

This is an OCR-Version from the book, I did this in order to keep the pagination of the original

New Aspects of Human Ethology

**Edited by Schmitt, A., Atzwanger, K., Grammer, K and Schaefer, K.
1997 PlenumPress: London, NewYork.**

THE COMMUNICATION PARADOX AND POSSIBLE SOLUTIONS

Towards a Radical Empiricism

Karl Grammer,¹ Valentina Filova,² and Martin Fieder¹

¹Ludwig-Boltzmann-Institute for Urban Ethology c/o Institute for Human
Biology/University of Vienna Althanstrasse 14 A-1090 Vienna/Austria

²Institute for Automation-Department for Pattern Recognition and Image Processing
Technical University Vienna Treitlstrasse 3 A-1040 Vienna/Austria

In the history of both animal and human ethology the direct observation of unstaged interactions in a natural habitat plays a critical role for methodological and theoretical considerations. Even when ethologists think that they already know much about adaptations and the ways in which they interact with the environment, the principles which have been involved in the evolution of increasingly complex human behaviour are still not very well understood.

A major reason for this lies in methodological problems connected with the observation and description and the nature of human behaviour itself. In order to assess causation and function of behaviour we rely on an "observational device." The process of information reduction which is applied to the study of behaviour results in highly variable observations. The assessment of meaning and function rarely produces reproducible results, and different signals especially in human communication seem to take many meanings which are context-specific. Partially this might be due to the observational approaches used for coding behaviour.

1. WHAT IS COMMUNICATION?

A straightforward definition of communication is not difficult. As a starting point we can define it as the transfer of information between two communicative units. Ethology has created many models for the process of information transfer. Basic to these approaches is the term "signal," an information carrier which is produced through encoding information in an

communication channel by a sender. This signal is decoded by a receiver who adds information to the signal and then decodes its meaning. In a classical ethological approach many of the signals are a result of evolutive constraints and work in a quasi automatic way. Although many human signals have been isolated as cultural universals (Eibl-Eibesfeldt, 1972) a closer look reveals high variability. For instance Ekman and Friesen (1971) propose cross-cultural universal facial signals for emotions. An observational approach to the study of emotions in every day life reveals a high variability in the production of patterns and pure emotion patterns which occur rarely (Grammer *et al.*, 1988).

The Lorenz-Tinbergen approach sees signals as discrete and deterministic: A sends a signal X and B decodes and returns Y. In this definition, visual, tactile, acoustical and verbal information are divided into units of meaning. Signals have a lexical structure: one entry in the lexicon has one specific meaning and one definite function. The basic assumption of this approach is that signals exist as independent units and sender and receiver share a common code. The signals themselves are considered as discrete units of movements each with a beginning and an end in time, which in turn can frame a static component like a posture. This is clearly demonstrated by the acoustic properties of laughter which separate it from speech (Provine and Young, 1991), or by bodily movements in interactions, like illustrating hand-movements (Ekman and Friesen, 1972). Thus, signals have a content which is different for each signal, and which makes them identifiable reliably. If so, signals have to show a certain form constancy which is necessary for identification and two different signs do not overlap in their meaning. Most of the times form constancy defines the relation of movements to each other, like the typical head movements connected with laughter (Grammer and Eibl-Eibesfeldt, 1989). If laughter occurs in interactions, the head is moved in a circular fashion away from the partner.

Another approach to signaling uses the same basic structure, but it takes the probabilistic nature of communication into account. It still assumes that signals are discrete: A sends X and B decodes and returns Y, or not (Argyle, 1988). These probabilistic models imply some law of summation over space and time. This means that A sends X and Z and T to B at the same time, or A sends X and Z and T sequentially. This model of communication is called "modulated communication" (Markl, 1985). Grammer (1995) proposed that "natural" signs sent in parallel, like age and sex, signals of dominance/ submission or emotions contain the decoding instruction for a signal. For instance age and sex of the sender can modulate the meaning of a smile from a "come-on" to a simple friendly gesture. A slightly different approach was suggested by Schleidt (1973) as tonic communication. He assumed that meaning could be encoded in the form of pulse rate modulation. The sender sends a signal of uniform height and duration repeatedly in distinctive intervals. The receiver then applies some kind of low pass filter in order to integrate the signals over time. The effect on the receiver then is a slowly accumulating, tonic one.

For the observation of communication behavioural categories are constructed as classes or prototypes of signals. These classes have to be stereotypical, homogenous and discrete in order to be reliable. In addition we want to avoid functional descriptions because the goal of ethological approaches is the description of function itself. Thus, in order to produce reliable results the reduction of information has to be enormous. An example for such a classical approach is provided by Grammer (1991) who paired strangers randomly and tried to find out if body movements and postures could predict self-reported interest in another person. The resulting clusters of movements proved to be inconsistent and unreliable in predicting interest in the other person.

Furthermore, single postures did not covary with reported interest in this study. But as soon as postures and vocal stimuli were combined, the situation changes dramatically.



Figure 1. Body postures and laughter. The figure shows body poses 2 seconds before two strangers laughed in a waiting room experiment (Grammer, 1991). The postures were constructed by regression analysis from 4547 postures. In (a) the most frequent postures during laughter are shown. In the right picture the view from the female is shown, the middle picture shows the observers view and the left picture shows the view from the male. In (b) the postures are shown males and female take during laughter which signals aversion and in (c) postures during laughter which signals approval and interest in the other person are shown. Interestingly there is an additive effect for the single posture elements in the different body parts: the more of the single elements are present, the higher is the correlation with the self-reported intentions (Wire frame models by A. Jtte).

It was possible to show that postures which are taken during laughter might well transport the meaning of laughter. The acoustic event laughter does not alone delineate interest from no interest. Highly interested males or females do not laugh more often than persons with no interest. Moreover, people who are together with strangers of the same sex laugh more often. In sum, there is no contextual evidence that laughter alone is a sexual signal. When combined with postures however, laughter may take different meanings on a continuum from rejection to appraisal of the partner.

Evidence from the consequences of particular signals reveals many contradictions. The same behaviour can have different meanings. *open legs* among females and *Hair Flip* (Fig. 2) where the hair is moved out of the face with the hand and the head tossed backward, indicates low female interest. If both behaviours are combined with laughter they covary with high interest. other behaviours may take different meaning when they are static or dynamic. The *Head Akimbo*, a behaviour where the breast is pushed out and the hands are folded behind the neck, is associated with high interest when it occurs as a



Figure 2. Hairflip a female mannerism. The Hairflip consists of a typical movement sequence which starts with a slight head tilt, followed by a head up movement. The hand reaches out into the hair and the head turns back into the starting position with gaze aversion. This movement is performed more often by females (Grammer, 1991) than by males.

movement during laughter, but with low interest when it occurs as posture before laughter (Grammer, 1991).

The communicative situation presents itself as unclear and ambiguous although the receiver seems to be able to decode the sender's intentions. Receivers are generally aware of what the sender wants to tell them. Thus the decoding of meaning in interactions can not be described with a simple signal oriented approach. As an alternative, we could speculate that meaning and intentions are communicated solely through the verbal channel by speech content itself. Krauss et al. (1981) showed that speech accompanying gestures were not related to the decoding of meaning. They assume that body and arm movements are results of speech production itself and propitiate speech production. In earlier rating experiment of politicians Krauss et al. (1981) had already shown that verbal information dominates visual and auditive information. If we agree with this approach it seems useless to search for signaling intentions in human non-verbal behaviour. In contrast to the above mentioned results, Mehrabian (1972) showed in series of experiments the relative role facial expression, vocal behaviour and speech content play in the perception of persons. Mehrabian comes to the general conclusion that non-verbal behaviour plays the main role for the decoding of meaning. He finds that the meaning of messages is determined to 55% by visual information, 38% by vocal information and only 7% by speech content. These relative relations have been replicated by Siddiqi et al. (1973) and Wallbott (1991).

If we look at the content of the information which is transferred between interactants, we find that facial information is used for decoding tendencies of dominance and positive affect (Rosenthal and Depaulo, 1979). Ekman et al. (1980) gave information from seven different communicative channels: only facial information, only bodily information, speech, filtered voice, transcribed speech content, and combinations of voice and speech, voice, body and speech. In this research none of the presented channels was dominating the others in the transfer of meaning.

3 What Is Communication for?

Communicative models are a description of a communicative process of information transfer. Most models will fail when it comes to explain the function of communication because of the nature of communication itself, and the tools which are applied to reduce the information, as we will show later. Social groups are complex structures and their main feature is that the goals of the members rarely are in accordance. Human groups can be seen as an agglomeration of conflicting interests. This fact ultimately may be the driving force behind the evolution of social intelligence. Proximately it may be the basic constraint for communication and thus the generation of signals in any channel of communication.

The probabilistic multi-meaning nature of human communication is present in verbal and in non-verbal communicative acts. Linguistic research shows that indirectness of verbalisation and verbal acts like "hedging" depend on the risk of the intended communicative act (Brown and Levinson, 1978). If the benefits for the sender are high and the costs for the receiver are also high, it is obvious that the risk of not reaching the pursued goal for the sender is also high. As a result, the sender has to use signals and actions which allow to manipulate the receiver in the sense of the sender. Evolutionary theorists have forwarded comparable ideas. openly presenting intentions in communication might not pay (Dawkins and Krebs, 1981) because the signal receiver might act directly against the sender's intentions. The sender thus would not be able to reach his/her goals. In addition, as soon as the receiver recognises the intentions of the sender the probability of deception might rise. This situation

drives any type of communication into manipulative efforts. The manipulative component of a signal has to force the receiver into a certain state where he is willingly accepting the goals of the sender, preferably without recognising that he was manipulated. This situation is the communicative paradoxon: showing intentions and not getting caught by a suspicious receiver. In this view the function of communication is manipulation and is used for risk dependent transfer of information. Thus, a prerequisite for any communicative model is the assessment of risk, which will be highly context dependent. Risk itself is created by the goal under quest, i.e. the imposition for the receiver, the relationship between the interactants and motivational factors. In our introductory example of the waiting room situation risk should be high for a person who develops interest in the other person. Risk is determined by the possible costs and benefits for both sender and receiver. Risk dependent communication allows the explanation of simple straightforward transfer of meanings (under low risk conditions) and highly ambiguous transfer of information in situations with high risk. In this view both verbal and non-verbal channels can be affected. So, any model of communication should take risk assessment into account. Non-verbal behaviour may be an important tool in high risk situations because of its non-binding standard, when compared to verbal behaviour (Grammer *et al.*, 1996). The contradiction in the results which show either dominance of verbal information over visual information or vice-versa lies in the fact that an independent rater of situations had no possible costs in such an experiment, nor had the sender of the signal. This trap forces research on communication to work under naturalistic conditions, that is to observe unstaged social interactions.

1.1.1. Direct Communication. How can the sender achieve the delicate task of risk dependent communication? The sender has to assess risk and to act accordingly. The production of signals then could be optimised. Signals used in a situation of high benefit and/or high costs for both the receiver and the sender should be easily decodable. In this case the only preconditions of effective signaling are low environmental noise, encoding error by the sender and decoding errors by the receiver. An important means to actively reducing these errors is to increase contrast in a signal. Another mechanism is to produce a signal repeatedly and constantly over time. Both processes lead to ritualisation of signals. A ritualised signal consists only of a few elements that are produced repeatedly and in a fixed sequence. The aim of ritualisation is to make a signal definite and unmistakable. Grammer and Eibl-Eibesfeldt (1989) have shown that laughter follows ritualisation principles. Under high risk conditions, female laughter becomes more stereotypic, the threshold for performance is lowered and it is accompanied by typical movement sequences. As soon as risk becomes more asymmetric and the possibility of deception rises, the communicative situation changes drastically.

