The Amount of Information on Emotional States Conveyed by the Verbal and Nonverbal Channels: Some Perceptual Data

Anna Esposito[1,2]

[1] Dipartimento di Psicologia, Seconda Università di Napoli
Via Vivaldi 43, 81100 Caserta, Italy
[2] IIASS, Via Pellegrino 19, 84019, Vietri sul Mare, SA, Italy
iiass.annaesp@tin.it, anna.esposito@unina2.it

Abstract. In a face-to-face interaction, the addressee exploits both the verbal and nonverbal communication modes to infer the speaker's emotional state. Is such an informational content redundant? Is the amount of information conveyed by each communication mode the same or is it different? How much information about the speaker's emotional state is conveyed by each mode and is there a preferential red communication mode for a given emotional state? This work attempts to give an answer to the above questions evaluating the subjective perception of emotional states in the single (either visual or auditory channel) and the combined channels (visual and auditory). Results show that vocal expressions convey the same amount of information as the combined channels and that the video alone conveys poorer emotional information than the audio and the audio and video together. Interpretations of these results (that seem to not support the data reported in literature proving the dominance of the visual channel in the emotion's perception) are given in terms of cognitive load, language expertise and dynamicity. Also, a mathematical model inspired to the information processing theory is hypothesized to support the suggested interpretations.

1 Introduction

It has been and remains difficult to define emotions, and this difficulty continues, even though there have been a lot of attempts to characterize emotional states. The proposed approaches to understand emotions are centered on finding insights on what emotions are, but the question still remains unanswered. James [44] was the first to posit the question and to indicate the bodily approach as a first coherent approach to explain emotions. According to James, emotions are the feeling of <<…*bodily changes that follow directly the perception of the exciting fact…*>>, whereas in Darwin's opinion [16] they are patterns of actions deriving from our evolutionary or individual past and may be <<…*of the least use…*>> in our modern context. Along the past centuries, the social role of emotions has been differently interpreted and they have been either considered as conflicting aspects of soul (Plato, 450 B.C.[60]) or essential in the context of the Ethic (Aristotele, 384 B.C.[5]) as well as disturbing passions (Descartes, 1649 [38]) or as playing a central role in the human behaviour

Y. Stylianou, M. Faundez-Zanuy, A. Esposito (Eds.): WNSP 2005, LNCS 4391, pp. 249 – 268, 2007.

[64]. Theory of emotions tends to privilege inner and psychological motivations but modern anthropologists conceive emotions as socially embedded responses that <<*…take on behavioural significance within a field of culturally interpreted person-person and person-situations relations...* [80]>>. During the last century, the study of emotions involved different scientific fields and today the effort among psychologists, neurologists, anthropologists, and moral philosophers is to converge towards a holistic theory of emotion. However, there is still the need to define emotions and below a working definition is proposed.

1.1 A "Working Definition"

A working definition is not "*the definition*". It is just a way to deal with a phenomenon that is not yet entirely understood to provide an orientation. Such a definition can be changed as soon as new relevant explanations are discovered. On this premise, the central condition of emotions is considered to be the evaluation of an event. A person is consciously or unconsciously evaluating an external or internal event that is relevant to him or to some of his concerns or goals. This is what the psychologists call the appraisal phase. At the core of this evaluation (or concurrently with this evaluation) there is a "readiness to act" and the "prompting of plans" [32-33] that allow the person to handle the evaluated event producing appropriate mental states accompanied by physiological changes, feelings, and expressions. This process can be schematized in the diagram reported in Figure 1 which is a personal interpretation of the feedback loop proposed by Plutchik [61-62].

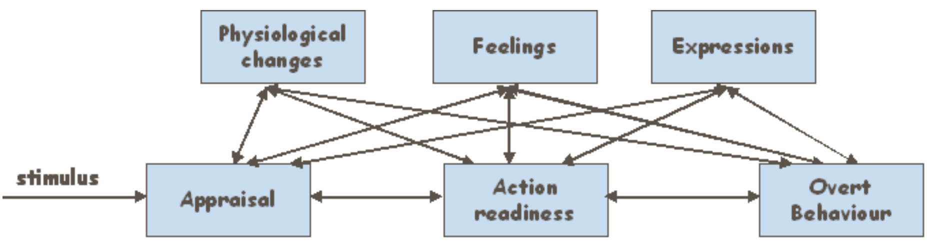

Fig. 1. A personal interpretation of the feedback loop proposed by Plutichik [61]

The above interpretation assumes that after the evaluation of an event there are several processes that take place concurrently to the readiness of action. They all affect each other and it is possible that none of them has a primary role. They are all the result of mental processes that take over on the current cognitive state to handle the subject's emotional state. This interpretation is in accord with the largely accepted definition of emotions proposed, among others, by Plutchik [61-62], Frijda [32-33], and Scherer [66, 70], even though it may differ by not assigning a priority to any of them. Our idea is that, being emotions experienced as a comprehensive manifestation of feelings, expressions, physiological changes, action readiness, and prompting of plans, assigning a subsidiary role to any of these signals would produce a restricted interpretation of their essence as a whole.

According to many authors (see also Oatley and Jenkins [58]) emotions have several functional roles, among these the most important in driving social interaction are:

1. They solicit appropriate reactions to events that happen unexpectedly and demand immediate action;
2. They act as signs (exhibited through a variety of comportments) made for the purpose of giving (consciously or unconsciously) notice of a plan of future intents;
3. They play a communicative function.

However, the above functions, do not completely define the role of emotions and why they should exist. At the date, a unique scientific theory of emotions is not yet available. Scientists are debating on divergent opinions ranging from evolutionary vs. socio-cultural theories [12-13, 22, 24, 29-30, 32-33, 40-41, 44, 48-51, 59, 61-62, 70]. It is not in the aims of this work to discuss on them and an interested reader can find very fascinating highlights in the book of Oatley and Jenkins [58]. Our primary aim is on the expressions of emotions, on the perceptual features exploited to recognize them from facial expressions, voice, and body movements, and on trying to establish if there is a preferential channel for perceiving an emotional state. These will be the arguments discussed in the following sections.

2 The Expressions of Emotions

We are not sure that the term "expressions" is clearly identifying what we call the "expressions of emotions". This is because, as stated above first, expressions cannot be separated from emotions as secondary processes, and second, in our opinion, also physiological changes and feelings are "expressions" of emotions. However, in this context, the term "expressions" is used to indicate what is perceived on the emotional states in the face-to-face interaction. In this sense, we restricted ourselves to consider the perceptual appearance of emotional states, i.e. the facial, the vocal, and the gestural expressions of emotions. Two of them, the facial, and the vocal, have received a lot of attention by many authors, while gestures have been, at the date, partially neglected. Verbal communication is not considered as an expression of emotions. Firstly, because it seems to involve less immediate and different cognitive processes; and secondly, because the expression of emotions through the verbal channel meets lexical difficulties due to the fact that often subjects can hardly label or recognize their emotional state. We will summarize the main findings about the expressions of emotions in these three domains and we will present some interesting perceptual data on the emotional information conveyed by the visual and the auditory channels.

