

# Ritual or Routine: Communication in Long-Term Relationships with Companions

Sabine Payr

Austrian Research Institute for Artificial Intelligence (OFAI)\*

Freyung 6/6  
A-1010 Vienna  
Austria

sabine.payr@ofai.at

## Abstract

On the basis of data extracted from a long-term experiment with a robotic user interface, the paper discusses prerequisites and functions of interaction rituals.

## 1 Introduction

The challenge of building domestic companions - virtual or robotic - is to design them in such a way that they can build and maintain social relationships with their users. Experiences so far invariably point out the users' complaints that social interaction with companions, as complex as it may seem at first, becomes repetitive and boring. The researchers' conclusion has been that much more - ideally infinite - variation and variability in behaviour has to be designed for. [Bickmore et al. 2009]

On the other end of the spectrum of "household companions", we find simple devices like robotic vacuum cleaners that do not lay any claim to being social. Still, their users develop attachment to and relationships with them, without finding their single-purpose activity too boring or repetitive over time.

What, then, is the solution to these contradicting observations? In other words, under what circumstances do users accept what kind and degree of repetitiveness? And: is repetitiveness synonymous with monotony, or is there more to it? Answers to these questions could provide valuable guidelines for the design of domestic companions.

This paper first presents and discusses two sequences of video data collected during the first stage of the SERA field study (section 2). Section 3 presents and discusses concepts and theories of interaction

rituals with regard to their preconditions, signs and outcomes. Section 4 discusses Interaction Ritual Theory from the point of view of emotional processes and outcomes, in particular in long-term relationships. In the concluding section 5, we will be able to formulate more concrete questions for future research on human-companion interaction.

## 2 Backstage and Front Performance

The two video clips that are the basis for this paper stem from the first round of data collection in the on-

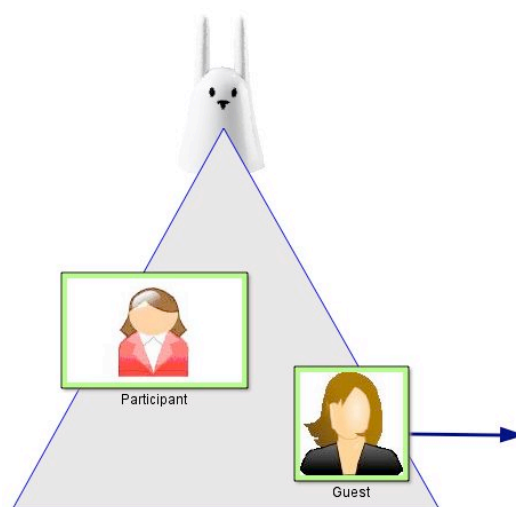


Fig. 1: Position of the participant and her guest in the second video sequence. The guest moves out of the camera's field of vision (grey triangle) during the interaction.

\* The Austrian Research Institute for Artificial Intelligence is supported by the Austrian Federal Ministry for Transport, Innovation, and Technology and by the Austrian Federal Ministry for Science and Research.

going EU project SERA (FP7, no. 231868). For a description of method and participants see [Klamer & Ben Allouch 2010, this volume]. The video recording was activated by the participants after a request by the

the Nabaztag ([www.violet.com](http://www.violet.com)), a rabbit-like robotic interface. The following two sequences were recorded on two consecutive days halfway through the experiment by one of the three participants. Both were the only recordings made on that particular day. In the transcripts, N stands for Nabaztag, P2 for the participant (female, age 50+), and (in the 2nd clip) G for a (female, somewhat younger) guest of P2, probably a friend. The diagram (Fig. 1) should help the reader to understand the spatial references in the transcript. The brackets mean:

(...) movements, non-verbal expressions, notes  
 [...] overlaps in the dialog  
 {...} interaction with N through buttons, switches

### Video 1: it1\_p2\_KOct08\_1803

*(P2 is alone, faces N)*

1 N: Are you feeling okay after today's activities?

2 P2: (slight frown, looks at N) {presses button - doesn't work}

3 N: (pause)

4 P2: (frown, looks up, sigh)

5 N: Press the buttons to say no or yes.

6 P2:  
{presses button}

7 N: Are you feeling okay after today's activities?

8 P2: {presses YES button several times} (then looks up, mouth slightly open, slight frown)

9 N: (short pause) Good.

10 P2: (closes mouth, nods, slight smile)

11 N: Do you think it was the right amount of activity for one day?

12 P2: (looks at N) {presses YES button several times}

13 N: Okay, great.

14 P2: (looks half sideways, slight smile)

15 N: Keep going with the activity plan.

16 P2: (looks at N, nods, slight smile)

17 P2: (waits a little, waves her hand at N - restrained gesture)

### Video 2: it1\_p2\_KOct09\_1041

*(G stands in the doorway opposite N, P2 stands beside it in profile = start position)*