1.1.2. Lying, Deception and Mind Reading. The first possibility is the use of deception in the sense of sending false information. There are some constraints connected to lying. An example are children who cry in conflicts (Grammer, 1992~. A crying child who is engaged in a conflict with another child receives support by a third child in most cases. The risk of the supporter is high in such a situation, because he might get attacked. The probability of getting support depends on the frequency the child cried in the past. If it cries too often and uses crying in a deceptive way (i.e. to receive support) he/she won't receive any longer support. The receiver only engages in support if the honesty of the signal is guaranteed. Thus the use of deception will depend from the frequency in which it is used and the costs and benefits connected to the interaction. Harper (1992) pointed out that deception can only occur when the frequency of deception is low, the signal has little

costs for the sender and high benefits for the receiver. This situation forces the receiver to apply "mind-reading" and try to find additional cues for possible detection of deception. If the sender tries to deceive the receiver, the sender will try to control his behaviour in order to avoid detection. Yet control is rarely complete. If the sender tries to control his emotions he for instance creates leaks in the rest of his non-verbal behaviour. The receiver then will be able to detect the deception (Ekman and Friesen, 1969). Therefore lying is not always a solution for the communicative paradox.

The second form of deception is withholding information. According to Harper (1992) this is the main form of deception. Even if its use is widespread, the signal sender has to clarify his intentions sooner or later, or he will not be able to reach his goals.

1.1.3. Direct Cognitive and Physiological Manipulation: ~ Smile Is Not Just a Smile. Direct manipulation of the cognitive apparatus or the physiology of the receiver can play a role. Lorenz (1973) and later Cosmides et al. (1992) proposed that our information processing apparatus was formed and optimized in the course of evolution. If our brains are optimized for adaptive information processing then these adaptive structures can be exploited. An example of this possibility is the perception of emotions. The signal receiver experiences the same physiological changes as the sender of an emotion (Ekman et al., 1983). Thus by sending signaling "emotion" the sender is able to influence the physiology of the receiver. In the case of a smile, this makes sense because emotions change the cognitive processing of social stimuli. Happy people process information less critically than sad people (Forgas, 1992). Thus, a smile does not only mean "I am Happy," it simply influences the information processing in the receiver in favour of the sender (Grammer, 1995). This also explains why smiling is not necessarily bound to emotions (Kraut and Johnston, 1979). Smiles are more reliably associated with social motivations than with emotional experience. Comparable physiological changes occur with the perception of olfactory stimuli. A female pheromone, i.e. copulin, which is produced in the vaginal secretion influences male processing of female attractiveness. Under the influence of female copuline, males judge female attractiveness more positively (Grammer et al., 1996).

Although direct manipulation of the receiver through signals might play a critical role for communication, it does not yet explain how senders can hide their intentions. This goal can not be achieved simply by doing nothing, because receivers will become suspicious. Again there are many possible solutions to such a goal.

1.1.4. Being Honestly Dishonest. The first solution is the sending of meaningful signals either out of context or with different motivational background. This motivational background might contradict the actual goal of the sender. Grammer and Kruck (1996) showed that this solution is not uncommon. Females who are not interested in males do not send negative signals, they send a mixture of 60% positive and 40% negative signals. In this case the negative message is hidden in positive signals. In this case the deceptive attempt is the avoidance of face-loss by the male and thus potential aggression towards the female, which could occur when the male is bluntly rejected.

1.1.5. Multimodal Combinations: Signals from Different Channels and Metacommunication. A second solution to the problem of intention hiding lies in combining signals from different sensory channels. For instance in the case of laughter, body movements or postures, each with a determined meaning itself, can be added, but also laughter quality (i.e. amount of vocalisation, number of bouts etc.), odour and touching can co-occur. Such combinations have the potential to create an almost infinite number of meanings of laugh

ter, from sexual enticement to mobbing (Grammer, 1991). Furthermore metacommunication, becomes possible through multimodal combinations. Laughter for instance might put everything else what happened in a "play-mode" which simply says "Look it is not serious what I am doing" (vanHooff, 1972). By doing so, multi-meaning combinations are possible, which allow almost endless combinations for communicative purposes.

In the case of multimodal combinations, communicative channels are interacting. If this occurs, there has to be cross-modal integration of information. By comparing the impact of the different parts of the signal the receiver has to arrive at a decision what actually is meant. This problem is well known in research on the interaction of non-verbal behaviour and verbal utterances. Eyebrow raises and slight vertical or lateral rotations of the head appear to serve both punctuating speech rhythm and the emphasizing utterances. Although eyebrow raises are a clearly identifiable and a cross-culturally constant and discrete signal (Eibl-Eibesfeldt, 1989, Grammer et al., 1988), it can take many different meanings when associated with speech. Thus we would reach an almost infinite number of possible combinations. This is also the case for combinations with of eye-brow raising with other facial muscle movements. Grammer et al. (1988) showed that the resulting patterns were highly variable. This makes it unlikely to reencounter the same combination of stimuli again and these could make decoding more difficult. The same could also account for the high variance in human non-verbal behaviour.

1.1.6. Higher order Combinations.. Signals from the Same Channels. Higher order combinations or the summation of signal combinations over time could also veil the senders intentions. Combinations can occur when simple or multimodally combined signals are combined again either sequentially or at the same time. It becomes even more complex when a member of the combination can be substituted by another signal with the same meaning. As a result the same meaning can arise through different combinations, which are not identical. Moreover if in such combinations not only the presence of a signal plays a role, even the absence can be a signal itself. Moore's (1985) description of female solicitation may serve as illustration. She found 51 behavioural units in the validation of her repertoire and any combination of at least 10 units of female behaviour could predict male approaches. How could a male receiver have managed to "evaluate" or "recognise" female interest? Indeed there are $1.3 \cdot 10^8$ possible combinations of 10 behaviour elements out of a repertoire of 51.

1.1. 7. Manipulation of the Time Structure.. Good Vibrations and the Generation of Noise. There is mixed evidence on the assessment of function when using discrete behavioural categories, thus syntactical rules could provide some cue for the interpretation of signals independent of the signal content. With a mathematically sophisticated method, Grammer et al. (in prep.) analysed behaviour sequences in male-female interactions. They found a highly complex interweaving of behavioural elements. Pairs created dance like movement patterns. The drawback is that these patterns are highly idiosyncratic. Not one pattern occurred twice in almost 8000 patterns identified on the pair level. The temporal organisation of these patterns varies with interest: they become more stereotypical when interest in the partner is high. Even more, the female initiates the patterns and the more patterns are present the better the male feels himself in the interaction. This suggests that temporal organisation itself could predict male-female-interest. Manipulation of time structure and its perception can also be reached by trying to achieve synchronisation. Although most attempts to empirically describe synchronisation empirically have not been successful. In contrast to the impossibility of description subjects can rate the degree of

synchronisation in interactions, and this ratings correspond to the subjective experience of interpersonal rapport among the interactants (Bernieri and Rosenthal, 1995).

Another solution to the hiding of intentions problem is the creation of "noise." That means sending many signals without attaching meaning to it, and hiding the meaningful signals in this "noise." The noise will make it almost impossible for the receiver to decode the real intentions of the sender. In this context, the concept of "protean behaviour" (Chance and Russell, 1959) could play a role. The concept proposes that behaving unpredictably and erratically will mask the signal sender's intentions like a prey trying to evade its predator. In the study cited above, the repeated and time constant patterns of behaviour are hidden in a continuous flow behaviour. A closer analysis of the data reveals that only the precise timing and not the content of the performed behaviour plays a role for the generation mutual understanding.

Another alternative is that rare events which are hidden in "noise" advertise intentions at "hot times" at "hot-spots." This means that one single signal with a distinctive and explicit meaning sent at the right time could signal the intentions of the sender. Thus one or two events which might differ from interaction to interaction can be enough for communicating intentions. This alternative is highly likely because individuals communicate stimulus information in a way which is adaptive for perceivers to detect, and perceivers detect this information, when they are attuned to it (MacArthur and Baron, 1983).

Mutual understanding can also be reached by sending information in a way the receiver can not consciously assess. This means sending unknown signals to the receiver or sending variants of signals the receiver is not likely to interpret. Thus, the receiver is forced to learn the shared code slowly. This hazard would lead to a highly variant shared code. The emphasis lies on the term "learning." The sender has to prepare the receiver slowly for receiving his actual intentions.

4 The Quality of Movements: A Neglected Dimension

Slow escalation and hiding intentions when at risk of failure on the sender's side is paired with a cognitive apparatus which tries to unravel the intentions on the receivers side. This situation will force communication to a level where the receiver might not be able to assess consciously the manipulation which is underway. The receiver then has to look for honest and develop strategies for "mindreading."

This situation is contrasted with a scientific research apparatus which seems not at all adequate for the analysis of such a paradoxical situation. It is obvious for any observer that behaviour is distinctive and that there are many levels and methods of description for signals. Categories can embody muscle movements or groups of such movements like in Ekman's and Friesen's Facial Action Coding system (Ekman and Friesen, 1978) or descriptions of limb movements like in the Berner System developed by Frey and Pool (1976) or even more complex units like walking. Muscle movement description is the most basic level of description. on the next level behaviour is already described by interpretative categories, even when the definition is highly operationalized. This interpretation sometimes can involve hypotheses on the function of a behaviour. The term "coy smile" is already a hypothesis on the function of behaviour although it describes a distinct motor pattern of head movements combined with a smile (Eibl-Eibesfeldt, 1989).

Common to all these approaches is that a continuous behaviour stream is forced into a series of event categories which might subsume comparable, but visually distinct behaviours. For example a non-verbal threat can be done in many ways. By raising an arm fast or slow, with fist clenched or not, the movement staying at the maximum flexion for a cer

tain time and going back fast or slowly. Any of the possible combinations will produce a different type of "threat." We also can transfer the movement combination itself to a leg or even a head movement: moving the head fast towards somebody else then staring at him and finally look slowly away. In every case we will produce an event of "threat" by using a certain movement configuration. A solution would be to describe all these different types of threat, with the result that the numbers of each type of threat events will become small and useless for statistical treatment. An additional problem at this level is the reliability of observations: in order to be identified reliably and to avoid the development of an observer bias, categories have to be unmistakable, stereotypical, homogeneous and discrete. This can lead to oversimplification and broad categories.

Thus a dilemma arises: observers surely can interpret and understand the behaviour of others and their possible intentions correctly but we might not be able to identify them on the bases of communication theories which use discrete signals.

