ÅSA ABELIN
JENS ALLWOOD
Department of Linguistics, Göteborg University, Göteborg, Sweden
E-Mail: abelin@ling.gu.se, jens@ling.gu.se

CROSS LINGUISTIC INTERPRETATION OF EXPRESSIONS OF EMOTIONS

Abstract
This study has three purposes: the first is to study if there is any stability in the way we interpret different emotions and attitudes from prosodic patterns, the second is to see if this interpretation is dependent on the listeners cultural and linguistic background, and the third is to find out if there is any reoccurring relation between acoustic and semantic properties of the stimuli.

Recordings of a Swedish speaker uttering a phrase while expressing different emotions was interpreted by listeners with different L1:s, Swedish, English, Finnish and Spanish, who were to judge the emotional contents of the expressions.

The results show that some emotions are interpreted in accordance with intended emotion in a greater degree than the other emotions were, e.g. “anger”, “fear”, “sadness” and “surprise”, while other emotions are interpreted as expected to a lesser degree. Furthermore emotions are interpreted with different degrees of success depending on the L1 of listeners; native listeners were the most successful. There is evidence that emotions with similar semantic features, e.g. “anger” and “dominance” or “fear” and “shyness” have similar acoustic features e.g. short duration and strong intensity (“anger” and “dominance”) or longer duration and weak intensity (“fear” and “shyness”). The results will be compared with interpretations of emotions expressed by Spanish speakers.

1. Introduction

There is a growing interest in comparing emotional prodosy in different languages. The question whether the expressions of emotions are universal or language specific also has importance for the question of how we interpret emotions. A recent cross-linguistic study of the interpretation of emotional prosody investigating the universality versus cultural relativity of emotions is Scherer, Banse and Wallbott (2001). This study, which was conducted in nine countries in Europe, the United States and Asia finds an overall recognition rate of 66% across all emotions and countries as well as a pattern of confusion which was very similar across all countries. It also found that accuracy decreased with increasing language dissimilarity from German. The approach of the present study of interpretation of Swedish emotional prosody is similar to the study of Scherer, Banse and Wallbott (2001), but is however conducted in a much smaller scale. The classic study by Williams and Stevens (1972) concerned acoustic analysis and the language is English. In an earlier study by Davitz and Davitz (1959) eight subjects expressed ten different emotions while reciting parts of the alphabet. Each subject was recorded, and the recordings were then played aloud to 30 judges, who were to identify the emotions they heard, from a list of the ten emotions. The results showed significant differences between the emotions, but also showed a great variation between the different judges, speakers and emotions. Studies of German emotional prosody have been made by e.g. Paeschke, Kienast and Sendlmeier (1999). They showed among other things that the intensity of syllable stress distinguished the group of so called non-excited emotions (e.g. sadness) from excited emotions (e.g. anger), and that there were emotion specific F0 features. Hayashi (1999) studied how Japanese speakers interpret emotions, expressed by Japanese, in dialogue. The interpretations were mostly in accordance with the intended emotions. The results indicate that F0 conveys substantial information about emotional states. Chung (1999) investigated vocal expressions for emotions in Korean and the perception of emotions for Korean, American and French listeners. The results show that the listeners agreed on the global judgement of the emotions.
In the study of interpretation of emotional prosody there is a need for both acoustic and semantic analysis. The result of an acoustic analysis is not independent of how the semantic analysis is done. This is particularly true when dealing with cross cultural interpretation. There are different theories for which emotions are the primary, see e.g. Woodworth (1938), Izard (1971), Roseman (1979). In the present study we have chosen some of the most commonly discussed emotions and attitudes, “joy”, “surprise”, “sadness”, “fear”, “shyness”, “anger”, “dominance” and “disgust” (in English translation) in order to cover several types. For studies of emotional prosody there is a need for further theoretic analysis of the concept ’emotion’, and of relations between different emotions.

