SOME OF THE VOCAL EMOTIONAL STATES

AN ACOUSTIC STUDY

By

Dr. Mervat Mohamed Ahmed Fashal

Phonetics Department

Faculty of Arts

1998

Contents:

1- Introduction
  1.1- Emotion Concepts.
  1.2- Classification of Emotions.

2- Studies of the Vocal Emotions.
  2.1- Experimental Methods.
  2.2- Analysis of the Parameters of Vocal Emotions.

3- The Experiment
  3.1- Subjects.
  3.2- Speech Material.
  3.3- Auditory Test.
  3.4- Acoustic Analysis.
    3.4.1- Graphic Representations
    3.4.2- Analysis Measurements.

4- Results
  4.1- Statistical Calculations of Measurements.
    4.1.1- Intensity Mean Values.
      Figures st/ unst. (P. 12)
    4.1.2- Fundamental Frequency Mean Values.
    4.1.3- Duration Mean Values.
    4.1.4- Correlation between the Parameters.
      Tables (1) (2)
    4.1.5- Fo change and Fo levels (table 3).

  4.2- Graphic Combination of Fo and Intensity Contours.
    4.2.1- Male M (Fig. 1,2).
    4.2.2- Male S (Fig. 3,4).
    4.3.3- Female G (Fig. 5,6).
    4.2.4- Female A (Fig. 7,8).

5- Discussion
  5.1- Dimensions of Auditory and Acoustic Differentiation.
  5.2- Normal.
  5.3- Angry.
  5.4- Astonishment.
  5.5- Sadness.

6- Conclusions

7- References.

Some of the Vocal Emotional States
An Acoustic Study

1- Introduction

1.1- Emotion Concepts

Emotion, in a simple definition, is the behavioral expression of feelings. Some psychiatric researchers consider emotion, affect and feelings as synonyms,[1] but they use emotion in the generic sense. Some others refer to the subjective component as affects or feelings, and to patterned behavioral expression as emotion[2].

Charles Darwins (1872) considered emotion expressions as "Vestiges of patterns of action no longer useful due to evolutionary development,", but later scientists[3] considered that Darwin's view was only a partial statement, and that emotion retains important functions.

Izard (1977) stated that a complete definition of emotion must take into account three basic processes: "The experience or conscious feeling of emotion, the processes that occur in the brain and nervous system, and the observable expressive patterns of emotion."

Emotions are different from moods,[4] emotions arise suddenly in response to particular stimuli, and last for seconds or minutes, while moods are more vague in nature, lasting for hours or days; emotions can broadly be considered as being concerned with changing something,[5] and moods with maintaining something.

---

(1) Hofling (1975).

(2) Alpert, M. & Rosen, A. (1990). - Kolb, L. & Brodie, H. (1982). - Redlich, A. and Freeman, D. (1966).

(3) Izard, C. (1971), (1977), Scherer (1981).

(4) Oatlay, K. (1989).

(5) Scherer, K. (1984), (1986).

1.2- Classification of Emotions

Many emotion theories use the concept of "basic emotions" and "nonbasic emotions". Ortony and Turner (1990), noted that basic emotions vary between 2 and 18 emotions, they also suggest that four of them are the most commonly accepted basic emotions which are: happiness, sadness, anger, and fear.

Perceptual studies[6] differentiated between facial emotion and vocal emotion, each of them has its own dimensions.

Kramer (1963) reviewed a number of studies which have demonstrated that various aspects of a speaker's physical and emotional state including age, sex, appearance, intellegence and personality can be identified by voice alone. The emotional state of a speaker thus produces an emotion "carrier wave" for the words spoken, but "It is not what he said but the way that he said it."[7]

2- Studies of the Vocal Emotions

2.1- Experimental Methods

In order to analyse features of the emotional voice, numerous differing methods have been reported in the literature, most of which have involved direct human analysis and most of the authors have used untrained speakers. Speech samples, however, become more natural if obtained from actors.[8]

The material read for analysis is also varied, three major experimental techniques have been used in the literature to assess the emotion content of speech:

1) Meaningless content, having speakers express emotions while reading letters of the alphabet[9].
2) Constant content, comparison of the same sentence given by speakers expressing different emotions[10].
3) Content - Free, electronic filtering of speech[11].

(6) Van Bezooijen, R., Otto, S. and Heenanen, T. (1983).
(7) Ortony, A. and Turner, T. (1990).
(8) Williams, C. and Stevens, K. (1972).
(9) Davitz, J. and Daivtz, L. (1959), Laukkannen, A. et al. (1996).
(10) Leinonen, L. & Hiltunen, T (1997), Fairbanks, G. & Hoaglin, L. (1941).
(11) Van Bezooijen, R. & Boves, L. (1986), Murray, I. and Arnott, (1993)

2.2- Analysis of the Parameters of Vocal Emotions

The most commonly referenced vocal parameters in the emotion literature are pitch (mean Fo, Fo change, Fo level and Fo contour), intensity, duration, tempo and voice quality. Acoustic differences of most of these parameters were examined in recent study by Leinonen (1997) using a data of one word, the name "Saara", which is expressed in 10 emotional connotations[12].

Pitch was also found to be an important emotional factor by Levin and Lord (1975) who noted that "an analysis of upper octave range measurements of Fo is sufficient to provide an indication of emotional changes". While Uldall (1960) concluded that the best indicator of the emotional content of an utterance was the pitch contour.

Also Scherer and Scherer (1981, P. 125) noted various vocal indicators of emotional expression, principally pitch, intensity and voice quality. They noted in recent work (1986, P. 145) that while pitch is important in emotion, voice quality is more important in differentiating discrete emotion. However, Ladd et al. (1985) suggested that the level of fundamental frequency, intonation contour and voice quality vary independently of each other.

Another research does indicate a possible correlation between speech rate and emotion[13]. Fónagy (1981) noted increased intensity leading to the shortening of consonants, liquids and nasals, and to the lengthening of vowels. Black (1961, P. 199) also related speech rate with emotion concentrating on the importance of intensiy, he noted from experiments that speech intensity increases along with pitch, and that soft speech was characterized by slow rate.

Davitz (1964) has divided emotions into "active" and "passive" groups, characterizing passive emotions as those with slow speech rate, low volume, low pitch, and more resonant

(12) "naming, commanding, angry, frightened, pleading, astonished, satisfied, admiring, scornful, and sad"

(13) Huttar, G. (1968)

timbre (voice quality), and characterizing active emotions by their fast speech rate, high loudness, high pitch, and "blazing" timbre.

In the communication of emotions, recent studies[14] have shown the importance of glottal excitation of waveform in conveying emotion. Ann-Laukkanen et al. (1996) aimed at elucidating the variation in the glottal waveform related to the the variation of other physical variables of vocal expression of emotion, using data of meaningless letters "paà pa" (the first syllable is stressed), which is repeated in five emotional states "neutral, surprise, enthusiastic, sadness and anger", resulting that stress production and vocal expression of emotions bring about simultaneous changes in many variables including Fo, SPL, subglottal pressure and glottal airflow waveform.

The studies interested in the voice quality consider the vocal fold vibration and the glottal flow waveform as minimal parameters to convery the emotions[15] and also to differentiate between the male and female glottal wave variations[16].

Cummings and Clements (1995) considered that the glottal excitation waveform is well known to be an important method of conveying prosodic information in speech. They used the parameters of the glottal excitation which are: the ratios of closing and opening slopes and durations of the glottal waveforms, rather than the other parameters of Fo, intensity, duration and intonation as prosodic factors. Their data was: 11 styles × 2 male speakers × 2 vowels, they found that from styled speech to normal speech glottal waveforms themselves change; in addition to changing in pitch, pitch contour, duration and intensity.