The solution we propose to this dilemma is that the categories which are used to break up the flow of behaviour are only a poor approximation to how the receiver processes information. Any behaviour is a change in a continuous information flow which could be noticed by a receiver. These changes in the information flow can be of various qualities. It is possible to encode information in the quality of a change. We propose that body movements and parts of body movements themselves are processed, i.e. the receiver does not perceive and summate behaviour in single categories "Legs open" or "Hair Flips." In contrast the receiver could assess elementary dimensions of behaviour like speed, acceleration and amount of movement or motion quality. The brain might not use categorical perception in the same way as we classically analyse behaviour. An example for different possible meanings that arise through speed differences in onsets of the events was given by Grammer *et al.* (1988) for "eye-brow-flashes" in a cross-cultural analysis. In the most common event the pattern starts with the contraction of the *M. corrugator supercilii*. This contraction disappears and in a fast movement the brows are lifted and a smile appears. The duration of the contraction of the *M. frontalis* and the *Pars palpebralis* which causes the eyebrow lift is variable and disappears slowly, whereas the smile caused by a contraction of the *M. zygomaticus major* stays on the face. The second pattern also starts with a contraction of the *M. corrugator supercilii*. But then, there is a slow lift of the brows, and the contraction of the *M. corrugator supercilii* does not disappear. Rarely a smile is added (See figure 3 for details). In this case, meaning arises through the combination of elements present and the dynamic properties of the elements.

The solution to the dilemma between event oriented categorisation of behaviour and possible qualitative differences in the same behaviour with different meaning can be found in the way visual information is processed. The information is not only reduced, but parallel new information is created. Processing of visual signals takes place on two levels: one is low level processing where the perceived information is recoded: colour, motion, depth, time integration of movements. on this level information is detected which seems to be necessary for spatial navigation. During high level processing, where pattern recognition occurs, a "world-model" or a priori knowledge like form, size, schemes is added. on this level, the brain compares the results it has got so far and then tries to come to a coherent interpretation of the world (Arbib and Hansen, 1987). Traditional behaviour research is working on the second level and neglected the first level which basically describes the quality and not the content of behaviour.

In order to assess qualitative aspects of body movement Johansson (1973, 1976) used a point light display fixed to the joints of his subjects and filmed them in the dark. The resulting films appeared as a configuration of bright points against a dark background. If such point-light clips are shown to raters, they can recognise sex and age of the

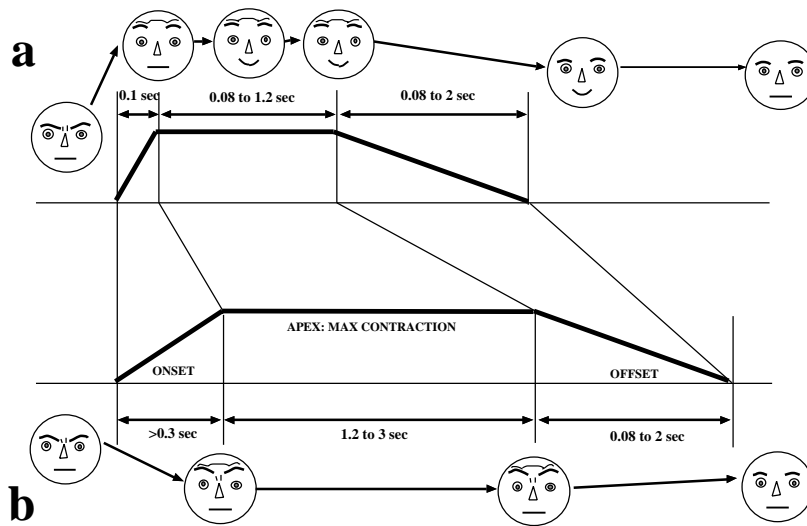


Figure 3. Time structure and patterns in the eye-brow-flash. The figure shows two different patterns in an eyebrow-flash. The prototypical pattern in (a) was described by Eibl-Eibesfeldt (1972). This pattern starts with a frown which disappears. The brows are lifted quickly and a smile is added. The brow raise disappears, while the smile can stay on the face. The second pattern (b) is completely different: It also starts with a frown, which does not disappear whilst the brow raise appears on the face. The brow raise onset duration is three times as long as in the first pattern and the patterns duration is much longer than in (a) (Grammer et al., 1989). In addition to the difference in combinations of muscle movements there is also a difference in the time structure which changes the quality of the expression.

subject (Cutting and Proffitt, 1981). This effect has been replicated several times (e.g. Runeson and Frykholm, 1983). If the point light displays are presented statically they are rated as a random display of points. observers are able to isolate abstract patterns of movements from such displays. Moreover, observers are able to detect effort, intentions and deception (Runeson and Frykholm, 1983), and if the point light displays are applied in faces, perceivers are able to detect information connected to emotion (Bassili, 1979). Gait and movement analysis has a considerable tradition in medicine where it is used to describe movement disorders. It has rarely been used in behaviour research.

Motion quality is a powerful descriptor which is used for signal decoding by humans. Yet there is a serious methodological drawback: point light displays can not be used in unstaged interactions, because they are an obvious research device. They make any stimulus person rather self conscious and alerts them to the variables of interest to the researcher (Berry et al., 1991). We thus developed a digital image analysis system which can be used to decode automatically speed, acceleration and size of moving body parts from digitised video images.

2. AUTOMATIC MOVIE ANALYSIS (AMA) AND MOTION-ENERGY-DETECTION (MED)

Machine understanding of human action has become a fascinating topic in computer sciences through the last years. Unfortunately the researchers have mostly avoided dealing with people in their natural environments (Pentland, 1995). In contrast to this, several re-

search groups have developed devices which are able to accomplish the basic task of making computers recognise who they are working with and to be sensitive to people's gestures and expressions. Computers who are able to track model human actions can be used for behaviour analysis and not only as new input devices. They can recognise faces (Moghaddam and Pentland, 1995), and facial expressions (Essa and Pentland, 1995). Unfortunately the approaches are made in a way that the dynamical component is lost and expressions are forced again in event categories, like surprise, happiness, disgust or anger with all the drawbacks of categorisation.

In a more general approach, Starner and Pentland (1995) and Pentland and Liu (1995) have developed a motion analysis system where the human body is modeled as a Markov device with a number of internal mental states, each with its own particular behaviour, and inter-state transition probabilities. Internal states are determined through an indirect estimation process, using the person's movement and vocalizations as measurements. These variables then are fed directly into a computer which then can observe a person's actions and respond accurately. Niyogi and Adelson (1995) developed a set of techniques which are capable of analysing the patterns which are generated by walking. These patterns are then translated into a stick-figure and the gait can be analyzed.

Unfortunately these devices are quite sophisticated and expensive and they have not been used for behaviour research. In order to assess the quality of movements and human expression we developed a system which can be used even in labs with a low budget. We dispensed with real time operation capabilities and turned to the analysis of digitised video. These procedure is called Automatic Movie Analysis (AMA). The advantages are clear: it is possible to repeat any type of analysis and control for artifacts.

2.1. The Programming Platform

Automatic Movie Analysis is a programming platform which can apply a row of sophisticated filters to videoimages sequentially. In this article, we will restrict our analysis on Motion Energy Detection (MED). This is a simple but elegant filtering method for the determination of qualitative aspects of behaviour. The method relies on the fact that pictures of movies (frames) are time dependent distributions of grey-scales or colour values. A video-picture consists of an array of pixels which take different values. In a grey-scale these values usually range from 0 (white) to 255 (black). If the camera view is static, single pixels will change their values when movement occurs. If there is no movement, the pixels stay at the same value. The solution is to make the difference between two pictures. If there was no movement, we will get an all white picture, if there is movement we will get grey values in those regions where movement occurred. The amount of movement will become visible through the amount of pixels which are not white.

This image differencing (Sonka *et al.*, 1993) works quite well if the camera position and the lighting conditions remain unchanged. The method can detect movements, not their direction. In this case AMA consists of the following steps (illustrated in Fig. 4):

1. *Digitising* of a movie in a range between 12.5 to 25 pictures a second (320 x 240 pixels frame size) in greyscales (Pixelvalue=[0..256] greys, where 0 is white and 256 is black).
2. *Making difference pictures* The arithmetic difference is calculated between two or more consecutive images of the movie (MED).
3. *Digital noise-reduction* Videonnoise is related to tape quality, camera resolution and light conditions. In a video picture, pixels will change their colour ran

domly. As a result "lonely pixels" will appear in the difference picture. The noise was reduced by the application of a median filter to the difference pictures, where a pixel received the median grey value of its nine surrounding neighbours.

4. *Error detection* Flashes result from poor videotape quality or technical problems of the recording equipment. If such flashes occur, there is an immediate high change in mean grey density in the overall picture. If this occurred the grey density for this frame (t) was replaced by the difference density between points $t-1$ and $t+1$.

5. *Calculate mean grey values* for one or more predefined regions of the picture. These regions have to be determined by hand. In a newer version of the program one or more persons can be separated automatically from the background.

6. *Standardising of mean grey values* Different viewing angles of the person and changes of body postures during the experiment lead to different overall values of mean grey density, because visible contours change. Thus the mean grey values were transformed to z-scores.

7. *Smoothing* At this stage, noise was still present and resulted in small but very fast and short changes. The scores then were smoothed with a 5-point moving average.

8. *Thresholding* Further analysis is possible if the continuous recordings are collapsed into events. A threshold method was used. When the grey level change was under a certain threshold, the grey level change was set to zero. The threshold was determined with an optimal thresholding method (Sonka et al., 1993).

Figure 4a shows a videosequence of a female "Hair-Flip" and a male watching her. In Figure 4b all difference pictures are shown. Figure 4c shows the transformed z-scores of the male and female from Fig. 4a. The male shows no movement, three movement clusters were identified (Burst 1, 2, 3). First a hand movement together with a head toss (Burst1, frame (17). Immediately after this hand movement the hair flip starts (Burst 2, frames 19 - 39). Finally, the female turns the head towards the male (Burst 3, frames 58-65).

Thus we have the following descriptors for movement quality: Number of bursts, their duration and the size of an event (the area enveloped by the burst). The number of elements in an event was calculated by counting the number of maxima and minima in the burst. A burst can have several elements which are produced by combining different movements. The number of elements describes the complexity of the movement. Finally we were able to record speed of movement change, which is the size of the burst divided by its duration.

2.2. Interactions between Strangers and Subliminal Manipulation

This procedure was applied by Grammer et al. (1996) to social interactions between strangers in a waiting room situation. In an experiment, strangers of both sexes met and interacted for 10 minutes while they were videotaped with a hidden camera. The experiment was made in Japan and in Germany. This situation can be described in terms of high risk of social non-acceptance; thus communication should be forced into a manipulative level. After the experiment, the subjects made a self report on their interest in their partner and how pleasant they found the interaction.