2.1 The Facial Expression of Emotions

Facial expressions as markers of emotions were first suggested by Darwin [16]. Darwin approached the study of facial expressions of emotions from a biological point of view considering them as innate "ancient habits" that can partially be modified by learning and imitation processes. This idea was debated by several authors, among whom the most representatives were Klineberg [48], La Barre [50], and Birdwhistell

[7], and more recently by White and Fridlund [29-30, 80], according to whom any behaviour, and therefore also facial expressions of emotions, are learned.

Darwin's approach was recovered and reinforced about a century later by Tomkins [77-79], which posits the basis for the research on the instrumental role of facial movements and facial feedbacks in experiencing emotions. According to Tomkins, emotions are the driving forces of our primary motivational system acting as amplifiers of what we receive from the environment, and our face is the most important expression of them. This is because facial expressions of emotions are universally shared. People from different cultures can easily recognize a happy or a sad face because of an "innate affective program" for each of what is considered "a primary emotion" that acts on the facial musculature to produce emotional facial expressions. On this research line, Ekman [21-23] and Izard [39-41] identified a small set of emotions (the primary ones) associated with a unique configuration of facial muscle movements and both developed some anatomically based coding systems (the Facial Action Coding System, (FACS) [20, 25-26], the Maximally Discriminative Facial Movement Coding System (MAX) and AFFEX [42-43] for measuring facial behaviours and identifying emotional expressions. The three systems are different both in the primary emotions they include (only six for Ekman - happiness, sadness, anger, fear, surprise, and disgust; nine for Izard) and in the expressions they assign to some of them. Sophisticated instruments, such as facial EMG (Electromyography) were used [12-13, 72-74] to asses the above facial coding systems when emotional information was not visually perceptible in facial actions. Nevertheless, such studies proved that distinction among primary emotions or more generally among negative and positive emotions, was not possible using either visible or EMG information on facial expressions.

At the date, the role of facial expressions of emotions in identifying and recognizing emotions is debated among two theories:

1. The "micro-expression hypothesis" which assumes that emotion-specific facial expressions do exist even when they are not visually perceptible and they emerge as visible expressions as the emotional intensity increases [72-74]. This means that an innate program that controls our facial musculature does exist and it changes our face when we are emotionally involved;
2. The "ethological perspective" that assumes that facial expressions are not directly related to emotional states, being signals associated to the communicative context [29-30]. The essence of this theory is that "we make faces". They are contextual to the moment and the particular situation. Facial expressions may have nothing to do with our emotional states.

2.2 The Vocal Expressions of Emotions

Emotions produce changes in respiration, phonation, and articulation which in turn affect vocalizations and the acoustic parameters of the corresponding signal [2, 4, 6, 8, 15, 17-19, 31, 34, 67-69, 71]. Measurements of emotional speech are derived from three perceptual cues such as loudness, pitch, and timing, which in turn are expressed in changes of acoustic variables such as: amplitude (which is the measure of sound pressure), and sound intensity (which is the measure of sound energy at different frequencies) for loudness; the fundamental frequency (F0) of the signal (measured in Hz), and some derived measures on F0 for pitch; and speaking rate, utterance and syllable lengths, empty and filled pauses for timing. The search for acoustic features that describe emotional states in speech is performed through long-term or supra-segmental measurements since emotions are expected to last more than single speech segments and therefore, sophisticated acoustic measurements have been proposed to detect vocal expressions of emotions. Apart from F0 frequency [45, 56], these measurements include the spectral tilting [45], the vocal envelope [71] - defined as the time interval for a signal to reach the maximum amplitude and decay to zero amplitude - the Long Term Average Spectrum (LTAS), the energy in different frequency bands, some inverse filtering measurements [47, 57], the F0 contour, the F0 maximum and minimum excursion, and the F0 jitter - random fluctuations on F0 values [2, 28, 63]. Exploiting this research line, the cited authors identified some acoustic correlates of vocal expressions for some primary emotions (see [63, 67]). However, these studies have shown that the perceptual and acoustic cues to infer emotional states from speech may be complex as much as using physiological variables and facial expressions, and that the acoustic and temporal features that can be used to unequivocally describe emotions have not yet been identified because of the many sources of variability associated with the speech signal.

2.3 The Gestural Expressions of Emotions

Gesturing is what people unconsciously do when approaching each other and understanding gestures is crucial to understand social interactions. The Oxford Dictionary of English defines a gesture as “a movement of the body, or any part of it, that is considered as expressive of thought or feeling”. This definition incorporates a very large set of body expressions and movements among these facial expressions but also face directions, hand movements, eye contact, posture, proxemics (study on varying patterns of physical proximity), touch, dressing, and so on. Therefore, excluding facial expressions, there is a lot that has been neglected about emotional gestures. Three fundamental questions remain uncovered in this field:

- What can gestures tell about human emotions?
- Which gestures are exhibited for a particular emotion?
- Do humans make use of such non verbal cues to infer a person’s emotional state and how?

At present there is limited knowledge about the extent to which body movements and gestures provide reliable cues to emotional states even though several authors have suggested that the use of body motion in interaction is part of the social system and therefore body movements not only have potential meaning in communicative contexts but can also affect the interlocutor’s behavior [3, 7, 9-10, 14, 36, 46, 52-55]. However, these studies were often concerned with gestures (in particular hand gestures) used jointly with verbal communication and therefore showing intimate relationships between the speaker’s speech and actions and the listener’s nonverbal behaviors [36, 46, 52]. These gestures may have nothing to do with emotions since according to Ekman and Friesen [20, 25-27] affective displays are independent from language. However, there is subjective evidence that variations in body movements (e.g. gait, body posture) and gestures (e.g. hand position) convey important information about people's emotions. To this aim, the author performed an informal perceptual experiment showing to about 50 people participating at a conference (Workshop on Nonlinear Speech Processing held in Crete, Greece, September 2005) two silent cartoon characters. One of them displayed a rapid body movement one step back, the other a rapid open hand movement that covered first the eyes and then the face. All the participants attributed to the characters an emotional state of fear and worry respectively, suggesting that they were able to use gestural information to infer emotional states.

At the date, relationships between bodily activity and emotional communicative functions remain a research field to be investigated even though several authors [2, 9-10, 14, 44, 54-55, 65] have drawn attention to the existence of relationships between body posture, head movements, hand gestures (in particular those named *adapters* by [27]) and emotions.

3 The Proposed Research Work

Two procedures have been employed to study emotional states:

- ❖ **The coding** procedure that investigates emotional features eliciting emotional states through actor portrayals, induction or clandestine recordings of naturally occurring emotions;
- ❖ **The decoding** procedure that evaluates the listeners' ability to recognize emotional states from speech, gestures and gaze expressions.

