestures and eye gazing. They can also be high-level meaning tags, handling conversation regulators or other communicative functions. Besides that, styles can be defined. Styles determine the eventual behavior of an ECA

because high-level meaning tags are translated to low-level behavior/signal tags by using information defined in meaning-to-behavior mapping dictionaries.

The motivation behind the development of GESTYLE is to create more believable ECAs and to be able to specify characteristics like personality, ethnicity and culture. These characteristics are important because they have an influence on the expressed behavior of the ECA and therefore on the user's response to the ECA [59, 77].

In GESTYLE every agent has a style declaration which specifies one or more *style dictionaries* to use. Style dictionaries contain rules which map meaning to behavior. Style dictionaries can be used to define a certain style in which factors like culture, gender and personality can be represented. For example, one style dictionary could represent the style of an Asian person while another style dictionary could represent the style of a 10-year old girl. Many different ECAs can be defined because style dictionaries can be combined to create a more complex style. Mapping conflicts between dictionaries are resolved by prioritizing dictionaries in the style declaration.

Besides selecting dictionaries to use, a style declaration can also define a manner definition, specifying motion characteristics of gestures (smooth, fast), and modality usage, specifying modality preference (preferring the left hand above the right hand). GESTYLE also provides the ability to dynamically change style during communication. The emotional state of an agent can change or there can be a change in situation. At runtime changes can be made to the style declaration by modifying the style dictionary configuration, their priorities, the modality preference or the manner characteristics.

2.2.2 GRETA

Another attempt to design an ECA with style is *GRETA* [64, 47], an expressive embodied conversation agent. The *GRETA* system generates behavior for a talking ECA. It takes as input text that is augmented with communicative functions described in APML [8]. To determine the nonverbal behaviors to express the system relies on a taxonomy of communicative functions proposed by Poggi [63]. The translations from the communicative functions described in APML to the animated behavior are defined by meaning-signal pairs which map functions to behaviors.

Besides this translation of meaning to signals, the GRETA system can also define a style for an ECA. The expressive behavior determining the style does not aim to model different factors like culture or personality. What is modeled however is the way a communicative function can be expressed quantitatively

and qualitatively. To achieve this, six expressivity dimensions are defined:

- *Overall Activation:* Quantity of movements across several modalities during a conversation turn.
- *Spatial Extent:* Amplitude of movements.
- *Temporal:* Duration of movements.
- *Fluidity:* Smoothness and continuity of overall movement.
- *Power:* Dynamic properties of the movement (weak vs. strong)
- *Repetitivity:* Tendency of rhythmic repeats of specific movements along specific modalities.

Using these expressivity dimensions, two agents performing the same communicative act can now behave very differently by assigning different dimension values to each agent. This expressivity control layer can act as a basis for modeling more complex factors, such as culture and personality.

Part II

LITERATURE

3. CULTURE AND COMMUNICATION

This chapter reviews the literature on culture and human communication that is relevant for this research. Because we are dealing with the *simulation* of human face-to-face communication with the goal of making ECAs more believable, a lot of the literature involves analysing, modeling and categorizing complex human factors involved in communication.

Culture is one of these factors and is a term used in many different situations. What is meant by culture in the scope of this research is described in section 3.1. Section 3.2 attempts to provide an overview of how humans go from goals they want to accomplish to beliefs they want to transfer to the actual verbal and nonverbal behavior they express during communication. Section 3.3 handles different factors that can have an influence on the process of going from beliefs to communicative behavior. After that, section 3.4 handles communicative functions in more detail which are for a large part responsible for the generation of communicative behavior. This chapter concludes with an overview of human nonverbal communicative behaviors in section 3.5. Throughout this chapter, the influence *culture* can have on the different topics is investigated.

3.1 Culture

The word *culture* can have many different meanings used in different contexts. In this context, culture is used in the form of behavioral norms and rules that are shared by people belonging to a group. The focus is also mainly on groups on a national level in the sense that cultural differences between nations are explored. A definition of culture is given here that is used throughout this document. Existing cultural categorizations like Hofstede's *cultural dimensions*, Hall's *contact* versus *noncontact cultures* and *country clusters* are reviewed. At last, the relationship between cultural dimensions and nonverbal behavior is being investigated.

3.1.1 Definition of Culture

There are many definitions for the word *Culture*, ranging from aesthetic, ethnographic and symbolic definitions. Below are listed a few classic definitions of culture to better understand the range of factors that are involved. Following that, a simpler definition of culture will be given that is most suited for the context of this research.

"Culture taken in its wide ethnographic sense is that complex whole which includes knowledge, belief, art, morals, law, custom, and any other capabilities and habits acquired by man as a member of society." - Edward B. Tylor, 1958 [1871] [73]

"A historically transmitted pattern of meanings and symbols, a system of inherited conceptions expressed in symbolic form by means of which men communicate, perpetuate and develop their knowledge about attitudes toward life." - Clifford Geertz, 1973 [22]

"A society's culture consists of whatever it is one has to know or believe in order to operate in a manner acceptable to its members, and to do so in any role that they accept for any one of themselves." - Ward Goodenough, 1957 [23]

"Culture consists in the shared patterns of behavior and associated meanings that people learn and participate in within the groups to which they belong" - Whitten and Hunter, 1976 [80]

In the context of this research, a more cognitive definition of culture is desired, with the focus on behavioral rules, knowledge, norms and meanings. Tylor's definition is too wide in the sense that it also includes the physical consequences of the cognitive factors that are part of a culture. The more cognitive definitions of Goodenough and Whitten and Hunter are more useful and address knowledge and behavior shared by a group as the key concept. The definition of culture that will be used throughout this text is:

"A culture is a configuration of learned behavioral rules, norms and meanings which are shared and transmitted by members of a particular group and therefore understandable and acceptable to its members."

3.1.2 Cultural Dimensions

Theories for cross-cultural communication is largely based on research looking at value differences and cultural dimensions. Some major contributors in this field are Edward T. Hall [26], Geert Hofstede [30], Fons Trompenaars [27] and Shalom Schwartz [71].

Cultural dimensions are mostly psychological dimensions that can be used to describe aspects of a specific culture. They can differentiate and measure differences between cultural groups. The next two sections will illustrate the most influential model of cultural dimensions by Geert Hofstede followed by Edward Hall's model of contact versus noncontact cultures.

3.1.3 Hofstede's Cultural Dimensions

Geert Hofstede is a Dutch organization psychologist. In the late 60's and begin 70's he conducted a massive survey on IBM employees in 72 countries. As a result he developed a cultural model identifying four dimensions of value perspectives between national cultures. Later he added a fifth dimension based on survey results from the Far East and Asia [30]. These five dimensions are:

Power Distance

This dimension determines the extent to which the less powerful members of a society accept and expect that power is distributed unequally. Countries with a shorter power distance tend to be more democratic in their approach to power. In countries with a large power distance politics tends to be more extreme. Countries with the greatest power distance include Malaysia, Guatemala, the Philippines and Panama. Among the lowest power distance countries are Austria, Israel, Scandinavian and Anglo-Saxon countries.

Individualism versus Collectivism

This dimension determines the degree to which individuals are integrated into groups. In an individualist nation everyone grows up to look after themselves and identity is based on the individual. The ties between the individuals are loose. Collectivist nations on the other hand base their societies on extended families and social networks define people's identities. Everything is organized into groups. Among the top individualist cultures are the USA, Australia, Canada, Great Britain and the Netherlands. High collectivist countries include

Guatemala, Ecuador, Venezuela, Panama and Colombia. Neutral countries include Turkey and Arab countries.

Masculinity versus Femininity

This dimension refers to the roles between the genders. Masculine societies tend to be more assertive and competitive, whereas feminine oriented societies tend to be more cooperative, modest and caring. In feminine countries men's and women's modest and caring values are the same, but in masculine countries there exists a gap between the values. The women are somewhat assertive and competitive, but not as much as the men. Highly masculine oriented societies include Japan, Austria, Italy and Venezuela. Highly feminine oriented societies are Norway, Sweden, Denmark and the Netherlands.

Uncertainty Avoidance

This dimension deals with a society's tolerance for uncertainty and ambiguity. Uncertainty-avoiding cultures try to minimize unstructured situations by strict laws, rules and security. Uncertainty-accepting cultures are more tolerant towards opinions other than their own. They try to have as few rules as possible. Cultures with the strongest uncertainty avoidance are Portugal, Greece, Japan, Israel and France. The lowest scoring countries are Sweden, Denmark, Singapore, Jamaica and the UK.

Long-term versus Short-term Orientation

This dimension is a result of a study among 23 countries conducted in Asia. Values associated with long-term societies are persistence, thrift, having a sense of shame and ordering relationships by status. Short-term orientation values include tradition, social obligations, personal stability and protecting one's face. The long-term societies are Asian countries with China at the top. Anglo-Saxon countries like the USA, Australia, New Zealand and Canada have the lowest score.

3.1.4 Contact versus Noncontact Cultures

Edward T. Hall stated that cultures can be distinguished, in part, by shared behavior regarding interaction distances, body orientation, gaze patterns, and frequency of touch [25]. This distinguishment led to the division of cultures into *contact cultures* and *noncontact cultures*. According to him, contact cultures

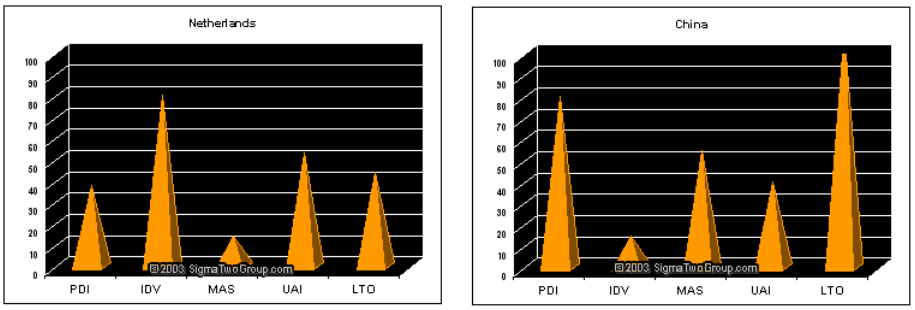


Fig. 3.1: A comparison of Hofstede's cultural dimensions between the Netherlands and China. Figure taken from [31].

prefer more immediate, affiliative or involving behaviors. His observations led him to conclude that Latin Americans, Arabs and French (south or east of Paris) belonged to *contact cultures* while North Americans, Northern Europeans and Japanese people belonged to *noncontact cultures*.

Watson classified 30 countries as either a *contact* culture or a *noncontact* culture. He found that contact cultures engaged in more gazing and had more direct orientations when interacting with others, less interpersonal distance and more touching [78].

3.1.5 Country Clusters

Another way to look at culture categorization is to divide the world into country clusters and dividing country clusters into single countries. Because this research is focussing on the influence of national cultures on behavior, it might proof useful to categorize countries into clusters. Not only is there a lot of information available about behavioral differences between countries, but also information about different country clusters.

An example of a country cluster categorization is the categorization of Ronen and Shenkar [66] will be used here. The clusters are *Anglo* (United States, Australia, United Kingdom), *Germanic* (Switzerland, Germany, Austria), *Nordic* (Denmark, Sweden, Norway), *Near Eastern* (Greece, Turkey, Iran), *Arab* (Saudi Arabia, Oman, Kuwait), *Far Eastern* (Indonesia, Thailand, Malaysia), *Latin American* (Argentina, Mexico, Peru), *Latin European* (France, Spain, Italy) and *Independent* (Brazil, Japan, India).

3.1.6 Relationship Cultural Dimensions and Behavior

In this research the cultural influence on behavior generation is explored. For this reason, it is worthwhile to look at different *cultural dimensions* that have been proposed and investigate the existence of a relationship between *cultural dimensions* and certain aspects of behavior.

Take for example the observed relationship between the cultural dimension of *Individualism/Collectivism* and *Proxemic Behavior*. Proxemic behavior in communication involves the distance people adhere when interacting. Such a relationship may state that people from individualistic cultures (the Netherlands, United States) adhere a larger distance towards others than people from collectivist cultures (China, Colombia).

Another observed relationship is the cultural dimension *Masculinity*. Women in low-masculinity cultures show more synchrony in their movement than those in high-masculinity cultures. Research has also shown that people from high *power distance* cultures tend to be more tense in body movement and tend to smile more towards superiors. They also are more aware that vocal loudness may be offensive to others.

3.2 From Beliefs to Behavior

What someone will communicate at a given time is dependent on his goals, beliefs and intentions. Everybody has beliefs about himself, about external objects and events. Communication is a way to transfer those beliefs to someone else, who in return might transfer his own beliefs. To achieve this, a person has to use his communication system to produce communicative signals like speech, words, gaze, facial expressions, gestures, body movements, etc.

Cognitive factors like high level goals, psychological needs, senses, physical needs and beliefs can all be the cause of the process of executing a communicative act. For example, the physical need of hunger might generate a plan of goals to accomplish. A high level goal might be *Get something to eat*. The planning of sub goals determine what actions are needed to accomplish the higher level goal. Sub goals might be *Get to the city*, *Get money from the bank* and *Find a restaurant*. Eventually the lowest level goals requiring communication are the cause for generating communicative acts. For example, when sitting at a table in a restaurant, one needs to communicate with a waiter in order to get food. This communication might start with the signalling of the waiter in a verbal or nonverbal manner.

Figure 3.2 illustrates this process from goals and beliefs to the planning of a

communicative act resulting in verbal and nonverbal multimodal behavior. The top part of the figure shows the cognitive factors representing goals, beliefs and needs which might lead to communication. Communication may be initiated by the planning of a communicative act. These processes which take place inside the brain are highly complex and fall outside the scope of this research.

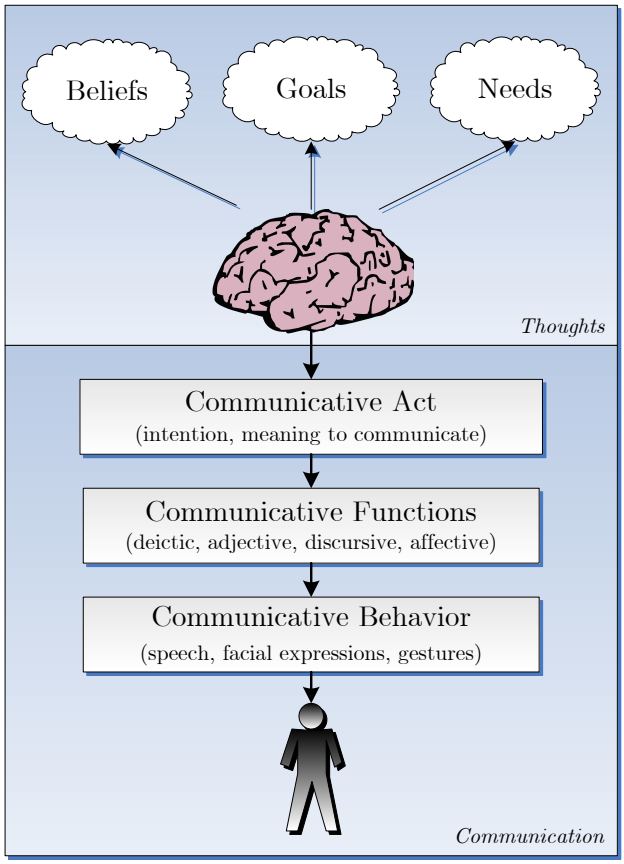


Fig. 3.2: From *Thoughts* to *Communication*.

The bottom part of the figure includes the elements involved in communication. The planning of a communicative act can result in different communicative functions which in turn can result in the planning of multimodal communicative behavior. For example, referring to the example described above, take the signalling of a waiter in a restaurant with the goal of ordering food and fulfilling the main goal of satisfying the physical need of hunger. A communicative act is planned to signal the waiter which results in the planning of functions to per-

form like *getting attention*. Different behaviors can be expressed to execute this function, for example in the form of performing a *hand gesture*, *gazing* at the waiter, *orienting* the upper body in the direction of the waiter, etc. Usually the expressed behavior involves multiple modalities like gestures, facial expressions and gazing.

There are many factors that influence these planning processes. Culture is one of them. For example, in one culture it may be a custom to aggressively *order* the waiter to come. In another culture it may be custom to politely *inform* the waiter that you want to order food. These different customs can result in the planning of different communicative functions. Culture can also have an influence on the behavior to express when ordering or informing the waiter. For example, in Spain it is a custom to signal a waiter by clapping the hands while in England it is a custom to raise your finger. Culture and other influences are described in the next section, followed by an overview of communicative functions and communicative behaviors.

The focus of this research is on these influences during the conversion from communicative functions to multimodal behavior. This conversion will be investigated from a cultural viewpoint, i.e. to investigate group behavior in similar communicative acts.

3.3 Influences

There are many different factors that influence what beliefs and meanings a person may communicate and how a person might achieve this by using verbal and nonverbal behavior. Factors like personality, culture, age and gender influence his choice of behavior. This is what makes a person unique in the expression of behavior. Other people around him also influence his behavior as well as the environment. Poggi et al. have designed a structured overview of all factors which influence the decision concerning which communicative signals to produce, both verbal and nonverbal [64]. This overview is given in figure 3.3.

There are two main types of influences: permanent and contingent factors. The former include factors that are relatively constant during communication and are always active inside the person. These factors are also known as *intrinsic influences*. The latter include factors which are dependent on the situation or context of the communication and are known as *contextual influences*. Contextual influences can further be divided into *external influences* (participants, environment) and *mental and emotional influences* [55]. The different influences are described in more detail below.

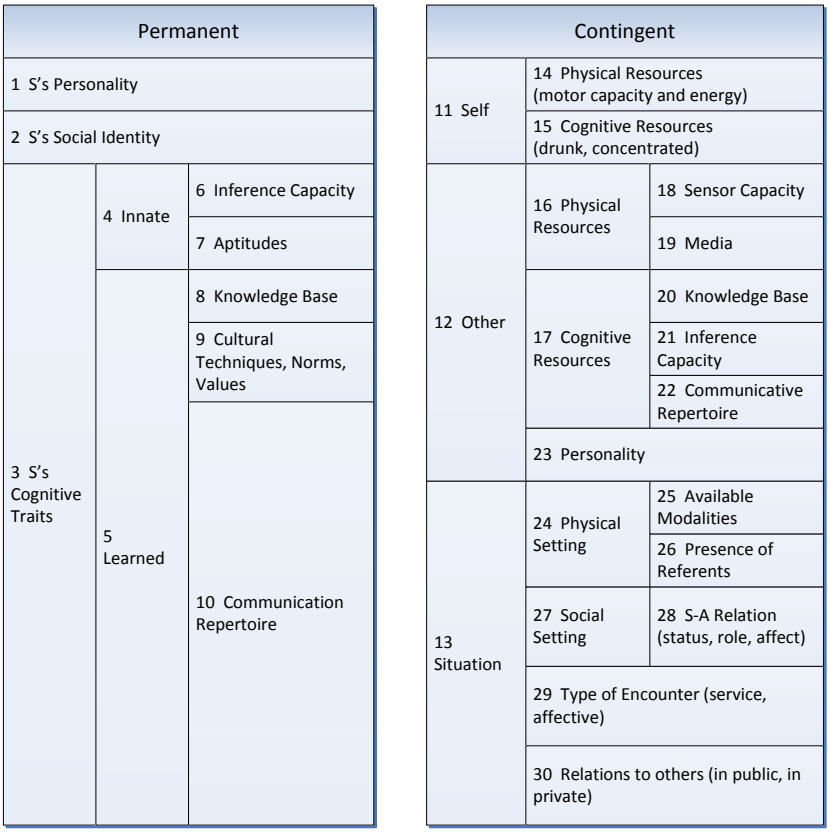


Fig. 3.3: Factors affecting the choice of communication [64].

3.3.1 Intrinsic Influences

Intrinsic, constant factors include personality, social identity (age, gender, cultural roots) and cognitive traits which can be innate or learned. Innate factors include the ability to make inferences and the different aptitudes a person owns. The learned factors are culturally dependent and include beliefs about the environment, cultural norms, values and rules and a communication repertoire containing verbal and nonverbal behaviors which a person has learned since birth. All the permanent factors together define a person's identity. All of these factors have an influence on the behavior that will be expressed. Also intrinsic factors like weight, height or even shoe size or hair type could have an influence on someone's behavior in a certain situation. Intrinsic influences are constant during a typical dialog session. The factors that are the most useful for this research are illustrated below.

Culture

One of the best examples to illustrate the influence of culture is the behavior people express when they greet each other. Other very influential factors in greeting are age, gender and relationship. In the United States it's proper to use a firm handshake with direct eye contact when greeting. Women may briefly hug other women and men may quickly kiss the cheek of a woman. In Argentina good friends will greet each other with a hug accompanied with a few pats on the back. Men from Saudi Arabia greet each other with a hug and a kiss on the cheek. In some Asian cultures the bow is a common greeting. Other greeting forms are the pressing of palms, the rubbing of noses or the nodding of the head.

The examples described here involve national cultures, but other types of cultures also have their influences. Also note that it's too simplistic to categorize people into one culture. People can belong to many cultures and subcultures at the same time which can also change from time to time. Throughout their lives, people join and leave groups which can all have their own norms and standards of how to behave.

More examples of cultural influences, specifically on nonverbal behavior expressions, are discussed in the overview of nonverbal behaviors in section 3.5.

Age

Age is another factor that can have a large influence on the generation of verbal and nonverbal behavior. Take for example the difference in word choice between children and adults, where children have a limited vocabulary compared

to adults. Elderly people usually talk slower than adults, they usually have slower movements and use a smaller gesture space. Children generally use more gestures than adults do and can be more expressive with their emotions.

Gender

Many studies have been conducted to explore the gender differences in nonverbal communication although some studies have mixed results. Burgoon, Buller and Woodall (1996) concluded that North American women engage in more eye contact during conversations than men [5]. Hanna and Wilson (1998) stated that women used fewer gestures than men and that they also used fewer gestures when they were with other women, but more gestures with men [28]. Burgoon, Buller and Woodall (1996) felt that the difference in gestures was not the frequency but the types of gestures. Hanna and Wilson (1998) concluded that women typically smile more than men. Experts in literature generally agree on the difference of personal space, where men require more personal space than women.

Personality

It's obvious that personality plays an important role in behavior. All aspects of personality determine the likelihood some behavior will be expressed. Personality researchers have proposed five basic dimensions of personality. This categorization has been supported by many researchers over the years including D.W. Fiske (1949), Norman (1967), Smith (1967), Goldberg (1981) and McCrae & Costa (1987). The five dimensions are usually being listed as *Extroversion* (excitability, sociability, talkativeness), *Agreeableness* (trust, altruism, kindness, affection), *Conscientiousness* (thoughtfulness, impulse-control, goal-directed), *Neuroticism* (emotional instability, anxiety, moodiness) and *Openness* (imagination, insight, interests).

Intelligence

Intelligence could also have an influence on communicative behavior. The influence on *verbal behavior* for example might be the size of a person's lexicon. More intelligent people usually have a larger lexicon and more knowledge of the use of more complex words and sentence structures. There also exist signs of the influence on *nonverbal behavior*. In one of his studies, Adam Kendon shows that people might become more intelligent if they use gestures on a repeated basis [39]. Although this implies an influence of gestures on intelligence instead of

the other way around, it still shows that there is a positive correlation between gestures and intelligence, namely that a higher intelligence might be associated with a more frequent use of gestures.

