

The Modulation of Cooperation and Emotion in Dialogue: The REC Corpus

Federica Cavicchio

Mind and Brain Center/ Corso Bettini 31,
38068 Rovereto (Tn) Italy
federica.cavicchio@unitn.it

Abstract

In this paper we describe the Rovereto Emotive Corpus (REC) which we collected to investigate the relationship between emotion and cooperation in dialogue tasks. It is an area where still many unsolved questions are present. One of the main open issues is the annotation of the so-called “blended” emotions and their recognition. Usually, there is a low agreement among raters in annotating emotions and, surprisingly, emotion recognition is higher in a condition of modality deprivation (i. e. only acoustic or only visual modality vs. bimodal display of emotion). Because of these previous results, we collected a corpus in which “emotive” tokens are pointed out during the recordings by psychophysiological indexes (ElectroCardioGram, and Galvanic Skin Conductance). From the output values of these indexes a general recognition of each emotion arousal is allowed. After this selection we will annotate emotive interactions with our multimodal annotation scheme, performing a kappa statistic on annotation results to validate our coding scheme. In the near future, a logistic regression on annotated data will be performed to find out correlations between cooperation and negative emotions. A final step will be an fMRI experiment on emotion recognition of blended emotions from face displays.

1 Introduction

In the last years many multimodal corpora have been collected. These corpora have been recorded in several languages and have being elicited with different methodologies: acted (such as for emotion corpora, see for example Goeleven, 2008), task oriented corpora, multiparty dialogs, corpora elicited with scripts or storytelling and ecological corpora. Among the goals of collection and analysis of corpora there is shading light on crucial aspects of speech production. Some of the main research questions are how language and gesture correlate with each other (Kipp et al., 2006) and how emotion expression modifies speech (Magno

Caldognetto et al., 2004) and gesture (Poggi, 2007). Moreover, great efforts have been done to analyze multimodal aspects of irony, persuasion or motivation.

Multimodal coding schemes are mainly focused on dialogue acts, topic segmentation and the so called “emotional area”. The collection of multimodal data has raised the question of coding scheme reliability. The aim of testing coding scheme reliability is to assess whether a scheme is able to capture observable reality and allows some generalizations. From mid Nineties, the kappa statistic has begun to be applied to validate coding scheme reliability. Basically, the kappa statistic is a statistical method to assess agreement among a group of observers. Kappa has been used to validate some multimodal coding schemes too. However, up to now many multimodal coding schemes have a very low kappa score (Carletta, 2007, Douglas-Cowie et al., 2005; Pianesi et al., 2005, Reidsma et al., 2008). This could be due to the nature of multimodal data. In fact, annotation of mental and emotional states of mind is a very demanding task. The low annotation agreement which affects multimodal corpora validation could also be due to the nature of the kappa statistics. In fact, the assumption underlining the use of kappa as reliability measure is that coding scheme categories are mutually exclusive and equally distinct one another. This is clearly difficult to be obtained in multimodal corpora annotation, as communication channels (i.e. voice, face movements, gestures and posture) are deeply interconnected one another.

To overcome these limits we are collecting a new corpus, Rovereto Emotive Corpus (REC), a task oriented corpus with psychophysiological data registered and aligned with audiovisual data. In our opinion this corpus will allow to clearly identify emotions and, as a result, having a clearer idea of facial expression of emotions in dialogue. In fact, REC is created to shade light on the relationship between cooperation and emotions in dialogues. This resource is the first

up to now with audiovisual and psychophysiological data recorded together.

2 The REC Corpus

REC (Rovereto Emotive Corpus) is an audiovisual and psychophysiological corpus of dialogues elicited with a modified Map Task. The Map Task is a cooperative task involving two participants. It was used for the first time by the HCRC group at Edinburg University (Anderson et al., 1991). In this task two speakers sit opposite one another and each of them has a map. They cannot see each other's map because they are separated by a short barrier. One speaker, designated the Instruction Giver, has a route marked on her map; the other speaker, the Instruction Follower, has no route. The speakers are told that their goal is to reproduce the Instruction Giver's route on the Instruction Follower's map. To the speakers are told explicitly that the maps are not identical at the beginning of the dialogue session. However, it is up to them to discover how the two maps differ.

