UNIVERSITY OF GOTHENBURG

How do different modes contribute to the interpretation of affective epistemic states?

How different mode of representation (*video*, *audio*, *video+audio* and *written words*) can influence the understanding and interpretation of *AES*

Stefano Lanzini

Master of Communication Thesis
Report No. 2013:071
ISSN: 1651-4769

University of Gothenburg
Department of Applied Information Technology
Gothenburg, Sweden, May 2013

Abstract

For every human being, it is essential to transmit information and advise other people about their own *affective epistemic states*. On the other hand, it is also indispensable for every person to be able to interpret the states of other people with whom they interact. By the term affective epistemic states (*AES*) we want to indicate all those states that involve cognition, perception and feeling, or as Schroder suggests "states which involve both knowledge and feeling" (Schroder, 2011).

The aim of this paper is to show how different mode of representation (video, audio, a combination of audio+video and written words) can influence the understanding and interpretation of *AES*. We also want to examine the effect of multimodality (using *visual* and *auditory sensory modality simultaneously*) compared to unimodality (using just *visual* or just *auditory sensory modality*). Even if some studies investigated the area of emotions and affective states, it is still very hard to find research that involves epistemic features. More studies are essential in order to understand the mechanism about how *AES* are interpreted by humans.

We conducted an experiment at the University of Gothenburg with 12 Swedish participants. Four recordings of first encounters face-to-face interaction, were displayed to each respondent. Each recording was shown in a different mode. The modes used for the experiment consisted of a transcriptions (T), a video with audio (V+A), a video without audio (V) and an audio recording (A). The recordings were all about two people that were meeting for the first time. We asked the respondents to identify which kinds of *AES* were displayed by people participating in the recording. Respondents were asked to motivate their answers.

Several interesting outcomes have been observed. Participants were able to interpret different *AES* when exposed to the same behavior in different modes. This means that when the same behavior is displayed in different modes, the respondent's perception is often influenced in different ways. The same *AES* can be shown through vocal and gestural behaviors, and it can be perceived by visual, auditory or both modalities together, depending on the modes displayed. We observed that *AES* are highly multimodal and the majority of the times, different behaviors are perceived differently, depending on if they were shown by multimodal or unimodal modes.

Keywords: Affective epistemic states, perception, multimodality, unimodality, communication.

Table of Contents

1. Introduction

This thesis discusses issues related to multimodal communication and in particular to the interpretation of *affective epistemic states*. By the term *affective epistemic states* (*AES*) we want to indicate all those states that involve cognition, perception and feeling, or as Schroder suggests "states which involve both knowledge and feeling" (Schroder, 2011). Since not so much research has been carried out to investigate states such as *understanding*, *interest*, *surprise* and *confusion*, the mechanism by which *AES* are interpreted by humans is still a relatively unexplored area. Even if some studies examine emotions and affective states, it is very hard to find research that investigates epistemic features. Therefore, to understand how humans show and interpret different *AES*, more research is necessary.

In order to study how people interpret *AES*, we have conducted an experiment at the University of Gothenburg. The aim of the experiment is to analyze and understand how people interpret *AES* when they are presented with different modes of communication (video, audio and writing) and the combination of them (video+audio). Moreover, we want to investigate the effect of multimodality compared to unimodality. In order to answer these questions, we presented a video+audio version, a video version, an audio version and a transcription of a recording of two people that were meeting for the first time. We asked the respondents to identify which kind of *AES* were displayed by people participant in the recording. Respondents were asked to motivate their answers.

It was possible to observe several interesting outcomes. Firstly, we observed that when the same behavior is displayed in different modes, the respondents' perception is often influenced in different ways. This means that the participants were able to interpret different *AES* when exposed to the same behavior. We can see that the answers depend on the modes of presentation. Since the same *AES* can be shown through both audio and video modes and perceived by auditory and visual modalities, we can claim that *AES* are highly multimodal, and they can be perceived by different senses. Therefore, we can say that multimodality, in comparison with unimodality, influences respondents' interpretation of *AES* in different ways.

This thesis is divided in six chapters. In the first chapter, we will illustrate the purpose of this work, our research questions and our hypotheses. Subsequently, a literature revision is presented. In this part we will present and explain previous research and studies conducted in the area of multimodal communication. Here we will explain what multimodality is and the meaning of the term *AES*. Moreover, we will also present and describe the different patterns of communication. Thereafter, we will explain the method used in order to carry out this research, which is then followed by a presentation of our data. Consequently, we will present our results in a discussion leading to future studies and possible applications. The last chapter of the thesis is a summary of our results.

1.1. Purpose of the research

The aim of this paper is to show how different modes of representation (video, audio and written words) can influence the understanding and interpretation of *AES*, i.e. states which simultaneously involve cognition, perception and feeling. We would like to answer at two different research questions. The two queries have different goals, but they are correlated to each other.

The first question is general, and it has a focus on how people interpret AES when they are displayed in different modalities.

> - *How do different modes contribute to the interpretation of affective epistemic states?*

The second question is more specific and is aimed of discovering the possible differences of interpretation of *AES*, when they are presented in unimodal and multimodal modes.

> - *What differences of information, concerning affective epistemic states, can you get in an interpretation of a multimodal communication mode in comparison with an interpretation of a unimodal communication mode?*

1.2. Hypothesis

In order to answer to our research questions we formulated two hypotheses:

1) Different modes contribute in different ways to the interpretation of *AES*.

2) If multimodality leads to more information about the interpretation of *AES* than unimodality, then viewers will be able to interpret and give more information about *AES* when they are shown a Video+Audio (V+A) recording than when they are shown a unimodal recording, Video, Audio, or Transcription (T). If multimodality does not lead to more information about the interpretation of *AES* than unimodality, we will obtain the same results when we show multimodal stimuli (A+V) as when we use unimodal stimuli (A or V or T). This would mean that multimodality is redundant for the interpretation of *AES*.

2. Different aspects of multimodal communication

For every human being, it is essential to transmit information and advise other people about their own *affective epistemic states*. On the other hand, it is also indispensable for every person to be able to interpret the states of other people with whom they interact. In this section we will present a general background about different studies made in the area of *multimodal communication* concerning how people communicate their *AES*. Below, we will focus mainly on two different aspects of multimodal communication: *verbal* and *non-verbal communication*, here we will describe their most common characteristics, the behaviors and actions associated with them and we will present some of the studies made to study these two kinds of communication.

2.1. Multimodal communication

In order to explain what communication is, it can be useful to give some basic definition. If we use a functional definition of communication we can say that communication is the sharing of content X between a sender Y and a recipient Z using an expression W and a medium Q in an environment E with a purpose/function F (Allwood, 2002). Therefore, concerning multimodality, we can claim that multimodal communication is the sharing of a content X from between sender Y and a recipient Z using more than one sensory modality W and more than one physical medium Q simultaneously, in an environment E with a purpose/function F (Partan, 2005).

If we think about this definition we can see that multimodal communication occurs very often in nature. In fact humans are animals with sense organs that are able to perceive and transmit different kinds of information simultaneously. Normal human face-to-face communication can also be defined as a multimodal communication, since people use more than one type of behavior to produce information and more than one sensory channel to receive information (Allwood, 2002). One of the most famous studies of multimodal communication, conducted by McGurk and MacDonald (1976) shows that our perception of vocal verbal sounds is influenced by observation of the articulatory gesture (for example the movements of the lips and tong). This phenomenon takes the name of its researcher, and it is often called "McGurk effect" (McGurk, 1976). Other research also shows a correlation between sound and visual signals. Kuhl and Meltzoff (1982) illustrate that infants are able to recognize the correspondence and find discrepancies between visual and auditory in speech sounds (Kuhl, 1982). This feature occurs also in primates, like macaque monkeys (Ghazanfar, 2003) and chimpanzees (Parr, 2004). Multimodal communication is an area of communication that has been studied during the last 30 years and it is still studied today since several aspects remain still unexplored. Even if several investigations were made in order to study how multimodality contributes to communicate emotions, fewer studies have been made to explore how multimodal communication can influence the interpretation of affective and epistemic states (especially epistemic). In our experiment, the participants will use the two sensory modalities vision and hearing. The visual modality will be used to visualize images and read text, while the auditive modality will be used to listen to voices.

2.2. Affective epistemic states

By the term *affective epistemic states* we refer to all states that involve cognition, perception and feeling or as Schroder suggested "states which involve both knowledge and feeling" (Schroder, 2011). In order to better understand this definition, it could be useful to explain the meaning of each word. If we consider the term *state*, we can say that it indicates the conditions of an animal or a thing a certain period of time. Therefore in our case, when we are talking about *AES*, we use the term *state* to refer to the precise, biological, physiological and psychological condition of a person,

during a specific period of time. The term *affective state* refer to all those situations where a person is expressing emotional dispositions and/or emotional attitudes (Allwood, Chindamo, Ahlsén, 2012). So mostly we can say that *affective states* are all those states where a person is feeling some kind of emotion or affection, for instance: happiness, joy, anger, sadness, nervousness. On the other hand, by the term *epistemic* we want to indicate all those situations where a person is feeling for example certainty or doubts towards the content of that specific situation (Leavitt, 1991). Both affective and epistemic states can be expressed vocally or with gesture or with the voice.
In order to explain this concept I want to give a very useful example. Let's say that Marco and Andrea are having a conversation where they are talking about the date for a meeting.

Marco: "When can we meet?"
Andrea: "I'm available on Tuesday?"
Marco: "Sorry, when?" Marco shakes his head.
Andrea: "Tuesday!"
Marco: "Oh Tuesday is perfect! I'm available too!" Marco smiles.

From this very short conversation we can observe three different epistemic states of Marco. Firstly, when he says: "Sorry, when?" he is feeling and showing uncertainty about Andrea's answer. He is showing his feeling of uncertainty through his words and his body gestures. On the other hand Marco is showing two other attitude. The first one is the attitude of understanding; he understood that Andrea is available on Tuesday. He shows this through his words "Oh Tuesday is perfect! …". He is showing also the feeling of happiness, through his facial expression, about the fact that he also is available the same day.

Often affective states blend with epistemic if they have a relation and are directed to an epistemic entity, i.e. being happy (angry, sad) that it rains (Allwood , Chindamo, Ahlsén 2012), and epistemic states blend with affective too. Actually it is very hard to find states that are just affective or epistemic. There is usually a reason that we are feeling a precise emotion (even imperceptible), and the reason can be known or not. In fact, very often when we do or don't understand, or we have some doubts about something, certain emotions inside us appear. These emotions can also be imperceptible.

2.3. Non-verbal communication

Non-verbal communication is a very important part of people's daily communication and it can in this case be interpreted by other people as more important than our verbal messages. Through *non-verbal communication* we are able to share a wide range of information. By the term *non-verbal communication* we indicate all those activities that don't include the action of producing words. Words can be produced both vocally and by gesture (e.g. deaf sign language). It is very common that non-verbal and verbal actions occur at the same time. Indeed a study about gestural feedback expressions shows that most of the times feedbacks are displayed simultaneously with vocal/verbal and gestural verbal and non-verbal actions (Allwood, & Cerrato, 2003). Gestures can be used together with the speech (*speech related gesture*) to illustrate, explicate, emphasize, point, regulate turn taking or to add information concerning what it is said verbally (Knapp, 2010). When we are producing verbal information, we are producing a determinate body movement in order to share information. Verbal symbols, in order to have a common meaning, must be shared by the sender and receivers (Knapp, 2010). For example, the simple action of nodding is a symbol that conventionally in many part of the world means the word *yes*. On the other hand, when we are producing an involuntary body movement, we are also producing a possible signal of an *AES*. People often nod spontaneously during a conversation which can be interpreted as a sign that can indicate the *AES* of understanding. The person whom I'm talking with is nodding, so probably

he/she understands what I'm saying. If speaker and listener are not attracted to each other and a common meaning is not shared, it is very easy to misunderstand this kind of signals.

2.3.1. Body movements, posture, body orientation and proximity

People's movements, posture and proximity are important factors of *non-verbal communication* during any kinds of interaction. Body movements are those movement that people make with their body: hands, arms, legs feet, trunk and head. All can suggest information about communicator's *state*, for example: feelings, mood, emotions, degree of attention and degree of liking the other interlocutor. Sometimes the observation of *non-verbal communication* behaviors is the best way to understand people's *AES*, especially when they are alone, since they don't use very often verbal communication and their actions are less influenced by social rules (Ekman, 1972). One of the firsts pioneers in this area was Charles Darwin, in his book "*The expression of the emotions in man and animals*" he created some categories in order to classify body movements and emotions (Darwin, 1872). In his classification, Darwin observed that the state of joy is commonly displayed with several body movements, such as: jumping, dancing, clapping hands, nodding, shaking during laughing, body erect and head upright (Darwin, 1872). In the past numerous studies have been made in order to study all these particular types of communication. The importance of body movements in the interpretation of *AES* is underlined by the study of Meeren (2005) about body movements and facial expression. According to the writer, when the facial expression and body movements are in contrast, people tend to give more importance to emotion expressed by body movements, than the facial expressions (Meeren, 2005). The capacity to recognize and interpret body movements is something inherent in our brain, and it is present in 3-months old babies (Gliga, 2005). Research about fear interpretation confirms that one brain interpret fear very fast and automatically prepares itself for action (de Gelder, 2004). We have a certain areas of the brain that are specialized visual, motor and emotional processing; these areas let humans to interpret emotions through body movements (Kana, 2011). The results of other studies about body movements and emotions, shows that specific patterns of body movement allow us to match specific emotions with specific body movements (Dael, 2012) (Wallbott, 1998). For example, bowed, fast and indirect movements, seem to refer to *rejection* (anger, antipathy, contempt, disgust) (de Meijer, 1989). On the other hand, *acceptance* (interest, joy, sympathy and admiration) is inferred by open, light, elastic and direct movements (de Meijer, 1989).

If we focus just on hands and arms we can see that also their movements can carry information about *AES*. For instance in Finnish sign language respondents were able to identify anger and neutrality from hand movements even if they were not able to interpret Finnish sing language (Hietanen, 2004). Ekman (1972) building on Efron (1941), in a very interesting paper about hand movements, made a classification of gestural *illustrators*. *Illustrators* movements (mostly hands movement) are all those movements that are just related to the speakers and his/her speech and are used to help the receiver to understand what it was said verbally (Ekman, 1972).

Ekman (1972) also divides gestures into *emblems* and *adaptors*. *Emblems* are all those gestures with a specific meaning, which are known by a very big group of people (i.e. culture, nation, class, etc.). Most of the times, a person is aware that he/she is showing an *emblem* gesture. *Adaptors* are all those movements used to achieve self needs or body needs, as for example: scratching, shaking because a certain psychological state or body changes. These movements can also involve objects, for instance: playing with a pen during a stress moment. These kinds of gestures are often displayed when people are alone, instead they are avoided or controlled when people are in public or observed (Ekman, 1972). Even if all these kinds of gestures are not specific for any part of the body, it is also true that they often involve the use of hands and arms.

Also head movements were studied in order to investigate the possible correlations between emotions and head movements. For example, holding the head up can be interpreted as a sign of superiority (Horstmann, 2011). Moreover it seems that there exists a relation between expression of affections, head movements and tone of voice. Up head movements are associated with states of happiness and high tone of voice. Unlikely down head movements are associated with states of anger and low tone voice (Horstmann, 2011). However head movements include also epistemic states, for example uncertainty can be displayed with a lateral movement of the speaker's head (McClave, 1990). It is also interesting to see the movements of speaker's head activate the listener's backchannels as vocal feedback and head nods (McClave, 1990). If we examine the relations between words and head movements, the results of Boholm's study (2011) show that words are frequently spoken after or at the same time as multimodal nods, the duration of the nods is longer with words than without and that certain types of nods are associated with certain words, as *m* repeated nods and *okay* with up nods. Also prosody is associated with head movements. In fact, it seems that single up nods are displayed with long sentences and increasing pitch, while down nods occur with dropping or flat pitch (Boholm and Lindblad, 2011).

Undoubtedly, another important feature of *non-verbal communication* is proximity. By the term proximity we refer to the physical distance between two or more communicators. According to Hall's cultural differences studies (1959, 1963), too much or too little distance between two communicators can often lead to negative feelings (Hall, 1959) (Hall, 1963). Proximity can also have some effect during video conferencing. Interaction becomes more informal when people appear close to the camera (Grayson, 1998). Also body orientation, the degree to which a communicator's trunk, shoulders and legs are rotated in the same direction, or in the opposite direction of another communicator (Mehrabian, 1969) can influence communicators' attitudes and feelings. During face-to-face interaction, people tend to experience a more positive attitude toward each other if the other speaker orients her/his body in their direction (Mehrabian, 2006). When two people are talking and a third person has to judge the degree of communicators' positive attitude, the head orientation has an greater influence than body orientation (Mehrabian, 2006).

2.3.2. Facial expressions

The face is a part of our body that is always visible and active during a face to face interaction. It is composed of different parts that even on their own can offer some information about people's state, such as: mouth, eyebrows and various facial muscles. With all its parts, facial expressions are some the most useful features to interpret people's emotions and psychophysical conditions. Unlike other features of communication, people still express their state through the face also when they are alone (Cohn, 2004). Because of the importance of this part of our body, several studies have been made in order to study the different kinds of facial expressions, and how they influence the interpretation of people's states (mostly affective). Some studies of faces and emotional expression show that some facial expressions are easier to interpret and store than others (Nomi, 2012) (Hansen, 1988) (Fox, 2000). For example, angry faces are easier to interpret in happy crowds than happy faces in angry crowds (Hansen, 1988) (Fox, 2000). Similar results are presented also in research about visual short-term memory (Jackson , 2009). A recent study shows that people need more time in order to interpret disgust or angry faces than happy, neutral and sad faces (Chen, 2011). However this last study used a different methodology then the first two. Using the same methodology in order to study facial emotions is very important since also displaying faces in different angles can influence the emotional state interpretation (Bruce, 1982). Another important outcome is that people recognize and react (mirroring) unconsciously to facial expression. This result is presented in Dimberg's (2000) research about reactions to emotional facial expressions, where respondents were exposed to emotional faces for a very short time span (Dimberg, 2000). In conclusion, we can say that even if there are several studies about how facial expression influence

emotions interpretation, more studies have to be done in order to explore more deeply how facial expressions influence the interpretation of epistemic states.

2.3.3 Laughing and smiling

The Mouth is one of the most important parts of the face. Thanks to this organ we can create sounds and express emotions. Through the mouth we are able to display smiles and produce laughs. These two important communication features have a big influence in the recognition of people's states. Both laughing and smiling are very common in most daily interaction with other people. In our experiment, laughs and smiles are the most common actions used by respondents. Moreover these behaviors were often used as explanation for the detection of different *AES*, such as: happiness, nervousness, confidence, interest and disinterest. This shows that different kinds of smiles, laughs and combinations of them occur in different situation, and that they can have different meanings and interactional functions.

People smile in several different situations, and this action can have different intentions. One of the most important is its social function. Indeed, even if most of the time we relate the action of smiling with a feeling of enjoinment, children's smiles are more connected with their social interactions than with their state of happiness (Schneider, 1992) (Soussignan, 1996). Duchenne was one of the most important researchers, studying the action of smiling and its functions. He described the characteristics to differentiate smiles caused by enjoyment from other kinds of smiles. According to Duchenne, we can talk about a smile as an index of happiness when there is a simultaneous contraction of the zygomatic major and orbicularis oculi muscles (Russell, 2003). However, even if some research seems to agree with Duchenne' s categorization of different kinds of smiles (Messinger, 2001) (Ekman, 1993), some other studies show how the "Duchenne's smile" can occur also in a situation of failure (Schneider, 1991) or it can be simulated (Gunnery, 2012). According to Frank and Ekman (1993) besides "Duchenne's smile" there are four more distinctive markers to differentiate the enjoyment smile from all the other kinds: "symmetrical action of the zygomatic major on both side of face, zygomatic major actions which are smooth and not irregular, duration of zygomatic major that is consistent from one enjoyment smile to the next, and synchronous action of the zygomatic major and the orbicularis oculi such that they reach maximal contraction at about the same time" (Frank, 1993). Still, smiles also have other functions, for instance they can also be used to reply to a previous laugh or to indicate a delicate argument (Haakana, 2010). Often a smile can be interpreted as laugh, but we have to remember that there is not any vocal element during the smile action (Ruch, 2001). However occasionally it can occur in a combination of laughs and smiles. Haakana (2010) shows that sometimes smiles occur just before laughs and that they can be interpreted as a pre-laughing signal (Haakana, 2010).