19 N: You are going out? Please could you press the video button

20 P2: (turns to G)  
(back to start pos)

21 N: on your way past? See you later.

22 P2: (turns to N, looks at N){presses button}

23 N: Recording on.

24 P2: (start pos, steps back, sottovoce) don't know why if I'm going out.

25 G: Pardon?

26 P2: (louder) don't know why if I'm going out.

27 G: Oooh?

28 P2: (steps forward, faces N, smiles) Hello, rabbit

29 P2: (bends) {puts keys on hook switch} (sottovoce) right

30 P2: (steps back out of picture)

31 N: Welcome home. If you don't mind the video recording being done, could you press the video button please.

32 G: (laughs, leaves the scene, but remains visible to P2)

33 N: Did you have a good time?

34 P2: (forward, turns to N) {presses YES button} (looks up toward camera)

35 P2: (half turns toward G) The YES button doesn't [work]

36 N: [good]

37 P2: (displays surprise): Oh ... (?rest not intelligible)

38 N: Were you doing some exercise?

39 P2: (laughs, looks back at N) {presses YES button} (looks away, smiles)

40 N: Ok, but remember that it's important

41 P2: (nods, glances at G, laughs)

42 N: to stick to the activity plan where possible

43 P2: (turns to N, stronger nod, smiles)

44 N: Don't forget to stop if you feel tired

45 P2: (grimace, strong head-shaking, smiles)

46 N: and take regular breaks.

47 P2: (strong nods)

48 P2: (grimace, steps away, toward G): what's a break

49 P2: (looks back at N, smiles): By-ye!

50 P2: (steps away)

51 G: (off: laughs)

The dialogs of the Nabaztag are scripted. The first dialog is activated after P2 has completed the final (scheduled) activity of the day. The first part of the second video shows the dialog that is activated when the participant goes out (= removes the keys from the sensitive hook) and no scheduled activity is due. P2 then puts the keys back on the hook to activate the dialog for those occasions where the participant comes home from non-scheduled activities. In content, this dialog is quite similar to that in video 1, as the Nabaz-

tag in this study is supposed to coach and monitor the subjects' physical activities.

The similarity in content allows to concentrate on the difference in P2's behaviour in the two videos. In Video 1, she responds to the Nabaztag's utterances with slight nods and smiles. We cannot be sure whether these are feedback to the utterances alone or also, partly, expressions of satisfaction that the device is functioning (after some previous technical trouble). For the first time during the study, a greeting can be observed: she waves her hand slightly in a good-bye gesture. In the second video, feedback, facial expressions and greeting are much more expressive. The greeting is now also expressed verbally, nods and head-shakes are pronounced, facial expression is exaggerated to the point of grimacing.

The main difference in the setting of the scene is the presence of a person in the second. The strongest impression one gets when watching these two videos is the contrast between the private and the public situation, or, to put it in Goffman's [1959] terms, the backstage and the front performance. Goffman noticed that the frontstage events are characterized by dramatization and idealization. Dramatization is clearly visible in this video: P2's facial expressions are more expressive than in everyday conversation, and much more than in her private interaction with the Nabaztag (Video 1). But what could be meant by "idealization"?

In Video 1, we see the private interaction between P2 and the Nabaztag. It is private in the sense that no other person is present: it is true that there is the camera recording, but it is not evident that P2 is aware of it, because her gaze and gestures are directed toward the Nabaztag and not at the camera. In this video, we see the interaction as the researchers have imagined it: a one-on-one interpersonal dialog.