Our map task is modified with respect to the original one. In our Map Task the two participants are sitting one in front of the other and are separated by a short barrier or a full screen. They both have a map with some objects. Some of them are in the same position and with the same name, but most of them are in different positions or have names that sound similar to each other (e. g. Maso Michelini vs. Maso Nichelini, see Fig. 1). One participant (the giver) must drive the other participant (the follower) from a starting point (the bus station) to the finish (the Castle).

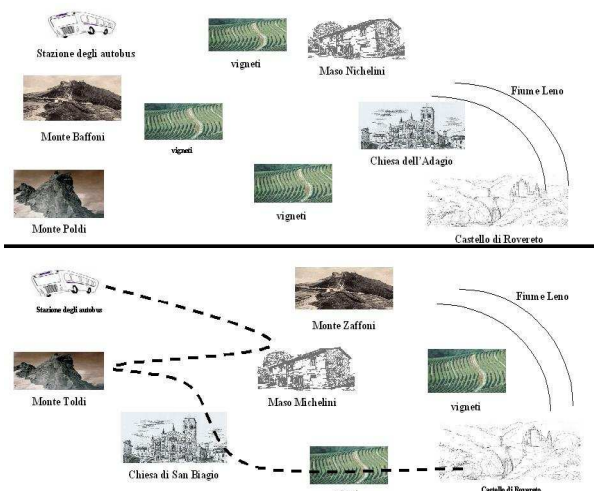


Figure 1: Maps used in the recording of REC corpus

Giver and follower are both native Italian speakers. In the instructions it was told them that they

will have no more than 20 minutes to accomplish the task. The interaction has two conditions: screen and no screen. In screen condition a barrier was present between the two speakers. In no screen condition a short barrier, as in the original map task, was placed allowing giver and follower to see each other's face. With these two conditions we want to test whether seeing the speakers face during interactions influences facial emotion display and cooperation (see Kendon, 1967; Argyle and Cook 1976; for the relationship between gaze/no gaze and facial displays; for the influence of gaze on cooperation and coordination see Brennan et al., 2008). A further condition, emotion elicitation, was added. In "emotion" condition the follower or the giver can alternatively be a confederate, with the aim of getting the other participant angry. In this condition the psychophysiological state of the confederate is not recorded. In fact, as it is an acted behavior, it is not interesting for research purpose. All the participants had given informed consent and the experimental protocol has been approved by the Human Research Ethics Committee of Trento University.

REC is by now made up of 17 dyadic interactions, 9 with confederate, for a total of 204 minutes of audiovisual and psychophysiological recordings (electrocardiogram and derived heart rate value, and skin conductance). Our goal is reaching 12 recordings in the confederate condition. During each dialogue, the psychophysiological state of non-confederate giver or follower is recorded and synchronized with video and audio recordings. So far, REC corpus is the only multimodal corpus which has psychophysiological data to assess emotive states.

The psychophysiological state of each participant has been recorded with a BIOPAC MP150 system. In particular, Electrocardiogram (ECG) was recorded by Ag AgC1 surface electrodes fixed on participant's wrists, low pass filter 100 Hz, at a 200 samples/second rate. Heart Rate (HR) has been automatic calculated as number of heart beats per minute. Galvanic Skin Conductance (SK) was recorded with Ag AgC1 electrodes attached to the palmar surface of the second and third fingers of the non dominant hand, and recorded at a rate of 200samples/second. Artefacts due to hand movements have been removed with proper algorithms. Audiovisual interactions are recorded with 2 Canon Digital Cameras and 2 free field Sennheiser half-cardioid microphones with permanently polarized condenser, placed in front of each speaker

The recording procedure of REC is the following. Before starting the task, we record baseline condition that is to say we record participants' psychophysiological outputs for 5 minutes without challenging them. Then the task started and we recorded the psychophysiological outputs during the interaction which we called task condition. Then the confederate started challenging the speaker with the aim of getting him/her angry. To do so, the confederate at minutes 4, 9 and 13 of the interaction plays a script (negative emotion elicitation in giver; Anderson et al., 2005):

- *"You driving me in the wrong direction, try to be more accurate!"*;

- *"It's still wrong, this can't be your best, try harder! So, again, from where you stop"*;

- *"You're obviously not good enough in giving instruction"*.

