The Function of Pitch Range Variations in Samples of Emotional Expressions in Persian

Elkhas Veysi

Associate Professor, Payam-e-Noor University

elkhas@yahoo.com

Abstract

This study aims at investigating the interface between emotion and intonation patterns (more specifically, duration and pitch amplitude of speech). To this end, the acoustic properties of spectral parameters related to speech prosody are investigated. The results of acoustic and Statistical analysis show that mean level and range of FO in the contours vary strongly as a function of the degree of activation of the portrayed emotions. The results also indicate that FO/pitch and other vocal aspects affected by emotion and/or emotional arousal in collected samples of speech in Persian.

Keywords: Pitch, Range, Tilt, Model, Emotion, Intonation Pattern, Spectral Parameters

1. Introduction

Over the past decades, the interface between emotion and intonation has been addressed by many researchers from different research backgrounds and is still a matter of sustained debates. Consequently phonologists researching intonation have switched their interest from mere description of pitch contours of individual clauses and sentences to larger chunks of speech. They have tried to establish the function of different prosodic features in the development of spoken interaction.

Phonologists were no longer interested in the grammatical functions of prosodic features but rather in how they contribute to emotional, attitudinal and also to the ever-changing nature of spoken discourse and its understanding. Communication is not merely an exchange of words. Linguistic, paralinguistic and non-verbal communication elements such as, co-speech gestures and non-speech sounds are part of the communication and all convey meaning (Lyons, 1995, p. 97).Such acoustic parameters of speech and prosodic cues as variation in pitch, intensity, speech rate, duration of vowel may be manipulated -exaggerated, diminished, sped up, or slowed down- to convey emotions and attitudes.

Pitch range variation can help the speaker and listener to encode and decode the full emotional speech. They can also provide paralinguistic information such as the speaker's, emotion and attitude toward the shared interlocutor and the situation. Indeed, all the acoustic cues contained in speech contribute to the interpretation of the message.

Considering the above-mentioned points the present research aims to investigate the pitch range variation conveying the expression of emotion and attitude in samples of real Persian speech. The acoustic parameters for such a description of emotion in speech are FO/pitch range and duration of vowel. To

this end, detailed acoustic information on four basic emotions (anger, sadness, joy and neutral) expressed in speech by six speakers (3 males and 3 females) are obtained.

Acoustic analysis is conducted in order to show the effectiveness of these parameters in emotion categorization in Persian based on the intonation model of the research. To derive quantitative international measures we have used the Tilt model. This paper presents an analysis of Fo contours in portrayed emotions, using the suggested model. The focus is on quantifying the Fo differences among the above-mentioned emotions. This study also analyses the use of intonation as an instrument for arousing various sensation in listeners. The aim was to verify which elements of intonation are more decisive to generate a specific sensation. Thus, in this study, the production aspect supplemented with the perception one; this is due to the fact that the complementarity of the production and the perception processes is the basis of the spoken communication process and firmly establishes the communicative importance of the parameters being studied.

As an interdisciplinary field, the findings of this study may have a range of applications in clinical, psychological, pedagogical or communicative research.

In what follows, the key researches conducted in non-Persian and Persian languages are reported.

2. Previous Studies

Studies relating to intonation and emotion that use parameter “pitch level” to analyze the relationship between the given variables are highly numerous. Acoustic investigation of paralinguistic information and emotion expressed in speech has recently received increasing attention.

Uldal (1964) applied 16 stylized Fo contours on five utterances. She showed that the emotional meaning attributed to different contours varies depending on the sentence carrying the contour.

Williams and Stevens (1972) conducted a study to demonstrate the importance of pitch range variation as a medium for expressing emotion in speech. In their study the importance of supplementing production with perception studies is also stressed.

Grabe, et al. (1997) in their research concluded that "the Frequency Code is widely used for the expression of affective meanings", whereas for Mozziconacci (1984) "the pitch level and excursion size of the pitch movements would be more related to the speaker's emotional state."

Mozziconacci (1984, pp. 79-163) also explored pitch range variations in production and perception of Dutch speech conveying six emotions or attitudes: joy, anger, sadness, boredom, fear and indignation against neutrality as a reference. In the first stage of the study 315 utterances (3 speaker×5 sentences×7 emotions×3 trials) successfully conveying emotion in speech were selected as speech material. This selection was made on the basis of perception test. An analysis of the subset of fourteen utterances (1 speaker×2 sentences ×7 emotions×1 trial) successfully conveying the emotion categories was carried out, involving the global measures, pitch range, pitch level and speech rate. The curve of pitch in all individual utterances was described in terms of the IPO model known as perceptual Analysis of Intonation. The measurements obtained in the production study were used as a source of inspiration for speech manipulations. In a series of experiments, Mozziconacci found that pitch level, pitch range and speech rate were systematically varied per emotion around the values found for these parameters in the original speech.