In Video 2, on the contrary, the participant performs the interaction as she sees it, or more precisely: as she wants others to see it. She creates a little drama presenting what, for her, counts as a good interaction. She would have several options for this performance of and with the Nabaztag: one would be to highlight its malfunctioning, another one would be to show its stupidity. Both of these elements are present in the first part of Video 2 (line 24 and line 35), and both would involve "taking sides" with the other person and a distancing from the Nabaztag. Instead, to get more of a performance, she initiates the "coming home" sequence by putting back the key. Her position, facing halfway between the Nabaztag and her friend, indicates that, for her, there are two "others" in this interaction, and she addresses the human and the machine in turns. Her focus of attention turns more to the Nabaztag as the interaction progresses. She elaborates a dramatic "peak" in it which starts when she finds that the button this time works perfectly (line 38). She turns to the Nabaztag, and nods and shakes her head in synchrony with the positive and negative statements of the dialog (lines 41 to 47). Compared with the first dialog, gestures are significantly longer and more expressive. Their most striking feature is that they take up and underscore the rhythm of the Nabaztag's

speech. After an "aside" to her friend (line 48), she closes the interaction with a verbal greeting (line 49) - which is the only such greeting we have recorded from this participant. This idealized interaction has many elements of what has been called a "ritual" in sociological literature.

### 3 Interaction rituals

#### 3.1 The ritual performance

"Two alternative conceptions of communication have been alive ... since this term entered common discourse", writes Carey [2009], and goes on to explain his distinction between the transmission and ritual views of communication.

When communication is viewed as transmission, it is understood in terms of sending, receiving, and distributing information, in general in metaphors of transportation and exchange of packaged goods (cf. the "conduit metaphor" of communication, [Reddy 1979]). Whereas, in the view of communication as ritual, it is connected with terms such as sharing and participation. It reminds of the etymological relationship of the term with communion or community. "A ritual view of communication is directed not toward the extension of messages in space but toward the maintenance of society in time" (ibid. p. 15). It does not primarily serve to impart information but to express shared beliefs and emotions.

The purpose of communication is not the transmission of information but the construction and maintenance of a meaningful cultural world. Communication is a symbolic process whereby reality is produced, maintained, repaired, and transformed. Carey [2009] illustrates the difference and the necessity to reconcile both views with the "news". What the audience finds in them is not only and not even primarily information but stories on the contending forces at work in the world. "Under a ritual view, then, news is not information but drama" that invites our participation. News are not consumed for their content, but for their promise to make the reader/spectator a member in the ongoing dramas and stories.

Goffman [1967, 1981] transformed Durkheim's analysis of ritual religious gatherings [1912] into the concept of encounter which he saw as the unit of interaction, and so brought the ritual from religion into everyday face-to-face interaction. Collins [2004] has an even broader concept of ritual. Drawing on Durkheim and Goffman, he resumes the necessary ingredients of a ritual as follows:

- co-presence
- boundaries
- common focus of attention
- sharing a common mood or experience

Where Goffman saw the stereotyped sequences of talk and other gestures (used e.g. to open, close, and repair) as the defining characteristics of rituals, Collins takes his model of interaction rituals to the whole of ordinary conversation and shows that all the characteristics of a ritual can be found here. Turn-taking, for exam-

ple, can only succeed smoothly when there is an underlying rhythmic coordination. Body movements and nonverbal behaviour are synchronized in successful interaction on such a subconscious level that even brainwaves are involved. In Conversation Analysis, such phenomena have been studied under the heading of "alignment" [Bateman 2006, Branigan 2006], but the subtleties cannot be detected with its methods. Instrumental analyses of conversations have shown that synchronization is correlated with a feeling of solidarity. The participants, in this rhythmic entrainment, do not react to each other - which would be too slow - but fall into the same rhythm so that they can anticipate the "beats" of the other's talk and turn.

Such a rhythmic coordination is performed by the participant in the second video. It is "performed" in the sense that it is dramatized: nods and head-shakes are slightly exaggerated, which becomes visible in comparison with the first video. By facing the Nabaztag and thus, for a few turns, excluding her friend from the interaction, she draws the boundaries of the interaction and acts "as if" she and the device had a mutual focus of attention. In a natural conversation, the unconscious process of alignment is the work of both participants. Here, it is the human alone who does the "job" of rhythmic entrainment by adapting to the Nabaztag.

A successful interaction ritual generates shared emotions and intensifies them: beside rhythmic entrainment, there is also emotional entrainment of whatever emotions there are. The participant in the video also shows slightly exaggerated facial expressions ("grimace" in the transcript) that reinforce nods and head-shakes with agreement and rejection.

