Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

The effect of age and gender on the acoustic analysis of anxious sound

Turgut Özseven [1, *], Muharrem Düğenci [2], Ali Doruk [3]

[1]Turhal Vocational School, Gaziosmanpaşa Univerity, Tokat, Turkey
[2]Departmant of Industrial Engineering, Karabük University, Karabük, Turkey
[3]Psychiatry, Gülhane Military Medical Academy, Ankara, Turkey

ARTICLE INFO

Article history:
Received 13 August 2016
Received in revised form
14 November 2016
Accepted 20 November 2016

Keywords:
Anxiety
Speech processing
Acoustic
Age and gender

ABSTRACT

The aim of this study is to investigate the effects of age and gender in sound reflection of anxiety with acoustic analysis. In the study, 148 speech records that express the emotions of the actors as anxiety and neutral were used as the data set. PRAAT software is used for acoustic analysis. The ANOVA method was used to analyze the data. The according to the results of statistical analysis, gender and age increased the count of acoustic parameters that affected of anxiety. The standard deviation of F0 increased too much, jitter local and jitter rap increased mid-range and other parameters did not change when examined changes based gender. The mean of F0, shimmer apq3 and number of unvoiced frame decreased to mid-range, the standard deviation of F0 and jitter local increased too much, the standard deviation of F3 and jitter rap increased to mid-range and other parameters did not change when examined changes based age. The changes occurring in emotions cause changes in sound by affecting respiratory and muscle tension. The anxiety has been changed according to gender and age because the number of parameters in the analysis based on the gender and age is more. The gender causes change in the speed of glottic cycle and this change increases with anxiety. In addition, vocal cords by both male and female occur irregularities and this case also differs according to age. The irregularities in intensity of sound in lower ages are being further while the pauses in the conversation with advancing age are increasing.

1. Introduction

Anxiety emerges in cases the threat or danger and people are prepared to take action in case of danger. However, if anxiety is severe than expected and takes longer than necessary and mind is constantly busy and disrupts people's daily functions, pathological anxiety is concerned. The somatic symptoms related to the respiratory system that occurs in the case of anxiety affects a person's sound system.

The objective and subjective methods are used for psychological diagnosis of the human voice, emotion recognition and identity person. Acoustic analysis is used for objective assessment of the sound with various parameters were obtained from the sound.

When examined the data used in studies on anxiety, available databases (Murray and Arnott, 1993; Ruiz et al., 1996) as well as studies using sound recordings collected by investigators (Protopapas et al., 1997; Drioli et al., 2003; Hagenaars and Minnen, 2005; Fuller et al., 1992; Diamond et al., 2010; Goberman et al., 2011; Weeks et al., 2012) are also available.

According to the results obtained in the studies, average value of F0 is increased in general in case of anxiety (Murray and Arnott, 1993; Ruiz et al., 1996; Drioli et al., 2003; Diamond et al., 2010; Goberman et al., 2011; Weeks et al, 2012; Ververidis and Kotropoulos, 2006). According to the study of non-verbal parts of speech, average value of F0 is decreased (Laukka et al., 2008). When the variability of F0 examined, the studies detected decrease (Hagenaars and Minnen, 2005) as well as increase (Goberman et al., 2011) are available. In case of anxiety, the speech time (Ververidis and Kotropoulos, 2006) and pause rate (Goberman et al., 2011; Laukka et al., 2008) are decreased, speech rate is very increased (Murray and Arnott, 1993).

* Corresponding Author.
Email Address: turgutozseven@gmail.com (T. Özseven)
https://doi.org/10.21833/ijaas.2016.12.003

According to the study of the speech energy value, it varies (normal, high, and low in the non-verbal part) (Murray and Arnott, 1993; Drioli et al., 2003; Laukka et al., 2008). The gender and age causes change in vocal cords the effective on the voice formation (i.e. F0 value) (Gerçeker et al., 2000). Accordingly, the various speech parameters such as speech rate, pause rate, speech energy varies by gender and age (Özseven and Düğenci, 2016). In a study of emotion recognition, seven different moods are used and determined that more reflect the emotions to speech the Middle Ages than young people, the male than female in all emotions (Paulmann et al., 2008). In addition, it has been concluded that vary according to the type of emotion of success emotion recognition (Paulmann et al., 2008). This information shows that gender and age cause changes on all emotions. However, this information does not provide information about whether difference between the effected rate from gender and age of emotions.