In Fig. 2 we show the results of a 1x5 ANOVA executed in confederate condition. Heart rate (HR) is confronted over the five times of interest (baseline, task, after 4 minutes, after 9 minutes, after 13 minutes). The times of interest are baseline, task, and after 4, 9 and 13 minutes, that is to say just after emotion elicitation with the script.

We find that HR is significantly different in the five conditions, which means that the procedure to elicit emotions is incremental and allows recognition of different psychophysiological states, which in turns are linked to emotive states. Mean HR values are in line with the ones showed by Anderson et al. (2005). Moreover, from the inspection of skin conductance values (Fig. 3) there is a linear increase of the number of peaks of conductance over time. This can be due to two factors: emotion elicitation but also an increasing of task difficulty leading to higher stress and therefore to an increasing number of skin conductance peaks.

As Cacioppo et al. (2000) pointed out, it is not possible to assess the emotion typology from psychophysiological data alone. In fact, HR and skin conductance are signals of arousal which in turns can be due both to high arousal emotions such as happiness or anger. Therefore, we asked participants after the conclusion of the task to report on a 8 points rank scale the valence of the emotions felt towards the interlocutor during the task (from extremely positive to extremely negative). On 10 participants, 50% of them rated the experience as quite negative, 30% rated the

experience as almost negative, 10% of participants rated it as negative and 10% as neutral.

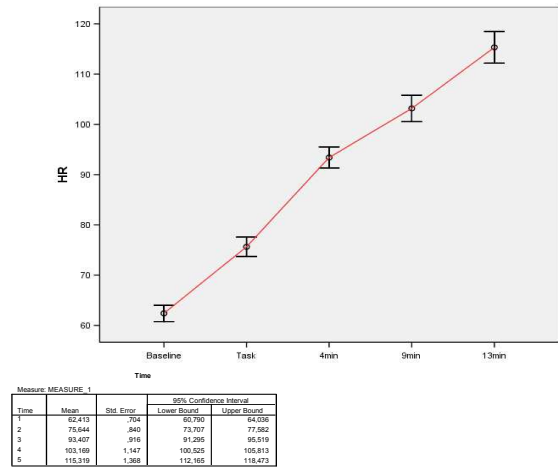


Figure 2: 1x5 ANOVA on heart rate (HR) over time in emotion elicitation condition in 9 participants

Participants who have reported a neutral or positive experience were discarded from the corpus.

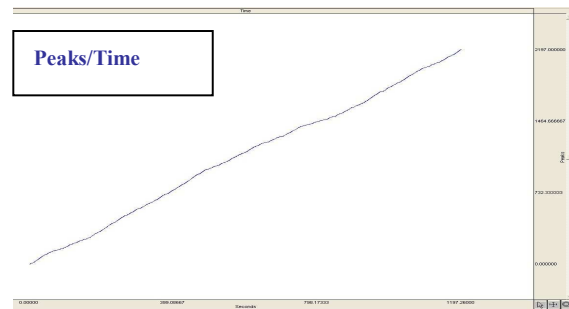


Figure 3: Number of skin conductance positive peaks over time in emotion elicitation condition in 9 participants

3 Annotation Method and Coding Scheme

The emotion annotation coding scheme used to analyze our map task is quite far from the emotion annotation schemes proposed in Computational Linguistic literature. Craggs and Woods (2005) proposed to annotate emotions with a scheme where emotions are expressed at different blending levels (i. e. blending of different emotion and emotive levels). In Craggs and Woods opinions' annotators must label the given emotion with a main emotive term (e. g. anger, sadness, joy etc.) correcting the emotional state with a score ranging from 1 (low) to 5 (very high). Martin et al. (2006) used a three steps rank scale of emotion valence (positive, neutral and negative) to annotate their corpus recorded from TV interviews.

But both these methods had quite poor results in terms of annotation agreement among coders.