### 3.2 Ritual and routine

A ritual, in the everyday meaning of the word, involves stereotyped actions such as prescribed formulas, costume, gestures, protocols. These props contribute to the core process, but they are neither necessary nor sufficient ingredients. Indeed, if a ceremony relies only on the formal rules and elements, it fails to become a ritual. Collins [2004] calls this sort of ritual "formal" and contrasts it with "natural" rituals. A formal ritual usually is repeated periodically to keep it alive. A natural ritual, on the other hand, can come off spontaneously without explicit concern, e.g. the rituals of everyday sociability such as greetings. The borders are fluid: a natural ritual can crystallize around fixed symbols whereby subsequent rituals of this kind are increasingly formalized. The difference between the two, then, is not that the natural ritual is always and completely new and spontaneous. In fact, greetings and formal politeness are strongly stereotyped and more or less formalized through repetition. Repetition can lead to routinization if the participants lose the shared focus of attention, but some repetition and take-up is necessary for rituals to confirm their symbolic value and to renew the "emotional energy" that is their outcome.

Bedtime rituals for small children are an example of interpersonal or intra-family culture. They tend to become highly repetitive in content, sequence of

events, even gestures and words. Their repetitiveness and similarity come themselves to be symbols of their meaning: the order and continuity of the world into the next day is ascertained, and the monsters of the night are effectively chased and banned. With their "magic" effect they come very close to the religious rituals described by Durkheim. What distinguishes them from mere routines is their emotional outcome. With Goffman, we could say that rituals are not repeated, but re-performed.

A routine is characterized, in contrast, by the lack of focused attention. Even if it is carried out by a group, the members act on their own as individuals (e.g. on the assembly line). Rituals can decay into routines when they lose their symbolic strength, while a familiar routine can by its repetition come to symbolize continuity itself and gain the attention of the participants, and so be "celebrated" as an emotionally gratifying ritual. Routines and rituals may share repetitiveness, but are nonetheless different in the level of attention and emotional outcome. Interaction rituals could be started spontaneously, but then be carried on with more or less variation, some will decay into routines while new ones will emerge.

## 4 Ritual and emotion

### 4.1 Emotional outcomes

The ritual as a source and catalyst of emotions has a long tradition in sociology. Already Durkheim [1912/1965] described the "emotional effervescence" as the outcome of ritual gatherings. Goffman [1967] noted that feelings of solidarity emerge in the encounter. In this line, Collins [2004] says that the long-term and most important outcome of an interaction ritual is "emotional energy". Emotional energy is more enduring than the varying transient emotions that can arise in a particular situation. The gain in positive emotional energy itself is the motivation for seeking and entering into interaction rituals. A common mood or shared feeling such as joy, anger, sadness etc. are ingredients and prerequisites of the interaction ritual. The sharing and coordination of these feelings by the group reinforces this transient emotion, but this is only the short-term effect. In the long term, what remains is what he calls an "energy": the feeling of attachment to the group, of solidarity and belonging. Collins thus makes an effort to actually ground social life in everyday interaction, to show how common conversation contributes to the (re)construction of society.

Seen from the perspective of emotion research, his concept of emotional energy is so general and all-inclusive that it risks to be empty: Collins collapses the two dimensions of valence and arousal into one by putting enthusiasm, confidence and good self-feelings at one end of the spectrum and depression, lack of initiative and negative self-feelings on the other. This leads him then to link the amount of emotional energy that individuals can take away from an interaction ritual to their dominance and power [see also Collins 1990]: the more powerful they are (e.g., a group

leader), the more emotional energy they get out of the ritual. This hypothesis serves well the sociologist, in that it allows to link interaction rituals to macro-social conflicts for power [cf. Turner & Stets 2005] but does not help much in the study of everyday interaction where power differences without conflicts are the norm.

In this regard, Affect Control Theory (ACT) [Heise 2002, 2004, MacKinnon 1994] offers a more differentiated approach to the emotional outcome of interactions. It starts out recognizing different social identities (roles) that come together, with different social and affective meanings, among which their perceived power. It goes on to state that what people seek in the interaction is confirmation of their respective identities. That the successful confirmation confers a good self-feeling remains implicit, but the outcome is doubtlessly emotional. Taking, as an example, a successful conversation between a customer and a call-center agent, Collins' model cannot well explain how both sides can come away equally satisfied from such an encounter. But both customer and agent can confirm their identities which is none other than reinforcing their solidarity and bonds with their respective social groups. While Collins is concerned mainly with in-group rituals, ACT allows us thus to take the idea of interaction rituals to out-group encounters.

