DIMENSIONS OF EMOTIONAL MEANING IN SPEECH

Cécile Pereira

Speech, Hearing and Language Research Centre,
Macquarie University, NSW 2109, Australia
cpereira@ling.mq.edu.au

ABSTRACT

A number of psychologists have conceptualized the communication of affect as three-dimensional (e.g. Davitz, 1964). The three dimensions proposed are arousal, pleasure and power. This paper reports the findings of a perception study where 31 normally-hearing subjects rated utterances said in the emotion of happiness, sadness, two forms of anger (hot and cold) and a neutral state on the dimensional scales of arousal, pleasure and power. Findings show that the concept of the dimensions of emotion is useful to describe and distinguish emotions; and that emotions with a similar level of arousal, and sometimes a similar level of power, share acoustic characteristics in terms of F0 range and mean, and particularly intensity mean. It is suggested that this contributes to perceived similarity between emotions, and consequently confusions, especially in the hearing-impaired.

1. INTRODUCTION

According to Osgood, Suci, and Tannenbaum's theory (1957) [1] and subsequent research (e.g., Davitz, 1964; Mehrabian and Russel, 1974) [2] [3], the communication of affect can be seen as having three major dimensions of connotative meaning, *arousal*, *pleasure* and *power*. The arousal dimension refers to the degree of intensity of the affect and ranges "from sleep to frantic excitement" (Pittam, 1994) [4]; this is also referred to as the activity, activation, intensitive or intensity dimension. The pleasure dimension refers to how positive or negative the affect is; this is also referred to as the evaluative, pleasantness, valence, valency, hedonic valence, or evaluation dimension. The power or potency dimension relates to the degree of power or sense of control over the affect, and helps distinguish emotions initiated by the subject from those elicited by the environment e.g., contempt versus fear; this has also been called the strength, dominance, confidence, or control dimension.

This paper reports a perceptual experiment where a group of normally-hearing subjects listened to utterances which previous listeners, normally-hearing and hearing-impaired, had listened to in order to identify the emotions [5] [6]. The new group of listeners rated the utterances they heard on the dimensional scales of emotional meaning, *arousal*, *pleasure* and *power*. The aims of this experiment were to describe the emotions of cold anger, hot anger, happiness, neutrality and sadness in terms of their dimensional ratings, to relate these ratings to the acoustic analysis [7] and to try to explain some of the confusions between emotions which previous listeners, including hearing-impaired ones, had made.

2. METHODOLOGY

The listening tape consisted of 40 sentences – two different productions (hereafter referred to as Token 1 and Token 2) of two utterances ("I'm going to move house" and "Two thousand five hundred and ten"), said by two actors in the emotions of happiness, sadness, two forms of anger (hot and cold) and a neutral state. There were 2 repetitions of these 40 utterances, thus creating 80 presentations. 31 normally-hearing listeners (18 men, 13 women) listened to the tape and rated each presentation on six Likert scales, adapted from Mehrabian and Russel (1974) [2]. Each of the three dimensions was represented by two scales defined by pairs of adjectives: *pleasure* was represented by happy/unhappy, and pleased/annoyed; *arousal* by agitated/calm and excited/apathetic; *power* by powerful/powerless and dominant/submissive.

The listeners choices on the Likert scales were coded from 1 to 6, corresponding to the ascending gradation of the scale e.g. 1 for very unhappy, 2 for unhappy, 3 for slightly unhappy, 4 for slightly happy, 5 for happy, and 6 for very happy. The listeners' responses were then aggregated over the two repetitions.

In order to study the acoustic manifestation of the dimensions of arousal, pleasure and power, Pearson's linear correlation coefficients were calculated between the mean dimensional ratings of the 40 different utterances in the data set and the previously calculated acoustic values (F0 range and mean, intensity mean, and duration) [7]. (The speech had been digitized at 20 kHz, with a 16 bit resolution, giving a bandwidth of 10 kHz. The F0 values were calculated with ESPS/WAVES, using an algorithm developed by Secrest and Doddington (1983). The average amplitude was calculated from the sample amplitudes of the speech signal using the root mean square method (RMS)). To be able to comment on the errors which listeners in a previous experiment on the identification of emotion had made [5] [6], difference scores were computed between the dimension ratings of the feelings intended by the speaker and the dimension ratings of the listeners' erroneous responses, in a manner similar to that employed by Davitz (1964) [2].

3. RESULTS

3.1 Dimensional ratings of pleasure, arousal and power

As two different scales represented each of the three dimensions, paired sample correlations were calculated to check how the scales of the one dimension related to each other. Within each

pair, scales correlated strongly, and it seemed then justified to pool the results of the three pairs of scales into the three dimensions of pleasure, arousal and power. Figures 1 and 2 show the mean rating on each dimension for each emotion and each speaker.