On the other hand, concerning the action of laughing, we can say that it is not just a human behavior, because other primates and mammals giggle (Duglas, 2003) (Preuschoft, 1995). Laughing is part of vocal communication, but it is still part of *non-verbal communication*. Laughing doesn't include the action of producing words. However it can happen that these two activities overlap. According to Duglas (2003) and Ruch (2001) laughing is not just a behavior but it is a social activity and social signs can work as social glue (Duglas, 2003) (Ruch, 2001). However, it is not clear when laughter appeared in the human species, but most probably it was a human characteristic before speech (Preuschoft, 1995). As we know, there are several varieties of laughs and like other actions, it can be spontaneous, voluntary. It can also be a mixture of controlled and spontaneous laugh. If we carefully examine spontaneous and intentions laughs with clinical observation, we can see that the two behaviors are actually very different (Ruch, 2001). When we are laughing spontaneously we are following an uncontrolled impulse, our self-awareness and self- attention are reduced and mostly it is described as a pleasant event (Ruch, 2001). On the other hand when we are

laughing intentionally, we just reproduce (or we try to reproduce) a similar sound to a spontaneous laughter. Even if we can reproduce a fake laugh we cannot reproduce the affective state that the spontaneous one is able to produce. Also in this case Duchenne's study of smiling faces gave a contribution to several investigations about laughing. In fact the "Duchenne smile" is often produced during the action of laughing (Ruch, 2001). This result strengthens even more the correlation between laughing and smiling. However one difference between smiling and laughing is that since laughing embraces also sound, it include also the activity of several muscles in different part of the body, most of which are involved in the respiratory/vocal behavior.

2.4. Verbal communication

By the term *verbal communication* we want to refer at all those actions by which people are able to produce words. Therefore *verbal communication* can be divided in other three sub-categories: *vocal verbal communication*, *written verbal communication* and *gestural verbal communication* (deaf people sings). In order to be part of verbal communication a category has to have a *syntactic*, *semantic* and *pragmatic* component. Chomsky (1957) defines *syntax* as "the study of the principles and processes by which sentences are constructed in particular languages" (Chomsky, 1957, p.11). In linguistics the term s*emantics* indicates the study of the meanings of linguistic expressions. *Pragmatic* studies how people use the language in a specific situation, or better how the context can influence conventional meanings of the words (Moeschler, 2009). Transcription of face-to-face communication, doesn't have to be considered *written verbal communication*, but as *vocal verbal communication* displayed in a written modality. A propriety of vocal verbal communication is the possibility to express concepts and *AES* also with *prosody*. By this term we indicate the intonation, cadence, rhythm and length of people's *vocal verbal communication*. However thanks to the IPA (*International Phonetic Alphabet*) symbols and rules it is partly possible reproduce the *prosody* of words and sentences in a written system. In the following two parts we will illustrate how *semantics, pragmatics, syntax* can influence people's interpretation of *AES*.

2.4.1. Words: pragmatics, semantics and syntax

When we are talking with people, the context is a very important factor, and it influences our *verbal communication*. Using certain kinds of words instead of other kinds, in specific situations it we can affect our interlocutor understanding of *AES*. Therefore the same words or sentences pronounced in two different situations can give different impressions about people's states. If we take our recordings as an example, we know that there are some rules that we have to respect when we are meeting a person for the first time. Just imagine this conversation between two people that are meeting for the first time.

1) Person (A): I was in Milano last week-end because work.
2) Person (B): really? I love that city.
3) Person (A): actually my parents are from Milano and I grew up there. (little smile)
4) Person (B): ok. (smile) I mean, I have been there so many times because work too. I like the restaurants, the shops, the monuments, it is just the car traffic that makes me mad, and of course I love the Duomo. (little smile)

Probably this may looks like a normal conversation between two strangers. Now if we have to say which kind of *AES* person (A) and person (B) are showing during the sentences 3) and 4), perhaps we can say that since both are smiling and talking about Milano probably the two people are pleasantly surprised because by having something in common.
Now imagine the same situation.

1) Person (A): I was in Milano last week-end.
2) Person (B): really? Awful city.
3) Person (A): actually my parents are from Milano and I grew up there. (little smile)
4) Person (B): ok. (smile) I mean, I have been there so many times because work too. I like the restaurants, the shops, the monuments, it is just the car traffic that makes me mad, and of course I love the Duomo. (little smile)

Now the situation is totally different, and if we have to say which kind of *AES* person (A) and person (B) are showing during the sentences 3) and 4), for sure we cannot say that the two people are pleasantly surprised. Probably person (A) can be a little disappointed about the fact that (A) doesn't like his city. On the other hand now the words of person (B) and the fact that he is smiling, can suggest that actually he is embarrassed about the situation. Therefore if we analyze two identical sentences from a *pragmatic* point of view, we can see how they can transmit different information about people's states because the context where they were said. Moreover also the word *smile* had two different meanings in the two different situations. Smiling because pleasant feelings and smiling because of embarrassment.

On the other hand, when we talk about *semantics* we are refer to a part of linguistics that studies the relation between a word and its meaning. Using one word or sentence insted a similar one may influence the interpretation of people's *AES*. For example now just imagine that you present an idea to your boss and he/she answers you back by written post: "bad idea: too risky !" or "terrible idea: too risky !". These two affirmations have a similar negative meaning however they transmit a different impression about the possible *AES* of your boss. In fact even if both answers are negative, using the word terrible instead bad, could suggest that the boss was bothered after considering your idea.

However also the right use of *syntax* can influence the interpretation of people's *AES*. For example feelings like anger and nervousness can affect the right grammatical production of a sentence. A person that is expressing a concept using the right *syntax*, most probably can transmit more secureness about his/her state than a person that is talking in an incorrect grammatical way.

2.4.2. Prosody

Without any doubt, words have a big importance in *verbal communication*. However also the tone of voice and intonation of how we speak different words can give information about people's *AES*. Often we consciously choose specific words to use in a specific sentence for a specific purpose. But how we speak: intonation, cadence, rhythm and length, are frequently displayed spontaneously. The idea that the tone of voice can influence the interpretation of the speaker's state is very old, and it has been found even in the manuals of rhetoric dating back to the Roman and Greek age (Scherer, 2003). In an experiment conducted by Banse and Scherer (1996), the data showed that vocal parameters indicate the level of intensity of spoken words, and they are representative of each different emotion. According to Scherer (2003) the physiological changes caused by the emotional stimulation, will affect speaker's respiration, phonation, and articulation. These changes will produce distinctive emotional patterns of acoustic parameters (Scherer, 2003). A study review about the relation between expressing emotion via voice and via music, shows that music uses the same voice-parameters that permit the expression of emotions in human-human communication (Juslin, 2003).

3. Methodology

In this part we describe step by step, which kind of methodology we used to conduct our investigation.

We recorded four different encounters of people who did not know each other and were meeting for the first time. Four recordings were displayed to each participant. Each recording was showed in a different mode. The modes used for the experiment consisted of a transcriptions (T), a video with audio (V+A), a video without audio (V) and an audio recording (A). 12 subjects took part in the experiment (6 male and 6 female). The different stimulus conditions were presented to the subjects in random order. The different modes, thus, have been shown in this way:

Participant 1: V+A (rec1), A (rec2), V (rec3), T rec4)

Participant 2: T (rec1), V+A (rec 2), V (rec4), A (rec3)

Participant 3: A (rec4), V+A (rec3), T (rec2), V (rec1)

and so on …

The subjects were asked to give an interpretation of which *AES* were expressed in the four modes. The participants were also asked to give an explanation for their answers, explaining which factors led them to think that a particular person in the recordings or in the transcription was expressing a particular *AES*.

3.1. Participants

The participants were selected following some specific criteria: they had to be native speakers of Swedish, they had to be at least 20 years old and they did not have any study background in the communication field. In the end we selected 12 participants (6 men and 6 women) and their identity is anonymous.

3.2. Ethical Considerations

All of the 12 participants that took part in this experiment were aware of the nature and the goal of the research. All of them voluntarily gave their consent to take part in the investigation. They were free to leave the experiment at any time they wanted. The participants' identity is anonymous, but it was revealed to the people that collected the data. Also participants in the recordings are anonymous, but some of them revealed their names and some personal information (address, birthplace) during the conversation, however they gave their permission to use the recordings for academic research.

3.3. Experiment procedure

The participants were interviewed one at a time. The study took place in different places and most of the time it was conducted at the university. During all the experiments, the researchers and the participant were alone in a room, in order to not be interrupted and to be focus just on the experiment. No one disturbed the experiments. To show the recording and the transcriptions the researcher used a laptop. The computer sound level was good and the participants did not encounter any problem in understanding the recordings. The laptop screen displaying the video was 12 inches.

The resolution of the videos was excellent but one video was a little blurred compared to the others. However also in this case the respondents did not indicate any problem concerning the video quality.

At the beginning of the experiment, the goal of the study was explained to all participants including what they had to do. The meaning of the term *AES* was briefly explained. The participants were told to interpret all those states which include both knowledge and feelings (Schroder, 2011). After this brief introduction, the experiment started.

The respondent were randomly showed one *video+audio* recording, one *video* without audio, one *audio* without any video and one *transcription*. One of the four version of the recordings was shown to each respondent, The recordings are all about two strangers that are meeting for the first time in a room and they were recorded The excerpts we showed are all about 2 minutes long, with a total average of 2 minutes and 7 seconds. In two of the recordings the beginning was cut, where the two persons were greeting each other. Each video will be analyzed more in depth.

The experimental procedure for *video+audio*, just *video* and just *audio* versions were identical. The procedure for showing the *transcription* was little different. The procedure for the first three was as follows: the recording was stopped after every three or four contributions. It means that each person in the video talked (or gave some kind of verbal contribution) at least once. The recordings were stopped with an average of 15.5 times, an average of 8.3 seconds. Every time that a recording was stopped, the respondent had to say which kind of *AES* were recognized (if he/she recognized some) and explain why he/she recognized it. When the *transcription* was displayed, the respondent had to indicate the part of the *transcription* where he was able to recognize some *affective epistemic-states*. The procedure followed for transcriptions will be presented in depth in the next section. The sessions of the experiment lasted around 75 minutes per participant.

3.4. Data coding

During the experiment all the respondents' answers were transcribed manually. Afterwards all the data was digitized in electronic form. Very different answers were given by respondents, they were free to use any words they wanted to describe the *AES* that they identified. Some answers were very specific, some others were very general. Moreover not all the respondents were able to give a complete answer. For example some of them just mentioned an *AES* without giving an exhaustive reason for their answer. Others were able to explain what was going on in the video and the actions of the people involved, without being able to give a complete description or "label" for the *AES* identified. It is important to mentiom that a lot respondents expressed their difficulties in finding the right words to describe or label the emotions, feelings and epistemic states. Despite this, to get answers without any influence, we preferred to not push the respondents to answer, but let them answer freely in the manner best suited to them.

Due to all these varieties of data we had to find ways to merge all similar answers. Several different *AES* have been found, such as: curiosity, sadness, thoughtfulness, aggressiveness, excitement, compassion, informative, sarcastic, proudness, reservedness. However we decided to focus on just on the most five common *AES*, which are:

- Happiness: we have included in this group all those states that concern the emotion or the feeling of happiness, contentment, gladness, joy and the enjoyment in doing a determinate thing.

- Interest: we have included in this group all those states that concern interest, curiosity, listening, seeking for more information and pay the attention to something.

- **Nervousness:** we have included in this group all those states that concern the emotion or the feeling of nervousness, uncomfortableness, insecurity, uneasiness, tension, embarrasses, shyness.

- **Confidence:** we have included in this group all those states that concern the emotion or the feeling of confidence, relaxation, comfortableness, security.

- **Disinterest:** we have included in this group all those states that concern disinterestedness, indifference, being tired, bored or being annoyed about something.

Since we had to deal with very different answers, and some of them were not complete, it was necessary to interpret the data. For example sometimes the respondents used the terms *listening* to refer at a certain *AES*. However this term can be used in different ways, so it is very important to understand why a respondent gave a certain kind of answers. For instance, if a respondent would say: " Person (A) is *listening* (B), he is nodding and giving verbal feedback". In this case most probably the answer *listening* mean that person (A) is *interested* in the conversation. It would be different if the respondent would answer: "Person (A) is *listening* just because he wants be polite". Most probably in this case person (A) is *not very interested* in the conversation, but he doesn't want to show it. So in this case the *AES* is the opposite of the previous one.

Also the data about how *AES* were displayed needed to be carefully interpreted. Also in this case we got very different answers by our respondents. Some of them were more or less precise than others. For instance a very complete and exhaustive answer could be this: " Person (A) is nervous because he is scratching his nose ". On the other hand a less precise answer could be: " Person (A) is nervous because he is moving his hand ".

In the ***complete tables***, are coded all the complete answers that we got from the respondents. In order to be reported in the table the answers have to have both the *AES* and a explanation associated with it.

3.5. Transcription

The transcriptions of the recordings were made by university students and employees. The original transcription shows all the words, verbal and nonverbal actions displayed by people involved in the recordings and some comments of the transcriber about their attitude, feelings and states as for example: anxiety, frustration, satisfaction, etc. Since the aim of this part of the study was to understand how written words can influence respondent's interpretation of *AES*, we had to modify the transcription and delete those comments that could influence respondent's interpretation.

Here we will present a part of a transcription without any transcriber's comments:

L: <1 ja >1 / <2 men det e0 <3 inte det >3 [10 <4 pause >4 >2]10
@ <1 general face: smile >1 ; <1 gaze: to interlocutor >1
<2 speech: exclaim >2 ; <2 general face: smile >2 ; <2 gaze: to interlocutor >2
<3 head gesture: repeated shake, emphasizing >3
<4 chuckle >4
K: [10 < nä >]10 <1 det började som e1 danska // med pålagrad värmländska >1
@ <1 speech: state >1 ;
<1 gaze: to interlocutor >1

L: < ja'a: okej >
< : up nod > ; < comment: up-down f0 >

Before starting the experiment the respondents were instructed how to read the transcription and what they were supposed to do. We underlined with two different colors the speech of the two people involved in the recording. After the @ sign and between < > signs, are displayed all of the person's verbal and nonverbal actions during his/her contribution. The numbers indicate the moment when a specific action was performed during the verbal contribution. Unlike the other representations, in the transcription there were no stops. In this case the respondents just had to answer which kind of *AES* was present in the transcription. They also had to indicate which part of the transcription suggests to them that *AES*.

This is an example of an answer that we got from one of our respondents:

L: < ja'a: okej >
< : up nod >

Respondent's answers: *Surprise*

In this case the combination of spoken words and body movement suggested to the respondents that (L) was surprised.

It is important to underline that due to the presence in the transcription of a high number of numerals, and since often words are transcribed as they sound and not as they are spelled, some of the respondents had problems in understanding the thread of the conversation presented in the transcriptions.

3.6. Presentation of recordings

Recording Figure 1

In order to give a general idea of what the stimulus recordings displayed, we will present each recording. Firstly we can say the recordings are all fairly different in showing *AES*. In some recordings, people are very active and exhibit a lot of verbal and non-verbal behaviors. In some others, people are quiet and show less behaviors.

In the recordings the camera recorded the two persons laterally, therefore we have to be aware that also the camera position can influence the respondents' perception of *AES*.

Recording V8602

Length: 2 minutes and 19 seconds.

Gender present in the recording: 2 females

Names: (A) person on the left side, (B) person on the right side

Number of stops: 20

Average time stop: 7 seconds

Brief description of the recordings: in this recordings two women are meeting for the first time. The quality of the video is good. It is clearly possible to see the faces of people involved when they turn to the camera. The participants' bodies are fully visible. The audio is good and it is very easy understand what people say and what they are talking about. Just in the beginning of the video there is some audio noise for a couple of seconds, probably due to some camera movements. The recording was displayed from the beginning and it has not been previously cut. The video starts with the two women that introduce themselves to each other. The person (B) starts immediately by posing a question to person (A). Since they pose questions to each other and there are no long pauses the two women seem engaged in the conversation and their interaction is quite good. However, just a few body movements are showed in this recording. Even if the people seem to be quite fluent in talking, their posture could suggest that they are little uncomfortable about the situation. Both women often look at the camera and in the end of the video person (A) asks a direct question to the cameraman.

STATES	V+A	Video	Audio	Transcription	Total
LAUGHING, SMILING – happy	7	9	2		**18**
GIVING FEEDBACK – interest	8		8		**16**
ASKING QUESTION – interest	7		6	2	**15**
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - confident	6		7	1	**14**
STANDING – nervous	8	6			**14**
LOUD VOICE – confident			10		**10**
EYE CONTACT – interest	8	1		1	**10**

Complete Table 1 – combinations of states/explanations reported at least ten times in the recording V8602

In **Complete Table 1** all the interview replies containing combinations states/explanation cited at least ten times are presented. These results suggest that the two women in the recording are *interested* in the conversation. In fact they are sharing feedback and asking questions to each other several times. Moreover they are laughing and smiling very often and for this reason the respondents claimed that they seem to be *happy*.

Recording V8640

Length: 2 minutes.

Gender present in the recording: 1 female and 1 male.

Names: (A) woman on the left side, (B) men on the right side

Number of stops: 13

Average time stop: 9,23 seconds

Brief description of the recording: in this recording a man and a woman are meeting for the first time. The quality of the video is very good and it is quite easy to see the faces of the people involved. The participants' bodies fully visible. The audio is good and it is very easy to understand what the two person say and what they are talking about. The level of the voice of the woman is a little lower than the voice of the man, but none of the respondents reported problems in hearing what she said. The recording was displayed from the beginning but it had been previously cut in the end, after two minutes. At the beginning of the video the two persons introduced themselves to each other. After the greeting, person (B) posed a question to person (A). The interaction between the two participants is good. They seem to enjoy the conversation, both of them are laughing and smiling quite often. Since several questions have been made by both parts, they also seem being interested in the conversation. Numerous body movements have been showed by both participants, this can mean that the two persons are nervous about the situation. In fact, both (A) and (B) are not very fluent in talking and by this characteristic we can deduce the presence of *nervousness* during the interaction.

STATES	V+A	Video	Audio	Transcription	Total
STUMBLING, HESITATIONS, REPETITION - Nervous	4		6	5	**15**
MEANING OF WORDS – Nervous	1		2	9	**12**
MOVING HANDS AND ARMS- Nervous	5	6		1	**12**
LAUGHING, SMILING – Nervous	1	3	5	2	**11**
CLEAR VOICE – INTEREST	1		10		**11**
LAUGHING, SMILING – HAPPY	2	5	4		**11**
STANDING – CONFIDENT	5	5			**10**

Complete Table 2 – combinations of states/explanations reported at least ten times in the recording V8640

The combinations of behaviors and *AES* in **Complete Table 2** suggest that the two people involved in the recording appear quite nervous. *Nervousness* is displayed in several ways: through voice, body movements and words. Also in this recording *happiness* is shown mostly by *laughing and smiling*. *Talking clearly* was selected as a sing of *interest*.

Recording V8644

Length: 2 minutes and 1 second.

Gender present in the recording: 1 female and 1 male.

Names: (A) men on the left side, (B) woman on the right side.

Number of stops: 15

Average time stop: 8.07 seconds

Brief description of the recording: in this recording a men and a woman are meeting for the first time. The quality of the video is very good and it is possible to see the faces of the people involved. The participants' bodies entirely were framed by the camera. Also in this case the audio is good and it is very easy understand what the two person say and what they are talking about. The level of the

voice of the man is little lower than the voice of the woman, but none of the respondents reported problems in hearing what he said. The recording had been previously cut at the beginning and at the end, for a total two minutes and one second. Since the video has been cut at the beginning there is no greeting between the two persons. Person (A) is talking with a low voice. Person (B) is giving feedbacks and following what (A) is saying. The interaction between the two participants seems to be difficult. Pauses occur very often and both people look quite shy. Rarely they have eye contact. Person (A) drives the conversation for almost all the video. Person (B) appears interested in the conversation because she is seeking more information about (A) and giving a lot of feedbacks through voice and body movements. Numerous non-verbal communication behaviors are shown by both participants and from this we can deduce that they are nervous about the situation. Nevertheless, they laughing and smiling quite often, so probably they are enjoying the conversation anyway.