Childers & Lee (1991) studies of the manifestation of voice quality in glottal flow waveforms have shown that pressed phonation seems to bring about a sharper pulse with long phase and low flow amplitude, whereas breathy phonation is characterized with symmetrical pulse with a short or no closed phase.

(14) Cummings and Clements (1990).

(15) Cummings and Clements (1995).

(16) Monsen & Engebretson (1977).

The sex of the speaker can cause significant differences in some speech emotional parameters. Abe (1980) noted that the difference between the pitches and pitch ranges of male and female voices is even more extreme if the speakers compared are excited. Many of these effects can be attributed directly to anatomical differences between the sexes. He also noted that tiredness (classed as an emotion by some authors) cause a loss of elasticity of articulatory muscles leading to a lowering of the voice and narrowing of the pitch range.

Eva Holmberg (1988)'s study provides new insight into functional differences between male and female voices and across intensity conditions (normal, soft and loud voices). Its results show a significant increase of pressure from soft to normal, and normal to loud voice for both males and females, also there is no significant male - female difference in pressure in any of the loudness conditions. Sharing studies in this concern have been made of how such a waveform varies with intensity, and of a waveform differences between male and female speakers (Monsen 1977, Glave, R. and Rietveld, A. 1975), Their results elucidated that male waveforms have been found more asymmetrical, with the closing portion shorter than the opening portion, while female waveforms have shown more equal duration of opening and closing portions.

Monsen and Engebretson (1977)'s experiment examined the glottal volume - velocity waveform of normal, loud and soft voice; falsetto and creaky voice (which is also called vocal fry). 5 males and 5 females were asked to produce and elongated neutral schwa in the sequence Λ-∂-∂ with changing stress position. Analysis of the collected data shows that the glottal sound source of the normal adult varies within a wide range, not only as to Fo and rms intensity but also as to the appearance and shape of the waveform, the phase spectrum and the intensity spectrum. This variation is systematic and may be accounted for in most part by taking into consideration 3 interacting factors: a) the sex of the speaker, b) the vocal mode or register in which the speaker phonates, and c) the linguistic function which the phonation is intended to fulfill

3- The Experiment

3.1- Subjects

4 adult speakers (2 males and 2 females) with no history of speech voice or hearing problems served as subjects. Their age ranged from 30 - 42 years.

3.2- Speech Material

The speech material consisted of one sentence [ʔàna - ʔànaʃi ʔoltkeda] with a repitition of the first word which is emphasized. The subjects were asked to produce the sentence in four vocal emotions: Normal, Sad, Astonished and Angry. They were asked also to repeat them about 20 times. For the productions to be as similar as possible to the subjects' natural speech, they were free to use a comfortable level for each condition, so that they are not given any instructions in this concern.

3.3- Auditory Test

10 graduate students judged the recorded data. They listened in small groups of 3 or 4 subjects at a suitable distance of the taperecorder, that the weakest utterances were audible to everybody. The judges were asked to choose the best 10 repititions of each emotional condition, which convey the intended connotations.

The connotations: "Astonished" and "Angry" were conveyed best, but "Sad" and "Normal" intents were somewhat confused by listeners. The analysis dealt with the best 10 repititions of judgements i.e which are agreed on by about 80% of the judges.

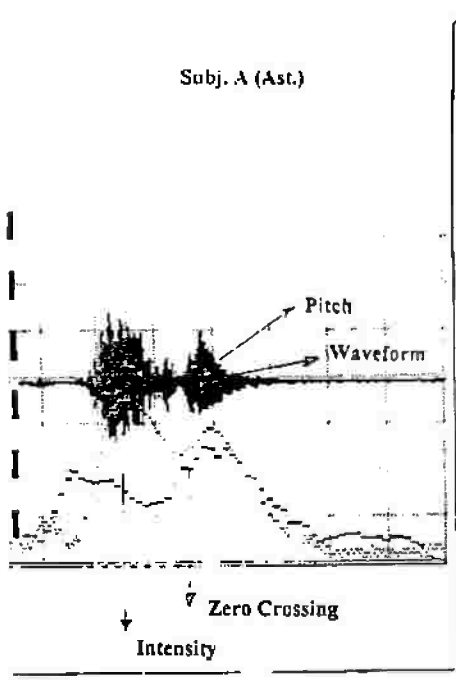
3.4- Acoustic Analysis

3.4.1- Graphic Representations

The instruments which are used in the acoustic analysis were; Kay Elemenrric Corp. Model 5500 DSP, connected with Computerized Speech Lab. (CSL) Model 4300. The representation of the graphs were traced using DSP 5500 which are shown as a combination of:

- Waveform
- Intensity Contour → with amplitude range: 72 dB
  " analysis attenuation: 5 dB
- Fo Contour → Frequency range: 4 Khz
- Time axis → Horizontally Time axis scale: 2 sec
  each square = 100 msec
- Zero Crossing

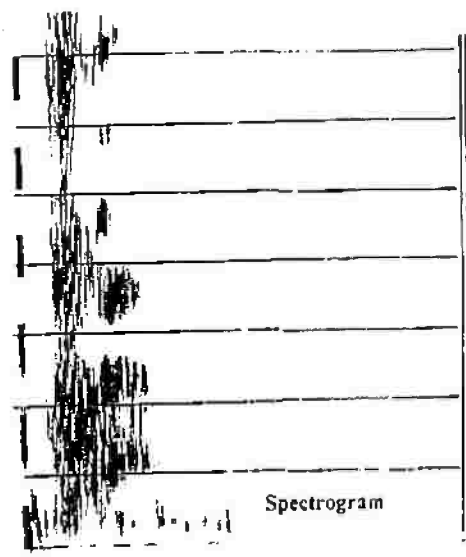
These analyses combination are included in the upper screen (channel 2), while the lower screen (channel 1) includes Sonagram of the same signal. (See for example the following graphic analysis for female A, Astonish)



KAY ELEMETRICS CORP. MODEL 5500
SIGNAL ANALYSIS WORKSTATION

Date: 3F 19FF Rec Off Time:12
Analysis by:

INPUT SETTINGS	Channel 2
Source	LEFT CONNECTORS
Frequency Range	DC - 4 KHz.
Input Shaping	FLAT
Buffer Size	4 SECONDS
ANALYSIS SETTINGS	Upper Screen
Signal Analyzed	CHANNEL 2
Analysis Format	COMBINATION
Transform Size	1024 pts. (15 Hz)
Time Axis	100ms (2sec)
Frequency Axis	FULL SCALE
Analysis Window	HAMMING
Averaging Set Up	NO AVERAGING
DISPLAY SETTINGS	Upper Screen
Freq. Divisions	0.000 Hz.
Dynamic Range	72 dB
Analysis Atten.	5 dB
Set Up Options Set	



INPUT SETTINGS	Channel 1
Source	LEFT CONNECTORS
Frequency Range	DC - 8 KHz.
Input Shaping	HI-SHAPE
Buffer Size	4.0 SECONDS
ANALYSIS SETTINGS	Lower Screen
Signal Analyzed	CHANNEL 1
Analysis Format	SPECTROGRAPHIC
Transform Size	100 pts. (300 Hz)
Time Axis	100ms (2sec)
Frequency Axis	FULL SCALE
Analysis Window	HAMMING
Averaging Set Up	NO AVERAGING
DISPLAY SETTINGS	Lower Screen
Freq. Divisions	1000. Hz.
Dynamic Range	42 dB
Analysis Atten	20 dB
Set Up Options Set	to: # 05

3.4.2- Analysis Measurements

The following measurements were excuted from the recorded and digitized signals using the Computerized Speech Lab. (CSL) Model 4300 programs of numerical analysis and statistical analysis:

1- mean intensity in (dB), which is got from "Energy Statistics" of the stressed vowel [$à_1$] and unstressed [$à_2$] of the word [?àna].