3.3.2 Contextual Influences

Contextual factors are dependent on the communication situation and differ in almost every situation. These influences can come from the person himself, from other people or from situational factors.

Self

Factors coming from the person himself are the dynamic factors falling under the *mental and emotional influences*. These include the physical, cognitive, mental and emotional state of the person. To give an example of a such an influence: a person with a lot of muscles aches the day after he did too much exercising will probably use his body differently by restricting certain movements. Also, a person who is very drunk will of course behave very differently: drunkenness can influence the use of paralinguistic features like articulation, volume and rate of speech; it can influence the choice of words because the ability to retrieve items from the lexicon may be affected; the gesture space will probably be larger because movements are harder to control, etc.

Previous research has shown that the emotional state of a person can also have an impact on physical behavior such as facial expressions, gestures and gaze [2, 19, 20]. For example a sad or depressed person may express less gestures, more averted gaze and a posture with fewer movements in a smaller gesture space than a happy or excited person.

Other

Furthermore, people behave differently when they communicate with different people. We take into account the physical and cognitive resources of the other person including his sensor capacity, beliefs, inference capacity and communication repertoire. For example, it would be useless to refer to some object by pointing at it when the other person is blind; we might restrict our use of curse words when the other person is very religious; our choice of words is differently when we talk to a child; we might even talk in a whole other language if the person we communicate with is coming from another country. Also we choose what we communicate with someone based on our knowledge of how that person might react, depending on how well we know the other person's personality and

identity. We also get influenced by the nonverbal behavior of the interlocutor. The frequency of gesturing, gaze behavior and smiling between the speaker and the listener are highly correlated [21].

Situation

At last situational factors can have a major influence on someone's behavior. We behave according to the physical and social setting we find ourselves in as well as the type of encounter and the relation to other people that might be present. Factors involved are the available modalities, the presence of others and the relationship we have with the other person. For example, we talk louder or use more gestures when we are in a very noisy environment; in a cold environment, interactants might have a more active posture to keep themselves warm; the presence of others might limit the topics we talk about or the emotions we express; if the interlocutor is someone with a high status we tend to be more polite; our behavior differs when we are at a job interview or at a party and we might not talk about adult topics when there are children present.

The relationship between the speaker and the addressee can influence the use of nonverbal behavior in various ways. Take for example Hall's theory about proxemics where he identifies four personal spaces people use while interaction. An intimate relationship between the interactants will result in the use of less space and more touching behavior. On the other hand, complete strangers usually prefer more space and are reluctant in touching.

An experiment conducted by Leffler, Gillespie and Conaty (1982) investigates the relationship between status and nonverbal behavior by studying interactions between teachers and students. Findings indicate that high status subjects (teachers) use a more direct orientation towards their lower status subjects (students), talked more, and attempted more interruptions. They also expressed more pointing and touching behavior [51].

As a final remark, note that the influence others and the environment can have on our choice of communication is dependent on the beliefs we have of others and the environment, even if they are not based on facts. For example, we address someone who speaks another language using our own language because we might think the person will understand us; a blind person may talk to no one if the person he was talking to just left; a person may insult someone from another culture and not know that the behavior he expressed is considered obscene in that culture. These issues exist because there is a lack of information or incorrect knowledge. This is usually the cause of many cross-cultural conflicts.

3.3.3 Complex whole of influences

The influence factors that have been described can influence the process of converting beliefs and meanings to communicative behavior at different phases. For example, the influence factors not only partly determine what the content is of a communicative act, i.e. *what* is being said, but also what the expressed behavior is, i.e. *how* the content is being expressed by verbal and nonverbal behaviors. In this research, the focus is mainly on the phase responsible for converting specific communicative functions and meanings to communicative behaviors as described in section 3.2.

The focus here is also mainly on the cultural influence, although it is required to look at factors other than culture as well. This is because behavior that is the result of a communicative function is usually determined by a combination of influence factors. To give an example, children in Western cultures might tease someone by sticking out their tongue. Adults usually don't choose this behavior. So the expressed behavior with the intention to tease someone can be dependent on the culture as well as the age of the person. In total, the behavior is dependent on more influence factors like gender, personality and emotional state but it would be impossible to consider every factor when analysing behavior. Usually it is sufficient to only look at one or a few factors to explain an expressed behavior. So in the example above, only the cultural background and the age of the person are relevant in explaining a behavior of *sticking out the tongue*. This of course does not mean that a child from a Western country always expresses this behavior with the intention of teasing someone. This would probably also be dependent on the emotional state of the person, the mood he's in, the target to which this act is directed, etc.

Many times a behavior could be the cause of multiple intentions. For example, the sticking out of the tongue could also be the result of *showing disgust*, which is one of the six basic universal emotions. On the other hand, in Polynesian culture this gesture is used as a greeting form. Note here that only one specific modality is described, namely the gesture of sticking out the tongue, but the behavior caused by an intention is usually a multimodal behavior using more modalities like facial expressions, posture, hand gestures, etc. For example, the intention of *showing disgust* is responsible for expressing multiple modalities of behavior. Besides sticking out the tongue, the person may also use gestures, facial expression or take on a more defensive posture. The Polynesian custom of greeting would probably be expressed in a more friendly manner in close proximity accompanied with a more friendly posture.

3.4 Communicative Functions

Communicative functions are the lowest-level functions which together fulfill the goal of a communicative act. They directly result in expressions of behavior.

Communicative functions can be divided into two types of information which contribute to the conversation, namely *propositional information* and *interactional information* [9]. Propositional information corresponds to the content and meaning of the conversation, which includes speech, gestures and other nonverbal communication that elaborate on a speech content. It contains meaningful information that moves the conversation forward. An example might be a person who is describing an object and using speech and gestures to describe properties of the object.

Interactional information is used as discourse functions for creating and maintaining a communication channel between participants. They perform regulatory functions such as turn-taking, giving feedback. An example might be a person requesting the turn by raising his finger and his eyebrows to indicate he wants to say something or a person who is giving positive feedback by nodding his head.

3.4.1 Poggi's categorization of meaning

Poggi et al. identify four broad classes of meaning which can be distinguished [63]. This categorization will be referred to throughout this document as the model of communicative functions. The categories are listed below. Figure 3.4 shows an overview of the meaning classes including the communicative functions belonging to them. The *type* column represents a possible subtype or value. The *vocal marker* column gives examples of possible usage in discourse and the *signal* column represents possible behavior expressions corresponding to the specific communicative function type.

Information about the Sender's beliefs

In this category of meanings, the sender provides information about his or her beliefs. Communicative functions belonging to this category are:

- Certainty functions: Senders mark the reliability of information by signaling how certain they are of it. They can do this in a verbal way by using phrases like *'maybe'*, *'of course'* or *'i suppose'*. They can also communicate it by using their eyes (squeezing, avoiding, gazing), eyebrows (raising, lowering) or body posture for example.

Class	Function	Type	Vocal Marker	Signal
Belief	Certainty	High	certain, sure, of course, definitely	lower inner eyebrow
		Low	uncertain, maybe, I suppose	raise inner eyebrow
	Adjectival	Intensification	big, long, large, great, wonderful, quite, very	large eye aperture
		Diminishment	small, subtle, tiny, difficult	small eye aperture
		Inclusivity	all, whole, everything	hand gesture, raise eyebrow
	Belief Relation	Contrast	but, on the other hand	beat gesture, increase pitch, raise eyebrow
		Cause-Result	because, as a result	
		Listing	and	
Intention	Performative	Inform	let me tell you about...	gaze
		Request	could you…?	tilt head, raise inner eyebrow, gaze
		Enquiry	do you…?	tilt head, raise eyebrow, gaze
	Topic-Comment	Topic shift	starting new topic	change posture, high energy
		Theme/rheme	adding new information to the discourse	raise eyebrow
	Deicic	Concrete	this, that, there	point, gaze at reference target
		Event		
		From Memory		pointing finger up, gaze at addressee
	Turn Allocation	Take	taking the turn	break eye contact
		Give	giving the turn	engage eye contact
		Request	requesting the turn	
Affect		Sadness	sad	raise inner eyebrow, look down, frown
		Happiness	happy	smile, raise cheek
Meta-Cognitive		Deciding	well	gaze up
		Remembering	i'm trying to remember...	gaze aside, close eyes

Fig. 3.4: An overview of communicative functions and its possible usage.

- Adjective functions: Senders signal physical properties of things to the receiver. The verbal way of signalling this information is by using adjectives like *small* or *big* for example. The nonverbal way could be by using the eyes or gestures to communicate the properties.

- Belief-relation functions: Senders signal the plan of how their discourse is arranged. Examples are item numbering, cause-result or contrast relations. The verbal way of signalling this is by using words like *because* (cause-result) or *but* (contrast). A nonverbal way for example is by using the fingers to number items or tilting the head the other way to signal contrast.

Information about the Sender's intentions

The second category handles the communicative acts where the sender expresses his or her intentions. The following types of functions belong to this category.

- Performative functions: Any communicative act with the goal of informing, requesting or questioning. Examples are providing information, giving advice, requesting an action by ordering or suggesting, asking a question etc.

- Topic-comment functions: In a conversation there is information that is being shared by the persons communicating. If the speaker provides information that is new, different or contrasting with previous information, he uses signals to illustrate this. Any communicative act performing this function belongs to this class. The signals are typically provided by looking at the interlocutor and raising the eyebrows. A sentence or phrase spoken by the speaker can be divided into a part which represent the information shared by both the speaker and the interlocutor and a part which represent information that is new to the interlocutor. The former part is referred to as the *topic* or *theme* of the sentence while the latter part is referred to as the *comment* or *rheme* of the sentence.

- Deictic functions: Communicative acts referring to something. Examples are referring to a concrete object, person or event that is either physically present or present in the working memory of the persons interacting. Nonverbal ways of referring to something can be achieved by gazing or pointing at something.

- Turn-allocation functions: Communicative acts which regulate the conversation. These include signals performing functions like requesting the turn, taking the turn or giving the turn.

Information about the Sender's affective state

This class handles the communicate acts where the sender informs about his or her emotional state. Emotions like love, anger, fear, terror, joy, sadness, surprise etc, can be expressed in a verbal or nonverbal manner. The expression does not have to be intentionally though. Emotional expressions can also be signalled unconsciously.

Meta-cognitive information about the Sender's metal actions

Belonging to this class are all communications concerning a thinking activity. They provide the receiver with information about sender's processes of thought. Verbal examples communicating those meta-cognitive signals are phrases like: *'I'm thinking'*, *'Let me see'* or *'Well...'*. A nonverbal example of signalling meta-cognitive activities are looking up or closing the eyes.

3.4.2 Cultural Influence on Communicative Functions

Culture has an influence on the process of generating communicative functions. For example, one study of cultural differences between Italian and Icelandic people in conversation concluded that Icelandic subjects *take turn* more frequently than Italian subjects [1]. Other examples might include difference in turn speed and duration. Some cultures are accustomed to have long monologues while other cultures are used to fast turn taking. Also, in one culture it may be a custom to intentionally show your emotions more often.

These influences act at a level of mental processes that fall outside the scope of this research. Therefore, these cultural differences will not be handled.

3.5 Communicative Nonverbal Behavior

Now that concepts like beliefs, intents and communicative functions have been handled, it's time to take a closer look at *communicative behavior*. Nonverbal communication is the largest element of communication. Different studies show that a large percentage of total communication is nonverbal as opposed to verbal communication. In one study the percentage of nonverbal communication is as

high as 93% of which 55% is through facial expression, posture and gesture and 38% is through tone of voice [57].

Nonverbal communication encompasses a wide range of behaviors which can be broken down into different categories, some of which are complete studies of their own. The different components will be described in the sections 3.5.1-3.5.7. Section 3.5.8 provides a categorization of nonverbal behaviors into different types of communicative functions as developed by Ekman and Friesen [19]. Cultural influences will be explored in each section.

3.5.1 Kinesics

Kinesics is the study of body language, consisting of all nonverbal behavior related to movement, using either part of the body or the body as a whole. Kinesics is the largest part of nonverbal behavior. It includes facial expressions, gestures, posture, locomotion, stance and other body movements. The various meanings communicated through these body movements can vary greatly among different cultures. A lot of times this can cause conflicts during cross-cultural communication. A body movement expressing something positive in one culture could be offensive in another culture. Some aspects of kinesics are described individually below.

Gestures

A gesture is the motion of the limbs or body made to express or help express thought or to emphasize speech. Different types of gestures exist. Gestures may be categorized according to their cognitive construction, whether and how they are represented in memory [6]. Some gestures are *coded* and represented in memory by linking a gesture expression (signal) to its meaning. These include emblems or beat gestures structuring discourse. Other gestures are creative and invented on the spot. Examples of *creative* gestures are iconic and metaphoric gestures describing concrete or abstract objects and actions [56, 10].

Looking at *coded* gestures, some vary between cultures like many emblems, others are biologically determined, for example those which are a ritualization of physiological movements like raising the fist up to show elation [16]. As for creative gestures, many of those are probably universal because they consist of imitating objects or actions. But when those objects or actions become culturally determined, the creative gestures will also become culturally dependent.

Facial Expressions

A facial expression is made up of contractions of facial muscles to create an expression. They are used to communicate different types of information. Examples are squeezing the eyes, raising or lowering the eyebrows, showing emotions, etc.

Some of these expressions can be universal in use while others are not. For example, facial expressions giving information about the world could be universally understood. For example squeezing the eyes to signal something is small or difficult. Facial expressions resulting from meta-cognitive activities like remembering something by raising the cheeks might be universal because meta-cognitive activities are biologically determined.

Ekman and colleagues have shown that the six basic emotions are universally understood [18, 20]. This also means that there may be other emotions that are culturally determined. Facial expressions resulting in affective displays are described in more detail in section 3.5.8.

Postures

A posture can be defined as the positioning of the limbs or the carriage of the body as a whole. Postures can consist of standing postures, sitting postures etc.

In different cultures people can have different kinds of postures or postures can have different meanings. In some cultures a certain posture might be considered poor or showing little confidence, whereas in other cultures they might be considered normal. Postures could also be determined by the climate or living environment of a culture's country.

Locomotion

Locomotion can be defined as the style of physical movement in space. This includes walking, running, etc.

Cultures and groups can have different styles of locomotion. For example, white American's usually bounce when they walk with swinging arms and rolling pelvis [40]. Also the average walking speed may differ between cultures [52].

3.5.2 Proxemics

Proxemics is the study of nonverbal behavior concerning the use of space and distance. The term was first introduced by Edward T. Hall in 1966 to describe distances between people as they interact [25]. He states that the social distance

between people correlates with the physical distance and he identifies four zones of personal space:

- Intimate space (6-18 inches), used among partners or very close relatives.
- Personal space (1.5-4 feet), used among good friends.
- Social space (4-12 feet), used among acquaintances.
- Public space (12 feet and more), used for public speaking.

Although the correlation between social distance and physical distance is generally universal, Hall notes that cultures maintain different standards of personal spaces. In countries where individualism and privacy are important factors (United States, Germany or the Netherlands) people usually use larger distances. In Latin or Arabic countries people use smaller distances while interacting. Also Southern European countries use smaller distances than Northern European countries. On the other hand, in Brazil the relationship between the interactants hardly influences their proxemic behavior and people usually stand less than a foot away while interacting, no matter what their relationship is. It is important to know these differences. People can feel uncomfortable or threatened when their space is being violated when others come to close, or they can feel offended or mistrusted when others stay too far.

3.5.3 Haptics

Haptics is the study of touching behavior. This includes every behavior that includes touching someone. Examples are a handshake, a kiss, holding hands, a pat on the shoulder, a kick, an arm brush, etc. Every touching behavior gives off a nonverbal message to the other person and can cause positive or negative feelings at the receiver.

Several researchers have proposed different categorizations for haptic behavior. Heslin (1974) has defined five haptic categories [29]:

- Functional/professional, used in professional or formal settings.
- Social/polite, used in social settings.
- Friendship/warmth, used between good friends.
- Love/intimacy, used in interpersonal relationships.
- Sexual/arousal, used in sexual settings.

Jones and Yarbrough have identified 18 different meanings of touch grouped in seven types: *Positive Affect*, *Playfulness*, *Control*, *Ritual*, *Hybrid*, *Task-Related* and *Accidental Touch* [37].

The most important difference of haptic behavior between different cultures is the frequency of a haptic behavior occurring during a conversation. Remland and Jones (1995) conducted a study recording the number of touching behavior of people from different countries. They found that touching was relatively rare in England (8%), France (5%) and the Netherlands (4%) compared to Italy (14%) and Greece (12.5%) [65]. In Arabic cultures exist a high degree of haptic behavior. It is not uncommon for Arab males to kiss each other on the cheek or walk hand in hand on the street. In countries with a low degree of haptic behavior, this may be seen as act of intimacy. These differences in touching behavior around the world have led to the idea of *contact* and *noncontact* cultures.

3.5.4 Oculesics

Oculesics is the study of eye behavior. This includes behaviors like eye contact, eye avoidance, gazing, glancing and all other eye movements. Eye behavior may perform different functions. It can have a cognitive function, for example many people glance away while they are thinking. Eyes can perform a monitoring function. We can monitor our communication impact by monitoring other peoples' feedback. The eyes can also perform regulatory functions like showing the willingness to respond by gazing or avoiding eye contact. At last the eyes can perform an expressive function, by offering insight to our emotions and feelings.

Eye behavior can have different meanings in different cultures. For example, lowering the gaze may convey respect in some cultures, whereas in other cultures it may be understood as evading or insulting. On the other hand, direct eye contact may be insulting in some cultures, whereas it conveys attention in others [53]. The frequency and duration of eye contact can also have different norms among cultures. American women may feel insulted or embarrassed when they are being looked at for a long time by Italian or French men, while Italian and French women may perceive American men as cold because of their short glance towards them [74].

3.5.5 Paralanguage

Paralanguage, also known as vocalics, is the study of vocal cues. It consists of features that accompany speech and contribute to communication by modifying

meaning or conveying emotion. Paralinguistic elements include the pitch, tone, volume, rate, rhythm, articulation and intonation of speech. The use of pause and silence is also considered to be part of paralanguage. Essentially all non-verbal cues that contribute to how something is being said can be considered as being part of paralinguistics. These vocal cues can be a very important part of communication and can change the meaning of words that are communicated. Take for example the notion of sarcasm. Saying something with a sarcastic tone can make the meaning the complete opposite of what is being said with words.

Cultural differences also exist in the use paralinguistic elements. People from Arab countries for example tend to speak louder and at a higher rate than American people. In Latin countries people use a softer volume and speak at a slower rate. Cross-cultural communication can cause misinterpretation when a certain volume of speech during communication is considered as aggressive or secretive. In Japan people often have long pauses during communication, while this can be very uncomfortable for Americans.

3.5.6 Chronomics

Chronomics is the study of time. It tells us something about our attitude towards time. Example behaviors influenced by our use of time are waiting, pausing, how fast or slow we express behavior, etc.

The attitude towards time can be very different in different cultures. For example, Americans are extremely time-conscious and are not used to waiting for longer periods of time. It is expected that people are in time for their appointments and a silence in a conversation is usually experienced as an uncomfortable situation. They also have a faster average walking pace than a lot of other cultures. Brazilians on the other hand have a more relaxed attitude towards time and a last minute change of plans is common. They have more patience in waiting situations. Their average walking pace is also slower.

A study comparing the pace of life in large cities from 31 countries around the world has shown that the pace of life was fastest in Japan and countries of Western Europe and the slowest in economically undeveloped countries. The pace was also faster in colder climates and individualistic cultures [52].

3.5.7 Miscellaneous

There are other aspects of nonverbal behavior which will not be covered here in detail and also will not be considered in the remainder of this document. For example, *Olfactics*, the study of smell. Odors when present can influence

or cause someone's behavior. Another aspect is *Adornment*, the way people present themselves through clothes, jewelry, hairstyle, etc.

3.5.8 Classification System

Ekman and Friesen (1969) [19] have developed a classification system which classifies body movements into five types which have certain communication functions. These types are described below. Each type gives examples of cultural differences in the expression of the function type in question.

Emblems

Emblems are nonverbal messages that correspond to a single word or phrase. The message has a meaning on its own and is usually communicated without its verbal counterpart.

Emblems are highly culturally dependent and can sometimes even have opposite meanings. For example, the V-sign (forefinger and middle finger erect), is the British sign for victory. In Australia this sign can be seen as insulting, it symbolises the number two in the USA and can be a symbol for peace in other cultures. The O-sign (a circle created with the thumbs and the forefinger) means 'zero' or worthless in France, whereas in a lot of cultures it is the sign for 'OK'. Thumbs up is another emblem and is a positive sign in a lot of western cultures but it is a very obscene sign in some Middle-Eastern countries and parts of West Africa. Because emblems are used sporadically and in a very specific context, they usually don't lead to large conflicts when they are wrongly interpreted during inter-cultural communication.

Illustrators

Illustrators are nonverbal cues that accompany or reinforce a verbal message. They are directly linked with words and unlike emblems have no meaning on their own. Illustrators encompass a wide range of different gestures. Examples include pointing to refer to something, drawing a picture in the air to clarify something or illustrating a verbal movement.

While some illustrative gestures are universally understood, a lot of cultural differences exists when using illustrators. Take for example pointing. In the United States people point with their index finger while this is impolite in the Middle and Far East. In Germany people point with their little finger, in Japan people point with their whole hand and Kiowa Indians don't use their arms at all but point with their lips.

Not only difference in style of illustrators exist between cultures, also difference in interpretation of the frequency of these gestures. In some Asian cultures extensive use of illustrators is interpreted as a lack of intelligence, whereas in Latin cultures the absence of illustrators can be interpreted as a lack of interest. Because these gestures are being used and interpreted more subconsciously than emblems, they can easily cause conflicts that can be harder to detect.

Affective Displays

Affective displays are movements of the body, but most notable facial movements that display a certain affective state or emotion like fear, joy, disgust or anger. Ekman and his colleagues have found that people from different cultures can identify the expressions of the six basic emotions: happiness, sadness, anger, fear, surprise and disgust [18, 20]. Keltner later added the expressions of embarrassment and shame [38].

Even though the facial expressions are universally understood, there are still cultural differences in the degree and frequency of these facial expressions. For example, Italians are very expressive when using gestures, including facial expressions, but Japanese people hardly express their emotions. So even though people from these cultures may feel the same degree of a certain emotion, the degree of expressiveness may be very different. There may also be cultural differences in the expressed emotion triggered by some event. The same event in one culture may lead to joy, while in another culture it may lead to sadness.

Another example of the cultural differences in the usage of emotion expressions is the result of a study involving two cultures: American and Japanese. Subjects from both cultures were shown films inducing different emotions. There were two different settings. In one setting there was an authoritative figure present, in the other there was not. In the presence of this authoritative figure, the Japanese subjects masked their negative expressions by positive ones. Some Americans did the same, but at a much lesser degree. This study proves that the difference in showing some expressions is not due to the difference in the emotional feeling, but due to the cultural differences [18]. Evidence from other studies are consistent, supporting the idea of universal *meaning* of expression but cultural differences in the *usage* of the expressions [41].

Regulators

Regulators are nonverbal cues which regulate, modulate and maintain the flow of information during interaction. These include cues regulating turn-taking,

speaking, listening and giving or requesting feedback. Behavior resulting from these cues include eye movements, nodding or shaking of the head, raising an eyebrow, etc.

Cultural differences in regulatory cues may be highly confusing during communication. For example while in most countries nodding the head means yes or shows agreement and shaking the head means no or shows disagreement, in Bulgaria the opposite is true.