Several studies on emotions have shown how emotional words and their connected concepts influence emotion judgments and their labeling (for a review, see Feldman Barrett et al., 2007). Thus, labeling an emotive display (e. g. a voice or a face) with a single emotive term could be not the best solution to recognize an emotion. Moreover researchers on emotion recognition from face displays find that some emotions as anger or fear are discriminated only by mouth or eyes configurations. Face seems to be evolved to transmit orthogonal signals, with a lower correlation each other. Then, these signals are deconstructed by the “human filtering functions”, i. e. the brain, as optimized inputs (Smith et al., 2005). The Facial Action Units (FACS, Ekman and Friesen, 1978) is a good scheme to annotate face expressions starting from movement of muscular units, called action units. Even if accurate, it is a little problematic to annotate facial expression, especially the mouth ones, when the subject to be annotated is speaking, as the muscular movements for speech production overlaps with the emotional configuration.

On the basis of such findings, an ongoing debate is whether the perception of a face and, specifically, of a face displaying emotions, is based on holistic perception or perception of parts. Although many efforts are ongoing in neuroscience to determine the basis of emotion perception and decoding, little is still known on how brains and computer might learn part of an object such as a face. Most of the research in this field is based on PCA-alike algorithms which learn holistic representations. On the contrary other methods such as non Negative Matrix Factorization are based on only positive constrains leading to part based additive representations. Keeping this in mind, we decide not to label emotions directly but to attribute valence and activation to nonverbal signals, “deconstructing” them in simpler elements. These elements have implicit emotive dimensions, as for example mouth shape. Thus, in our coding scheme a smile would be annotate as “)” and a large smile as “+)”. The latter means a higher valence and arousal than the previous signal, as when the speaker is laughing.

In the following, we describe the modalities and the annotation features of our multimodal annotation scheme. As an example, the analysis of emotive labial movements implemented in our annotation scheme is based on a little amount of signs similar to emoticons. We sign two levels

of activation using the plus and minus signs. So, annotation values for mouth shape are:

- o open lips when the mouth is open;
- closed lips when the mouth is closed;
-) corners up e.g. when smiling; +) open smile;
- (corners down; +(corners very down
- !cornerup for asymmetric smile;
- O protruded, when the lips are rounded.

Similar signals are used to annotate eyebrows shape.

3.1 Cooperation Analysis

The approach we have used to analyze cooperation in dialogue task is mainly based on Bethan Davies model (Bethan Davies, 2006). The basic coded unit is the “move”, which means individual linguistic choices to successfully fulfill Map Task. The idea of evaluating utterance choices in relation to task success can be traced back to Anderson and Boyle (1994) who linked utterance choices to the accuracy of the route performed on the map. Bethan Davies extended the meaning of “move” to the goal evaluation, from a narrow set of indicators to a sort of data-driven set. In particular, Bethan Davies stressed some useful points for the computation of collaboration between two communicative partners:

- social needs of dialogue*: there is a minimum “effort” needed to keep the conversation going. It includes minimal answers like “yes” or “no” and feedbacks. These brief utterances are classified by Bethan Davies (following Traum, 1994) as low effort, as they do not require much planning to the overall dialogue and to the joint task;
- responsibility of supplying the needs of the communication partner*: to keep an utterance going, one of the speakers can provide follow-ups which take more consideration of the partner’s intentions and goals in the task performance. This involves longer utterances, and of course a larger effort;
- responsibility of maintaining a known track of communication or starting a new one*: there is an effort in considering the actions of a speaker within the context of a particular goal: that is, they mainly deal with situations where a speaker is reacting to the instruction or question offered by the other participant, rather than moving the discourse on another goal. In fact the latter

is perceived as a great effort as it involves reasoning about the task as a whole, beside planning and producing a particular utterance.