2. Method

2.1. Speech material and recording

In order to isolate the contribution of prosdy to the interpretation of emotions and attitudes we chose one carrier phrase in which the different emotions were to be expressed. The carrier phrase was salt sill, potatismos och pannkakor (salted herring, mashed potatoes and pan cakes). The idea was that many different emotions can be held towards food and that therefore the sentence was quite neutral in respect to different emotions. The phrase was spoken expressing different emotional states by a native Swedish male speaker, and tape recorded.

2.2. Listener tests

The recordings of the expressed emotions were presented to native speakers of Swedish (35 subjects), English (12 subjects), Finnish (23 subjects) and Spanish (23 subjects). They were told to write down the emotional state expressed in each utterance of the carrier phrase. They could choose any expression they liked, and where the listeners hade written the answer in their L1 the answers were later translated into Swedish.

2.3. Analysis

The answers given by the listener groups were classified into semantic fields. An example of how the answers were grouped is the field of "anger" containing e.g. the words (in translation) *angry, anger, mad, wrath*, and the field of "fear" containing e.g. the words *afraid, fear, scared, frightened, timid, horror, dismay, alarm*. Some subjects displayed great fantasy in their answers, as *suburban kitchen without beer*, making some answers a bit difficult to classify, but the method of free choice was still considered to give greater validity to the answers. Another problem, specific for cross language investigations like this, is that there often are no one-to-one correspondences between lexemes in different languages. The Spanish word *espantado* which expresses simultaneous "surpise" and "fear" has no equivalent in the three Swedish surprise words *förvånad* (emotional reaction towards something unexpected), *överraskad* (who has become surprised, and usually glad, by something unexpectedly happening) and *häpen* (very, and noticably, surpised). The details of this example shows the importance of doing a careful semantic analysis of different languages while comparing responses. Yet another question is that of cultural relativity of emotions. It is obvious that emotions are expressed (also non verbally) in different ways in different cultures, but are also certain emotions experienced in certain cultures but not in others? These questions are not answered in this paper, but deserve future attention.

3. Results

3.1. Interpretation of the emotions

The interpretations of some of the emotions are presented in the following tables. The figures show percent of answers from each language. The reason why percent does not always add up to 100% is that miscellaneous answers are not shown.

Happiness	**Swedish**	**English**	**Finnish**	**Spanish**
happy	89	50	35	22
content	0	8	4	22
surprised	3	0	9	0
sad	0	8	17	35

Table 1: The interpretations of "happiness".

The Swedish interpretations of "happiness" show quite a great agreement with intended emotion. Among the other languages Spanish is the most deviant. 35% of the Spanish speakers interpret "happiness" as "sadness". 17% of the Finnish speakers interpret the emotion as "sad". The Swedish interpretations of "surprise" show great agreement with intended emotion. Among the other languages Spanish is again the most deviant. 30% of the Spanish speakers interpret "surprise" as "anger".

Surprise	**Swedish**	**English**	**Finnish**	**Spanish**
surprised	74	42	65	22
happy	9	0	0	0
content	0	0	0	9
sad	0	0	9	0
afraid	0	0	0	17
angry	0	0	9	30

Table 2: The interpretations of "surprise".

The emotion "sadness" is quite unanimously interpreted as expected. Many of the Swedish speakers interpreted the emotional expression as "disappointed", which is also the case for the Finnish speakers. It is possible that the speaker did not succeed in expressing "sadness" with the most typical Swedish prosody.

Sadness	**Swedish**	**English**	**Finnish**	**Spanish**
sad	69	100	70	91

Table 3: The interpretations of "sadness".

The emotion "anger" is also quite unanimously interpreted as expected. The Swedish and English listeners interpreted the emotion as either "angry" or "dominating" – two semantically related emotional states. The Finnish speakers were most successful.

Anger	**Swedish**	**English**	**Finnish**	**Spanish**
angry	66	50	87	78
dominating	17	25	0	9
irritated	0	0	0	9

Table 4: The interpretations of “anger”.

Finnish and Spanish on the one hand differ from Swedish and English on the other in the interpretation of the emotion “fear”.