Ladd et al. (1985) explored prosodic cues to emotional states. It was indicated that not only Fo range and voice quality had strong effects on listener's influences of the arousal-related state of the speaker, but also on inference of cognitively related attitudes and emotions. Ladd (1985) also came to the conclusion that three prosodic cues: Fo range, voice quality and type of pitch contour function independently of each other for conveying emotion and attitude in speech.

Scherer (1990) conducted the most detailed study. He expressed the need for a combination of production and perception studies in order to make progress in understanding emotional vocal communication. Scherer (1990) applied different techniques to simulate standard emotions like anger, sadness and joy to test how well subjects can identify those emotions from their acoustic signals. Conducting another study on Dutch and Hungarian languages Scherer (1995) by testing assumptions concerning intonation theories, reached conclusions very similar to Mozziconacci's (1995).

Tickle (1990) conducted cross-cultural studies of emotional expressions. She distinguishes the two influences of biological factors, leading to the expectation of universal expression of emotions across cultures, and culturally determined factors, leading to expectations of influences in expression between cultures.

Laver (1995) explored the prosodic function of emotional speech based on linguistic and a paralinguistic components. He said that the choice of contour would be more related to the type of sentence, while the pitch level and excursion size of the pitch movements would be more related to the speaker's emotional state. The role of intonation contour in conveying emotional state and meaning was also demonstrated in Grab et al. (1997). In their study, an orthogonal design was used, combining high and low preheads with three

Dutch pitch accents: H*L, H*LH and L*H, which resulted in six experimental intonation contours. A perceptual experiment was conducted in order to test which of the contours would best convey friendliness, aloofness, irritation, uncertainty and politeness.

The study of correlation between prosodic features and emotional speech has been scarcely conducted in Persian. Veysi (1386/2007) in his study conducted an experimental research seeking various degrees of correspondence between prosodic cues (duration, pitch level and intensity) and paralinguistic information in some samples of speech in Persian. His analysis of all received data revealed strict correlation between prosodic features and paralinguistic information in Persian.

The present study along with the other related reviewed explorations promises new source of information for the investigation of long-time concerns in related fields.

3. Methodology

3.1. Material and Recordings

The experiment was set out to determine the extent to which the variables – pitch level and contour type-affected the arousal of different sensations in the listeners. To do so, a corpus prepared at two stages, i.e. production and perception. First, the production of samples is used to construct the data. The advantage of this method is that it contributes us to control and compare the acoustic variation in the same context for all suggested emotions in all participants. The disadvantage of this method is due to the fact that the given method of constructing data may have unnatural aspects. The corpus used in this study consisted of 150 emotional expressions. The samples analyzed in this study were uttered with any of the four emotions, i.e., anger, joy, sadness and

neutral. 10 participants (5 males and 5 females) among the students from Payame Noor University of Ahvaz (aged between19-30) with no voice pathologies or abnormalities were used for producing the utterances. Each selected utterance was spoken several times in a neutral voice and several times with the related emotions of anger; joy and sadness and then they were recorded. The actors were not given instructions on how the sentences was to be read in terms of intensity, speed ,duration of vowel or pauses. To achieve realistic portrayal of emotion, the speakers were instructed to imagine a situational context in which the sentences could appear. At the second stage, the recorded utterances were played to the listeners in separate sessions. Subsequently, these recordings were subject to evaluation by a sample comprising 5o students majoring in English translation. 25 males and 25 females, chosen randomly and without intonation knowledge. They were of varying ages and sexes. The listeners were instructed to make any utterance that sounded as acted utterance or emphasized as unnatural. The emotional content was evaluated using 5 categories: neutral, joy, sadness, anger and unnatural or not recognizable emotional content. Only sample sentences recognized by at least 80% of all listeners were used for the parameter analysis. The utterances used for parameter analysis in the present research were evaluated as natural and the emotional content was deemed unambiguous. Once all the stages had been over, the tests were chosen and were significantly different for each emotion in terms of pitch level and contour pitch.

3.2. Measurement of Acoustic Parameters

The analysis of recorded samples was made by using acoustic analysis software *praat,* broadly used for prosodic studies. Considering the model of intonation

in the present study, the rise and fall connection model, such parameters as pitch range and duration were measured from the corresponding label files.

The pitch contours of each utterance were measured using *Praat* (Boersma and Weenink, 2006) speech processing software. Global level statistics related to Fo (maximum-minimum) were calculated from Fo contours in each peak. All speakers' samples were acoustically analyzed via the pitch tracks and spectrogram using the mentioned software. As this stage was over, all scales were put in a matrix. Then, Statistical analysis of all measures related to each sample utterance contributed us to hypothesize that significant differences would be found between the acoustic cues used by the speakers to express emotions: neutral, sadness, joy and anger. Linear Regression was used to test the correlation between the variables.

Since this research is conducted in the framework of a model of intonation, it is appropriate to provide a brief description of the tilt model which we appealed to as the basis of this research.