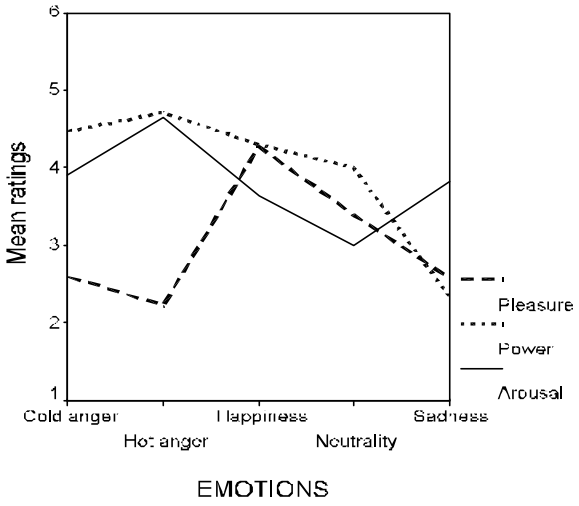

Figure 1: Dimensional ratings of the emotions expressed by the female speaker.

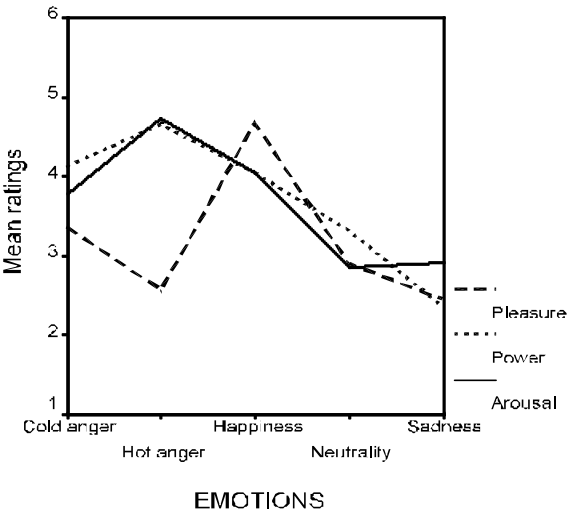

Figure 2: Dimensional ratings of the emotions expressed by the male speaker.

Looking at the dimension of arousal on these two figures, one can see that the shapes are very similar, except for sadness. The emotion rated the highest was hot anger, followed in the female speaker by cold anger, sadness, and happiness, and in the male speaker by happiness, cold anger and sadness, and ending in both speakers with neutrality.

In the dimension of pleasure the shape of the line plots are very similar for both speakers. The highest rated emotion was happiness followed by, for the male, cold anger, neutrality, hot anger/sadness and, for the female, neutrality, cold anger, sadness then hot anger.

On the dimension of power the ranking of emotions was almost the same in both speakers, i.e., in descending order, hot anger, cold anger, happiness, neutrality and sadness.

Analysis of variance showed that the emotions were significantly different from each other on each dimension for both speakers as well as for each speaker individually. Paired-sample *t*-tests on the means of each emotion per speaker per dimension showed that most of the emotions were significantly different from each other on each dimension. All emotions were significantly different from each other on at least two of the dimensions.

3.2. Dimensional ratings and acoustic characteristics

Table 1 presents the correlations for each speaker between the mean of the acoustic values (F0 range and mean, intensity mean, and duration) and the mean ratings on the three dimensions of emotion.

The dimension of arousal correlated positively and significantly with F0 range and mean, and mean intensity. This was for both female and male speakers, but especially so for the male speaker. In the male speaker, the dimension of pleasure, correlated significantly with F0 mean. The dimension of power correlated positively and significantly with F0 range and mean, and mean intensity in the male speaker. Intensity mean also correlated positively and significantly with power in the female speaker.

Table 1: *Pearson's linear correlation coefficients between judgments on dimensions and acoustic measurements*

Scale	F0 mean	F0 range	RMS mean	duration
arousal				
female	.633 **	.461*	.712 **	.101
male	.811 **	.754 **	.868 **	.268
pleasure				
female	- .170	- .019	- .322	.114
male	.501*	. 310	.418	- .035
power				
female	- .059	.358	.524*	.294
male	.686 **	.715 **	.800 **	.176

* Correlation is significant at the 0.05 level
** Correlation is significant at the 0.01 level

3.3. Dimension ratings and errors of identification

In all three groups, the difference scores over all emotions, speakers, utterances and tokens were significantly different in each dimension, with, as can be seen in Table 2, the smallest amount of difference on the arousal dimension, followed by the power dimension, and ending with the pleasure dimension. In the hearing-aid users the mean difference scores were larger than in the normally-hearing listeners, and were larger again in the cochlear implant users.

Table 2: *Difference scores (means and standard deviations) on dimensions of arousal, pleasure and power for all errors of identification of emotion*

	Mean	*SD*
arousal		
normally-hearing listeners	.51	.14
hearing-aid users	.61	.10
cochlear implant users	.67	.09
pleasure		
normally-hearing listeners	.82	.16
hearing-aid users	1.04	.12
cochlear implant users	1.20	.12
power		
normally-hearing listeners	.69	.15
hearing-aid users	.87	.16
cochlear implant users	.97	.14

In the next step of this analysis, the difference scores for the major confusions in identification of emotion which the normally-hearing listeners, the hearing-aid group and the cochlear implant group had made [5] [6] were examined. This showed that, in the majority of cases, the two emotions confused were closer on the arousal dimension than on the other two dimensions – and this was common to all listener groups. However, there were some pairs of emotions which were confused without being closer on the arousal dimension.