Fear	**Swedish**	**English**	**Finnish**	**Spanish**
afraid	66	42	9	30
sad	0	0	61	48

Table 5: The interpretations of “fear”.

“Shyness” was not successfully interpreted by anybody. This might depend on the fact that “shyness” is more dependent on social context and maybe, as a consequence of this, not so clearly expressed in voice. The answers given, “sad” and “afraid”, are semantically similar. On the other hand, “sadness” and “fear” were not interpreted as “shyness” but this might be because “shyness” is a less basic emotion than the other two.

Shyness	Swedish	English	Finnish	Spanish
sad	11	8	30	30
afraid	34	8	0	17
surprised	0	0	9	0

Table 6: The interpretations of “shyness”.

As well as for the emotion “shyness” the Finnish and Spanish speakers were more deviating in interpreting the emotion “dominance”. “Dominance” or rather “resoluteness” was interpreted as “anger” by a majority of the Spanish speakers. Still “anger” is semantically and acoustically close to “dominance”.

Dominance	Swedish	English	Finnish	Spanish
angry	4	0	35	57
dominating	71	67	43	9
irritated	0	0	0	9

Table 7: The interpretations of “dominance”.

The answers for “disgust” seem to be the most haphazard. Either the speaker did not succeed in expressing the emotion, or “disgust” is a more complicated, more periferal or less basic emotion. However, as seen below, the results are better with a different semantic analysis.

Disgust	Swedish	English	Finnish	Spanish
sad	0	0	17	30
surprised	6	0	9	0
happy	6	0	0	9
irritated	0	0	9	0

Table 8: The interpretations of “disgust”.

Table 9 gives a measure of how homogeneous the answers for the different emotions were. A low % means that a high degree of miscellaneous answers were given. The emotions that had less miscellaneous answers among native Swedish listeners were “happiness”, “surprise”, “anger” and “dominance”.

Emotions	Swedish	English	Finnish	Spanish	mean%/emotion
happy	92	66	65	79	76
surprised	83	42	83	78	72
sad	69	100	70	91	83
angry	83	75	87	96	85
afraid	66	42	70	78	64
shy	45	16	39	47	37
dominant	81	67	78	79	76
disgusted	12	0	35	39	22

Table 9: Percent answers for each emotion for the different languages and mean percent answers for each emotion for all languages.

The mean percent answers within main categories for each emotion shows that, for all four languages, “anger” and “sadness” are the most easily interpreted, while “shyness” and “disgust” are the least interpretable.

3. 2 Semantic analysis

As we saw above, each emotion was not interpreted as the intention was, but they were often interpreted as an emotion which was, in some way, semantically similar to the intended, as when “shyness” was interpreted as “sadness” or “fear”. For this reason it is interesting to study the intended and interpreted emotions semantically. The basis for the analysis is a comparison of similarities and differences, between the terms in question, which could be relevant from an emotional perspective; this means that the analys of meaning of the terms is only partial. A convenient way of comparing terms within a semantic field is to find some of the properties and dimensions which both unite and distinguish between the terms in question. In the literature there are several suggestions; Wundt (1896) suggested 3 basic dimensions: lust – non-lust, relaxation – tension and calm – excitation, and Schlosberg (1954) who suggesed lust – non-lust, attention – rejection and sleep – tension. Osgood (1966) suggested degree of lust, degree of activation and degree of control which are similar to what he suggested in his semantic differential: evaluation, activity and strength. Frijda (1970) adds another dimension self assured – insecure. Roseman (1979) suggests the following 5 dimensions need, occurence of a certain state, probability, type of cause and legitimacy. In the present study we have chosen the dimensions lust – non-lust, active – passive, secure – insecure. We use these dimensions in a binary way, but a more thorough analysis would of course require a finer grading. These three dimensions we do not use maximally for each term, only where they seem to be appropriate. “Happiness” is analysed as +lust, +activity; “surprise” as +lust, + activity, – secure; “sadness” as –lust, – activity; “fear” as –lust, –security; “shyness” as –lust, –activity and –security, “anger” as –lust, +activity, +security; “dominance” as +activity, +security and “disgust” as –lust, +security.