STATES	V+A	Video	Audio	Transcription	Total
LAUGHING, SMILING - HAPPY	2	16	10	2	**30**
LAUGHING, SMILING - Nervous	7	2	10	2	**21**
LOOKING AWAY, DOWN - Nervous	10	5		2	**17**
MEANING OF WORDS - Nervous			4	8	**12**
NODDING - INTEREST	2	9			**11**
MOVING HANDS AND ARMS- Nervous	5	3		2	**10**
EYE CONTACT - INTEREST		10			**10**

Complete Table 4 – combinations of states/explanations reported at least ten times in the recording V8644

Complete Table 4 shows that the two participants involved in the recording appear quite nervous. Also in this case *nervousness* is displayed through both verbal and non-verbal behaviors. Similarly to the other recordings, laughing and smiling were selected by respondents as an indicator of *happiness*. *Interest* was identified through *head movements* and *eye contact.*

Recording V8604

Length: 2 minutes and 8 seconds

Gender present in the recording: two females

Names: (A) woman on the left side, (B) woman on the right side

Number of stops: 14

Average time stop: 8.36 seconds

Brief description of the recording: in this recording two women are meeting for the first time. The quality of the video is not perfect, but it is good enough in order to see the faces of the people's involved. However, none of the respondents referred problems about the video quality. The participants' bodies entirely were fully visible. The audio is good and it is very easy to understand what the two person say and what they are talking about. The level of the voice of both women is quite good. Also in this case, the video had been previously cut at the beginning and at the end, for a total two minutes and eight seconds. Since the video had been cut at the beginning there is no greeting between the two persons. The participant's interaction doesn't seem very smooth. Pauses occur quite often and both people seem to be little uncomfortable and nervous. Several body movements are shown by both participants. The recording begins with person (A) posing a question

to person (B). Person (B) is responding to the question, and she is asking more information in turn. In the beginning person (A) drives the conversation, but successively person (B) takes control of the discussion. The conversation goes better when one of the participant is talking about some specific argument. Very often the two women are laughing and smiling. However the participants seems to be quite involved in the conversation, in fact both the women are asking questions to each other and giving quite few feedback.

STATES	V+A	Video	Audio	Transcription	Total
LAUGHING, SMILING - Nervous	3		8	8	**19**
LAUGHING, SMILING - HAPPY	7	4	2	5	**18**
EYE CONTACT - INTEREST		15			**15**
LOOKING AWAY, DOWN - Nervous	2		5	7	**14**
LEANING ON THE WALL - CONFIDENT	2	11			**13**
MEANING OF WORDS - Nervous	3		5	5	**13**
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - CONFIDENT			11		**11**
MEANING OF WORDS - INTEREST	4		6	1	**11**
GIVING FEEDBACK - INTEREST	6		2	2	**10**

Complete Table 5 – combinations of states/explanations reported at least ten times in the recording V8604

Complete Table 5 shows a mixture of *AES*. Also in this case *happiness* is displayed mostly through *laughing and smiling*. *Confidence* is shown by posture and verbal communication. Talking about own experiences and creating coherent sentences can transmit to other people a sense of secureness. An high level of *nervousness* is displayed during the conversation, mostly by *laughing and smiling*, but also by verbal and non-verbal behaviors. The *AES* of *interest* it was identified by *eye contact*, *words* and *vocal feedback*.

4. Results

In this section we will illustrate the results of our research. Firstly, we will illustrate the most common behaviors, actions and sounds that influenced the understanding of *AES*. We will see how the same behavior, action, sound and word can be interpreted in different ways, and therefore associated to different *AES*. We will also see how, some behaviors, actions, sounds and words are just typical of certain *AES*. We will focus on each *AES* and we will discuss how different modes influence their interpretation. According to the participants' answers, our data presents several *AES*, but we will focus on those most cited by our respondents: *happiness*, *interest*, *nervousness*, *confidence* and *disinterest*.

In the **Complete Table** it is possible to see the numbers of times the same *AES* is displayed in different modes. *V+A* is the mode with the highest number of *AES*, and it is followed by VIDEO mode, AUDIO mode and TRANSCRIPTION. If we observe the TOTAL number of interpretations, we can see that the V+A multimodal mode, in comparison with the other unimodal mode, permits to interpret an higher number of *AES*.

STATES	V+A	Video	Audio	Transcription	Total
NERVOUS	142	80	73	76	372
INTEREST	96	94	81	21	292
CONFIDENCE	75	82	103	16	276
HAPPINES	46	65	36	15	162
DISINTEREST	18	8	24	9	59
TOTAL	377	329	317	137	1161

Complete Table - the table shows how many times the same AES is displayed in different modes.

In the following section we will present the 4 most common behaviors, actions and sounds perceived in each recording, divided into each *AES* (**Complete Tables 1**). In this way it will be possible to observe if respondents were influenced by behaviors, actions and sounds in similar ways in different recordings. We will also present another table (**complete Table 2**) where we will illustrate all the behaviors, actions and sounds that were cited at least 10 times in total. Due to the large amount of data in the table, in this paper we decided to show just the most common taxonomies. The full tables are available in the appendix.

4.1. Nervousness

If we think about nervousness, we can classify it as an affective state but, it is possible to find several epistemic features in this state. In fact, also in this case, people are nervous because of some specific reason that may or may not be known. People in the video were probably nervous because they were in a new situation, where they had to talk with a person that they didn't know.

Nervousness is the most common state found in our data. The respondents gave several explanations about the reason why they interpreted *nervous* behaviors.

Nervouss, insecure, tense, embarrassed 8644	V+A	V	A	T	Total
LAUGHING, SMILING - NERVOUSS	7	2	10	2	21
LOOKING AWAY, DOWN - NERVOUS	10	5		2	17
MEANING OF WORDS - NERVOUS			4	8	12
MOVING HANDS AND ARMS – NERVOUS	5	3		2	10

Nervouss, insecure, tense, embarrassed 8640	V+A	V	A	T	Total
STUMBLING, HESITATIONS, REPETITION - NERVOUS	4		6	5	15
MEANING OF WORDS - NERVOUS	1		2	9	12
MOVING HANDS AND ARMS - NERVOUS	5	6		1	12
LAUGHING, SMILING - NERVOUSS	1	3	5	2	11

Nervouss, insecure, tense, embarrassed 8602	V+A	V	A	T	Total
STANDING - nervous	8	6			14
STUMBLING, HESITATIONS, REPETITION - nervous	4		4	1	9
MOVING HANDS, ARMS - nervous	1	6			7
MOVING A LOT - nervous	1	3			4
LAUGHING, SMILING - nervous			4		4

Nervouss, insecure, tense, embarrassed 8604	V+A	V	A	T	Total
LAUGHING, SMILING - Nervous	3		8	8	19
LOOKING AWAY, DOWN - Nervous	2	5		7	14
MEANING OF WORDS - Nervous	3		5	5	13
MOVING A LOT - Nervous	7			1	8

Nervousness Complete Tables 1 – The tables show the four most common behaviors reported in each recording, as signs of *nervousness*

TOTAL Nervous, insecure, tense, embarrassed	V+A	Video	Audio	Transcription	Total
LAUGHING, SMILING	11	5	27	12	**55**
LOOKING DOW, AROUND, AT CAMERA	15	16	0	11	**42**
MEANING OF WORDS	4	0	13	23	**40**
STUMBLING, HESITATIONS, REPETITION	12	0	12	11	**36**
MOVING HANDS AND ARMS	11	17	0	3	**31**
MOVING A LOT	11	8	0	2	**21**
STANDING	14	6	0	0	**20**
HANDS IN THE POCKET, HOLDING	10	5	0	4	**19**
MOVING FEET	11	4	0	1	**16**

Nervousness complete Table 2 - the table shows the most common behaviors reported at least ten times, as signs of *nervousness*.

Laughing and smiling are the most common behaviors selected by our respondents. These two behaviors appear in the first four positions in all the recordings. In the **Nervousness complete Table 2** we can notice that both *verbal* and *non-verbal communication* behaviors are cited several times by the participants. *Looking around* (and therefore avoiding eye contact) is another common behavior that

influenced the interpretation of *nervousness*. The perception of this *AES* is also influenced by *semantic* and *pragmatic* features of communication. According to our data, the *words* used during the conversation are very useful in order to understand if a person is nervous or not. The way people speak has big importance too, and behaviors such as: *stumbling, repetitions and hesitations* have been mentioned by several respondents as symptoms of *nervousness*. Concerning *body movements*, we can say that both *moving too much* and *not moving at all* can be considered as signs of *nervousness*. However if we sum all the body movements (hands, feet and moving a lot) we can see that bodily communication is extremely important to interpret *nervousness*.

Nervousness complete tables divided by modalities

Nervous, insecure, tense, embarrassed	V+A
LOOKING DOW, AROUND, TO CAMERA	15
STANDING	14
STUMBLING, HESITATIONS, REPETITION	12
MOVING HANDS AND ARMS	11
MOVING A LOT	11
LAUGHING AND SMILING	11
MOVING FEET	11
HANDS IN THE POCKET, HOLDING	10
GENERAL MOVEMENTS	5
SCRATCHING TOUCHING BACK, FACE, ETC	5
TWISTING BODY, SHAKING	5
MEANING OF WORDS	4
TALKING TOO MUCH	4
VOICE	3
LEANING BACKWARD	2
LOUD VOICE	2
LAUGHING POLITE	2
TALKING SLOW	2
BREAK AFTER GREETING	2
EYE CONTACT	1
NODDING	1
SHORT ANSWERS AND FEEDBACKS	1
GIVING VOCAL FEEDBACKS	1
LEANING FORWARD	1
THE OTHE PERSON DOESN'T GIVE FEEDBACK	1
MIRRORING	1
NOISY VOICE	1
STOP TO TALK	1
BECAUSE THE OTHER PERSON IS NERVOUS	1
DOWN EYEBROW	1
TOTAL	**142**

Nervous, insecure, tense, embarrassed	Video
MOVING HANDS AND ARMS	17
LOOKING DOW, AROUND, TO CAMERA	16
MOVING A LOT	8
STANDING	6
LAUGHING AND SMILING	5
HANDS IN THE POCKET, HOLDING	5
MOVING FEET	4
GENERAL MOVEMENTS	4
SCRATCHING TOUCHING BACK, FACE, ETC	4
STOP TO EXPLAIN, TO SPEAK	3
TWISTING BODY, SHAKING	2
MOVING LEGS, FEET	2
ARMS DOWN	2
LEANING BACKWARD	1
EYE CONTACT	1
NODDING	1
LIFTING HEELS	1
LOW EYEBROW	1
SHAKING	1
TWISTING ARMS	1
TOTAL	**85**

Nervous, insecure, tense, embarrassed	Transcription
MEANING OF WORDS	23
LAUGHING AND SMILING	12
STUMBLING, HESITATIONS, REPETITION	11
LOOKING DOW, AROUND, TO CAMERA	11
HANDS IN THE POCKET, HOLDING	4
GIVING VOCAL FEEDBACK	3
MOVING HANDS AND ARMS	3
SHORT ANSWERS AND FEEDBACK	2
MOVING A LOT	2
TALKING TOO MUCH	2
MOVING FEET	1
TWISTING BODY, SHAKING	1
HE DOESN' T WANT TO INTERUPT	1
TOTAL	**76**

Nervous, insecure, tense, embarrassed	Audio
LAUGHING AND SMILING	27
MEANING OF WORDS	13
STUMBLING, HESITATIONS, REPETITION	12
LOUD VOICE	2
SHORT ANSWERS AND FEEDBACK	2
LOW VOICE	2
TALKING FAST - NERVOUS	2
VOICE	1
GIVING FEEDBACK	1
NEUTRAL VOICE	1
SHORT VOCAL FEEDBACK	1
SPEAKING FAST	1
TAKLING LITTLE	1
TONE OF VOICE UP AND DOWN	1
BREATHING DEEPLY	1
TOTAL	**68**

4.1.1. How different modes influenced the interpretation of nervousness

According to our results, we can see that the *V+A mode* leads to more interpretations of *nervousness* than all the other modes (*V+A* 142 interpretations, *Video* 85, *Audio* 68, *Transcription* 76). During displaying the *V+A mode*, respondents perceived both *movements* and *sounds*, therefore it seems that multimodality in this case adds much more information than unimodality. If we compare the unimodal modes to each other, we can see that the *video mode* got the highest number of interpretations. This result is confirmed also in *V+A mode*, where the respondents selected mainly *body movements*, such as*: looking around*, *standing*, *moving arms*, *hands and foot*. The majority of *bodily movements* were recognized more easily in the *V+A mode*, when *video* and *audio* were displayed together. Therefore, the respondents selected many more times the same behavior in *V+A mode* than in the *video* mode. The only movements that got more identifications in the *video mode* than in *V+A mode* are the *movements of hands, arms and head* (*movements of hands*, 17 interpretations in *video* and 11 in *V+A*) (*looking down around and/or in the camera*, 15 interpretations in *V+A* and 16 interpretations in *video*). In this case *laughing and smiling* because of *nervousness* is very hard to recognize without the support of the *sound* (just 5 interpretations in *video mode*). Moreover, this behavior is harder to recognize when it is displayed in *V+A mode* than in *audio mode* (11 interpretations in *V+A* and 27 interpretations in *audio mode*). Therefore, when we are talking about *laughing and smiling* we can say that the interpretation of *nervousness* is strongly influenced by the presence of both *sounds and video images*. This result is also confirmed by the similar numbers of interpretations in *V+A mode* and *transcription* (11 interpretations in *V+A mode* and 12 in *transcription*). If we compare *video mode* with the other unimodal modes, it seems that respondents are less likely to interpret smiling and laughing as sings of *nervousness* when they are perceived just through the *visual modality*. We have a similar situation if we focus on *spoken words*. According to our data, respondents are more likely to recognize *vocal words*, as signs of *nervousness*, when they are perceived without *video images*. In fact, they got much more interpretations in *audio* and in *transcription* modalities than in *V+A* (23 interpretations in

transcription, 13 in *audio* and just 4 in *V+A*). It is possible that *prosody* and *vocal sounds* have a strong influence on people perception. The same spoken words, perceived by visual *written words* in the *transcription*, will be perceived as signs of another *AES* when they were perceived by just *auditory modality*. This hypothesis can be confirmed by the other *AES* discussions, where *spoken words* have more influence in *audio mode* than in *transcription*. However, the presence of *video images* doesn't influence the perception of *hesitations, stumbling and repetitions*. These actions got the same number of interpretations when they were displayed in *V+A* and *audio* modes (12 interpretations in *V+A* and *audio*). Similar outcomes are also present in the *transcription* (11 interpretations). In the end, we can say that multimodality sometimes increases significantly the influence of non-verbal behaviors, and sometimes decreases significantly the influence of vocal behaviors. However due to the possibility of receiving both *video images* and *sounds* together, it seems that multimodality increase the possibility to perceive *vocal sounds* and *movements* as sings of the *AES* of *nervousness*. Since there are not so many studies made on interpretation of *nervousness* it is difficult give an explanation to this characteristic of multimodality. However, it seems that several actions, movements and sound that we produce when we are *nervous* are involuntary and very hard to mask, therefore *nervousness* is probably very hard to hide and therefore quite easy to interpret.

4.2. Happiness

Besides being an affective state, *happiness* can also have an epistemic component. If we think about the feeling of happiness, from our daily life, we can observe that people are always happy for some reason and sometimes have some doubts about the real reason of their happiness. The epistemic aspect of happiness is related to the object of happiness and reason for happiness.

During our data collection we were able to get several answers about how and why different respondents interpreted a particular behavior or action as happy. In this section we will present the most common behaviors, actions, sounds and words that influenced the participants' interpretation of the *AES* of *happiness*.

Happy, Glad, Enjoying the Conversaion 8644	V+A	V	A	T	Total
LAUGHING, SMILING - HAPPY	2	16	10	2	30
MOVING HANDS AND ARMS- HAPPY		5			5
GENERAL MOVEMENTS HAPPY		4			4
LOUD VOICE - HAPPY			3		3

Happy, Glad, Enjoying the Conversaion 8640	V+A	V	A	T	Total
LAUGHING, SMILING - HAPPY	2	5	4		11
MEANING OF WORDS - HAPPY	2		3		5
EYE CONTACT - HAPPY	1	4			5
LEANING FORWARD - HAPPY	1	1			2

Happy, Glad, Enjoying the Conversaion 8604	V+A	V	A	T	Total
LAUGHING, SMILING - HAPPY	7	4	2	5	18
NODDING - HAPPY		2		2	4
LOUD VOICE - HAPPY	2		2		4
LEANING FORWARD - HAPPY	2	2			4

Happy, Glad, Enjoying the Conversaion 8602	V+A	V	A	T	Total
LAUGHING, SMILING - happy	7	9	2		18
LOUD VOICE - happy	3		1		4
AGREE - happy		2			2
MOVING HANDS, ARMS - happy	2				2

Happiness **Complete Tables 1 – The tables show the four most common behaviors reported in each recording, as signs of *happiness***

TOTAL Happy, Glad, Enjoying the conversation	V+A	Video	Audio	Transcription	Total
LAUGHING, SMILING	18	34	18	7	**77**
LOUD VOICE	6		6		**12**
MEANING OF WORDS	2		5	3	**10**

Happiness **complete Table 2 - the table shows the most common behaviors reported at least ten times, as sign of *happiness*.**

As we can observe by the **Happiness complete Tables 1** the answer "*laughing and smiling*" is the most common in all four recordings. However there are other behaviors that influenced the interpretation of *happiness*. According to **Happiness complete Table 2** *speaking loudly* is another behavior that permits to interpret *happiness*. In fact, also this characteristic is cited in all four recordings, but in one it is cited just once (check the appendix for the full tables). The category "*laughing and smiling"* is the most common in all modes. Concerning laughing, we can say that it is the only behavior that can be perceived through *visual* and *auditory modalities*. If we observe this behavior just in *Video mode*, we can note that respondents that watched the video recording without any sound, recognized these behaviors nearly twice as often in comparison with those that watched the *V+A* mode. Instead, if we observe the numbers of *audio mode* and *V+A*, we can notice that they are quite similar. "*Laughing and smiling"* is the most common choice in both modes and it was selected the same number of times. As previously mentioned, *loud voice* is another common explanation about how people show their happiness. It was selected 6 times in the *V+A mode* and 6 times in the *audio mode*. The *AES* of *happiness* can be interpreted also by the *meaning of words*. In this case we want to indicate a

precise pattern of *verbal communication*, the respondents directed the answer to *happiness* because of the semantic meaning of the words pronounced during the conversation. This answer was selected 2 times in *V+A*, 5 times in *audio* and 3 times in *transcription* modes.