The first two minutes of the video clips were analysed first with traditional methods and then with AMA. On the one hand, traditional analyses yielded cultural differences be

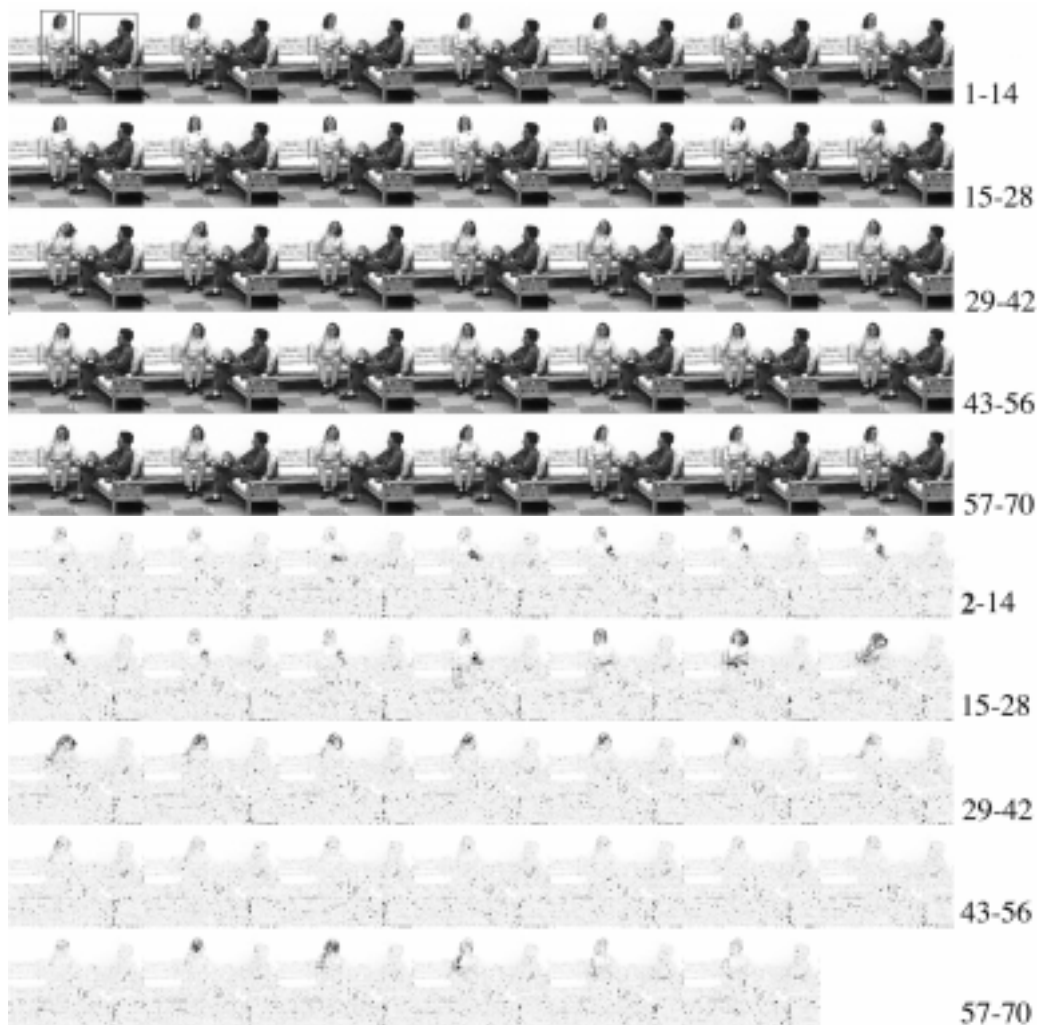
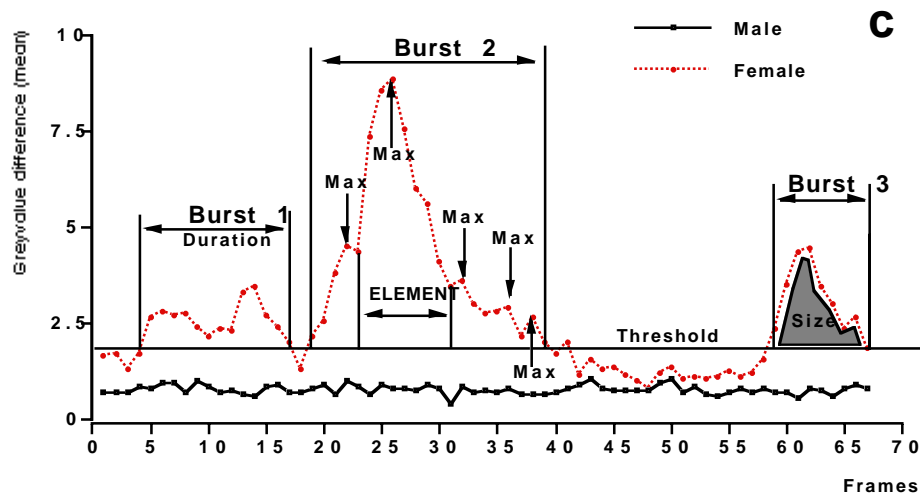


Figure 4. A female Hairflip analyzed with AMA. Instance of a female Hairflip preceded by hand movement. In (a), the original film sequence is shown (70 frames, every second frame skipped). The rectangles in frame 1 delimit the areas of interest for which difference-pictures were calculated (b). Note the changes in greydensity (e.g. head movement). In (c), terminology and the movement descriptors resulting from the automatic movie analysis (AMA) are shown: Number of bursts (total amount of movement), and duration, size and speed of bursts (size divided by duration, i.e. greydensity change per s). The number of elements (number of maxima) within a burst is considered to be a measure of complexity since elements normally are produced by different movements. The three movement bursts of the woman identified by the threshold method are in good agreement with the film sequence. The first burst depicts a hand movement accompanied by a head toss (frames ~17), whereas the immediately following burst represents the hair flipping (frames 19-39). During the third burst, after a pause of about 1.5 s, she turns her head towards him (frames 58-67).

tween Japan and Germany. A typical Japanese behaviour is "nodding" which rarely occurs with comparable high frequencies in Germany. Moreover, when the frequencies of the generic behaviour codes were compared to male and female interest, no significant correlations were found. On the other hand, AMA reached similar results in Japan and in Germany. Females changed the quality of their behaviour when they had high interest in the male. These qualitative changes were not due to mere nervousness or excitement. The



females actually moved more, but showed smaller and slower movements. These qualitative changes give an impression of slow and determined movements where the single parts were accentuated. Males reacted to these qualitative modifications positively and experience the situation more pleasant although their interest in the partner is not affected. In addition, males who perceive the situation positively talk more. Thus we can conclude that it is not the content of a non-verbal behaviour, it is the quality of the movement which actually holds the information about the interest of the sender.

one objection could be that the movements are generated by speech and lively conversation between the interactants. This was not confirmed. The amount of speech did not correlate with qualitative changes in the females behaviour nor with female interest. When the correlations between movement data and interest were corrected for speech, it became clear that non-verbal behaviour is the means of communication in this situation. Thus in real life situations where risk is present, non-verbal behaviour plays the main role in communication. The evolutionary theory behind this explanation is the fact that females actually have a greater risk in male-female interactions than males (Trivers, 1972). But it is not only the risk of losing investment: actually the risk of being deceived by a male is quite high. In a questionnaire study by Tooke and Camire (1991) 60% percent of the males reported that they had used deception in such interactions. Thus it seems logical that females would try to manipulate the males slowly without revealing their intentions in order to gather information about the male's behaviour tendencies. This is possible when the male feels pleasant and starts to talk and reveal information.

A second even more interesting result is the fact that behaviour is not a simple continuous flow of movements. It is definitely structured into single bouts (see Figure 5). Times of movement and non-movement alternate. Thus for a closer analysis we will look at the quality of the bouts themselves, and take a look which information qualitative changes might provide.

2.3. Showing off: Simple Movements and Their Communicative Value

In order to accomplish this task, video-material from an observational study on female cycle and self-presentation (Grammer et al., 1996) was used. The starting point of

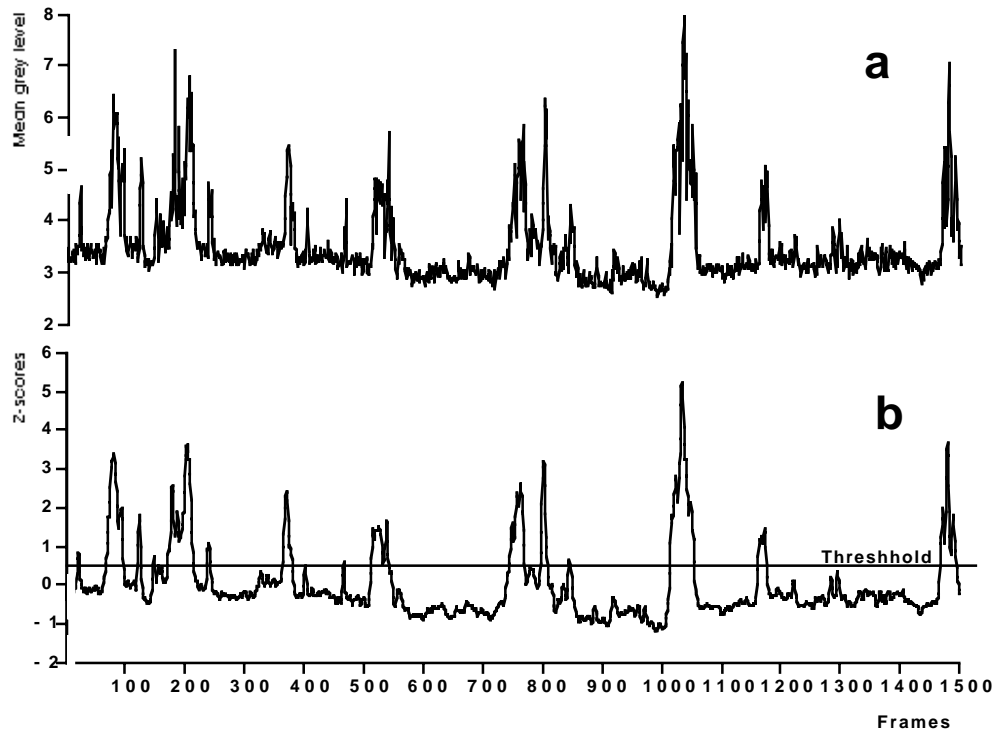


Figure 5. Movements from a two-minute interaction. Greydensity changes of a person throughout a whole film sequence of two minutes. In (a), greydensities obtained by averaging all greyvalue differences of a difference-picture are plotted (end of step S, cf. text for details). In (b), the data are z-transformed (step 6), smoothed (step 7), and a threshold has been calculated (step 8). AMA identified IS bursts.

this study was the fact that in humans we find almost complete female sexual cripsis. This means that in contrast to most non-human primates, humans show no obvious behavioural or visual sign for ovulation. There has been a lot of speculation about the function of female sexual cripsis and it is proposed that it leads to male-female bonding, that it promotes active female choice or that it may give a chance to females to induce sperm-competition (Alexander and Noonan, 1979, Benshoof and Thornhill, 1979, Baker and Bellis, 1995). By hiding the point where conception is quite likely, females would force the male to stay near them to ensure conception. The male would do so to be sure that the female conceives his child and not that of another male. The fact that the male has to stay near the female could promote the emergence of a male-female relationship and lead to male investment in the offspring. The induction of sperm-competition and geneshopping only works if females use extra-pair copulations at the time of maximal likelihood of conception. This is actually the case and leads to relative paternity security of about 90% (Baker and Bellis, 1995) The function of sperm-competition is thought to be either the optimisation of offspring quality and, if a relationship already exists, to obtain genetically variable offspring. Grammer et al. (1996) showed that females who had a steady partner and who came alone to an anonymous discotheque, showed more skin and wore tighter clothes when oestrogen levels were high. Moreover, these females knew consciously that they were signaling availability through clothingstyle.