Whatever was the exploited procedure, emotions and the related perceptual cues to infer them, have always been investigated, to our knowledge, considering separately the three expressive domains discussed above, even though, some studies have suggested that one of them could be preferential with respect to the others. In particular, some studies sustain that facial expressions are more informative than gestures and vocal expressions [21, 23, 37, 39], whereas others suggest that vocal expressions are more faithful than facial expressions in expressing emotional states since physiological processes, such as respiration and muscle tension, are naturally influenced by emotional responses [4, 11, 66-69]. It should also be noted that while the power of the vocal expressions in expressing emotional states has been tested dynamically along the time dimension because speech is intrinsically a dynamic process, for the facial expressions, a long established tradition attempts to define the facial expression of emotion in terms of qualitative targets - i.e. static positions capable of being displayed in a still photograph. The still image usually captures the apex of the expression, i.e. the instant at which the indicators of emotion are most marked. However, in the daily experience, also emotional states are intrinsically dynamic processes and associated facial expressions are varying along time. Is dynamic visual information still emotionally richer than auditory information? To our knowledge, at the date, there are no data in literature that attempt to answer this question.

At the light of the above considerations, our research proposal is to use the decoding procedure to try to answer the following questions:

- ❖ Which of the two domains is most expressive of emotional states?
- ❖ Does the perception of the speaker's emotional state require both vocal and visual information or is emotional information "redundantly" transmitted both by the visual and the auditory channel?

- Is there any preferential channel and can it be culturally specific?
- Could a channel be preferential for an emotional state with respect to another in a given culture?
- Which channel is preferential for a foreign culture?

Answering these questions will allow to identify emotional features of some basic emotions through a cross-modal evaluation of the visual and auditory channel. The final goal is to identify which modality should be considerably enhanced (reducing computational costs) in the development of embodied conversational agents, i.e. computer interfaces represented by lifelike human or animal characters capable of performing believable actions and naturally reacting to human users.

To answer the above questions we set up a series of perceptual experiments that evaluated the subjective perception of emotional states in the single (either visual or auditory channel) and the combined channels (visual and auditory). We exploited video clips extracted from movies (in our case, Italian movies). This choice allowed to overcome two critiques generally moved to perceptual studies of the kind proposed: 1) the use of video clips avoided the stillness of the pictures in the evaluation of the emotional visual information; 2) even though the emotions expressed in the video clips were still simulations under studio conditions (and may not have reproduced a genuine emotion but an idealization of it) they were able to catch up and engage the emotional feeling of the spectators and therefore we were quite confident of their perceptual emotional contents.

3.1 Materials

The collected data are based on extracts from Italian movies whose protagonists were carefully chosen among actors and actresses that are largely acknowledged by the critique and considered capable of giving some very real and careful interpretations. The final database consists of audio and video stimuli representing 6 basic emotional states: *happiness*, *sarcasm/irony*, *fear*, *anger*, *surprise*, and *sadness*. We introduced sarcasm/irony to substitute the emotional state of disgust, since after one year of movie analysis only 1 video clip was identified for this emotion.

For each of the above listed emotional states, 10 stimuli were identified, 5 expressed by an actor and 5 expressed by an actress, for a total of 60 audio and video stimuli. The actors and actresses were different for each of the 5 stimuli to avoid bias in their ability to portray emotional states. The stimuli were selected short in duration (the average stimulus' length was 3.5s, SD = ± 1s). This was due to two reasons: 1) longer stimuli may produce overlapping of emotional states and confuse the subject's perception; 2) emotional states for definition cannot last more than a few seconds and then other emotional states or moods take place in the interaction [58]. Consequently, longer stimuli do not increase the recognition reliability and in some cases they can create confusion making the identification of emotions difficult, since in a 20 seconds long video clip, the protagonist may be able to express more than one and sometimes very complex emotions.

Care was taken in choosing video clips where the protagonist's face and the upper part of the body were clearly visible. Care was also taken in choosing the stimuli such that the semantic meaning of the sentences expressed by the protagonists was not clearly expressing the portrayed emotional state and its intensity level was moderate.

For example we avoided to include in the data, sadness stimuli were the actress/actor were clearly crying or happiness stimuli where the protagonist was strongly laughing. This was because we wanted the subjects to exploit emotional signs that could be less obvious but that were generally employed in every natural and not extreme emotional interaction. From each complete stimulus - audio and video - we extracted the audio and the video alone coming up with a total of 180 stimuli (60 stimuli only audio, 60 only video, and 60 audio and video).

The emotional labels assigned to the stimuli were given first by two expert judges and then by three naïve judges independently. The expert judges made a decision on the stimuli carefully exploiting emotional information on facial and vocal expressions described in the literature [4, 20-21, 25-26, 42-43, 66-68] and also knowing the contextual situation the protagonist is interpreting. The naïve judges made their decision after watching the stimuli several times. There were no opinion exchanges between the experts and naïve judges however, the final agreement on the labelling between the two groups was 100%. The stimuli in each set were then randomized and proposed to the subjects participating at the experiments.

3.2 Participants

A total of 90 subjects participated at the perceptual experiments: 30 were involved in the evaluation of the audio stimuli, 30 in the evaluation of the video stimuli, and 30 in the evaluation of the video and audio stimuli. The assignment of the subjects to the task was randomly. Subjects were required to carefully listen and/or watch the experimental stimuli via headphones in a quite room. They were instructed to pay attention to each presentation and decide as quickly as possible at the end, which of the 6 emotional states was expressed in it. Responses were recorded on a matrix paper form (60x8) where the rows listed the stimuli's numbers and the columns the 6 possible emotional states plus an option for any other emotion not listed in the form, plus the option that they recognized a *neutral* state in the presentation. Each emotional label given by the participants as an alternative to one of the six listed emotions was included in the emotional classes listed above only if criteria of synonymity and/or analogy were satisfied otherwise it was included in a class labelled "*others*".

For each emotional stimulus, we computed the frequency response distribution among the 8 emotional classes under examination (*happiness*, *sarcasm/irony*, *fear*, *anger*, *surprise*, *sadness, others,* and *neutral*) and the percentage of correct recognition.

3.3 Results

Table 1 reports the confusion matrix for the audio and video condition. The numbers are percentages computed over the number of the subject's correct answers to each stimulus and averaged over the number of the expected correct answers for each emotional state.

The data displayed in Table 1 show that, in the audio and video condition, the higher percentage of correct answers - 75% - was for **sadness** (in 14% of the cases confused with the *irony*), followed by 64% of correct answers for **irony** (in 12% of the cases confused with *surprise*). **Anger** reached 60% (confusion was made for 11%

Table 1. Confusion matrix for the audio and video condition. The numbers are percentages computed considering the number of subject's correct answers over the total number of expected correct answers (300) for each emotional state.

Audio and Video	*Sad*	*Ironic*	*Happy*	*Fear*	*Anger*	*Surprise*	*Others*	*Neutral*
Sad	**75**	14	0	2	2	0	4	3
Ironic	2	**64**	5	3	7	12	3	4
Happy	1	29	**50**	3	0	11	0	6
Fear	19	2	2	**48**	4	15	7	3
Anger	5	10	1	11	**60**	0	8	5
Surprise	0	8	5	14	2	**59**	0	12

Table 2. Confusion matrix for the video condition. The numbers are percentages computed considering the number of subject's correct answers over the total number of expected correct answers (300) for each emotional state.