The changes according to both moods of speech-related features were investigated based on gender and age using speech records in the case of anxiety and neutral. In addition, the affected rate the anxiety than neutral of gender and age was investigated. The acoustic analysis and statistical methods were used to investigate the relationship between anxieties of the effect.

2. Materials and methods

In this paper, 148 speech expressions contain anxiety and neutral emotions in EMO-DB database are used. Berlin Database of Emotional Speech (EMO-DB) was obtained by expression by actor's different emotions. Speech records are 16 bit mono and 16 kHz sampling frequency (Burkhardt et al., 2005). The distribution of the used data is given in Table 1.

Table 1: The distribution of used data

	Male	Female	Number of Speakers	Number of Sentence	20-30 Age	31-35 Age
Neutral	39	40	10	10	29	50
Anxiety	36	33	10	10	19	50

PRAAT (Boersma, 2002) software used for acoustic analysis and 17 acoustic parameters are obtained of each speech record. Acoustic parameters used in this study are given in Table 2.

Table 2: Acoustic parameters used in this study

Acoustic Parameters	Mean	Standard Deviation
F0, F1, F2, F3	✓	✓
Jitter (Local, Rap), Shimmer (Local, apq3)	✓	
HNR	✓	
Unvoiced frame, Voiced break	✓	
Energy	✓	✓

F0 is defined as the number of opening and closing per second of glottis and the number of vibrations per second of the vocal cords (Sarımehmetoğlu, 2012). F0 increases when tension and thickness of vocal cord is constant and subglottic pressure increased (Sarımehmetoğlu, 2012). F0 decreases with increasing age because it is inversely proportional to the mass of the vocal cords (Akçam et al., 2004; Çevik, 1999; Franz and Aharinejad, 1994). The change of F0 is related to change of glottic cycle (Sarıca, 2012). The value of F0 is between 100-150 Hz and 150-250 Hz respectively in adult male and female while an average of 220-240 Hz in pre-adolescent girls and boys (Sarıca, 2012). Jitter and shimmer is commonly used in speech analysis, and these values are increased in patients with voice disorders (Göksel and Topaloğlu, 2009). The irregular closure and asymmetric vibrations in the vocal cords are evaluated by jitter. Shimmer used to examine irregularities and changes in sound intensity (Sarımehmetoğlu, 2012). Unvoiced frame includes a ratio to the entire speech time of the time waited in silent during people speaking. To increase the ratio of silent stand of mood disorders is expected. Voiced break is the number of pause in speech. Formant (F1, F2, and F3) resonance in vocal tract and it provides information about the quantitative properties of vocal tract. HNR is ratio noise energy of the total energy of the harmonics which F0 and its folds.

Statistical methods: Testing for normality is made according to value of skewness and kurtosis. If these values between -1.5 and +1.5, distribution of data is considered normal (Tabachnick and Fidell, 2012). The ANOVA was used to determine whether this is effective in distinguishing of neutral and anxiety emotions of the acoustic parameters with normal distribution. Statistical analyzes were performed on the IBM Statistics 20 software and 95% confidence level ($p < 0.05$) was used.

3. Results

The normality test and ANOVA results are given in Table 3. According to Table 3, standard deviation of F0, mean of F1 and standard deviation of energy do not have a normal distribution. These parameters were tried to normalize with logarithmic transformation. However, standard deviation of energy is removed from the analysis for providing normal distribution after normalization.