2- mean Fo in (Hz), which is got from "Pitch Statistics" of [$à_1$] and [a_2].

3- mean duration in (msec) from "Time Domain" of each vowel after segmentation.

4- Fo pitch analysis from "Impulse Markers" of each period of the waveform of the sentence [?ànalli? oltkeda] and then from the "numerical pitch analysis" which gives the analysed data range. Fo change, and Fo level is derived from all of them.

All the above analyses were measured for 4 speakers × (= multiplied by) emotional conditions × 10 selected productions.

Analysis concentrated on two syllables taken to represent stress production of the first emphasized word [?$à_1$na$_2$] in the utterance [?àna-?analli?oltkeda].

4- Results

4.1- Statistical Calculations of Measurements

Summary statistics were calculated for each of the measured parameters in the three loudness conditions (Sad, Normal, Astonished/ Angry), for males and females separately see (table 1 and table 2). Statistical analysis was performed to test the physical variables of vocal expression of emotions: male/ female differences, stressed/ unstressed syllables, and the correlation between the acoustic factors of intensity, duration and Fo.

4.1.1- Intensity Mean Values

a- The ranges of intensity in respect to the different cases in both males and females were as follows:

Sad range from	[51 dB : 59 dB]
Nor.	[53 dB : 63 dB]
Ast.	[65 dB : 70 dB]
Ang.	[70 dB : 72 dB]

From the intensity values, it is observed that regular tendency in loudness of stressed vowels increased from Sad, Normal, Astonished to the highest intensity in Angry case.

b- Male/ Female Comparison of the intensity of the different cases showed slight value differences between males and females in the cases of Sad and Normal, but intensity values seem to be similar in the loud voice (astonished/ angry).

c- Stressed/ Unstressed Vowels indicated that Sad and Normal were similar in the values of dB, to some extent, while they showed significant differences in Astonishment and Anger (see tables 1.2) (Figures 1: 8).

In Anger cases intensity values of the stressed syllables were overall the loudest, while in Astonishment stressed and unstressed values were sometimes equal. (Male S. Ast. and Fem. G. Ast).

4.1.2- Fundamental Frequency Mean Values

a- Fundamental Frequency during stressed/ unstressed vowels showed overall rise in Fo during the stressed vowel versus unstressed. There is one exception, however, in only two cases Male S and Male M, Ast.) which showed that Fo of unstressed vowel was slightly higher than stressed vowel.

b- Male/ Female Mean Fo comparison showed an overall rise in Female's Mean Fo than Male's:

Sad $\bar{x}$ Fo in Females was 186 Hz while in Males was 127 Hz.

Normal $\bar{x}$ Fo in Females was 221 Hz while in Males was 158 Hz.

Astonish $\bar{x}$ Fo in Females was 250 Hz while in Males was 236 Hz.

Angry $\bar{x}$ Fo in Females was 256 Hz while in Males was 234 Hz.

These results proved that Male/ Female differences in Fo were extremely significant, especially in the Sad and Normal cases, while in loud vocal emotions "Astonish and Angry", Fo values seem to be similar.

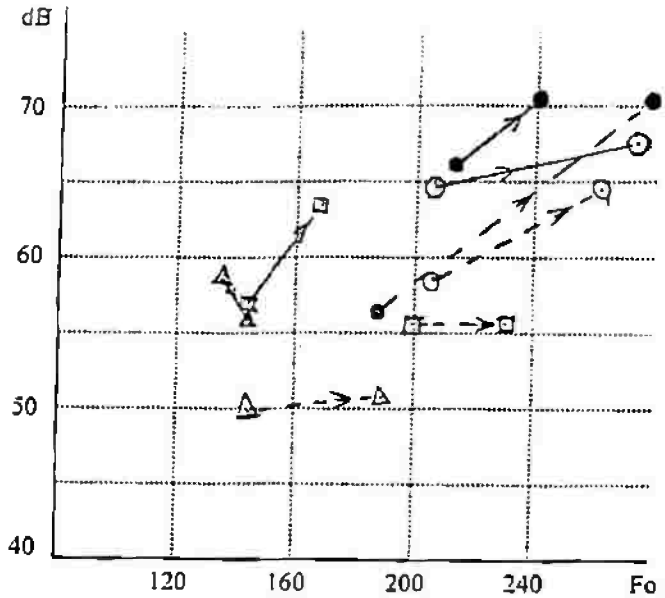


Fig (1) - - - - - Fem. A
——— Male M

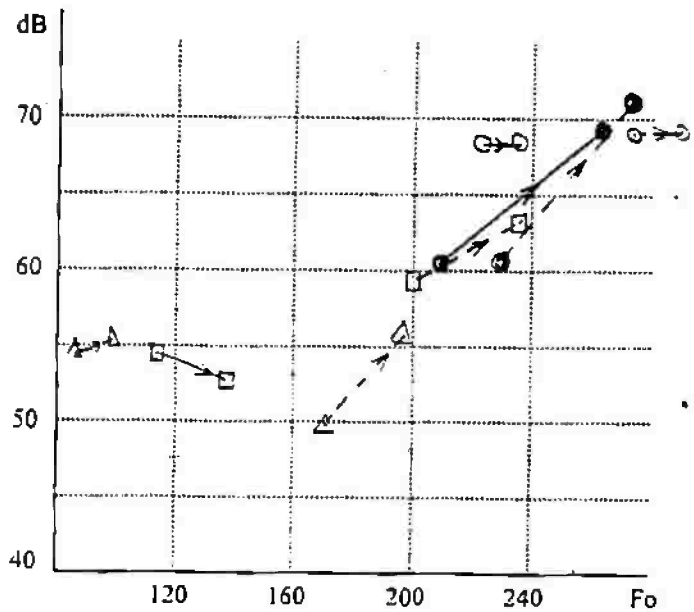


Fig (2) - - - - - Fem. G
——— Male S

□ = Nor. ▽ = Sad
○ = Ast. • = Ang.

Arrows point from unstressed to stressed values, through the mean values of Fo and of dB.

Statistical Calculations

Subjects			*Phonation Modes of Emotions*			
			Normal	Sad	Astonish	Angry
Males	M	$\bar{x}$ Hz	170.5 Hz	142	255	224
		$\bar{x}$ dB	63.8 dB	59	68.5	70.5
		$\bar{x}$ ms	88. ms	75	96	107
	S	$\bar{x}$ Hz	146	112	222	245
		$\bar{x}$ dB	53	56	96.8	70
		$\bar{x}$ ms	73	69	75	99
Females	G	$\bar{x}$ Hz	222	187	259	257
		$\bar{x}$ dB	64.6	56.6	70	72.6
		$\bar{x}$ ms	74	64	67	74
	A	$\bar{x}$ Hz	220	185	242	255
		$\bar{x}$ dB	56	51	65.5	70.8
		$\bar{x}$ ms	64	43	79.8	100

table (1) Mean values of Fo. dB and duration of the stressed vowel [à] in the word [ʔàna] of the four modes of emotions (Normal, Sad, Astonish and Angry). 2 male and 2 Female subjects.

Subjects			*Phonation Modes of Emotions*			
			Normal	Sad	Astonish	Angry
Males	M	$\bar{x}$ Hz	150	150	200	203
		$\bar{x}$ dB	57	57	65	66.8
		$\bar{x}$ ms	70	70	70	80
	S	$\bar{x}$ Hz	123	132	215	205
		$\bar{x}$ dB	55	55	69.7	61
		$\bar{x}$ ms	62	63	88	52
Females	G	$\bar{x}$ Hz	198	172	263	220
		$\bar{x}$ dB	60	50	70	61
		$\bar{x}$ ms	45	84	77	88
	A	$\bar{x}$ Hz	187	102	200	173
		$\bar{x}$ dB	56	50	59	56
		$\bar{x}$ ms	57	40	70	86

table (2) Mean values of Fo dB and duration of the unstressed vowel of the word [ʔàna].