Video	*Sad*	*Ironic*	*Happy*	*Fear*	*Anger*	*Surprise*	*Others*	*Neutral*
Sad	**49**	9	6	9	3	4	7	13
Ironic	8	**49**	10	2	5	4	8	14
Happy	1	24	**61**	3	1	4	2	4
Fear	5	2	3	**59**	14	8	6	3
Anger	3	6	1	9	**68**	4	3	6
Surprise	4	7	6	11	13	**37**	4	18

of the cases with *fear*, and 10% with *irony*), and **surprise** 59% of correct answers (confusion was made for 14% of the cases with *fear*, and 12% with a *neutral* expression). In all cases the percentage of correct recognition was largely above the chance. Surprisingly, **happiness** and **fear** were not very well recognized. **Happiness** had 50% of correct answers but was significantly confused with *irony* in almost 30% of the cases and partially confused with *surprise* in 11% of the cases. The percentage of correct answers for **fear** was 48% and it was confused with *sadness* and *surprise* by almost 20% and 15% of the cases respectively.

Table 2 reports the confusion matrix for the video condition. The numbers are percentages computed over the number of the subject's correct answers to each stimulus and averaged over the number of the expected correct answers for each emotional state.

The data displayed in Table 2 show that, in the video condition, the higher percentage of correct answers - 68% - was for **anger** (the higher confusion, 9%, was with *fear*), followed by 61% of correct answers for **happiness** (significantly confused - 24% - with *irony*). **Fear** reached 59% of correct answers (confusion was made in 14% of the cases with *anger*). S**adness** and **irony** obtained both the 49% of correct answers. The highest confusion was on *neutral* expressions that had 13% and 14% for sadness and irony respectively. **Surprise** was not well recognized getting only 37% of correct answers. The higher confusion was spread out among *anger* -13%- *fear* -11%- and *neutral* -18%.

Table 3. Confusion matrix for the audio condition. The numbers are percentages computed considering the number of subject's correct answers over the total number of expected correct answers (300) for each emotional state.

Audio	*Sad*	*Ironic*	*Happy*	*Fear*	*Anger*	*Surprise*	*Others*	*Neutral*
Sad	**67**	7	0	7	1	1	4	13
Ironic	9	**75**	7	2	2	4	0	1
Happy	5	25	**48**	8	1	5	1	7
Fear	12	3	6	**63**	6	5	2	3
Anger	2	9	0	10	**77**	0	1	1
Surprise	5	21	3	19	4	**37**	2	9

Table 3 reports the confusion matrix for the audio condition. Again, the numbers are percentages computed over the number of the subject's correct answers to each stimulus and averaged over the number of the expected correct answers for each emotional state.

The data displayed in Table 3 show that, in the audio condition, the higher percentage of correct answers - 77% - was for **anger** (the higher confusion, 10%, was with *fear*), followed by 75% of correct answers for **irony** (confused in 9% of the cases with *sadness*). **Sadness** obtained 67% of correct answers and in 13% of the cases was confused with a neutral expression. **Fear** got 63% (confusion was made in 12% of the cases with *sadness*), and **happiness** reached 48% of correct answers. Happiness was significantly confused with irony in 25% of the cases. **Surprise** was not well recognized also in the audio condition getting only 37% of correct answers. It was significantly confused with *irony* and *fear* in 21% and 19% of the cases respectively.

The emotional state better identified in all of the three conditions was anger (60% in the audio and video, 68% in the video, and 77% in the audio). Sadness and irony were easily recognized in the audio and video (75% and 64% respectively) and in the audio (67% and 75% respectively) conditions but hardly recognized in the video alone (49% and 49 % respectively).

Happiness was better identified in the video alone (61%) than in the audio (48%) and in the audio and video (50%) condition. Fear was better perceived in the audio (63%) and video (59%) than in the audio and video condition (49%), whereas, surprise was more easily identified in the audio and video (59%) than in the audio (37%) and video (37%) alone.

A comparative display of the data discussed in Tables 1, 2, and 3 is reported in Figure 2. Figure 2 shows the number of correct answers for each emotional state (sadness in Figure 2a., irony in Figure 2b., happiness in Figure 2c., fear in Figure 2d., anger in Figure 2e., and surprise in Figure 2f.) and for the audio, video and audio and video conditions. On the x-axis are reported the emotional labels and on the y-axis is reported the number of correct answers obtained for each emotional state.

An ANOVA analysis was performed on the above data considering the *condition* (audio alone, video alone, and audio and video) as a between subject variable and the *emotional state* as a six level (a level for each emotional state) within subject variable. The statistic showed that *condition* plays a significant role for the perception of emotional states ($F(2, 12) = 7.701$, $\rho=.007$) and this did not depend on the different

