The Effects of Noise on Acoustic Parameters

*[1]Turgut Özseven and [2]Muharrem Düğenci
[1]Turhal Vocational School, Gaziosmanpaşa University, Turkey
*[2]Faculty of Engineering, Department of Industrial Engineering Karabük University, Turkey

Abstract

Acoustic parameters are attributes gathered from the environment or from living beings, used for objective evaluation on sound signals. The basic acoustic parameters are pitch, harmonic noise, shimmer and jitter. Noise is defined as unwanted signal which cause disruptions in the signal during applications such as communication, measuring and signal processing. In this study, we have added white noise to a sample conversation and examined the differences before and after added noise in pitch, jitter, shimmer, harmonic noise rate, noise harmony rate, format frequencies (F1, F2, F3) and energy density parameters.

Key words: Acoustic analysis, signal noise, signal processing

1. Introduction

Acoustic is a branch of physics interested in sound and psycho-acoustic is interested in the psychological answers of acoustics. Speaking acoustics include both physical and psychological studies and are interested in the sound's acoustic properties [1]. Since sensory evaluation differs for each person, objective evaluation methods are being used [2]. Acoustic sound analysis is one of the methods used to evaluate sound disruptions objectively [3]. Acoustic analysis is an objective, noninvasive, inexpensive method that provides data in a short time and there are free software programs that can be used for analysis [4]. Acoustic analysis can be used not only for human voice, but also for acoustic analysis in different environments such as machines, vehicles and under water. The basic acoustic parameters are pitch, harmonic noise, shimmer and jitter. Pitch is the number of cycles tone curves opening and closing in a second. It reports the depth of the sound. Jitter is the parameter showing the difference between periods. Shimmer is the periodic variation between the amplitude peaks.

Noise is defined as unwanted signal which cause disruptions in the signal during applications such as communication, measuring and signal processing [5]. Noise lowers the system's performance in systems including conversation analysis. There are several noise types based on the source of the noise and the spectral structure of the signal [5]. Of these, white noise is defined as a random noise process that has equal force on all frequencies [5, 6]. It is a random noise with a flat power spectrum. Theoretically it includes all frequencies in equal power [7].

In this study, we have added white noise to a sample conversation and examined the differences before and after added noise in pitch, jitter, shimmer, harmonic noise rate, noise harmony rate,

*Corresponding author: Address: Turhal Vocational School, Gaziosmanpaşa University, 60300, Tokat TURKEY. E-mail address: turgut.ozseven@gop.edu.tr, Phone: +903562757839 Fax: +903562761448

format frequencies (F1, F2, F3) and energy density parameters. The analyses have been conducted through the PRAAT software [8].

Pitch is the number of vibrations tone curves per second. The time between two vibrations is called a period. It reports the depth of the sound. It is measured as cycles/second and defined with Hertz (Hz). This value is about 220-240 Hz for pre-adolescence girls and boys, while it is around 100-150 Hz and 150-250 Hz for adult men and women, respectively [9].

Jitter is the parameter showing the difference between periods. Periods include the unwanted irregularities [2]. The differences in the jitter show the pathology of tone curves.

Shimmer is the periodic variation between the amplitude peaks. It shows the relative change in short intervals between the amplitude of the sound wave.

Harmonic-to-Noise Ratio (HNR) are parameters located in the sound spectrum other than the harmonics which occur at multiples of the fundamental frequency. In a complex sound, the integer multiples of basic frequencies create harmonics. HNR is the rate of the total energy of basic harmonic frequencies and its multiples to the noise energy. The value is in dB.

Noise-to-Harmonic Ratio (NHR) is obtained through the modification of the HNR parameter. The value changes in accordance with the noise in the sound, unlike HNR.

Format is the resonance of the audio path. There is an infinite number of formats in theory, however in practicality only 3 or 4 formants include important information. The formants are defined with numbers such as F1, F2 and F3. F1 includes high pitched frequencies such as “ee” and “u”, and low sounds such as “a” and “ae” are in the higher frequencies of this formant. So the F1 frequency is inversely proportional to the language stroke height language stroke height of vowels. On the other hand deep vowels such as “a” and “u” share the lower frequencies of F2 and “ee” and “ae” high vowels share the higher frequencies of F2. So, the vocal articulation of F2 shows differences in a posterior-anterior dimension [1].