Adapters

Adapters are body movements that people use to adapt to an environment, to release bodily tension, to feel more comfortable or to perform a specific physical function. Adapters are performed at a low level of awareness. Examples include changing posture, yawning, scratching your head or tapping your fingers.

Because people can be unaware of carrying out an adaptive movement, it could easily be the cause of misinterpretation when talking to someone from another culture. For example, showing the soles of your shoes while taking on a more relaxed seating position could be very insulting because showing your shoe sole is a very offensive emblem in many Arabic cultures.

4. THE SAIBA FRAMEWORK

The last chapter gave an overview of the process of how humans go from beliefs they want to communicate to the expression of multimodal behavior and what factors are involved in this process. This chapter introduces the use of ECAs to simulate this process.

Embodied Conversational Agents (ECAs) are becoming increasingly adept at having multimodal interactions and trying to achieve the goal of simulating human-to-human interactions. Many capabilities are required by these ECAs which has resulted in more and more sophisticated architectures. Some of these capabilities include the interpretation and generation of verbal and nonverbal behavior, the ability to follow the norms that regulate social interaction, the modeling of cognitive and emotional processes and the animation of the agent's body [76]. The focus of the problem this framework addresses is on the generation of verbal and nonverbal behavior that carries out an agent's communicative function. The framework consists of processing stages interconnected with XML based interfaces, which describes the agent's function and behavior.

The proposed multimodal behavior generation framework will be outlined in section 4.1. The *Functional Markup Language (FML)* describes the agent's communicative functions and will be described in section 4.2. The *Behavior Markup Language (BML)* describes the agent's communicative behavior and will be described in section 4.3. Section 4.4 describes the proposed *BML Gesticon*, which is a repository for predefined behavior described in BML. This chapter concludes with an overview of the most recent developments of the SAIBA/FML/BML research project in section 4.5.

4.1 SAIBA

The SAIBA framework (Situation, Agent, Intention, Behavior, Animation) is the result of an international effort to unify a multimodal behavior generation framework for Embodied Conversational Agents. It was designed to identify the problem of multimodal generation, to define planning stages involved in

multimodal generation and to identify the knowledge structures that mediate between those stages [43]. Its key design goal was to make a clear distinction between an agent's communicative function and its behavior. This would lead to support for realizing the same communicative function in multiple ways, depending on the physical state of the agent, the environment or different social en cultural settings [76].

The SAIBA framework consists of three processing stages which lay down a general structure for every multimodal behavior generation system. The framework is illustrated in figure 4.1. The first processing stage consists of the planning of a communicative function, followed by the planning of a multimodal realization of this function (behavior planning), ending with the realization of the planned behaviors. The stages are bi-directionally linked to each other. One stage delivers input to the next stage and gives feedback to the previous stage. The focus of this framework is not on the internal structure of each stage but rather on the data that is being processed at each stage together with the information flow between the stages and how this information should be represented.

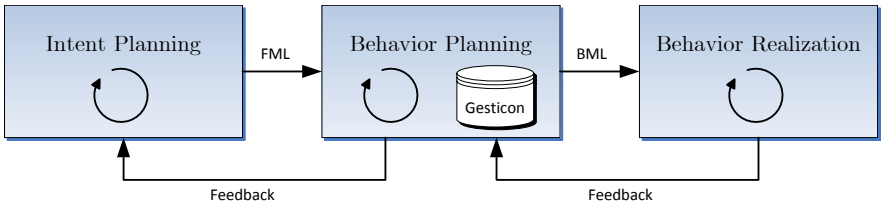


Fig. 4.1: SAIBA framework for multimodal generation.

Two languages have been proposed to account for the information flow between the planning stages. These languages are independent of a particular application or domain and independent of any graphics and sound model. The first language is called the *Functional Markup Language* (FML). It is used as the interface between the Intent Planning and Behavior Planning stages. Its function is to describe the communicative function of an agent without reference to any physical behavior. For example conversational regulators such as turn taking can be specified in FML. Not only is it possible to define basic semantic units associated with a communicative event, it is also possible to annotate these units further with properties that further describe the communicative event.

The second language is called *Behavior Markup Language* (BML). This language is used as the interface between the Behavior Planning and Behavior

Realization stages. In BML, multimodal behavior can be described that can be used to control an agent. The behavior realization depends on the particular realization model and can be very diverse. Animations for example can be procedural or fixed and chosen from a repository. Sounds can be generated by a text-to-speech engine or played from file. Therefore the descriptions that can be specified by BML are of a particular level of detail where descriptions are independent of any realization model. The framework also presents a Gesticon, that can be used by the Behavior Planner. This Gesticon is a dictionary which could contain predefined BML behavior definitions.

The next sections will describe *FML*, *BML* and the *Gesticon* in more detail.

4.2 Functional Markup Language

The *Functional Markup Language (FML)* specifies the communicative functions of an ECA. These include performative, discursive, affective, adjective, deictic and meta-cognitive functions as described in section 3.4.

As of this point, there isn't an official specification of FML and the language is a work in progress. An initial set of tags representing communicative functions have been specified. It is expected to grow as researchers develop more complete models of human communicative behavior.

This document provides an unofficial FML specification which proposes some tags, attributes. The specification is heavily based on the communicative function model described Poggi et al. [63]. The FML specification proposal can be found in appendix A.

4.3 Behavior Markup Language

The *Behavior Markup Language (BML)* specifies the verbal and nonverbal communicative behaviors of an ECA. This includes speech, body movements, lip movements, gestures, animations, i.e. every behavior that could be the result of a communicative function. The elements of BML roughly correspond to the body parts involved in the behavior. They can be further defined through the use of attributes. Just like FML, BML is a work in progress. The current set of BML elements used for behavior specification are described below. For a detailed specification of the BML elements see appendix B.

BML Elements

- **Body:** This element handles body movements including approaching, orientations and postures.
- **Torso:** This element handles the orientation and shape of the spine and the shoulders. A posture name and an animated transition that gets played before the posture is assumed can be specified.
- **Legs:** This element handles movements of the body elements downward from the hip (pelvis, hip, legs, knee, toes and ankle). A posture name and an animated transition that gets played before the posture is assumed can be specified.
- **Head:** This element handles all head movements like nodding, shaking, tossing or orienting the head towards a specified target. The speed as well as the number of repeats (in case of a nod, shake or toss) can be specified.
- **Gesture:** The gesture element can be used for gesture movements like pointing, reaching, depicting, signaling or beat gestures. In case of a point or a reach gesture a target can be specified.
- **Face:** The face element handles all facial expressions. It controls the eyebrows, the eyelids and the mouth. Both sides of the face can be controlled. Also a reference can be given of an Action Unit (AU) of the Facial Action Coding System (FACS) developed by Ekman and Friesen in 1978.
- **Gaze:** The gaze element controls the eye behavior, indicating where the character is looking. The amount to use the eyes, neck, spine and body can be specified when looking at the character as well as an angle offset together with an offset direction.
- **Lips:** This element controls the lip movement. A viseme (a basic unit of speech in the visual domain) can be specified together with an articulation value. Also mouth flapping (opening and closing of the mouth) can be controlled.
- **Speech:** This element can specify the text to be spoken, a reference to an external object containing speech and the type of the referred object.

Timing

Strict timing constraints can be specified when defining behavior using BML. Every behavior can have up to six phases defined by seven points in time. These phases are illustrated in figure 4.2. The first phase occurring between the synchronization points **start** and **ready** is the initial movement of the behavior into gesture space. This is the movement the agent needs to perform to enter the state and position from which the behavior can start. The movement at the end of the behavior to go back to the neutral or previous state occurs between **relax** and **end**. The actual behavior occurs between **ready** and **relax**. The most effortful phase of the behavior takes place between **stroke-start** and **stroke-end** with **stroke** being the point of maximum effort.

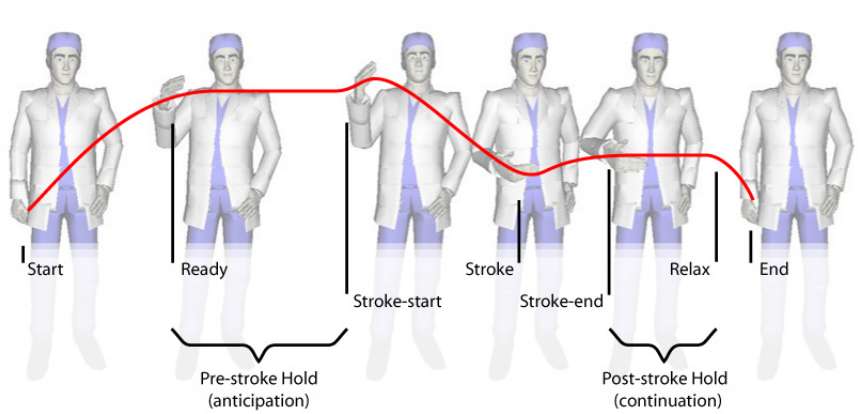


Fig. 4.2: Synchronization points of Communicative Behavior. Figure taken from [43].

These points can be specified in two ways. They can be time-points, describing the time calculated from the first time-point of the behavior, i.e. when the behavior started. They can also refer to another synchronization point from another behavior. This makes the creation of complex multimodal behavior possible. For example consider the composite behavior of nodding the head in combination with a beat gesture. It's possible to start the nod behavior when the beat gesture is at its stroke point by synchronizing the start point of the nod behavior and the stroke point of the beat gesture. However, it's not necessary to specify every synchronization point in a behavior. It's sufficient to specify only a start and end time, but even these attributes are optional. If no points are specified, the BML processor just starts the behavior as soon the behav-

ior is available. It is possible for the BML processor to receive multiple BML elements at the same time that cause conflicts. These conflicts can be solved through the use of the priority attribute of the conflicted BML elements. The BML processor then executes the element with the highest priority.

4.4 BML Gesticon

The *BML Gesticon* is a behavior repository with entries that are defined in BML. The *Gesticon* representation is derived from the work of Pelachaud, Krenn and Pirker [46, 45]. The original representation of the Gesticon also contained a meaning/function description for each Gesticon behavior. This has been excluded in the *BML Gesticon*, because in this framework, meaning should be handled by the *FML* language. The *BML Gesticon* is in an early stage of development and its specification is dependent on future work and research.

4.5 Current Status

During the most recent BML workshops of *Vienna 2006* and *Paris 2007*, different aspects of BML have been discussed. A proposed addition to the BML specification is the addition of *Description Levels* inside each BML tag. A description level can be used to specify more detailed behaviors than the BML core level 0. These more detailed descriptions could be represented in a different xml-based behavior representation level like MURML [44] for gestures or SSML [35] for speech.

Most recently, the naming of the core BML elements have been reviewed and a few new elements have been proposed. It is good to remember that BML is a language in the making and therefore in this research it is tried not to be dependent too much on any details of BML.

Part III

FINDINGS

5. CULTURE INFORMATION MODEL

Now that the literature on communicative functions and communicative non-verbal behavior as well as the cultural influences have been reviewed, a start can be made to look more closely at the different types of cultural information that would be interesting to model in ECAs. An example of what is meant by *cultural information* in this chapter might be information like: *"Italians use a lot of gestures"* or *"It is common for Arabic people to touch each other during communication"* or *"In France, it not uncommon for a man to gaze at a woman for a longer period of time when passing by"*. Each of the examples above talk about the behavior people from a specific culture express. These could be seen as *Behavior Rules*, which give information about a commonly shared pattern of behavior in a certain culture. We only concern ourselves with *this* type of information, i.e. information about *behavior expressions*. Information on communicative functions and meanings is not considered. For example, this excludes information like: *"Icelandic people take-turn more frequently than Italians [1]"*. This kind of information talks about the cultural differences of aspects which *cause* behavior. As explained earlier, this kind of information falls outside the scope of this project.

In this chapter an information model will be presented in which communicative cultural behavior expression information can be categorized and behavior rules can be designed. This model will be used to illustrate clear distinctions between different types of information and will help to give a clear overview of the information categories that can be used. This might proof to be very useful when attempting to model cultural information into a cultural behavior generation framework from ECAs.

To stay within the scope of this research, the focus will be on information concerning cultural differences of the expression of behavior and not so much on the cultural differences of meanings that can be interpreted by perceiving behavior. To explain this, take for example the emblem gesture sign which is represented by a circle created with the forefinger and thumb. This gesture may be used by *Americans* to signal they're OK. On the other hand, this sign is an

obscene gesture in many *European cultures*, having sexual implications. Now, the expressions of these meanings are handled by this model, but not the cross-cultural conflicts that may arise when the expressions are interpreted by people from different cultures. In other words, only *behaviors* caused by meanings are handled and not *meanings* that may be interpreted from observed behaviors.

In section 5.1 a *Culture Information Model* will be proposed that will guide the design of a cultural behavior generation framework for ECAs. Section 5.2 provides templates for every information category in the information model. A template can be used to structure the information of one cultural communication aspect and to create a *Cultural Behavior Rule*.

5.1 Information Model

In this section an information model for cultural aspects of communication is presented. This model creates the ability to design so-called *Cultural Behavior Rules* which apply to specific situations or domains. The model is illustrated in figure 5.1. The different components of the model are described in the sections below. Examples will be given to get a better understanding of the model.

Note that not all examples that are given in this chapter are universally recognized facts about the cultures in question. Some examples are conclusions of cultural studies or experiments, but other studies might have different conclusions. Some examples have also been oversimplified. The goal here is not to concern ourselves with the correctness of the information, but to provide a model in which all different *types* of information can be modeled.

The first distinction that is made in the model is that every information is either *Behavior Information* or *Style Information*. *Behavior Information* contains a description of some *(Multimodal) Behavior* that is being expressed, while *Style Information* contains a description of some *Style* aspect of behavior that is being expressed. In other words, *Behavior Information* describes *what kind* of behavior is being expressed while *Style Information* describes *how* or *how often* some behavior is being expressed.

To clarify this distinction, some examples of *Behavior Information* are:

- *"Maoris, the indigenous people of New Zealand, traditionally greet by pressing their noses against each other."*
- *"People from Bulgaria nod their head to show disagreement."*
- *"The gesture of putting your fingers and thumbs together, kissing your fingertips and then opening your hand is a typical French gesture to express*

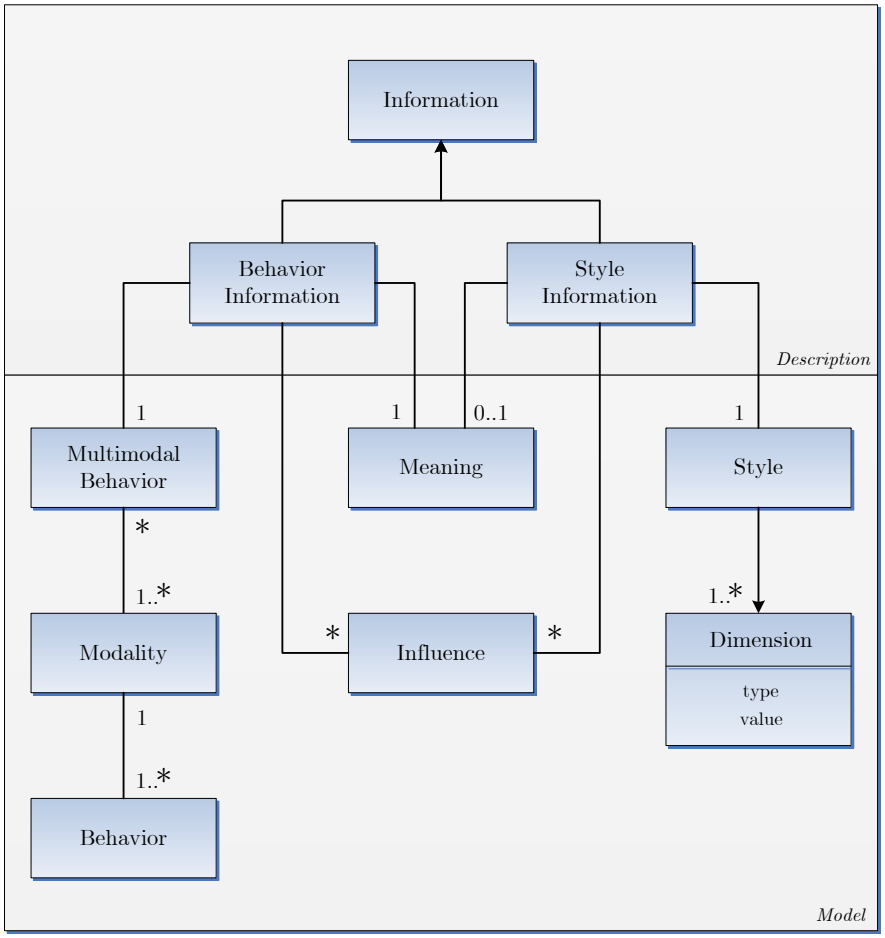


Fig. 5.1: The Culture Information Model.

how delicious something is."

These examples all describe a certain behavior that is being expressed: *pressing of the nose*, *nodding of the head* and *performing an emblem* respectively. They also describe a certain meaning that is communicated: *greeting*, *showing agreement* and *expressing deliciousness* respectively.

Some examples of *Style Information* are:

- *"Middle Eastern people express their hand gestures more expansive than Americans do."*
- *"Israelis typically speak loud and fast."*
- *"People from India point with their chin."*

These examples don't describe a certain behavior but describe a *style* of behavior. They describe *how* a behavior is being expressed. The styles are: *'a large spatial extent'*, *'loud and fast'* and *'the use of'* respectively. The behaviors to which the styles are applied are: *all hand gestures*, *speech* and *pointing with the chin*.

From the model, it can be seen that each type of information can be modeled using three *classes*. Besides the specification of *Multimodal Behavior* and *Style*, they also contains links to additional information, namely *Meaning* and *Influences*. The *Multimodal Behavior* and *Style* specifications are described in section 5.1.1 and section 5.1.2 respectively. The *Meaning* and *Influences* information is described in section 5.1.3 and section 5.1.4 respectively. At last section 5.1.5 describes the usage of styles, i.e. when and on what behavior they can be applied.

5.1.1 Multimodal Behavior

Each behavior described in a *Behavior Information* is defined in the model as a *Multimodal Behavior* element. Each *Multimodal Behavior* description can contain behaviors across one or more modalities. Modalities include for example gestures, facial expressions, eye movements, etc. For each modality that is used in the information, one or more behaviors can be specified belonging to that modality. Take for example the following information: *"People emphasize words by raising their eyebrows, looking at the interlocutor and performing a beat gesture"*. In this example a multimodal behavior is given, describing behavior across three different modalities, namely *facial expressions*, *eye movements* and *hand gestures*.

5.1.2 Style

In the scope of the model described in this chapter, *Style* is defined as a combination of one ore more dimension values. An example of some dimension types are the *expressivity dimensions* as described by Pelachaud et al. [62] consisting of *spatial extent*, *temporal*, *fluidity*, *power* and *repetitivity*. Style dimensions can also include dimensions representing paralinguistic features like *volume*, *rate*, *articulation* or *pitch*. Another dimension could be *activation* which determines if a behavior should be expressed or not. Using these dimensions, it is possible to define a style for some behavior which can be used to tune previously defined behavior to a desirable style.

Styles on their own are not useful. They need to be applied to one or more behaviors. For example, the style dimension *spatial extent* can be applied to behaviors from different modalities, making it possible to change the expressiveness of facial expressions or the gesture space. Note that some style dimensions might work differently for different modalities. Referring to the example above, the *spatial extent* dimension is applied to facial expressions by changing muscle contractions, while for the gestures, the different joint angles might be changed. Other style dimensions can only be applied to a specific behavior. For example, paralinguistic style dimensions can only be applied to speech.

Styles are a powerful tool because they can be applied to a whole range of different behaviors in different situations. How styles can be applied to behaviors is described in section 5.1.5.

Dimension Scales

A style is made up of any number of dimension values handling different dimensions. Dimensions are represented as a scale ranging from zero to one. Take for example the vocal volume dimension where a *zero value* results in no volume and a *one value* results in full volume. The issue that arises is the meaning of a value on the dimension scale. Is it a linear scale or a logarithmic scale? What represents the maximum value? This can be different for each dimension type and each modality to which the dimension should be applied. The main advantage of using dimensions is that many different types of styles can be defined as long as they can be specified on a dimension scale. How a scale might be interpreted by each dimension type depends on the styles that are specified and falls outside the scope of this research. A few example dimensions are shown in figure 5.2.

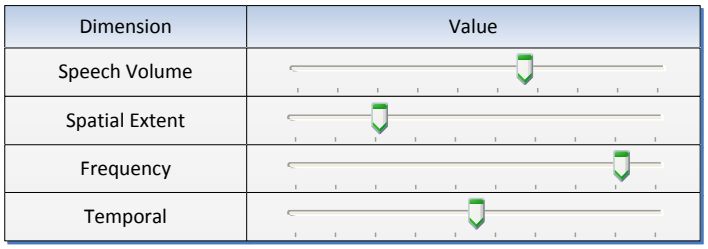


Fig. 5.2: A few examples of Style Dimensions.

5.1.3 Meaning

What is meant by *Meaning* in this model is that the information that is being modeled can contain a description of a communicative function or meaning that is being communicated. If this information is present, it describes the behavior or style being expressed *only* when communicating a specific meaning.

From the model in figure 5.1 it can be seen that *Behavior Information* must always contain a *Meaning* while this is optional for *Style Information*. This is best explained by showing some examples. Take for example the *Behavior Information*: *"Italians use the 'bag hand' gesture when emphasizing something"*. This example not only gives us information about the specific behavior to express (the 'bag hand' gesture), but also *when* to express this behavior, namely when emphasizing something. If this meaning was excluded, the information would not be very useful.

Style Information on the other hand does not necessarily need to have a *meaning* associated with it. Take for example the information: *"People from Arabic cultures gaze much longer at their partners than Americans do"*. Now this information describes a style (*much longer*) and the behavior to which to apply this style (*gazing*). It does not give us any information for which *meaning* to apply this style. It just states, every time people from Arabic cultures gaze at their partners, they do this much longer that Americans do.

Although *Style Information* doesn't have to be associated with a *meaning*, it is possible to include one. Take for example the information: *"In Japan, when greeting someone of a higher status, people bow deeper and longer"*. Here, the described style is *deeper and longer*, the behavior to which to apply this style is *bowing*, but it also describes for what meaning this style needs to be applied, namely *greeting*.

5.1.4 Influences

The influences described here correspond to the influences described in section 3.3. To recapitulate, these factors tell us by who and in what situation some behavior or style is expressed. These factors can be *intrinsic factors*, like personality, culture, age or gender. They can also be *contextual factors*, like the cognitive state the speaker is in, the relationship with the interlocutor or the social setting of the conversation. A simple model of these influences is illustrated in figure 5.3.

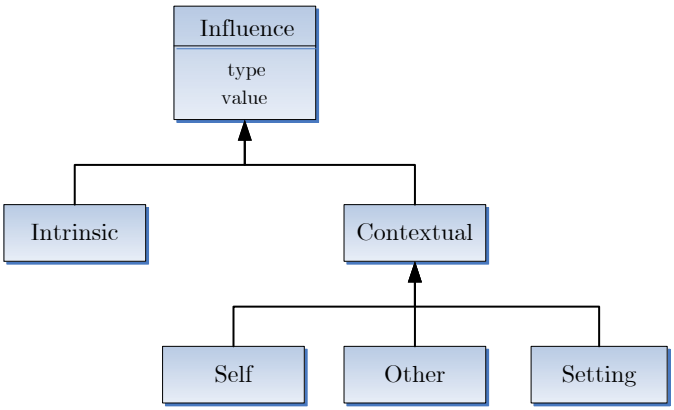


Fig. 5.3: The Influence Model.