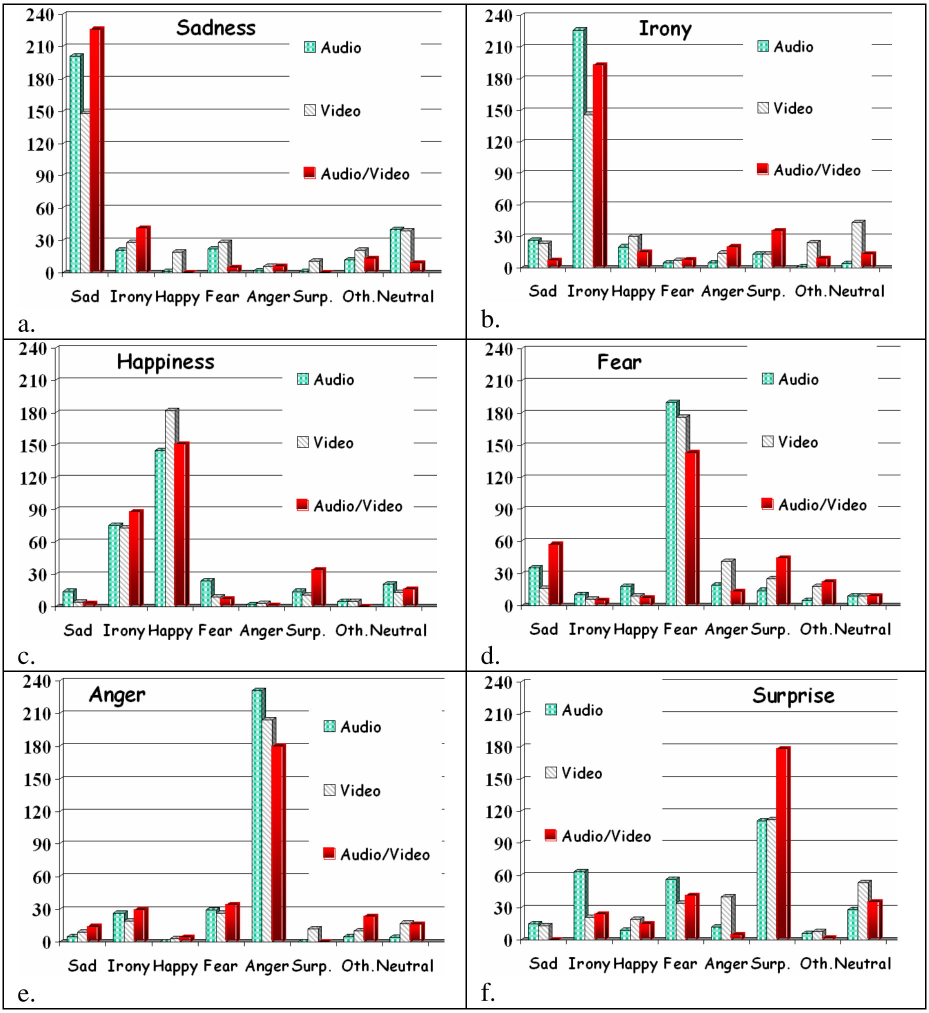

Fig. 2. Histograms of the number of correct answers given by the subjects participating to the perceptual experiments for the six emotional states under examination and the three different perceptual conditions (gray for audio, white for video, and black for audio and video – color would be different in a color version of the paper)

emotional states (F(5, 60) = .938, ρ=.46) since no interaction was found between the two variables (F(10, 60) = 1.761, ρ=.08). Moreover, female stimuli were more affected (F(2, 12) = 4.951, ρ=.02) than male stimuli (F(2, 12) = 1.141, ρ=.35) by condition, even though gender was not significant (F(1, 12) = .001, ρ=.98) and no interaction was found between condition and gender (F(2, 12) = .302, ρ=.74). In the details, we found that there was not a significant difference in the perception of the emotional states for the audio alone and the audio and video condition (F(1, 8) = .004, ρ=.95). Significant differences were found between the video alone and the audio and video condition ($F(1, 8) = 10.414$, $\rho=.01$) in particular for sadness ($F(1, 8) = 5.548$, $\rho=.04$), and the video and audio alone ($F(1, 8) = 13.858$, $\rho=.005$). In this particular case, significant differences were found for sadness ($F(1,8) = 8.941$, $\rho=.01$), fear ($F(1,8) = 7.342$, $\rho=.02$), and anger ($F(1,8) = 9.737$, $\rho=.01$).

A methodological concern that could be raised at this point is whether the ability of the actors and/or actresses in expressing emotions through vocal expressions may have affected the results making the audio stimuli more emotionally expressive than the video. However, if this was the case, it would be necessary to explain why in the combined audio and video condition subjects did not exploited the same vocal cues they have perceived in the audio alone to infer the protagonist emotional state. We will return to this argument in the final discussion. Nonetheless, further research is necessary to evaluate the above possibility, and to determine whether the present results generalize to other geographically and socially homogeneous groups of subjects.

4 Discussion

The statistical data discussed above and displayed in Tables 1, 2, and 3, and in Figure 2 show a really unexpected trend and posit the basis for new interpretations of the amount of perceptual information that are caught in the three different conditions. According to common sense it should be expected that subjects will provide the highest percentage of correct answers to all the emotional states in the audio and video condition, since it can be assumed that the combined channels contain more or otherwise redundant information than the audio and video alone. However, the experimental data did not support this trend. The subjects perform similarly, or at least, there is not a significant difference in perceiving emotional states either exploiting only the audio or the audio and video together, whereas they are less effective in recognizing emotional states through the video alone. In particular, in the video alone, the percentage of correct answers for sadness is significantly lower than those obtained in the audio alone and in the audio and video together, suggesting that the visual channel alone conveys less information's on sad perceptual cues. Moreover, in the video alone, the percentage of correct answers for fear and anger is significantly lower than that obtained in the audio alone, attributing to the visual channel also a minor effectiveness in communicating perceptual cues related to these two emotional states. Furthermore, there is not a significant difference between the audio alone and the audio and video together on the subject's perception of the emotional states under consideration. Therefore, it appears that the vocal and the visual communication modes convey a different amount of information about the speaker's emotional state. The vocal or audio channel can be considered as preferential since it conveys the same information as the combined audio and video channels. In some cases it also seems to be able to resolve ambiguities produced by combining information derived from both the channels. For example, for anger, we had a stimulus that was correctly identified by 18 subjects in the audio alone, whereas it was confused with fear by almost all the subjects in the audio and video condition. Even though in a non statistically significant manner, audio beats audio and video for anger, irony and fear, whereas happiness appears to be more effectively transmitted by the video alone than the audio and the audio and video together. Surprise seems to be the only emotional state that to be clearly perceived requires information from both the audio and video channels. We can conclude that the audio channel is preferential for most of the emotional states under examination and that, the combined audio and visual stimuli do not help in improving the perception of emotional states, as it was expected.

4.1 The Proposed Explanation

These results first of all suggest a nonlinear processing of emotional cues. In perceiving emotions, expressive information is not added (this is true even in the case of surprise, since the percentage of correct answers in the combined audio and video channels is not the sum of those obtained in the audio and video alone) over the amount of emotional cues provided. Instead, as the number of emotional cues increase (assuming that when combining together the video and the audio, it must increase) the subject's cognitive load increases since she/he should concurrently elaborate gestures, facial and vocal expressions, dynamic information (due to the fact that the video and audio information evolved along the time), and the semantic of the message. These concurrent elaborations deviate the subject's attention from the perceived emotion bringing to a perception that may go beyond the required task since subjects are more inclined to attempt to identify the contextual situation of the protagonist (and the identification of the contextual situation may be harder than the identification of the emotional state due to shortness of the stimuli) than the expressed emotional state. This could be seen from the number of alternative labels given in each task. The fewer contextual cues they have, the fewer alternative labels are given. In fact, in the audio, the total number of alternative labels was 13, against the 64 of the video and the 49 of the video and audio together.

The explanation provided above is not in disaccord with reported data [12-13, 21, 25-26, 37, 39, 42-43, 72-74] demonstrating that facial expressions are more powerful than vocal expressions in eliciting emotional states, since these data are based on still photographs. As we already said in a previous section, still images usually capture the apex of the expression, i.e. the instant at which the indicators of emotions are mostly marked. In this case, the subject's cognitive load is reduced since only static and highly marked emotional information must be processed. These highly marked emotional indicators reproduced in the photos are like emblems universally shared because of the possible existence of an innate motor program [12-13, 72-74] that controls our facial musculature and changes our faces when we are emotionally involved. And their meaning is also universally shared explaining why we make faces also when we are not emotionally but contextually involved in emotional situations.