Material and methods are discussed in the second part of the study and the experimental results are given in the third part.

2. Materials and Method

The sample sound record was taken from the Berlin Database of Emotional Speech [10]. The recordings have a 16 kHz sample frequency. The time-amplitude image of the sound recording is given in Figure 1.

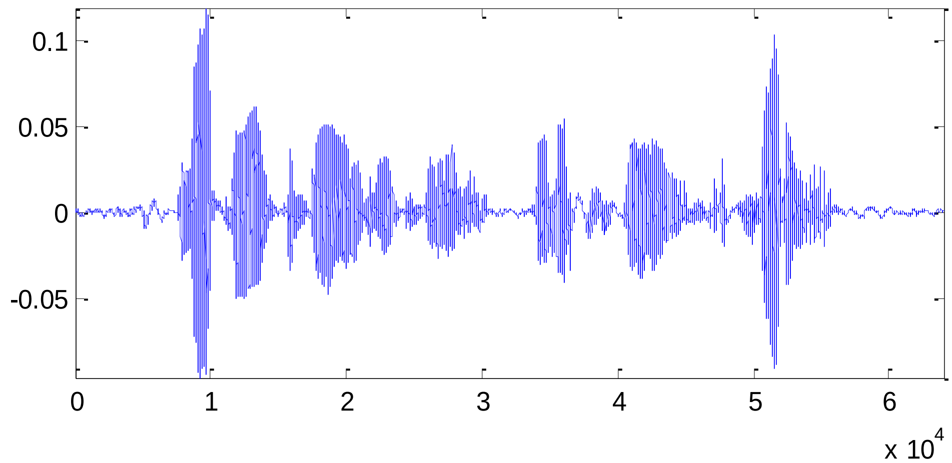

Figure 1. Example speech

Four different white Gaussian noises with, respectively, 1dB, 5dB, 10dB and 15dB SNR values were added through MATLAB on the sample recording.

The PRAAT software has been designed and developed by Paul Boersma and David Weenink of the Amsterdam University and can be used for free in acoustic sound analysis. This software is especially designed for phonetic analysis and is completely free.

Acoustic analysis was done with the PRAAT software and the difference in pitch, jitter, shimmer, harmonic noise rate, noise harmony rate, format frequencies (F1, F2, F3) and energy density parameters and how these change according to the noise have been examined. 25 ms was used as window width and Hamming was used as a window to extract the acoustic properties. The acoustic parameters used were given on Table 1.

Table 1. The Acoustic Parameters Used [8]

Pitch (Hz)	
	Min F0, Max F0, Mean F0, Median F0, Standard Deviation F0
Jitter (%)	
	Local, rap, ppq5, ddp
Shimmer (%)	
	Local, apq3, apq5, apq11, dda
HNR-NHR (dB)	
	Mean HNR, Mean NHR
Formant Frequency (Hz)	
	Mean F1, Mean F2, Mean F3
Energy (dB)	
	Min, Max, Mean

3. Results

The acoustic analysis of the sample record used in the study was done with the PRAAT software and its time-amplitude image after having added 4 different noises is shown in Figure 2, Figure 3, Figure 4 and Figure 5.