As can be seen in the *Information Model* in figure 5.1, a behavior or style description can contain zero or more influences. The more influences are described, the more specific the information description becomes. Take for example the hypothetical information: *"Dutch male children are accustomed to maintain a large distance when they interact with an adult of a high status"*. This information gives a lot of influences. The described style (large distance) should only be applied when the actor is *Dutch*, is a *male* and is a *child*, while the interlocutor is an *adult* who has a *higher status* than the actor.

On the other hand, information might also be universal and general with very few or no influences. Take for example: *"People take on a more active posture when the environment is very cold"*. This rule should only be applied when the *environmental temperature* is low. Or an example with no influences: *"People lower their eyebrows when they express anger"*. This example of *Behavior Information* only contains information about the behavior to express and the communicative function specifying when to express the behavior. There are

no additional factors that influences this behavior.

5.1.5 Applying Styles

Styles can be applied to different configurations of behaviors in different situations. In which situation a style needs to be applied depends on the information concerning the *Meaning* and *Influences* that is given in the *Style Information*. Take for example the information: *"In Puerto Rico people maintain less visual contact to show respect"*. The style described here (*less activation*) should only be applied in the situation where the communicative function is *showing respect* and the cultural identity of the person in question is *Puerto Rican*.

Every *Style Information* needs to provide a behavior or group of behaviors to which the style needs to be applied. For example, in the information illustrated above the behavior is *eye contact*. This involves a single behavior consisting of the eyes looking at something. There are also styles which could be applied to groups of behavior, for example to all *gestures* or *facial expressions*. Different types of groups could be defined. These different groups are outlined below.

All Behavior

First of all, if desired, a style can be applied to every behavior that is expressed across every modality. Take for example the *Style Information*: *"Introvert people express less verbal and nonverbal behavior than extrovert people"*. Here, the style is *less activation* and it should be applied to all behavior, both verbal and nonverbal, making it less likely that a behavior will be expressed.

Unimodal Behavior

Next, a style can be applied to a specific unimodal behavior. For example, it is possible to specify the *fluidity* of a single specific hand gesture like waving, or the *spatial extent* of a specific facial expression like showing anger.

Multimodal Behavior

A style can also be applied to a specific multimodal behavior. This might be useful in situations where a style is desired for a specific multimodal behavior in some situation. For example, it is possible to modify the *temporal* aspects of a behavior consisting of a specific hand gesture in combination with a specific facial expression.

Modality

Further, styles can be applied to all behaviors belonging to a specific modality. Modalities here can range from high level modalities (face, gaze, gesture, posture, head) to low level modalities identifying specific body parts. Extra information could also be modeled like applying styles to the left or right side of high or low level modalities. For example, it is possible to specify the *fluidity* of all hand gestures or the *power* of all eye movements, but also the *use of* the forefinger or the *restriction* of the use of the left arm.

Behavior with Meaning

Besides categorizing behavior into modalities, behaviors can also be categorized into the meanings they transfer. These could be low level communicative functions like expressing emotions, but also higher level meanings like grouping all behaviors representing an emblem. Another example is the grouping of all behaviors involving ritualistic touches like greetings and departures.

5.2 Information Templates

In this section, information templates are introduced. These templates provide a way to structure pieces of information that contain some cultural aspect of communication. There are two templates, one to use for *Behavior Information* and one to use for *Style Information*.

With these templates, a huge range of cultural behavior information can be modeled. The most notable cultural differences that can be specified using these templates include: the expression of emblem signs used in different cultures, different greeting forms between people that differ in culture, relationship, age or gender; the general amount of gestures used, gesture space, frequency of haptic behavior, proxemic standards, paralinguistic features, gaze behavior, the use of facial expressions, restriction or necessity of specific body parts and much more.

The *Information Model* is used here for cultural aspects, but the functionality and domain of the model does not have to be restricted to just cultural aspects. Basically all information about nonverbal behavior in some given situation (which is determined by the *Meaning* and *Influence* information) can be modeled. For example, simple universal rules concerning communication that could be modeled might be: *"People talk louder when they are in a noisy environment"* or *"People are more likely to use gestures and other nonverbal*

behavior when they are in a noisy environment" or *"People that are very tired are more likely to talk slower"*. So the *Information Model* described here might proof to be useful for modeling other types of information as well.

The templates in combination with some examples are illustrated in the sections below.

Behavior Information	
Information	People from Bulgaria shake their head to show agreement.
Meaning	Show Agreement
Intrinsic Influences	Culture: Bulgaria
Contextual Influences	
Multimodal Behavior	Shake Head

Behavior Information	
Information	The 'pianist'-gesture is typically an Icelandic gesture used to emphasize things.
Meaning	Emphasize
Intrinsic Influences	Culture: Icelandic
Contextual Influences	
Multimodal Behavior	'Pianist'-Gesture

Behavior Information	
Information	In Iraq, people might put their right hand over their heart while slightly bowing when respectfully greeting someone.
Meaning	Respectful Greeting
Intrinsic Influences	Culture: Iraq
Contextual Influences	
Multimodal Behavior	Hand over Heart in combination with a slight Bow.

Behavior Information	
Information	Japanese people may mask their negative emotions with a smile when someone of a higher status is present.
Meaning	Negative Affect
Intrinsic Influences	Culture: Japanese
Contextual Influences	Presence: Higher Status Person
Multimodal Behavior	Smile

Fig. 5.4: Behavior Information Template Examples.

Style Information		
Information	Asian people use less gesture space	
Meaning		
Intrinsic Influences	Culture: Asian	
Contextual Influences		
Style Dimensions	Spatial Extent	
Apply to	☐ All Behavior ☐ Unimodal Behavior ☐ Multimodal Behavior ☑ Modality ☐ Behavior with Meaning	Gestures

Style Information		
Information	In India, when interacting with someone of an equal or higher status, people point with their whole hand.	
Meaning		
Intrinsic Influences	Culture: Indian	
Contextual Influences	Relationship: equal or higher status	
Style Dimensions	Use Of	
Apply to	☐ All Behavior ☐ Unimodal Behavior ☐ Multimodal Behavior ☑ Modality ☐ Behavior with Meaning	Hand

Style Information		
Information	People from Arabic cultures use more positive affect touches.	
Meaning		
Intrinsic Influences	Culture: Arabic	
Contextual Influences		
Style Dimensions	Frequency	
Apply to	☐ All Behavior ☐ Unimodal Behavior ☐ Multimodal Behavior ☐ Modality ☑ Behavior with Meaning	Positive Affect Touches

Fig. 5.5: Style Information Template Examples.

6. EXTENDED BEHAVIOR GENERATION FRAMEWORK

After having created a model in which different types of cultural information can be modeled, it is time to see how it would be possible to use this information to adapt ECAs in such a way that they can behave according to some cultural norm. In this chapter a behavior generation framework for ECAs will be introduced. This framework, which will be called the *Extended Behavior Generation Framework* includes components necessary for behavior generation in ECAs. Part of the framework uses the components of the *SAIBA Framework* described in chapter 4. It acts as a starting point for the design of the *Culture Framework* which is outlined in chapter 7.

As the name implies, the main purpose of the framework is to generate verbal and nonverbal behavior for ECAs. Every communicative act starts with beliefs, meaning or intents the agent wants to communicate. The reason behind the communicative act, the goals and events which cause the act, fall outside the framework. In short, the framework handles the process from the planning of a communicative act using a combination of verbal and nonverbal behaviors which are available to the ECA, to the planning and realizing of the behavior that has been chosen.

As said, the basis of the framework is the *SAIBA Framework*. The same three stages of behavior generation are included in this framework, namely the *Intent Planner*, the *Behavior Planner* and the *Behavior Realizer*. Some internal components of the stages have been added. It is tried to comply with the design of the *SAIBA Framework*. For this reason, the function and behavior languages *FML* and *BML* are used as example languages. If desired, other function and behavior representation languages could be used as well. The most important extension of the *SAIBA Framework* is the addition of a *Communication Repertoire* in which behavioral rules of an agent can be stored, i.e. meaning to behavior mapping rules. Every behavior an agent expresses will have to come from this component. This extension will be an important requirement for the *Culture Framework*, which job it will be to fill this *Communication Repertoire* with behavioral rules that might conform to a specific culture.

First some requirements for the *Extended Behavior Generation Framework* will be listed in section 6.1. In section 6.2 the framework's design will be illustrated after which every component will be individually discussed. Section 6.2.3 shows some possibilities for implementing an optional *Communicative Function Generator* which could be used to automate the process of adding a range of communicative functions to dialog texts to be spoken by ECAs.

6.1 Requirements

Communication Repertoire

As noted earlier, the main requirement of the framework is a place where all behavior an agent is able to express can be stored. This place is the agent's *Communication Repertoire*. Referring to the model of influence factors from figure 3.3, the *Communication Repertoire* contains all learned behavior, verbal and nonverbal. This is the place where all the verbal and nonverbal behavior learned and mastered by a person is stored.

The *Communication Repertoire* inside a behavior generation framework could be realized by a meaning-to-signal dictionary which can be queried when an ECA needs to generate appropriate behavior for some communicative function or meaning. A different content of such a dictionary results in different behavior of an ECA. This is very useful in designing believable agents which become more believable when the content of their *Communication Repertoire* is specifically adapted to a certain identity consisting of a person's culture, age, personality, gender, etc. How this communication repertoire will be modeled in the context of this research will become clear in the next section.

Multimodal Behavior

To make ECAs more believable, they need the ability to express multimodal behavior. This means an ECA can express multiple verbal and nonverbal behaviors concurrently, like speech, facial expressions, posture, gestures, etc. To account for this ability, the framework needs to be able to map a single function element (FML) to multiple behavior elements (BML). This multimodal behavior described by multiple behavior elements could be stored as a single behavior *macro*. To reuse these macros, they could be stored inside a *gesticon* as described in the SAIBA framework.

Variability

To make the behavior of ECAs more variable and less predictable the framework could incorporate an element of chance in deciding what behavior to choose. An ECA wouldn't be very believable if it chooses the same behavior every time it wants to express an intent in the same context. Therefore the framework must be able to choose between different behavior elements or macros based on some specified probability when deciding what behavior to express for a given function.

Reusability

It would be desirable if a specific mapping from function elements to behavior elements could be reused by multiple ECAs in different systems or domains. Therefore such a mapping should be stored in some easy file format. The obvious choice for this file format is xml because most function and behavior representation languages are also in xml format.

6.2 Design

The *Extended Behavior Generation Framework* is illustrated in figure 6.1. Below, the different modules and aspects of the framework are discussed. In the discussions, FML and BML are being used as example languages.

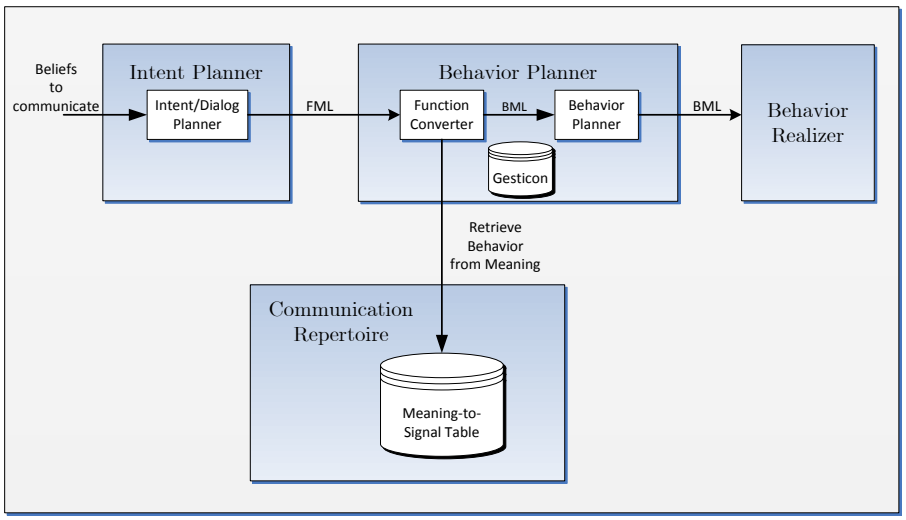


Fig. 6.1: The Extended Behavior Generation Framework.

6.2.1 Intent/Dialog Planner

To use the *SAIBA Framework* for the purpose of this research, internal designs of the SAIBA stages are given. The *Intent/Dialog Planner* is an internal module of the parent *Intent Planner* and is a very important module in the framework because its main purpose is to convert abstract beliefs and meanings an ECA wants to communicate to a concrete combination of words or semantic units and FML, i.e. it plans the next communicative act of the discourse. This module is considered to be a black box and the internal workings are beyond the scope of this research. One option is that the act is generated by some representation of the agent's mind. Another option is that the agent's conversation is fixed in the form of scripts that have been created beforehand.

An example of a communicative act in FML planned by the *Intent/Dialog Planner* is illustrated in figure 6.2.

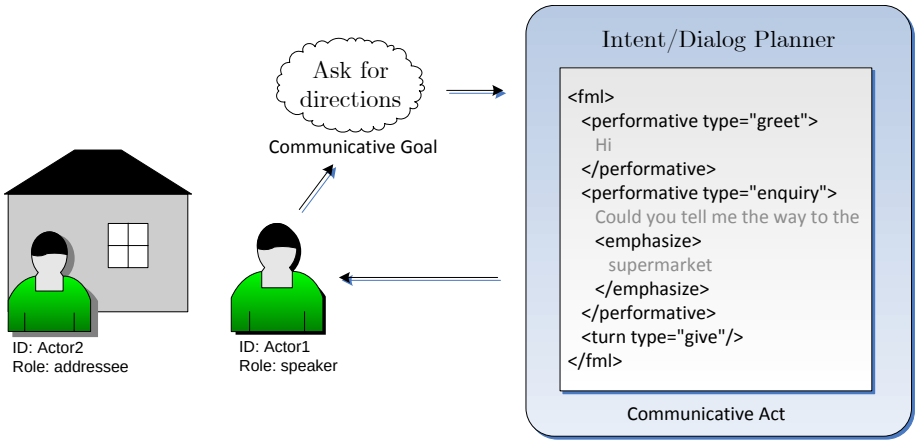


Fig. 6.2: Example of generating a communicative act.

6.2.2 Communication Repertoire

The *Communication Repertoire* is the place where the whole range of verbal and nonverbal behavior an ECA is able to express is stored. This module represents the communicative behavioral abilities of an ECA which *could* reflect the identity of the ECA, where he comes from, what he has learned, etc.

In the design, the *Communication Repertoire* consists of a storage module named the *Meaning-to-Signal Table*. This module can be seen as a table which maps meaning (FML) to behavior (BML). The content of the table defines all

```
<MeaningMappingRule key="UID" name="STRING">
   <description>
      ...description...
   </description>
   <meaning type="FML">
      ...fml element...
   </meaning>
   <signals>
      <option type="BML" probability="0.6" dependency="true">
         ...BML block specifying (multimodal) behavior...
      </option>
      <option type="BML" probability="0.4" dependency="true">
         ...BML block specifying (multimodal) behavior...
      </option>
      <option type="BML" probability="0.8" dependency="false">
         ...BML block specifying (multimodal) behavior...
      </option>
   </signals>
</MeaningMappingRule>
```

Fig. 6.3: A Mapping Rule Definition.

behavioral aspects of an ECA.

All requirements for this framework can be handled by the presence of this table: multimodal behavior can be achieved by mapping one FML element to multiple BML elements, called a *BML block*; variability can be achieved by providing different BML options with certain probabilities; reusability can be achieved by defining a table format, preferably in xml.

Meaning-to-Signal Table

Here a format for a mapping table is proposed which is able to convert a communicative *meaning*, in this case described in FML, to a communicative *signal* (behavior), here described in BML. The format is chosen in such a way that it is independent of the function and behavior languages used.

The mapping table consists of rules. Each rule is able to convert a single meaning element to one or more signal options. For every signal option a multimodal behavior can be specified consisting of one or more behavior elements. The structure of an entry in the mapping table is illustrated in figure 6.3.

Every mapping rule has a unique key and a name. The *MeaningMappingRule* element contains a *description* element, which can be used to give the rule a human readable description. After that follows a *meaning* element with a *type*

attribute which determines the language in which the meaning is specified, which is *FML* in the example above. At last there is the *signals* element which can contain one or more *option* elements. Just like the *meaning* element, the *option* element contains a *type* attribute, identifying the language in which the behavior is specified, which is *BML* in this case. Inside every *option* element there is a *BML block* in which one or more *BML elements* can be specified which make up the (multimodal) behavior. Alternatively a *BML block* can also reference a *BML macro* which could exists in a *Gesticon* for BML.

Every BML option is associated with a probability. This is the probability that this behavior will be expressed. To have more control over the multimodality a dependency flag is available. The dependency flag gives the option of which multimodal behaviors can be expressed simultaneously and which multimodal behaviors exclude other behaviors. The probabilities of all the behaviors of a mapping rule whose dependency flag is set to *true* must add up to 1.0. These behaviors cannot occur simultaneously and a choice will be made which one of them will be expressed based on their probabilities. A BML option with its dependency flag set to *false* doesn't have this limitation and the choice whether or not to express this behavior is independent of the other options and is determined solely on its own probability. In order to define a BML option which represents no behavior, an empty *option* element can be used with some probability and its dependency flag set to true. Figure 6.4 shows an example of a mapping for the *emphasize* element. This FML element can be used by an agent during communication to emphasize some phrase or word.

From the code, it can be noted that there are three options. The first two options depend on each other. There is a 70% change that a *beat gesture* will be expressed and a 30% change that the *beat gesture* will not be chosen. Besides that, the *eyebrow raise* will be expressed because its probability is *1.0* and this option is not dependent on other options. The options that have been chosen will be merged together and sent back to the *Function Converter*. Note that every option can specify a *multimodal* behavior if necessary, increasing the flexibiliy and possibilities of behavior specification.

The Use of Options

In the mapping table format described above, the possibility of choosing different signal options is introduced. The reason for this is to provide the ECA with more variety in expressing behavior. In reality, the exact behavior a person will express is not a matter of chance, but is usually dependent on influence factors, like the ones described in section 3.3. Many times it cannot be explained what

```
<MeaningMappingRule key=”a0” name=”emphasize”>
   <description>
      Emphasizing a word expressed by raising the eyebrows
      and/or performing a beat gesture
   </description>
   <meaning type=”FML”>
      <emphasize/>
   </meaning>
   <signals>
      <option type=”BML” probability=”0.7" dependency=”true”>
         <bml>
            <gesture type=”beat”/>
         </bml>
      </option>
      <option type=”BML” probability=”0.3" dependency=”true”>
         <bml/>
      </option>
      <option type=”BML” probability=”1" dependency=”false”>
         <bml>
            <face type=”eyebrows” side=”both” shape=”flat” amount=”0.75"/>
         </bml>
      </option>
   </signals>
</MeaningMappingRule>
```

Fig. 6.4: A Mapping Rule Example.

factors lead to what different behavior expressions. Therefore the use of these options is provided in the *Meaning-to-Signal Table* to still be able to give the ECA a variety in behavior expression if the influence of factors is not known or too complex.

6.2.3 Function Converter

The *Function Converter* receives a communicative act from the *Intent/Dialog Planner*. This act consists of words annotated with FML. The main goal of the *Function Converter* is to convert the whole act to BML elements. Words have to be converted to *BML Speech* elements while FML tags have to be converted to other BML tags. This last conversion is done by using the *Meaning-to-Signal Table* in the *Communication Repertoire*.

After the conversion, the output of the *Function Converter* consists solely of BML elements. This output will be transferred to the *BML Planner*.

Mapping Example

An example of converting a communicative act described in FML to BML is illustrated in figure 6.5. Figure 6.6 shows part of a *Meaning-to-Signal Table* in a simplified form that was used for this conversion. Timing and synchronization constraints have been omitted in this example. Note that this is just one possible outcome by the *Function Converter*. Another time the output might be different because of the different probabilities in the mapping table.

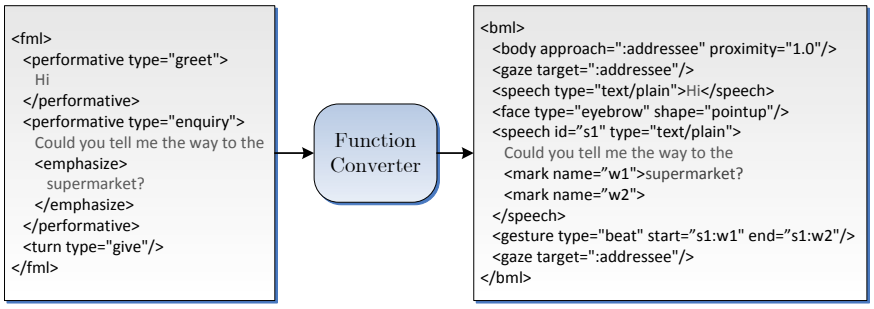


Fig. 6.5: Example of converting FML to BML.

Meaning	Signal Option	Probability	Dependency
greet	approach gaze orient head up	1.0 1.0 0.5	false false false
enquiry	raise inner eyebrow head tilt	1.0 0.6	false false
emphasize	beat - raise eyebrow	0.7 0.3 1.0	true true false
give turn	gaze	1.0	false
...	...	...	...

Fig. 6.6: Meaning Mapping Table Example.

6.2.4 BML Planner

The *BML Planner* takes care of synchronizing all behavior using the synchronization and timing constraints. At this point, any BML that refers to a behavior in the *Gesticon* will be retrieved and integrated. When absolute timing processing is finished, the *BML Planner* sends the BML to the *Behavior Realizer*, which executes the BML.

6.3 Communicative Functions Generator

There are possibilities to automate the process of adding communicative functions to dialog texts for ECAs. Single words or phrases may directly relate to expressions of communicative functions. These relationships could be implemented by annotating text with function tags by using rules that can be applied to words and phrases. A few examples of these rules have been implemented in a system for nonverbal behavior generation for ECAs [50].

Figure 6.7 shows some communicative functions described in FML that could be associated with certain vocal markers. These associations could be used when implementing a *Communicative Functions Generator*. The figure looks a lot like figure 3.4 but the focus here is on showing relationships between words and FML. Therefore this excludes functions that are performed nonverbally.

Some vocal markers in the figure correspond directly to FML tags, like *certainty*, *adjectival* or *meta-cognitive* functions. Other vocal markers like *belief-relation* and *performative* functions correspond to FML tags but a more extensive knowledge of the syntax structure is required here. For example using the *cause-result* relation function, FML has to know exactly what part of the sentence is the *cause* and what part is the *result*. Last, some vocal markers

Class	Function	FML	Vocal Markers
Belief	Certainty	<certainty amount="1.0">	of course (not), definitely (not), yes, no
		<certainty amount="0.0">	maybe, uncertain, perhaps
	Adjectival	<adjectival type="diminish">	small, tiny, subtle, difficult
		<adjectival type="intensify">	big, long, large, great, really, very, quite, wonderful, huge, fantastic,
		<adjectival type="inclusivity">	all, whole, everything, full
Intension	Belief-Relation	<relation type="contrast">	but, on the other hand, however
		<relation type="cause">	because..., ...as a result
		<relation type="result">	...because, as a result...
		<relation type="listing">	...and...
	Performative	<performative type="inform">	let me tell you...
		<performative type="request">	could you…?
		<performative type="enquiry"">	...?
		<performative type="assumption">	i suppose..., i guess..., i assume...
	Deictic	<reference>	this, that, there
Meta-Cognitive		<cognitive type="think">	i'm thinking, let me think
		<cognitive type="deciding">	well
		<cognitive type="word search">	um, uh, well

Fig. 6.7: Associations between FML and vocal markers.

correspond to *deictic* functions who need knowledge about objects or events to which they refer in order to be performed.

