Double level analysis of the Multimodal Expressions of Emotions in Human-Machine Interaction

MultiModal Corpora Symposium
Marrakech, May 27th, 2008

Jean-Marc Colletta [1], Ramona Kunene [1], Aurélie Venouil [1], Anna Tcherkassof [2]

[1]Lidilem, Université Stendhal, BP25, 38040 Grenoble Cedex 9, France
[2] LIP-2CS, Université Pierre Mendès France, BP47, 38040 Grenoble Cedex 9, France

e-mail: Jean-marc.colletta@u-grenoble3.fr

ACADEMIE DE GRENOBLE

Sciences du Langage LIDILEM
UNIVERSITE Stendhal GRENOBLE 3

Introduction

Today, there is a growing interest for the multimodal complexity processes of oral communication

The crucial issue is to take into account the multimodal expressions of thought, emotions, mental states, and to annotate for these clues

We present an annotation device designed for a study to improve the detection and characterisation of expressions of emotions in human-machine interaction
>>> Le Chenadec, Maffiolo, & Chateau, 2007; Le Chenadec, Maffiolo, Chateau & Colletta, 2007

Our talk is in four parts:

1. Theoretical considerations
2. The background of our annotation work
3. The coding scheme
4. Validation considerations

Theoretical questions (1) : language, emotions & mental states

- Linguistic attributes often betray the emotional as well as the mental state of the speaker's mind

  >>> Galati & Sini, 2000; Plantin, 2003; Tutin, Novakova, Grossmann & Cavalla, 2006

- In addition, all aspects of prosody may contribute to express emotional and mental states: pitch, intensity, speech rate, hesitations, grunts, etc.

  >>> Keller et al., 2003; Scherer, Bänziger & Greandjean, 2003; Shochi, Aubergé & Rilliard, 2006; Shochi, 2008

- But language does not wholly and directly express emotions and mental states:

- Social context, social relationships, social constraints, language task parameters lead us to control our verbal and vocal expression of emotions and mental states

  >>> Ekman, 1989; Goffman, 1973, 1974; Caffi & Jeanney, 1994

- The verbal and vocal contribute to the expression of emotions and mental states together with body language cues

  >>> Cosnier, 1994, Tcherkassof, 1997

Theoretical questions (2) : gesture, emotions & mental states

- All bodily movements may help express attitudes, emotional and mental states:
- Attitudes and postures were correlated to mental states, pathologies and emotional dispositions

  >>> Feyereisen & De Lannoy, 1985; Feldman & Rimé, 1991; Cosnier, 1994
- The gaze contributes to the expression of emotion and its appearance is correlated to levels of awareness

  >>> Brossard, 1992
- The facial expressions exteriorise the whole range of emotions and feelings

  >>> Ekman, 1989; Feldman & Rimé, 1991
- Some gestures (self centred gestures, adaptators) appear more frequently in stressful situations and are correlated to anxious states and certain affects

  >>> Mahl, 1967; Scheflen, 1973

Theoretical questions (3) : the ETIC/EMIC approach of body behaviour

- Pike's distinction between ETIC / EMIC approaches of body language
  >>> Pike, 1967
- The ETIC approach emphasizes the physical aspects of body movements and relies on microanalytical descriptions
- The EMIC approach emphasizes the symbolic aspects of body movements and relies on comprehensive macroanalytical descriptions
- The ETIC approach is useful for gesture studies and research in computer sciences
  >>> Calbris, 2003; Kendon, 2004; Mancini & Pelachaud, 2005; Kipp et al., 2006, 2007
- The EMIC approach is useful for linguistics and psychology studies of multimodal communication and the relationship between speech and gesture
  >>> McNeill, 1992, 2005; Pléty, 1993; Bouvet, 2001; Goldin-Meadow, 2003; Colletta, 2004
- Both are relevant for the detection and characterisation of expressions of emotions and mental states as well as for the creation of Embodied Conversational Agents
  >>> see the BML/FML distinction: Kopp et al. 2006; Vilhjalmsson et al., 2007

The background of our annotation work (1)

- We present an annotation device designed for a study on the fusion of multimodal information to improve the detection and characterisation of expressions of emotions in human-machine interaction >>> *Le Chenadec, Maffiolo & Chateau, 2007*.