4. Intonation Modeling

There are many different approaches to model intonation in speech analysis. One classification divides various models into acoustic, perceptual and linguistic models. Acoustic models aim to reproduce the intonation patterns in a compact way. The perceptual models concentrate on those intonation events, which are the most relevant perceptually. The linguistic approach treats intonation contours as a part of the linguistic structure. Considering intonation data concerning emotional speech within the theoretical framework of approaches to intonation contributes us to understand which meaningful variations, as mentioned in [3.2] help to convey emotion in speech. The model for our study was selected due to its compactness and assumed suitability for

the acoustic analysis of isolated sentences. The chosen model is the *tilt* model (Taylor: 1998, 2000 and 2009) for processing pitch range variation. In the following section brief details of the tilt model provided.

4.1. The Intonation Model

In this section we introduce the intonation model which is designed specifically to facilitate automatic intonational processing of speech. In this model, intonation is represented as a linear sequence of events, which can be pitch accents or boundary tones. In the model, intonation is characterized by a sequence of phonetic intonational events. The basic unit in the model is the *intonational event.* Events occur as instants with nothing between them ,as opposed to segmental based phenomena where units occur in a contagious sequence.The basic types of intonational events are *pitch accents* (denoted by the letter a) and *boundary tones* (denoted by the letter b). Pitch accents are Fo excursions associated with syllables which are used by the speaker to give some degree of emphasis to a particular word or syllable. In the *tilt* model, boundary tones are rising Fo excursions which occur at the edges of intonational phrases. A combination of event *ab* occurs when a pitch accent and boundary tone occur so close to one another that only a single pitch movement is observed. Each event has a rise and fall component which can vary in size. Some events have a zero rise or zero fall component indicating that they only have a fall or a rise respectively. The middle of the event is defined as the end of the rise component or start of the fall component. Each event is characterized by tilt parameters, which describe its Fo shape. This model can be regarded as a phonetic model of intonation in that it describes the intonational phenomena observable in an Fo contour.

In this study, the acoustic parameters of the first three peaks, representing rise and fall components, were calculated. Events *a* and *b* in each part are measured separately. Intonational events are represented in figure (1).

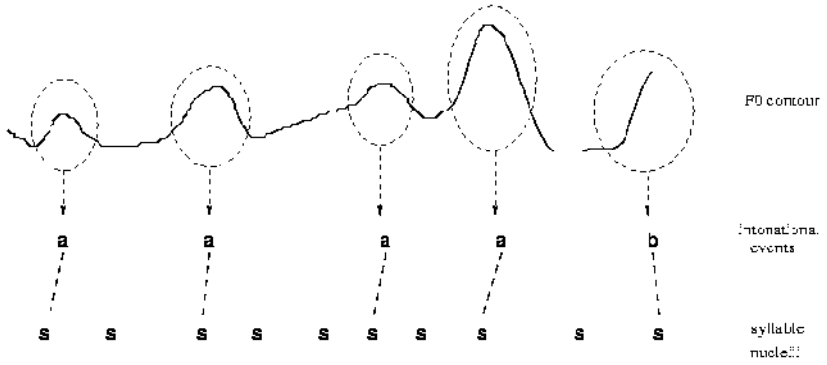

Figure 1. ***Intonational Events in the Intonational Model*** (Taylor, 2009, p.122)

The model facilitates automatic intonational processing. It follows up different stages. The first stage in automatic analysis is to find the events from the waveform. Waveforms are parameterized into Fo and energy. The parameters for each of the events are derived. This stage uses an algorithm which examines each event and fits rise or fall shapes by minimizing the error between original contour and the fitted shape. The results of this process are that each event is now described as a rise shape, a fall shape or a rise followed by a fall shape. This parameterization produces a representation in terms of the model, known as the rise/fall connection (RFC model).

5. Analysis of Data

Each utterance is analyzed in terms of emotion and sex of the participants as independent variables and peaks and the rise/fall half of each peak as dependent variables. The null hypothesis is considered with regard to the

interface between the given two variables supported by descriptive and inferential statistics. The significance level given in different tables will determine the kind of the hypothesis. The sig. level is set to be less than 0.05 ($p<0.05$). A diagram of the average of each parameter in both halves of the same peak will also be shown.

We also aimed to seek the correlation between the intended emotions and the given acoustic parameters in the rise half and fall half of the first peak of all given samples taken out of the emotional speech based on the method noted in [2-4]. We restricted our work to present descriptive statistical results of the given variables in the first peak of one utterance, say the fourth one. For limitation of space we will confine our work to the results made for other peaks in the given utterances.

Notice the following sample utterances:

(1) befarm?id! Hame montazere £oma budim.
Come in! Every body's been waiting for you.

(1) jeddi migi! Az ? u da:vat be kar kardan
You're kidding! He was called to a job.

(2) VaGe?an!?emruz mehmun darim.
Really! We have guest today.