Following Traum (1994), speakers tend to engage in lower effort behaviors than higher ones. Thus, if you do not answer to a question, the conversation will end, but you can choose whether or not to query an instruction or offer a suggestion about what to do next. This is reflected in a weighting system where behaviors account for the effort invested and provides a basis for the empirical testing of dialogue principles. The use of this system provides a positive and negative score for each dialogue move. We slightly simplified the Bethan Davies’ weighting system and propose a system giving positive and negative weights in an ordinal scale from +2 to -2. We also attribute a weight of 0 for actions which are in the area of “minimum social needs” of dialogue. In Table 1 we report some of the dialogue moves, called cooperation type, and the corresponding cooperation weighting level. There is also a description of different type of moves in terms of Grice’s conversational rules breaking or following. Due to the nature of the map task, where giver and a follower have different dialogue roles, we have two slightly different versions of the cooperation annotation scheme. For example “giving instruction” is present only when annotating the giver cooperation. On the other hand “feedback” is present in both annotation schemes. Other communicative collaboration indexes we codify in our coding scheme are the presence or absence of eye contact through gaze direction (to the interlocutor, to the map, unfocused), even in full screen condition, where the two speakers can’t see each other. Dialogue turns management (turn giving, turn offering, turn taking, turn yielding, turn concluding, and feedback) has been annotated as well. Video clips have been orthographically transcribed. To do so, we adopted a subset of the conventions applied to the transcription of the speech corpus of the LUNA project corpus annotation (see Rodriguez et al., 2007).

3.2 Coding Procedure and Kappa Scores

Up to now we have annotated 9 emotive tokens of an average length of 100 seconds each. They have been annotated with the coding scheme previously described by 6 annotators. Our coding scheme has been implemented into ANVIL software (Kipp, 2001). A Fleiss’ kappa statistic (Fleiss,

1971) has been performed on the annotations. We choose Fleiss’ kappa as it is the suitable statistics when chance agreement is calculated on more than two coders. In this case the agreement is expected on the basis of a single distribution reflecting the combined judgments of all coders.

Cooperation level	Cooperation type
-2	No response to answer: breaks the maxims of quality, quantity and relevance
-2	No information add when required: breaks the maxims of quality, quantity and manner
-2	No turn giving, no check: breaks the maxims of quality, quantity and relevance
-1	Inappropriate reply (no giving info): breaks the maxims of quantity and relevance
0	Giving instruction: cooperation baseline, task demands
1	Question answering y/n: applies the maxims of quality and relevance
1	Repeating instruction: applies the maxims of quantity and manner
2	Question answering y/n + adding info: applies the maxims of quantity, quality and relevance
2	Checking the other understands (ei sei? Capito?): applies the maxims of quantity, quality and manner
2	Spontaneous info/description adding: applies the maxims of quantity, quality and manner

Table 1: Computing cooperation in our coding scheme (from Bethan Davies, 2006 adapted)

Thus, expected agreement is measured as the overall proportion of items assigned to a category k by all coders n.

Cooperation annotation for giver has a Fleiss’ kappa score of 0.835 ($p < 0.001$), while for follower cooperation annotation is 0.829 ($p < 0.001$). Turn management has a Fleiss kappa score of 0.784 ($p < 0.001$). As regard gaze, Fleiss kappa score is 0.788 ($p < 0.001$). Mouth shape annotation has a Fleiss kappa score of 0.816 ($p < 0.001$) and eyebrows shape annotation has a Fleiss kappa of 0.855 ($p < 0.001$). In the last years a large debate on the interpretation of kappa scores has wide-spread. There is a general lack of consensus on how to interpret those values. Some authors (Allwood et al., 2006) consider as reliable for multimodal annotation kappa values between 0.67 and 0.8. Other authors accept as reliable only scoring rates over 0.8 (Krippendorff, 2004) to allow some generalizations. What is clear is that it seems inappropriate to propose a general cut off point, especially for multimodal annotation where very little literature on kappa agreement has been reported. In this field it seems more necessary that researches report clearly the method they apply (e. g. the number of coders, if they code independently or not, if their coding relies only manually).

Our kappa scores are very high if compared with other multimodal annotation results. This is because we analyze cooperation and emotion with an unambiguous coding scheme. In particular, we do not refer to emotive terms directly. In fact every annotator has his/her own representation of a particular emotion, which could be pretty different from the one of another coder. This representation will represent a problem especially for annotation of blended emotions, which are ambiguous and mixed by nature. As some authors have argued (Colletta et al., 2008) annotation of mental and emotional states is a very demanding task. The analysis of non verbal features requires a different approach if compared with other linguistics tasks as multimodal communication is multichannel (e.g. audiovisual) and has multiple semantic levels (e.g. a facial expression can deeply modify the sense of a sentence, such as in humor or irony).