7. CULTURE FRAMEWORK

At this point two models have been designed. On one side there is the *Information Model* described in chapter 5 which can be used to categorize and model cultural information about behavior in communication. On the other side there is the *Extended Behavior Generation Framework* for ECAs described in chapter 6.

The *Culture Framework* described in this chapter tries to create a bridge between the two models which will make it possible for an ECA to adopt the cultural information which has been modeled using the *Information Model* and to store it to his own communication repertoire. With the behaviors stored in his communication repertoire, the ECA can become a more believable agent who can behave according to his culture, age or gender. The more behavior and style rules designed specifically for his identity factors, the more believable the ECA can become.

This framework can be used to generate verbal and nonverbal behaviors from communicative functions an ECA wants to perform. The range of nonverbal behavior stored in the ECA's *Communication Repertoire* can be adapted to his identity (culture, gender, age). The framework can also be used to specify the contents of an ECA's *Communication Repertoire* using an authoring tool which enables the modeling of information conforming to the *Information Model*. Defining the behavior content of an ECA is performed during an ECA's design phase.

In order to achieve these requirements, three layers have been designed which combined together become the *Culture Framework*. A simplified overview of the framework can be seen in figure 7.1.

The *Information Processing Layer* includes the authoring tool which can be used to define any behavior or style generation rules. Each rule can specify some identity properties an ECA must have in order to use the rule. The authoring tool stores the generation rules into corresponding tables in the *Behavior Storage Layer*. This layer stores behavior information that can be used by ECAs having different identities. The main goal of this layer is to provide an ECA with

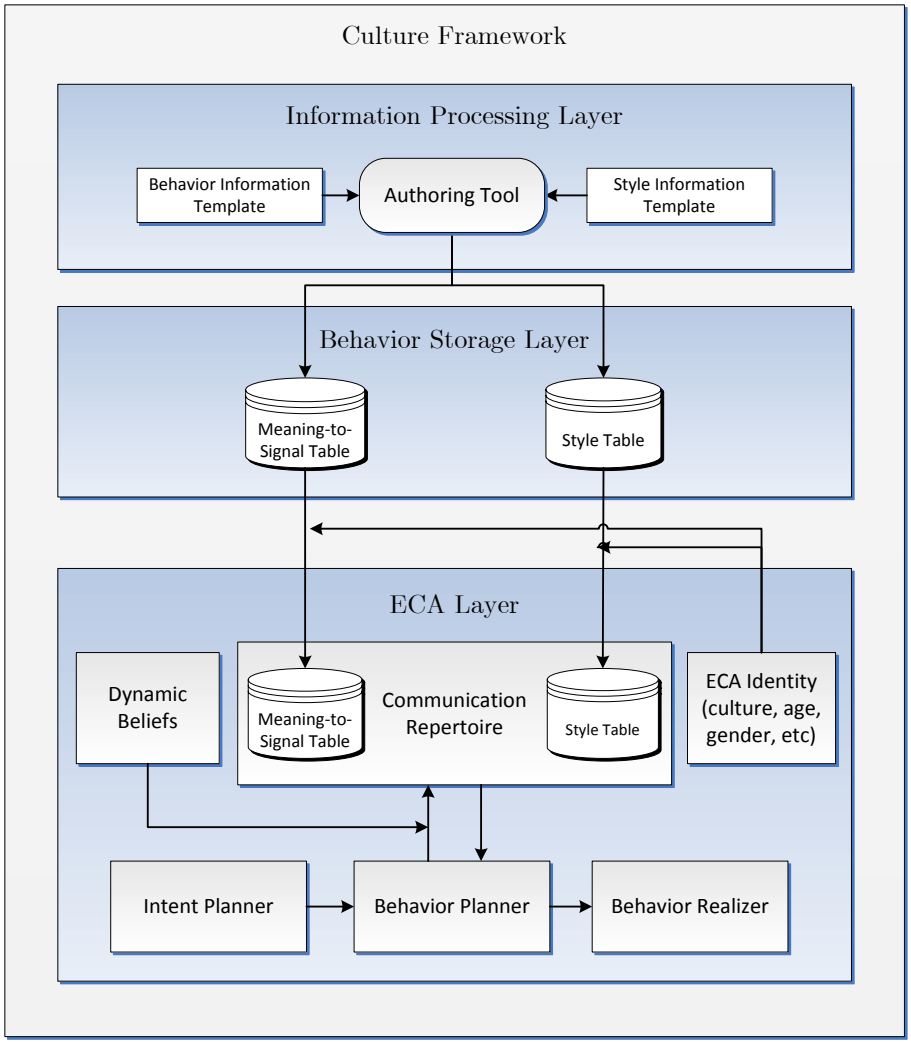


Fig. 7.1: Overview of the Culture Framework.

behavior and style generation rules in the design phase of the ECA. An ECA can retrieve these rules by providing the storage layer with an *identity*. The storage layer retrieves all rules belonging to that identity and transfers them to the *Communication Repertoire* in the *ECA Layer* into the corresponding tables.

Starting at the top layer, the *Information Processing Layer* is described in section 7.1. The *Behavior Storage Layer* is described in section 7.2. It acts as an intermediate layer which on one side can provide a storage facility for all behaviors that can be modeled in the *Information Processing Layer* and on the other side can provide multiple ECAs with behaviors that conform to their specified identity. The *ECA Layer* is described in section 7.3. The *ECA Layer* is an extension of the *Extended Behavior Generation Framework* and includes the support for an ECA of have a communication repertoire which is adapted to his identity and the ability to choose behaviors that conform to his beliefs about the environment and the people around him.

An important part of the *Culture Framework* is the ability to use *styles* as defined in the *Information Model* in section 5.1.2. One interesting feature is the range of information modeling that could be achieved when styles could be applied to different types of behaviors, like single movements, modalities, behaviors with meaning or other groups of behaviors that could be categorized on some aspect. Another feature that should be explored is the different types of behavior that *could* be tuned using styles and how these styles might be implemented. These topics are discussed in section 7.4.

7.1 Information Processing Layer

This section handles the layer of the *Culture Framework* which provides a way for designers to create behavior and style rules for ECAs. This is the *Information Processing Layer* which is illustrated in figure 7.2. This layer is responsible for processing any data that can be modeled using the *Information Model* from chapter 5 and storing it in a format that can be recognized by the *Behavior Storage Layer*. This way, the data and therefore the behavior and style rules that can be modeled can become available for ECAs supporting the *Culture Framework*. This layer consists of two components, namely the *Information Model Templates* described in section 7.1.1 and an *Authoring Tool* described in section 7.1.2.

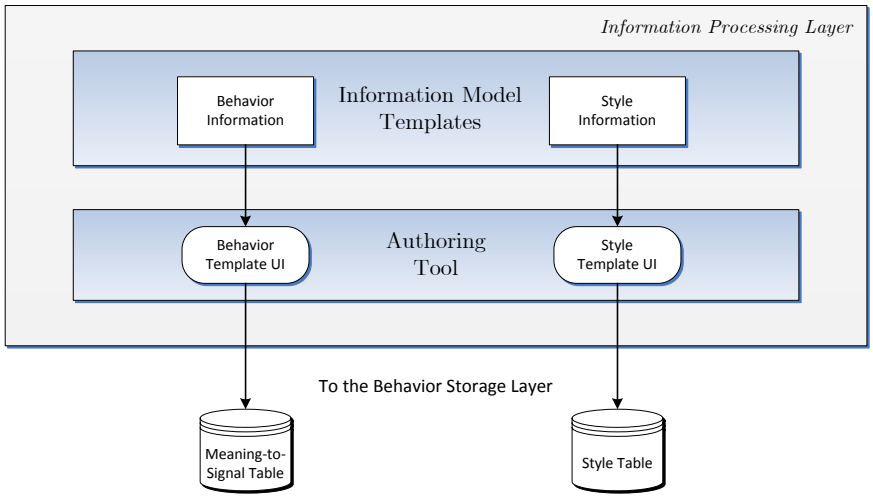


Fig. 7.2: The Information Processing Layer of the Culture Framework.

7.1.1 Information Model Templates

Two templates can be seen in the layer, namely the *Behavior Information* template and the *Style Information* template. These templates together provide the ability to model a huge range of behavioral communication information. The templates and the type of information that can be modeled using the templates is described in detail in chapter 5.

7.1.2 Authoring Tool

An *Authoring Tool* can be used to process the input of the *Behavior* and *Style Information Templates*. To achieve this the tool must have two types of user interfaces that a user can use to insert the desired information. The requirements and usage of the two user interfaces is described below.

Behavior Template User Interface

This user interface can be used to store *Behavior Information*. When saved, the information can be stored in a *Meaning-to-Signal Table* in the form of an xml file.

The process of creating a behavior rule consists of the following actions a user has to perform:

- Provide a description of the behavior rule.

- Provide the ECA identity preconditons.
- Provide the communicative function associated with this rule.
- Provide any preconditions involving the ECA's contextual influences.
- Provide the behavior choices to choose from with their probabilities and provide the multimodal behavior of each choice.

In figure 7.3 an example graphical user interface is illustrated that could be used to design behavior rules for ECAs. In the example, a behavior rule for an ECA representing an Italian adult is specified. FML and BML are used for meaning and behavior representation languages respectively. Note that the specification of multiple options with assigned probabilities has been provided. This is done to comply with the *Meaning-to-Signal Table* as described in chapter 6. The reason to use options is described in section 6.2.2.

Style Template User Interface

This user interface can be used to store *Style Information*. When saved, the information can be stored in a *Style Table* in the form of an xml file.

The process of creating a style rule consists of the following actions a user has to perform:

- Provide a description of the style rule.
- Provide the ECA identity preconditons.
- Provide an optional communicative function associated with this rule.
- Provide any preconditions involving the ECA's contextual influences.
- Provide the style specification consisting of style dimension values.
- Provide the behaviors to which this style should be applied.

In figure 7.4 an example graphical user interface is illustrated that could be used to design style rules for ECAs. In the example, a style rule for an ECA representing an Arabic person is specified. FML and BML are used for meaning and behavior representation languages respectively.

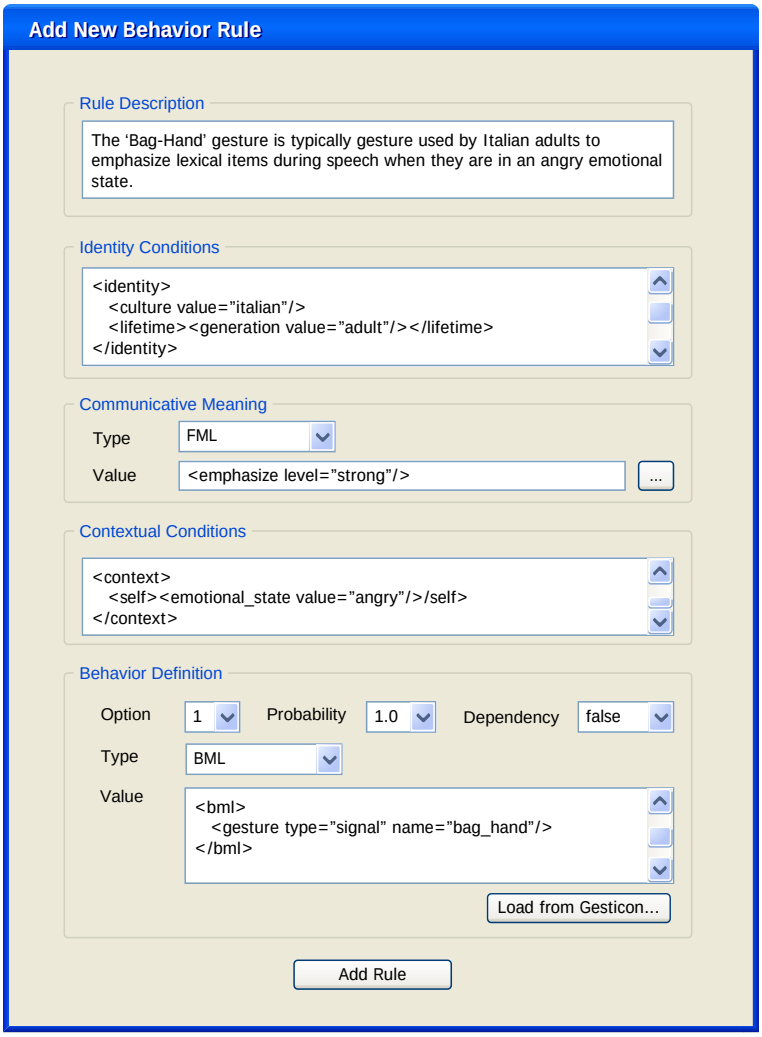


Fig. 7.3: A User Interface Example for Storing Behavior Information.

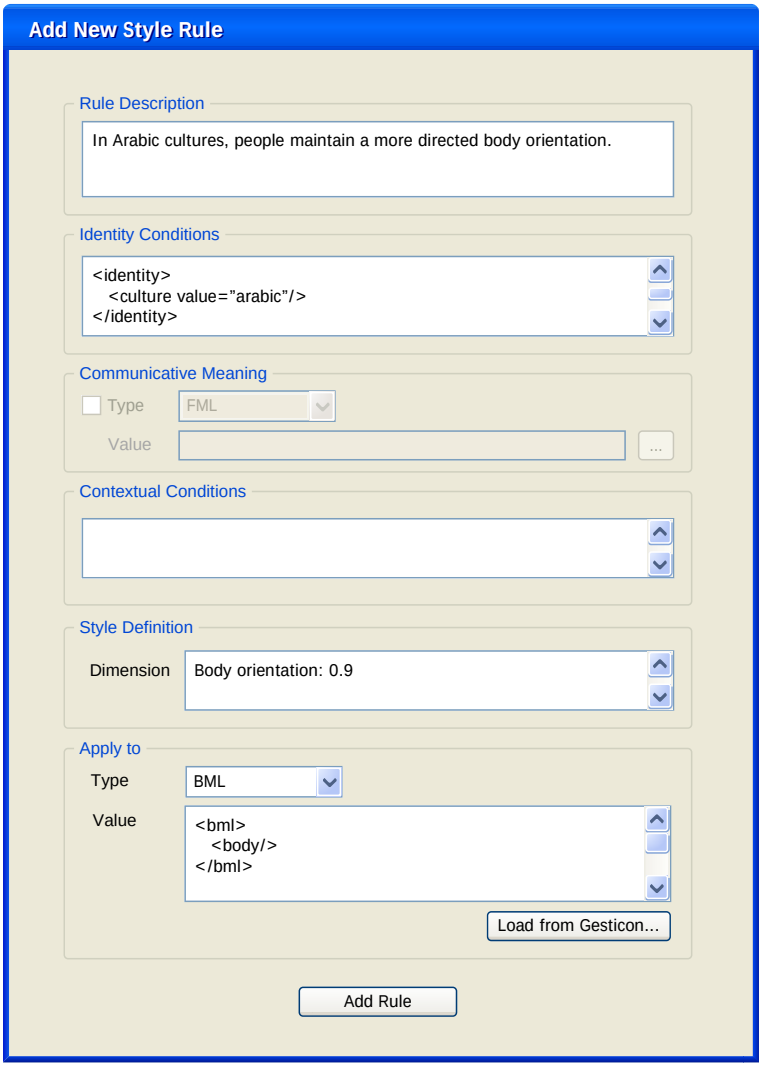


Fig. 7.4: A User Interface Example for Storing Style Information.

7.2 Behavior Storage Layer

The *Behavior Storage Layer* of the *Culture Framework* provides a storage facility where behavior and style rules can be stored for ECAs with a specific identity (culture, age, gender). This layer is an intermediate layer between the *Information Processing Layer* and the *ECA Layer* and can provide multiple ECAs with behavior and style rules. The main advantage of this layer is the ability to store communicative behaviors tied to identity factors which ECAs can use. These rules can be reused by other ECAs that have the same values for the identity factors. For example, if many rules have been stored in this layer designed specifically for an ECA belonging to some specified culture, then other ECAs of the same culture can also use these rules. Other rules could be designed specifically for ECA's belonging to a certain age-span, a gender, a country cluster or a combination of two or more factors. The specification of these *identity* factors depends on the taxonomy that is used, which is usually domain dependent.

The internal design of the *Behavior Storage Layer* is illustrated in figure 7.5. At the top there is the *Behavior Database*. The *Behavior Database* provides a storage facility for *Behavior Information* and *Style Information* which can be modeled using the *Information Processing Layer* as described in the previous section. At the bottom is the *Behavior Provider* which is a front-end to the *Behavior Database* and can be used to query the database and retrieve any desired data. This provider can be requested to retrieve all behavioral data that conforms to some given *ECA Identity*. This data that can be retrieved is compatible with the contents of an ECA's *Communication Repertoire*, namely in the form of a *Meaning-to-Signal* and *Style* table.

The *Behavior Database* is described in section 7.2.1 and the *Behavior Provider* is described in section 7.2.2.

7.2.1 Behavior Database

The *Behavior Database* component provides storage for two types of tables, namely the *Meaning-to-Signal Table* and the *Style Table*. These tables are the same as the corresponding tables in an ECA's *Communication Repertoire* except that each table entry here has some extra preconditions that must be met in order to retrieve the rule from the database. These preconditions involve the *ECA Identity* factors as described in section 7.3.1. This means that these entries are specifically intended to be used by ECAs conforming to some identity. For example, specific rules can be stored here for ECAs having the identity of a

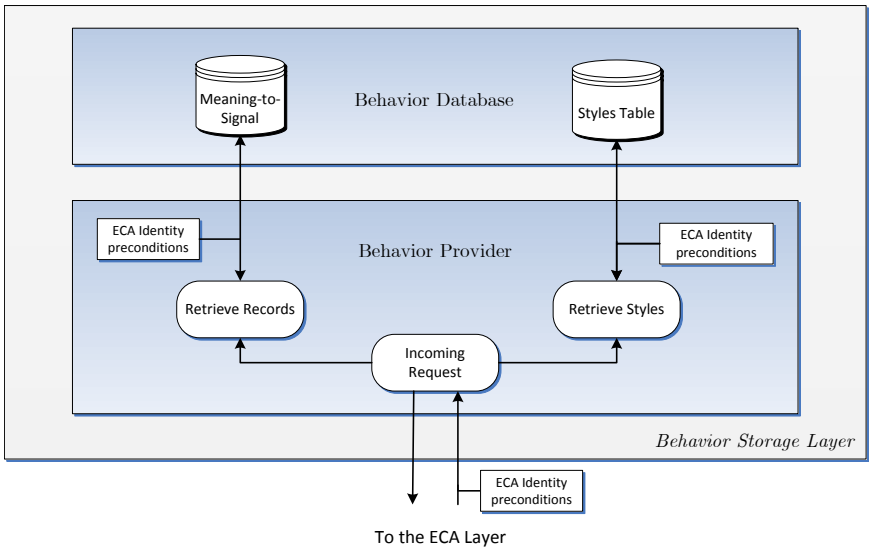


Fig. 7.5: The Behavior Storage Layer of the Culture Framework.

male, an adult male, or an adult male from a Western country, depending on the number of preconditions that have been specified. The less preconditions there are specified, the larger the range of ECAs identities that can use the rule.

Part of an example content of the *Behavior Database* can be seen in figure 7.6. The preconditions determining the ECA's intrinsic influences (section 3.3.1) have been called *Identity Conditions* whereas the preconditions concerning the ECA's contextual influences (section 3.3.2) have been called *Context Preconditions*.

A few culture and gender differences have been modeled in the *Meaning-to-Signal Table*. Cultural differences involve the behavior expression with the intention of showing agreement and disagreement which can have opposite meanings in some cultures. Further the difference of an emblem gesture usage in America and Iraq has been modeled. At last a few different greeting forms between men and women in Arabic cultures have been modeled.

Some more cultural differences have been modeled in the *Style Table*. These concern proxemic behavior differences between America and Iraq, the expressiveness of facial expressions in Japan, the rate of speech in Israel and the use of gestures in Italy. Besides that, two other styles have been defined. One states that men use more power in their use of gestures than women and the other states that elderly persons are a bit slower in all nonverbal behaviors.

Identity Preconditions	Meaning	Context	Signal Option
	express agreement		nod head
	express disagreement		shake head
culture/county=”bulgaria”	express agreement		shake head
culture/county=”bulgaria”	express disagreement		nod head
culture/county=”america”	nonverbally insult someone		extent middle finger
culture/county=”iraq”	nonverbally insult someone		thumbs up
gender=”male”	greet	other/gender/male	handshake
gender=”male”	greet	other/gender/female	kiss on cheek
gender=”male” culture/cluster=”arab”	greet	other/gender/male	kiss on cheek
gender=”male” culture/cluster=”arab”	greet	other/gender/female	handshake

Meaning-to-Signal Table

Identity Preconditions	Meaning	Context	Style Dimensions	Apply to
			proxemics: 0.5	approach
culture/country=”america”			proxemics: 0.9	approach
culture/country=”iraq”			proxemics: 0.2	approach
culture/country=”japan”			spatial extent: 0.1	facial expressions
culture/country=”israel”			rate: 0.7	speech
culture/country=”italy”			activation: 0.8	gestures
gender=”male”			power: 0.6	gestures
lifetime/generation=”elder”			temporal: 0.7	gestures

Style Table

Fig. 7.6: Example content of the Behavior Database.

7.2.2 Behavior Provider

The *Behavior Provider* provides an interface between the *Behavior Database* and the *ECA Layer* of any ECA that requires behavior information from the database. The *ECA Layer* can request this behavior with the goal of filling his own *Communication Repertoire* with behavior adapted to his identity. To achieve this, the *ECA Layer* must provide the *Behavior Provider* with an *identity* specification. The *Behavior Provider* in turn searches the two tables inside the *Behavior Database* to retrieve all entries who's identity preconditions are met by the ECA's *identity* specification. The retrieved entries stripped of their identity preconditions are transferred to the corresponding tables in the *Communication Repertoire* of the *ECA Layer*. The ECA now contains behavior and style rules adapted to his identity.

To give an example, take for example an ECA who is required to adapt his behavior to a 70 year old Iraqi male. Assume that the *Behavior Database* is filled with the content illustrated in figure 7.6. The ECA will consult the *Behavior Provider* to see if there are rules in the *Behavior Database* that comply with his identity. The rules the *Behavior Provider* will retrieve from the database and transfer to the ECA are the yellow highlighted entries in figure 7.7. Note that the entries involving greetings expressed by males are not highlighted even though the ECA in question has a male identity. This is because there are more detailed greeting rules available which concern Arabic males. These entries have a higher priority because more conditions have been met.

7.3 ECA Layer

In this section the *ECA Layer* of the *Culture Framework* will be discussed. The design of this layer is illustrated in figure 7.8. The bottom three components, namely the *Intent Planner*, the *Behavior Planner* and the *Behavior Realizer*, work exactly the same as described in the *Extended Behavior Generation Framework*. The difference here concerns the content and retrieval of the behavior contained in the communication repertoire. The components involved in this process, namely the *ECA Identity*, the *Dynamic Beliefs*, the *Communication Repertoire* and the *Function Converter* will be discussed in section 7.3.1-7.3.4.