Non-verbal behaviour can be used strategically when somebody tries to impress another person. In the context of self-presentation non-verbal behaviour is irrepressible and impactful. It is off-the-record and it is difficult to describe verbally (DePaulo, 1992). Many studies on self-presentation have shown that people can successfully make clear to others, using non-verbal cues, the internal state that they are actually experiencing and that they also can convey the impression of a state that they are not really experiencing. In this process many non-verbal cues can play a role, even body movements, postures (Cunningham, 1977) and gait (Montepare et al., 1987) can be used as signals. According to DePaulo (1992) women are more concerned with self-presentation than men, they are non-verbally more involved and spontaneously more expressive.

out of these results we can hypothesise that the quality of behaviour may also change in order to signal sexual availability. The reason for unobtrusive signaling in this case would be enforced by the fact that active female choice has to exist for the induction of sperm-competition. Thus risk is also present in such a situation because signaling sexual availability will obviously attract all males. The risk is that one might attract the wrong males thus the signaling level should be subtle and not exaggerated.

2.3.1. Turning Around – An instance of Self-Presentation. During this study 123 females (mean age 23) were filmed from front and back in order to determine the clothing style. The females had to turn around on a command issued by a female experimenter standing behind the camera. This turn-around movement has no communicative value on the first sight. In this situation however we had onlookers (male and female) and a videocamera. According to Goffman (1959) self-presentation will take place as soon as a stage for the presentation is present. In this type of self-presentation public self-consciousness plays also a role. This is defined as "an awareness of and a responsibility to the impressions that are being made on others" (Scheier and Carver, 1981, p.198). We thus assumed that the turning-around movements can be used as communicative tools. At the side of the subject, a male or female was placed for holding a blackboard with an identification number on it. The female subjects all were non-pilltakers, who either were singles or who had a steady partner but came alone to the discotheque. The result is a sex of experimenter (male/female) times type of female (single/paired but alone) design. Oestrogen levels which can give a hint on cycle state were determined by enzyme immuno assay from saliva. In addition skin exposure and tightness were analysed by Grammer et al. (1996) with digital methods and a rating scheme for tightness.

The turning movement was digitised in a 320x240 pixels size at 25 frames a second. Digitisation started with the beginning of the turning movement and ended when the turning movement came to a halt. These digital videos were analyzed with Automatic movie analysis, Motion energy detection was applied for each subject on the screen, the experimenter and the female separately. Duration of the movement, the number of bursts in the movement (the basic units of the movement), complexity (the number of single maxima), the maximum speed of the movement, and the information content (amount of grey value change per frame) were recorded (see figure 6). The resulting movement graphs then were played back parallel to the movies. This procedure allowed a frame exact determination of the beginnings and the endings of the movements by comparing the movement graph to the videopictures. The play back also showed that the movements could be divided into several parts (see figure 6 and table 1). Basically a turn-around can start with an intention movement prior to the actual turning. We find different types of movement initiation. The head can introduce the movement as can one foot, one knee or a turn of the upper body. Then the turning around itself follows. When the turning ends, the standing position can

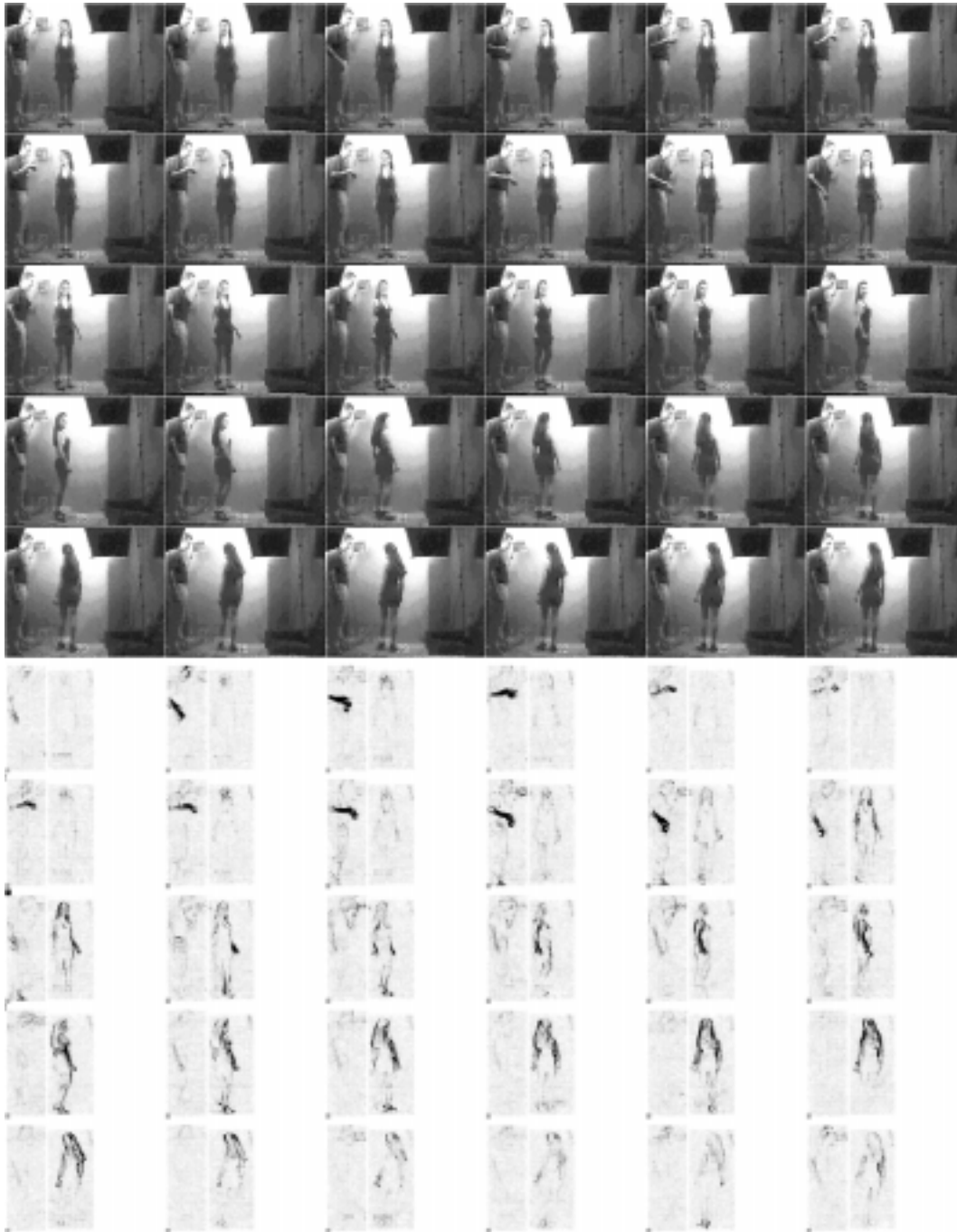
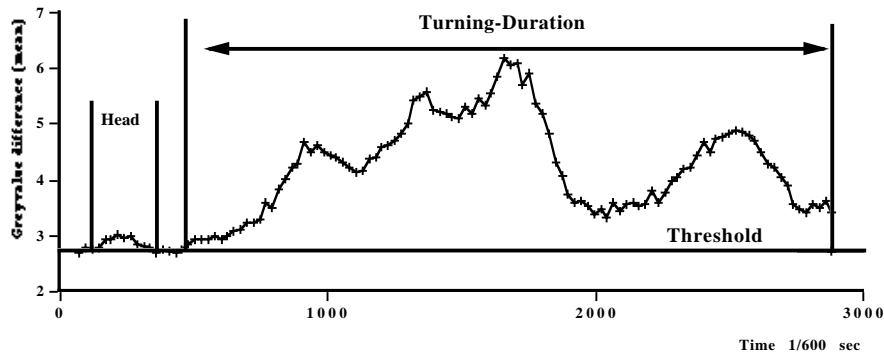


Figure 6. Female reaction to a command: Turning-around movements. Sequence (a) shows the digitized pictures for a reaction to a command. The female command giver is standing at the camera. on the command "Please turn around" the female turns around. On the left side of each picture the "stimulus male" can be seen. Sequence (b) shows the difference pictures with the regions for the stimulus and the subject which are used to calculate the mean grey density. In (c) these grey density levels are presented as z-scores. The movement has three phases: phase I where an initiation movement is made. Phase II where the actual turn around movement happens and phase III where an additional movement is added to the turn around (see also text). The description parameters for the movement are the same as in Fig.4c. The difference to Figure 4 is that the threshold is calculated dynamically for each 10 frames in order to deal with shifts of grey density changes.



be corrected or additional movements can be added. These movements may consist of hip swaying, hair-flip arm swaying or moving the body from the left to the right. Fig. 6 shows a breast-presentation movement, where the upper body turns over after the stop of the actual turning movement and the body silhouette becomes visible from the side. In addition a typical "breast out-shoulder-back movement" occurs. These four types form only the main classes, the additional movements can be done by any movable body part. The movements also were coded "traditionally" with the observation categories presented in Table 1.

Table 1. Behaviour occurring during turn-around movements

Movement phase	Behaviour type
Phase I: Intention movements	Head down
	Shoulders up
	Lift arms
	Turn head
	Others
Start of movement	Left/Right foot
	Upper body
	Arms
Phase II: Turn around	Number of steps needed
Phase III: Additional final movement	Position correction
	Hair Flip
	Hip Sway
	Move Shoulders
	Arms sway
	Arms lift
	Sway trunk
	Side presentation
	Head turn
	Head tilt
	Others

Table 2. Differences in movement quality between the experimental groups

Group	N	Duration		Bursts		Complexity		Information		Maximal speed	
		mean	std	mean	std	mean	std	mean	std	mean	std
Paired-FS	20	32.8	19.0	3.4	2.3	14.1	11.6	47.9	39.4	2.7	1.0
Paired-MS	14	19.3	6.4	1.5	0.7	6.3	3.2	29.5	8.7	3.3	1.1
Single-FS	57	30.1	17.9	2.8	2.3	12.5	10.7	37.0	14.9	2.8	1.1
Single-MS	29	28.0	15.0	2.5	2.1	12.4	10.1	31.9	10.4	2.8	1.2
K-W Anova		p=0.11		p=0.10		p=0.12		p=0.02		p=0.33	
Total	120	26.8	16.8	2.74	2.3	12.1	10.31	36.7	20.5	2.9	1.1

(FS: female as stimulus, MS: male as stimulus, Paired: female having partner but coming alone, Single: female with no partner)

2.3.2. Movement Quality and oestrogen Levels. In a first step a conventional statistical analysis was performed in order to see if the four experimental groups are qualitatively different. Table 2 shows the results which actually propose that the four groups can be differentiated on the basis of partner status and stimulus person in only respect to information content of the movement. The highest amount of information is present when the female is confronted with a female stimulus. on the first glance this contradicts the showing off hypothesis.

If we go one step further, we find a high interrelation between the recorded movement parameters themselves. Duration, the number of bursts, complexity and information are positively correlated in all groups ($n=120$, $r_s=0.65$ to 0.84) and correlate negatively with speed (-0.45 to -0.87). That is short movements have fewer bursts, are less complex and have less information, and their speed is high. These interrelations suggest that there are physical constraints on movements.

In a next step, we correlated oestrogen levels with the movement parameters. The Bonferroni corrected correlations show that only females confronted with a male stimulus change their behaviour quality. Single females make slower and more complex movements. Both single and paired females show more information per time in their movements with increasing oestrogen levels (Table 3).