4.1.3- Duration mean Values

a- The ranges of duration in respect to the different cases of both Males and Females were as follows:

- Sad	[43 msec	- 75 ms]	with	the	average 61 ms
- Normal	[64 ms	- 88 ms]	"	"	" 74 ms
- Astonish	[67 ms	- 96 ms]	"	"	" 80 ms
- Angry	[74 ms	- 107 ms]	"	"	" 95 ms

The results showed regular increase in the stressed vowel length according to the soft, normal and loud (Ast. and Angry) voices.

b- Stressed/ Unstressed vowel duration

Stressed vowels indicated longer duration than unstressed in most cases, while in a few they showed that unstressd vowels were longer than stressed (Male S, Ast.) and (Fem. G, Ast., Sad and Angry) (table 1,2).

c- Male/ Female differences.

Males showed longer duration in all cases than Females, in stressed and unstressed syllables.

4.1.4- Correlation Between the Parameters

The results showed systematic correlation between the values of $\bar{x}$ Fo, $\bar{x}$ dB and $\bar{x}$ ms according to the soft, normal and loud voices in males and females i.e. increasing the energy of the phonation mode increased Fo, intensity and the duration of the vowels.

Sad	$\bar{x}$ Fo = 156 Hz	$\bar{x}$ dB = 56 dB	$\bar{x}$ ms = 61 ms
Normal	$\bar{x}$ Fo = 187 Hz	$\bar{x}$ dB = 60 dB	$\bar{x}$ ms = 74 ms
Astonish/ Angry	$\bar{x}$ Fo = 244 Hz	$\bar{x}$ dB = 70 dB	$\bar{x}$ ms = 89 ms

4.1.5- Fo Change and Fo Level Among the Emotions

Through the production of the sentence [?analli? oltkeda] Fo is measured. Among stressed and unstressed vowels in the sentence there are low values and high values of Fo. From all the emotional cases, an extraction of the lowest and highest Fo is made, and by statistical calculations of these values, Fo change

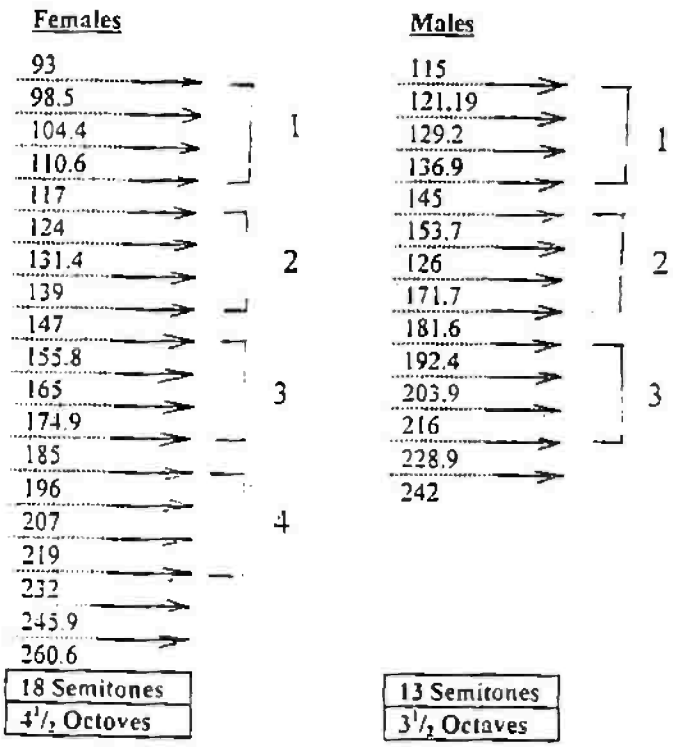
and Fo level is accounted. This was done is order to recognize the melodic range in which the females and males moves.

For females:	Fo change	=	263 - 93 = 170 Hz
	Fo level	=	18 semitones
For males:	Fo change	=	245 - 115 = 130 Hz
	Fo level	=	13 semitones

Table (3) indicated the difference between the voice range of females and males. From the calculations, females move from soft to loud emotion in a wider range than males about I octave. Also Fo change through the emotions of females were higher than males about 40 Hz.

Table (3)

Fo Level of Females and Males



4.2- Graphic Combination of Acoustic Parameters

4.2.1- Pitch Contour

a) Normal

Normal pitch almost appears on the stressed syllable, pitch level in normal has, as expected, a median level see: (Fig 8 Fem. A, Fig 6 Fem. G) - in a few cases - has slightly lower than the median e.g. (Fig 2 Male M) and even has very low and it is similar to "Sad" see (Fig 4 Male S).

b) Sad

Pitch level in "Sad" is generally low (see Fig 1 Male M. Fig 3 Male S and Fig 7 Fem. A). In some cases, however, Fo in sad is weak to the extent that pitch contour totally disappears see (Fig 5 Fem. G).

c) Angry

Pitch contour of Anger has a rigid and straight rising - in various levels - in all cases. In Female G and A it is at very high level and in Males M and S is at high level respectively, see Angry Figures- It is an abrupt rising on the stressed syllable i.e the curve form of:

d) Astonish

Pitch contour has different shapes which are quite different from Angry. They are summarized as follows:

1- → $\grave{a}_1$ straight in mid level, ↑ a_2 rising at high level
(Fig 4, Male S)

2- ↑ $\grave{a}_1$ rising in mid level, ↑ a_2 rising at high level
(Fig 4, Male M)

3- ↑ $\grave{a}_1$ rising at high level, ↑ a_2 rising at high level
(Fig 6, Fem. G)

4- ↑ $\grave{a}_1$ rising at high level, → a_2 straight in mid level
(Fig 8, Fem. A)

4.2.2 Intensity Curve

a) Normal

Intensity curve has a median level in all cases, and in both males and females see (Fig 2 Male M, and Fig 4 Fem. G).

b) Sad

Intensity curve always has a lower level than Normal, but sometimes it is slight rise to be similar to Normal e.g in (Male S and Fem; A).

c) Astonish/ Angry

In loud voice i.e. "Astonish" and "Angry", intensity contour was very loud on the first stressed syllable, while it was slightly lower on the second unstressed one. Sometimes the stressed and unstressed syllables seem to be roughly similar except for a small difference in loudness see (Male S, Ast. and Male M, Ast.). In a few exceptional cases, the unstressed syllable was louder than the stressed see (Fem. G, Ast.).

4.2.3 Correlation of Pitch and Intensity

In loud voice the increase in loudness of intensity contour is accompanied by rising of pitch contour, but this is not necessarily the case in low and normal voice. However, pitch contour level is sometimes lower than intensity contour level see (Male S. Sad and Nor., Male M. sad).

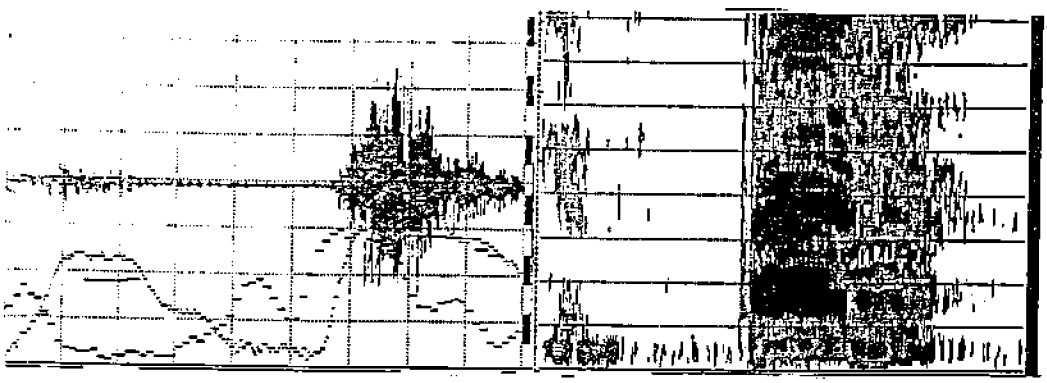


Fig. (1) Male. M (Sad, Ang.)