When dynamic is involved, audio is better. This is because audio is intrinsically dynamic and we are used to processing it dynamically. Moreover, listeners are used to dynamically vocal emotional cues. Therefore, the cognitive load is just reduced and they perform better. Moreover, we must here add a further speculation. Our idea is that audio is better for native speakers of the language, since natives rely on paralinguistic and suprasemental information that is strictly related and unique to their own language. They learned this information by themselves when they were emotionally involved and observed the changes in their vocal expressions. They may have learned how a given word is produced in a particular emotional state, even though the word is not semantically related to the emotional state. Consequently, they may be able to capture very small supra-segmental and paralinguistic cues that are not noticeable to a non native speaker. This is not to say that vocal expressions of emotions are learned. We are not making this point since we do not have data on that. The uncontrolled changes in the vocal motor program when we are emotionally involved may be innate as those in our facial musculature. However, changes in intonation and other paralinguistic information that derive from changes in the vocal motor control may have been learned together with the language. Hence, native speakers may be able to catch up more emotional information in the audio than in the video when the emotions are expressed in their own language. Nonetheless, they may rely more on visual rather than vocal information when they are requested to catch up emotional states in a foreign language. This is because visual cues may share more commonalities being more universal than vocal cues since the visual processing of inputs from the outside world exploits neural processes common to all human beings whereas, languages require a specific expertise or training to be processed.

At the light of the above considerations, we may expect that using a foreign language, we will find that the percentage of correct answers will not vary significantly between the video alone and the combined audio and video channels and that significant differences will be reported for the audio alone. We are actually testing this hypothesis (and we have already some preliminary results) through a set of perceptual experiments similar to those reported above, with the differences that in this case the protagonists of our clips speak a foreign language and therefore in the audio alone condition the participants cannot rely on their language expertise. To this aim, we checked that the subjects participating at the experiments did not know the language.

Dynamicity and language expertise is also able to explain why subjects performed poorer in the audio and video alone than in the combined channels for surprise. Surprise is an almost instantaneous state, is like a still image. The short duration of this emotional state allow the listener to exploit both the visual and auditory information without increasing the cognitive load and avoid to diverge her/his attention from the perceived emotion.

The discussion on the language expertise however, rises a more fundamental question on how much of the expressions of emotions are universally shared. Data reported in literature support a universal interpretation of emotional facial expressions [12-13, 21-23, 39, 72-74] but these data were based, as we already pointed out, on still images. In our case the question is how much of the dynamic features of emotional faces are universal? We have already speculated that the perception of emotional vocal expressions can strictly depend on the language expertise and that, in such a case, visual cues are more universally recognized than vocal cues. However, we have not considered that dynamic emotional visual stimuli can also be culture specific. This means that we may expect subjects to perform better on culturally close visual emotional cues than on culturally distant ones. The above proposed experiments could be able to answer also to this question, and quantify how much dynamical emotional information we are able to capture in a foreign face-to-face interaction. They should also be able to identify how much is universally and how much is culture specific in the expressions of emotions.

4.2 A Mathematical Model Fitting the Experimental Data

The above results are also interesting from an information processing point of view. Information processing theory is a research branch of communication theory that attempts to quantify how well signals encode meaningful information and how well systems process the received information. In this context, *entropy* is defined as a measure of "self-information" or "uncertainty" transmitted by a given signal source as it can be considered in our case, both the audio and video channels. Mathematically the entropy is described by the equation:

$$H(X) = -\sum_{n=1}^{N} p(X_n)\log p(X_n) \tag{1}$$

Where X_n, $n=1, \ldots, N$ are the symbols that can be emitted by the source and $p(X_n)$ the probability for a given symbol to be emitted. In the Shannon view, the entropy is a measure of the complexity or the information contained in the set of symbols and its value is highly affected by the symbol's probability distribution. The entropy model can be extended to more than one source defining what is called the *joint entropy*, which, for two sources is mathematically described by the equation:

$$H(X,Y) = -\sum_{n=1}^{N}\sum_{m=1}^{M} p(X_n,Y_m)\log p(X_n,Y_m) \tag{2}$$

Where X_n, $n=1, \ldots, N$, and Y_m, $m=1, \ldots, M$ are the symbols that can be emitted by the sources *X* and *Y*, and $p(X_n, Y_m)$ the joint probability distribution.

Theoretically, it can be proved that the joint entropy of two different sources is greater than that of a single source, and therefore, it can be expected that increasing the number of sources the amount of information transmitted over a communication channel will consequently increases [75]. The theoretical definition of entropy does not take into account the informative value of the source, i.e. the meaning the signal can convey and what the receiver assumes to be informative about the received signal. Consequently, entropy measure does not consider cases where the source can be only noise, and/or cases where extraneous information can be exploited by the receiver. In our context entropy cannot be considered an objective measure of the information transmitted by the audio and video channels. A more adequate measure could be the mutual information that takes into account the effects of the transmission channel on the received signal. Mutual information is a measure of the dependence between the source and/or the receiver and it produces a reduction of the "uncertainty" or the amount of information transmitted by the source. Mutual information is mathematically expressed by the equation:

$$I(X,Z) = -\sum_{n=1}^{N}\sum_{k=1}^{K} p(X_n,Z_k)\log\frac{p(X_n,Z_k)}{p(X_n)p(Z_k)} \tag{3}$$

Mutual information could be an appropriate mathematical model to describe our results, since being always less greater or equal to the entropy of the source it can explain why by combining audio and video channels the information about the emotional states does not increases. However, mutual information requires a complete knowledge on how the receiver filters (or transforms) the information emitted by the source. In our case this information is completely unknown, since we do not have at the date a clear idea on how our brain process information received by visual and auditory channels. Moreover, mutual information cannot explain our hypothesized differences in perception between native and non native speakers and between close and distant cultural backgrounds.

A more adequate model could be the *information transfer ratio* proposed by Sinanović and Johnson [76] that quantifies how the receiver affects the received information by measuring the distance between (the probability distributions of) the actions performed by the receiver for given received inputs and (the probability distributions of) the received inputs. Assuming that emotional information encoded through the audio and the video represents the input to a processing system, the information transfer ratio value can be quantified as the ratio of the distances between the information transmitted by the sources and the distances between the outputs (for those sources) of the processing system. Information transfer ratio is described by the following equation:

$$\gamma_{XY,Z}(\alpha_0,\alpha_1) \equiv \frac{d_Z(\alpha_0,\alpha_1)}{d_X(\alpha_0,\alpha_1)} + \frac{d_Z(\alpha_0,\alpha_1)}{d_Y(\alpha_0,\alpha_1)} \tag{4}$$

Where $\alpha 1$ indicates the changes in the emotional contents of the sources with respect to a reference point α_0, which depends not on the signals itself but on their probability functions $p_X(x;\alpha_0)$, $p_X(x;\alpha_1)$, $p_Y(y;\alpha_0)$, $p_Y(y;\alpha_1)$, and $p_Z(z;\alpha_0)$, $p_Z(z;\alpha_1)$. *X* and *Y* are the sources (audio and video), *Z* the processing systems, and *d(…)* indicate an information theoretic distance (for example, the Kullback-Leibler distance [1]) that obeys to the Data Processing Theorem [35], which roughly states that the output of any processing system cannot contain more information than the processed signal.