Table 3: The results of statistical analysis

Acoustic Parameters	Skewness	Kurtosis	Emotion	Emotion+Age	Emotion+Gender
			p (ANOVA)		
F0 mean	0.634	-0.231	0.000*	0.000*	0.000*
F0 standard deviation	1.210	1.832	0.055	0.000*	0.000*
F1 mean	0.957	2.595	0.000*	0.000*	0.000*
F2 mean	0.435	0.434	0.000*	0.000*	0.000*
F3 mean	0.459	0.355	0.003*	0.000*	0.000*
F1 standard deviation	0.151	-0.197	0.211	0.000*	0.000*
F2 standard deviation	-0.016	-0.202	0.000*	0.000*	0.000*
F3 standard deviation	-0.139	-0.274	0.054	0.000*	0.000*
Jitter Local	0.404	-0.675	0.505	0.000*	0.000*
Jitter Rap	0.871	0.079	0.000*	0.000*	0.000*
Shimmer Local	0.123	-0.215	0.405	0.113	0.000*
Shimme apq3	0.465	-0.020	0.618	0.210	0.000*
HNR	0.102	-0.370	0.071	0.039*	0.000*
Unvoiced Frame	0.186	0.442	0.475	0.166	0.003*
Voiced Break	0.730	-0.047	0.681	0.835	0.871
Energy mean	-0.024	-0.077	0.105	0.001*	0.000*
Energy standard deviation	8.934	91.289			

*p<0.05

The analyses were performed on three groups (emotion, emotion+age and emotion+gender). The average value of the acoustic parameters was used to examine the changes in anxiety based on gender and age. The change was detected compared to neutral emotion. The gender-based and age-based changes in the acoustic parameters are given in Table 4 and Table 5.

Table 4: The gender-based changes in the acoustic parameters

Acoustic Features	Male			Female			Difference of Change
	Neutral	Anxiety	Change	Neutral	Anxiety	Change	
F0 mean	115.88	170.58	⇧	186.07	270.13	⇧	⇔
F0 standard deviation	27.92	37.15	⇧	43.64	40.93	⇔	⇧
F1 mean	779.85	878.72	⇗	707.50	758.21	⇔	⇔
F2 mean	1919.05	2038.77	⇔	1916.28	1949.8	⇔	⇔
F3 mean	2989.40	3094.98	⇔	3032.12	3019.09	⇔	⇔
F1 standard deviation	511.71	538.43	⇔	487.27	481.39	⇔	⇔
F2 standard deviation	605.03	574.27	⇔	556.42	479.49	⇘	⇔
F3 standard deviation	514.22	499.94	⇔	432.34	397.51	⇔	⇔
Jitter Local	1.98	2.05	⇔	1.20	0.95	⇘	⇗
Jitter Rap	0.52	0.37	⇩	0.21	0.12	⇩	⇗
HNR	9.92	9.45	⇔	12.82	12.24	⇔	⇔
Energy mean	75.71	76.10	⇔	77.81	76.37	⇔	⇔

⇩: very decrease, ⇘: decrease, ⇧: very increase, ⇗: increase, ⇔: no change

According to Table 4, acoustic parameters vary according to gender for both anxiety and neutral. The change column in the table shows the rate and direction of change in the parameters when switching anxiety emotion from neutral emotion. The difference of change column in the table shows the rate and direction between changes. In the case of anxiety for male; mean of F0 and standard deviation of F0 increased too, jitter rap much decreased, mean of F1 increased and there is no change in the other parameters. In the case of anxiety for female; mean of F0 very increased, jitter rap very decreased, standard deviation of F2 decreased, jitter local decreased and there is no change in the other parameters. When gender-based difference of change examined; standard deviation of F0 very increased, jitter local and rap increased and there is no change in other parameters.

According to Table 5, acoustic parameters vary according to age for both anxiety and neutral. The mean of F0 and standard deviation of F0 very increased, jitter local increased, jitter rap decreased and there is no change in the other parameters for anxiety and 20-30 age range. The mean of F0 very increased, mean of F1 decreased, standard deviation of F2 decreased, jitter local decreased, jitter rap very decreased and there is no change in the other parameters for anxiety and 31-35 age range. When age-based difference of change examined; mean of F0, shimmer apq3 and unvoiced frame decreased; standard deviation of F0 and jitter local very increased; standard deviation of F3 and jitter rap increased and there is no change in other parameters.