- Main objectives: to collect a wide range of multimodal expressions of emotional and mental states expressed freely and spontaneously.

- Data collected during an experiment based on the Wizard-of-Oz methodology:
  - interaction between a human and a virtual character
  - repetition of a play (three scenes of *Don Quixote de la Mancha)*
  - each subject had to give his cue to the virtual character as Don Quixote.
  - cues from the virtual actor controlled by the experimenter in real time
  - simulation of an autonomous system.

- In order to elicit spontaneous emotional expressions of users, different system bugs were designed

- Some bugs were clearly related to a system's failure, other bugs were perceived by the subjects to be a result of their mistakes of their cues

- All subjects were expected to express the emotional feelings and mental states they experienced as a result of being confronted with these bugs

The background of our annotation work (2)

- Eighteen actors took part in the experiment
- Interview recordings lasted 1h15mn for each participant
- Multimodal behaviour of each participant recorded with two digital cameras (head-only and upper body) + a microphone
- Additional data :
  - Immediately after the interaction with the virtual character, each subject was asked to comment what he/she felt.
  - A subsequent interview was conducted with a close relative of each subject who was asked to comment on the behaviour of the user using the same method.
  - A categorisation experiment was conducted with 90 third-party observers who had to attribute an emotional or cognitive value to the observed behaviour and indicate his/ her starting and ending time.

Annotation grid for the verbal + prosodic aspects (PRAAT) (1)

Linguistic	Commentary	commentary on the interaction
→	Error	error in syllable or word pronounciation
→	Unexpected articulation	pronounciation of an unsual final syllable with a silent "e.
→	False start	a "*" attached to the word + annotate the complete word sequence
	Elision	presence of elision
→	Recovery	reformulation of a portion of a speech segment
	Repetition	repetition of a portion of a speech segment
→	Incomprehensible words	transcription of an impossible word or speech segment
Dialogue	Repétition	repetition of the identical
	Reformulation	repetition of response with other terms

Annotation grid for the verbal + prosodic aspects (PRAAT) (2)

Prosody	Silent pause	pause in the middle of a speech segment
	Intelligible pause	silence voluntarily added in the middle of a speech segment
	Pause filler	"euh" ou "hum"
Sounds	Sounds of the system	
	Speech cuts	the virtual actor cuts the live actor's speech
	cough, throat, mouth	cough, throat clearing, noise made by the mouth
	Laugh	
	Exhalation, breath, sigh	
	Inhalation	

Annotation grid for the non verbal aspects: principles (1)