7.3.1 ECA Identity

In this framework it is assumed that the behavior of an ECA that will be generated is adapted to his identity (culture, gender). This means that the

Identity Preconditions	Meaning	Context	Signal Option
	express agreement		nod head
	express disagreement		shake head
culture/county="bulgaria"	express agreement		shake head
culture/county="bulgaria"	express disagreement		nod head
culture/county="america"	nonverbally insult someone		extent middle finger
culture/county="iraq"	nonverbally insult someone		thumbs up
gender="male"	greet	other/gender/male	handshake
gender="male"	greet	other/gender/female	kiss on cheek
gender="male" culture/cluster="arab"	greet	other/gender/male	kiss on cheek
gender="male" culture/cluster="arab"	greet	other/gender/female	handshake

Meaning-to-Signal Table

Identity Preconditions	Meaning	Context	Style Dimensions	Apply to
			proxemics: 0.5	approach
culture/country="america"			proxemics: 0.9	approach
culture/country="iraq"			proxemics: 0.2	approach
culture/country="japan"			spatial extent: 0.1	facial expressions
culture/country="israel"			rate: 0.7	speech
culture/country="italy"			activation: 0.8	gestures
gender="male"			power: 0.6	gestures
lifetime/generation="elder"			temporal: 0.7	gestures

Style Table

Fig. 7.7: Content of the Behavior Database with rules highlighted that comply to the identity of a 70 year old Iraq male.

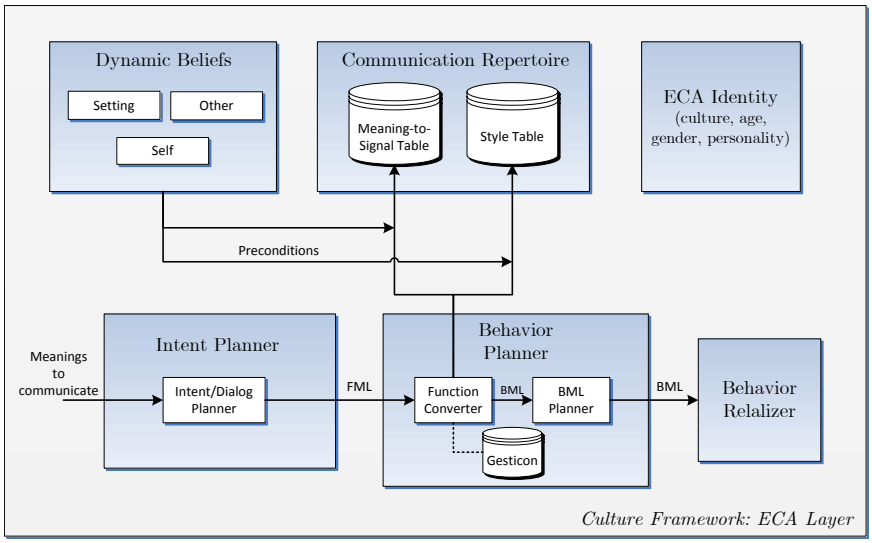


Fig. 7.8: The ECA Layer of the Culture Framework.

content of the *Communication Repertoire* consists of adapted behavior that conforms to the ECA's identity as desribed in the previous section.

When designing an adapted ECA it is required to know some identity factors he needs to represent, like culture, age, gender, personality, etc. When these are known, behavior and style rules can be retrieved from the *Behavior Storage Layer* which conform to his specified identity. These identity values can be stored in the *ECA Identity* component. They correspond to the intrinsic influences as described in the *Information Model* in section 5.1.4. Aspects that can be stored here are the agent's *personality*, *age*, *gender* or *culture*. Figure 7.9 illustrates two examples of different ECA identities that could be specified. The taxonomies used here are examples. Normally, the taxonomies to use are highly domain dependent. For example, in one domain it might be required to design a lot of behavior rules for ECAs having different personalities. This requires a taxonomy in which more personality dimensions can be specified. Other domains require focussing on other properties.

7.3.2 Dynamic Beliefs

Besides the agent's identity, also beliefs the agent has about himself, others around him and the environment should have an influence on the choice of behavior. For example in reality, people behave differently when communicating

```
<identity>
  <culture>
    <cluster value=”Arab”/>
    <country value=”Iraq”/>
  </culture>
  <gender value=”male”/>
  <lifetime>
    <generation value=”adult”/>
    <age value=”45"/>
  </lifetime>
  <personality>
    <extroversion value=”0.2"/>
  </personality>
</identity>
```

```
<identity>
  <culture>
    <cluster value=”Far East”/>
    <country value=”Indonesia”/>
  </culture>
  <gender value=”female”/>
  <lifetime>
    <generation value=”child”/>
    <age value=”9"/>
  </lifetime>
  <personality>
    <extroversion value=”0.8"/>
  </personality>
</identity>
```

Fig. 7.9: Example of ECA Identity Specifications.

with a good friend in contrast to communicating with a complete stranger. For ECAs to become more believable, they too should be able to express different behaviors in different situations. In addition, mental and physical resources can influence behavior as well. For example, there can be a big behavioral difference between people who are very energetic and people who are physically and mentally very tired. Again, ECAs should be able to reflect these situations in order to improve their believability.

The factors described here correspond to the contextual influences as described in the *Information Model* in section 5.1.4. These factors are dynamic and can be different between different communication encounters, but can also be changed during the evolution of a communication. For example, an agent might become angry because of something his interlocutor said to him. This change should be reflected in his future behavior until some other change might occur.

Because these factors are dynamic, they can have different values during a communication encounter. This means that if the agent needs to be able to behave differently for different values of these dynamic beliefs, he needs to have behavior rules that were designed for these different situations. Therefore, retrieval of behavior from the *Communication Repertoire* can be dependent on the values of these different dynamic factors.

The values of these factors are stored in the *Dynamic Beliefs* component. This component has been divided into beliefs the agent has about himself (emotional, physical, mental state), about the participants of the communication (personality, culture, gender, age) and about the setting (physical, social, type of encounter, presence of others, relationships). Behavior decision can be influenced by one or a combination of these factors. Returning to the design of the *ECA Layer*, when the agent consults the *Communication Repertoire* to find the correct behavior given a meaning, the agent uses his *Dynamic Beliefs* to further filter the behaviors to choose from and eventually retrieves the behavior which is in accordance with his intentions and *Dynamic Beliefs*. Figure 7.10 shows an example of a possible configuration of the dynamic factors. Again, the size and choice of the taxonomy to use is domain dependent.

7.3.3 Communication Repertoire

At last there is the *Communication Repertoire* from which the actual behaviors an ECA expresses can be fetched. This repertoire is an extension of the communication repertoire from the *Extended Behavior Generation Framework* and contains two tables which are outlined below.

```
<context>
  <self>
    <physical_state value="energetic"/>
    <mental_state value="concentrated"/>
  </self>
  <other>
    <gender value="male"/>
    <generation value="adult"/>
  </other>
  <setting>
    <physical_setting value="outdoor"/>
    <relationship value="stranger"/>
  </setting>
</context>
```

Fig. 7.10: Example of ECA Beliefs Specification.

Meaning-to-Signal Table

The same *Meaning-to-Signal* mapping table is available here. However in this table, every rule can contain extra preconditions specifying the contextual influences which must be met in order to apply the rule and retrieve the specified multimodal behavior. These preconditions correspond to the factors inside the *Dynamic Beliefs* component. For example, there may be rules specified which should only be applied if the addressee is of a specific gender. In many cultures, men greet other men differently than they greet women and vice versa. Using preconditions, different greeting rules could now be specified if desired.

It is not required for rules to have preconditions. Rules without preconditions apply to the whole range of dynamic beliefs the agent may have, i.e. it doesn't matter how the agent feels, to whom he's talking and in which setting the conversation is situated. Rules with preconditions that are met have priority over other rules mapping the same meaning but which contain no preconditions. However, all preconditions a rule includes must be met in order to retrieve the corresponding behavior. To see if the preconditions of a rule are met, the values of the *Dynamic Beliefs* component are compared to the preconditions. An example of part of a *Meaning-to-Signal Table* handling the communicative function of greeting is illustrated in figure 7.11

This example table contains five different behavior rules for greeting. The ECA using this table has knowledge of how to greet people from Japan and America. He also has norms about how to greet female friends and male friends, regardless of their cultural roots. In all other cases, he has a standard way of greeting people by shaking their hands. Every greeting gesture is preceded

Meaning	Context	Signal Option
greet		approach 1.0p
		handshake 1.0p
greet	other/culture/country="japan"	approach 1.0p
		bow 1.0p
greet	other/culture/country="america"	approach 1.0p
		handshake 1.0p
greet	other/generation="adult" other/gender="female" setting/relationship="friend"	approach 1.0p
		kiss on cheek 1.0p
		hands on shoulder 1.0p
greet	other/generation="adult" other/gender="male" setting/relationship="friend"	approach 1.0p
		warm handshake 1.0p
		pat on shoulder 1.0p

Fig. 7.11: Part of a Meaning-to-Signal Table for Greeting.

with an approach which makes sure the body is in the desired proximity of the interlocutor.

However, an issue may arise when the ECA needs to greet a Japanese female friend. The mapping table has two different rules for this case, one rule for greeting Japanese people and one rule for greeting female friends. There isn't a special rule for greeting Japanese female friends. These kind of issues may be resolved by prioritizing the contextual factors depending on the domain for which the table is used. For example, in one domain the *cultural* factor might be the most important factor, while in other domains the *relationship* factor is more important. Alternatively, the ECA can fall back to its default greeting behavior with no preconditions. The choice being made is thus domain dependent. So the main advantage of using preconditions is the *possibility* of specifying detailed behavior in different situations.

Style Table

The second extension this component introduces is a *Style Table*. In this table *style rules* can be defined. These correspond to the styles described in section 5.1.2 discussed as part of the *Information Model*. To recapitulate, styles can be applied to a (multimodal) behavior, to a modality or to behaviors with a specific meaning.

An entry in the *Style Table*, i.e. a style rule, consists of one or more style dimensions with dimension values, an optional meaning, optional preconditions and some behavior definition to which the style must be applied. All these

factors are described in detail in chapter 5.

During the behavior rules retrieval of an ECA, zero or more style rules can be applied. This is dependent on the communicative function an ECA wants to express and the values of the contextual preconditions. Every time the *Behavior Planner* consults the *Communication Repertoire* to retrieve a certain behavior, the *Meaning-to-Signal Table* will be searched to find the correct entry. At the same time, the *Style Table* will be searched to retrieve all the styles that should be applied at that moment. How the behavior and styles that have been retrieved are combined to form the final behavior the ECA will express is described in the next section.

Figure 7.12 illustrates part of an example *Style Table*. Five styles have been defined, each for different situations. The first two styles define proxemic behavior of the ECA and states that when approaching a friend the ECA will stand closer than when approaching a stranger. The next style states that when the ECA is feeling tired, he will talk a bit slower and softer. The fourth style defines a general property of the ECA, namely that he isn't very expressive in his facial expressions. The last style defines a specific style for a specific situation, namely that when the ECA greets a Japanese person of a high status *and* when he does this by bowing, *then* he will express this behavior longer and with more spatial extent, resulting in a longer and deeper bow. When the meaning condition (greeting) in this style rule was not specified, then this style would be applied to every expressed bow, also if it resulted from other communicative functions other then greetings.

Meaning	Context	Style Dimensions	Apply to Signals
-	setting/relationship="stranger"	proxemics: 0.7	approach
-	setting/relationship="friend"	proxemics: 0.4	approach
-	self/physical_state="tired"	vocal rate: 0.4 vocal volume: 0.4	speech
-	-	spatial extent: 0.1	facial expressions
greet	other/culture="japan" setting/relationship="high_status"	temporal: 0.6 spatial extent: 0.6	bow

Fig. 7.12: Part of a Style Table.

7.3.4 Function Converter

As described in the *Extended Behavior Generation Framework*, the function of the *Function Converter* is to convert a communicative act described in some meaning representation language like FML to a series of communicative behaviors described in some behavior representation language like BML.

The added functionality in the *Culture Framework* is the retrieval of *style rules* from a *Style Table* as described in the previous section. These *style rules* have to be applied to behaviors retrieved from the *Meaning-to-Signal Table*. The processes inside the *Function Converter* are illustrated in figure 7.13.

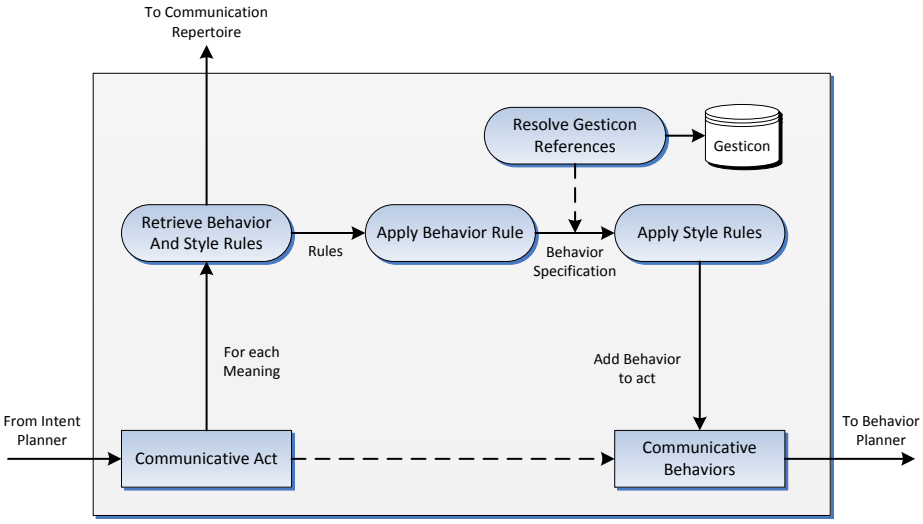


Fig. 7.13: Internal Design of the Function Converter.

To begin with, the *Function Converter* receives a communicative act from the *Intent Planner*. This act contains words augmented with communicative functions and has to be converted to communicative behaviors. To achieve this, the *Function Converter* will convert each communicative meaning inside the act to a communicative behavior using behavior and style rules that are retrieved from the *Communication Repertoire*.

First the retrieved behavior rule will be applied. This process involves choosing a (multimodal) behavior from a choice of behaviors using assigned probabilities. After a (multimodal) behavior has been chosen, any *Gesticon* reference will be resolved if necessary. At this point the behavior specification is complete. Now any style rule that needs to be applied to each behavior is processed. The

result is a multimodal behavior in which some behavior expressions are augmented with style information. The behavior expressions are added to the list of communicative behaviors already generated. When all meanings have been converted, the list of behaviors is send to the *Behavior Planner* who will handle any synchronization and timing constraints.

The conversion of a communicative act to communicative behavior will be made clear using an example. Take for example the scenario of two Arab male friends meeting each other on the street. One of them starts a communicative act of greeting the other one and asking him how his wife is doing. The scenario with the communicative act is illustrated in figure 7.14.

```
<fml>
  <performative type="greet">
    Hi
  </performative>
  <performative type="enquiry">
    How is your
    <emphasize>
      wife
    </emphasize>
    doing?
  </performative>
  <turn type="give"/>
</fml>
```

Fig. 7.14: Example scenario of a communication between two Arab males.

The ECA's *Identity*, *Dynamic Beliefs*, *Meaning-to-Signal Table* and the *Style Table* of the ECA performing the communicative act are shown in figure 7.15. The identity and beliefs specification states that the ECA is an 45 year old Iraqi male which beliefs his interlocutor is also an adult Iraqi male who he considers a friend. The way this ECA will express his act will depend on the contents of the behavior and style tables. The yellow highlighted entries in the tables are the rules that need to be retrieved to convert this communicative act to behavior. The preconditions of the *greeting* rule have been met using the *Dynamic Beliefs* specification, namely that the ECA is interacting with an Arabic male friend.

In this example, a few Arabic cultural communicative behaviors have been modeled. In the *Meaning-to-Signal* table one Arabic behavior has been modeled, namely that it's not uncommon for Arabic male adult friends to kiss each other on the cheek when they greet. In the *Style* table other factors have been modeled. People from Arabic cultures, for example, gaze long and direct at their partners. They also have a very direct body orientation, little interper-

```
<identity>
  <culture>
    <cluster value="arab"/>
    <country value="iraq"/>
  </culture>
  <gender value="male"/>
  <lifetime>
    <generation value="adult"/>
    <age value="45"/>
  </lifetime>
 </identity>
```

ECA Identity

```
<context>
  <other>
    <gender value="male"/>
    <generation value="adult"/>
    <culture>
      <cluster value="arab"/>
      <country value="iraqi"/>
    </culture>
  </other>
  <setting>
    <relationship value="friend"/>
  </setting>
</context>
```

ECA Dynamic Beliefs

Meaning	Context	Signal Option	Probability	Dependency
greet		handshake	1.0	false
greet	other/gender="male" other/culture/cluster="arab" setting/relationship="friend"	approach	1.0	false
		gaze	1.0	false
		kiss cheek	1.0	false
greet	other/gender="female" other/culture/cluster="arab" setting/relationship="friend"	approach	1.0	false
		gaze	1.0	false
		handshake	1.0	false
enquiry		raise inner eyebrow	1.0	false
		head tilt	0.6	false
emphasize		beat	0.7	true
		<empty>	0.3	true
		raise eyebeow	1.0	false
give turn		gaze	1.0	false
...		...	...	...

Meaning-to-Signal Table

Meaning	Context	Style Dimensions	Apply to Signals
		orientation: 0.9	body, head
		spatial extent: 0.8	facial expressions
		spatial extent: 0.6	hand gestures
		temporal: 0.8	gaze
		vocal volume: 0.7	speech
	setting/relationship="stranger"	proxemics: 0.4	approach
	setting/relationship="friend"	proxemics: 0.1	approach

Style Table

Fig. 7.15: ECA Identity, Dynamic Beliefs and contents of the Communication Repertoire.

sonal space and tend to speak with a relative loud voice [24, 79]. Furthermore styles have been specified to increase the expressiveness of facial expressions and the spatial extent of hand gestures.

First the *behavior rules* are applied and the communicative meaning is converted to communicative behavior, in this case a conversion from FML to BML. Next the *style rules* are applied to each single behavior element. The result is shown in figure 7.16.

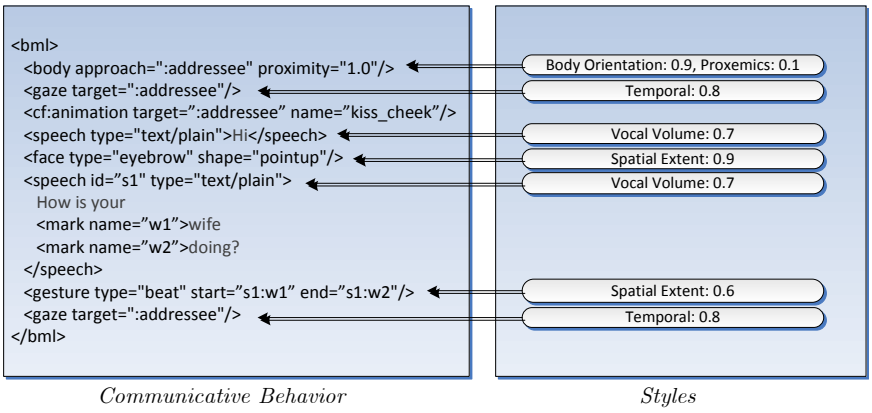


Fig. 7.16: Mapping Result of the Communicative Act.

The issue that remains is how styles can be implemented and how can they be applied to the different behavior elements. This can depend on the modality in question. A single style can be applied differently to each modality, as described in section 5.1.2. Also, it can depend on the behavior representation language that is used. Each language could have its own mechanism for adding styles to behavior. It might also be dependent on the domain in which the language is used. A discussion of implementing and applying styles is given in the section below.

7.4 Styles

The use of styles is a very important feature of the *Culture Framework* providing a huge range of information modeling if used correctly. This huge range of modeling is achieved on one hand by the ability to define styles using only dimensions and on the other hand by the ability to apply styles to different groups of behaviors. The possibilities for using and applying styles requires some more discussion.

The issues involved in the implementation of different styles are outlined in section 7.4.1 and the issues involved in applying styles to different behaviors is described in section 7.4.2.

7.4.1 Style Implementation

Most styles that have been discussed in this document work directly on the expressed behavior, like the *expressivity dimensions* or *paralinguistic features*. The implementation of these styles are fully dependent on the behavior representation language that is used.

Another style that might be interesting to use but which does not work directly on the expressed behavior is a style handling the frequency of certain types of behavior. This style might be represented by a value on a *frequency dimension* which determines the chance of occurrence. The implementation of this style is discussed below followed by the implementation of a few other styles in BML.

Style for Frequency Regulation

As said above, a style could be defined that specifies the *frequency of occurence* of certain types of behavior. This style would be very interesting when modeling some cultural differences. Take for example the frequency of gestures. In Italy people use relatively many gestures while this may be seen as a lack of intelligence in Britain. Therefore, it would be desirable to be able to regulate the frequency of gestures expressed by ECAs.

One important question that arises is if this observed frequency of some behavior is caused by the generation of nonverbal behavior or that it is caused by the generation of communicative functions *resulting* in nonverbal behavior. Take for example the lack of gesture usage. Is there a lack of gestures because there isn't any communicative function planned that could result in a gesture or is the communicative function planned, but is the expression of the gesture suppressed. Because the planning of communicative acts falls outside this research, a solution is proposed which implements the latter option.

One solution for implementing *frequency styles* is to use the mechanism of *options* when designing *behavior rules* as described in section 6.2.2. Options were introduced to provide an ECA with more variety in behavior expressions and not be concerned with too complex behavior modeling. Every option is associated with a probability stating the chance that an option will be chosen.

The *frequency style* might be implemented in the form of influencing these

probabilities and therefore influencing the chance a behavior will be expressed. How the probabilities might be influenced can be dependent on the behaviors to which the frequency style applies.

Style Implementation in BML

At this moment, BML doesn't have any special mechanism for implementing *styles* as defined in this document. A few styles however could be implemented using current BML element attributes.

Proxemic behavior for example could be implemented using the *proximity* attribute of the *body* element specifying the distance of approach towards the interlocutor. *Body orientation* could be implemented using the *angle* attribute of the *body* element specifying the offset angle of final facing towards the interlocutor. These examples however might require some variations in their values during the execution of a communicative act. People don't always orient their body towards the interlocutor at one specific angle. This orientation may vary throughout the conversation, so the orientation could be seen as a sort of average value. To handle these issues a special mechanism has to be introduced.

Styles representing paralinguistic features like rate, volume or pitch could be implemented using SSML [35] in case the BML processor recognizes SSML in *speech* elements.

Other styles could be implemented by designing attributes for BML elements using a custom namespace. This namespace has to be recognized by the BML processor that is used. For example, the *spatial extent* dimension applied to gestures could be implemented by adding a *space* attribute to the *gesture* element. The usages of a few expressivity dimensions is shown in figure 7.17.

```
<bml xmlns:cf=”http://www.eca.edu/cultureframework”>
  <gesture type=”beat” cf:space=”0.7" cf:fluidity=”0.8" cf:power=”0.4”/>
</bml>
```

Fig. 7.17: BML Style implementation using namespaces.