4. Discussion

The human voice occurs the result of changing with larynx structure and mouth movements of air pressure discharged from the lungs. Acoustic analysis used for the objective evaluation of voice disorders by various parameters were obtained

from the voice. Anxiety emerges in cases the threat or danger and people are prepared to take action in case of danger.

Table 5: The age-based changes in the acoustic parameters

Acoustic Features	20-30 Age			31-35 Age			Difference of Change
	Neutral	Anxiety	Change	Neutral	Anxiety	Change	
F0 mean	126.72	167.53	⇧	165.75	237.44	⇧	⇘
F0 standard deviation	25.59	41.99	⇧	41.85	37.80	⇔	⇧
F1 mean	788.60	849.06	⇔	716.89	810.45	⇘	⇔
F2 mean	1980.03	2028.72	⇔	1881.47	1983.87	⇔	⇔
F3 mean	3074.46	3084.40	⇔	2974.24	3048.91	⇔	⇔
F1 standard deviation	523.76	545.23	⇔	485.17	498.20	⇔	⇔
F2 standard deviation	589.97	562.63	⇔	574.87	516.13	⇘	⇔
F3 standard deviation	491.41	531.54	⇔	461.94	420.32	⇔	⇗
Jitter Local	1.77	2.07	⇗	1.47	1.31	⇘	⇧
Jitter Rap	0.45	0.35	⇘	0.31	0.21	⇩	⇗
Shimmer Local	12.14	12.69	⇔	11.35	11.67	⇔	⇔
Shimmer apq3	4.16	3.95	⇔	3.69	3.94	⇔	⇘
HNR	10.80	10.36	⇔	11.74	10.95	⇔	⇔
Unvoiced Frame	37.81	34.71	⇔	34.65	37.53	⇔	⇘
Energy mean	75.71	76.66	⇔	77.39	76.06	⇔	⇔

⇩: very decrease, ⇘: decrease, ⇧: very increase, ⇗: increase, ⇔: no change

In this paper, the effects on the gender and age of anxiety were examined by acoustic parameters. The severity of anxiety has been ignored on the study. Therefore, the limitation of the study cannot rule out the effect of the severity of anxiety on the changes in gender-related. Because, the severity of anxiety in female is usually more.

According to the statistical analysis, the number of parameters increases on anxiety when gender and age be included. This result shows that vary according to gender and age of anxiety. In both of the results of gender and age based; F0, F3, jitter, shimmer and unvoiced frame showed change.

According to the results, the change of F0 shows that change according to the gender of the glottic cycle, and this value increases case of anxiety than neutral. In addition, the value of mean of F0 is high than the reference values set for adults in anxiety. Jitter local decreased in female while don't change for male and this change is more than neutral in anxiety. Jitter rap is decreased too much for both male and female, and the difference of change by gender-based is a medium level. This state of the jitter parameter shows that the more irregular of vocal cords of both male and female in anxiety. However, these irregularities moderate rate varies by gender. Jitter local showed decreasing on 31-35 age range while increasing on 20-30 age range. This change is more than neutral in anxiety. Jitter rap showed decrease on both 20-30 and 31-35 age range. This result is an indication of different of irregularities in the vocal cords by age. The change of shimmer and unvoiced frame is very low for neutral and anxiety. However, pause rate shows the changes by age group. The elderly people more pausing in anxiety while the young people more pausing in neutral. The volume of the young people is slightly higher, but not much differ by age in anxiety while the volume of the elderly people is high in neutral.

References

Akçam T, Bolu E, Merati AL, Durmus C, Gerek M and Ozkaptan Y (2004). Voice changes after androgen therapy for hypogonadotrophic hypogonadism. The Laryngoscop, 114(9): 1587–1591.

Boersma P (2002). PRAAT, a system for doing phonetics by computer. Glot International, 5(9/10): 341–345.

Burkhardt F, Paeschke A, Rolfes M, Sendlmeier WF and Weiss B (2005). A database of German emotional speech. In the Ninth European Conference on Speech Communication and Technology (Interspeech '05), Lisbon, Portugal: 1517–1520.