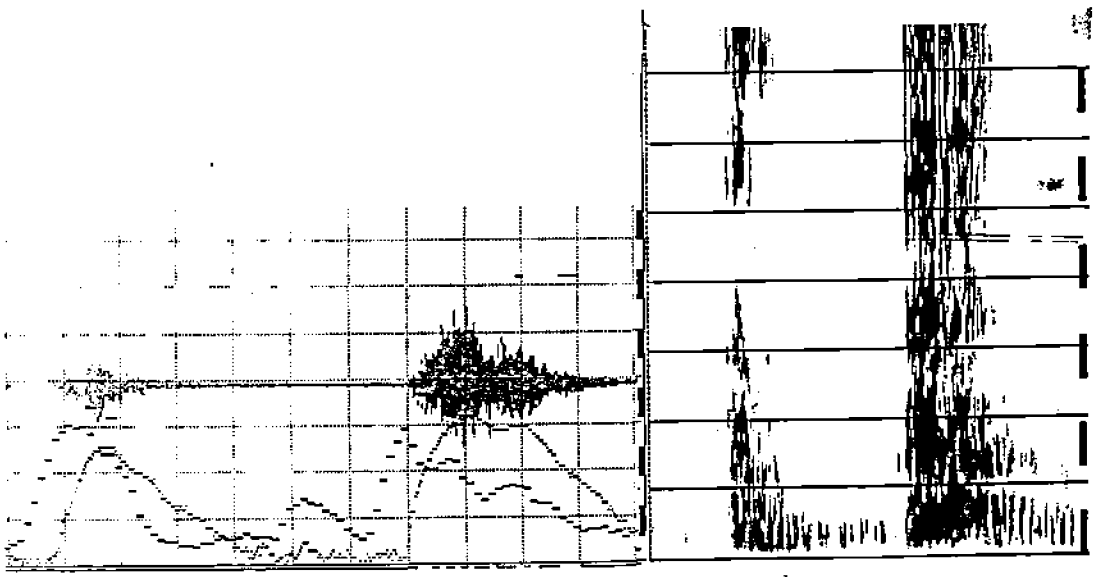


Fig. (2) Male. M (Nor., Ast.)

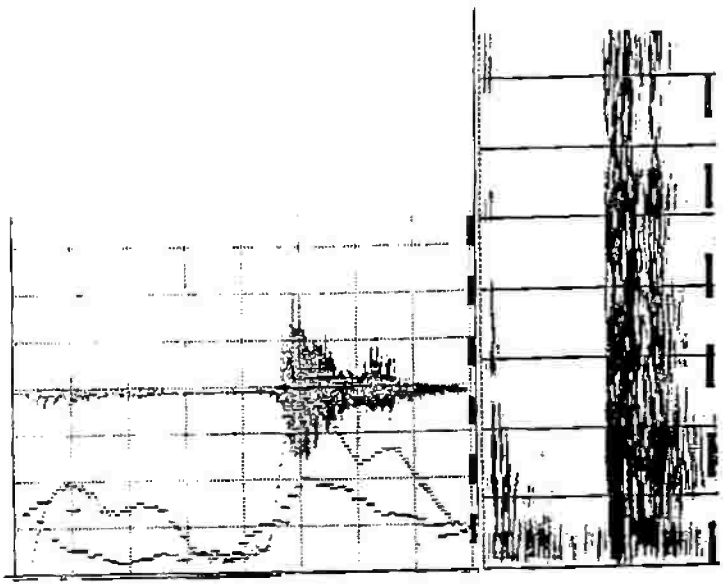


Fig. (3) Male. S (Sad, Ang.)

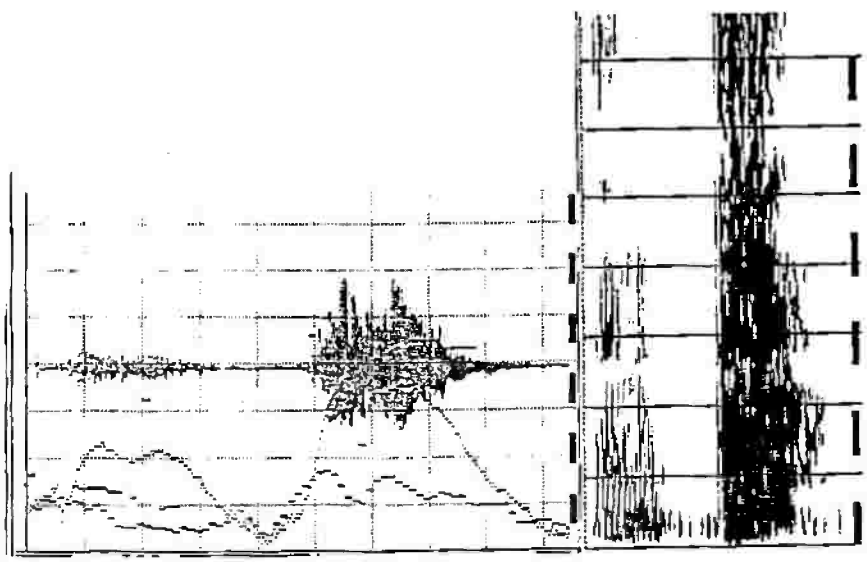


Fig. (4) Male. S (Nor., Ast.)

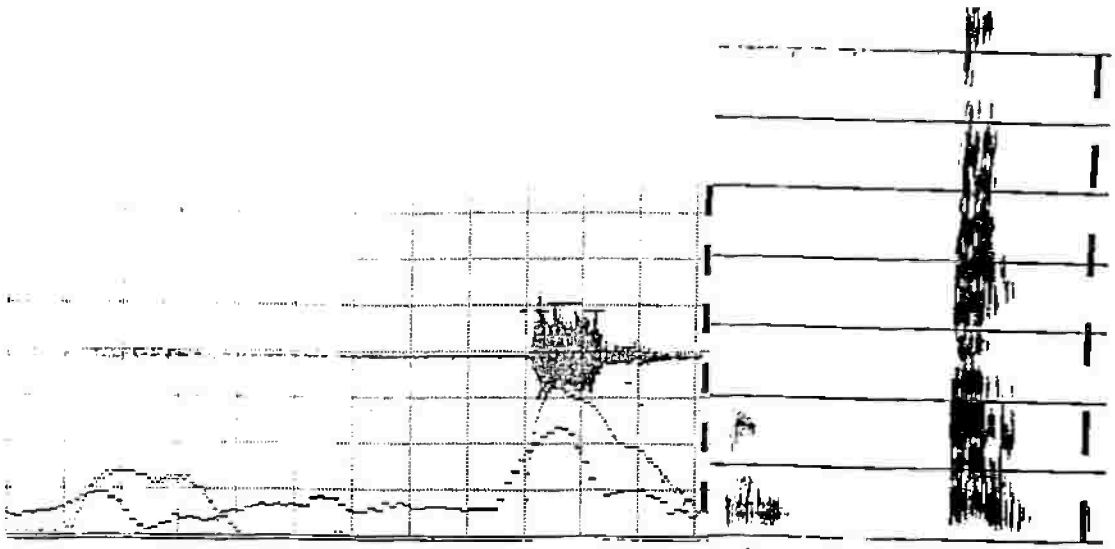


Fig. (5) Fem. G (Sad, Ang.)