Happiness complete tables divided by modalities

Happy, Glad, Enjoying the conversation	V+A
LAUGHING, SMILING	18
LOUD VOICE	6
LEANING FORWARD	3
MEANING OF WORDS	2
MOVING HANDS AND ARMS	2
CLEAR VOICE	2
RECOGNIZING	2
LEANING BACKWARD	2
EYE CONTACT	1
OPEN GESTURE	1
LIGHT VOICE	1
LOOKING DOW, AROUND, TO CAMERA	1
MIRRORING	1
TAKING INITIATIVE	1
TALKING FAST	1
VOICE GOING UP AND DOWN	1
WAITING THE END OF THE SENTENCE	1
TOTAL	46

Happy, Glad, Enjoying the conversation	Audio
LAUGHING, SMILING	18
LOUD VOICE	6
MEANING OF WORDS	5
LOW VOICE	2
TONE OF VOICE	2
CLEAR VOICE	1
ASKING QUESTION	1
SPEAKING FAST	1
TOTAL	36

Happy, Glad, Enjoying the conversation	Video
LAUGHING, SMILING	34
MOVING HANDS AND ARMS	5
GENERAL MOVEMENTS	5
NODDING	4
EYE CONTACT	4
LEANING FORWARD	3
AGREE	2
MOVING HANDS TO EXPLAIN	2
RECOGNIZING	1
OPEN GESTURE	1
LEANING ON THE WALL	1
LOOKING UP	1
SHOULDER DOWN	1
TOUCHING CHIN	1
TOTAL	65

Happy, Glad, Enjoying the conversation	Transcription
LAUGHING, SMILING	7
MEANING OF WORDS	3
NODDING	2
MOVING HEAD FORWARD	2
HE/SHE KNOW WHAT HE/SHE IS TALKING ABOUT	1
TOTAL	15

4.2.1. How different modes influenced the interpretation of happiness

In our data it was possible to observe that *happiness* is displayed in different ways. Firstly we can say that *video mode* (65 interpretations) seems to provide more cues for *happiness* than *audio mode* (36 interpretations) and *transcription* (15 interpretations). Moreover *video mode* got a higher number of interpretation than *V+A mode*. Therefore it looks like that for the respondents it is more easy to interpret *happiness* in *unimodal video mode* than the *multimodal V+A mode* (46 interpretations). However, if we look with more attentions at the results, we can see that actually the *AES of happiness* is strongly related to the presence of the *sound*. The respondents' interpretation of

this state is strongly influenced by *laughing and smiling* and they are interpreted in different ways depending on the mode to which they were displayed. *V+A* and *audio modes* gave the same numbers of interpretations (18 times each mode) of *laughing and smiling*, while *Video mode* gave almost twice as many (34 times). This means that these behaviors were interpreted in different ways when they were displayed with or without the presence of *sound*. Thus, when *laughing and smiling* are displayed in *V+A mode*, it seems that our respondents paid more attention to *sounds* than *video images*. This result is also confirmed by the same number of *loud voice* answers, in *audio modality* and *V+A modality*. According to Duchenne, the real smile because happiness consists of in the simultaneous contraction of the zygomatic major and orbicularis oculi muscles (Russell, 2003). By our results, it seems that people are more likely to interpret a smile as an Duchenne smile, when it is displayed just through the *video mode* without the presence of *sound*. Also the *prosodic component* of verbal behaviors gave its contribution in the interpretation of this *AES*. Concerning the volume of the voice, we can say that *loud voice* was the second most common behavior (it was mentioned 6 times in the *V+A* and 6 times in the *audio* modes). According to Juslin and Laukka (2003) a "medium–high voice intensity/sound level" is considered a sign of happiness (Juslin, 2003). In conclusion, people perceive behaviors, sounds and actions depending on the mode to which they are displayed. Therefore we can say that different modes influence the interpretation of *AES of happiness* in different ways. *Laughing and smiling* are the most common behaviors interpret by our respondents in each mode. Compared with other modes, *written words* don't seem to have a big influence on the interpretation of happiness. This could be related to the fact that respondents are strongly influenced by *sounds* and *volume of voice*, *V+A mode* got 10 interpretations more than *audio mode*, mostly because the possibility to also see body movements. Therefore, in this case we can say that multimodality could work as a filter, allowing us a better interpretation of the real *AES of happiness*. During the presentation of *video mode*, several *video images* of behaviors and actions were interpreted as signs of *happiness*. In *V+A mode*, these behaviors and actions are perceived and filtered by *auditory modality* and often interpreted as other *AES*. Even if the combination of *video images* and *sounds* permits the perception of both vocal and non-vocal behaviors, if we pay attention to the numbers in our results, we can notice that the influence of the *video images* decreases significantly, when they are displayed together with *sound*.

4.3. Interest

Since the *state of interest* involves primarily understanding and knowledge, it can be defined more as an epistemic state than an affective state. However, when we are in this particular state, it can create inside us some kind of affect too. For example, watching a dramatic film can create certain kinds of emotion. These feelings are strongly dependent on our level of interest in the film. A uninteresting dramatic film will create different kinds of emotion than an interesting dramatic film.

During our data collection, we got several answers about how and why our participants selected *interest* as *AES*. In this part we will present the most common behaviors, actions, sounds and words that influenced the participants' interpretation of the *AES* of *interest*.

Interested, Curios, Listening 8604	V+A	V	A	T	Total
EYE CONTACT - INTEREST		15			15
MEANING OF WORDS - INTEREST	4		6	1	11
GIVING VOCAL FEEDBACK - INTERST	6		2	2	10
NODDING – INTEREST	2	4			6

Interested, Curios, Listening 8644	V+A	V	A	T	Total
NODDING – INTEREST	2	9			11
EYE CONTACT - INTEREST		10			10
LAUGHING, SMILING - INTEREST		1	7		8
GIVING VOCAL FEEDBACK - INTERST	2		5		7

Interested, Curios, Listening 8602	V+A	V	A	T	Total
GIVING VOCAL FEEDBACK - INTERST	8		8		16
ASKING QUESTION - INTEREST	7		6	2	15
EYE CONTACT – INTEREST	8	1		1	10
NODDING – INTEREST	2	6		1	9

Interested, Curios, Listening 8640	V+A	V	A	T	Total
CLEAR VOICE - INTEREST	1		10		11
EYE CONTACT - INTEREST	5	3		1	9
LAUGHING, SMILING - INTEREST	3	4	2		9
LOUD VOICE - INTEREST			9		9

Interest Complete Tables 1 - The tables show the four most common behaviors reported in each recording, as signs of *interest*

TOTAL Interested, curios, listening	V+A	Video	Audio	Transcription	Total
EYE CONTACT	13	29	0	2	**44**
VOCAL FEEDBACK	16	0	15	3	**34**
NODDING	7	23	0	2	**32**
ASKING QUESTIONS	11	0	10	5	**26**
LOUD VOICE	7	0	16	0	**23**
LAUGHING, SMILING	6	7	9	0	**22**
MEANING WORDS	4	0	10	5	**19**
CLEAR VOICE	5	0	10	0	**15**

Interest complete Table 2 - the table shows the most common behaviors reported at least ten times, as sign of *interest*.

Interest complete Table 2 shows that e*ye contact* is most common behavior selected by our respondents and it is also present in all the **Interest complete Tables 1**. Another behavior that got a high number of interpretations, is *giving a vocal feedback*, such as: yes, ja, no, haha, mmmm. This behavior was cited 16 times in the *V+A mode* and 15 in the *audio mode*. *Nodding* is another common movement and it is cited 32 times in total. There are also other common behaviors, such

as: *Asking questions*, *talking loudly* and *clearly*, *laughing*. Also in this case, the *semantic meaning of verbal words*, influenced the interpretation of the *AES* of *interest*.

Interest complete tables divided by modalities

TOTAL Interested, curios, listening	V+A
VOCAL FEEDBACK	16
EYE CONTACT	13
ASKING QUESTIONS	11
NODDING	7
LOUD VOICE	7
LAUGHING AND SMILING	6
CLEAR VOICE	5
MOVING HANDS	4
MEANING WORDS	4
MOVING FORWARD	3
MOVING HANDS TO EXPLAIN	3
VOICE	3
OPEN MOUTH	2
SPEAKING FAST	2
ANSWERING FAST	2
BACKWARD	1
GENERAL MOVEMENTS	1
HANDS IN THE POCKET, HOLDING	1
OPEN GESTURE	1
RISING EYEBROW	1
GIVING HAND STRAIGHT	1
LOOKING AWAY, DOWN	1
ANSWERING - INTEREST	1
TOTAL	**96**

TOTAL Interested, curios, listening	Transcription
ASKING QUESTIONS	5
MEANING WORDS	5
VOCAL FEEDBACK	3
EYE CONTACT	2
NODDING	2
HANDS IN THE POCKET, HOLDING	1
MOVING HEAD FORWARD	1
SHORT ANSWER	1
WHISPERING	1
TOTAL	**21**

TOTAL Interested, curios, listening	Video
EYE CONTACT	29
NODDING	23
LAUGHING AND SMILING	7
FORWARD	5
MOVING HANDS TO EXPLAIN	4
MOVING HANDS	3
BACKWARD	3
GENERAL MOVEMENTS	3
FACIAL EXPRESSIONS	3
MOVING HEAD FORWARD	3
STANDING	3
AGREE	2
OPEN MOUTH	1
HANDS IN THE POCKET, HOLDING	1
OPEN GESTURE	1
RISING EYEBROW	1
HOLDING HANDS. HANDS BACK	1
STOP TO GESTURING	1
TOTAL	**94**

TOTAL Interested, curios, listening	Audio
LOUD VOICE	16
VOCAL FEEDBACK	15
ASKING QUESTIONS	10
CLEAR VOICE	10
MEANING WORDS	10
LAUGHING AND SMILING	9
AGREE	2
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT	2
TALKING FAST	2
VOICE	1
SPEAKING FAST	1
ANALIZYING	1
LOW VOICE	1
TRY TO GUESS	1
TOTAL	**81**

4.3.1. How different modes influenced the interpretation of interest

Firstly, we can say that if we compare the sum of all interpretations in each mode, it seems that the *video mode* (94 interpretations) has more influence than *audio mode* (81 interpretations) and *transcription* (21 interpretations). Moreover, multimodality, in comparison with the *visual modality*, seems to be quite redundant since *V+A mode* (96 interpretations) got almost the same number of interpretation of *video mode*. However, if we look to our results with more attention, we can observe that in this case, when *sounds* and *video images* are displayed simultaneously, the *sound* has a stronger influence on respondent's interpretation. In fact, if we take into account e*ye contact* and *nodding* actions, we can observe that they are recognized much more often in *video mode* (*eye contact* 29 interpretations, *nodding* 23 interpretations) than in *V+A mode* (*eye contact* 13 interpretations, *nodding* 7 interpretations). Boholm's (2011) and McClave's (1999) studies about head movements show that very often *nodding* and *vocal feedback* are displayed (almost) simultaneously (Boholm, 2011) (McClave, 1990). According to their findings, we can deduce that in this case, the respondents paid more attention on *sounds* than *visual movements*. Moreover, if we look the data of *V+A mode*, we can notice that several vocal behaviors got a very high number of interpretations and some of them don't seem to be affected by the presence of *video mode*. In fact *vocal feedbacks* got almost the same number of interpretations in *V+A mode* (16 interpretations) and in *audio mode* (15 interpretations). The actions of *asking questions* got similar results (11 interpretations in *V+A mode* and 10 in *audio mode*). On the other hand, behaviors like *talking clearly* (5 interpretations in *V+A mode*, 10 interpretations *Video mode*) and *loudly* (7 interpretations in *V+A mode*, 16 interpretations in *Audio mode*) got many more interpretations in *audio mode* than in *V+A mode*. So also in this case, the combination of *video images* and *acoustical modalities* reduces the influence of some *bodily movements* and *vocal behaviors*, such as: *eye contact*, *nodding*, *meaning of words*, *talking loudly* and *clearly*. However, if we look at respondents' answers, we can notice that when *sounds* and *video images* are displayed together, people pay attention to both *vocal behaviors* and *body movements*. Concerning *written words*, it is hard drawing conclusions because of limited responses. However, despite the small number of answers, we can notice that in the *transcription*, respondents focused mainly on *vocal behaviors* and *spoken words* rather than *bodily gesture*. In the end, if we compare multimodality with *video modality*, we can see that even if we got a similar number of interpretations, actually people perceived both *sounds* and *video images*, therefore we can say that multimodality, compared with *video modality*, is redundant only in term of number of interpretations.

4.4. Confidence

By the term *confidence* we refer to all those states that in some way concern the feeling of being relaxed and sure about ourselves, our actions and our words. We can say that being calm and serene is strongly related with believing in ourselves. This state can be considered the opposite of *nervousness*. It embraces both *affective* and *epistemic* characteristics. *Affective* because confidence is a positive sensation and makes people feels good. *Epistemic* because *confidence* like *nervousness* is caused by the understanding and knowing of something. Often, people are confident about a specific situation, about what they say, what they know or what they think to know.

Confidence got a very high level of interpretation. This *AES* can be interpreted through different patterns of *verbal* and *non-verbal communication*. The tables below show the most common behaviors that influenced the interpretations of this *AES*.

Confident, Calm, Relaxed 8644	V+A	V	A	T	Total
LAUGHING, SMILING - CONFIDENT	2	4	3		9
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - CONFIDENT	3	1	2		6
MEANING OF WORDS - CONFIDENT	2		2		4
MOVING HANDS TO EXPLAIN- CONFIDENT	3	1			4
VOICE - CONFIDENT	2		2		4
STANDING - CONFIDENT		3			3
TALKING A LOT, MORE - CONFIDENT			3		3
EYE CONTACT - CONFIDENT	2	1			3

Confident, Calm, Relaxed 8640	V+A	V	A	T	Total
STANDING - CONFIDENT	5	5			10
MEANING OF WORDS - CONFIDENT	2		4	1	7
EYE CONTACT - CONFIDENT	2	3			5
TAKING INITIATIVE - CONFIDENT			3		3
GIVING VOCAL FEEDBACK CONFIDENT			2	1	3
LAUGHING, SMILING - CONFIDENT			3		3

Confident, Calm, Relaxed 8604	V+A	V	A	T	Total
LEANING ON THE WALL - CONFIDENT	2	11			13
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - CONFIDENT			11		11
MEANING OF WORDS - CONFIDENT	2		3	3	8
LAUGHING, SMILING – CONFIDENT	2		4		6

Confident, Calm, Relaxed 8602	V+A	V	A	T	Total
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - CONFIDENT	6		7	1	14
LOUD VOICE - confident			10		10
CLEAR VOICE - confident			9		9
MOVING HANDS TO EXPLAIN - confident	5	3			8

Confidence Complete Tables 1

Confident, Calm, Relaxed	V+A	Video	Audio	Transcription	Total
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT	9	1	20	1	**31**
LAUGHING, SMILING	6	8	11	0	**25**
MEANING OF WORDS	8	0	9	7	**24**
STANDING	8	8	0	0	**16**
LEANING ON THE WALL	2	11	0	0	**13**
EYE CONTACT	4	8	0	1	**13**
LOUD VOICE	0	0	13	0	**13**
CLEAR VOICE	1	0	11	0	**12**
MOVING HANDS TO EXPLAIN	8	4	0	0	**12**
LEANING FORWARD	0	11	0	0	**11**

Confidence complete Table 2 - the table shows the most common behaviors reported at least ten times, as sign of *confidence*.

If we look at the **Confidence complete Table 2** the most common reason for respondents to interpret *confidence* is "*people in the recording know what they are talking about*". This is a quite generic answer, because it includes *words*, *verbal* and *non-verbal communication*. However, if we observe our data, we can easily understand that people focused just on spoken *words* and not on body movements (we got just one selection in the *video mode*). In fact, according to our data, the answer "*meaning of words*", got a quite high result. Also how people speak (*clear* and *high level* of *voice*) can give information about the state of *confidence* and also in this case, l*aughing and smiling*, *body movements* and *posture* has been selected several times as signs of *confidence*.

Confidence complete tables divided by modalities

Confident, Calm, Relaxed	V+A
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT	9
MEANING OF WORDS	8
MOVING HANDS TO EXPLAIN	8
STANDING	8
LAUGHING, SMILING	6
HANDS IN THE POCKET, HOLDING	5
EYE CONTACT	4
MOVING HANDS, ARMS	4
VOICE	3
LEANING ON THE WALL	2
LOW VOICE	2
OPEN GESTURE	2
ANSWERING	1
GIVING HAND STRAIGHT	1
HEAD MOVEMENTS	1
MOVING HEAD	1
RISING EYEBROW	1
SHOULDER DOWN	1
SLOW MOVEMENTS	1
ASKING QUESTION	1
CLEAR VOICE	1
GENERAL MOVEMENTS	1
LOOKING DOW, AROUND, TO CAMERA	1
LEANING BACKWARD	1
NODDING	1
STANCE	1
TOTAL	75

Confident, Calm, Relaxed	Video
LEANING ON THE WALL	11
LEANING FORWARD	11
STANDING	8
LAUGHING, SMILING	8
EYE CONTACT	8
GENERAL MOVEMENTS	6
MOVING HANDS, ARMS	5
MOVING HANDS TO EXPLAIN	4
NODDING	4
LEANING BACKWARD	3
STANCE	3
LOOKING DOW, AROUND, TO CAMERA	2
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT	1
HANDS IN THE POCKET, HOLDING	1
OPEN GESTURE	1
AGREE	1
ARMS DOWN	1
BACK STRAIGHT	1
MIRRORING	1
SCRATCHING TOUCHING BACK, FACE, ETC	1
STRONG HANDSHAKE	1
TOTAL	82

Confident, Calm, Relaxed	Audio
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT	20
LOUD VOICE	13
LAUGHING, SMILING	11
CLEAR VOICE	11
MEANING OF WORDS	9
TAKING INITIATIVE	6
VOICE	5
ASKING QUESTION	4
FLAT TONE OF VOICE	4
TALKING A LOT, MORE	4
SPEAKING STRAIGHT	3
AGREE	2
TALKING SLOW	2
LOW VOICE	2
GIVING VOCAL FEEDBACK	2
OVERLAPPING	2
ADDING MORE INFO	1
EXPRESSING OPINION	1
STOP TO LAUGHING	1
TOTAL	**103**

Confident, Calm, Relaxed	Transcription
MEANING OF WORDS	7
TAKING INITIATIVE	2
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT	1
SPEAKING STRAIGHT	1
GIVING VOCAL FEEDBACK	1
EYE CONTACT	1
NODDING	1
SHORT ANSWER	1
STRAIGHT ANSWER	1
TOTAL	**16**

4.4.1. How different modes influenced the interpretation of confidence

Like the other *AES*, we observed that different modes influenced in different ways the interpretation of *confidence*. Behaviors, actions and sounds have been perceived differently, depending on the modes by which they were displayed. The *Audio mode* got the highest number of interpretations (103 interpretations), followed by *video mode* (82 interpretations), *V+A mode* (75 interpretations) and *transcription* (16 interpretations). In this case, it seems that *auditory modality* is the most powerful modality to trigger an interpretation of confidence. Concerning multimodality, we can say that since it displays *video images* and *sounds* together, respondents have more options in choosing behaviors, actions and sounds that can influence the interpretation of *confidence*. However, if we compare the numbers of interpretations, multimodality seems to work as a filter for many behaviors, that have more influence when they are displayed in unimodal modes. These results are very interesting and if we observe and compare them, we can make several reflections. If we focus on the results of *V+A mode*, we can notice that the most common answer is "*he/she knows what he/she is talking about*". It is also the most common answer in the *audio mode*, with almost twice as many occurrences (9 times in *V+A mode* and 20 times in *audio mode*). Similar results are visible for other behaviors and actions, such as: "*laughing and smiling*" (6 interpretations in *V+A mode* and 11 in *audio mode*), "*clear voice*" (1 selection in *V+A mode* and 11 in *audio mode*) Therefore, also in this case the *video images* affect reducing the *auditory* perception of *confidence*. The interpretation of this state decreases significantly, when it is perceived by *video images* and *sounds* together. Also the *sound* negatively influences the interpretation of *confidence* when this *AES* is displayed simultaneously with *video images*. In fact, the interpretations of *non-verbal communication* behaviors decreases in number when they are displayed with both *video images* and

sounds, than with just *sounds*. For example, some respondents said that people in the video looked *confident* because they were "*leaning on the wall*" or because of "*eye contact*". The first behavior got 11 interpretations in the *video mode* and just 2 in the *V+A mode*, while the second one got 8 interpretations in *video mode* and just 4 in *V+A mode*. We also notice that when *video mode* is displayed, respondents pay a lot attention to posture and head's position. The *meaning of the words* is the answer that got more interpretations in the *transcription* (7 interpretations). Therefore, in the end we can say that the semantic meaning of the *words* has a big influence in the interpretation of *confidence* in *V+A*, *audio* and *transcription modes*. Spoken words influence much more the respondents' interpretation when they are heard in *audio mode*, than in other modes. The behaviors *moving hands to explain* and *holding hands, or hands in the pocket*, got more interpretations in *V+A mode* than in *video mode*. This means that multimodality amplifies the influence of just these two behaviors.