The final goal of this research is performing a logistic regression on cooperation and emotion display. We will also investigate speakers' role (giver or follower) and screen/no screen conditions role with respect to cooperation. Our predictions are that in case of full screen condition (i. e. the two speakers can't see each other) the cooperation will be lower with respect to short screen condition (i. e. the two speakers can see each other's face) while emotion display will be wider and more intense for full screen condition with respect to short barrier condition. No predictions are made on the speaker role.

4 Conclusions and Future Directions

Cooperative behavior and its relationship with emotions is a topic of great interest in the field of dialogue annotation. Usually emotions achieve a low agreement among raters (see Douglas-Cowie et al., 2005) and surprisingly emotion recognition is higher in a condition of modality deprivation (only acoustic or only visual vs. bimodal).

Neuroscience research on emotion shows that emotion recognition is a process performed firstly by sight, but the awareness of the emotion expressed is mediated by the prefrontal cortex. Moreover a predefined set of emotion labels can influence the perception of facial expression. Therefore we decide to deconstruct each signal without attributing directly an emotive label. We consider promising the implementation in computational coding schemes of neuroscience evidences on transmitting and decoding of emotions. Further researches will implement an experiment

on coders' brain activation of to understand if emotion recognition from face is a whole or a part based process.

References

- Allwood J., Cerrato L., Jokinen K., Navarretta C., and Paggio P. 2006. A Coding Scheme for the Annotation of Feedback, Turn Management and Sequencing Phenomena. In Martin, J.-C., Kühnlein, P., Paggio, P., Stiefelbogen, R., Pianesi, F. (Eds.) *Multimodal Corpora: From Multimodal Behavior Theories to Usable Models*: 38-42.
- Anderson A., Bader M., Bard E., Boyle E., Doherty G. M., Garrod S., Isard S., Kowtko J., McAllister J., Miller J., Sotillo C., Thompson H. S. and Weinert R. 1991. The HCRC Map Task Corpus. *Language and Speech*, 34:351-366
- Anderson A. H., and Boyle E. A. 1994. Forms of introduction in dialogues: Their discourse contexts and communicative consequences. *Language and Cognitive Process*, 9(1):101 - 122
- Anderson J. C., Linden W., and Habra M. E. 2005. The importance of examining blood pressure reactivity and recovery in anger provocation research. *International Journal of Psychophysiology* 57(3): 159-163
- Argyle M. and Cook M. 1976 *Gaze and mutual gaze*, Cambridge: Cambridge University Press
- Bethan Davies L. 2006. Testing Dialogue Principles in Task-Oriented Dialogues: An Exploration of Cooperation, Collaboration, Effort and Risk. In *University of Leeds papers*
- Brennan S. E., Chen X., Dickinson C. A., Neider M. A. and Zelinsky J. C. 2008. Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition* 106(3): 1465-1477
- Ekman P. and Friesen WV. 1978. *FACS Facial Action Coding Scheme. A technique for the measurement of facial action*, Palo Alto, CA: Consulting Press
- Carletta, J. 2007. Unleashing the killer corpus: experiences in creating the multi-everything AMI Meeting Corpus, *Language Resources and Evaluation*, 41: 181-190
- Colletta, J.-M., Kunene, R., Venouil, and A. Tcherkasof, A. 2008. Double Level Analysis of the Multimodal Expressions of Emotions in Human-machine Interaction. In Martin, J.-C., Patrizia, P., Kipp, M., Heylen, D., (Eds.) *Multimodal Corpora: From Models of Natural Interaction to Systems and Applications*, 5-11
- Craggs R., and Wood M. 2004. A Categorical Annotation Scheme for Emotion in the Linguistic Content of Dialogue. In *Affective Dialogue Systems*, Elsevier, 89-100