(3) be salamat! Farda mibinamet.
Bye! See you tomorrow.

(5) javad ?az safar barga£t xune!
Javad returned from a trip!

In the following figures, the spectrographs of the first peak of the sample 1 for three emotions i.e., joy; anger and sadness, compared with neutral expression as a baseline are given:

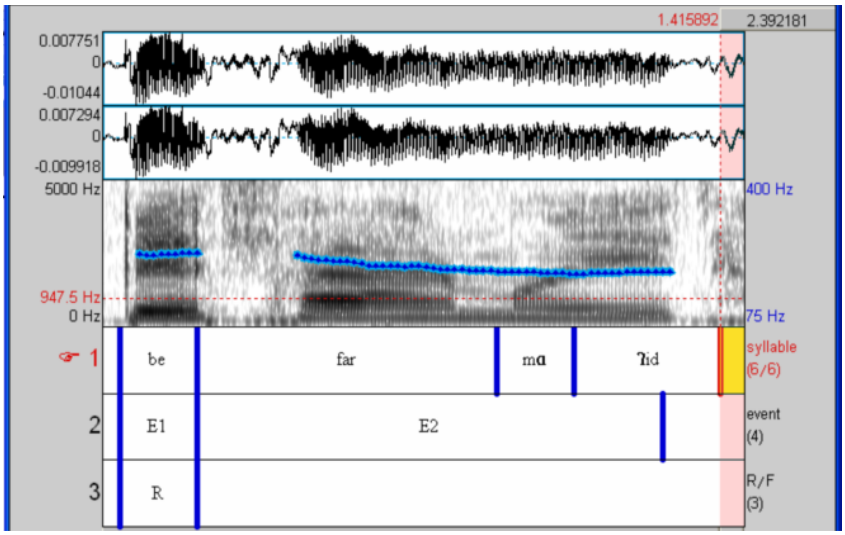

Figure 3. *The Spectrograph of the First Peak, Befarm?Id, for Neutral Utterance of Sample 1*

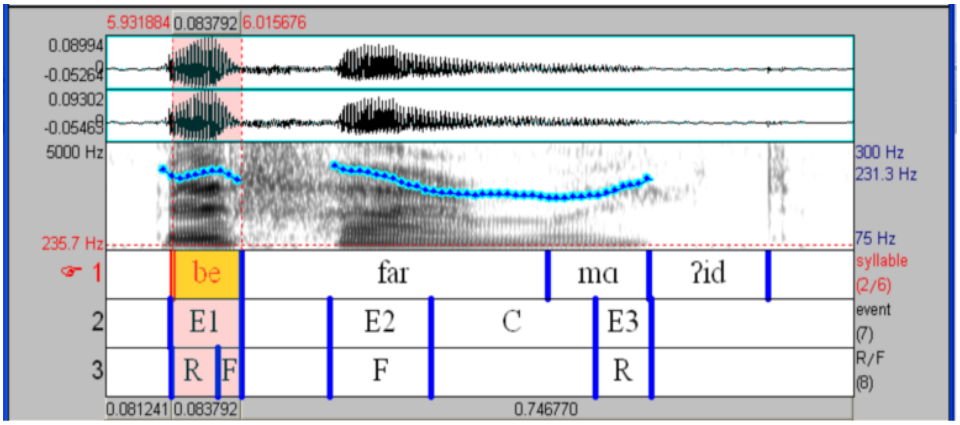

Figure 4. *The Spectrograph of the First Peak, Befarm?id , for Sadness Expression of Sample 1*

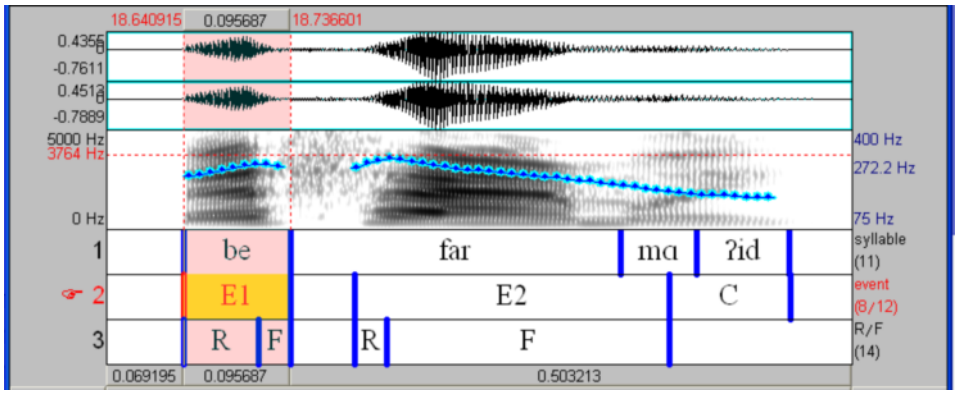

Figure 5. *The Spectrograph of the First Peak, Befarma?id, for Anger Expression of Sample 1*