4.5. Disinterest

Like all the other states, also the *AES* of *disinterest* has one affective and one epistemic component. Every time that we are in the state of *disinterest*, in some way we feel emotions that often are negative, such as: boredom or bother. When we have *disinterest* toward something, we have to make inferences about this something anyway. The epistemic part consists of the reflections that we have to make, to understand and realize if we are interested or not. In the end, the result of our reflections will produce a certain kind of emotion.

Due to the small number of answers received, in this case we will present a more general view about the *AES* of *disinterest*. Even if this particular *AES* didn't get numerous citations, we can observe several interesting outcomes.

Disinterested, Bored	V+A	Video	Audio	Transcription	Total
VOICE GOING UP AND DOWN			8		8
GIVING VOCAL FEEDBACK	3		1	3	7
LOW VOICE	3		4		7
MEANING OF WORDS	1			3	4
SPEAKING SLOW	1		3		4
SHORT CONTRIBUTION	1		3		4
LOOKING DOW, AROUND, TO CAMERA		2			2
MOVING HANDS, ARMS	1	1			2
FLAT TONE OF VOICE	1		1		2
LEANING BACKWARD		2			2
MOOVING LEGS, FEET		2			2
NO RESPONDING, NO FEEDBACK	1		1		2
TOO MANY REPETITION				1	1
BACK STRAIGHT	1				1
BREATHING DEEPLY			1		1
FAKE LAUGHING	1				1
GENERAL MOVEMENTS	1				1
LOUD VOICE			1		1
MOVING A LOT		1			1
NO TALKING SO MUCH				1	1
NOT CLEAR VOICE	1				1
RASPY VOICE	1				1
STANDING	1				1
TWISTING BODY				1	1
VOICE			1		1
TOTAL	18	8	24	9	59

Disinterest Complete table 1 – all the behaviors reported as signs of *disinterest*

Firstly, we can easily note that the most common behaviors concern mostly *vocal communication*, such as: *vocal feedback* and *prosody* (tone of voice, speech's length and velocity). If we focus on how different modes influenced the interpretation of *disinterest*, we see that *audio* is the mode that got the highest number of interpretations. Differently, *video* is the *mode* that got the smallest

number of interpretations. These outcomes are also confirmed by *V+A mode*, where the highest number of interpretations refer to *sounds*, such as: *tone of voice* and *vocal feedback*. We can also say that the state of *disinterest* is harder to interpret when it is displayed with both *sounds* and *video images* together, than instead with just *sounds*. Concerning *transcription* we can observe that respondents paid more attention to *meaning of words* and *vocal feedback*.

4.5.1. How different modes influenced the interpretation of disinterest

If we take a look at the tables about *disinterest* we can deduce that *sound* features, more than gesture feature trigger an interpretation of *AES* of *disinterest* (24 interpretations). Respondents seem to be very influenced by vocal characteristics and actions. Since *V+A modality* got 18 interpretations, we can say that multimodality, in comparison to unimodal modes, transmits a different kind of information. Also in this case it seems that multimodality works as filter, and it decreases the influence of the *auditory modality*. Even if multimodality gives the opportunity to get both vocal and non-vocal behaviors, respondents paid more attention to vocal behaviors. In fact, *video mode* got just 8 interpretations, therefore it looks like people are hardly influenced by *video images*. Similar results are presented in *transcription* (9 interpretations), where it seems that respondents focused a lot on vocal behaviors.

4.6. Thoughtfulness and understanding

Above we have focused on the most common *AES* interpretations observed during our experiment. However several other *AES* have been reported by our participants. These states have not been studied deeply in this thesis due to the low number of interpretations. Despite of the small number of interpretations it is, still possible to observe some interesting outcomes. Here we will briefly present two *AES* that can give to us some interesting information. These states are *thoughtfulness* and *understanding.*

Concerning *thoughtfulness* we can say that this state is strongly dependent on the action of thinking. The epistemic part of this state consists of the fact that we have to know something in order to think about it. Moreover when we think about a precise thing, our reasoning will produce some affects in our body. The kinds of emotions that this action produces depend on what we are thinking and on the level of understanding that we have about that precise subject.

Thoughtfulness	V+A	Video	Audio	Transcription	Total
Looking up	2	3			**5**
Looking away		1			**1**
General movements	1				**1**
Low voice	1				**1**
Meaning of words	1				**1**
Laughing			1		**1**
TOTAL	**5**	**4**	**1**	**0**	**10**

Thoughtfulness complete table 1 - all the behaviors reported as signs of *thoughtfulness*

Since the *AES* of *understanding* mostly concerns the nature of knowledge, it could be just thought of as an epistemic state; however it has also an affective part. In fact, this state can also create some affective states, such as: happiness, sadness or anger depending on the meaning of the message and how we elaborate it.

Understanding	V+A	Video	Audio	Transcription	Total
Smiling		3			**3**
Looking around, in the camera		2			**2**
Nodding		1			**1**
Answering			1		**1**
Standing	1				**1**
Meaning of words	1				**1**
TOTAL	**2**	**6**	**1**	**0**	**9**

Understanding Complete Table 1 - all the behaviors reported as signs of *understanding*

This two *AES* have been interpreted in different ways. For example in **Thoughtfulness complete table 1** we can observe that the *AES* of *thoughtfulness* was mentioned 5 times in *V+A mode*, 4 times in *video mode*, 1 time in *audio mode* and it did not get any interpretations in *transcription*. Therefore it seems that *V+A mode* can transmit more information in comparison with the other modes. The **Understanding Complete Table 1** shows that the *AES of understanding* can be easily interpreted, when it is displayed just by *video images*. Respondents had more difficulties when it was displayed in *V+A*, *transcription* and *audio modes*.

4.6.1. How different modes influenced the interpretation of thoughtfulness and understanding

In our results, we can observe that there is a precise movement that permitted the interpretation of *thoughtfulness*, and it is *looking up* (or *looking away*). Therefore, we can say that in this case the influence of *video images* is stronger than the *sounds* and *written words*. However, since in *V+A modality* (5 interpretations), the respondents selected both bodily movements and vocal behaviors, it seems that multimodality can provide some more information than at least *audio mode* (1 interpretation). The *transcription modality* did not get any interpretation.

In the case of *understanding*, the *video mode* (6 interpretations) has an higher number of interpretations (2 interpretations in *V+A mode* and 1 selection in *audio mode*). In fact, we can also observe that the head is the most important part of the body in order to interpret the *AES* of *understanding*. S*miles*, *eye directions* and *head movements* are all actions that belong to the head. However, these behaviors are not interpreted when they are displayed in a multimodal way. Therefore, in this case, we can claim that multimodality works as a filter. In the end, the *transcription modality* did not get any interpretation.

5. Discussion

Considering the *visual sensory modality*, we can say that it works differently when we are looking at *video images* or *written words*. Respondents are more likely to interpret *AES* through *video images* than *words*. Therefore, a behavior displayed by *written words* will be more hard to perceive than seeing the same behavior displayed through *video images*. When a person read a *word*, he/she has to think and make inferences in order to create a visual image in the mind. For example, if we read the word laughing, we can imagine a person that is laughing, but this behavior can be executed in an infinite number of different ways. This mechanism can especially create problems when we talk about feelings, emotions or every kind of abstract concept. Therefore, a single word cannot precisely describe a behavior, an action or a state. For many *multimodal* behaviors, the *sound* acts as filter and decrease the influence of *visual modality*. On the other hand, the *sound* acts as an amplifier and increases the interpretation of *nervousness* when behaviors are displayed through *video images*.

If we focus on *vocal words* we can easily observe that word have more influence when they are perceived as *vocal sounds* than as *written words*. Only in the case of the interpretation of *nervousness*, respondents were more influenced by the *semantic meaning of words* when they were displayed through *written form* than through *acoustic sounds*. When *video images* are displayed together with the *sound*, we can observe that very often, they reduce the focus on *semantic meaning of words* pronounced by video participants. This characteristic was observed in all the *AES* except in the *AES* of *happiness*.

Concerning multimodality we can say that it often works as a filter for *auditory modality* and *visual modality*. This means that when *video images* and *sounds* are perceived together, one modality affect and decrease the perception of the other modality. However, if we focus on behaviors and movements, they can be easier perceived and interpreted when they are displayed through a multimodal mode than through *written words*. But if we compare *V+A mode* with *video mode* and *audio mode*, we can observe that multimodality works as a filter when we interpret: *confidence*, *interest*, *happiness*, *disinterest*, *understanding* and just few times in *nervousness* (*hands movements/position* and *meaning of words*). This means that when a specific behavior is displayed, respondents are more likely to interpret it as a signal of a precise *AES* in an unimodal mode than in a multimodal mode. In our data about the *AES of happiness* we could observe that *laughing and smiling* got a higher number of interpretations in *video mode* than in *V+A mode*. On the other hand, if we pay attention to the *nervousness*'s data, we can see that respondents are more likely to interpret this *AES* through *sounds* than *video images*. For example, if we observe a person that is *laughing* because of *nervousness* and we utilize just the *visual modality*, we are more likely to interpret him/her as a *happy* person even if he/she is *nervous*. But if we add the *auditory modality*, we will also perceive the *sound* of a *nervous laughing*. In this case, the combination of *sounds* and *video images* allows us to have another kind of perception of *laughing*, and therefore we will have a better chance to interpret this behavior as a sign of *nervousness* than a sign of *happiness*. In this way the number of interpretations of *V+A mode* will increase in the case of *nervousness* and decrease in the case of *happiness*. Multimodality works as an amplifier when *nervousness* and *thoughtfulness* are displayed, but also for some behaviors of *confidence*, such as: *hands movements* and *hands position*. In conclusion, if we compare multimodality and unimodality, we can notice that in some *AES*, multimodality got a smaller number of interpretations. However, this doesn't mean that multimodality provides less information, but it means that since we perceive behaviors, sounds and actions through a combination of two different modalities, *auditory* and *visual* (*video images*), multimodality, in comparison with unimodality, provides a different kind of information. We will get another kind of information, when behaviors, sounds and actions have been displayed through *written words*. If we observed the *AES* of *interest*, we noticed that multimodality got similar number

of interpretations in total in *A+V mode* and *Video mode* (96 interpretations in *V+A mode*, 94 in *Video mode*). Therefore, if we just consider numbers, we can claim that in this case multimodality is quite redundant with *visual modality*. However, we have to be aware that multimodality, in comparison with *visual modality* can transmit different kinds of information.

5.1. Possible future studies and applications

The understanding of how *AES* are perceived by humans is a fairly unexplored area. In fact, just a small number of studies have investigated states, such as: *understanding*, *interest*, *surprise*, *confusion*, etc. Even if only a few *AES* were presented in this paper, we have to be aware that many other *AES* can be studied in depth. Previously, we saw that almost all *AES* are highly multimodal and they can displayed by *sounds* and *gesture*, and perceived by different modalities. In the future, it would be interesting to focus just on one behavior, and investigate how people's interpretation of *AES* is influenced when the same behavior is displayed in a different modality, for example: smiling and laughing. In our results it was possible to see that *written words* also influenced the interpretation of *AES*, therefore it would be useful to research in depth the difference about how *words* perceived by *sound* and perceived by *written words*, can influence in the interpretation of *AES*. It would be also interesting to know if really exist a difference among similar words (such as: nervous, uneasy, embarrassed, shy) and to understand if they can be all classified as different categories of the same *AES* or if they are actually different *AES*. Since we just took into consideration Swedish people, another interesting subject could be the observation of how cultural differences influence how people display and interpret *AES*. Since the personality of each participant could also affect the interpretation of the different *AES*, this could be an interesting research area to investigate. Research about *AES* can be very useful in the development of software and technical aids for blind and deaf people. For example, it would be interesting to see if blind and deaf people interpret *AES* in a different way from others. In the end, we can say that even if some studies investigated the area of emotions and affective states, it is still very hard to find research that involves epistemic features. Consequently it is essential to pursue this research area, in order to better understand how people show and interpret different *AES.*

5.2. Limitation of the research

Even if this study gave very interesting outcomes there are still some limitations. We did not focus on personality or personal background, but we have to be aware that people have very different characteristics, which could influence the interpretation of *AES*. Moreover, it is very hard to understand if all the respondents understood the real meaning of the term *AES*. Since we got very different kinds of answers from the participants, we could probably assume that majority understood the meaning of the term. Another limitation of this research could be the understanding of the transcriptions. In fact, a lot of respondents had difficulty with the large amount of numbers and notes in the transcription. This made it hard for them to understand what it was going on in the recording. Unfortunately, each section of the experiment took quite long time, an average of 75 minutes. Some participants referred that in the last part of the experiment it was extremely hard to remain concentrated on the recording. Therefore the last mode presented could be influenced by this factor. Another limitation of this research is the difficulty that the participants had in describing *AES* in a language that was not their mother tongue. We have to remember that since we are talking about interpretation of *AES* of other people, the respondents can also misinterpret what a person really feels. Consequently, we don't know if people in the recording were really in the state interpreted by our respondents. However, we are much more interested in why the respondents think that the person in the recording is in that specific *AES*. The position of the camera may also be a limitation. People in the video were recorded laterally, and their faces were completely visible just

when they were turning the head to the camera. Therefore not showing their faces in frontal view could have influenced the interpretation of some *AES.*

6. Conclusion

The aim of this study was to explore people's interpretation of *AES*. To investigate this area we formulated the following research questions:

- *How do different modes contribute to the interpretation of affective epistemic states?*

- *What differences of information, about affective epistemic states, can you get in an interpretation of multimodal communication mode in comparison with an interpretation of unimodal communication mode?*

To answer the first question, we formulated a hypothesis. Since just few studies have been done in this area, it was very hard to formulate a answer. However, based on our research background we claimed that the different modes contribute in different ways to the interpretation of *AES*. This means that when the same behavior, sound or action is displayed in different modes, the respondent's perception and interpretation are often influenced in different way. This means that our participants can perceive the same behavior, sound or action in different ways, and give it different meanings depending on the modes with which the behavior, sound or action was displayed. In conclusion, since different modes affect ways the perception of the same behavior in different, we can say that the hypothesis formulated is confirmed.

The hypothesis to answer the second research question claimed that if multimodality leads to more information about the interpretation of *AES* than unimodality, then viewers will be able to interpret and give more information about *AES* when they are shown through Video+Audio (V+A) recording than when they are shown through unimodal recording, Video, Audio, or Transcription (T). If multimodality does not lead to more information about the interpretation of *AES* than unimodality, we will obtain the same results when we show multimodal stimuli (A+V) as when we use unimodal stimuli (A or V or T). This would mean that multimodality is redundant for the interpretation of *AES*. Observing our data, we can realize that it is possible to interpret the same *AES* in different modalities. In fact, very often the same *AES* has been found in different modalities, by different respondents. Therefore, we can easily say that *AES* are highly multimodal. The same *AES* can be shown through vocal and gestural behaviors, and it can be perceived by visual, auditory or both the modalities together, depending on the modes displayed. If we observe the **complete table on pag. 22** we can see that actually the multimodal mode, in comparison with the other unimodal modes, got the highest number of interpretations. Therefore we can claim that the hypothesis is confirmed. But, if we consider each *AES* separately, we can see that very often multimodality works as a filter for one *modality* or both *modalities*: *visual* and *auditory*. This means that when we are watching *visual images* and hearing *sounds* simultaneously, often one modality decrease the perception of the other modality. Due to the difference of quantity of information provided by multimodality and unimodality, the perception of a particular behavior or sound can change and lead to different interpretations of *AES*, depending on if we are using one or more sensory modalities. Just few times the combination of *sounds* and *video images* increases the perception of both *visual* and *auditory modalities*. In the majority of the times, different behaviors are perceived differently, depending on if they were shown in multimodal or unimodal modes. Redundancy has been observed just in the case of the *AES* of *interest*, when the multimodal mode was compared with *video mode*.

References:

Allwood, J. (2002). Bodily Communication - Dimensions of Expression and Content. *Multimodality in Language and Speech Systems*. Björn Granström, David House and Inger Karlsson (Eds.). Dordrecht: Kluwer Academic Publishers, pp. 7-26.

Allwood, J., and Cerrato, L. (2003). A study of gestural feedback expressions. In P. Paggio et al. (Eds.). *Proceedings of the First Nordic Symposium on Multimodal Communication* (pp. 7–22).

Banse, R., and Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology*. 70, 614-636.

Boholm M. and Lindblad G. (2011). Head movements and prosody in multimodal feedback. *Nealt proceedings series.* 15.

Bruce, V. (1982). Changing faces: Visual and nonvisual coding processes in face recognition. *British, Journal of Psychology*. 73. (1). p.105-116

Chen W., Lander K., and Hong Liu C. (2011). Matching Faces with Emotional Expressions. *Front Psychol*. 2. (206).

Chindamo, M, Allwood, J & Ahlsén, E. (2012). Some suggestions for the study of stance in communication. *Proceedings of IEEE SocialCom Amsterdam 2012*, 3-5.

Chomsky N. (1957). *Syntactic structures*. Moutin, The Hague, p.11

Cohn J.F. and Ekman P. (2004). Measuring facial action by manual coding, facial EMG, and automatic facial image analysis. *Handbook of Nonverbal Behavior Research Methods in the Affective Sciences*. Harrigan J.A., Rosenthal R., Scherer K.R. (Eds.). 2003. pp. 9–64

de Gelder B, Snyder J, Greve D, Gerard G, Hadjikhani N. (2004). *Fear fosters flight: a mechanism for fear contagion when perceiving emotion expressed by a whole body*. 101. (47). p. 16701-16706

de Meijer M (1989). The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal Behavior*. 13. (4). p. 247 - 268

Dael N., Mortillaro M., Scherer K. R. (2012). Emotion Expression in Body Action and Posture. *Emotion*. 12. (5). p. 1085-1101

Darwin, C. (1872). *The expression of the emotions in man and animals*. London: Murray. University of Chicago Press, 1965.

Dimberg U., Thunberg M. and Elmehed K., (2000). Unconscious Facial Reactions to Emotional Facial Expressions. *Psychological Science*. 11. (1). P. 86-89

Duglas, K. (2003). What are you laughing at?. *New Scientist*. 180. p. 72-73

Ekman P. and Friesen W. (1972). Hand Movements. *The Journal of communication*. 22. p. 353-374

Ekman Paul and Davidson Richard J (1993). Voluntary smiling changes regional brain activity. *Psychological Science*. 4. (5). p.343-345

Fox E., Lester V., Russo R., Bowles R.J., Pichler A. and Dutton K. (2000). Facial Expressions of Emotion: Are Angry Faces Detected More Efficiently?. *Cognition & Emotion*. 14. (1). p. 61-92

Frank M. G. and Ekman P. (1993). Not all smiles are created equal: The differences between enjoyment and other smiles. *Humor: The International Journal for Research in Humor*. p 6 (1). p. 9-26.

Ghazanfar, A. A. and N. K. Logothetis. (2003). Facial expressions linked to monkey calls. *Nature*. 423. p. 937-938.

Gliga, T. and Dehaene-Lambertz, G. (2005). Structural encoding of body and face in human infants and adults. *Journal of Cognitive Neuroscience*. 17. p.1328-1340.

Grayson D. and Dr. Coventry L. (1998). The Effects of Visual Proxemic Information in Video Mediated Communication. *SIGCHI Bulletin* 30. (3). p. 30-39

Gunnery Sarah D., Hall Judith A. and Ruben Mollie A. (2012). The Deliberate Duchenne Smile: Individual Differences in Expressive Control. *Nonverbal Behavior*. 37. (1). p .29-41

Haakana Markku (2010). Laughter and smiling: Notes on co-occurrences. *Journal of Pragmatics*. 42. (6). p. 1499–1512.

Hall, E. T. (1959). *The silent language*. New York: Fawcett.

Hall, E. T. (1963). A system for the notation of proxemics behavior. *American Anthropologist*. 65. (5). p. 1003-1026.

Hansen, C.H. and Hansen, R.D. (1988). Finding the face in the crowd: An anger superiority effect. *Journal of Personality and Social Psychology*. 54. (6) 917- 924.

Hietanen J. K., Leppänen J. M. and Lehtonen U. (2004). Perception of emotions in the hand movement quality of Finnish sign language. *Journal of Nonverbal Behavior*. 28. (1). p. 53-64

Horstmann G. and Ansorge U. (2011). Compatibility between tones, head movements, and facial expressions. *Emotion.* 11. (4). p. 975–980.