Alternatively, the BML specification has a mechanism available for adding multiple levels of descriptions to BML elements. Description levels can be used to give more detailed descriptions of behavior. Level 0 consists of core BML, where every higher level description is able to give more detailed behavior specifications. Every level should be self-contained and able to be processed on its own. To implement some style dimensions, a level 1 description could be

designed handling some basic style features. The nice thing about using *level descriptions* is that there is no need to adapt the BML core specification by adding more attributes. An example of using level descriptions is illustrated in figure 7.18.

```
<bml>
  <gesture type=”beat”>
    <description level=”1" type=”CultureFramework”>
      <type>beat</type>
      <space>0.7</space>
      <fluidity>0.8</fluidity>
      <power>0.4</power>
    </description>
  </gesture>
</bml>
```

Fig. 7.18: BML Style implementation using description levels.

All these implementation options would have to be carefully considered and evaluated to see if they would achieve the desired goal of implementing the style dimension in question the way it was intended. Eventually, for these styles to work properly, the *Behavior Realizer* should be able to interpret these implementation options.

7.4.2 Applying Styles to Behavior

The way styles can be applied to behavior is totally dependent on the behavior and function representation languages that are used. Styles can only be applied to certain groups of behavior if the representation language used has knowledge of that group. Below, styles that are applied to different groups of behaviors are discussed. These are the same groups as described in the part of applying styles in the *Information Model* in section 5.1.5.

Styles for (Multimodal) Behaviors

Applying styles to specific (multimodal) behaviors should not pose many issues because these behaviors can be specified by any representation language using one or more elements. These elements can be augmented with the desired style.

Styles for Modalities

Applying styles to specific modalities might be more problematic. There are however some possibilities. Looking at BML, some elements directly correspond to modalities, like the *speech* element handling all verbal behavior or the *face* element handling all facial expressions or the *gaze* element handling eye behavior. Therefore, when using BML, it is relatively easy to apply a style to a certain modality if that modality is represented by one element. For example it's possible to apply styles to all facial expressions or to all eye movements by augmenting the *face* and *gaze* element respectively with the style in question.

More issues arise when style rules are designed to limit or state the necessity of the use of a very specific modality. Take for example the modeling of the *Style Information*: *"In India, it is impolite to point at someone or something with your left hand"*. Another specific rule that has been observed is: *The Kiowa Indians use their lips to point at things.*

The possibilities of modeling these interesting cases again depends on the behavior representation language used. If there would exist an element in the language for pointing, these style rules might be implemented by augmenting the element with additional information on how to point, including the use or limitation of specific body parts.

Styles for Behaviors with Meaning

Applying styles to other kinds of behavior groups can be trickier. Take for example behavior groups made up of behaviors that fulfill a certain communicative function. These include for example the classification system of Ekman and Friesen described in section 3.5.8 which classifies body movements into five types: *emblems*, *illustrators*, *affective displays*, *regulators* and *adapters*. Another example is the categorization of haptic behaviors by Jones and Yarbrough into seven groups of meanings: *Positive Affect*, *Playfulness*, *Control*, *Ritual*, *Hybrid*, *Task-Related* and *Accidental Touch*.

Interesting things could be modeled if styles could be applied to such groups, like controlling the style of *regulator gestures* or *control touches*. This would involve styles being applied to behaviors resulting from certain *meaning* groups. One way to apply a style to a specific communicative function that can be described in FML could be to attach the style to every (multimodal) behavior that resulted from the conversion of FML to BML. For example, FML has an *affect* element which includes all emotion expressions. A style could then be defined and applied to all BML elements resulted from the conversion of the

affect element of FML. However, these examples involve very low level communicative functions. Higher level function modeling might become available in future research on FML.

The range of style modeling is thus dependent on the representation languages used. This can make it hard to model interesting notable cultural differences like the *frequency of haptic behavior* for example which could be defined as a style. This would only be possible if the behavior representation language would have the knowledge of which behaviors involve touch and which don't. On the other hand, it is probably more interesting to model frequencies of haptic behavior fulfilling some *communicative meaning*. For example the modeling of the frequency of *positive affect* touches. This kind of information modeling might be possible in the future when there is a more complete model of communicative functions and a corresponding representation language identifying these meanings.

8. EVALUTATION

Now that the *Culture Framework* has been described, an evaluation has to be made to see what the possibilities are for modeling and implementing cultural differences in communicative behavior for ECAs. In this chapter these possibilities are explored for communicative functions in section 8.1 and for communicative behavior in section 8.2. Most attention will be given to behaviors where culture has a large influence on. Besides cultural differences, behaviors that are universally recognized are handled as well. For ECA design, it would be desirable to equip an ECA with a basic set of behaviors that can be used for every ECA, no matter to what cultural identity he has to conform.

8.1 Communicative Functions

8.1.1 Emblems

A very clear difference between cultures in behavior expression is the expression of emblem gestures. The expressions for the same meaning can be very different and one expression can sometimes have very opposite meanings in different cultures. We have seen that emblems are *coded* gestures, meaning they are represented in memory by linking a gesture expression to its meaning. From this follows that these gestures can be coded in the framework using the *Meaning-to-Signal Table*.

8.1.2 Affective Displays

An important difference in affective displays between cultures concerns the expressiveness of the displays. While in some cultures people are very expressive when showing emotion, people from other cultures hardly express any emotion. Note that affective displays don't have to be expressed using only facial expressions. Other modalities like body posture can also contribute in displaying an emotional feeling.

The expressiveness of behavior that results from affective displays could be modeled using a *spatial extent* style dimension where a zero value shows no

expression and a one value shows the maximum possible expression. This style then has to be applied to all behaviors generated after mapping this communicative meaning to signals.

8.1.3 Greetings

We have seen that greeting forms between cultures can vary a lot, involving different modalities and different forms in different situations. Implementing different greeting forms is a typical example of using the *Meaning-to-Signal Table* in the framework to map a greeting function to the desired signal or form depending on the culture.

8.2 Communicative Behavior

8.2.1 Gestures

Notable cultural differences in using gestures are the frequency of gestures and the gesture space that is used to perform a gesture. These properties could be modeled using style dimensions like *frequency* and *spatial extent*. The *frequency* dimension could be applied to all behavior rules involving gestures. The *spatial extent* dimension could be applied to all gestures.

8.2.2 Facial Expressions

Most facial expressions are universally understood, like facial expressions resulting from meta-cognitive functions or facial expressions showing one of the universally recognized emotions. However, the difference between cultures is usually in the expressiveness of the expressions. In some cultures, people might use their facial muscles more than in other cultures. This difference is especially noticed when looking at the expression of affective displays through facial expressions.

Just like the expressiveness of affective displays can the expressiveness of facial expressions be modeled by a *spatial extent* style dimension. However, here only facial expressions are influenced, instead of the whole body for affective displays. This style could be applied to all facial expressions.

8.2.3 Proxemics

Many differences in the use of personal space and distance between cultures have been observed. These differences range from very close distances observed

in many collectivist and contact cultures to large distances observed in individualist and noncontact cultures. These differences could be modeled using a *proxemics* style dimension whose value indicates a culture's norm for keeping distance during interaction. This style could be applied to all behaviors involved in moving the body in space.

8.2.4 Haptics

Differences in haptic behavior among cultures mainly concerns the frequency of behavior where touching is involved. This difference is an important property in the division of cultures into *contact* and *noncontact* cultures. To influence the occurrence of behavior where touching is involved a *frequency* style dimension could be used that applies to all behavior rules with touching behavior.

8.2.5 Oculesics

When covering differences concerning eye behavior, few things can be noticed. First, cultures can have different meanings for different eye behaviors. To give one example, in one culture lowering the eye gaze conveys respect while in another culture it may seem as evading. This difference in meaning could be modeled in the *Meaning-to-Signal Table*.

Second, there are also differences in style. In certain situations, some cultures gaze longer or more frequent than other cultures. These differences could be modeled using the style dimensions like *temporal* and *frequency* to tune these aspects of eye behavior. The styles would then have to be applied to a certain eye behavior in a specified situation.

8.2.6 Paralanguage

A few differences is the use of paralinguistic features have been observed. Cultural differences are visible in the rate and volume of speech, but also in the use of pauses. These differences could all be modeled using style dimensions like *vocal rate*, *vocal volume* and *use of silence*. These styles then have to be applied to all verbal behavior.

8.2.7 Chronomics

Cultures can have a very different attitude towards time. It can be difficult though to point out exactly what nonverbal behaviors are affected by this property. Does a more relaxed attitude results in doing everything slower? Probably

not, but for some behaviors this might be true. One example is the average walking pace. The style dimension that could be used for this behavior is the *temporal* dimension. This style could be applied to any behavior involved in moving the body to another location.

8.3 Styles for Culture Categorizations

In this section the use of *styles* will be evaluated when modeling cultural categorizations. The different categorizations have been described in section 3.1 and are handled below separately. Note that the style dimension values in this section are just examples to show differences in behavior between cultural categorizations.

8.3.1 Cultural Dimensions

The first cultural categorization that is handled uses Hofstede's cultural dimensions. It is difficult to detect a correlation between a cultural dimension and a nonverbal behavior dimension. One dimension though often shows differences in nonverbal behavior, namely *Individualism/Collectivism*. These differences are illustrated in figure 8.1.

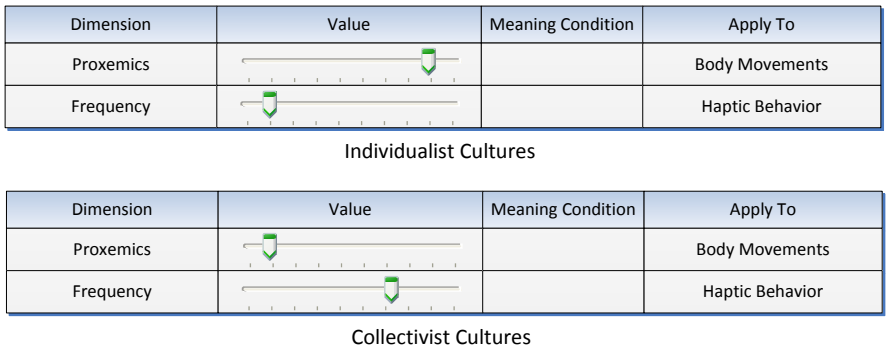


Fig. 8.1: Example Styles for Individualist and Collectivist Cultures.

8.3.2 Contact versus Non-Contact Cultures

The second cultural categorization concerns Hall's division of cultures into *contact cultures* and *non-contact cultures*. Different behavioral styles have been observed between these two categories. They look a lot like differences between

individualist and *collectvist* cultures, altough here a few more dimensions have been added. These are illustrated in figure 8.2.

Dimension	Value	Meaning Condition	Apply To
Proxemics			Body Movements
Frequency			Haptic Behaviors
Temporal			Eye Contact
Orientation			Body

Contact Cultures

Dimension	Value	Meaning Condition	Apply To
Proxemics			Body Movements
Frequency			Haptic Behaviors
Temporal			Eye Contact
Orientation			Body

Non-Contact Cultures

Fig. 8.2: Example Styles for Contact and Non-Contact Cultures.

8.3.3 Country Clusters

The last cultural categorization to evaluate is the categorization of cultures into country clusters as described in section 3.1.5. Some studies have shown that there may exist shared behavioral styles among people belonging to the same country cluster. Take for example the study where American students interviewed individuals from different country clusters, namely *African*, *Asian*, *European*, *Latin American* and the *Middle East* [61]. The interviewers needed to describe their perceptions of five nonverbal cues including *voice*, *space*, *eye behavior*, *facial expressions* and *hand gestures*. Some of the results of the study have been illustrated using *styles* in figure 8.3 and figure 8.4. Note that these dimension values are all *relative* to the norms of American Culture.

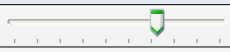
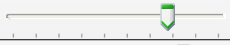
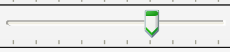
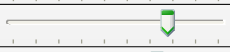
Dimension	Value	Meaning Condition	Apply To
Speech Volume			Speech
Speech Rate			Speech
Spatial Extent		Affect	Facial Expressions
Frequency		Affect	Smile
Proxemics			Body Movements
Temporal			Eye Contact
Frequency			Gestures
Spatial Extent			Gestures

Asian Cultures

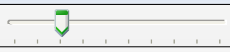
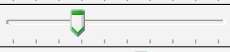
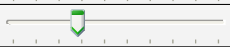
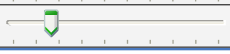
Dimension	Value	Meaning Condition	Apply To
Speech Volume			Speech
Speech Rate			Speech
Proxemics			Body Movements
Frequency			Eye Contact
Temporal			Eye Contact
Spatial Extent		Affect	Facial Expressions
Frequency		Affect	Smile

European Cultures

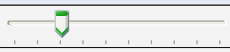
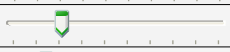
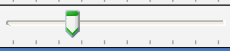
Fig. 8.3: Example Styles for different Country Clusters.

Dimension	Value	Meaning Condition	Apply To
Spatial Extent			Hand Gestures
Frequency			Eye Contact
Temporal			Eye Contact
Speech Rate			Speech
Spatial Extent		Affect	Facial Expressions
Frequency		Affect	Smile
Proxemics			Body Movements

Middle Eastern Cultures

Dimension	Value	Meaning Condition	Apply To
Speech Volume			Speech
Speech Rate			Speech
Frequency			Eye Contact
Spatial Extent		Affect	Facial Expressions
Frequency		Affect	Smile
Proxemics			Body Movements
Frequency			Hand Gestures
Spatial Extent			Hand Gestures

Latin-American Cultures

Dimension	Value	Meaning Condition	Apply To
Speech Volume			Speech
Speech Rate			Speech
Proxemics			Body Movements
Frequency			Eye Contact

African Cultures

Fig. 8.4: Example Styles for different Country Clusters.

9. APPLICATIONS

The *Culture Framework* is designed to be applicable in a broad range of domains. Any ECA system using communicative functions to generate communicative behavior could use the framework. Using the mapping table, customized function-to-behavior mappings can be made. In addition, a range of behaviors can be tuned at once using style dimensions.

The range of behavioral information that can be modeled depends on the function and behavior representations and the taxonomies used for the ECA identity and context factors. Further, the range of style dimensions that can be used depends on the mechanisms available for implementing these different types of dimensions (expressivity, paralinguistic features, frequency tuning, etc.)

A useful application for the framework could be a system using Human-ECA interactions which needs to be used by different groups of people. Studies have shown that people have more trust in ECAs to which they can relate and who behaves according to their norms. For example an ECA designed to provide people with information might be more accepted by the user if the ECA behaves according to the user's desired expectations. This should be taken into consideration when designing the behavior of an ECA. It could be desired for some ECA systems to be able to adapt the ECA's nonverbal behavior when it must be used by people from two or more completely different cultures. The same might be the case for people who differ in age, gender or even personality.

Another useful application for the framework could be a multi-agent system in which many different ECAs need to be implemented. These ECAs could be made more believable by letting them behave according to some identity. Furthermore, they could be designed to behave differently in a range of contexts where they communicate with different ECAs in different settings.

In some of these systems the behavior storage layer of the framework might be useful. Using the storage facilities where behaviors can be stored that need to be used by ECAs with different identities (culture, age, gender), behaviors could be designed for many ECAs at once. Styles could be specified that apply to a range of behaviors for multiple ECAs having different identities.

When it is important to model ECAs for different cultures, a cultural expert could aid in the specification of behavior and style rules when designing a culturally adapted ECA.

10. FUTURE WORK

Future work for this research covers different fields of research which will be handled separately below.

Style Dimensions

Because style dimensions are defined in such a way they can be used for a broad range of behavior tuning, more research would have to be done on identifying dimensions that might be useful. A few dimensions have been giving as examples in this document: dimensions to tune an ECA's expressivity, paralinguistic features, proxemic behavior, haptic behavior or behavior frequencies.

First, the range of other potentially useful dimensions must be investigated. Any extra information that needs to be assigned to (multimodal) behaviors and can be modeled using a dimension scale could be implemented using style dimensions. Second, when considering different dimension types, possibilities for using and implementing these styles in the *Culture Framework* must be investigated. What type of scale must be used for a dimension? How to implement the style? Does the style influences functions, behaviors or function-to-behavior mappings? What are the possibilities for implementing style dimensions in the behavior representation language that is used?

Application of Styles on Functions

Cultural differences in communication not only concern the surface level of communication (communicative behavior). It has been shown that they also concern the function level of communication. In some culture, people perform more long monologues while in other cultures people perform a very fast turn-taking. Some groups might perform more cognitive functions while other groups think less and talk more. These examples though are more applicable to personality differences. There could also be differences in the content of the conversation, where people from some cultures talk more about certain topics than people from other cultures. It would have to be investigated to what level it would

be worth implementing these differences. Also the implementation of these differences using styles could prove to be very complicated, because this would involve having access to an ECA's communicative goals and beliefs in order to have an influence on his communicative functions.

Cross-Cultural Conflicts

Future research might also include research on cross-cultural conflicts during communication. Research could be done on the origin and cause of these conflicts and a way to react or recover from the conflict. For example, we have seen that some emblem gestures can have opposite meanings in different cultures. Possibilities for generating a reaction in the ECA perceiving the behavior could be investigated. The ECA might interpret the perceived behavior differently depending on his cultural roots. This reaction could be in the form of updating certain beliefs, changing the emotional state of the ECA or generating a responsive communicative act.

FML and BML

In this research we have used FML and BML as example languages in the *Culture Framework*. Concerning these languages, still more research needs to be done on extending or modifying the current FML and BML capabilities of modeling communicative functions and behaviors. Currently, the focus of attention is on BML research, while FML is left more or less unattended.

Extending FML to a more broader but also higher-level model of communicative functions depends on the desire to influence an ECA's behavior more specific. For example, the current proposed FML specification contains a *performative* element which can be used to specify a *request* function. The specification of other higher-level functions which fall under this element might be desirable, like *politely ask* or *order*. Also, current FML lacks high-level functions which result in the expression of emblems like the *OK* sign or the *thumbs up* sign.

For these additional functions more experience is needed with FML. Therefore, attempts should be made to use FML and BML in ECA systems to discover further issues or drawbacks of the current specifications.

11. CONCLUSION

The *Culture Framework* is built around the main research solution for modeling important influences that different cultures can have on communicative behavior into a behavior generation framework for ECAs. Behavior generation in this framework is accomplished by mapping communicative functions to communicative behaviors using mapping and style tables containing generation rules. The solution involves the division of generation rules into *behavior* rules and *style* rules. Behavior rules define *what* (multimodal) behavior will be expressed for certain communicative function whereas style rules can provide additional properties of certain behaviors describing *how* or *how often* they will be expressed. Styles can consist of dimensions modifying the expressivity of nonverbal behaviors or paralinguisitc properties, but also properties handling the frequency or activation of certain (multimodal) behaviors or modalities. Styles are a powerful tool because they can be applied to all behaviors, unimodal behaviors, multimodal behaviors, specific modalities or to behaviors conveying a specific communicative function.

The combination of these rules provide the ability to model a large range of cultural differences in communicative behavior. The *Culture Framework* intergrates this solution in a behavior generation framework for ECAs. In addition, the framework provides a layer which handles the specification of the behavior and style generation rules by implementing the proposed *Information Model* into an authoring tool for behavior specification for ECAs. An intermediate behavior storage layer provides the ability to store generation rules that need to be used by ECAs having a certain identity (culture, age, gender). When equiping an ECA with behaviors, this layer can be used to retrieve generation rules that conform to the ECA's desired identity.

The framework was originally designed to model cultural differences in communicative behavior generation for ECAs. However, in human-to-human communication, there are many other factors besides culture that can have an influence on the use of communicative behavior. What behavior to express when performing a certain communicative function is usually influenced by a com-

plex mixture of factors, one of which may be the cultural roots of a person. Other factors include age, gender or relationships between people. Therefore, the framework to design had to be able to handle more complex configurations of factors influencing the behavior generation process. The design that resulted from these requirements provides the ability to model a lot more communicative behavior than just culturally dependent communicative behavior.

The range of information that can be modeled is fully dependent on the taxonomies that are used to specify the intrinsic and contextual influences on nonverbal behavior generation. The taxonomies to use can be highly domain dependent and will have to be specified when using the framework. Using the right taxonomies, any desired behavior can be modeled, depending on the modeling capabilities of the function and behavior representation languages that are used.

APPENDIX

A. FML SPECIFICATION PROPOSAL

A.1 Introduction

This Appendix contains the current FML specification with some modifications. Most part of the specification is taken from the FML specification developed at the *USC/Information Science Institute in Marina Del Rey [33].* Modifications and additions to the original specification are made in order to give a more complete model of communicative functions as described by Poggi et al. [63]. Unlike BML, FML is not being given attention at the moment and a formal specification does not exist by this time of writing. Therefore the specification described here should be seen as an optional specification in which some new tags, attributes or naming conventions are proposed. Alterations made to the ISI specification are outlined below each element.

A.2 FML Reference Proposal

The FML specification consists of two types of tags. The first type is used to identify the evolution of the discourse turn. They control the flow of the turns, topics and phrases. These are known as interactional tags. These tags are illustrated in section A.2.1.

The second type is used to give additional information about semantic units in the form of sentence parts, phrases or words which already have been structured by the first type. These are known as propositional tags. They are illustrated in section A.2.2.

Note that each FML element has an *ID* attribute associated with it which is not shown in the reference. This *ID* should be unique in the scope of each communication block and can be used to refer to a specific FML tag.

The ISI specification of FML contained the element ***coping*** which identifies a coping strategy that is employed by the speaker. This element has been removed because it's not a communicative function as seen by Poggi et al. [63]. A coping strategy can be seen as a dynamic influence factor (no. 15: cognitive resources as illustrated in figure 3.3).

A.2.1 Interactional Tags

element **TURN**

Description	Describes the speaking turn.				
Attributes	Name	Type	Use	Default	Description
	TYPE*	Name	required		The type of turn operation. [TAKE, GIVE, KEEP, REQUEST, ACCEPT]
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

* More **type** attribute values have been added.

element **TOPIC**

Description	Marks the current topic of discussion, reflecting discourse structure.				
Attributes	Name	Type	Use	Default	Description
	SHIFT*	int	required		Specifies the percentage of topicshift occurring.
	TOPICID**	string	required		An identifier specifying the topic.
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

* This attribute has been added.
** This attribute has been renamed from *ID* to *TOPICID*.

element **PERFORMATIVE***

Description	The speech act being performed.				
Attributes	Name	Type	Use	Default	Description
	TYPE	string	required		The type of speech act category. [INFORM, REQUEST, ENQUIRY]
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

* This element was previously known as ***intent***.

element **THEME**

Description	The part of a clause that ensures cohesion with the previous contribution.				
Attributes	Name	Type	Use	Default	Description
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

element **RHEME**

Description	The part of a clause that contains the new contribution to the discourse.				
Attributes	Name	Type	Use	Default	Description
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

element **CONTENT**

Description	Doesn't represent a communicative function, but specifies the actual content of the communicative act.

