Adverse Effect of Noise on Voice Perturbation Estimates: A Comparison of Three Voice Analysis Programs

Seste Bozulmanın Belirlenmesi Üzerine Gürültünün Olumsuz Etkisi: Üç Ses Analiz Programının Karşılaştırılması

Mehmet Akif KILIÇ, MD,[a]
Haldun OĞUZ, MD,[b]
Fatih ÖĞÜT, MD[c]

[a]Department of Otolaryngology, Kahramanmaraş Sütçü İmam University Faculty of Medicine, Kahramanmaraş
[b]Department of Otolaryngology, Ankara Training and Research Hospital, Ankara
[c]Department of Otolaryngology, Ege University Faculty of Medicine, İzmir

Geliş Tarihi/*Received:* 26.06.2010
Kabul Tarihi/*Accepted:* 14.12.2010

This paper was presented at the 3rd World Voice Congress in İstanbul, 2006.

Yazışma Adresi/*Correspondence:*
Haldun OĞUZ, MD
Ankara Training and Research Hospital, Department of Otolaryngology, Ankara,
TÜRKİYE/TURKEY
drhoguz@gmail.com

ABSTRACT Objective: The aim of this study is to investigate the influence of noise on voice perturbation outputs obtained by three different voice analysis programs, and to show the software more immune to noise. **Material and Methods:** For this purpose, 10 natural and 10 semi-synthetic voice samples were recorded. They were mixed with environmental noise and white noise. The unmixed and mixed signals were analyzed by Dr. Speech, MDVP and Praat. Three frequency perturbation outputs (jitter percent, relative average perturbation and pitch perturbation quotient) and two amplitude perturbation outputs (shimmer percent and amplitude perturbation quotient) were obtained. The unmixed and mixed signals were compared. Correlations were calculated between natural unmixed and natural mixed ones. **Results:** The frequency perturbation outputs obtained by Dr. Speech and Praat seemed to be less affected by noise. However, the amplitude perturbation values measured by three systems were severely affected by noise, but Praat's performance found mildly better than the others'. Correlation analyses for 10 items (five parameters x two noise situations) revealed that there were perfect correlations (r= 1.000) regarding six items for Praat, three items for Dr. Speech and one item for MDVP. **Conclusion:** As a conclusion, it was decided that Praat was the most immune one to noise among three analysis systems, followed by Dr. Speech and MDVP, respectively.

Key Words: Voice; voice quality

ÖZET Amaç: Üç farklı ses analiz programınca elde edilen seste bozulma değerlendirmesi üzerine gürültünün etkisini araştırmak ve hangi yazılımın gürültüden daha az etkilendiğini bulmak. **Gereç ve Yöntemler:** Bu amaçla 10 doğal ve yarı-sentetik ses örneği kaydedildi. Bunlar ortam gürültüsü ve beyaz gürültü ile karıştırıldı. Karışmış ve karıştırılmamış sinyaller Dr. Speech, MDVP ve Praat ile analiz edildi. Üç frekans bozulma çıktısı (jitter yüzdesi, rölatif ortalama pertürbasyon ve perde pertürbasyon oranı) ve iki amplitüd bozma çıktısı (shimmer yüzdesi ve amplitüd pertürbasyon oranı) alındı. Karışmış ve karışmamış sinyaller karşılaştırıldı. Karışmamış doğal ve karışmış doğal olanlar arasında korelasyonlar hesaplandı. **Bulgular:** Dr. Speech ve Praat ile elde edilen frekans bozulma çıktıları gürültüden daha az etkilenmiş görüldüler. Bununla birlikte, amplitüd bozulma değerleri üç sistemde de gürültü tarafından ciddi şekilde etkilenmişlerdi fakat Praat'ın performansı diğerlerinden hafifçe daha iyi bulundu. On maddenin korelasyon analizi (5 parametre x 2 gürültü durumu) sonucunda Praat için altı maddede, Dr. Speech için üç maddede ve MDVP için bir maddede mükemmel korelasyon gözlendi (r= 1.000). **Sonuç:** Sonuç olarak üç analiz sistemi arasında gürültüye en bağışık sistem olarak Praat belirlendi, bunu sırasıyla Dr. Speech ve MDVP izliyordu.