Jackson M.C., Wu C., Linden D. E. J. and Raymond J. E. (2009). Enhanced Visual Short-Term Memory for Angry Faces. *Journal of Experimental Psychology: Human Perception and Performance*. 35 (2). p. 363–374

Juslin P. N. and Laukka P. (2003). Communication of Emotions in Vocal Expression and Music Performance: Different Channels, Same Code?. *Psychological Bulletin*. 129. (5). p. 770–814.

Kana R. K. and Travers B. G. (2011). Neural substrates of interpreting actions and emotions from body postures. *Soc Cogn Affect Neurosci.* 7. (4). p. 446-456.

Knapp M. and Hall J. (2010). *Nonverbal Communication in Human Interaction*. 7th Edition, Wadsworth Cengage Learning.

Kuhl, P. K., and A. N. Meltzoff. (1982). The bimodal perception of speech in infancy. *Science.* 218. (4577). p. 1138-1140.

Leavitt, J. (1991). Review of "Evidentiality: the linguistic coding of epistemology" by Wallace Chafe; Johanna Nichols, *Language*. 67. (1). p. 133-141.

McClave E.Z. (1990). Linguistic functions of head movements in the context of speech. *Journal of Pragmatics.* 32. (7). p. 855-878.

McGurk, H.,and J. MacDonald. (1976). Hearing lips and seeing voices. *Nature*. 264. p. 746–748.

Meeren H. K. M., Van Heijnsbergen C. R. J., De Gelder B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proc Natl Acad Sci U S A*.. 102. (45). p. 16518–16523.

Mehrabian A. (1969). Significance of posture and position in the communication of attitude and status relationship. *Psychological Bulletin*.. 71. (5). p. 359-372

Mehrabian A. (2006). Orientation behaviors and nonverbal attitude communication. *Journal of Communication*. 17 (4). p. 324-332.

Messinger, D.S., Fogel, A., Dickson, K. (2001). All smiles are positive, but some smiles are more positive than others. *Developmental Psychology*. 37. (5). p. 642-653.

Moeschler J. (2009). Pragmatics, propositional and non-propositional effects: can a theory of utterance interpretation account for emotions in verbal communication?. *Social Science Information*. 48. (44). p. 447-464

Nomi J. S., Rhodes M. G. & Cleary A. M. (2012). Emotional facial expressions differentially influence predictions and performance for face recognition. *Cognition & Emotion*. 27. (1). p. 141-149

Parr, L. A. (2004). Perceptual biases for multimodal cues in chimpanzee (Pan troglodytes) affect recognition. *Animal Cognition*. 7. (3). p. 171-178.

Partan, S. R. and Marler. P. (2005). Issues in the classification of multimodal communication signals. *American Naturalist*. 166. (2). p. 231-245.

Preuschoft, S. (1995). *Laughter' and 'smiling' in Macaques--An Evolutionary Perspective*. Utrecht: University of Utrecht.

Ruch W and Ekman P (2001). The expressive pattern of laughter. In A.W. Kaszniak (Eds.) *Emotion qualia, and consciousness*. Word Scientific Publisher. Tokyo. p. 426-443

Russell J.A., Bachorowski J., and Fernandez-Dols J. (2003). Facial and Vocal Expressions of Emotion. *Ann. Rev. Psychology*. 54. p. 329-349.

Schneider K. and Josephs I. (1991). The expressive and communicative functions of preschool children's smiles in an achievement-situation. *Journal of Nonverbal Behavior*. 15. (3). p. 185-198

Schneider K, Unzner L. (1992). Preschoolers' attention and emotion in an achievement and an effect game: a longitudinal study. *Cognition & Emotion*. 6. (1). p. 37-63

Scherer K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*. 40. (1-2). p. 227-256.

Schroder M. et al. (2011). Building Autonomous Sensitive Artificial Listeners. *IEEE Trans. Affective Computing*. 9. (1). p. 1

Soussignan R, Schaal B. (1996). Forms and social signal value of smiles associated with pleasant and unpleasant sensory experience. *Ethology*. 102. (8). p. 1020-1041

Wallbott H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology*. 28. (6). p. 879-896.

Reference Note

Efron, D. (1941). *Gesture and environment*. New York: King's Crown.

UNIVERSITY OF GOTHENBURG

APPENDIX

How do different modes contribute to the interpretation of affective epistemic states?

How different mode of representation (*video*, *audio*, *video+audio* and *written words*) can influence the understanding and interpretation of *AES*

Stefano Lanzini

Master of Communication Thesis
Report No. 2013:071
ISSN: 1651-4769

University of Gothenburg
Department of Applied Information Technology
Gothenburg, Sweden, May 2013

Contents

Nervousness

Nervous, insecure, tense, embarrassed	V+A	Video	Audio	Transcription	Total
LAUGHING AND SMILING	11	5	27	12	**55**
LOOKING DOW, AROUND, AT CAMERA	15	16	0	11	**42**
MEANING OF WORDS	4	0	13	23	**40**
STUMBLING, HESITATIONS, REPETITION	12	0	12	11	**36**
MOVING HANDS AND ARMS	11	17	0	3	**31**
MOVING A LOT	11	8	0	2	**21**
STANDING	14	6	0	0	**20**
HANDS IN THE POCKET, HOLDING	10	5	0	4	**19**
MOVING FEET	11	4	0	1	**16**
GENERAL MOVEMENTS	5	4	0	0	**9**
SCRATCHING TOUCHING BACK, FACE, ETC	5	4	0	0	**9**
TWISTING BODY, SHAKING	5	2	0	1	**8**
TALKING TOO MUCH	4	0	0	2	**6**
SHORT ANSWERS AND FEEDBACK	1	0	2	2	**5**
GIVING VOCAL FEEDBACK	1	0	1	3	**5**
LOUD VOICE	2	0	2	0	**4**
VOICE	3	0	1	0	**4**
LEANING BACKWARD	2	1	0	0	**3**
STOP TO EXPLAIN, TO SPEAK - NERVOUS		3			**3**
LOW VOICE			2		**2**
TALKING FAST - NERVOUS			2		**2**
EYE CONTACT	1	1	0	0	**2**
NODDING	1	1	0	0	**2**
MOVING LEGS, FEET		2			**2**
ARMS DOWN - NERVOUS		2			**2**
LAUGHING POLITE	2				**2**
TALKING SLOW	2				**2**
BREAK AFTER GREETING	2				**2**
NEUTRAL VOICE			1		**1**
SHORT FEEDBACK			1		**1**
SPEAKING FAST			1		**1**
TAKLING LITTLE			1		**1**
TONE OF VOICE UP AND DOWN			1		**1**
BREATHING DEEPLY			1		**1**
HE DOESN' T WANT INTERUPT				1	**1**
LIFTING HEELS		1			**1**
LOW EYEBROW		1			**1**
SHAKING		1			**1**
TWISTING ARMS		1			**1**
LEANING FORWARD	1				**1**
THE OTHE PERSON DOESN'T GIVE FEEDBACK	1				**1**
MIRRORING	1				**1**

NOISY VOICE	1				**1**
STOP TO TALK	1				**1**
BECAUSE THE OTHER PERSON IS NERVOUS	1				**1**
DOWN EYEBROW	1				**1**
TOTAL	**142**	**85**	**68**	**76**	**372**

Nervous, insecure, tense, embarrassed 8644	**V+A**	**Video**	**Audio**	**Transcription**	**Total**
LAUGHING, SMILING - Nervous	7	2	10	2	**21**
LOOKING AWAY, DOWN - Nervous	10	5		2	**17**
MEANING OF WORDS - Nervous			4	8	**12**
MOVING HANDS AND ARMS- Nervous	5	3		2	**10**
SCRATCHING TOUCHING BACK, FACE, ETC - Nervous	5	3			**8**
STUMBLING, HESITATIONS, REPETITION - Nervous	4			2	**7**
HANDS IN THE POCKET, HOLDING - Nervous	3	2		1	**6**
MOVING FEET - Nervous	4	2			**6**
GENERAL MOVEMENTS - Nervous	3	2			**5**
TWISTING BODY, SHAKING - Nervous	2	2			**4**
TALKING TOO MUCH - Nervous	2			2	**4**
STOP TO EXPLAIN, TO SPEAK - NERVOUS		3			**3**
MOVING A LOT - Nervous	2				**2**
ARMS DOWN - NERVOUS		2			**2**
VOICE - Nervous	1		1		**2**
GIVING VOCAL FEEDBACK - Nervous	1			1	**2**
THE OTHE PERSON DOESN'T GIVE FEEDBACK - Nervous	1				**1**
LIFTING HEELS - NERVOUS		1			**1**
BREATHING DEEPLY - NERVOUS			1		**1**
TAKLING LITTLE - NERVOUS			1		**1**
TONE OF VOICE UP AND DOWN - NERVOUS			1		**1**
LEANING BACKWARD - NERVOUS	1				**1**
BECAUSE THE OTHER PERSON IS NERVOUS - Nervous	1				**1**
EYE CONTACT - NERVOUS		1			**1**
MIRRORING - Nervous	1				**1**
DOWN EYEBROW - Nervous	1				**1**
TOTAL	**54**	**28**	**18**	**20**	**121**

Nervous, insecure, tense, embarrassed 8640	**V+A**	**Video**	**Audio**	**Transcription**	**Total**
STUMBLING, HESITATIONS, REPETITION - Nervous	4		6	5	**15**
MEANING OF WORDS - Nervous	1		2	9	**12**
MOVING HANDS AND ARMS- Nervous	5	6		1	**12**
LAUGHING, SMILING - Nervous	1	3	5	2	**11**
LOOKING AWAY, DOWN - Nervous	3	3		2	**8**

MOVING A LOT - Nervous	1	5		1	**7**
MOVING FEET - Nervous	7				**7**
HANDS IN THE POCKET, HOLDING - Nervous	4	2			**6**
TWISTING BODY- Nervous	3			1	**4**
GENERAL MOVEMENTS - Nervous	2	2			**4**
LOUD VOICE - Nervous	1		2		**3**
GIVING VOCAL FEEDBACK - Nervous			1	2	**3**
TALKING FAST - NERVOUS			2		**2**
TALKING SLOW - Nervous	2				**2**
VOICE - Nervous	2				**2**
SHORT ANSWERS AND FEEDBACK - Nervous	1			1	**2**
TALKING TOO MUCH - Nervous	2				**2**
NODDING - Nervous	1				**1**
TWISTING ARMS - Nervous		1			**1**
HE DOESN' T WANT INTERUPT - Nervous				1	**1**
LEANING BACKWARD - NERVOUS	1				**1**
LEANING FORWARD - NERVOUS	1				**1**
STOP TO TALK - Nervous	1				**1**
EYE CONTACT - NERVOUS	1				**1**
TOTAL	**44**	**22**	**18**	**25**	**109**

Nervous, insecure, tense, embarrassed 8604	**V+A**	**Video**	**Audio**	**Transcription**	**Total**
LAUGHING, SMILING - Nervous	3		8	8	**19**
LOOKING AWAY, DOWN - Nervous	2	5		7	**14**
MEANING OF WORDS - Nervous	3		5	5	**13**
MOVING A LOT - Nervous	7			1	**8**
STANDING - Nervous	6				**6**
HANDS IN THE POCKET, HOLDING - Nervous	2			3	**5**
STUMBLING, HESITATIONS - Nervous			2	3	**5**
MOVING FEET - Nervous		2		1	**3**
MOVING HANDS - Nervous		2			**2**
LOUD VOICE - Nervous	1				**1**
NEUTRAL VOICE - Nervous			1		**1**
SHORT FEEDBACK - Nervous			1		**1**
NOISY VOICE - Nervous	1				**1**
TOTAL	**25**	**9**	**17**	**28**	**79**

Nervous, insecure, tense, embarrassed 8602	**V+A**	**Video**	**Audio**	**Transcription**	**Total**
STANDING - nervous	8	6			**14**
STUMBLING, HESITATIONS - nervous	4		4	1	**9**

MOVING HANDS, ARMS - nervous	1	6			**7**
MOVING A LOT - nervous	1	3			**4**
LAUGHING, SMILING - nervous			4		**4**
MEANING OF WORDS - nervous			2	1	**3**
LOOKING DOW, AROUND, AT CAMERA - nervous		3			**3**
SHORT ANSWER - nervous			2	1	**3**
HANDS IN THE POCKET, HOLDING - nervous	1	1			**2**
MOVING LEGS, FEET - nervous		2			**2**
LAUGHING POLITE - nervous	2				**2**
LOW VOICE - nervous			2		**2**
BREAK AFTER GREETING - nervous	2				**2**
SCRATCHING - nervous		1			**1**
SHAKING - nervous		1			**1**
LEANING BACKWARD - nervous		1			**1**
LOW EYEBROW - nervous		1			**1**
NODDING - nervous		1			**1**
SPEAKING FAST - nervous			1		**1**
TOTAL	**19**	**26**	**15**	**3**	**63**

Happiness

Happy, Glad, Enjoying the conversation	V+A	Video	Audio	Transcription	Total
LAUGHING, SMILING	18	34	18	7	77
LOUD VOICE	6		6		12
MEANING OF WORDS	2		5	3	10
MOVING HANDS AND ARMS	2	5			7
LEANING FORWARD	3	3			6
NODDING		4		2	6
EYE CONTACT	1	4			5
GENERAL MOVEMENTS		5			5
CLEAR VOICE	2		1		3
RECOGNIZING	2	1			3
AGREE		2			2
LEANING BACKWARD	2				2
MOVING HEAD FORWARD				2	2
TOTAL LOW VOICE			2		2
MOVING HANDS TO EXPLAIN		2			2
VOICE			2		2
OPEN GESTURE	1	1			2
ASKING QUESTION			1		1
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT				1	1
LEANING ON THE WALL		1			1
LIGHT VOICE	1				1
LOOKING DOW, AROUND, TO CAMERA	1				1
LOOKING UP		1			1
MIRRORING	1				1
SHOULDER DOWN		1			1
SPEAKING FAST			1		1
TAKING INITIATIVE	1				1
TALKING FAST	1				1
TOUCHING CHIN		1			1
VOICE GOING UP AND DOWN	1				1
WAITING THE END OF THE SENTENCE	1				1
TOTAL	46	65	36	15	162

Happy, Glad, Enjoying the conversation 8644	V+A	Video	Audio	Transcription	Total
LAUGHING, SMILING - HAPPY	2	16	10	2	30
MOVING HANDS AND ARMS- HAPPY		5			5
GENERAL MOVEMENTS - HAPPY		4			4
LOUD VOICE - HAPPY			3		3
NODDING - HAPPY		2			2
MEANING OF WORDS - HAPPY				1	1
RECOGNIZING - HAPPY		1			1

LOOKING UP - HAPPY		1			**1**
TOUCHING CHIN - HAPPY		1			**1**
LOW VOICE - HAPPY			1		**1**
MOVING HANDS TO EXPLAIN- HAPPY		1			**1**
TOTAL	**2**	**31**	**14**	**3**	**50**

Happy, Glad, Enjoying the conversation 8640	**V+A**	**Video**	**Audio**	**Transcription**	**Total**
LAUGHING, SMILING - HAPPY	2	5	4		**11**
MEANING OF WORDS - HAPPY	2		3		**5**
EYE CONTACT - HAPPY	1	4			**5**
LEANING FORWARD - HAPPY	1	1			**2**
VOICE - HAPPY			2		**2**
LOUD VOICE - HAPPY	1				**1**
CLEAR VOICE - HAPPY			1		**1**
RECOGNIZING - HAPPY	1				**1**
TAKING INITIATIVE - HAPPY	1				**1**
TALKING FAST - HAPPY	1				**1**
VOICE GOING UP AND DOWN - HAPPY	1				**1**
TOTAL	**11**	**10**	**10**	**0**	**31**

Happy, Glad, Enjoying the conversation 8604	**V+A**	**Video**	**Audio**	**Transcription**	**Total**
LAUGHING, SMILING - HAPPY	7	4	2	5	**18**
NODDING - HAPPY		2		2	**4**
LOUD VOICE - HAPPY	2		2		**4**
LEANING FORWARD - HAPPY	2	2			**4**
MEANING OF WORDS - HAPPY			1	2	**3**
MOVING HEAD FORWARD - HAPPY				2	**2**
LEANING BACKWARD - HAPPY	2				**2**
ASKING QUESTION - HAPPY			1		**1**
CLEAR VOICE - HAPPY	1				**1**
RECOGNIZING - HAPPY	1				**1**
OPEN GESTURE - HAPPY	1				**1**
MIRRORING - HAPPY	1				**1**
WAITING THE END OF THE SENTENCE - HAPPY	1				**1**
GENERAL MOVEMENTS - HAPPY		1			**1**
LEANING ON THE WALL - HAPPY		1			**1**
SHOULDER DOWN - HAPPY		1			**1**
TOTAL	**18**	**11**	**6**	**11**	**46**

Happy, Glad, Enjoying the conversation 8602	V+A	Video	Audio	Transcription	Total
LAUGHING, SMILING - happy	7	9	2		**18**
LOUD VOICE - happy	3		1		**4**
AGREE - happy		2			**2**
MOVING HANDS, ARMS - happy	2				**2**
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - happy				1	**1**
LOW VOICE - happy			1		**1**
MEANING OF WORDS - happy			1		**1**
OPEN GESTURE - happy		1			**1**
MOVING HANDS TO EXPLAIN - happy		1			**1**
CLEAR VOICE - happy	1				**1**
LIGHT VOICE - happy	1				**1**
SPEAKING FAST - happy			1		**1**
LOOKING DOW, AROUND, TO CAMERA - happy	1				**1**
TOTAL	**15**	**13**	**6**	**1**	**35**

Confidence

Confident, Calm, Relaxed	V+A	Video	Audio	Transcription	Total
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT	9	1	20	1	**31**
LAUGHING, SMILING	6	8	11	0	**25**
MEANING OF WORDS	8	0	9	7	**24**
STANDING	8	8	0	0	**16**
LEANING ON THE WALL	2	11	0	0	**13**
EYE CONTACT	4	8	0	1	**13**
LOUD VOICE	0	0	13	0	**13**
CLEAR VOICE	1	0	11	0	**12**
MOVING HANDS TO EXPLAIN	8	4	0	0	**12**
LEANING FORWARD	0	11	0	0	**11**
MOVING HANDS, ARMS	4	5	0	0	**9**
TAKING INITIATIVE	0	0	6	2	**8**
TONE OF VOICE	3	0	5	0	**8**
GENERAL MOVEMENTS	1	6	0	0	**7**
HANDS IN THE POCKET, HOLDING	5	1	0	0	**6**
NODDING	1	4	0	1	**6**
ASKING QUESTION	1	0	4	0	**5**
LOW VOICE	2		2		**4**
SPEAKING STRAIGHT			3	1	**4**
FLAT TONE OF VOICE	0	0	4	0	**4**
LEANING BACKWARD	1	3	0	0	**4**
STANCE	1	3	0	0	**4**
TALKING A LOT, MORE	0	0	4	0	**4**
AGREE		1	2		**3**
GIVING FEEDBACK			2	1	**3**
LOOKING DOW, AROUND, TO CAMERA	1	2	0	0	**3**
OPEN GESTURE	2	1	0	0	**3**
OVERLAPPING			2		**2**
TALKING SLOW	0	0	2	0	**2**
ADDING MORE INFO			1		**1**
ANSWERING	1				**1**
ARMS DOWN		1			**1**
BACK STRAIGHT		1			**1**
EXPRESSING OPINION			1		**1**
GIVING HAND STRAIGHT	1				**1**
HEAD MOVEMENTS	1				**1**
MIRRORING		1			**1**
MOVING HEAD	1				**1**
RISING EYEBROW	1				**1**
SCRATCHING TOUCHING BACK, FACE, ETC		1			**1**
SHORT ANSWER				1	**1**
SHOULDER DOWN	1				**1**

SLOW MOVEMENTS	1				**1**
STOP TO LAUGHING			1		**1**
STRAIGHT ANSWER				1	**1**
STRONG HANDSHAKE		1			**1**
TOTAL	**75**	**82**	**103**	**16**	**276**