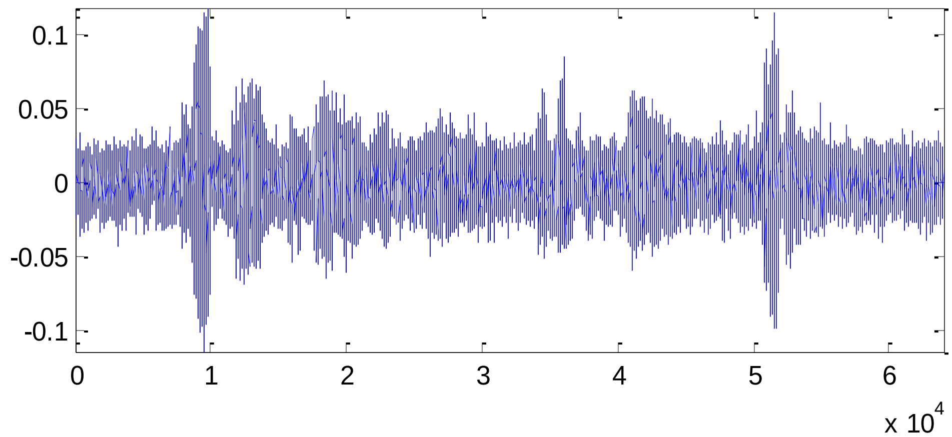

Figure 2. The recording after a 1dB white noise addition

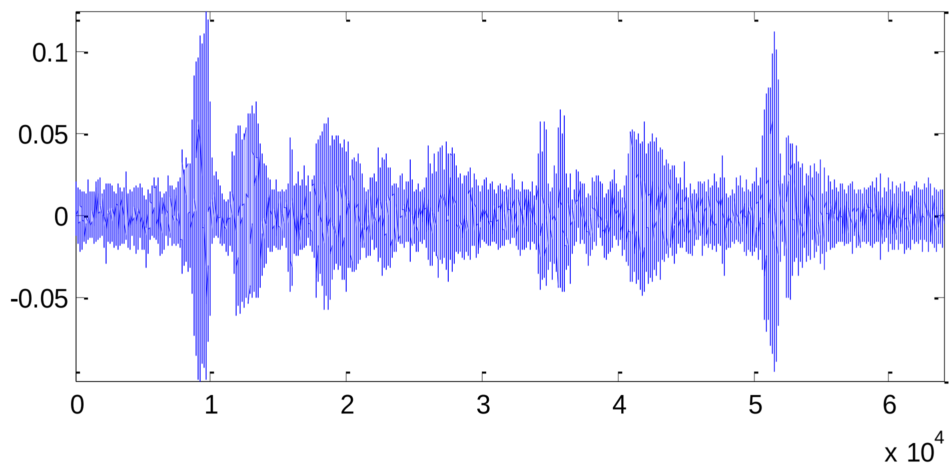

Figure 3. The recording after a 5dB white noise addition

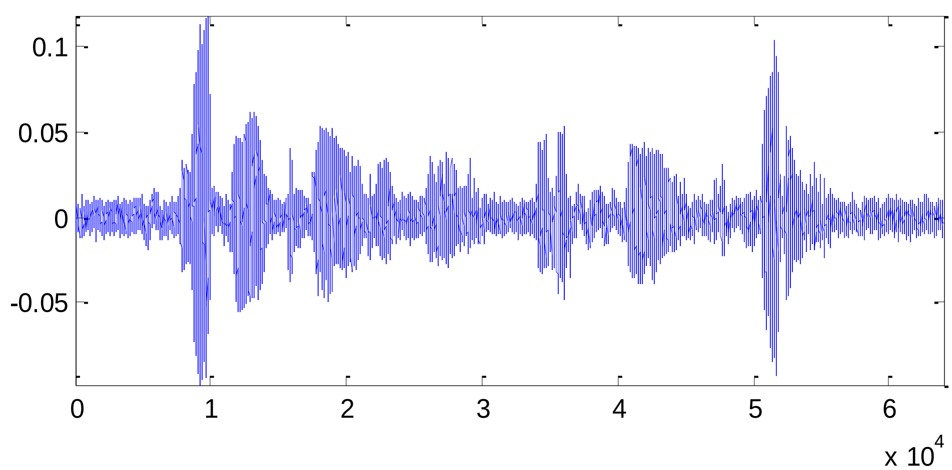

Figure 4. The recording after a 10dB white noise addition

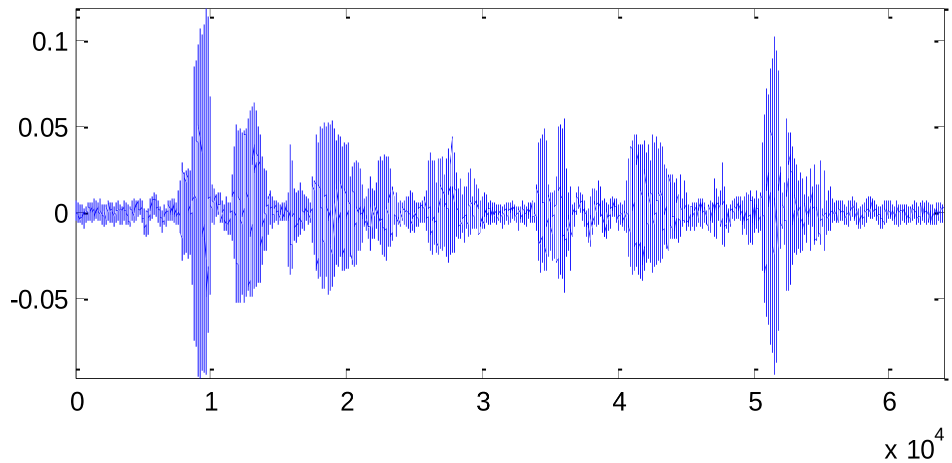

Figure 5. The recording after a 15dB white noise addition

The differences of pitch value according to the noise in the result of the study have been shown in Figure 6.