Anahtar Kelimeler: Ses; ses özelliği

Turkiye Klinikleri J Med Sci 2011;31(2):427-31

Dysphonic voices are evaluated with voice analysis systems to obtain objective comparable outputs. The reliability of acoustic voice measurements are influenced by many factors such as: (1) the recording system, (2) the microphone, (3) analysis window, (4) fundamental fre-

doi:10.5336/medsci.2010-19461

quency (F0) extraction methods, (5) algorithms for determining voice statistics, (6) environmental noise, (7) and internal computer noise.[1] In this study, we were concerned with the effects of noise (resulting both from environment and data acquisition system) on perturbation estimates from three different programs, and behaviors of them in the event of different noise situations.

In noisy conditions, automatic analysis systems may not be able to distinguish between the signal and extraneous noise. Noise-related variations might be misinterpreted as a part of the voice signal, resulting in errant estimates of perturbation.[1] Although recording the voice sample in a sound-treated booth is ideal for perturbation measurement, it may not be available in every clinic; therefore clinicians have to rely on the analysis results obtained in imperfect environments. A perfect voice analysis system should not be affected by external noise. It is important to determine whether different software and hardware configurations might be more or less vulnerable to the effects of external noise.

The aim of this study is to investigate the influence of noise on voice perturbation outputs obtained by three different voice analysis programs, and to reveal the software more immune to noise.

MATERIAL AND METHODS

Ten male adults, who had no laryngeal pathology, provided the voice samples. They were volunteers from the staff of Sütçü Imam University. Their ages ranged between 20 and 42, with a mean age of 27.3 years.

The stimuli used in this study consisted of six groups of voice samples:

- Natural unmixed (NU) samples: The subjects uttered sustained [a] vowels in a noiseless and anechoic environment, and their utterances were recorded by using an electret condenser microphone (Philips SBC ME400) positioned 15 cm away from the subjects' mouth at a sampling rate of 44.100 Hz and a 16 bit resolution. The intensity of voice samples was 60 dB, at average. For this purpose, a PC with Sound Blaster Audigy 2ZS sound card and Adobe Audition program (Version 1.5) were used. The first segment of 0.5 s was deleted, the next segment of 3 s was saved.

- Semi-synthetic unmixed (SU) samples: These samples were prepared using Adobe Audition software. A sample period was precisely extracted from a steady state portion of each natural unmixed sample, later this period was multiplied to form a stimulus of 3 s duration, thus, generating a perturbation-free voice sample.

- Natural (NP) and semi-synthetic (SP) samples mixed with policlinic noise: The environmental noise of 3 s duration was captured in our ENT policlinic room during working hours by the same microphone and hardware/software equipment. The recorded noise's average root mean square (RMS) power was calculated by Adobe Audition software and found as -51.13 dB. Later, the noise sample was mixed with natural and semi-synthetic samples.

- Natural (NW) and semi-synthetic (SW) samples mixed with white noise: The white noise of 3 s duration was generated using Adobe Audition software, and this noise sample's average RMS power was adjusted to -51.13 dB by decreasing its amplitude. Later, the noise sample was mixed with the natural and semi-synthetic samples.

Perturbation analyses were carried out using three different analysis systems: Dr. Speech Vocal Assessment (Tiger Electronics, Massachusetts, USA), MDVP for Multi-Speech (Kay Elemetrics, NJ, USA), and Praat (Paul Boersma and David Weenink). Three frequency perturbation values (jitter percent-Jitt, relative average perturbation-RAP, pitch perturbation quotient-PPQ) and two amplitude perturbation values (shimmer percent-Shim and amplitude perturbation quotient-APQ) were measured. All three systems were operated on the same computer.

The unmixed and mixed signals were compared by Wilcoxon test. Spearman's Rho correlations were calculated between natural unmixed and natural mixed ones. Statistical significance was defined using p values less than 0.05. All statistical analyses were performed using SPSS software (Version 10.0).

RESULTS

Means and standard deviations for MF0 (mean fundamental frequency), Jitt, RAP, PPQ, Shim and APQ of the NU data for each software are shown in Table 1.