The first three ingredients for a ritual (see above) can be present also in such out-group interactions, but we have to ask whether persons with different identities (i.e., with different group memberships) can "share a common mood or experience". Goffman's dramaturgical approach gives us a hint to what they can have in common: the participants share the *performance* of their respective acts. They can have in common the awareness of the stage, the roles, backstage and front. With such a modified view of interaction rituals, the theory can become relevant for human-machine interaction where fundamental differences between participants are obvious.

#### 4.1 Rituals in long-term relationships

Companions should ideally build and maintain long-term relationships with their owners. In apparent contradiction to the findings from long-term experiments with agents and robots [e.g. Bickmore & Picard 2005, see also Klamer & Ben Allouch 2010, this volume], commonsense and experience tell us that human-human relationships are far from being without repetitiveness. There are both rituals and routines, and they evolve to take up a significant part of the communication in everyday interactions. The example of the bedtime ritual is only of them. It can be safely assumed that the longer and (spatially) closer a relationship is, the more the proportion of rituals and routines in interaction will grow. People living together do not reinvent their daily interactions from scratch every morning. Cognitive economy is one factor that leads to a preference for similar situations, uncertainty avoidance another one. Where repetitive interactions can be qualified as rituals, however, they directly contribute to emotional well-being.

The role of rituals in the emotional life of long-term interpersonal relationships has not yet been studied in detail. While exchange theories [cf. Eimler et al. 2010, this volume] are based on a *trading* metaphor of emotional cost and benefit, Interaction Ritual Theory would rather be based on a *production* metaphor, because interaction rituals can generate, out of situation, co-presence and mutual attention, an emotional surplus, i.e. the feeling of belonging (bond) that is at the centre of human relationships, regardless of their content.

In the SERA field study, semi-structured interviews were conducted with all three participants [Klamer & Ben Allouch 2010, this volume]. As in comparable studies [like e.g. Bickmore & Picard 2005], subjects qualified the interaction with the companion as repetitive and rather boring. On the one hand, this result confirms that the companion is not considered a mere "machine" in the same way as, for example, a coffee-maker or vacuum cleaner: from machines, we do not expect variation. On the contrary: deviations from usual behaviour are irritating and considered as errors. On the other hand, it leaves open the question what users expect and are ready to accept in a companion. Do they indeed expect potentially endless variation, as for example from a radio or other mediating devices that do not rely on in-built content? Or would they accept a certain degree of routine which opens the possibility to develop ritual practices in which the companion is involved? Or else, would a companion that appears and behaves more machine-like, lower the expectations of variety?

We cannot answer these questions yet. We have shown, however, that more behavioural variety in companions is only one possible conclusion to draw from user feedback. There are only a handful of studies on long-term use of companions to date, so that the development of habits, routines and rituals, and the embedding of such devices into everyday practices is an uncharted area on the map. The risk is that it will remain so if researchers do not adapt their methods of inquiry and their questions to this challenge.

The SERA field study offers a unique possibility for this research because it collects observational data (video recordings of interactions over time) instead of relying on subjective data in the form of interviews or questionnaires.

## 5. Resume and Outlook

In the second video presented here, the participant performs a ritual with what she imagines as the ideal companion, and we can take these hints into our research agenda:

- Re-performance vs. repetition: why, how and when exactly do users notice and criticize repetitiveness?
- A certain repetitiveness of behaviour is a prerequisite for the development of rituals, but not monotony. The pattern of behavioral differentiation will have to be anything from "variations over a theme" to a song with stanzas and chorus. What amount of repetitiveness is acceptable, and is it related with ap-

pearance, user expectations, and functions of the companion?

- An interaction ritual is a mutual effort and a joint action. The participant in the experiment adapts to the Nabaztag in the performance unilaterally, but would this be sufficient in a long-term relationship, or should a companion be able to contribute through (adaptive, "performing" ...) behaviour?
- Rhythmic entrainment and subverbal alignment will require speech generation which adapts, to a certain degree, to the speed, voice and beat of the individual human speaker [Suzuki et al. 2003]. It is an open question whether absolute voice qualities [Nass & Brave 2005] are more important than these (user-)relative features.
- Co-presence: is there a difference in the evolution of interaction rituals between physically (robots) vs. virtually (agents) embodied companions?
- What social roles and how much time/space will owners give their companions in long-term everyday use? How can companions accommodate the wide variety of user attitudes? Or should owners rather be able to contribute actively to their "social configuration"?