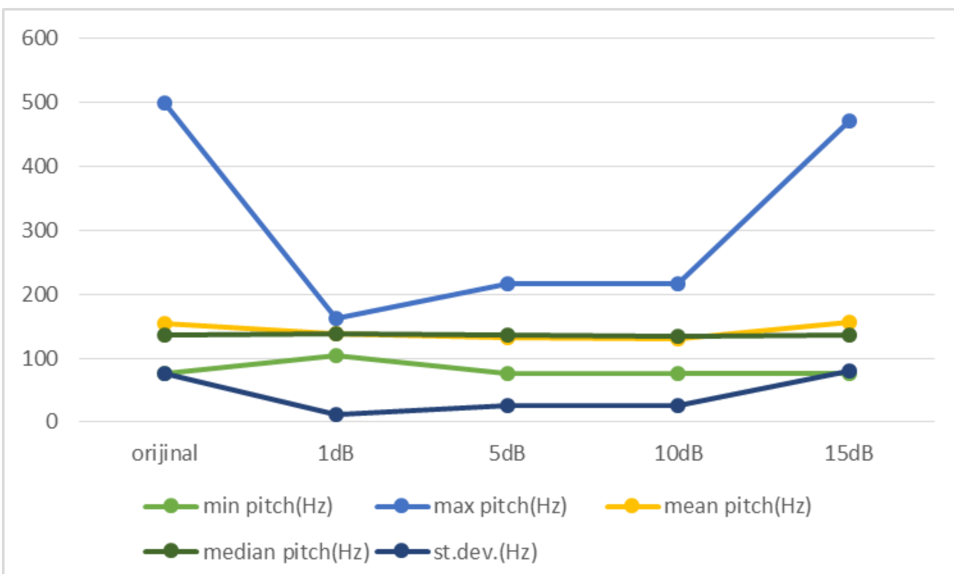

Figure 6. The differences in pitch parameters after the noise addition

When examining the differences in the pitch parameter, we see that there is no significant change in the noise median and mean values. The max and standard deviation is affected severely by high noise. The min value is only affected by very loud noise.

The differences of jitter value according to the noise in the result of the study have been shown in Figure 7.