Statistical analyses revealed that the frequency perturbation outputs obtained by Dr. Speech were not affected by noise. Praat seemed to be mildly affected by noise, only one parameter was affected (SU was different from SP for Jitt). The comparison results of the unmixed data with mixed data can be seen in Table 2. The relationship of frequency perturbation values according to Dr.Speech, MDVP and Praat are shown in Figures 1-3, respectively.

The amplitude perturbation values measured by Dr. Speech and MDVP of the natural data were severely affected by noise. However, Praat was more resistant to noise, there was no difference between NU data and NP data for Shim, neither was there a difference between NU data and NW data for APQ. The amplitude perturbation values of semi-synthetic samples seemed to be more severely affected by noise. The comparison results of the unmixed data with mixed data can be seen in Table 3.

Correlation analyses of the natural unmixed and mixed data for 10 items (five parameters × two noise situations) revealed that there were perfect positive correlations (r= 1.000) regarding six items for Praat, three items for Dr. Speech and one item for MDVP. Spearman's Rho correlation coefficients between the NU and NP data, the NU and NW data can be seen in Table 4.

DISCUSSION

In the scientific literature, there are several investigations comparing acoustic voice analysis systems. Some of them compare the results obtained by different voice analysis systems.[2-6] Some of them compare the effects of noise on voice quality estimates.[7,8] The first group of investigations generally focus on the differences that emerge from hardware and analysis protocol such as pitch extraction methods. The second group of studies generally focus on the effect of noise originated from different sources on the same analysis system. However, there are only a few investigations that compare behaviors of the different analysis systems in the event of different noise situations.

Boersma[9] compared performances of Praat and MDVP using five different samples with determined jitter and shimmer values (jitter and shimmer percent with 0.001, 0.01, 0.1, 1 and 10%). He obser-

TABLE 1: Mean values and standard deviations of MF0 and perturbation measures of the NU data for each software (MF0 values are in Hz, others in percent).

	Dr. Speech		MDVP		Praat	
Parameters	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
MF0	134	18	133	18	133	18
Jitt	0.264	0.124	0.385	0.114	0.274	0.115
RAP	0.148	0.074	0.218	0.071	0.145	0.069
PPQ	0.167	0.062	0.222	0.061	0.158	0.066
Shim	1.812	0.465	2.008	0.354	1.938	0.522
APQ	1.580	0.266	1.700	0.260	1.628	0.364

TABLE 2: p values obtained after comparison of frequency perturbation values of the unmixed and mixed data by Wilcoxon test.

	Dr. Speech			MDVP			Praat		
	Jitt	**RAP**	**PPQ**	**Jitt**	**RAP**	**PPQ**	**Jitt**	**RAP**	**PPQ**
NU-NP	0.285	0.317	0.317	0.086	0.020	0.041	0.317	0.317	0.317
NU-NW	1.000	0.317	1.000	0.057	0.066	0.013	1.000	1.000	1.000
SU-SP	1.000	1.000	1.000	0.005	0.005	0.005	0.008	0.317	0.317
SU-SW	1.000	1.000	1.000	0.005	0.005	0.005	0.317	1.000	1.000

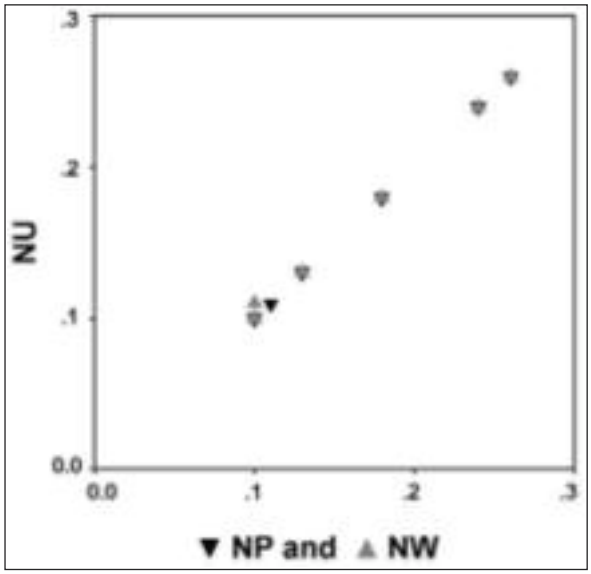

FIGURE 1: The relationship between PPQ values of unmixed and mixed data measured by Dr. Speech.