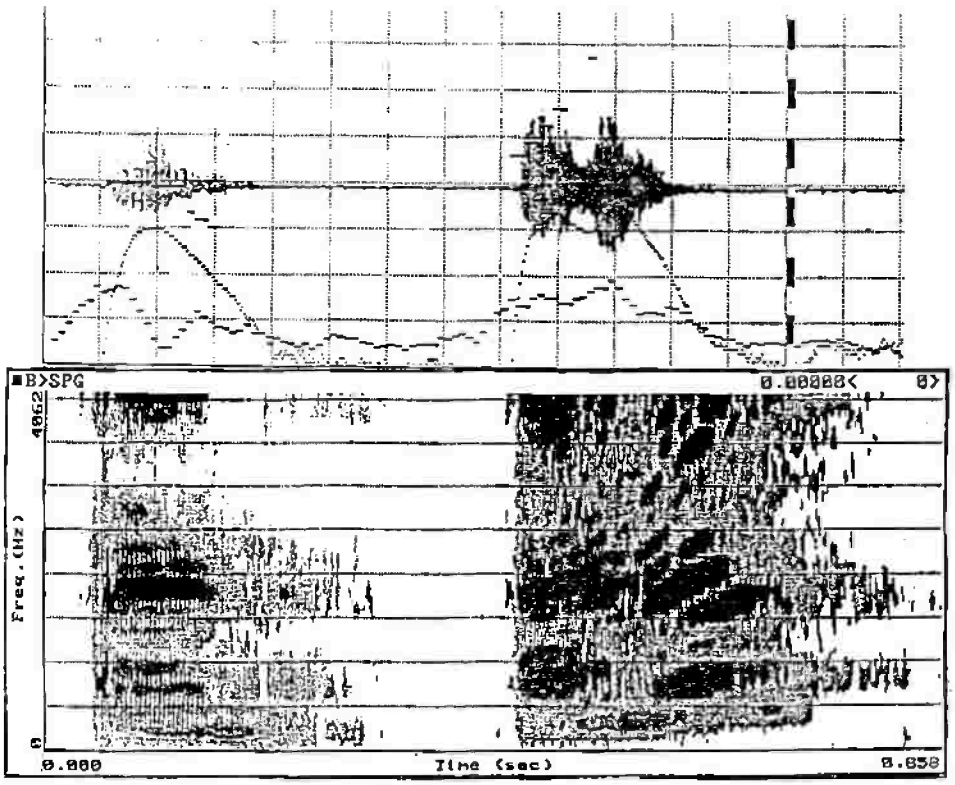


Fig. (6) Fem. G (Nor., Ast.)

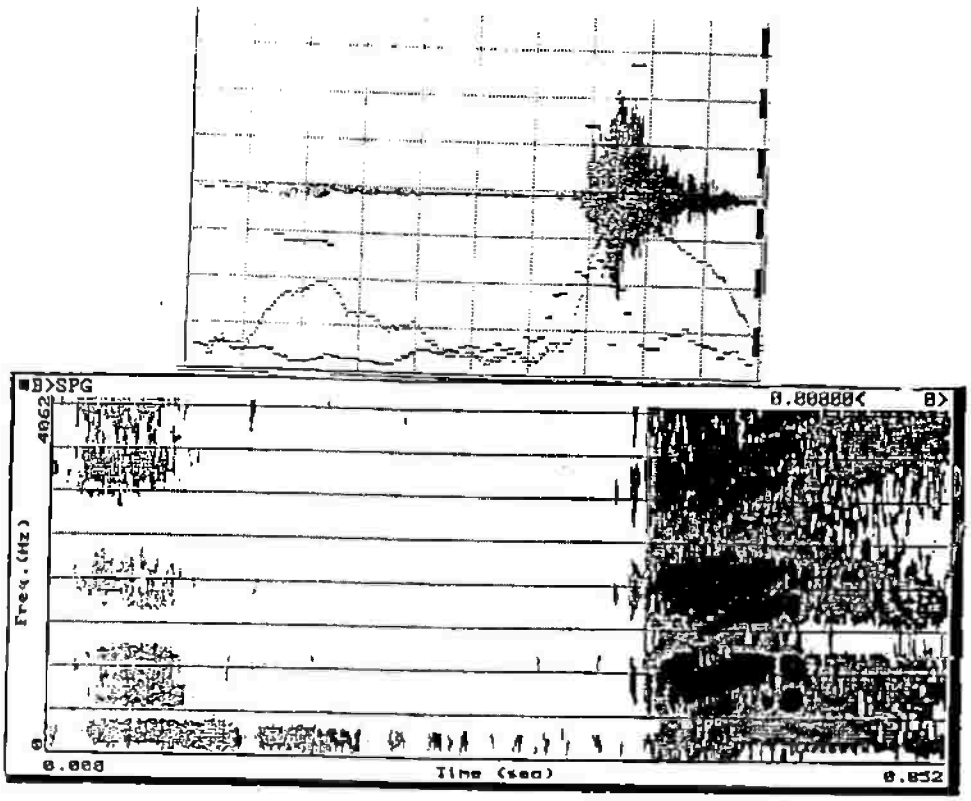


Fig. (7) Fem. A (Sad, Ang.)

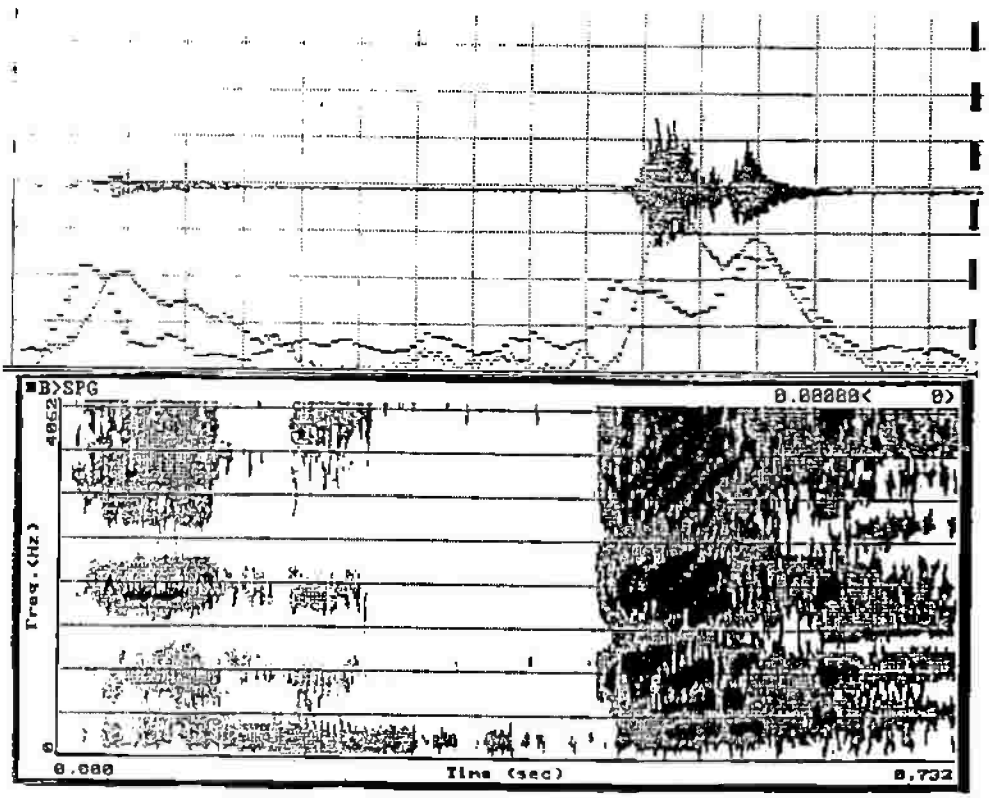


Fig. (8) Fem. A (Nor., Ast.)