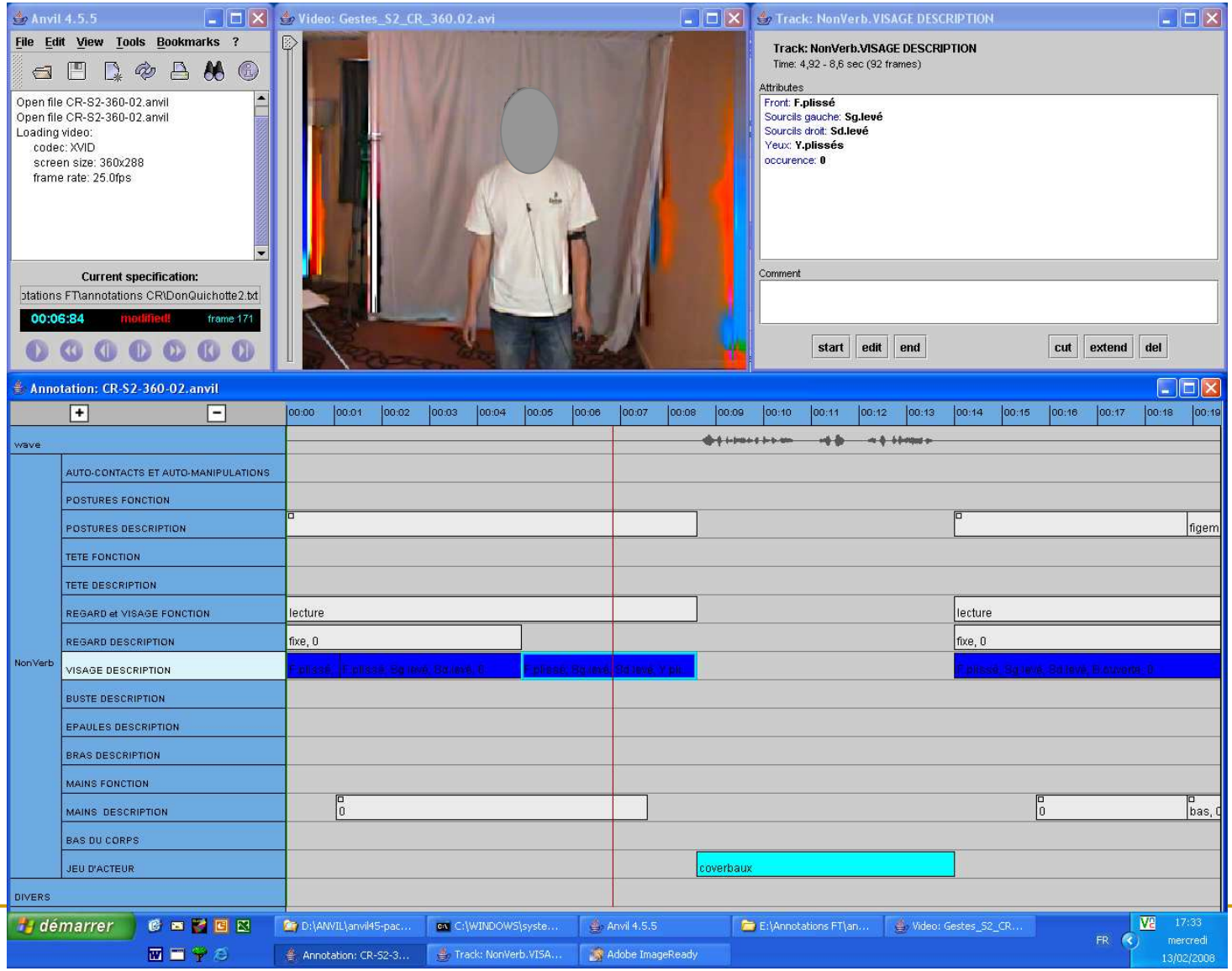
- Divided in **16 tracks** representing all parts of the human body:
  - Self-centered gestures and manipulation movements
  - Posture attitudes and changes (2 tracks)
  - Head gestures (2 tracks)
  - Gaze direction and changes (2 tracks)
  - Facial expressions (2 tracks)
  - Torso movements
  - Shoulders movements
  - Arms location and movements
  - Hand gestures (2 tracks)
  - Lower body movements
  - Gestures performed while acting

Annotation grid for the non verbal aspects: principles (2)

- **ETIC parameters** (11 tracks):
  - Body part and location of movement
  - Direction
  - Characteristic
  - Shape
  - Speed
  - Frequency of occurrence

- **EMIC considerations** (5 tracks):
  - General behaviour or attitude
  - Significant head movement
  - Gaze behaviour
  - Significant facial expression
  - Coverbal hand gesture

Illustration of the ANVIL window



Annotation grid for the non verbal aspects (1)

Type of annotation	**Name of phenomenon**
1- Self-contact gestures & auto-manipulations	Action: scratch/ touch/ twist/ handle
	Body part location: hair/temple/brow/glasses/ nose
2- Posture (function)	Comfort/ stretching
3- Posture (description)	Pattern : swaying/ complex movement/ freezing
	Leg movements: frontward/ backwards/ left/ right
	Speed: slow/ normal/ fast