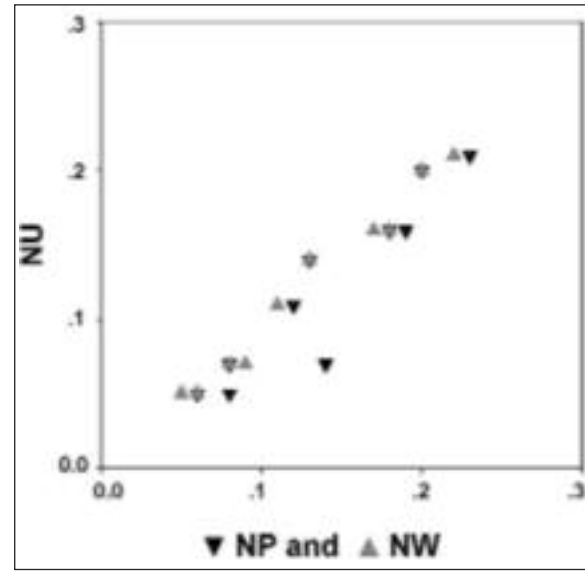

FIGURE 2: The relationship between PPQ values of unmixed and mixed data measured by MDVP.

ved that adding 1% white noise to those samples, deteriorated the performance of MDVP more than Praat. Our findings are in agreement with Boersma's study.

Carson et al.[1] investigated the effect of noise on perturbation parameters using three different software/hardware combinations, namely MDVP for CSL 4300B, MDVP for Multi-Speech and CSpeechSP. They mixed the collected voice samples with computer fan noise to create four different signal-to-noise (S/N) levels (50 dB-unmixed, 25 dB, 20 dB and 15 dB). They compared jitter values across S/N level, four of six comparisons were significant for CSL, three of six comparisons were significant for Multi-Speech and four of six comparisons were significant for CSpeechSP; they also compared shimmer values, six of six comparisons were significant for all systems was MDVP for Multi-Speech the most resistant system to noise. That was the most vulnerable system in our study.

Deliyski et al.[10] investigated the effect of noise on voice quality measurements using five different hardware types and six different microphone types. In that study, three softwares were used: MDVP, Praat and TF32. The voice samples were mixed with five different noise types at eight different S/N ratio levels. They observed that noise seriously affected the reliability of MDVP. They also observed that noise had an equal impact on normal and abnormal phonation.

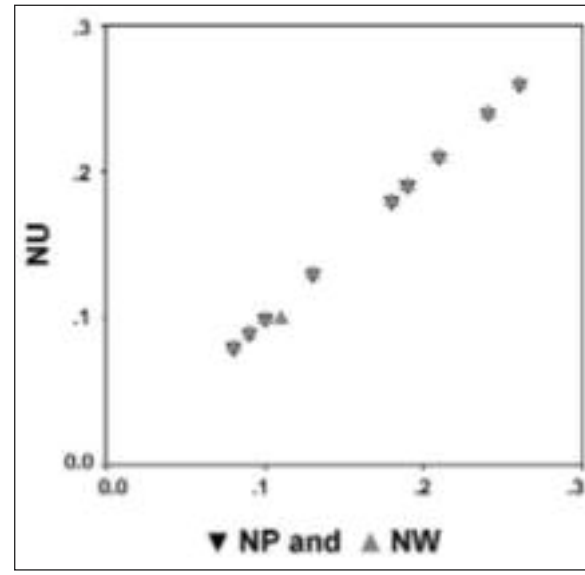

FIGURE 3: The relationship between PPQ values of unmixed and mixed data measured by Praat.

Deliyski et al.[11] investigated the influence of data acquisition environment on voice quality measurements. They commented that the previous observations which infererred that professional data acquisition devices were more vulnerable to noise did not reflect the facts. They suggested, in fact,

TABLE 3: p values obtained by comparison of amplitude perturbation values of the unmixed and mixed data by Wilcoxon test.