Çevik S (1999). Koro Eğitimi Yönetimi ve Teknikleri. 2nd Edition, Yurt Renkleri Yayınevi, Ankara, Turkey.

Diamond GM, Rochman D and Amir O (2010). Arousing primary vulnerable emotions in the context of unresolved anger:'Speaking about' versus 'speaking to'. Journal of Counseling Psychology, 57(4): 402-410.

Drioli C, Tisato G, Cosi P and Tesser F (2003). Emotions and voice quality: experiments with sinusoidal modeling. In the Voice Quality: Functions, Analysis and Synthesis (VOQUAL'03), Geneva, Switzerland: 127-132.

Franz P and Aharinejad S (1994). The microvasculature of the larynx: a scanning electron microscopic study. Scanning Microscopy, 8(1): 125–30.

Fuller BF, Horii Y and Conner DA (1992). Validity and reliability of nonverbal voice measures as indicators of stressor-provoked anxiety. Research in Nursing & Health, 15(5): 379–389.

Gerçeker M, Yorulmaz İ and Ural A (2000). Ses ve konuşma. K.B.B. Ve Baş Boyun Cerrahisi Dergisi, 8(1):71-78.

Goberman AM, Hughes S and Haydock T (2011). Acoustic characteristics of public speaking: Anxiety and practice effects. Speech Communication, 53(6): 867–876.

Göksel AO and Topaloğlu İ (2009). Endolarengeal mikrocerrahi uygulanan hastalarda ses kalitesinin akustik ve spektrografik analiz ile değerlendirilmesi. The Turkish Journal of Ear Nose and Throat, 19(5): 253–258.

Hagenaars MA and Minnen A (2005). The effect of fear on paralinguistic aspects of speech in patients with panic disorder with agoraphobia. Journal of Anxiety Disorders, 19(5): 521–537.

Laukka P, Linnman C, AAhs F, Pissiota A, Frans Ö, Faria V, Michelg A, Appel L, Fredrikson M and Furmark T (2008). In a nervous voice: Acoustic analysis and perception of anxiety in social phobics speech. Journal of Nonverbal Behavior, 32(4): 195–214.

Murray IR and Arnott JL (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. The Journal of the Acoustical Society of America, 93(2): 1097–1108.

Özseven T and Düğenci M (2016). Sayısal filtrelerin akustik parametreler, cinsiyet, yaş ve duygu durumu üzerindeki etkileri. Pamukkale University Journal of Engineering Sciences, PAJES-00922. (In Press) https://doi.org/10.5505/pajes.2016.00922

Paulmann S, Marc DP and Sonja AK (2008). How aging affects the recognition of emotional speech. Brain and Language, 104(3): 262-269.

Protopapas A and Lieberman P (1997). Fundamental frequency of phonation and perceived emotional stress. The Journal of the Acoustical Society of America, 101(4): 2267–2277.

Ruiz R, Absil E, Harmegnies B, Legros C and Poch D (1996). Time-and spectrum-related variabilities in stressed speech under laboratory and real conditions. Speech Communication, 20(1): 111–129.

Sarıca S (2012). Ses Analizinde Kullanılan Akustik Parametreler. Ph.D Dissertation, University of Kahramanmaraş Sütçü İmam, Kahramanmaraş, Turkey.

Sarımehmetoğlu EA (2012). Vokal Nodülü Olan Pediatrik Olgularda Ses Kalitesinin Akustik Ses Analizi ve 'Pediatric Voice Handicap Index İle Değerlendirilmesi. M.Sc. Thesis, University of Gazi, Ankara, Turkey.

Tabachnick BG and Fidell LS (2012). Using Multivariate Statistics, 6th Edition, Pearson, Boston, USA.

Ververidis D and Kotropoulos C (2006). Emotional speech recognition: Resources, features, and methods. Speech Communication, 48(9): 1162–1181.

Weeks JW, Lee CY, Reilly AR, Howell AN, France C, Kowalsky JM and Bush A (2012). The Sound of Fear: Assessing vocal fundamental frequency as a physiological indicator of social anxiety disorder. Journal of Anxiety Disorders, 26(8): 811–822.