Happiness

“Happiness” is thus analysed as +lust, +activity. In other words, the relevant meaning here, of “happiness” is lust-filled activity. This analysis connects “happiness” with 3 of the answers (miscellaneous answers also considered): “surprised”, “excited” and “drunk”. If we look at the component +lust, we also get an association to “content”. Accepting this analysis the following proportion of answers given in the tables above associated with “happiness” will be: for Swedish 92%, English 66%, Finnish 65% and Spanish 79%. What we see is that with this analysis the interpretations of “happiness” are more successful. This goes for all the emotions of this study: the classification of the answers by using the three semantic dimensions gives a higher (or the same) percent answers in accordance with intended emotion.

Disgust

For a semantic analysis of all the terms see Abelin & Allwood (1999). In this paper, just one more should be mentioned. “Disgust” is analyzed as -lust, +secure. In the answers given the –lust factor is predominant and can easily be connected with disgust 1 : "disappointment", "dissatisfied", "bored", "irritated", "bitter", "nervous", "sad", "tired".

3.3 Discussion of the interpretations of the different language groups

The different language groups did not show the same success in interpreting the intended emotion of the speaker. The Swedish listeners performed best according to expectations for most of the emotions. An interesting result is that sadness and anger, in spite of getting approximately as high percent expected interpretation among the Swedes as the other emotions, are those emotions that are interpreted better by the other language groups (the only exception being the English interpretations of anger). The implications of this are unclear. Another question arising from the results is why not all the Swedes interpreted the emotions correctly. A possible answer to this is that all emotions are, more or less, dependent on context for their interpretation. Other possible answers are that the informants were stressed by the test situation, or they have difficulties finding the right word for their interpretation of the emotional expression or that the emotions were not unambiguously expressed. However, it must be pointed out that the Swedes, in most cases, had chosen semantically related emotions when not answering as expected. It is also reasonable to suppose an interaction between prosody and non-verbal signals such as facial expressions (cf. Massaro, 2000); this is however not investigated in the present study. It is not to be expected that all the information about the emotion is present in the acoustic signal. The proportion between the information conveyed by the acoustic signal and the information in the visual signal might even differ between different emotions, and of course the proportions can also vary between different languages. On the other hand, a part of the non-verbal signal, namely the smile, is also coded in the acoustic signal (cf. Aubergé and Lemaître, 2000). Yet another type of interaction to be considered is that between prosodic and lexical information. Dubost and Tzu-ting (1999) showed that the perception of emotions in Mandarin was more difficult if a specific lexical marker was missing. In French, however, lexical markers seemed to be optional while intonation was decisive for interpretation of emotional expressions.

3.4 Acoustic analysis

The recordings of the eight emotions were studied in an oscillogram. The durations of the utterances were measured with and without the pauses between the noun phrases. The intensity of the registrations of the different utterances were graded in relation to each other. The duration of raised intensity was not measured, only the dominating peak. Intonation was also estimated with respect to if F0 is steady, rises or falls. F0-levels of the different utterances were also graded in relation to each other.

3.4.1 F0-variation

"Happiness", "fear", "shyness" and to some extent "sadness" show similarities. F0 is even and quite high, in relation to the other emotions. "surprise", "anger", and "dominance" have a strongly varying F0. ("Anger" and "dominance" differ from "surprise" mainly in being shorter.)

3.4.2 Intensity

The emotional expressions with the overall highest intensity levels are "anger", "surprise", "disgust" and "dominance". The weakest ones are "sadness" and "shyness".

3.4.3 Duration

The emotions group themselves a little differently depending on whether the length of the pauses between the noun phrases are counted or not. The emotions with the longest durations (without pauses) are "happiness", "disgust" and "surprise". Thereafter come "sadness", "fear" and "shyness". "Anger" and "dominance" have the longest durations. Expression of "shyness" and "sadness" make use of long pauses between noun phrases.