2.3.3. Movement Quality and Stimulus Reaction. These results suggest that single females react toward a male stimulus by changing the quality of their behaviour. But how does the stimulus male perceive it? When we correlate the stimulus male's behaviour quality with those of the subjects, we do not find any significant correlation. Thus it seems that males do not notice the female's qualitative behaviour changes directly. It is also possible to directly test if the male reacts to female oestrogen levels. The stimulus male's behaviour depends on the paired female's oestrogen levels. He changes the duration of his

Table 3. Female movement quality and oestrogen levels

Group	N	Duration		Bursts		Complexity		Information		Maximal speed	
		r_s	p	r_s	p	r_s	p	r_s	p	r_s	p
Paired-FS	20	-0.14	n.s.	-0.19	n.s.	-0.19	n.s.	-0.13	n.s.	0.25	n.s.
Paired-MS	14	0.27	n.s.	0.34	n.s.	0.19	n.s.	0.61	0.05	0.04	n.s.
Single-FS	57	0.05	n.s.	0.04	n.s.	0.01	n.s.	0.14	n.s.	-0.09	n.s.
Single-MS	29	0.21	n.s.	0.33	n.s.	0.52	0.02	0.52	0.02	-0.45	0.02

(FS: female as stimulus, MS: male as stimulus, Paired: female having partner but coming alone, Single: female with no partner)

movements ($n=14$, $r_s=0.68$) makes more bursts ($r_s=0.66$) and makes more complex movements ($r_s=0.57$). The stimulus female does not show any significant reaction to other females oestrogen levels.

But as we know, there is no direct coupling between the male's and the female's qualitative behaviour changes, i.e. stimulus males' movement quality does not change parallel to subjects qualitative change. -Maybe the males use other sources for information, like skin showing or tightness of clothes.

If this is the case, the reaction of the stimulus males might not be due to qualitative changes in female behaviour at all. The reaction of the males could be a reaction to the exalted sexual signaling through clothing style. Thus we controlled the correlations between stimulus behaviour and oestrogen levels for skin with partial correlations. The correlations for the male stimulus reactions to unpaired females oestrogen levels do not disappear ($df=11$, duration: r_s partialized=0.64, $p=0.02$, complexity: 0.62, $p=0.02$).

So far the results suggest that the behavioural changes are actually changes which occur together with high oestrogen levels. We may conclude that females who develop interest in a male signal high oestrogen levels. If this assumption is true, then we should expect that these changes are present in .all females with high oestrogen levels, and that it is impossible to suppress these changes completely. Nevertheless, under the right stimulus conditions, females could either fake or superelevate them suggesting a cognitive accessibility.

Females with higher oestrogen levels show higher information content in their movements when they are confronted with the stimulus male but only in the case of paired females the male reacts. We have found the highest values of information content when a female stimulus is present but there is no difference between paired and single females when a male stimulus is present (See Table 2).

If we look back at the considerations about possible communicative mechanisms, we find many possible solutions for this. The most obvious one is that the simple physical measures we used for movement description are not adequate, or there is a summation of different features over time. So far we can exclude at least multimodal communication where skin showing and movement quality add up.

2.3.4. A Neural Network Approach for the Analysis of Movement Quality: Parallel Distributed Processing. In recent years, connectionism has become a focus of research in a number of disciplines. Neural networks represent a special kind of information processing: connectionist systems simulated by a computer consist of many primitive cells which are working in parallel and are connected via directed links. This forms an analogy to the human brain: the cells are analogous to neurons and the links are the connections between those neurons. The main processing principle of these cells is the distribution of activation patterns across the links similar to the basic mechanisms of the brain. Information processing in the brain is based on the transfer of activation from one group of neurons to others through synapses. In analogy to activation passing in biological neurons each unit receives a net input that is computed from the weighted output of prior units with connections leading to this unit. However, the most current neural networks do not try to closely imitate biological reality.

In neural networks "knowledge" is distributed through the activation of cells and the weighting of the links. The networks are organised by training. In supervised training, the network "learns" a set of patterns together with their classification by repeated presentation. Through this "learning process," classical logical conclusions are replaced by vague and associative recalls. This is of advantage in all cases where no set of clear logical rules

Table 4. Network classification of movements according to oestrogen levels

Pattern classified as	Original oestrogen level associated with pattern		
	High	Middle	Low
High	25	0	1
Middle	0	16	16
Low	0	7	11
None	0	3	0
% correct	100%	61.5	57.1

can be given. After learning, the neural network can be able to classify unlearned patterns correctly or not. In the first case we then can assume that in the patterns is at least some information present which is common to certain classes. Unfortunately it is very difficult to recall the information the network has used for classification.

Neural networks thus can be used to look if information is present in a pattern which then can be used to classify these patterns. A neural network analysis was applied to the raw data from the "showing-off" study above. The network was constructed as a time delayed network (TDNN, Waibel, 1989) on the SNNS-Simulator (Zell, 1994). The network embodied 10 (features) times 24 (total delay length) input units, 120 hidden units (receptive field) and 3 output units for low, middle and high oestrogen levels. Time delayed networks do not use a static presentation of patterns and they can be used for the independent recognition of features within a larger pattern. The update algorithm forces the network to train on time/position independent detection of sub-patterns. However, there is no specific set of rules on how to construct a network, and building networks heavily relies on trial and error. Thus, the fact that it is not possible to train a network does not mean that there is no information to learn.

We applied two basic training methods. First the network was trained with data from single females with male stimulus. The validation was done with single females with female stimulus. The data from paired females with male and female stimulus were then tested for classification analysis. Second, the training was done with data from paired females with male stimulus, the validation was done with paired females with female stimulus, and the testing with single females with male or female stimulus.

The classification results showed astonishing stability: 66% of cases from method one and 70% of the cases from method Two were classified correctly. A closer look reveals that the wrong classifications were due to the fact that only low and middle oestrogen levels were classified sometimes incorrectly as either low or middle but never as high. With both methods combined high oestrogen levels were classified 100% correct. Thus the TDNN was able to discriminate between high and middle/low oestrogen levels correctly (See Table 4) using MED data from videopictures processed through AMA.

In order to isolate the movement prototypes, we calculated the mean movement curves for the three classes of oestrogen levels. Figure 7 shows the results. The three curves for high, middle and low oestrogen look different but when tested the only significant difference is in the information content. Lowest oestrogen shows lowest information content in movement (Median:34). Middle oestrogen shows middle information (Median:36) and highest oestrogen level shows the highest information content (Median:39, K-W 1-Way Anova, $p=0.118$). This result finally brings us back to coding with discrete categories.

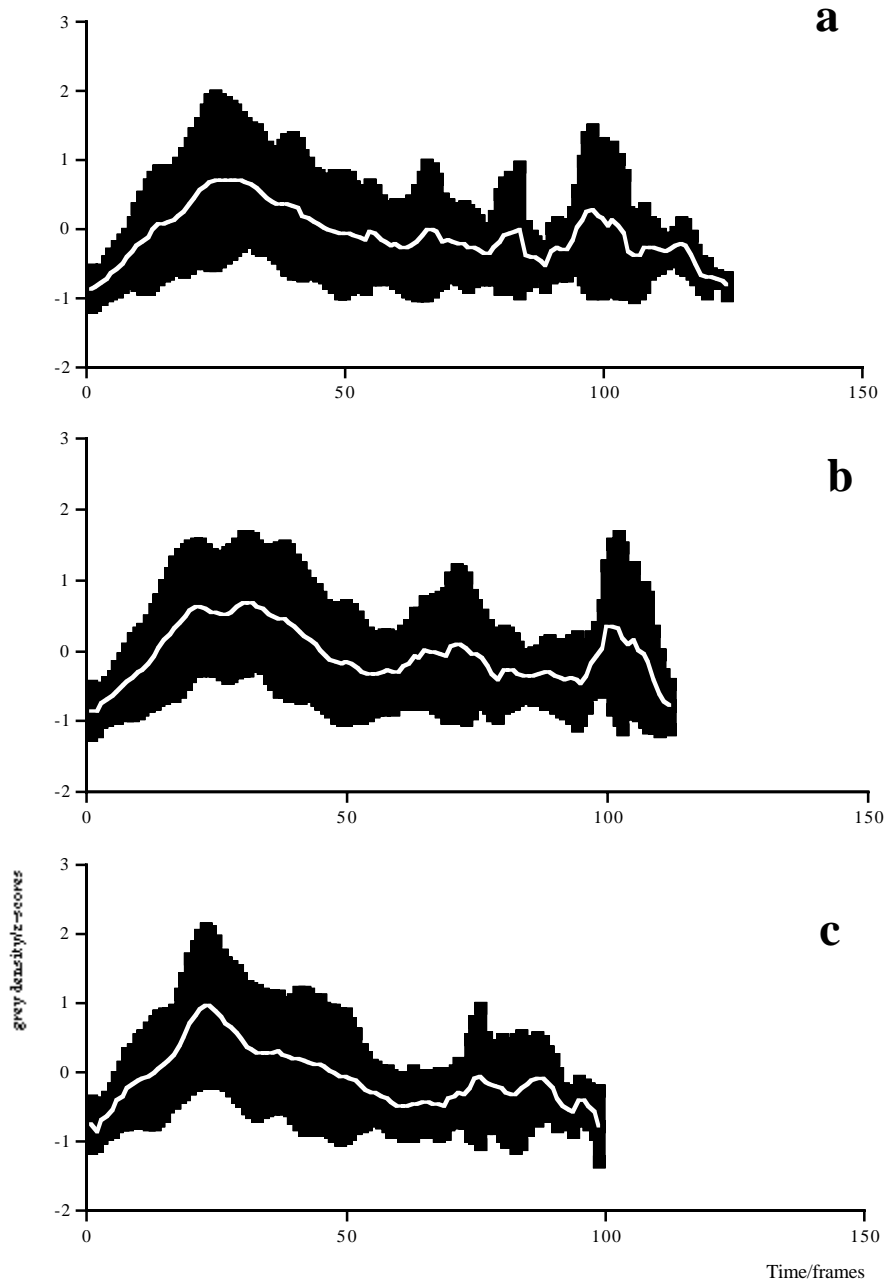


Figure 7. Movement prototypes for oestrogene levels in females. This figure shows the mean movement curves for the three oestrogen levels: (high (a), middle (b) and low (c)). The black bars indicate the standard deviation for each frame. The three movement phases and the apex (See Table I) are indicated by dashed lines.

2.3.5. *Discrete Coding of Movement Patterns.* In order to find out if certain discrete patterns as described in Table 1 are connected to oestrogen levels a traditional coding was applied to the digitised videos. It turned out that in none of the three phases discrete codes could separate between oestrogen levels. The exception was an additional movement in phase three. The presence of one or more additional movements occurred significantly more often under high oestrogen levels. This relation was independent from the content of the behaviour (Median-test, $p=0.02$).

So far we can conclude that it is possible to describe intentions in communicative acts with the help of qualitative changes in movements. Yet it is still unclear which changes are present, because MED only crudely describes qualitative changes on a holistic level. Single movement features are not captured by this method. Yet our hypothesis is confirmed that under high risk conditions, communicative acts are forced to a level where it is only difficult to assess them with generic coding methods. This situation has led us to a series of new developments which we currently pursue.