5. Discussion

5.1- Dimensions of Auditory and Acoustic Differentiation

Each category included samples on which more than 80% of the listeners agreed, among the emotional categories, can be embedded in a single word.

Most "Astonish" and "Angry" intentions were perceived as intended, and many of them were agreed on by over 90% of the judgements.

The distributions of auditory judgements showed that "Sad" was somewhat confused with "Normal". The results suggested that the confused categories share some acoustic features. Although "Astonish" and "Angry" acoustic features were somewhat similar, but by the auditory judgements they were always correctly identified. This is possible by specific patterns of intonation and specific qualities not related, only, to variation of mean Fo and sound pressure.

Most variations of mean Fo, Fo range, Fo contour, intensity and duration took place during the stressed and unstressed segments of the emphasized word [?àna]. The single word carrier was probably more suitable than connected speech for the expression of feelings, these observations are in line with the suggestions made of: Leinonen et al. (1997) "The name: saára", Cummings and Clements (1995) "Two Vowels", Monsen and Engeb. (1977) "stressed/ unstressed syllables of schwa the sequence of [ʌ-ə-ə], Holmberg et al. (1988) "nonsense syllable sequences of [pa], and Laukkanen et al. (1996) "nonsense st./ unst. [paàpa]".

Ladd (1985) suggested that single word also facilitates the speakers' task which would be used naturally for emotive signalling, and it is possible that connotations could be distinguished using a segment alone.

A speaker's effort to increase the sound pressure from normal to loud as "Astonish" and "Angry" is accompanied by rise of Fo, whcih is in agree with the results of (Monsen 1977, Holmberg 1988) and disagree with Leinonen et al- (1997) who suggested that:

"The increased effort in emotive speech was however, not accompanied by a reflexive increase in Fo as during the regulation of the intensity of neutral speech: for instance, when changing from "naming" to "commanding" voice, some speakers increased sound pressure without increasing Fo.

From the literature on the vocal expression of emotion in general, different authors have studied different groups of emotions using different techniques. It is clear that numerous factors in the voice contribute to the vocal expression of emotion, but there is considerable agreement about which ones are the most important. It seems, however, that the pitch envelope (i.e level, range, shape, and Timing of the pitch contour) is the most important parameter in differentiating between emotions.

Moreover, the speech spectrum carries much more information about the glottal waveform than sound pressure, Fo and cues induced by variations in the vocal tract shape with the determination of spectral criteria which would allow a good separation among different speech modes.

5.2- Normal

It is the "neutral" voice mode, i.e. it exhibits the voice in the natural condition which is far from any unusual effort. Acoustic parameters of the normal voice were in median level if compared with soft and loud voices. The differences between male/ female sex appears clearer in the normal voice than in loud or soft voices. see statistical calculation (4.1) which showed the big differences between Normal $\bar{x}$ Fo of females = 221 Hz and of males = 158 hz, in comparison to loud voice e.g. "Angry" $\bar{x}$ Fo of females = 256 Hz and of males = 234 Hz.

This finding is in disagreement with Abe (1980) who suggested that "pitch range, change, level of male/ female voices is even more extreme if speakers compared in excited speech".

5.3- Angry

The general dictionary definition of anger is "extreme displeasure" but its use in this study and also in the literature

seems to commonly tend toward agression rather than simply displeasure. Agression has been among the best - conveyed connotations in the present study, also in other studies on emotive speech (Banse & Scherer 1996) and (Van Bezooijen 1984). In the present study, comparing anger with the three other emotions, angry connotation was expressed with the highest speech pressure, and with high mean Fo and the longest duration see statistical calculations (4.1) and it is shown as high pitch of straight melodic line by all the subjects and the highest intensity contour especially on the stressed syllable see graphic representations in (4.2).

Fónagy and Magdics (1963, P. 297) noted that "Anger is generally expressed on a mid pitch level and is characterized by a straight, rigid melodic line leeping up a fourth, a fifth, or a sixth interval at the beginning of the phrase. They also noted (1981) that increasing intensity showed shortening of the consonants and lengtheing of vowels.

Comparing "anger" with the other emotions that they studied, Williams and Stevens (1972, P. 1248) also noted that "Anger" produced the highest observed pitch "at least half an octave above neutral" and the widest observed pitch range although they found that utterance duration were longer, indicating a reduction of speech rate.

Cumminges and Clements (1995) in their study of glottal source of emotions noted, "of all the 11 speech styles, angry has the most distinct and complete closed - glottis segment, the amplitude of the pulse is very high. This is expected because the acoustic energy generated is maximized with sharp glottal closure".

5.4- Sadness

In comparison to neutral speech (i.e normal), sadness has been found to exhibit normal or lower than normal average pitch, average intensity and it has a narrow pitch range. It is characterized by a short vowel duration, a decrease in intensity and attenuation of fundamental frequencies which sometimes is unable to show pitch contour especially in the unstressed syllables. Females proportionally confused sadness with normal

emotions more than males but it is characterized in some of their productions by breathy voice. Fonagy (1963) noted, in the case of sadness, a decrease in intensity, download inflections, "resonants" timbre, "slurred" enunciation and rhythm with irregular pause. Davitz (1964) and Black (1961) noted that soft speech was characterized by slow rate.

5.5- Astonishment

In the present study, all the "Astonish" intents were correctly conveyed, although the varied differences of pitch configurations of the various productions see (4.2.1) pitch contour of astonish. The good conveyance os astonishment was probably due to intonation contours.

As a common description of the "Astonish" graphs, intonation contour has an abrupt rise to high level especially on the stressed syllable accompanied by a rise intensity curve and again rising to high pitch level and high intensity level on the unstressed syllable but less than the stressed syllable. Astonishment showed the widest pitch range among the other emotions. The finding is in agreement with other observations on surprise speech (Van Bezooijong 1984, Hutter 1968 and Laukkanen 1996)

Fónagy & Magdics (1963, p. 293) noted that in surprise, the voice suddenly glides up (or up - and - down) to a high level within the stressed syllable, then... falls to the mid level (Joyful Surprise) or to a lower level (Stupefaction) in the last syllable. also "the beginning of the phrase bears a strong stress, the following syllables run down weakly", "tempo is restrained" and "the voice is breathy" Ortony and Turner (1990, P. 317) suggest that it is not, in fact, an emotion in itself, but is important in "The elicitation and intensification of emotions"

6- Conclusion

Human emotion is a complex field of study, this is due to the intercorrelation between physiologyical, psychological, physical, linguistic and semantic factors. Numerous voice variables contribute to the vocal expression of emotions but the results suggested that the most important physical factors by which the intended connotations conveyed were: the sex of the speaker, the stress, the pitch envelope of the production.

The results showed that:

1- Male- Female differences are significant factors to distinguish the vocal expressional emotions. Moreover female samples conveyed the connotations better.

2- Male - Female differences in Fo were extremely significant especially in Sad and Normal, while in loud excited vocal emotion these differences are eliminated.

3- There is a significant increase of Intensity from soft to normal and from normal to loud voice for both males and females, but there is no significant male - female differences in Intensity in any of the emotional conditions.

4- Increasing Fo is accompanied by increasing in intensity and duration.

5- Generally duration of the vowels is shorter in soft voice, and it seems to be the less effective variable.

6- Females move in wider melodic voice range than males, i.e., female pitch level is higher about 1 octave than males.

7- Pitch contour is the best indicator - among the other parameters - in the excited expression of the loud voices. For this reason, Angry and Astonish were the best conveyed intentions.