Confident, Calm, Relaxed 8640	V+A	Video	Audio	Transcription	Total
MEANING OF WORDS - CONFIDENT	2		4	1	**7**
TAKING INITIATIVE - CONFIDENT			3		**3**
STANDING - CONFIDENT	5	5			**10**
NODDING - INTEREST	1				**1**
GIVING HAND STRAIGHT - CONFIDENT	1				**1**
TALKING SLOW - CONFIDENT			1		**1**
ASKING QUESTION - CONFIDENT			2		**2**
CLEAR VOICE - CONFIDENT			1		**1**
GIVING FEEDBACK - CONFIDENT			2	1	**3**
FLAT TONE OF VOICE - CONFIDENT			2		**2**
GENERAL MOVEMENTS - CONFIDENT		2			**2**
STANCE - CONFIDENT		1			**1**
HANDS IN THE POCKET - CONFIDENT		1			**1**
LAUGHING, SMILING - CONFIDENT			3		**3**
STANCE - CONFIDENT		2			**2**
LOUD VOICE - CONFIDENT			1		**1**
VOICE - CONFIDENT			2		**2**
EYE CONTACT - CONFIDENT	2	3			**5**
TOTAL	**11**	**14**	**21**	**2**	**48**

Confident, Calm, Relaxed 8604	V+A	Video	Audio	Transcription	Total
MEANING OF WORDS - CONFIDENT	2		3	3	**8**
TAKING INITIATIVE - CONFIDENT			2	1	**3**
STRAIGHT ANSWER - CONFIDENT				1	**1**
SHORT ANSWER - CONFIDENT				1	**1**
ASKING QUESTION - CONFIDENT			2		**2**
CLEAR VOICE - CONFIDENT			1		**1**
LOOKING THE CAMERA - CONFIDENT	1				**1**
LAUGHING, SMILING - CONFIDENT	2		4		**6**
STANCE - CONFIDENT	1				**1**
LOUD VOICE - CONFIDENT			2		**2**
LEANING FORWARD - CONFIDENT		1			**1**
LEANING BACKWARD - CONFIDENT		1			**1**
VOICE - CONFIDENT	1		1		**2**

EYE CONTACT - CONFIDENT		1			**1**
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - CONFIDENT			11		**11**
LEANING ON THE WALL - CONFIDENT	2	11			**13**
ADDING MORE INFO - CONFIDENT			1		**1**
BACK STRAIGHT - CONFIDENT		1			**1**
OVERLAPPING - CONFIDENT			2		**2**
TOTAL	**9**	**15**	**29**	**6**	**59**

Confident, Calm, Relaxed 8602	**V+A**	**Video**	**Audio**	**Transcription**	**Total**
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - confident	6		7	1	**14**
LOUD VOICE - confident			10		**10**
CLEAR VOICE - confident			9		**9**
MOVING HANDS TO EXPLAIN - confident	5	3			**8**
LAUGHING, SMILING - confident	2	4	1		**7**
MOVING HANDS, ARMS - confident	4	3			**7**
MEANING OF WORDS - confident	2			3	**5**
HANDS IN THE POCKET, HOLDING - confident	5				**5**
LEANING FORWARD - confident		5			**5**
NODDING - confident		4		1	**5**
LEANING FORWARD - confident		5			**5**
EYE CONTACT - confident		3		1	**4**
LOW VOICE - confident	2		2		**4**
GENERAL MOVEMENTS - confident		4			**4**
SPEAKING STRAIGHT - confident			3	1	**4**
STANDING - confident	3				**3**
LEANING BACKWARD - confident	1	2			**3**
AGREE - confident		1	2		**3**
TAKING INITIATIVE - confident			1	1	**2**
OPEN GESTURE - confident	2				**2**
LOOKING DOW, AROUND, TO CAMERA - confident		2			**2**
SHOULDER DOWN - confident	1				**1**
MOVING HEAD - confident	1				**1**
STRONG HANDSHAKE - confident		1			**1**
MIRRORING - confident		1			**1**
SPEAKING A LOT - confident			1		**1**
SLOW SPEAK - confident			1		**1**
EXPRESSING OPINION - confident			1		**1**
TOTAL	**34**	**38**	**38**	**8**	**118**

Interest

Interested, curios, listening	V+A	Video	Audio	Transcription	Total
EYE CONTACT	13	29		2	44
VOCAL FEEDBACK	16		15	3	34
NODDING	7	23		2	32
ASKING QUESTIONS	11		10	5	26
LOUD VOICE	7		16		23
LAUGHING AND SMILING	6	7	9		22
MEANING WORDS	4		10	5	19
CLEAR VOICE	5		10		15
FORWARD	3	5			8
MOVING HANDS TO EXPLAIN	3	4			7
MOVING HANDS	4	3			7
BACKWARD	1	3			4
GENERAL MOVEMENTS	1	3			4
AGREE		2	2		4
VOICE	3		1		4
FACE EXPRESSIONS		3			3
MOVING HEAD FORWARD		3			3
STANDING		3			3
OPEN MOUTH	2	1			3
HANDS IN THE POCKET, HOLDING	1	1		1	3
SPEAKING FAST	2		1		3
OPEN GESTURE	1	1			2
RISING EYEBROW	1	1			2
ANSWERING FAST	2				2
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT			2		2
TALKING FAST			2		2
HOLDING HANDS. HANDS BACK		1			1
STOP TO GESTURING		1			1
GIVING HAND STRAIGHT	1				1
LOOKING AWAY, DOWN	1				1
ANSWERING	1				1
ANALIZYING			1		1
LOW VOICE			1		1
TRY TO GUESS			1		1
MOVING HEAD FORWARD				1	1
SHORT ANSWER				1	1
WHISPERING				1	1
TOTAL	96	94	81	21	292

Interested, curios, listening 8644	V+A	Video	Audio	Transcription	Total
NODDING - INTEREST	2	9			**11**
EYE CONTACT - INTEREST		10			**10**
LAUGHING, SMILING - INTEREST		1	7		**8**
GIVING VOCAL FEEDBACK - INTEREST	2		5		**7**
LOUD VOICE - INTEREST	1		3		**4**
ASKING QUESTION - INTEREST	1		2		**3**
MEANING OF WORDS - INTEREST				2	**2**
MOVING HANDS TO EXPLAIN- INTEREST	1				**1**
LOW VOICE - INTEREST			1		**1**
GENERAL MOVEMENTS - INTEREST	1				**1**
WHISPERING - INTERST				1	**1**
VOICE - INTEREST			1		**1**
LOOKING AWAY, DOWN - INTEREST	1				**1**
TOTAL	**9**	**20**	**19**	**3**	**51**

Interested, curios, listening 8640	V+A	Video	Audio	Transcription	Total
CLEAR VOICE - INTEREST	1		10		**11**
EYE CONTACT - INTEREST	5	3		1	**9**
LAUGHING, SMILING - INTEREST	3	4	2		**9**
LOUD VOICE - INTEREST			9		**9**
NODDING - INTEREST	1	4		1	**6**
MOVING HANDS - INTEREST	3	2			**5**
ASKING QUESTION - INTEREST				3	**3**
LEANING FORWARD - INTEREST	2	1			**3**
STANDING - INTEREST		3			**3**
FACE EXPRESSIONS - INTERST		3			**3**
VOICE - INTEREST	3				**3**
HANDS IN THE POCKET, HOLDING - INTEREST	1	1		1	**3**
LEANING BACKWARD - INTEREST	1	1			**2**
TALKING FAST - INTEREST			2		**2**
RISING EYEBROW - INTEREST	1	1			**2**
GIVING VOCAL FEEDBACK - INTEREST				1	**1**
MEANING OF WORDS - INTEREST				1	**1**
OPEN MOUTH - INTEREST		1			**1**
MOVING HANDS TO EXPLAIN- INTEREST	1				**1**
GIVING HAND STRAIGHT - INTEREST	1				**1**
MOVING HEAD FORWARD - INTEREST				1	**1**
GENERAL MOVEMENTS - INTEREST		1			**1**
TOTAL	**23**	**25**	**23**	**9**	**80**

Interested, curios, listening 8604	V+A	Video	Audio	Transcription	Total
EYE CONTACT - INTEREST		15			**15**
MEANING OF WORDS - INTEREST	4		6	1	**11**
GIVING VOCAL FEEDBACK - INTEREST	6		2	2	**10**
NODDING - INTEREST	2	4			**6**
ASKING QUESTION - INTEREST	3		2		**5**
LAUGHING, SMILING - INTEREST	3				**3**
LEANING FORWARD - INTEREST		3			**3**
OPEN MOUTH - INTEREST	2				**2**
LOUD VOICE - INTEREST	1		1		**2**
SHORT ANSWER - INTEREST				1	**1**
ANSWERING - INTEREST	1				**1**
LEANING BACKWARD - INTEREST		1			**1**
MOVING HANDS - INTEREST	1				**1**
TRY TO GUESS - INTEREST			1		**1**
ANALIZYING - INTEREST			1		**1**
TOTAL	**23**	**23**	**13**	**4**	**63**

Interested, curios, listening 8602	V+A	Video	Audio	Transcription	Total
GIVING VOCAL FEEDBACK - interest	8		8		**16**
ASKING QUESTION - interest	7		6	2	**15**
EYE CONTACT - interest	8	1		1	**10**
NODDING - interest	2	6		1	**9**
LOUD VOICE - interest	5		3		**8**
MEANING OF WORDS - interest			4	1	**5**
MOVING HANDS TO EXPLAIN, TO POINT - interest	1	4			**5**
AGREE - interest		2	2		**4**
CLEAR VOICE - interest	4				**4**
MOVING HEAD FORWARD - interest		3			**3**
SPEAKING FAST - interest	2		1		**3**
LAUGHING, SMILING - interest		2			**2**
ANSWERING FAST - interest	2				**2**
OPEN GESTURE - interest	1	1			**2**
LEANING FORWARD - interest	1	1			**2**
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - interest			2		**2**
GENERAL MOVEMENTS - interest		2			**2**
LEANING BACKWARD - interest		1			**1**
MOVING HANDS - interest		1			**1**
HOLDING HANDS. HANDS BACK - interest		1			**1**
STOP TO GESTURING - interest		1			**1**
TOTAL	**41**	**26**	**26**	**5**	**98**

Disinterest

Disinterested, Bored	V+A	Video	Audio	Transcription	Total
VOICE GOING UP AND DOWN			8		8
GIVING VOCAL FEEDBACK	3	0	1	3	7
LOW VOICE	3	0	4	0	7
MEANING OF WORDS	1	0	0	3	4
SPEAKING SLOW	1	0	3	0	4
SHORT CONTRIBUTION	1	0	3	0	4
LOOKING DOW, AROUND, TO CAMERA	0	2	0	0	2
MOVING HANDS, ARMS	1	1	0	0	2
FLAT TONE OF VOICE	1		1		2
LEANING BACKWARD		2			2
MOOVING LEGS, FEET		2			2
NO RESPONDING, NO FEEDBACK	1		1		2
A LOT OF REPETITION				1	1
BACK STRAIGHT	1				1
BREATHING DEEPLY			1		1
FAKE LAUGHING	1				1
GENERAL MOVEMENTS	1				1
LOUD VOICE			1		1
MOVING A LOT		1			1
NO TALKING SO MUCH				1	1
NOT CLEAR VOICE	1				1
RASPY VOICE	1				1
STANDING	1				1
TWISTING BODY				1	1
VOICE			1		1
TOTAL	18	8	24	9	59

Disinterested, Bored 8644	V+A	Video	Audio	Transcription	Total
LOW VOICE - DISINTEREST	1		2		3
NO RESPONDING, NO FEEDBACK - DISINTEREST	1		1		2
SHORT AND FAST FEEDBACK - DISINTEREST	1				1
VOICE - DISINTEREST			1		1
LAUGHING SHORT - DISINTEREST			1		1
MEANING OF WORDS - DISINTEREST				1	1
BREATHING DEEPLY - DISINTEREST			1		1
LOUD VOICE - DISINTEREST			1		1
GIVING FEEDBACK - DISINTEREST			1		1
FAKE LAUGHING - DISINTEREST	1				1
GENERAL MOVEMENTS - DISINTEREST	1				1
TOTAL	5	0	8	1	14

Disinterested, Bored 8640	V+A	Video	Audio	Transcription	Total
SLOW VOICE - DISINTEREST			2		**2**
NO TALKING SO MUCH - DISINTEREST				1	**1**
LOOKING AWAY, DOWN - DISINTEREST		1			**1**
TWISTING BODY- DISINTEREST				1	**1**
LOW VOICE - DISINTEREST			1		**1**
TOTAL	**0**	**1**	**3**	**2**	**6**

Disinterested, Bored 8604	V+A	Video	Audio	Transcription	Total
VOICE GOING UP AND DOWN - DISINTEREST			8		**8**
GIVING FEEDBACK - DISINTEREST	3			1	**4**
MEANING OF WORDS - DISINTEREST				2	**2**
FLAT TONE OF VOICE - DISINTEREST	1		1		**2**
RASPY VOICE - DISINTEREST	1				**1**
MOVING HEAD - DISINTEREST	1				**1**
SHORT QUESTION - DISINTEREST			1		**1**
SHORT ANSWERS - DISINTEREST			1		**1**
SLOW VOICE - DISINTEREST			1		**1**
LOW VOICE - DISINTEREST			1		**1**
STANDING - DISINTEREST	1				**1**
BACK STRAIGHT - DISINTEREST	1				**1**
TOTAL	**8**	**0**	**13**	**3**	**24**

Disinterested, Bored 8602	V+A	Video	Audio	Transcription	Total
LOW VOICE - DISINTEREST	2				**2**
MOOVING LEGS, FEET - DISINTEREST		2			**2**
LEANING BACKWARD - DISINTEREST		2			**2**
GIVING FEEDBACK - DISINTEREST				2	**2**
MEANING OF WORDS - DISINTEREST	1				**1**
SPEAKING SLOW - DISINTEREST	1				**1**
NOT CLEAR VOICE - DISINTEREST	1				**1**
MOVING A LOT - DISINTEREST		1			**1**
MOVING HANDS, ARMS - DISINTEREST		1			**1**
LOOKING DOW, AROUND, TO CAMERA - DISINTEREST		1			**1**
A LOT OF REPETITION - DISINTEREST				1	**1**
TOTAL	**5**	**7**	**0**	**3**	**15**

Thoughtfulness and Understanding

Thoughtfulness	V+A	Video	Audio	Transcription	Total
Looking up	2	3			**5**
Looking away		1			**1**
General movements	1				**1**
Low voice	1				**1**
Meaning of words	1				**1**
Laughing			1		**1**
TOTAL	**5**	**4**	**1**	**0**	**10**

Understanding	V+A	Video	Audio	Transcription	Total
Smiling		3			**3**
Looking around, in the camera		2			**2**
Nodding		1			**1**
Answering			1		**1**
Standing	1				**1**
Meaning of words	1				**1**
TOTAL	**2**	**6**	**1**	**0**	**9**

RECORDING V8602

STATES V8602	V+A	Video	Audio	Transcription	Total
LAUGHING, SMILING - happy	7	9	2		**18**
GIVING FEEDBACK - interest	8		8		**16**
ASKING QUESTION - interest	7		6	2	**15**
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - confident	6		7	1	**14**
STANDING - nervous	8	6			**14**
LOUD VOICE - confident			10		**10**
EYE CONTACT - interest	8	1		1	**10**
CLEAR VOICE - confident			9		**9**
STUMBLING, HESITATIONS - nervous	4		4	1	**9**
NODDING - interest	2	6		1	**9**
MOVING HANDS TO EXPLAIN - confident	5	3			**8**
LOUD VOICE - interest	5		3		**8**
LAUGHING, SMILING - confident	2	4	1		**7**
MOVING HANDS, ARMS - confident	4	3			**7**
MOVING HANDS, ARMS - nervous	1	6			**7**
MEANING OF WORDS - confident	2			3	**5**
HANDS IN THE POCKET, HOLDING - confident	5				**5**
LEANING FORWARD - confident		5			**5**
NODDING - confident		4		1	**5**
LEANING FORWARD - confident		5			**5**
MEANING OF WORDS - interest			4	1	**5**
MOVING HANDS TO EXPLAIN, TO POINT - interest	1	4			**5**
EYE CONTACT - confident		3		1	**4**
LOW VOICE - confident	2		2		**4**
GENERAL MOVEMENTS - confident		4			**4**
SPEAKING STRAIGHT - confident			3	1	**4**
MOVING A LOT - nervous	1	3			**4**
LAUGHING, SMILING - nervous			4		**4**
AGREE - interest		2	2		**4**
CLEAR VOICE - interest	4				**4**
LOUD VOICE - happy	3		1		**4**
Concentrate – Eye contact	3				**3**
Concentrate – Standing	3				**3**
STANDING - confident	3				**3**
LEANING BACKWARD - confident	1	2			**3**
AGREE - confident		1	2		**3**
MEANING OF WORDS - nervous			2	1	**3**
LOOKING DOW, AROUND, TO CAMERA - nervous		3			**3**
SHORT ANSWER - nervous			2	1	**3**
MOVING HEAD FORWARD - interest		3			**3**

SPEAKING FAST - interest	2		1		**3**
Understanding - smiling		3			**3**
Informative - moving hands		2			**2**
Wondering - moving hands		2			**2**
LOW VOICE - disinterest	2				**2**
MOOVING LEGS, FEET - disinterest		2			**2**
LEANING BACKWARD - disinterest		2			**2**
GIVING FEEDBACK - disinterest				2	**2**
Agree – nodding		2			**2**
Surprise – giving feedback			2		**2**
TAKING INITIATIVE - confident			1	1	**2**
OPEN GESTURE - confident	2				**2**
LOOKING DOW, AROUND, TO CAMERA - confident		2			**2**
HANDS IN THE POCKET, HOLDING - nervous	1	1			**2**
MOVING LEGS, FEET - nervous		2			**2**
LAUGHING POLITE - nervous	2				**2**
LOW VOICE - nervous			2		**2**
BREAK AFTER GREETING - nervous	2				**2**
LAUGHING, SMILING - interest		2			**2**
ANSWERING FAST - interest	2				**2**
OPEN GESTURE - interest	1	1			**2**
LEANING FORWARD - interest	1	1			**2**
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - interest			2		**2**
GENERAL MOVEMENTS - interest		2			**2**
AGREE - happy		2			**2**
MOVING HANDS, ARMS - happy	2				**2**
Understanding - looking around, in the camera		2			**2**
Aggressive – moving legs, feet		1			**1**
Compassionate for B – laughing polite			1		**1**
Conservative – general movements (small)		1			**1**
Energetic – moving a lot	1				**1**
Excited – open gesture		1			**1**
Overexcited – loud voice	1				**1**
Pushing-speaking fast			1		**1**
Recognition - laughing			1		**1**
Restrictive – touching hands close to the belly		1			**1**
Showing off - meaning of words				1	**1**
Want to be understandable - meaning of words				1	**1**
Willing to take the situation – moving hands		1			**1**
MEANING OF WORDS - disinterest	1				**1**
SPEAKING SLOW - disinterest	1				**1**
CLEAR NOT VOICE - disinterest	1				**1**

MOVING A LOT - disinterest		1			**1**
MOVING HANDS, ARMS - disinterest		1			**1**
LOOKING DOW, AROUND, TO CAMERA - disinterest		1			**1**
A LOT OF REPETITION - disinterest				1	**1**
Agree – smiling		1			**1**
Surprise – meaning of words			1		**1**
Surprise – going back	1				**1**
Explaining – flat voice			1		**1**
Explaining - meaning of words				1	**1**
Superior – giving feedback			1		**1**
Superior – asking questions			1		**1**
Thoughtful – general movements	1				**1**
Thoughtful – low voice	1				**1**
Thoughtful – meaning of words	1				**1**
Thoughtful – smile			1		**1**
Threatened – speaking fast			1		**1**
SHOULDER DOWN - confident	1				**1**
MOVING HEAD - confident	1				**1**
STRONG HANDSHAKE - confident		1			**1**
MIRRORING - confident		1			**1**
SPEAKING A LOT - confident			1		**1**
SLOW SPEAK - confident			1		**1**
EXPRESSING OPINION - confident			1		**1**
SCRATCHING - nervous		1			**1**
SHAKING - nervous		1			**1**
LEANING BACKWARD - nervous		1			**1**
LOW EYEBROW - nervous		1			**1**
NODDING - nervous		1			**1**
SPEAKING FAST - nervous			1		**1**
LEANING BACKWARD - interest		1			**1**
MOVING HANDS - interest		1			**1**
HOLDING HANDS. HANDS BACK - interest		1			**1**
STOP TO GESTURING - interest		1			**1**
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - happy				1	**1**
LOW VOICE - happy			1		**1**
MEANING OF WORDS - happy			1		**1**
OPEN GESTURE - happy		1			**1**
MOVING HANDS TO EXPLAIN - happy		1			**1**
CLEAR VOICE - happy	1				**1**
LIGHT VOICE - happy	1				**1**
SPEAKING FAST - happy			1		**1**
LOOKING DOW, AROUND, TO CAMERA - happy	1				**1**