3. FUTURE DEVELOPMENTS AND ANALYSIS

The description of the whole human body and its moving parts seems to be an unsolvable endeavour. In recent years however, digital analysis of human movements and bodies has moved far away from simple MED approaches. Basically, all methods which have been used up to now are derivatives from two approaches. Either the contours of moving or non-moving objects are separated from a background, or the displacement of pixels or groups of pixels are calculated as optical flow (Sonka *et al*, 1993). For instance, when we want to look at emotions, a method for surface analysis of the face has to be developed in contrast to a three dimensional tracking method for an arm. The assessment of movements will differ from the assessment of postures and a method to translate the static states will also be necessary. Basically a body has to be separated from its background and then divided into its segments. This means that the body has to be dissolved into head, face, arms, body and legs. Each of these parts can then be described separately. Interestingly enough, there are many approaches to solve the task of body movement tracking. The isolation of body parts including head and face does not pose a problem. This task has been solved repeatedly (Pentland, 1995). Kakadiaris, Metaxas and Bajcsy (1994) for instance proposed an integrated approach to segmentation, shape and motion estimation of complex articulated objects which can also be used for human bodies.

The best results in the human body tracking and action recognition are achieved at MIT Media Lab (Maes *et al.*, 1995). The ALIVE "Artificial Life Interactive Video Interface" allows wireless full-body interaction between the human participants and a rich graphical world inhabited by autonomous agents. Agents are modelled as autonomous behaving entities that have their own sensors and goals and that can interpret the actions of the human and react to them in "interactive time." Vision routines compute figure/ground segmentation and analyse the user's silhouette to determine the location of the head, hands, and other parts of the body in a colour image. This self-calibrating stereo person tracker can recover the 3D shape and motion of the hands and head of the moving person (Pentland, 1995).

our new developments work on a model base. The human body can be modelled as a system of objects connected together by joints with one or more degrees of freedom. Tracking motion of human body can be formulated as the real-time visual tracking of kinematic chains. A kinematic object is a collection of objects connected by joints. With

each object a local co-ordinate system can be associated to specify its 3D position and orientation. Since the objects are connected instead of using a six dimensional vector for each object to describe its 3D position and orientation, the joint parameters can be used to define the mutual relationships of the objects and degree of freedom. We refer to these parameters as kinematic parameters. For modelling the shape of the objects different models can be used ranging from line, plane, to more sophisticated surface models. We refer to these parameters as shape parameters. As the objects project onto the image plane, the image data may reflect the texture of the objects surface, the object contour, the optical flow if the objects are in motion, etc. We refer to this as image features. The problem of tracking articulated objects can be viewed as an estimation of the objects kinematic and shape parameters from image features. our aim is to isolate the (movement) vectors for the joints of the model. This first will give a position vector for all body parts and second we will get a movement vector for the head, which allows to gather data for the assessment of gaze direction, the shoulders, the elbows, the wrists, the lower body, the thighs, the knees and the ankles. These movement vectors then will be applied to a rendered simultaneously moving model. Speech will be processed simultaneously for loudness and frequency, thus allowing comparisons between movements and speech.

This analysis will produce a continuous data-stream that can be analysed in various ways. Each posture is defined as an unique set of vectors and each movement through unique changes of these vectors. Its advantage is that it does not need any interpretation on a higher level through an observer nor a computer based expert system which tries to reinterpret the movements. The vector data can be fed directly to neural nets for pattern recognition, and the patterns can be verified through rating studies. Future applications are person-recognition from gait, the monitoring of therapy success and the comparison of quality of signals in different species, under different contexts and physiological conditions.

4. COMMUNICATION THEORY AND PHYSIOLOGICAL STATES

Although it seems that we are just proposing a new method, this approach has consequences not only for the observation but also for the explanation of behaviour.

The main advantage of this approach, when it is compared to those using conventional coding methods, is that no presupposition on the structure and content of behaviour is made. This frees us from restrictions of conventional coding methods. With the use of conventional codes we can only find what we have put in the codes; behavioural codes are already hypotheses about behaviour. Behavioural codes are categories which represent many, sometimes different behavioural events. Although these categories correspond to the basic construction principles of our brain, which uses prototypes in order to reduce environmental information (Rosch, 1978), the assumption that communication works on the same level may be wrong. Indeed signals can be organised as prototypes but this is not necessarily so. We have shown that communication between humans can work on a level where no categorisation exists. The fallacy of looking at communication with the principles of the apparatus involved in it leads to a false and incomplete understanding of the nature of communication.

With this approach we possess the almost complete data-stream and we can look at how the brains of both the receiver and sender actually construct communicative reality. our approach allows the manipulation of stimuli which can be tested against reality. We will propose a new communication theory which is in its nature multi-modal and multilayered with different channels and many possible communicative principles.

The starting point for such a theory is that like the evolution of intelligence, the evolution of human communication has its basic constraints in machiavellism. Human brains are devices for processing information. We can suppose that there was differential survival and reproduction connected to optimal information processing (Lorenz, 1973, Cosmides *et al.* 1992). If there are adaptations to optimal information processing, these adaptations can be exploited. Thus, communication research has to deal with the constraints and possibilities presented by these adaptations. Future studies should look for adaptive information processing structures which could be exploitable through communication. We suppose that low-level processing of information is at least one possibility. This means that levels where the basic information is extracted from visual stimuli could be exploited. Comparable approaches could be made with the complexity of a stimulus. The less complex a stimulus is, the easier it could be decoded, producing higher levels of excitation in the brain. This is basically an open field but it can not be mastered with traditional coding and research methods. The problems which are connected to any communication theory we have shown in the introduction starting from the possibility of deception and ending with the possibility that noise is used tactically to veil intentions are avoided by our methodological approach. The method can even deal with repeated meaning encoded in pulse rate modulation when small changes or movements are repeated in time as shown in the waiting room study.

We propose that there are multi-layered processing mechanisms. The top layer for processing holds consciously accessible information. The bottom layers can not be assessed directly and controlled. on the top layer communication is actually an accessible information exchange about the real world with its social and ecological aspects. We can tell each other what we think about each other or gossip about others (Dunbar, 1993), we can create and use non-verbal signals like gestures (Morris *et al.*, 1979) differently in different cultures and we are able to lie and detect lies.

on the other hand we are able to veil our intentions by many measures like the creation of noise. We try to manipulate each other's physiological states and influence information processing in our social environment. This is the basic assumption of a new communication theory: brains are able to exploit others brain's functions and structures in order to manipulate them. These manipulations are intended and planned, and conscious access to these plans is not necessary for their realisation. Qualitative changes of behaviour which are present under different oestrogen levels can be used intentionally when male brains have been selected for detection of qualitative changes in behaviour caused by oestrogen levels which promise stable female cycles and successful reproduction. Such an approach do not need necessarily innate behaviours. Just basic construction principles for "what brains like." If our brain perceives "approaching speed" as danger then any fast movement toward another will be interpreted as threat. In such a way individuals could learn to use the same behaviours again and again. This would explain the wide variety and individual differences.

We are able to show that on the sender's side information might be encoded in the quality of behaviour. Females seem to do so under high risk conditions. This corresponds to the fact that females are more sensitive to the production and decoding of non-verbal behaviour (Rosenthal and Depaulo, 1979). The fact that communication is goal directed and depends on the pursued goals and possible risk of not achieving the goal has been neglected so far.

The results from qualitative movement changes in both studies have theoretical consequences. So far it is the first time where it has been shown that females try to manipulate male perception directly. Moreover, there should be at least some conscious

assessment of cycle state, because females signals can become more obvious when a stimulus is present and when the female is at an ovulatory stage. This underlines the hypotheses that female sexual crispis has indeed the function of promoting active female choice and thus can be used to induce sperm competition.

Showing off and subliminal manipulation is a means to manipulate the perception of one's self through others, and there is no need to assume that this is done consciously. In this article we have shown that the principles are comparable: changing the quality of behaviour, so that the receiver actually can not access the changes directly. This brings up another principle of manipulative communication. The sender has to avoid that the receiver might be able to learn. We can suppose that there is pressure on learning signals very fast in order to assess other's intentions early and reliably. Thus, communication should be variant and use different means in the same situations constantly. This leads to a model of parallel distributed processing for the decoding of meaning. The results on the classification of movements through neural nets propose such a model, although the model is only a poor approximation of human parallel processing.

Classical communication theories also do not account for the fact that signaling is not only about external information, but also about internal states and the manipulation of internal states which are encoded in behaviour quality. An exception to these hypotheses seems to be emotions which can be produced as signals. The problem in this is that the nature and signal value of emotions are unclear and it is not known to what extent qualitative changes in facial muscle movements affect emotional interpretation. There are some hints that actual movement quality is the cue which could be used for decoding information and not the actual configuration of muscle movements. Emotions among expressive actors are recognised easier than emotions among non-expressive actors (Wallbott, 1990). The solution to the communicative paradox thus lies in the possibility to observe the actual nature of communication with the help of new methods. Only behaviour recordings which are free from interpretation and which produce direct data are useful for the detection of communicative principles.

ACKNOWLEDGMENT

Funded by the Jubiliumsfond of the Austrian National Bank, P5676.

REFERENCES

- Alexander, R. D. & Noonan, K. M. 1979. Concealment of ovulation, parental care, and human social evolution. In: *Evolutionary biology and human social organization* (Ed. by N. A. Chagnon & W. G. Irons), pp. 436-453. Duxbury: North Scituate.
- Arbib, M. A. & Hansen, A. R. 1987. Vision, Brain and Cooperative Computation: an overview. In: *Vision, Brain and cooperative computation* (Ed. by M. A. Arbib & A. R. Hansen), pp. 1-16. Cambridge MA: The MIT Press.
- Argyle, M. 1988. *Bodily communication*. London: Methuen.
- Baker, R. R. & Bellis, M. A. 1995. *Human sperm competition. Copulation, masturbation and infidelity*. London: Chapman and Hall.
- Bassili, J. N. 1979. Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *J. Personality Soc. Psych.*, 37, 2049-2058.
- Benshoof, L. & Thornhill, R. 1979. The evolution of monogamy and concealed ovulation in humans. *J. Soc. Biol. Struct.*, 2, 95-106.