4. GENERAL DISCUSSION

The strong correlation of the two scales in each dimension was consistent with expectation since these were intended to represent the same dimension of emotion.

The placement of hot anger and (for the male speaker) sadness, at the positive and negative ends, respectively, of the dimension of arousal parallels Davitz's (1964)[2] results on anger and despair. The relatively high arousal rating of sadness for the female speaker seems consistent with a previous acoustic analysis [7], which reported values for the F0 mean and range which were higher than those of neutrality.

The high rating of happiness and the low rating of sadness and hot anger on the dimension of pleasure seem to match the findings of Davitz (1964) [2] and also of Scherer (1974) [8]. The relatively high rating of cold anger for the male speaker may have been induced by F0 mean and range values higher than those of neutrality.

The ranking of anger and happiness on the dimension of power concurs with Davitz's findings. Sadness in our experiment is at the negative end of the power dimension, whereas Davitz had 'despair' at the very positive end, before anger and happiness. In the absence of the definition of Davitz's despair or of his contextual paragraph for the test sentence, this is difficult to explain, since despair is usually understood as denoting a high degree of hopelessness, but relatively little power over the affect.

The fact that all emotions of the corpus were significantly different from each other on at least two of the dimensions of arousal, power and pleasure, lends support to the concept of the dimensions of emotions as a useful tool for the description of emotions.

The high, positive and significant correlation of arousal with F0 and RMS is similar to Davitz' findings of a correlation with pitch and loudness (1964) [4]. However, our results differ from Davitz in relation to duration, which in the present study did not correlate with arousal, and with regard to the power dimension, which in this research correlated positively with F0 and RMS in the male speaker, whereas in Davitz there was no such relationship.

The distance in terms of the ratings of the emotions mistaken for each other was smallest on the arousal dimension, followed by the dimensions of power then pleasure – and this was common to all listener groups. The larger distances and standard deviations in all dimensions found in the hearing-aid users and to a greater extent in the cochlear implant users are likely to be a reflection of the greater difficulty of the task for these two groups and of the greater heterogeneity of their results. Our analysis of specific confusions confirmed the results found over all emotions, but also indicated, unlike Davitz's findings, that some pairs of emotions were confused which were not closer on the arousal dimension than on the other two dimensions – although, admittedly, these were close in all dimensions, or were closer on power without being very distant in arousal.

Based on his findings that emotions mistaken for each other were closer on the arousal dimension than on either the dimension of pleasure or that of power, and that there was a significant relationship between the auditory characteristics of the emotions and their ratings on arousal, Davitz (1964) [2] concluded that emotions with a similar level of arousal sound similar in loudness, pitch, timbre and rate, and this is why listeners confuse them.

Our results support Davitz' conclusion, but this support needs to be qualified. The results of the present research suggest that emotions with a similar level of arousal, and sometimes a similar level of power, share acoustic characteristics in terms of F0 range and mean, and particularly intensity mean, and this must contribute to their perceived similarity and consequent confusion, especially in the hearing-impaired.

REFERENCES

1. Osgood, Charles E., George J. Suci, and Percy H. Tannenbaum. 1957. *The measurement of meaning.* University of Illinois Press: Urbana.

2. Davitz, Joel R. 1964. Auditory correlates of vocal expression of emotional feeling. In *The communication of emotional meaning*, ed. J.R. Davitz, 101–112. New York: McGraw-Hill.

3. Mehrabian, Albert and James Russel. 1974. *An approach to environmental psychology*. Cambridge: MIT Press.

4. Pittam, Jeffery. 1994. *Voice in social interaction.* Thousand Oaks: Sage.

5. Pereira, Cécile. 1996. Angry, happy, sad or plain neutral? The identification of vocal affect by hearing-aid users. In *Proceedings of the Sixth Australian International Conference on Speech Science and Technology*. Adelaide.

6. Pereira, Cécile. 1999. The perception of vocal affect by cochlear implantees. In *Cochlear Implants*, eds Waltzman, Susan, and Noel Cohen. New York: Thieme.

7. Pereira , Cécile, and Catherine Watson. 1998. Some acoustic characteristics of emotion. In *Proceedings of the Fifth International Conference on Spoken Language Processing.* Sydney.

8. Scherer, Klaus R. 1974. Acoustic concomitants of emotional dimensions: judging affect from synthesised tone sequence. In *Non verbal communication: Readings with commentary*, ed. E. S. Weitz, New York: Oxford University Press, 105–11.

9. Pittam, Jeff and Klaus Scherer. 1993. Vocal expression and communication of emotions. In *The handbook of emotion*, ed. M. Lewis, & J. Haviland, 185–197. New York: Guildford Press.