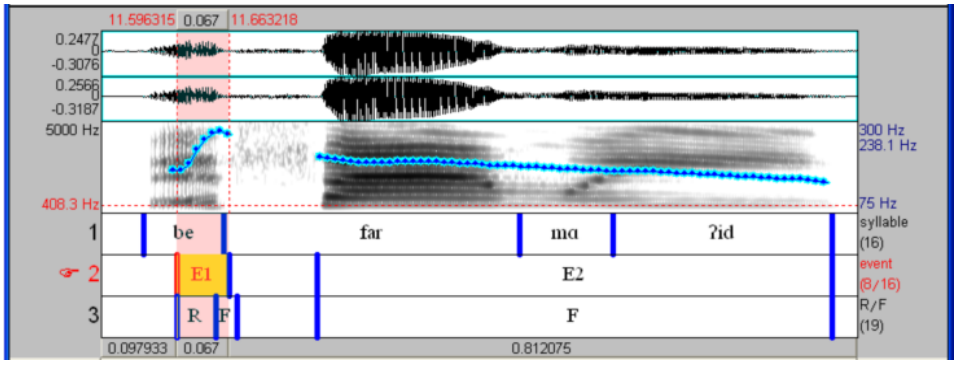

Figure 6. *The Spectrograph of the First Peak, Befarm?id , for Joy Expression of Sample 1*

In what follows, the interface between the acoustic parameters: duration, pitch range and tilt and the portrayed emotions will be explored in details.

5.1. Duration and its Interface with the Emotions

Statistical results of the variables of one peak are initially presented in the same (fourth) utterance for all intended emotions. Then, total results of all samples will be illustratively presented.

A- Null hypothesis about the duration: the duration of utterance in the first peak (first / second half) of the fourth sample to be bound to emotions e.g. sadness, joy and anger in males and females showed no significant difference.

Counter hypothesis: the duration in the first peak (first/second half) of the same utterance to be related to the emotions sadness, joy and anger in males and females showed significant difference.

Statistical data of duration in the first peak of the fourth utterance "**befarmâ?id! hame montazere šomâ budim.**" is given in the table below.

Table 1.*Statistical Data of Duration and its Correlation with Intended Emotions in the First Peak of the Utterance (1)*

Half	Sex	Neutral	Emotion	Estimation of coefficient	Variance of estimation	Standardized coefficient	Statistical hypothesis	Significance level
first half	Female	.618	sadness	- 230	.67	- .590	- 3.345	.007.5
			Joy	- .298	074	- .765	- 463	.002
			anger	- .426	.065	- 1.098	- 6.656	1
	Male	.532	sadness	- .168	.180	- .327	- .957	.366
			joy	- .239	.169	- .449	- 1.337	.199
			anger	- .388	.169	- .751	- 2.205	.049
second half	Female	.142	sadness	.342	.101	.807	3.324	.010
			joy	.386	.103	.915	3.754	.006
			anger	.192	.103	.454	1.864	.099
	Male	.706	sadness	- .410	.182	- . 725	- 2.259	.059
			joy	- .327	.182	- .578	- 1.800	.111
			anger	- .378	.167	- 684	- 2.112	.072

As seen in table 1, significant difference was found between the duration and the emotions e.g. joy and sadness in females in the first peak of the fourth sample. The results summarized in table 2 also indicate the significant correlation between the two given variables.

Table 2.*Statistical Data of the Total Significant Correlation Between the Duration and Intended Emotions in Three Peaks of all Samples*

Utterance	Peak	Half	Sex	Emotion	Neutral	Coefficient	Standardized Coefficient	Significance Level
1	2	2	Female	Joy	.214	.162	.788	.027
1	3	1	Female	Anger	.425	-.233	-.699	.027
2	1	1	Female	Joy	.178	.465	.844	.019
2	1	1	Female	Anger	.189	.410	.727	.031
2	2	1	Female	Sadness	.129	.207	.917	.013
3	2	1	Male	Sadness	.127	.194	.975	.010
3	3	1	Male	Sadness	.131	.284	.729	.045
4	1	1	Female	Sadness	.618	-.229	-.589	.009
4	1	1	Female	Joy	.619	-.298	-.775	.006
4	1	1	Female	Anger	.618	-.419	-1.088	00.6
4	1	2	Female	Sadness	.129	.340	.808	.014
4	1	2	Female	Joy	.127	.384	.919	.008
4	3	2	Female	Sadness	.159	.357	.923	.016
5	1	1	Female	Anger	.184	.228	.990	.008
5	1	1	Male	Joy	.359	-.054	-.525	.038
5	1	1	Male	Anger	.348	-.118	-1.060	.007
5	3	2	Male	Sadness	.217	-.089	-.907	.004
5	3	2	Male	Joy	.211	-.058	-.6040	.035
5	3	2	Male	Anger	.211	-.096	-.888	.006

As seen in table 2 there is a significant correlation between the two variables in the given samples. Thuso, the null hypothesis was rejected.