- Bernieri, E. J. & Rosenthal, R. 1991. Interpersonal coordination: behaviour matching and interactional synchrony. In: *Fundamentals of Nonverbal Behavior Part K Interpersonal Processes* (Ed by Feldman and Rime), pp. 401- 431. Harvard: Harvard University Press.
- Berry, D. S., Kean, K. J., Misovich, S. J. & Baron, R. M. 1991. Quantized displays of human movement: a methodological alternative to the point light display. *J. Nonverb. Behav.*, 15, 1-97.
- Brown, P. & Levinson, S. 1978. Universals in Language Usage: politeness phenomena. In: *Questions and Politeness. Strategies in Social Interaction.* (Ed. by E. Goody), pp. 56-289. Cambridge: Cambridge Univ. Press.
- Chance, M. R. A. & Russel, W. M. S. 1959. Protean displays: a form of allaesthetic behaviour. *Proc. Zool. Soc. London*, 132, 65-70.
- Cosmides, L., Tooby, J. & Barkow, J. H. 1992. Evolutionary psychology and conceptual integration. In: *The adapted mind* (Ed. by L. Cosmides, J. Tooby & I. H. Barkow), pp. 3-18. Oxford: Oxford University Press.
- Cunningham, M. R. 1977. Personality and the structure of the nonverbal communication of emotion. *J. Personality*, 45, S64-S84.
- Cutting, J. E. & Proffitt, D. E. 1981. Gait perception as an example of how we may perceive events. In: *Intersensory perception and sensory integration* (Ed. by R. D. Walk, & D. E. Proffitt), pp. 249-273. New York: Plenum Press.
- Dawkins, R. & Krebs, J. R. 1981. Signale der Tiere: Information oder Manipulation In: *Eco-Ethologie.* (Ed. by J. R. Krebs & N. B. Davies), pp. 222-242. Berlin und Hamburg: Parey.
- De Paulo, B. M. 1992. Nonverbal Behavior and Self-Presentation. *Psych. Bull.*, 111/2, 203-243.
- Dunbar, R. I. M. 1993. Coevolution of neocortical size, group size and language in humans. *Behav. Brain. Sci.*, 16, 681-735
- Eibl-Eibesfeldt, I. 1972. Similarities and differences between cultures in expressive movements. In: *Nonverbal communication.* (Ed. by R. A. Hinde), pp. 297-312. Cambridge: Cambridge University Press.
- Eibl-Eibesfeldt, I. 1989. *Human Ethology.* New York: Aldine de Gruyter.
- Ekman, P. & Friesen, W. V. 1969. Nonverbal leakage and clues to deception. *Psychiatry*, 32, 88-106.
- Ekman, R. & Friesen, W. V. 1971. Constants across cultures in the face of emotion. *J. Personality Soc. Psych.*, 17, 124-129.
- Ekman, R. & Friesen, W. 1972. Hand Movements. *J. Communication*, 22, 353-374.
- Ekman, R. & Friesen, W. 1978. *Facial Action Coding system.* Palo Alto, CA: Consulting Psychologists Press.
- Ekman, R., Friesen, W. V., O'Sullivan, M. & Scherer, K. R. 1980. Relative importance of face, body, and speech in judgements of personality and affect. *J. Personality Soc. Psych.*, 38, 270-277.
- Ekman, R., Levenson, R. W. & Friesen, W. V. 1983. Autonomous nervous activity distinguishes among emotions. *Science*, 221, 1208-1209.
- Essa, I. & Pentland, A. 1995. Facial expression recognition using a dynamic model and motion energy. Int'l Conference on Computer Vision, Cambridge, MA, June 20-23, 1995.
- Forgas, J. R. 1992. Affective Influences on Partner Choice. Role of Mood in Social Decisions. *J. Personality Soc. Psych.*, 61/5, 708.
- Frey, S. & Pool, J. 1976. A New Approach to the Analysis of Visible Behaviour. *Forschungsberichte aus dem Psychologischen Institut der Universität Bern.* Bern.
- Goffman, E. 1959. *The presentation of self in everyday life.* New York: Doubleday.
- Grammer, K. 1989. Human Courtship: Biological Bases and Cognitive Processing, In: *The sociobiology of sexual and reproductive Strategies* (Ed. by A. Rasa, C. Vogel & E. Volland), pp. 147-169. London: Chapman and Hal I.
- Grammer, K. 1992. Intervention in conflicts among children: context and consequences. In: *Coalitions and alliances in humans and other animals.* (Ed. by A. Harcourt & F. deWaal), pp. 259-283. Oxford: Oxford University Press.
- Grammer, K. 1991. Strangers meet: laughter and nonverbal signs of interest in opposite-sex encounters. *J. Nonverb. Behav.*, 14, 209-236.
- Grammer, K. 1995. *Signale der Liebe*. 3., neu bearbeitete Auflage. München: dtv-Wissenschaft.
- Grammer, K. & Eibl-Eibesfeldt, I. 1989. The ritualisation of laughter. In: *Natürlichkeit der Sprache und der Kultur* (Ed. by W. A. Koch), pp. 192-214. Bochum: Brockmeyer.
- Grammer, K., Honda, M. & Schmitt, A. 1996. Human courtship: digital image analysis of body movements. *J. Personality Soc. Psych.*, under revision.
- Grammer, K., Jütte, A. & Fischmann, B. 1996. Der Kampf der Geschlechter und der Krieg der Signale. In: *Sexualität im Spiegel der Wissenschaft.* Edition Universitat, Stuttgart: Hirzel. In press.
- Grammer, K. & Kruck, K. 1991. Decision making in opposite sex-encounters: love at first sight ?. Kyoto, 22nd International Ethological Conference.

- Grammer, K. & Kruck, K. 1996. Female control and female choice. In: *When women want sex: perspectives on female sexual initiation and aggression*. (Ed. by B. Anderson & C. Struckmann-Johnson) New York: Guilford Press.
- Grammer, K., Kruck, K. & Magnusson, M. 1996. The courtship dance: mathematical algorithms for pattern detection in non-verbal behaviour. *J. Nonverb. Behav.* (under revision).
- Grammer, K., Schiefelvel, W., Schleidt, M., Lorenz, B. & Eibl-Eibesfeldt, I. 1988. Patterns on the Face: brow movements in a crosscultural comparison. *Ethology*, 77, 279 - 299.
- Harper, D. G. C. 1992. Communication. In: *Behavioural ecology. An evolutionary approach*. (Ed. by J. R. Krebs & N. B. Davies), pp. 347 - 398. Oxford: Blackwell.
- Johansson, G. 1973. Visual perception of biological motion and a model of its analysis. *Perception & Psychophysics*, 14, 201-11.
- Johansson, G. 1976. Spatio-temporal differentiation and integration in visual motion perception. *Psychol. Res.*, 38, 379-393.
- Kakadiaris, I. A., Metaxas, D. & Bajcsy, R. 1994. Active part-decomposition, shape and motion estimation of articulated objects: A physics-based approach. *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 980-984. Seattle, Washington.
- Krauss, R. M., Apple, W., Morency, N. L., Wenzel, C. & Winton, W. 1981. Verbal, vocal, and visible factors in judgements of another's affect. *J. Personality Soc. Psychol.*, 40, 312-320.
- Kraut, R. E. & Johnston, R. E. 1979. Social and emotional messages of smiling: An ethological approach. *J. Personality / Soc. Psychol.*, 37, 1539-1553.
- Lorenz, K. 1973. *Die Rückseite des Spiegels*. München: Piper.
- MacArthur, L. Z. & Baron, R. M. 1983. Toward an ecological theory of social perception. *Psychol. Rev.* 90, 215-238.
- Maes, P., Darrell, T., Blumberg, B. & Pentland, A. 1995. The ALIVE system: wireless, full-body interaction with autonomous agents. *Proc. Computer Animation*, IEEE Press, April 1995.
- Malatesta, C. A. & Izard, C. E. 1984. The facial expression of emotion: Young, middle-aged, and other adult expressions. In: *Emotion in adult development* (Ed. by C. Z. Malatesta & C. E. Izard), pp. 253-273. Beverly Hills, CA: Sage.
- Markl, H. 1985. Manipulation, modulation, information, cognition: some of the riddles of communication. *Fortschritte der Zoologie*, 31, 163-194.
- Mehrabian, A. 1972. *Nonverbal communication*. Chicago: Aldine.
- Moghaddam, B. & Pentland, A. 1995. Probabilistic visual learning for object detection. Int'l Conference on Computer Vision, Cambridge, MA, June 20-23 1995.
- Montepare, J. R., Goldstein, S. B. & Clausen, A. 1987. The identification of emotions from gait information. *J. Nonverb. Behav.* 11, 33-42.
- Moore, M. M. 1985. Nonverbal courtship patterns in women: context and consequences. *Ethol. Sociobiol.*, 6, 237-247.
- Morris, D., Collett, B., Marsh, P. & O'Shaughnessy, M. 1979. *Gestures, their origins and distribution*. London: Jonathan Cape.
- Pentland, A. 1995. Machine understanding of human action. M.I.T. Media Laboratory Perceptual Computing Section Technical Report N0.350, Sept. 1995. Appeared: *7th Int'l Forum on Frontier of Telecommunication Technology*, Nov. 1995, Tokyo, Japan.
- Pentland, A. & Liu, A. 1995. Toward augmented control systems. IEEE Intelligent Vehicle Symposium 95, September 25-26, Detroit, MI.
- Provine, R. R. & Young, Y. L. 1991. Laughter: a stereotyped human vocalization. *Ethology*, 89, 115-124.
- Rosch, E. H. 1978. Principles of Categorization. In: *Cognition and Categorization* (Ed. by E. Rosch & D. Lloyd), pp. 27-48. Hillsdale: Erlbaum.
- Rosenthal, R. & Depaulo, B. M. 1979. Sex differences in eavesdropping on non-verbal cues. *J. Personality: Soc. Psychol.*, 37, 273-285.
- Runeson, S. & Frykholm, G. 1983. Kinematic specification of dynamics as an informational basis for person- and action perception. Expectation, gender recognition, and deceptive intention. *J. Exp. Psychol.*, 112, 585-615.
- Scheier, M. F. & Carver, C. S. 1981. Private and public aspects of self. In: *Review of personality and social psychology*, Vol. 2 (Ed. by L. Wheeler), pp. 189-216. Beverly Hills, CA: Sage.
- Schleidt, W. M. 1973. Tonic communication: continuous effects of discrete signs in animal communication systems. *J. Theoret. Biol.*, 42, 369-386.
- Siddiqi, J.A., Schwind, H.L. & Voss, H.G. 1973. Irrelevanz des Inhalts Relevanz des Ausdrucks. *Z. Experimentelle und Angewandte Psychologie*, 20, 472-488.

- Sonka, M., Hlavac, V. & Boyle R. 1993. *Image Processing, Analysis and Machine Vision* London: Chapman and Hall.
- Starnier, T. & Pentland, A. 1995. Visual recognition of american sign language using hidden Markov models *Proc Int'l Workshop on Automatic Face- and Gesture-Recognition*, Zurich, Switzerland, June 26-28, 1995.
- Tooby, J. & Cosmides, L. 1990. On the universality of human nature and the uniqueness of the individual: the role of genetics and adaptation. *J Personality*, 58, 1.
- Tooke, W. & Camire, L. 1991. Patterns of Deception in Intersexual and Intrasexual Mating Strategies. *Ethol. Sociobiol.*, 12, 345-345.
- Trivers, R. L. 1972. Parental investment and sexual selection. In: *Sexual selection and the descent of man 1871-1971*. (Ed. by B. Campbell), pp. 136-179. Chicago: Aldine.
- Van Hooff, J. A. R. A. M. 1972. A Comparative Approach to the Phylogeny of Laughter and Smile. In: *NonVerbal Communication*. (Ed. by R. A. Hinde), pp. 209-241. Cambridge: Cambridge University Press.
- Waibel, A., Hanazawa, T., Hinton, G., Shikano, K. & Lang K.J. 1989. Phoneme recognition using time-delay neural networks. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(13), 328-339.
- Wallbott, H. G. 1990. *Mimik im Kontext*. Göttingen: Verlag für Psychologie Dr. C.J. Hogrefe.
- Wallbott, H. G. 1991. The Emotional in Social Psychology and the Social in Emotion Psychology: An Overview Concerning the Intersection Between Social Psychology and Emotion Psychology. *Z. für Sozialpsychologie*, 22(1), S3-6S.
- Zell, A. 1994. *Simulation neuronaler Netze*. Bonn: Addison-Wesley.