Companions need not and should not mimic human-human relationships. They are devices that satisfy certain needs of their owners and have their uses and functions in the owner's life. When they play a role in the owner's health, well-being and independent living, however, they assume a role that goes far beyond that of, say, a vacuum cleaner, and they have to be able to maintain that role over a longer period. In this light, it becomes essential to investigate how long-term relationships are built and re-built on the micro-level of conversational interaction.

## Acknowledgment

The research leading to these results has received funding from the European Community's Seventh Framework Programme [FP7/2007-2013] under grant agreement no. 231868 and from the Austrian Federal Ministry of Science and Research under its Additional Funding Programme.

## References

- [Bateman 2006] John A. Bateman: A social-semiotic view of interactive alignment and its computational instantiation: a brief position statement and proposal. In *How People Talk to Computers, Robots, and Other Artificial Communication Partners*, edited by K. Fischer: SFB/TR 8 Spatial Cognition.
- [Bickmore et al. 2009] Timothy Bickmore, D. Schulmann and C. Sidner: Issues in Designing Agents for Long Term Behavior Change. CHI'09 Workshop on Engagement by Design.
- [Bickmore & Picard 2005] Timothy Bickmore, Rozalind Picard: Establishing and Maintaining Long-Term Human-Computer Relationships. ACM Transactions on Computer Human Interaction (ToCHI), 59(1):21-30.
- [Branigan & Pearson 2006] Holly Branigan, Jamie Pearson: Alignment in Human-Computer Interaction. In *How People Talk to Computers, Robots, and Other Artificial Communication Partners*, edited by K. Fischer: SFB/TR 8 Spatial Cognition.
- [Carey 2009] James W. Carey: *Communication as Culture*. Revised edition. New York: Routledge.
- [Collins 1990] Randall Collins: Stratification, Emotional Energy, and the Transient Emotions. In *Research Agendas in the Sociology of Emotions*, edited by T. D. Kemper. New York: State University of New York Press.
- [Collins 2004] Randall Collins: *Interaction Ritual Chains*. Princeton, NJ: Princeton University Press.
- [Eimler et al. 2010] Sabrina Eimler, Nicole Krämer, Astrid von der Pütten: Prerequisites for Human-Agent and Human-Robot Interaction: Towards an Integrated Theory. In Trappl, R. (ed.): Proceedings of EMCSR 2010. Vienna.
- [Goffman 1959] Erving Goffman: *The Presentation of Self in Everyday Life*. New York: Doubleday.
- [Goffman 1981] Erving Goffman: *Forms of Talk*. Philadelphia: University of Pennsylvania Press.
- [Heise 2002] David Heise: Understanding Social Interaction with Affect Control Theory. In *New Directions in Contemporary Sociological Theory*, edited by J. Berger and M. Zelditch: Boulder, CO: Rowman and Littlefield.
- [Heise 2004] David Heise: Enculturating Agents with Expressive Role Behavior. In *Agent Culture. Human-Agent Interaction in a Multicultural World*, edited by S. Payr and R. Trappl: Mahwah (NJ): Lawrence Erlbaum Associates.
- [Klamer & Ben Allouch 2010] Tineke Klamer, Somaya Ben Allouch: Acceptance and Use of a Zomorphic Robot in a Domestic Setting. in Trappl, R. (ed.): Proceedings of EMCSR 2010. Vienna.
- [MacKinnon 1994] Neil J. MacKinnon: *Symbolic Interaction as Affect Control*. Albany: State University of New York Press.
- [Nass & Brave 2005] Clifford Nass, Scott Brave: *Wired for Speech*. Cambridge, MA: MIT Press.
- [Suzuki et al. 2003] N. Suzuki, Y. Takeuchi, K. Ishii, M. Okada: Effects of echoic mimicry using hummed sounds on human-computer interaction. *Speech Communication* 40: 559-573.
- [Turner & Stets 2005] J. H. Turner, J. E. Stets: *The Sociology of Emotions*. Cambridge: Cambridge University Press.