The total result show that the differences statistically considered to be significant are not the same for all utterances. Significant difference was only found for 25% of the related cases which was mostly observable for the emotions: joy and anger in females.

5.2. Pitch Range and Its Interface with the Emotions

In the following part the correlation between two variables, the pitch range and the given emotions, will be investigated in the first sample.

A- Null hypothesis about the pitch range: the pitch range of utterance in the first peak (first/second half) of the first sample to be bound to emotions e.g., sadness, joy and anger in males and females showed no significant difference.

Table 3. *Statistical Data of the Pitch Range and Its Correlation with the Intended Emotions in the First Peak of Utterance (4)*

Half	Sex	Neutral	Emotion	Estimation of coefficient	Variance of estimation	Standardized coefficient	Statistical hypothesis	Significance level
first half	Female	83.850	sadness	4.985	29.331	.023	.170	.869
			Joy	200.417	29.331	.933	6.833	.000
			Anger	-13.683	29.331	-.064	-.467	.653
	Male	30.405	sadness	-12.205	11.400	-.268	-1.071	.316
			Joy	12.728	11.400	.279	1.116	.297
			Anger	31.295	11.400	.687	2.745	.025
second half	Female	35.467	Sadness	93.000	66.543	.388	1.398	.200
			Joy	221.800	66.543	.926	3.333	.010
			Anger	131.000	66.543	.547	1.969	.085
	Male	22.567	sadness	17.933	13.404	.274	1.338	.218
			Joy	26.000	13.404	.398	1.940	.088
			Anger	67.933	13.404	1.040	5.068	.001

As represented in table 3 the difference between the pitch range and the emotions e.g., joy and anger for females and males in the first and second half of the fourth utterance is significant thus, the null hypothesis is refuted and the given parameter is significant statistically.

The results summarized in table 4, in appendix, reveal some further cases of correlation between the pitch range and the intended emotions in three peaks of all given sample.

5.3. Rise and Fall of Pitch Contours and Its Correlation with the Emotions

We previously stated in [5] how the tilt parameters for each of the events are derived and calculated. To have more control over the variables under analysis, the previous method, considering the sig. level and the null hypothesis is appealed to analyze the data. As it has been done for [5-2] and [5-1] we primarily base our analysis on the first peak of sample (4).

A- Null hypothesis about the tilt: the tilt of utterance in the first peak (first /second half) of the fourth sample to be bound to emotions e.g., sadness, joy and anger in males and females showed no significant difference.

Counter hypothesis: the tilt in the first peak (first/second half) of the same utterance to be related to the emotions sadness, joy and anger in males and females showed significant difference.

The following table shows the descriptive statistics of the tilt and its effect on the intended emotions.

Table 5. *Statistical Data of the Tilt (Rise/Fall) and Its Significant Relation with Sad in the First Peak of the Utterance (4)*

Half	Sex	Neutral	Emotion	Estimation of coefficient	Variance of estimation	Standardized coefficient	Statistical hypothesis	Significance level
rise (tilt)	Female	.383	sadness	- .499	.193	- .613	- 2.590	.032
			Joy	- .297	.193	- .366	- 1.544	.161
			anger	- .809	.193	- .995	- 4.201	.003
	Male	.086	sadness	- .509	.182	- .849	- 2.802	.023
			joy	- .158	.182	- .264	- .872	.409
			anger	- .273	.182	- .455	- 1.501	.172
fall (tilt)	Female	.651	sadness	- .722	.127	- .815	- 5.677	.000
			joy	- .876	.127	- .988	- 6.881	.000
			anger	- .874	.127	- .986	- 6.871	.000
	Male	- .085	sadness	.144	.298	.191	.483	.642
			joy	- .001	.298	- .002	- .004	.997
			anger	- .223	.298	- .297	- .749	.475

Table 5 indicates that there exists correlation between the given variables. Statistical results in table 5 indicate that the difference between the tilt and

intended emotions: joy, anger and sadness for females in the first half rise and fall rise of utterance (4) is significant. Thus the null hypothesis is rejected, that is, counter hypothesis is confirmed.

The total findings in table (6) show more cases of correlation between the rise –fall and the intended emotions.