3.4.4. Summary of acoustic analysis

There are then, acoustically, similarities between certain expressions of emotions. "Anger" and "dominance" resemble each other in having short durations and strong intensity. "Anger" and "dominance" have the semantic analysis +activity, +security. "Fear" and "shyness" have medium duration, weak or medium intensity and F0-variation, and have the semantic analysis -lust, -security. The emotions that have been confused are similar acoustically and semantically. "Anger" and "dominance" are often confused for each other. This is also the case with fear, "sadness" and "shyness". Thus, similarity of semantic dimensions (at successful production and interpretation) is coupled with similarity of acoustic dimensions. The acoustic dimensions are independent from each other but also seem to co-occur, which could be a result of e.g. general excitation of the speech apparatus in certain emotional states. Another aspect on how acoustics correlate with emotions is brought up by e.g. Montero, Gutiérrez-Arriola, Colás, Enríquez and Pardo (1999), who showed that, for Spanish, the contribution of prosody to the recognition of the uttered emotion greatly varies, with "sadness" and "surprise" being more supra segmental, and "joy" and cold "anger" being more marked by segmental voice characteristics.

4. Conclusion

The conclusions are, that the Swedish listeners generally performed as expected to a higher degree than the listeners of the other language groups. Furthermore, the listeners with other first languages showed variation in the interpretations of the different emotional expressions; no one language showed a consistently different performance for all the emotions, which could be the case if Swedish grammatical prosody had been interfering. It is not excluded that certain emotions, e.g. "sadness" and "anger" can be expressed – and interpreted – in similar ways in different languages while other emotions, e.g. "joy" are, at least in some manifestations of these emotions, more language specific. It is also probable that there is, in natural speech, an interaction between prosody, non-verbal communication, lexico-grammatical information and different contextual factors.

References

Abelin, Å., Allwood, J., 1999, *Tolkning av svensk känsloprosodi – en kontrastiv studie*, Dept. of Linguistics, Göteborg (ms)

Chung, S.–J., 1999. 'Vocal expression and perception of emotion in Korean'. In*: Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, pp. 969–972

Davitz & Davitz, 1959, 'The communciation of feelings by content free speech'. *Journal of communication,* 9, 6-13, 1959

Frijda, N.H., 1970, 'Emotion and Recognition of Emotion' In M. Arnold (ed) *Feelings and Emotions*. New York. Academic Press, New York.

Hayashi, Y., 1999, 'Recognition of vocal expression of emotions in Japanese: using the interjection eh' In: *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, pp. 2355–2358

Izard, C., 1971, *The Face of Emotion*, Appleton Century Crofts, New York.

Osgood, C.E., 1966, 'The Semantic Differential Technique in the Comparative Study of Culture'. *American Anthropologist*, 66, 171-200.

Montero, J.M., Gutiérrez-Arriola, J., Colás, J., Enríquez, E. and Pardo, J.M., 1999, 'Analysis and Modeling of Emotional Speech in Spanish' In: *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, pp. 957–960

Paeschke, Kienast, A., M., Sendlmeier W.F., 1999, 'F0-contours in emotional speech' In*: Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, pp. 929–932

Roseman, I., 1979, *Cognitive Aspects of Emotion and Emotional Behavior*. Paper read at the 87th Annual convention of the American Psychological Association. Sept. 4. 1979. New York, unpublished manuscript from the Department of Psychology, Yale University.

Scherer, K. R., Banse. R., Wallbott, H. G., 2001, 'Emotion inferences from vocal expression correlate across languages and cultures' *Journal of Cross-Cultural Psychology*, 32, 76–92.

Schlosberg, H., 1954, 'Three Dimensions of Emotion' *Psychological Review*, 61, 81-88.

Williams, U., & Stevens, K.N., 1972, 'Emotions and speech: some acoustical correlates'. *JASA*, 52, 1238–1250.

Woodworth, R.S., 1938, *Experimental Psychology*. Holt, New York.

Wundt, W., 1896, *Grundriss de psychologie*. W. Englemann, Tyskland.