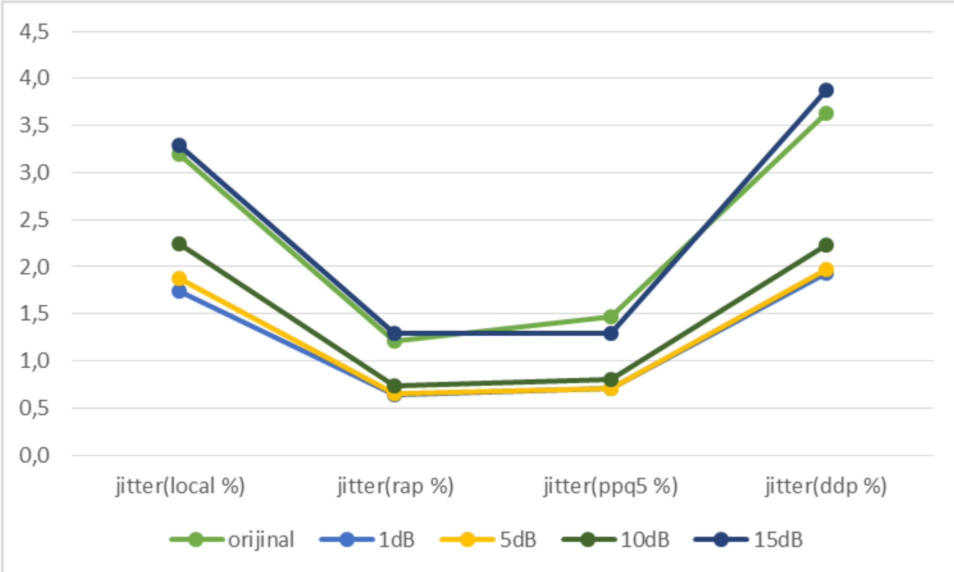

Figure 7. The differences in jitter parameters after the noise addition

According to Figure 7, while the jitter parameters are severely affected by high and medium level noise, we see that low level noise had no significant impact.

The differences of shimmer value according to the noise in the result of the study have been shown in Figure 8.

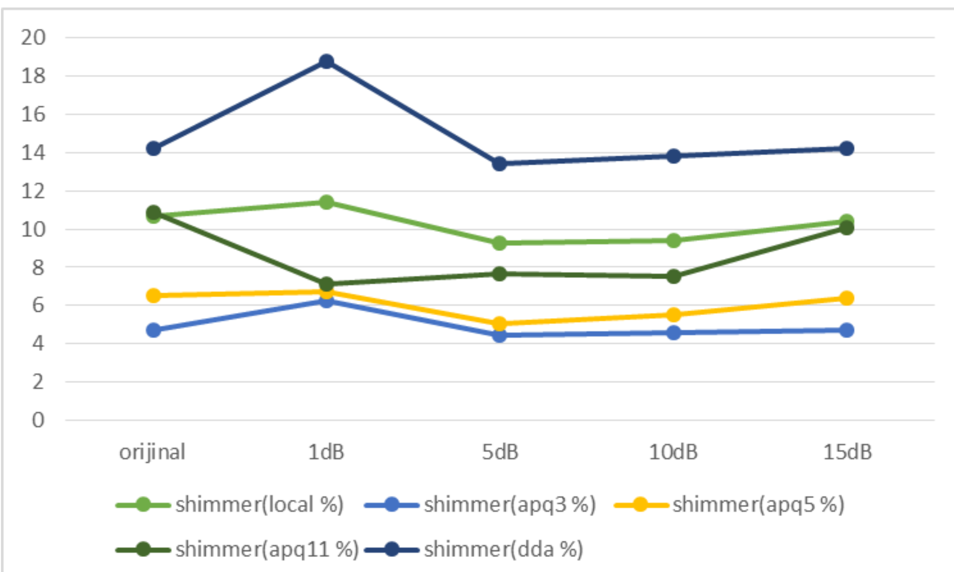

Figure 8. The differences in shimmer parameters after the noise addition

According to Figure 8, we see that the shimmer parameter was generally not significantly affected by noise. However while the shimmer (dda) parameter was not affected by low or high level noise, we see medium level noise affected this parameter.

The differences of HNR-NHR value according to the noise in the result of the study have been shown in Figure 9.