Table 6. *Statistical Data of the Total Significant Correlation between the Tilt and the Emotions: Sadness, Joy and Anger in all Samples*

Utterance	Peak	Tilt	Sex	Emotion	Neutral	Variance of Coefficient	Standardized Coefficient	Significance Level
1	1	Rising	Female	Sadness	.576	-.564	-.734	.041
1	1	Rising	Female	Anger	.576	-.581	-.755	.037
1	3	Rising	Female	Sadness	.407	-.572	-1.000	.002
1	3	Rising	Female	Anger	.407	-.402	-.703	.016
1	3	Falling	Female	Anger	.367	-.504	-.771	.039
3	1	Rising	Female	Joy	-.888	.838	.870	.010
3	1	Rising	Female	Anger	-.888	.653	.678	.032
3	1	Rising	Male	Sadness	-.574	.934	.712	.044
3	1	Rising	Male	Anger	-.574	1.058	.806	.027
3	3	Rising	Female	Sadness	-.534	.568	.909	.009
3	3	Rising	Female	Joy	-.534	.452	.723	.026
4	1	Rising	Female	Sadness	.383	-.499	-.613	.032
4	1	Rising	Female	Anger	.383	-.809	-.995	.003
4	1	Rising	Male	Sadness	.086	-.509	-.849	.023
4	1	Falling	Female	Sadness	.651	-.722	-.815	.000
4	1	Falling	Female	Joy	.651	-.876	-.988	.000
4	1	Falling	Female	Anger	.651	-.874	-.986	.000
4	3	Rising	Male	Sadness	-.127	-.374	-.537	.004
4	3	Rising	Male	Joy	-.127	.342	.491	.006
4	3	Rising	Male	Anger	-.127	.303	.435	.011
5	1	Falling	Female	Anger	-.396	.634	.731	.048
5	2	Rising	Female	Joy	.308	-.356	-.798	.019
5	2	Rising	Female	Anger	.308	-.379	-.850	.014

The findings summarized in table (6) indicate that the differences statistically regarded to be significant are not the same for all samples. Significant difference was only found for 26% of the related cases which was generally observable for the intended emotions e.g. sadness, joy and anger. As it can be seen in table (6) the level of significance of the parameter tilt for the given emotions in females is more than males in all samples. There exists significant relation between the tilt (rise), rather than the tilt (fall) and the given emotions in all samples. Hence, the null hypothesis is rejected. In 64% of the other related cases no statistically significant difference was found between the tilt (rise/fall) and the intended emotions.

6. Results and Discussion

The study revealed that pitch range conveys substantial information about intended emotional states i.e., joy and anger and this parameter is more important than duration and tilt. Thus, among the given prosodic features pitch range is shown to be pivotal and introduces the function of these features as a kind of paralinguistic behavior. The pitch range parameter in most cases (90%) was statistically meaningful in deviating from neutral emotion to joy and anger. The duration in 25% and the tilt in 26% of the cases showed significant difference with the intended emotions in this research. These results are partly in accordance with the results of Mozziconacci (1995) and Scherer (1990) which showed that joy and anger correlated with higher pitch range in emotional speech in Dutch and Hungarian languages.

Turning to the results and total findings, it can thus be concluded that among the prosodic features, pitch range and tilt respectively and to a lesser extent duration showed statistically significant difference with the intended emotions.

The results of the study also showed that emotions with different features i.e., anger and joy or sadness and neutral have similar acoustic cues.

It is worth quoting that in some cases no significant difference was found between the intended emotions and the given parameters. It may be due to this fact that all emotions are, more or less, dependent on context for their interpretation. As Scherer (1995) said it is not to be expected that all the information about emotion is present in the acoustic signal, but they play the key role in emotional speech. We have finally come to this conclusion that research makers from a variety of background, e.g., psychology can fruitfully integrate prosody into their research. Such studies with greater scopes and larger set of samples can be promisingly a further step toward finding the nature of emotional speech in Persian.

References

Boersma, P., Weenink, D., (2006). Praat Speech Processing Software. *Institute of phonetic science of the University of Amsterdam*. http://www.praat.org

Grabe, E., Gussenhoven, C., Hann, J., Marsi, E., Post, B. (1997). Preaccentual pitch and speaker attitude in Dutch. *Language and Speech,* (41), 64-86.

Ladd, D. R., Silverman, K. E. A., Tolkmitt, F., Bergman, G., Scherer, K. R. (1985). Evidence for the independent function of Intonation Contour type, voice quality and Fo range in signaling speaker affect. *Journal of the Acoustical Society of America*. (78), 436-445.

Laver, J. (1995),

The Phonetic description of paralinguistic phenomena. *Proceeding of the XIIIth International Congress on Phonetic Sciences.* Stockholm, Sweden, 144-153.

Lyons, J. (1995). *Linguistic semantics: An introduction*. London: Cambridge University Press.

Mozziconacci, J. L. (1984). *Prosody and emotion*. The Netherland: Leiden University.

Mozziconacci, J. L. (1995). Pitch Variations and emotion in Speech. *Proceeding of the XIIIth International Congress on Phonetic Sciences*, Stockholm, Sweden, Vol 1, 178-182.

Scherer, K. R. (1990). *Emotional Expression in Speech and Music*. London: MacMillan.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice , (9)*, 234-49.

Taylor, A. P. (1992). A Phonetic model of English intonation: *A thesis submitted for the degree of Doctor of Philosophy,* UK: University of Edinburg.