- Douglas-Cowie E., Devillers L., Martin J.-C., Cowi R., Savvidou S., Abrilian S., and Cox C. 2005. Multimodal Databases of Everyday Emotion: Facing up to Complexity. In *9th European Conference on Speech Communication and Technology (Interspeech'2005)* Lisbon, Portugal, September 4-8, 813-816
- Feldman Barrett L., Lindquist K. A., and Gendron M. 2007. Language as Context for the Perception of Emotion. *Trends in Cognitive Sciences*, 11(8): 327-332.
- Fleiss J. L. 1971. Measuring Nominal Scale Agreement among Multiple Coders *Psychological Bulletin* 11(4): 23-34.
- Goeleven E., De Raedt R., Leyman L., and Verschuere, B. 2008. The Karolinska Directed Emotional Faces: A validation study, *Cognition and Emotion*, 22:1094 -1118
- Kendon A. 1967. Some Functions of Gaze Directions in Social Interaction, *Acta Psychologica* 26(1):1-47
- Kipp M., Neff M., and Albrecht I. 2006. An Annotation Scheme for Conversational Gestures: How to economically capture timing and form. In Martin, J.-C., Kühnlein, P., Paggio, P., Stiefelhofen, R., Pianesi, F. (Eds.) *Multimodal Corpora: From Multimodal Behavior Theories to Usable Models*, 24-28
- Kipp M. 2001. ANVIL - A Generic Annotation Tool for Multimodal Dialogue. In *Eurospeech 2001 Scandinavia 7th European Conference on Speech Communication and Technology*
- Krippendorff K. 2004. Reliability in content analysis: Some common misconceptions and recommendations. *Human Communication Research*, 30:411-433
- Magno Caldognetto E., Poggi I., Cosi P., Cavicchio F. and Merola G. 2004. Multimodal Score: an Anvil Based Annotation Scheme for Multimodal Audio-Video Analysis. In Martin, J.-C., Os, E.D., Kühnlein, P., Boves, L., Paggio, P., Catizone, R. (eds.) *Proceedings of Workshop Multimodal Corpora: Models Of Human Behavior For The Specification And Evaluation Of Multimodal Input And Output Interfaces*. 29-33
- Martin J.-C., Caridakis G., Devillers L., Karpouzis K. and Abrilian S. 2006. Manual Annotation and Automatic Image Processing of Multimodal Emotional Behaviors: Validating the Annotation of TV Interviews. In *Fifth international conference on Language Resources and Evaluation (LREC 2006)*, Genoa, Italy
- Pianesi F., Leonardi C., and Zancanaro M. 2006. Multimodal Annotated Corpora of Consensus Decision Making Meetings. In Martin, J.-C., Kühnlein, P., Paggio, P., Stiefelhofen, R., Pianesi, F. (Eds.) *Multimodal Corpora: From Multimodal Behavior Theories to Usable Models*, 6--9
- Poggi I., 2007. *Mind, hands, face and body. A goal and belief view of multimodal communication*, Berlin: Weidler Buchverlag
- Reidsma D. Heylen D., and Op den Akker R. 2008. On the Contextual Analysis of Agreement Scores. In Martin, J.-C., Patrizia, P., Kipp, M., Heylen, D., (Eds.) *Multimodal Corpora: From Models of Natural Interaction to Systems and Applications*, 52--55
- Rodríguez K., Stefan K. J., Dipper S., Götze M., Poesio M., Riccardi G., and Raymond C., and Wisniewska J., 2007. Standoff Coordination for Multi-Tool Annotation in a Dialogue Corpus. In *Proceedings of the Linguistic Annotation Workshop at the ACL'07 (LAW-07)*, Prague, Czech Republic.
- Smith M. L., Cottrell G. W., Gosselin F., and Schyns P. G. 2005. Transmitting and Decoding Facial Expressions. *Psychological Science* 16(3):184-189
- Tassinari L. G. and Cacioppo J. T. 2000. The skeleto-motor system: Surface electromyography. In LG Tassinari, GG Berntson, JT Cacioppo (eds) *Handbook of psychophysiology*, New York: Cambridge University Press, 263-299
- Traum D. R. 1994. A Computational Theory of Grounding in Natural Language Conversation, PhD Dissertation. urresearch.rochester.edu