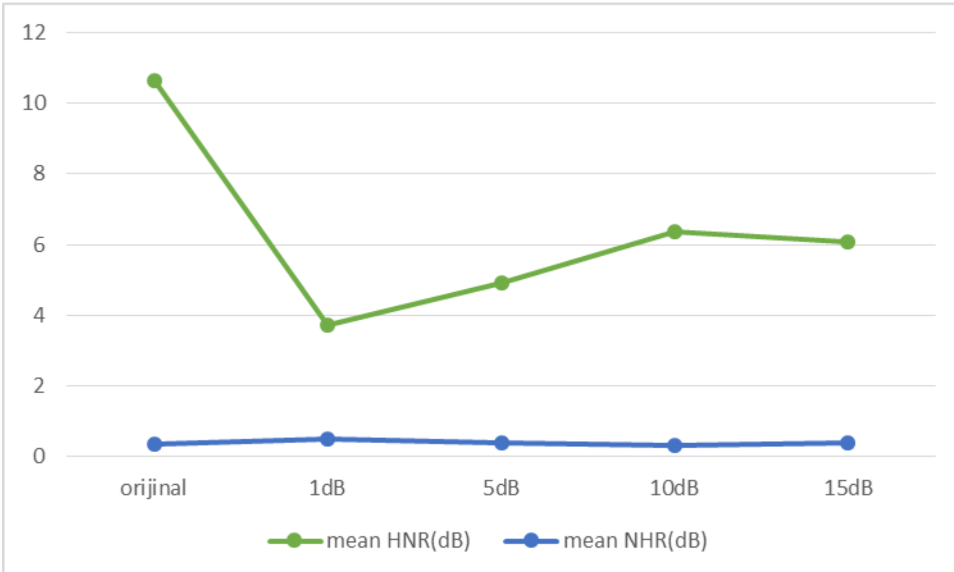

Figure 9. The differences in HNR-NHR parameters after the noise addition

According to Figure 9, the changes in HNR parameters with a voice addition are as expected and low HNR was observed at high noise.

The differences of formant frequency according to the noise in the result of the study have been shown in Figure 10.

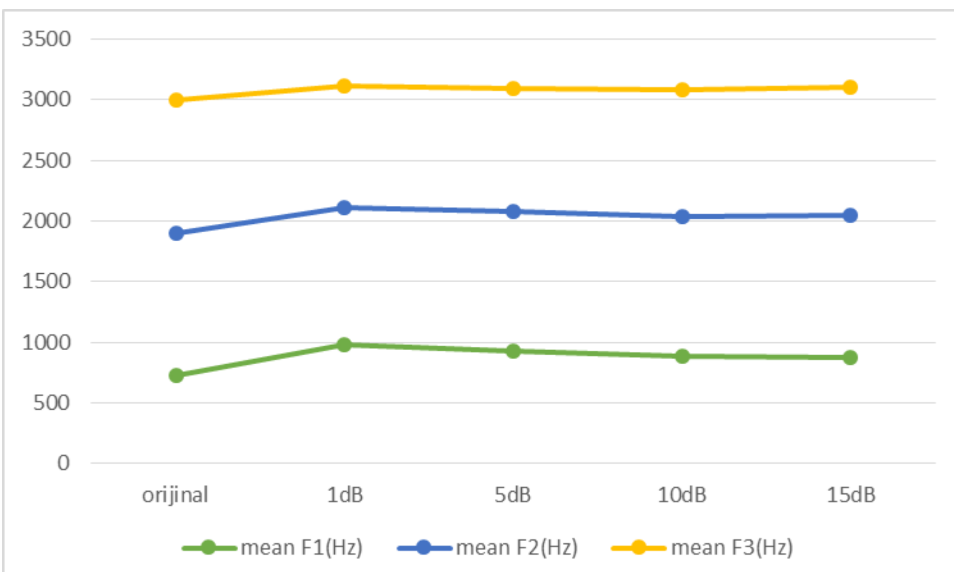

Figure 10. The differences in formant frequencies after the noise addition

According to Figure 10, F1 got affected by noise the most. Other formant frequencies didn't show any significant changes.

The differences of energy value according to the noise in the result of the study have been shown in Figure 11.