The information transfer ratio quantifies the ability of the receiver to process the received information and predicts in the case of multiple sources that the overall information transfer ratio may be smaller than the individual ratios for each input signal. This is due to the arbitrary filtering action performed by the receiver (the human being) whose transfer function is in our case unknown, even though its ability to catch emotional information can be derived by the number of the subject's correct answers. In simple words, in the complex world of emotion's perception the whole, in some cases, seems not to be more than the sum of its parts.

5 Conclusions

The present paper reports on perceptual data showing that native speakers rely more on the auditory than on the visual channel to infer information on emotional states, and that the amount of emotional information conveyed does not increases combining auditory and visual information. An attempt is made to explain this phenomenon in terms of the subject's cognitive load and information processing theory.

Acknowledgements

This research has been partially supported by the project *i-LEARN: IT's Enabled Intelligent and Ubiquitous Access to Educational Opportunities for Blind Student,* at Wright State University, Dayton, Ohio, USA. Acknowledgment goes to Tina Marcella Nappi for her editorial help. Acknowledgement also goes to Professors Maria Marinaro and Olimpia Matarazzo for their useful comments and suggestions. In particular, Olimpia has greatly helped in the selection of stimuli. The students Gianluca Baldi, Eliana La Francesca, and Daniela Meluccio are greatly acknowledged for running the perceptual tests.

References

1. Ali, S.M., Silvey, S.D.: A general class of coefficients of divergence of one distribution from another. Journal of Royal Statistic Society, 28, 131-142 (1966)
2. Apple, W., Hecht, K.: Speaking emotionally: The relation between verbal and vocal communication of affect. Journal of Personality and Social Psychology, 42, 864–875 (1982)
3. Argyle, M.: Bodily Communication. New York: Methuen & Co (1988)
4. Banse, R., Scherer, K.: Acoustic profiles in vocal emotion expression. Journal of Personality & Social Psychology 70(3), 614-636 (1996)
5. Barnes, J.: Aristotele complete work. Revised Oxford translation in 2 volumes, Princeton University press (1984)
6. Bachorowski, J.A. (1999). Vocal expression and perception of emotion. Current Directions in Psychological Science, 8, 53–57.
7. Birdwhistell, R.: Kinesies and context. Philadelphia, University of Pennsylvania Press (1970)
8. Breitenstein, C., Van Lancker, D., & Daum, I. The contribution of speech rate and pitch variation to the perception of vocal emotions in a German and an American sample. Cognition & Emotion, 15, 57–79 (2001)
9. Burgoon, J.K.: Nonverbal signals. In Knapp, J. L., Miller, G. R. (eds.). Handbook of interpersonal communication, Thousand Oaks, CA: Sage, 229-285 (1994)
10. Burleson, B.R.: Comforting messages: Features, functions, and outcomes. In Daly, J.A., Wiemann, J.M. (eds.), Strategic interpersonal communication, Hillsdale, NJ: Erlbaum, 135-161 (1994)
11. Burns, K. L., Beier, E. G.: Significance of vocal and visual channels in the decoding of emotional meaning. Journal of Communication, 23, 118–130 (1973)
12. Caccioppo, J. T., Klein, D. J: Bernston, G.C., Hatfield, E.: The psychophysiology of emotion. In Haviland, M., Lewis, J. M. (eds) Handbook of Emotion, New York: Guilford Press, 119-142 (1993)
13. Caccioppo, J. T., Bush, L.K., Tassinary, L.G.: Microexpressive facial actions as a functions of affective stimuli: Replication and extension. Personality and Social Psychology Bullettin, 18, 515-526 (1992)
14. Corraze, G.: Les communications nonverbales. Presses Universitaires de France, Paris (1980)
15. Cosmides, L.: Invariances in the acoustic expressions of emotions during speech..Journal of Experimental Psycology, Human Perception Performance, 9, 864-881 (1983)
16. Darwin, C.: The expression of the emotions in man and the animals (1872). Reproduced by the University of Chicago, Chicago press (1965)

17. Davitz, J. R.: Auditory correlates of vocal expression of emotional feeling. In Davitz J. R. (ed.), The communication of emotional meaning, New York: McGraw Hill, 101-112 (1964)
18. Davitz, J.: The communication of emotional meaning. McGraw-Hill (1964)
19. Davitz, J. R.: Auditory correlates of vocal expression of emotional feeling. In Davitz J. R. (ed.), The communication of emotional meaning, New York: McGraw Hill, 101-112 (1964)
20. Ekman, P., Friesen, W.V., Hager, J.C.: The facial action coding system. Second edition. Salt Lake City: Research Nexus eBook.London: Weidenfeld & Nicolson(2002)
21. Ekman, P. Facial expression of emotion: New findings, new questions. Psychological Science, 3, 34-38 (1992)
22. Ekman, P.: An argument for basic emotions. Cognition and Emotion, 6, 169-200 (1992)
23. Ekman, P.: The argument and evidence about universals in facial expressions of emotion. In H. Wagner, H., Manstead, A. (eds.). Handbook of social psychophysiology, Chichester: Wiley, 143-164 (1989)
24. Ekman, P.: Expression and the nature of emotion. In Scherer, K., Ekman, P. (eds), Approaches to emotion, Hillsdale, N.J.: Lawrence Erlbaum, 319-343 (1984)
25. Ekman, P. Friesen, W. V.: Facial action coding system: A technique for the measurement of facial movement. Palo Alto, Calif.: Consulting Psychologists Press (1978)
26. Ekman, P., Friesen, W.V.: Manual for the Facial Action Coding System, Palo Alto: Consulting Psychologists Press (1977)
27. Ekman, P., Friesen, W.V.: Head and body cues in the judgement of emotion: A reformulation. Perceptual Motor Skills 24: 711-724 (1967)
28. Frick, R.: Communicating emotions: the role of prosodic features. Psychological Bullettin, 93, 412-429 (1985)
29. Fridlund, A.J.: The new ethology of human facial expressions. In Russell J.A., Fernandez-Dols J. (eds.), The psychology of facial expression Cambridge: Cambridge University Press, 103-129 (1997).
30. Fridlund, A.J.: Human facial expressions: An evolutionary view. San Diego: CA Academic press (1994)
31. Friend, M.: Developmental changes in sensitivity to vocal paralanguage. Developmental Science, 3, 148–162 (2000)
32. Frijda, N.H.: Moods, emotion episodes, and emotions. In Haviland, M., Lewis, J. M. (eds) Handbook of Emotion, New York: Guilford Press, 381-402 (1993)
33. Frijda, N.H.: The emotions, Cambridge University press (1986)
34. Fulcher, J.A.: Vocal affect expression as an indicator of affective response. Behavior Research Methods, Instruments, & Computers, 23, 306–313 (1991)
35. Gallager, R.G.: Information theory and reliable communication. John Wiley & Son (1968)
36. Goldin-Meadow, S.: Hearing gesture: How our hands help us think. The Belknap Press at Harvard University Press (2003)
37. Graham, J., Ricci-Bitti, P.E., Argyle, M.: A cross-cultural study of the communication of emotion by facial and gestural cues. Journal of Human Movement Studies, 1, 68-77 (1975)
38. Haldane, E.L., Ross, G.R.: The philosophical work of Descartes. New York: Dover, current edition (1911)
39. Izard, C.E.: Innate and universal facial expressions: Evidence from developmental and cross-cultural research. Psychological Bulletin, 115, 288–299 (1994)
40. Izard, C.E.: Organizational and motivational functions of discrete emotions. In Lewis M., Haviland J. M (eds.), Handbook of emotions New York: Guilford Press, 631–641 (1993)