8- Intensity in Angry is higher than Astonish, but Astonish has various pitch configurations with wide melodic levels than Angry

9- "Anger" and "Astonish" intents had higher sound pressure than normal, but "Sad" is sometimes lower than "Normal" and sometimes equals.

10- Stressed vowels always bear the discrete physical indicators for the discrimination of vocal emotions.

11- Mean Fo and Intensity during the first stressed vowel were the highest for Astonish and Anger which in these respects differed from all other categories.

References

1- Abe, I. (1980) "How Vocal Pitch Workes?" in the Melody of Language, edited by L.R Waugh and C.H. Van Schooneveld (Univ. Park, Baltimore).

2- Alpert, M. & Rosen, A. (1990) "A Semantic Analysis of the Various Ways that the Terms "Affect," "Emotion." and "Mood" are Used". J. Commun. Disord. 23, 237 - 246.

3- Banse, R., and Scherer, K. (1996) "Acoustic Profiles in Vocal Emotion Expression" J. Pers. Soc. Psychol. 70, 614 - 636.

4- Black, J. (1961) "Relationships Among Fundamental Frequency Vocal Sound Pressure, and Rate of Speeking" Lang. & Sp. 4, 196 - 199.

5- Cahn, J. (1988) "From Sad to Glad: Emotional Voices" Proceedings of Speech Tech.' 88 New York, 35 - 36.

6- Childers, D. and Lee, C. (1991) "Vocal Quality Factors: Analysis, Synthesis and Perception" JASA 90, 2394 - 2410.

7- Cummings, K. and Clements, M. (1990) "Analysis of Glottal Waveforms across Stress Styles". In Proce. Intern. Conf. on Acoustics. Speech and Signal Processing. 3 rd - 6th A.

8- Cummings, K. and Clements, M. (1995) "Analysis of the Glottal Excitation of Emotionally Styled and Stressed Speech" JASA 98. PP. 88 - 98.

9- Darwin, C. (1872). The Expression of Emotions in Man and Animals. Reprinted (1965), Chicago: University of Chicago Press.

10- Davitz, J. and Davitz, L. (1959) "The Communication of Feelings by Content- Free Speech" J. Commun. 9, 6 - 13.

11-Davitz, J. (1964) "Auditory Correlates of Vocal Expressions of Emotional Feeling" in the Communication of Emotional Meaning, edited by J.R. Davitz (Mc Graw - Hill, New York), 101 - 112.

12-Dawes, R. and Kramer, E. (1966) "A Proximity Analysis of Vocally Expressed Emotion" Percept. Mot. Skills 22, 571 - 574.

13-Fairbanks, G. and Hoaglin, L. (1941) "An Experimental Study of the Durational Characteristics of the Voice During the Expression of Emotion" Speech Monograph 8, 85 - 91.

14-Fant, C. (1970) Acoustic Theory of Speech Production, Mouton, the Hauge.

15-Fant, G. (1979) "Glottal Source and Excitation Analysis" Speech Trans. Lab. Quar. Prog. & Status Report, 1. 85 - 107.

16-Fónagy, I. (1981) "Emotions, Voice and Music" in Research Aspects on Singing, edited by J. Sundberg (Royal Swedish Academy of Music No 33) PP. 51 - 79.

17-Frick, R. (1985) "Communicating Emotion: the Role of Prosodic Features" Psychol. Bull. 97, 412 - 429.

18-Glave, R. and Rietveld, A. (1975) "Is the Effort Dependance of Speech Loudness Explicable on the Basis of Acoustical Cues?" JASA 58. 4. P. 875.

19-Hofling, C. (1975) "Textbook of Psychiatry for Medical Practice, 3rd ed. Philadelphia: J.B. Lippincott.

20-Holmberg, E., Hillman, R., Perkell, J. (1988) "Glottal Airflow and Transglottal Airpressure Measurements for Male and Female Speakers in Soft, Normal and Loud Voice" JASA 84. 2. P 511.

21-Huttar, G. (1968) "Relations between Prosodic Variables and Emotions in Normal American English Utterances" J.SP. Hear. Res. 11, 481 - 487.

22-Izard, C. (1971) The Face of Emotion. New York: Appleton - Century - Crofts.

23-Izard, C. (1977) Human Emotions. Plenum, New York.

24-Johnson, W., Emde, R., Scherer, K. and Klinnert, M. (1986) "Recognition of Emotion from Vocal Cues" Arch. Gen. Psychol. 43, 280 - 283.

25-Kolb, L. and Brodie, H. (1982) Modern Clinical Psychiatry, 10th ed. Philadelphia P. 24.

26-Kramer, E. (1963) "Judgement of Personal Characteristics and Emotions from Nonverbal Properties of Speech" Psychol. Bull. 60, 408 - 420.

27-Ladd, D., Silverman, K., Tolkmitt, F. Bergmann, G. and Scherer, K. (1985) "Evidence for the Independent Function of Intonation Contour Type, Voice Quality, and Fo range in Signaling Speaker Affect" JASA, 78, 435 - 444.

28-Laukkanen, A., Vilkman, E., Alku, P., and Oksanen, H. (1996) "Physical Variations Related to Stress and Emotional State: a Preliminary Study". J. Phon. 24, 313 - 335.

29-Leinonen, L., Hiltunen, T., Linnankoski, I. and Laakso, M. (1997) "Expression of Emotional - Motivational Connotations with a One - Word Utterance" JASA, 102, 3, 1853 - 1862.

30-Levin, H., and Lord, W. (1975) "Speech Pitch Frequency as an Emotional State Indicator" IEEE Trans. Syst. Man. Cybern. SMC - 5, 259 - 273.

31-Lieberman, P. and Michaels, S. (1962) "Some Aspects of Fundamental Frequency and Envelope Amplitude as Related to the Emotional Content of Speech" JASA 34, 922 - 927.

32-Monsen, R. and Engebretson, A. (1977) "Study in Variations in Male and Female Glottal Wave" JASA 62, 981 - 993.

33-Murray, I. and Arnott, J. (1993) "Toward the Simulation of emotion in Synthetic Speech: A Review of the Literature on Human Vocal Emotion" JASA, 93 (2) P. 1097

34-Oatlay, K. (1989) "The Importance of Being Emotional" New Science 123 PP. 33 - 36.

35-Ortony, A. and Turner, T. (1990) "What's Basic about Basic Emotions?" Psychol. Rev. 97, 315 - 331.

36-Redlich, A. and Freeman, D. (1966) The Theory of Practice of Psychiatry. New York: Basic Books.

37-Scherer, K. (1981) "Speech and Emotional States", in Speech Evaluation in Psychiatry, Edited J.K. Darby.

38-Scherer, K. (1984) "Emotion as a Multicomponent Process: A Model and Some Cross - Cultural Data" Rev. Personal Social Psychol. 5. 37 - 63.

39-Scherer, K. (1986) "Vocal Affect Expression: A Review and a Model for Future Research" Psychol. Bull. 99, 143 - 165.

40-Sulc, J. (1977) "To the Problem of Emotional Changes in Human Voice" Act. Nerv. 1 Super. 19, 215 - 216.

41-Uldall, E. (1960) "Attitudinal Meanings Conveyed by Intonation Contours" Lang. & SP. 3, 233 - 234.

42-Van Bezooijen, R., Otto, S. and Heenan, T. (1983) "Recognition of Vocal Expressions of Emotion: A three - Nation Study to Identify Universal Characteristics" J. Cross - Cultural Psychol. 14, 387 - 406.

43-Van Bezooijen, R. and Boves, L. (1986) "The Effect of Low - Pass Filtering and Random Splicing on the Perception of Speech" J. Psychol. In Linguis. Res. 15, 403 - 417.

44-Williams, C. and Stevens, K. (1972) "Emotions and Speech: Some Acoustical Correlates" JASA 52, 1238 - 1250.