Annotation grid for the non verbal aspects (2)

4- Head (function)	Movement : nod/ shake/ beat/ deictic
5- Head (description)	Tilted high/ low
	Turn: left/ right
	Complex movement: front / backward
	Single movement: up / down
	Single movement: front / backward
	Single tilt: left/ right
	Single side-turn: left/right
	Speed: slow/ normal/ fast

Annotation grid for the non verbal aspects (3)

6- Gaze (function)	Characterisation: waiting/ reading/ staring/ scanning
7- Gaze (description)	Direction: up/ down
	Direction: left/right
	Movement: sweeping/ rolling eyes
	Speed: slow/ normal/ fast
8- Face (function)	Smile, laughter/ biting/ pursing/ licking lips/pouting
9- Face (description)	Brows: frowning
	Left eyebrow: raising / frowning
	Right eyebrow: raising/ frowning
	Eyes: closing / opening/ wide opening/ rolling/ blinking/ winking

Annotation grid for the non verbal aspects (4)

10- Torso (description)	Movement: forward/ backward
	Movement: left/right
	Unsteady movement
	Bend: forward/ backward
	Turn:left/ right
	Twist: left/ right
	Side: left/ right
	Position: protruded/ retracted
	Speed: slow/ normal/ fast
11- Shoulders (description)	Identification: left/ right/both
	Description: shrugging/ sagging
	Number: left/ right/ both
	Occurrence: 0 to 5
	Speed: slow/normal/fast

Annotation grid for the non verbal aspects (5)

12- Arms (description)	Left-arm direction: going up/down, moving sideways, forwards, backwards, to the side, up, not moving
	Left-arm position: bent, half-bent, stretched out 
	Right-arm direction: going up / down, moving sideways, forwards, backwards, to the side, up, not moving
	Right-arm position: bent, half-bent, stretched out
	Both arms action: crossing
	Occurrence: 0 to 5 
	Speed: slow/ normal/ fast

Annotation grid for the non verbal aspects (6)

13- Hands (function)	Deictic, beat, iconic, metaphoric, interactive
14- Hands (description)	Left hand action: rotation, opening, closing
	Left-hand direction: up/ down/left/ right/ forward/ backward
	Left-palm direction: Left-hand direction: up/ down/left/ right/ forward/ backward
	Right hand action: rotation, opening, closing
	Right-hand direction: up/ down/left/ right/ forward/ backward
	Right-palm direction: up/ down/left/ right/ forward/ backward
	Occurrence: 0 to 5
	Speed: slow/ normal/ fast

Annotation grid for the non verbal aspects (7)

15- Lower body	Free comments
16- Acting	Mime, exaggerated gestures and expressions

The non verbal annotation validation process

Main objective: to verify and validate the non verbal annotations

The non verbal transcription and annotation was carried out by **two independent coders**

A **third coder** had to finalise the annotation from choices made by both coders, and decide in case of disagreement

A two-stage process (double independent coding + decision stage) cannot ensure that the analysis procedure is a hundred percent conclusive.

To annotate for emotions and mental states is to observe the whole body, including the problem of identifying the movements, which does not arise when we annotate for precise gestures (e.g., the coverbal hand gestures)

In conclusion

The quest for emotions and human mental state expressions shows that each and every part of the body has a story to reveal.

The grid used on *ANVIL* enabled us to annotate this rather complex set of facial expressions, posture changes, gaze changes, self-centred gestures, head gestures, etc.

Our analysis procedure aimed at using the double level (*ETIC / EMIC*) annotation.

The missing puzzle remains in the cross-validation from several data sources: next stage will consist in comparing our annotated data to the results from the categorisation experiment conducted with 90 third-party observers…

Aknowledgements

This research was conducted and financed by the France Telecom R & D, Lannion.

The authors thank Valérie Maffiolo and Gilles Le Chenadec for the designing of the experiment and for contributing to the creation of the annotation grid.