Understanding - nodding		1			**1**
TOTAL	126	128	96	23	373

RECORDING V8604

STATES V8604	V+A	Video	Audio	Transcription	Total
LAUGHING, SMILING - Nervous	3		8	8	**19**
LAUGHING, SMILING - HAPPY	7	4	2	5	**18**
EYE CONTACT - INTEREST		15			**15**
LOOKING AWAY, DOWN - Nervous	2	5		7	**14**
LEANING ON THE WALL - CONFIDENT	2	11			**13**
MEANING OF WORDS - Nervous	3		5	5	**13**
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - CONFIDENT			11		**11**
MEANING OF WORDS - INTEREST	4		6	1	**11**
GIVING FEEDBACK - INTEREST	6		2	2	**10**
VOICE GOING UP AND DOWN - DISINTEREST			8		**8**
MEANING OF WORDS - CONFIDENT	2		3	3	**8**
MOVING A LOT - Nervous	7			1	**8**
LAUGHING, SMILING - CONFIDENT	2		4		**6**
STANDING - Nervous	6				**6**
ASKING QUESTION - INTEREST	3		2		**5**
HANDS IN THE POCKET, HOLDING - Nervous	2			3	**5**
STUMBLING, HESITATIONS - Nervous			2	3	**5**
GIVING FEEDBACK - DISINTEREST	3			1	**4**
LEANING FORWARD - HAPPY	2	2			**4**
LOUD VOICE - HAPPY	2		2		**4**
NODDING - HAPPY		2		2	**4**
voice up - Agree			3		**3**
TAKING INITIATIVE - CONFIDENT			2	1	**3**
LAUGHING, SMILING - INTEREST	3				**3**
LEANING FORWARD - INTEREST		3			**3**
MEANING OF WORDS - HAPPY			1	2	**3**
MOVING FEET - Nervous		2		1	**3**
MEANING OF WORDS - DISINTEREST				2	**2**
FLAT TONE OF VOICE - DISINTEREST	1		1		**2**
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - Satisfy			2		**2**
meaning of words - Surprise				2	**2**
ASKING QUESTION - CONFIDENT			2		**2**
LOUD VOICE - CONFIDENT			2		**2**
OVERLAPPING - CONFIDENT			2		**2**
TAKING INITIATIVE - CONFIDENT			2		**2**
VOICE - CONFIDENT	1		1		**2**
LOUD VOICE - INTEREST	1		1		**2**
OPEN MOUTH - INTEREST	2				**2**
LEANING BACKWARD - HAPPY	2				**2**

NODDING - HAPPY		2			**2**
MOVING HEAD FORWARD - HAPPY				2	**2**
MOVING FEET - Nervous		2			**2**
MOVING HANDS - Nervous		2			**2**
Meaning of words - Explaining	1		1		**2**
BACK STRAIGHT - DISINTEREST	1				**1**
LOW VOICE - DISINTEREST			1		**1**
MOVING HEAD - DISINTEREST	1				**1**
RASPY VOICE - DISINTEREST	1				**1**
SHORT ANSWERS - DISINTEREST			1		**1**
SHORT QUESTION - DISINTEREST			1		**1**
SLOW VOICE - DISINTEREST			1		**1**
STANDING - DISINTEREST	1				**1**
meaning of words - Exposing, guessing			1		**1**
meaning of words - Secretive			1		**1**
meaning of words - Seems strange	1				**1**
GENERAL MOVEMENTS (SMALL) - Answering	1				**1**
Laughing - Try to be polite	1				**1**
Looking up - Try to remember				1	**1**
Low tone of voice - Sad	1				**1**
SHORT ANSWERS - No surprise, no excited			1		**1**
eye contact - informative				1	**1**
Looking the camera - surprise				1	**1**
open eyes - Surprise	1				**1**
Up nod - Surprise				1	**1**
ADDING MORE INFO - CONFIDENT			1		**1**
BACK STRAIGHT - CONFIDENT		1			**1**
CLEAR VOICE - CONFIDENT			1		**1**
EYE CONTACT - CONFIDENT		1			**1**
LEANING BACKWARD - CONFIDENT		1			**1**
LEANING FORWARD - CONFIDENT		1			**1**
LOOKING THE CAMERA - CONFIDENT	1				**1**
STANCE - CONFIDENT	1				**1**
SHORT ANSWER - CONFIDENT				1	**1**
STRAIGHT ANSWER - CONFIDENT				1	**1**
ANALIZYING - INTEREST			1		**1**
ANSWERING - INTEREST	1				**1**
LEANING BACKWARD - INTEREST		1			**1**
MOVING HANDS - INTEREST	1				**1**
TRY TO GUESS - INTEREST			1		**1**
SHORT ANSWER - INTEREST				1	**1**
ASKING QUESTION - HAPPY			1		**1**
CLEAR VOICE - HAPPY	1				**1**
GENERAL MOVEMENTS - HAPPY		1			**1**
LEANING ON THE WALL - HAPPY		1			**1**
MIRRORING - HAPPY	1				**1**

OPEN GESTURE - HAPPY	1				**1**
RECOGNIZING - HAPPY	1				**1**
SHOULDER DOWN - HAPPY		1			**1**
WAITING A FINISHS THE SENTENCE - HAPPY	1				**1**
LOUD VOICE - Nervous	1				**1**
NEUTRAL VOICE - Nervous			1		**1**
NOISY VOICE - Nervous	1				**1**
SHORT FEEDBACK - Nervous			1		**1**
answering - Understanding			1		**1**
Meaning of words - Understanding	1				**1**
Standing - Understanding	1				**1**
flat tone of voice - Explaining	1				**1**
Laughing - Positive				1	**1**
NODDING - positive				1	**1**
flat tone of voice - Reserve	1				**1**
low tone of voice - Reserve	1				**1**
Low tone of voice - Tired			1		**1**
low tone of voice - Tired			1		**1**
low tone of voice - Upset			1		**1**
meaning of words - Upset			1		**1**
TOTAL	92	58	94	60	304

RECORDING V8640

STATES V8640	**V+A**	**Video**	**Audio**	**Transcription**	**Total**
STUMBLING, HESITATIONS, REPETITION - Nervous	4		6	5	**15**
MEANING OF WORDS - Nervous	1		2	9	**12**
MOVING HANDS AND ARMS- Nervous	5	6		1	**12**
LAUGHING, SMILING - Nervous	1	3	5	2	**11**
CLEAR VOICE - INTEREST	1		10		**11**
LAUGHING, SMILING - HAPPY	2	5	4		**11**
STANDING - CONFIDENT	5	5			**10**
EYE CONTACT - INTEREST	5	3		1	**9**
LAUGHING, SMILING - INTEREST	3	4	2		**9**
LOUD VOICE - INTEREST			9		**9**
LOOKING AWAY, DOWN - Nervous	3	3		2	**8**
MOVING A LOT - Nervous	1	5		1	**7**
MOVING FEET - Nervous	7				**7**
MEANING OF WORDS - CONFIDENT	2		4	1	**7**
HANDS IN THE POCKET, HOLDING - Nervous	4	2			**6**
NODDING - INTEREST	1	4		1	**6**
MOVING HANDS - INTEREST	3	2			**5**
MEANING OF WORDS - HAPPY	2		3		**5**
EYE CONTACT - INTEREST	1	4			**5**
EYE CONTACT - CONFIDENT	2	3			**5**

looking up - Try to remember, THOUGHTFULLY	2	3			**5**
TWISTING BODY- Nervous	3			1	**4**
GENERAL MOVEMENTS - Nervous	2	2			**4**
meaning of words - Excited				4	**4**
LOUD VOICE - Nervous	1		2		**3**
GIVING FEEDBACK - Nervous			1	2	**3**
ASKING QUESTION - INTEREST				3	**3**
LEANING FORWARD - INTEREST	2	1			**3**
STANDING - INTEREST		3			**3**
FACE EXPRESSIONS - INTEREST		3			**3**
VOICE - INTEREST	3				**3**
HANDS IN THE POCKET, HOLDING - INTEREST	1	1		1	**3**
TAKING INITIATIVE - CONFIDENT			3		**3**
GIVING FEEDBACK - CONFIDENT			2	1	**3**
LAUGHING, SMILING - CONFIDENT			3		**3**
STANCE - CONFIDENT		3			**3**
HOLDING HANDS - Explaining		3			**3**
TALKING FAST - NERVOUS			2		**2**
TALKING SLOW - Nervous	2				**2**
VOICE - Nervous	2				**2**
SHORT ANSWERS AND FEEDBACK - Nervous	1			1	**2**
TALKING TOO MUCH - Nervous	2				**2**
LEANING BACKWARD - INTEREST	1	1			**2**
TALKING FAST - INTEREST			2		**2**
RISING EYEBROW - INTEREST	1	1			**2**
LEANING FORWARD - HAPPY	1	1			**2**
VOICE - HAPPY			2		**2**
ASKING QUESTION - CONFIDENT			2		**2**
FLAT TONE OF VOICE - CONFIDENT			2		**2**
GENERAL MOVEMENTS - CONFIDENT		2			**2**
VOICE - CONFIDENT			2		**2**
SLOW VOICE - DISINTEREST			2		**2**
MOVING HANDS - EXPLEANING		2			**2**
MOVING HANDS - SHOWING OFF		2			**2**
MOVING A LOT - EXCITED				2	**2**
HOLDING ARMS AND HANDS - ENDING THE DISCUSSION		2			**2**
MOVING HEAD - Surprise	2				**2**
GIVING FEEDBACK - AGREE				2	**2**
LAUGHING, SMILING - EXCITED				2	**2**
NODDING - Nervous	1				**1**
TWISTING ARMS - Nervous		1			**1**
HE DOESN' T WANT INTERUPT - Nervous				1	**1**
LEANING BACKWARD - NERVOUS	1				**1**
LEANING FORWARD - NERVOUS	1				**1**
STOP TO TALK - Nervous	1				**1**
EYE CONTACT - NERVOUS	1				**1**

GIVING FEEDBACK - INTEREST				1	**1**
MEANING OF WORDS - INTEREST				1	**1**
OPEN MOUTH - INTEREST		1			**1**
MOVING HANDS TO EXPLAIN- INTEREST	1				**1**
GIVING HAND STRAIGHT - INTEREST	1				**1**
MOVING HEAD FORWARD - INTEREST				1	**1**
GENERAL MOVEMENTS - INTEREST		1			**1**
LOUD VOICE - HAPPY	1				**1**
CLEAR VOICE - HAPPY			1		**1**
RECOGNIZING - HAPPY	1				**1**
TAKING INITIATIVE - HAPPY	1				**1**
TALKING FAST - HAPPY	1				**1**
VOICE GOING UP AND DOWN - HAPPY	1				**1**
NODDING - CONFIDENT	1				**1**
GIVING HAND STRAIGHT - CONFIDENT	1				**1**
TALKING SLOW - CONFIDENT			1		**1**
CLEAR VOICE - CONFIDENT			1		**1**
HANDS IN THE POCKET - CONFIDENT		1			**1**
LOUD VOICE - CONFIDENT			1		**1**
NO TALKING SO MUCH - DISINTEREST				1	**1**
LOOKING AWAY, DOWN - DISINTEREST		1			**1**
TWISTING BODY- DISINTEREST				1	**1**
LOW VOICE - DISINTEREST			1		**1**
meaning of words - AGREE				1	**1**
NO SMILING - TRY TO BE COOL		1			**1**
GENERAL MOVEMENTS - UNDERSTANDING		1			**1**
STANDING - EXPLEANING		1			**1**
GENERAL MOVEMENTS (SMALL) - RESERVE		1			**1**
MOVING HANDS - ANSWERING		1			**1**
LOUD VOICE - PRIDE			1		**1**
DOWN-NOD - ENCOURAGING				1	**1**
NODDING - UNDERSTANDING				1	**1**
VOICE GOING UP AND DOWN - EXCITED			1		**1**
Meaning of words - AWKWARD				1	**1**
LAUGHING, SMILING - AGREE				1	**1**
meaning of words - RECOGNITION	1				**1**
giving feedback - NO EXCITED				1	**1**
Standing - RESERVED		1			**1**
giving feedback - SURPRISE	1				**1**
rising eyebrow - SURPRISE	1				**1**
TOTAL	96	90	77	54	317

RECORDING V8644

STATES V8644	V+A	Video	Audio	Transcription	Total
LAUGHING, SMILING - HAPPY	2	16	10	2	**30**
LAUGHING, SMILING - Nervous	7	2	10	2	**21**
LOOKING AWAY, DOWN - Nervous	10	5		2	**17**
MEANING OF WORDS - Nervous			4	8	**12**
NODDING - INTEREST	2	9			**11**
MOVING HANDS AND ARMS- Nervous	5	3		2	**10**
EYE CONTACT - INTEREST		10			**10**
LAUGHING, SMILING - CONFIDENT	2	4	3		**9**
SCRATCHING TOUCHING BACK, FACE, ETC - Nervous	5	3			**8**
LAUGHING, SMILING - INTEREST		1	7		**8**
STUMBLING, HESITATIONS, REPETITION - Nervous	4			2	**7**
GIVING FEEDBACK - INTEREST	2		5		**7**
HANDS IN THE POCKET, HOLDING - Nervous	3	2		1	**6**
MOVING FEET - Nervous	4	2			**6**
HE/SHE KNOWS WHAT HE/SHE IS TALKING ABOUT - CONFIDENT	3	1	2		**6**
GENERAL MOVEMENTS - Nervous	3	2			**5**
MOVING HANDS AND ARMS- HAPPY		5			**5**
MOVING HANDS - EXPLEANING		5			**5**
TWISTING BODY, SHAKING - Nervous	2	2			**4**
TALKING TOO MUCH - Nervous	2			2	**4**
LOUD VOICE - INTEREST	1		3		**4**
GENERAL MOVEMETNS - HAPPY		4			**4**
MEANING OF WORDS - CONFIDENT	2		2		**4**
MOVING HANDS TO EXPLAIN- CONFIDENT	3	1			**4**
VOICE - CONFIDENT	2		2		**4**
LAUGHING, SMILING - AGREE	4				**4**
STOP TO EXPLAIN, TO SPEAK - NERVOUS		3			**3**
ASKING QUESTION - INTEREST	1		2		**3**
LOUD VOICE - HAPPY			3		**3**
STANDING - CONFIDENT		3			**3**
TALKING A LOT, MORE - CONFIDENT			3		**3**
EYE CONTACT - CONFIDENT	2	1			**3**
LOW VOICE - DISINTEREST	1		2		**3**
LOW TONE OF VOICE - UNDERSTANDING			3		**3**
MEANING OF WORDS - TRY TO UNDERSTAND				3	**3**
MOVING A LOT - Nervous	2				**2**
ARMS DOWN - NERVOUS		2			**2**
VOICE - Nervous	1		1		**2**
GIVING FEEDBACK - Nervous	1			1	**2**
MEANING OF WORDS - INTEREST				2	**2**

NODDING - HAPPY		2			**2**
MOVING HANDS - CONFIDENT		2			**2**
FLAT TONE OF VOICE - CONFIDENT			2		**2**
NO RESPONDING, NO FEEDBACK - DISINTEREST	1		1		**2**
meaning of words - SARACASTIC				2	**2**
LOW EYEBROW - EXPLEANING		2			**2**
FLAT TONE OF VOICE - EXPLEANING			2		**2**
CLOSE MOUTH - EXPLEANING		2			**2**
NOT UNDERSTADING - SHORT CONFIRMATIONS				2	**2**
giving feedback - AGREE	2				**2**
THE OTHE PERSON DOESN'T GIVE FEEDBACK - Nervous	1				**1**
LIFTING HEELS - NERVOUS		1			**1**
BREATHING DEEPLY - NERVOUS			1		**1**
TAKLING LITTLE - NERVOUS			1		**1**
TONE OF VOICE UP AND DOWN - NERVOUS			1		**1**
LEANING BACKWARD - NERVOUS	1				**1**
BECAUSE THE OTHER PERSON IS NERVOUS - Nervous	1				**1**
EYE CONTACT - NERVOUS		1			**1**
MIRRORING - Nervous	1				**1**
DOWN EYEBROW - Nervous	1				**1**
MOVING HANDS TO EXPLAIN- INTEREST	1				**1**
LOW VOICE - INTEREST			1		**1**
GENERAL MOVEMENTS - INTEREST	1				**1**
WHISPERING - INTERST				1	**1**
VOICE - INTEREST			1		**1**
LOOKING AWAY, DOWN - INTEREST	1				**1**
MEANING OF WORDS - HAPPY				1	**1**
RECOGNIZING - HAPPY		1			**1**
LOOKING UP - HAPPY		1			**1**
TOUCHING CHIN - HAPPY		1			**1**
LOW VOICE - HAPPY			1		**1**
MOVING HANDS TO EXPLAIN- HAPPY		1			**1**
RISING EYEBROW - CONFIDENT	1				**1**
ASKING QUESTION - CONFIDENT	1				**1**
CLEAR VOICE - CONFIDENT	1				**1**
HEAD MOVEMENTS - CONFIDENT	1				**1**
GENERAL MOVEMENTS (MORE NATURAL) - CONFIDENT	1				**1**
SCRATCHING TOUCHING BACK, FACE, ETC - CONFIDENT		1			**1**
ARMS DOWN - CONFIDENT		1			**1**
STOP TO LAUGHING - CONFIDENT			1		**1**
ANSWERING - CONFIDENT	1				**1**
OPEN MOVEMENTS - CONFIDENT		1			**1**

SLOW MOVEMENTS - CONFIDENT	1				**1**
SHORT AND FAST FEEDBACK - NO INTERST	1				**1**
VOICE - CONFIDENT			1		**1**
LAUGHING SHORT - NO INTERST			1		**1**
MEANING OF WORDS - DISINTEREST				1	**1**
BREATHING DEEPLY - DISINTEREST			1		**1**
LOUD VOICE			1		**1**
GIVING FEEDBACK - DISINTEREST			1		**1**
FAKE LAUGHING - DISINTEREST	1				**1**
GENERAL MOVEMENTS - DISINTEREST	1				**1**
LOKING AWAY - THOUGHTFULLY		1			**1**
ASKING QUESTIONS - TRY TO UNDERSTAND			1		**1**
FAKE LAUGHING - TRY TO BE RALX			1		**1**
SIGHING - UPGIVING				1	**1**
RASPY VOICE - DEFENSIVE			1		**1**
TAKING TIME - CONCERN			1		**1**
LOW TONE OF VOICE - EMPHATIC			1		**1**
LOW TONE OF VOICE - SAD			1		**1**
LOW TONE OF VOICE - CONCERN			1		**1**
LOOKING UP - TOUGHTFULLY				1	**1**
STUMBLING, HESITATIONS - TIRED				1	**1**
Meaning of words - TAKING DISTANCE			1		**1**
meaning of words - OPEING UP			1		**1**
meaning of words - EXPLEANING			1		**1**
NOT SMILING - TRY TO UNDERSTAND		1			**1**
GIVING FEEDBACK - TRY TO UNDERSTAND	1				**1**
OPEN ARMS - REQUESTING		1			**1**
LAUGHING, SMILING - EXPLEANING		1			**1**
EXCLAMATION- REQUESTING		1			**1**
TOTAL	98	108	89	37	333