41. Izard, C.E.: Basic emotions, relations among emotions, and emotion–cognition relations. Psychological Review, 99, 561–565 (1992)
42. Izard, C.E., Dougherty, L.M., Hembree, E.A.: A system for identifying affect expressions by holistic judgments. Unpublished manuscript. Available from Instructional Resource Center, University of Delaware (1983)
43. Izard, C.E.: The maximally discriminative facial movement coding system (MAX). Unpublished manuscript. Available from Instructional Resource Center, University of Delaware (1979)
44. James W.: What is an Emotion? First published in Mind, 9, 188-205 (1884). See http://psychclassics.asu.edu/James/emotion.htm
45. Junqua, J.C.: The Lombard reflex and its role on human listeners and automatic speech recognizers. JASA, 93(1), 510-524 (1993)
46. Kendon, A.: Gesture : Visible action as utterance. Cambridge University Press (2004)
47. Klasmeyer, G., Sendlmeier W. F.: Objective voice parameters to characterize the emotional content in speech. In Proceedings of ICPhS 1995, Elenius, K., Branderudf, P. (eds), 1,182-185, Arne Strömbergs Grafiska (1995)
48. Klinerberg, O.: Emotional expression in chinese literature. Journal of Abnormal and Social Psychology, 33, 517-520 (1938)
49. Klinnert M.D., Campos J.J., Sorce J.F., Emde R.N., Svejda M.: Emotions as behaviour regulators: Social referencing in infancy. In Plutchik R,. Kellerman H. (eds), Emotions,: Theory, research, and experience, 57-86, New York, Academic Press (1983)
50. La Barre, W.: The cultural basis of emotions and gestures. Journal of Personality, 16, 49-68 (1947)
51. Lumsden, C.J., Wilson E.O.: Genes mind and culture: The co-evolutionary process. Harvard University Press (1981)
52. McNeill, D.: Gesture and thought. Chicago: University of Chicago Press (2005)
53. Mehrabian, A.: Orientation behaviors and non verbal attitude communication. Journal of Communication, 17(4), 324-332 (1967)
54. Millar, F.E., Rogers, L.E.: A relational approach to interpersonal communication. In Miller, G.R. (ed.), Explorations in interpersonal communication Beverly Hills, CA: Sage, 87-105 (1976)
55. Miller: Explorations in interpersonal communication. London: Sage, (1976)
56. Mozziconacci, S.: Pitch variations and emotions in speech. In Proceedings of ICPhS 1995, Elenius, K., Branderudf, P. (eds), 1, 178-181, Arne Strömbergs Grafiska (1995)
57. Nushikyan, E. A.: Intonational universals in texual context. In Proceedings of ICPhS 1995, Elenius, K., Branderudf, P. (eds),1, 258-261, Arne Strömbergs Grafiska (1995)
58. Oatley, K., Jenkins, J. M.: Understanding emotions. Oxford, England: Blackwell (1996)
59. Panksepp, J.: Affective neuroscience: A conceptual framework for the neurobiological study of emotions. In Strongman K (ed.) International Review of Studies of Emotions, Chichester, England: Wiley, 1, 59-99 (1991)
60. Plato: The Republic. (375 BC). Harmondworth, Middlesex: Penguin, current edition (1955)
61. Plutchik, R.:Emotion and their vicissitudes: Emotions and Psychopatology. In Haviland, M., Lewis, J. M. (eds) Handbook of Emotion, New York: Guilford Press, 53-66 (1993)
62. Plutchik, R.: The emotions. Lanham, MD: University Press of America (1991)
63. Pittam, J., Scherer, K.R.: Vocal expression and communication of emotion. In Haviland, M., Lewis, J.M. (eds) Handbook of Emotion, New York: Guilford Press, 185-197 (1993)
64. Ricouer, P.: The voluntary and the involuntary. Translation of Kohak K. Original work 1950, Northwestern University Press (1966)

65. Rosenberg, B.G., Langer, J.: A study of postural-gestural communication. Journal of Personality and Social Psychology 2(4), 593-597 (1965)
66. Scherer, K: Vocal communication of emotion: A review of research paradigms. Speech Communication 40, 227-256 (2003)
67. Scherer, K.R., Banse, R., Wallbott, H.G.: Emotion inferences from vocal expression correlate across languages and cultures. Journal of Cross-Cultural Psychology, 32, 76–92 (2001)
68. Scherer, K.R., Banse, R., Wallbott, H.G., Goldbeck, T.: Vocal cues in emotion encoding and decoding. Motivation and Emotion, 15, 123–148 (1991)
69. Scherer, K.R: Vocal correlates of emotional arousal and affective disturbance. In Wagner, H., Manstead, A. (eds.) Handbook of social Psychophysiology New York: Wiley, 165–197 (1989)
70. Scherer, K.R.: Experiencing Emotion: A Cross-cultural Study, Cambridge University Press, Cambridge (1982)
71. Scherer, K.R., Oshinsky, J.S.: Cue utilization in emotion attribution from auditory stimuli. Motivation and Emotion, 1, 331–346 (1977)
72. Schwartz, G.E., Weinberger, D.A., Singer, J.A.: Cardiovascular differentiation of happiness, sadness, anger, and fear following imagery and exercise. Psychosomatic Medicine, 43, 343–364 (1981)
73. Schwartz, G.E., Ahern, G.L., Brown, S.: Lateralized facial muscle response to positive and negative emotional stimuli. Psychophysiology, 16, 561-571 (1979)
74. Schwartz, G.E., Fair, P.L., Salt, P., Mandel, M.R., Klerman, G.L.: Facial muscle patterning to affective imagery in depressed and non-depressed subjects. Science, 192, 489- 491 (1976)
75. Shannon, C.E., Weaver, W.: Mathematical Theory of Communication. US: University of Illinois Press (1949)
76. Sinanović, S., Johnson, D.H.: Toward a theory of information processing. Submitted to Signal Processing. http://www-ece.rice.edu/~dhj/cv.html#publications (2006)
77. Tomkins, S.S.: Affect, imagery, consciousness. The positive affects. New York: Springer, 1 (1962)
78. Tomkins, S.S.: Affect, imagery, consciousness. The negative affects. New York: Springer, 2 (1963)
79. Tomkins, S.S.: Affect theory. In Scherer, K.R., Ekman, P.(eds.), Approaches to emotion, Hillsdale, N.J.: Erlbaum, 163-196 (1984)
80. White, G.M.: Emotion inside out The anthropology of affect. In Haviland, M., Lewis, J. M. (eds) Handbook of Emotion, New York: Guilford Press, 29-40 (1993)