Taylor, A. P. (1993). Acoustic recognition of Intonation from Fo Contours using the rise/fall connection model. *In Processing of Eurospeech*, Berlin, 164-178.

Taylor, A. P. (1995). The rise/fall Connection model of Intonation. *Speech Communication, 15*,168-187.

Taylor, A. P. (1998). The tilt Intonation model. *Centre for speech Technology Research.* University of Edinburg, UK.

Taylor, A. P. (2000). Analysis and Synthesis of Intonation using the tilt model. *Journal of the Acoustical Society of America. 107*(3), 1703-1718.

Taylor, A. P. (2009). *Text to Speech synthesis*. Cambridge: Cambridge University Press.

Tickle, A. (1990), Cross-Language Vocalization of emotion, methodological issues. *Proceedings of ICPHS 99*. San Francisco, USA, 305-308.

Veisy, E. (1386/2007). Bæræsi-ye rabbet-ye beyn-e vizhegihay-e avashenakhti væ bæyane ævatef dær farsi: motale?-ye mowredi. (Investigating the relationship between applied acoustic features and the expression of emotions in Persian: A case study)", In *the 7th Linguistics Conference, Iran,* 143-165.

Uldall, E.(1964). *Dimensions of meaning in intonation*. London:Longman

Williams, S.C.E, and Stevens, K.N. (1972). Emotion and Speech: Some acoustical factors. *Journal of the Acoustical Society of America*. 52, 1240-1252.

Appendix

Table 4. *Statistics of the Total Significant Correlation Between the Pitch Range and the Intended Emotions in all Given Samples*

Utterance	Peak	Half	Sex	Emotion	Neutral	Variance of Coefficient	Standardized Coefficient	Significance Level
1	1	1	Female	Joy	127.567	207.657	.738	.015
1	1	2	Female	Joy	29.133	280.567	.906	.007
1	2	1	Female	Joy	44.481	198.219	.932	.003
1	2	1	Male	Anger	16.250	62.350	.909	.006
1	2	2	Female	Joy	63.867	180.867	.742	.020
1	2	2	Female	Anger	63.867	158.467	.650	.035
1	2	2	Male	Anger	23.100	82.333	.765	.023
1	3	1	Male	Anger	19.450	79.083	.988	.000
1	3	2	Male	Anger	21.800	31.667	.740	.027
2	1	1	Female	Joy	45.443	137.157	.809	.003
2	1	1	Female	Anger	45.443	152.190	.897	.002
2	1	1	Male	Anger	44.115	53.852	.782	.046
2	1	2	Female	Joy	26.467	153.067	.943	.000
2	1	2	Female	Anger	26.467	123.283	.760	.001
2	1	2	Male	Anger	23.067	47.000	.835	.028
2	2	1	Female	Sadness	35.199	77.035	.523	.033
2	2	1	Female	Joy	35.199	156.168	1.061	.001
2	2	1	Female	Anger	35.199	103.068	.700	.009
2	3	1	Female	Joy	25.667	154.600	.724	.039
2	3	1	Female	Anger	25.667	175.567	.822	.023
2	3	2	Female	Joy	22.867	139.833	.731	.034
2	3	2	Female	Anger	22.867	161.833	.846	.018
2	3	2	Male	Anger	20.633	71.100	.896	.008
3	1	1	Female	Joy	10.100	55.433	.915	.005
3	1	1	Male	Joy	6.733	24.133	.715	.012
3	1	1	Male	Anger	6.733	32.600	.966	.002

3	2	2	Female	Joy	29.433	136.467	.738	.047
3	3	1	Female	Sadness	9.150	96.950	.430	.037
3	3	1	Female	Joy	9.150	241.783	1.073	.000
3	3	2	Female	Joy	33.933	263.233	1.069	.000
4	1	1	Female	Joy	83.850	200.417	.933	.000
4	1	1	Male	Anger	30.405	31.295	.687	.025
4	1	2	Female	Joy	35.467	221.800	.926	.010
4	1	2	Male	Anger	22.567	67.933	1.040	.001
4	2	1	Female	Joy	69.788	142.645	.772	.042
4	3	1	Female	Sadness	32.350	81.750	.623	.030
4	3	1	Female	Joy	32.350	133.150	1.014	.003
4	3	2	Female	Joy	16.933	155.000	.995	.004
5	1	1	Female	Joy	21.900	118.967	.998	.001
5	1	1	Male	Joy	36.343	65.120	.991	.000
5	2	1	Female	Joy	38.678	101.089	.918	.009
5	2	2	Female	Joy	20.500	124.800	1.042	.001
5	2	2	Female	Anger	20.500	77.300	.646	.011
5	3	1	Female	Joy	25.367	89.567	.914	.005
5	3	2	Female	Joy	20.733	73.567	.946	.003
5	3	2	Female	Anger	20.733	57.467	.739	.010