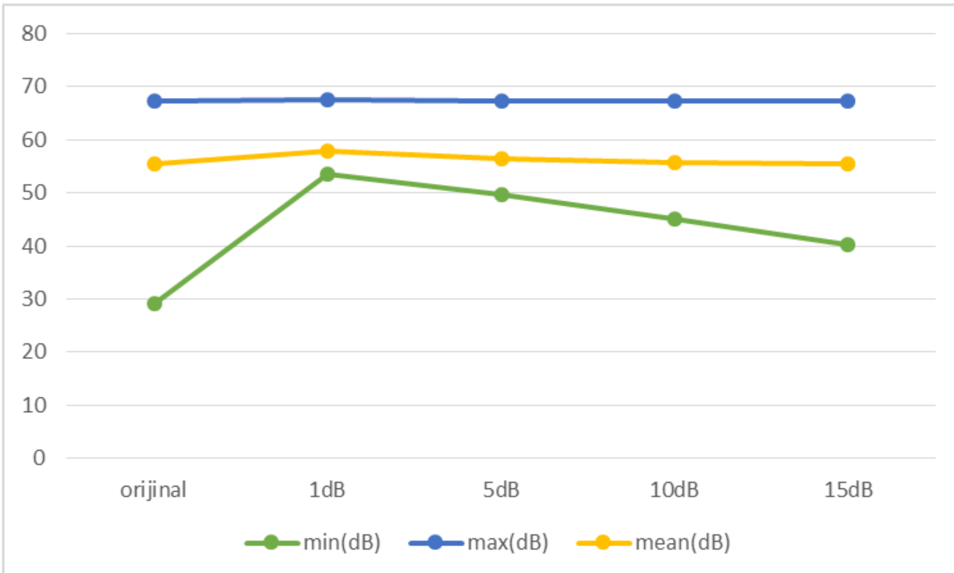

Figure 11. The differences in energy parameters after the noise addition

According to Figure 11, adding noise to the signal only created a difference in the min energy value.

4. Discussion

Even though the acoustic parameter numbers are quintessentially limited, the parameter numbers can be increased through the statistical variations of the basic parameters. In the last 20 years, several filtering methods have been recommended to reduce the noise. However the noise reduction methods shouldn't cause changes in the main signal's structure in order to diminish the noise completely. That's why many methods provide a reduction of 10-20 dB noise [11]. Significant changes in pitch parameters have been seen in max values and changes in SNR values higher than 15dB disappear. In literature, studies including acoustic analyses majorly use pitch parameters. Pitch parameters by themselves aren't affected by low noise and therefore don't require preprocess. This is also valid for jitter, shimmer, formant and energy parameters. If one uses a higher value than 15dB to reduce noise in acoustic analysis, preprocess is not required.

Conclusions

The results of the study have shown that the pitch parameter changes in noises with high min value, whereas noises with high and medium level max, min and std. dev. values are affected as well. For pitch parameters a change in median values based on noise was not observed. It shows no change in noises with low jitter parameters, whereas a difference can be observed in high and medium noises. Significant differences can't be seen in shimmer parameters based on noise except the apq11 value. As expected, the average HNR value drops as the noise increase. As for formant frequencies, we see no significant changes in them except F1. Lastly, when looking at the energy parameters we see only significant changes in the min value.

References

[1] Rezaei N, Abolfazl S. An Introduction to speech sciences (Acoustic Analysis of Speech). Iranian Rehabilitation Journal 2006;5-14: 4(4).

[2] Kılıç MA, Okur E. Comparison of fundamental frequency and perturbation data analyzed by CSL and Dr.Speech systems. Journal of Otolaryngology Specialized 2001;152-157-8.

[3] Özbal EA. Septum devisyonlu hastaların septoplast operasyonu öncesi ve sonrası akustik ses analizi ile değerlendirilmesi. Thesis, Şişli Etfal Hospital, İstanbul, 2008.

[4] Sataloff RT. Treatment of voice disorders. San Diego: Plural Publishing; 2005

[5] Vaseghi SV. Advanced digital signal processing and noise reduction. John Wiley & Sons; 2008.

[6] Elmas E, Güler NF. The audio dimension of Radio and TV broadcasting and the improvement studio acoustic. Journal of Advanced Technology Sciences 2013; 50-61:2(2).

[7] Villanueva-Luna AE, Jaramillo-Nuñez A, Sanchez-Lucero D, Ortiz-Lima CM, Aguilar-Soto JG, Flores-Gil A, May-Alarcon M. De-noising audio signals using MATLAB wavelets toolbox. 2011.

[8] Boersma P, Weenink D. Praat: Doing phonetics by computer. Retrieved from http://www.praat.org/, 2015.

[9] Sarıca S. Acoustic parameters that used in voice analysis. Thesis, Kahramanmaraş Sütçü İmam University Faculty of Medicine, Kahramanmaraş, 2012.

[10] Burkhardt F, Paeschke A, Rolfes M, Sendlmeier WF, Weiss B. A database of German emotional speech. In Interspeech 2005; 1517-1520:25(1).

[11] Virtanen T, Singh R, Raj B. Techniques for noise robustness in automatic speech recognition. John Wiley & Sons; 2012.