A.2.2 Propositional Tags

element **CERTAINTY***

Description	Describes how certain and assertive the speaker is about a specific unit.				
Attributes	Name	Type	Use	Default	Description
	AMOUNT	float	required	0.5	The level of certainty
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

* This element has been added.

element **ADJECTIVAL***

Description	Specifies an adjectival function that can be used to give additional information about a semantic unit.				
Attributes	Name	Type	Use	Default	Description
	TYPE	Name	required		The type of adjectival. [DIMINISH, INTENSIFY, INCLUSIVITY]
	LEVEL	Name	required	MEDIUM	The intended strength of the adjectival. [STRONG, MEDIUM, WEAK]
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

* This element has been added.

element **EMPHASIZE**

Description	Emphasizes a word or semantic unit.				
Attributes	Name	Type	Use	Default	Description
	LEVEL	Name	required	MEDIUM	The intended strength of the adjectival. [STRONG, MEDIUM, WEAK]
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

element **RELATION***

Description	Specifies the rhetorical relation in which the enclosed text is involved in.				
Attributes	Name	Type	Use	Default	Description
	TYPE	Name	required		The type of rhetorical relation. [CONTRAST, CAUSE, RESULT, LISTING]
	RELATIONID	string	required		The unique ID of the belief group. All elements with the same belief ID are considered to be in relation.
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

* This element has been added. The original specification contained a separate ***contrast*** element which has now been included here as subclass.

element **ILLUSTRATION**

Description	Indicates a feature of a discourse entity that could be clarified through illustration.				
Attributes	Name	Type	Use	Default	Description
	DESCRIPTION	string	required		Describes the elements of the illustration that can contribute to a depiction by a gesture for example.
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

element **REFERENCE**

Description	Identifies a discourse entity through a textual or visual reference.				
Attributes	Name	Type	Use	Default	Description
	TYPE	Name	required		The type of the entity being referred to. This is visual if the entity is visible in the environment and textual if the entity occurred in previous speech. [VISUAL, TEXTUAL]
	REFID*	string	required		The unique ID of the entity being referred to.
	SOURCE	string	optional		If textually evoked, this is the unique ID of the speaker who last referred to the entity.
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

* This attribute has been renamed from *ID* to *REFID*.

element **AFFECT**

Description	Describes the affective state of the speaker.				
Attributes	Name	Type	Use	Default	Description
	TYPE*	Name	required		The affect being displayed. [HAPPINESS, SADNESS, ANGER, FEAR, SURPRISE, DISGUST, EMBARRASSMENT, SHAME]
	STANCE	Name	optional	INTENTED	Describes whether the emotion is intentionally given off or involuntarily leaked. [INTENDED, LEAKED]
	TARGET	string	optional		The ID of the person who is being targeted by the affective behavior.
	EVENT	string	optional		Event that triggered the change in affective state.
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

* For now only the eight basic emotions that are universally recognized are included. This can be extended in the future.

element **COGNITIVE***

Description	Describes a meta-cognitive activity associated with a specific unit.				
Attributes	Name	Type	Use	Default	Description
	TYPE	Name	required		The type of cognitive activity. [THINK, REMEMBER, INFER, DECIDE]
	START	string	required		Time marker that identifies when this communicative function takes effect.
	END	string	required		Time marker that identifies when this communicative function is no longer in effect.

* This element has been added.

B. BML REFERENCE

B.1 Introduction

This Appendix contains the current BML specification. It includes the specification updates from the HUMAINE WP10 workshop in Vienna, November 6-8, 2006 [42]. Note that the language is sill under development changes to the tags may be proposed in the future. This specification contains only the actual behavior tags and not the synchronization and event mechanisms. For the full specification see the official BML specification [32].

B.2 BML Behavior Tags

element **HEAD**

Description	Movement of the head independent of the eyes (nodding, shaking, tossing and orienting).				
Attributes	Name	Type	Use	Default	Description
	TYPE	Name	required		The category of head movement. [NOD, SHAKE, TOSS, ORIENT]
	AMOUNT	float	optional	0.0	(NOD, SHAKE TOSS) The extent of the movement where 1.0 is fully extended and 0.0 is least extended.
	REPEATS	int	optional	1	(NOD, SHAKE TOSS) Number of times the basic head motion is repeated.
	TARGET	string	optional		(ORIENT) The world ID of the reference target.
	ANGLE	float	optional	0.0	(ORIENT) Orients the head this number of degrees in the specified direction from the current head position. If a target is also given, the orientation is relative to the orientation towards the target.
	DIRECTION	Name	optional	RIGHT	(ORIENT) Direction of orientation angle. [RIGHT, LEFT, UP, DOWN, ROLLRIGHT, ROLLLEFT]

element **TORSO**

Description	Movement of the orientation and shape of the spine and shoulder.				
Attributes	Name	Type	Use	Default	Description
	POSTURE	Name	required		The name of the posture to assume.
	TRANSITION	Name	optional		The name of the animated transition that gets to get played before final posture is assumed.

element **FACE**

Description	Movement of facial muscles to form certain expressions (eyebrow, eyelid, mouth movements.)				
Attributes	Name	Type	Use	Default	Description
	TYPE	Name	required		The part of the face being controlled. [FACS, EYEBROWS, EYELIDS, MOUTH]
	AMOUNT	float	optional	0.5	The amount of movement where 0.0 is the lowest (or closed) position and 1.0 is the highest (or open) position.
	SIDE	Name	optional	BOTH	Which side of the face is being controlled. [BOTH, LEFT, RIGHT]
	AU	int	optional	0	(FACS) The Action Unit (AU) reference number for a Facial Action Coding System (FACS) expression.
	SHAPE	Name	optional	FLAT	(EYEBROWS) The shape given to the eyebrows. [FLAT, POINTUP, POINTDOWN]
	SEPARATION	float	optional	0.5	(EYEBROWS) The horizontal distance of the eyebrows from the center of the forehead where 0.0 is the shortest distance and 1.0 is the furthest distance.
	LID	Name	optional	BOTH	(EYELIDS) Where both upper and lower eyelids are affected. [BOTH, UPPER, LOWER]
	SHAPE	Name	optional	FLAT	(MOUTH) The shape given to the mouth. [FLAT, SMILE, LAUGH, PUCKER, FROWN]

element **GESTURE**

Description	Coordinated movement with arms and hands, including pointing, reaching, emphasizing (beating), depicting and signaling.				
Attributes	Name	Type	Use	Default	Description
	TYPE	name	required		The category of gesture movement [POINT, REACH, BEAT, DEPICT, SIGNAL].
	NAME	string	optional		The name of a gesture needed for a DEPICT or a SIGNAL gesture.
	TARGET	string	optional		The WorldID of a reference target for POINT and REACH gestures.

element **GAZE**

Description	Coordinated movement of the eyes, neck and head direction, indicating where the character is looking.				
Attributes	Name	Type	Use	Default	Description
	TARGET	string	optional		The World ID of the reference target.
	ANGLE	float	optional	0.0	Orients the gaze this number of degrees in the specified direction from the current gaze orientation. If a target is also given, the orientation is relative to the orientation towards that target.
	DIRECTION	Name	optional	RIGHT	Direction of the orientation angle. [RIGHT, LEFT, UP, DOWN, UPRIGHT, UPLEFT]

element **BODY**

Description	Full body movement, generally independent of the other behaviors. Types include overall orientation, position and posture.				
Attributes	Name	Type	Use	Default	Description
	APPROACH	string	optional		The World ID of a target place or thing to approach prior to assuming an indicated posture.
	PROXIMITY	float	optional	1.0	How close to approach the target, where 1.0 is 'typical' distance for that target (defined elsewhere), and 0.0 is up against the target
	FACE	string	optional		The World ID of a reference target for final facing.
	ANGLE	float	optional	0.0	The offset angle of final facing, where 0.0 fully faces the reference target.
	POSTURE	Name	optional		The name of the posture to assume.
	TRANSITION	Name	optional		The name of the animated transition that gets played before final posture is assume.

element **LEGS**

Description	Movements of the body elements downward from the hip: pelvis, hip, legs including knee, toes and ankle.				
Attributes	Name	Type	Use	Default	Description
	POSTURE	string	required		The name of the posture to assume.
	TRANSITION	string	optional		The name of the animated transition that gets played before final posture is assumed.

element **SPEECH**

Description	Verbal and paraverbal behavior, including the words to be spoken (for example by a speech synthesizer), prosody information and special paralinguistic behaviors (for example filled pauses).				
Attributes	Name	Type	Use	Default	Description
	TYPE	string	optional	text/plain	MIME type or other string identifying the type of contents or referred object.
	REF	Name	optional		Refers to speech data if not contained within the speech element.
	TEXT	string	optional		Unprocessed element to promote legibility with external or encoded types.

element **LIPS**

Description	This element is used for controlling lip shades including the visualization of phonemes for audiovisual speech.				
Attributes	Name	Type	Use	Default	Description
	VISEME	Name	required		The name of a viseme to be displayed. It will blend with any expression specified in the FACE element.
	ARTICULATION	float	optional	0.5	The extent to which visemes are clearly articulated, where 0.0 represents sloppy and 1.0 represents hyper articulation.
	FLAPPING	boolean	optional	false	If true, keeps the mouth oscillating between the viseme and the closed position.

BIBLIOGRAPHY

[1] A. Agliati, A. Vescovo, and L. Anolli. Conversation patterns in icelandic and italian people: Similarities and differences in rhythm and accomodation. In L. Anolli, S. Duncan Jr., M.S. Magnusson, and G. Riva, editors, *The hidden structure of interaction. From neurons to culture pattern*, pages 223–236. Amsterdam: IOS Press, 2005.

[2] M. Argyle and M. Cook. *Gaze and Mutual Gaze*. Cambridge University Press, Cambridge, 1976.

[3] GRETA Group: Embodied Conversational Agent Research at the IUT Montreuil: http://www.iut.univ paris8.fr/greta/.

[4] R. Brislin. *Understanding culture's influence on behavior*. Hartcout Brace COllege Publishers, New York, 1993.

[5] J. K. Burgoon, D.B. Buller, and W.G. Woodall. *Nonverbal communication: The unspoken dialogue*. New York: McGraw-Hill Companies, Inc., 1996.

[6] E. Magno Caldognetto and I. Poggi. Creative iconic gestures: some evidence from aphasics. In R. Simone, editor, *Iconicity in Language*. John Benjamins, Amsterdam, 1995.

[7] B. De Carolis, V. Carofiglio, and C. Pelachaud. From discourse plans to believable behavior generation. In *Proceedings of INLG'02*, 2002.

[8] B. De Carolis, C. Pelachaud, I. Poggi, and M. Steedman. Apml, a mark-up language for believable behavior generation. In H. Prendinger and M. Ishizuka, editors, *Life-like Characters. Tools, Affective Functions and Applications*, pages 65–85. Springer, 2004.

[9] J. Cassell. More than just a pretty face: Embodied conversational agents. *Communications of the ACM*, 43(4):70–78, 2000.

[10] J. Cassell. Nudge nudge wink wink: Elements of face-to-face conversation for embodied conversational agents. In E. Churchill, editor, *Embodied Conversational Agents*. MIT Press, Cambridge, MA, 2000.

[11] J. Cassell, T. Bickmore, M. Billinghurst, L. Campbell, K. Chang, H.H. Vilhjálmsson, and H. Yan. Embodiment in conversational interfaces: Rea. In *Proceedings of ACM CHI 99 Conference on Human Factors in Computing Systems*, volume 1 of *Virtual Reality and Embodiment*, pages 520–527, 1999.

[12] J. Cassell, J. Sullivan, S. Prevost, and E. Churchill. *Embodied Conversational Agents*. MIT Press, Cambridge, MA, 2000.

[13] J. Cassell and H.H. Vilhjálmsson. Fully embodied conversational avatars: Making communicative behaviors autonomous. *Autonomous Agents and Multi-Agent Systems*, 2(1):45–64, 1999.

[14] J. Cassell, H.H. Vilhjálmsson, and T. Bickmore. Beat: The behavior expression animation toolkit. In *Proceedings ACM SIGGRAPH*, pages 477–486, Los Angeles, 2001.

[15] B. de Carolis, C. Pelachaud, I. Poggi, and F. de Rosis. Behavior planning for a reflexive agent. In B. Nebel, editor, *Proceedings of the seventeenth International Conference on Artificial Intelligence (IJCAI-01)*, pages 1059–1066, San Francisco, CA, 2001. Morgan Kaufmann Publishers, Inc.

[16] F. de Rosis, C. Pelachaud, and I. Poggi. Transcultural believability in embodied agents: a matter of consistent adaptation. In S. Payr and R. Trappl, editors, *Agent Culture: Designing Human-Agent Interaction in a Multicultural World.* Laurence Erlbaum Associates, New York, 2004.

[17] F. de Rosis, I. Poggi, C. Pelachaud, V. Carofiglio, and B. De Carolis. From greta's mind to her face: Modeling the dynamics of affective states in a conversational embodied agent. *International Journal of Human-Computer Studies*, 59:81–118, 2003.

[18] P. Ekman. Universals and cultural differences in facial expressions of emotion. In J. Cole, editor, *Nebraska Symposium on Motivation 1971*, volume 19, pages 207–282. University of Nebraska Press., 1972.

[19] P. Ekman and W.V. Friesen. The repertoire on non-verbal behavior: Categories, origins, usage and codings. *Semiotica*, 1:49–98, 1969.

[20] P. Ekman and W.V. Friesen. Constants across cultures in the face and emotion. *Personality and Social Psychology*, 17:124–129, 1971.

[21] P. Feyereisen and J. de Lanno. *Gestures and Speech: Psychological investigations*. Cambridge University Press., 1991.

[22] C. Geertz. *The Interpretation of Cultures: selected essays*. New York: Basic Books., 1973.

[23] W. Goodenough. Cultural anthropology and linguistics. *Georgetown University Monograph Series on Language and Linguistics*, 9:167–173, 1957.

[24] E.T. Hall. A system for the notation of proxemic behavior. *American Anthropologist*, 65:1003–1026, 1963.

[25] E.T. Hall. *The Hidden Dimension*. Archor Books, 1966.

[26] E.T. Hall. *Beyond Culture*. Anchor Books, 1976.

[27] C. Hampden-Turner and F. Trompenaars. *Riding the Waves of Culture: Understanding Diversity in Global Business*. McGraw-Hill, 1997.

[28] M.S. Hanna and G.L. Wilson. *Communicating in business and professional settings*. New York: McGraw-Hill Companies, Inc., 1998.

[29] R. Heslin. Steps toward a taxonomy of touching. In *Paper presented at the annual convention of Midwestern Psychological Association*, Chicago, IL, 1974.

[30] Geert Hofstede. *Cultures and Organizations: Software of the mind*. McGraw-Hill, 2004.

[31] Geert Hofstede Cultural Dimensions: http://geert hofstede.com.

[32] BML Specification: http://wiki.mindmakers.org/projects:bml:main.

[33] Information Science Intistute: http://www.isi.edu.

[34] Virtual Human Markup Language (VHML): http://www.vhml.org.

[35] Speech Synthesis Markup Language: http://www.w3.org/TR/speech synthesis/.

[36] K. Isbister and C. Nass. Consistency of personality in interactive characters: verbal cues, non-verbal cues, and user characteristics. *International Journal of Human-Computer Studies*, 53(2):251–267, 2000.

[37] S.E. Jones and A.E. Yarbrough. A naturalistic study of the meanings of touch. *Communication Monographs*, 52:19–56, 1985.

[38] D. Keltner and B.N. Buswell. Evidence for the distinct form of embarrassment, shame, and guilt: A study of recalled antecedents and facial expressions of emotion. *Cognition and Emotion*, 10:155–171, 1996.

[39] A. Kendon. Human gesture. In T. Ingold and K.R. Gibson, editors, *Tools, Language and Cognition*, pages 43–62. Cambridge University Press, 1993.

[40] M. S. Kirch. Non-verbal communication across cultures. *The Modern Language Journal*, 63(8):416–423, 1979.

[41] M.L. Knapp and J.A. Hall. *Nonverbal Communication in Human Interaction*. Harcourt Brace College Publishers, Fort Worth, Texas, 1997.

[42] S. Kopp. Protocol gesticon/bml meeting, humaine wp10 ws, vienna, nov 6-8 2006.

[43] S. Kopp, B. Krenn, S. Marsella, A.N. Marshall, C. Pelachaud, H. Pirker, K. Thórisson, and H.H. Vilhjálmsson. Towards a common framework for multimodal generation: The behavior markup language. In J. Gratch, M. Young, R. Aylett, D. Ballin, and P. Olivier, editors, *Intelligent Virtual Agents, 6th International Conference, IVA 2006, Marina Del Rey, CA, USA, August 21-23, 2006, Proceedings*, volume 4133 of *Lecture Notes in Computer Science*, pages 205–217. Springer, 2006.

[44] A. Krandsted, S. Kopp, and I. Wachsmuth. Murml: A multimodal utterance representation markup language for conversational agents. In *Proceedings of the AAMAS Workshop on Embodied conversational agents Lets specify and evaluate them!*, Bologna, 2002.

[45] B. Krenn, M. Lamolle, C. Pelachaud, H. Pirker, and M. Schröder. Humaine: Report on representation languages. Technical report, Information Society Technologies, 2006.

[46] B. Krenn and H. Pirker. Defining the gesticon: Language and gesture coordination for interacting embodied agents. In *Proceedings of the AISB-2004 Symposium on Language, Speech and Gesture for Expressive Characters*, University of Leeds, UK, 30-31 March 2004. Convention for the Study of Artificial Intelligence and the Simulation of Behavior.

[47] M. Lamolle, M. Mancini, C. Pelachaud, S. Abrillian, J.-C. Martin, and L. Devillers. Contextual factors and adaptive multimodal human-computer interaction: Multi-level specification of emotion and expressivity in embodied conversational agents. In A. Dey, B. Kokinov, D. Leake, and R. Turner,

editors, *Modeling and Using Context: 5th International and Interdisciplinary Conference*, pages 225–240. Springer, Berlin, 2005.

[48] Tactical Iraqi Language and Culture Training System: http://www.tacticallanguage.com.

[49] E.-J. Lee and C. Nass. Does the ethnicity of a computer agent matter? an experimental comparison of human-computer interaction and computer-mediated communication. In *Proceedings of the Workshop on Embedded Conversational Characters Conference*, pages 123–128, Lake Tahoe, CA, 1998. ACM Press. New York.

[50] J. Lee and S. Marsella. Nonverbal behavior generator for embodied conversational agents. In *Proceedings of the 6th International Conference on Intelligent Virtual Agents*, Marina del Rey, CA, 2006.

[51] A. Leffler, D.L. Gillespie, and J.C. Conaty. The effects of status differentiation on nonverbal behavior. *Social Psychology Quarterly*, 45:153–161, 1982.

[52] R.V. Levine and A. Norenzayan. The pace of life in 31 countries. *Journal of Cross-Cultural Psychology*, 30(2):178–205, 1999.

[53] C.W. Ling. Crossing cultural boundaries. *Nursing*, 72(3):32–34, 1997.

[54] M. Mancini, B.Hartmann, and C. Pelachaud. Non-verbal behaviors expressivity and their representation. Technical report, Univerity of Paris 8, 2004.

[55] V. Maya, M. Lamolle, and C. Pelachaud. Influences on embodied conversational agent's expressivity: Toward an individualization of the ecas. Technical report, University of Paris VIII, 2004.

[56] D. McNeill. Hand and mind: What gestures rveal about thought. *Language*, 70(2):345–350, 1992.

[57] A. Mehrabian. *Silent Messages*. Wadworth Publishing Company, 1971.

[58] H. Morton, H. McBreen, and M. Jack. Experimental evaluation of the use of ecas in ecommerce applications. In Z.M. Ruttkay and C. Pelachaud, editors, *From brows to trust: evaluating embodied conversational agents*, pages 293–321. Kluwer Academic Publishers, 2004.

[59] C. Nass, K. Isbister, and E.-J. Lee. Truth is beauty: researching embodied conversational agents. In *Embodied conversational agents*, pages 374–402. MIT Press, 2000.

[60] H. Noot and Z.M Ruttkay. Gesture in style, 2003.

[61] T.H. Ostermeier. Differences in meanings for nonverbal cues and ease/difficulty in intercultural listening, 2000.

[62] C. Pelachaud, V. Maya, and M. Lamolle. Representation of expressivity for embodied conversational agents. In *Workshop on Balanced Perception and Action in ECAs, at AAMAS'04*, New York, 2004.

[63] I. Poggi, C. Pelachaud, and F. de Rosi. Eye communication in a conversational 3d synthetic agent. *AI Communications*, 13(3):169–181, 2000.

[64] I. Poggi, C. Pelachaud, F. de Rosis, V. Carofiglio, and B. De Carolis. Greta: A believable embodied conversational agent. In O. Stock and M. Zancanaro, editors, *Multimodal Intelligent Information Presentation*, volume 27 of *Text, Speech and Language Technology*, pages 3–25. Springer Netherlands, 2005.

[65] M.S. Remland and T.S. Jones. Interpersonal distance, body orientation, and touch: Effects of culture, gender and age. *Journal of Social Psychology*, 135(3):281–297, 1995.

[66] S. Ronen and O. Shenkar. Clustering countries on attitudinal dimensions: A review and synthesis. *The Academy of Management Review*, 10(3):435–454, 1985.

[67] D. Rousseau and B. Hayes-Roth. Personality in synthetic agents. Technical report, Stanford Knowledge Systems Laboratory, 1996.

[68] Z.M. Ruttkay and H. Noot. Variations in gesturing and speech by gestyle. *International Journal of Human-Computer Studies, Special Issue on 'Subtle Expressivity for Characters and Robots'*, 62(2):211–229, 2005.

[69] Z.M. Ruttkay, C. Pelachaud, I. Poggi, and H. Noot. Exercises of style for virtual humans. In L. Canamero and R. Aylett, editors, *nimating Expressive Characters for Social Interactions, Advances in Consciousness Research Series*. John Benjamins Publishing Company, to appear.

[70] Z.M. Ruttkay, V. van Moppes, and H. Noot. The jovial, the reserved and the robot. In *Proceedings of the AAMAS03 Ws on Embodied Conversational Characters as Individuals*, Melbourne, Australia, 2003.

[71] S.H. Schwartz. Beyond individualism/collectivism: New cultural dimensions of values. In U. Kim, H.C. Triandis, C. Kagitcibasi, S. Choi, and G. Yoon, editors, *Individualism and Collectivism: Theory, method and application*, pages 85–119. Sage, Thousand Oaks, CA, 1994.

[72] K. Thórisson. Gandalf: An embodied humanoid capable of real-time multimodal dialogue with people. In *Proceedings of the First International Conference on Autonomous Agents*, pages 536–537, Mariott Hotel, Marina Del Rey, CA, February 5-8 1997. ACM Press.

[73] E. P. Tylor. *The Origins of Culture and Religion in Primitive Culture*. New York: Harper and Brothers, 1958.

[74] M.F. Vargas. *Louder than words: An Introduction to Non-Verbal Communication*. Ames, Iowa State University Press., 1986.

[75] H.H. Vilhjálmsson and J. Cassell. Bodychat: autonomous communicative behaviors in avatars. In *AGENTS '98: Proceedings of the second international conference on Autonomous agents*, pages 269–276, Minneapolis, Minnesota, United States, 1998. ACM Press.

[76] H.H. Vilhjálmsson and S. Marsella. Social performance framework.

[77] J. Walker, L. Sproull, and R. Subramani. Using a human face in an interface. In *Proceedings of CHI'94*, pages 85–91, 1994.

[78] O.M. Watson. *Proxemic Behavior: A cross-cultural study*. Mounton de Gruyter, 1970.

[79] O.M. Watson and T.D. Graves. Quantitative research in proxemic behavior. *American Anthropologist*, 68:971–985, 1966.

[80] D.E. Whitten and P. Hunter. *The Encyclopedia of Anthropology*. Harper and Row, 1976.