	Dr. Speech		MDVP		Praat	
	Shim	APQ	Shim	APQ	Shim	APQ
NU-NP	0.007	0.012	0.016	0.021	0.153	0.050
NU-NW	0.012	0.024	0.008	0.037	0.026	0.084
SU-SP	0.005	0.005	0.005	0.005	0.005	0.005
SU-SW	0.005	0.005	0.005	0.005	0.005	0.005

amplitude and frequency perturbations were measurements of noise in the investigated voice sample; so, being more vulnerable to noise meant more sensitivity to accurately measure the voice perturbations. However, this does not reflect the usual conditions a voice clinician faces in every day life. Even in a sound-treated room, one can not ignore the potential environmental noise sources such as the vibration sources (elevators, air-conditioning, etc) or routine professional and non-professional clinic equipment (computer fan, electricity, heating, etc). Thus, the software we use in the voice clinic should have the chance to accurately discriminate the noise source in the signal. It should not be affected from environmental noise and accurately measure the voice perturbation that originates from glottal noise. The search for the best voice analysis software will continue until a calibrated reference is available.

TABLE 4: Spearman's Rho correlations between perturbation values of the natural unmixed and mixed data (Correlations for semi-synthetic data could not be computed because at least one of the variables was constant)

	Parameters	NU-NP	NU-NW
Programs	Jitt	0.979	1.000
	RAP	1.000	0.991
Dr. Speech	PPQ	0.991	1.000
	Shim	0.952	0.988
	APQ	0.964	0.976
	Jitt	0.994	0.942
	RAP	0.998	0.945
MDVP	PPQ	0.991	0.933
	Shim	0.997	1.000
	APQ	0.979	0.982
	Jitt	0.997	1.000
	RAP	1.000	1.000
Praat	PPQ	0.997	1.000
	Shim	0.985	0.997
	APQ	1.000	1.000

CONCLUSION

This study revealed that Praat was the most immune system to noise among three analysis systems, followed by Dr. Speech and MDVP, respectively. Therefore, among these three softwares, Praat can be preferred when a sound-treated booth is not available for voice recording.

REFERENCES

1. Carson CP, Ingrisano DR, Eggleston KD The effect of noise on computer-aided measures of voice: a comparison of CSpeechSP and the Multi-Dimensional Voice Program software using the CSL 4300B Module and Multi-Speech for Windows. J Voice 2003;17(1):12-20.
2. Bielamowicz S, Kreiman J, Gerratt BR, Dauer MS, Berke GS. Comparison of voice analysis systems for perturbation measurement. J Speech Hear Res 1996;39(1):126-34.
3. Karnell MP, Scherer RS, Fischer LB. Comparison of acoustic voice perturbation measures among three independent voice laboratories. J Speech Hear Res 1991;34(4):781-90.
4. Karnell MP, Hall KD, Landahl KL. Com- parison of fundamental frequency and perturbation measurements among three analysis systems. J Voice 1995;9(4):383-93.
5. Kılıç MA, Okur E. [Comparison of fundamental frequency and perturbation data analyzed by CSL and Dr. Speech systems]. The Turkish Journal of Ear Nose and Throat 2001;8(3): 152-7.
6. Perry CK, Ingrisano DRS, Blair WB. The influence of recording systems on jitter and shimmer estimates. Am J Speech-Lang Pathol 1996;5(2):86-90.
7. Ingrisano DRS, Perry CK, Jepson KR. Environmental noise: a threat to automatic voice analysis. Am J Speech-Lang Pathol 1998;7(1): 91-6.
8. Perry CK, Ingrisano DR, Palmer MA, McDonald EJ. Effects of environmental noise on computer-derived voice estimates from female speakers. J Voice 2000;14(2):146-53.
9. Boersma P. [Measuring voices with Praat]. Stem, Spraak- en Taalpathologie 2004;12(4): 237-51.
10. Deliyski DD, Shaw HS, Evans MK. Adverse effects of environmental noise on acoustic voice quality measurements. J Voice 2005;19(1):15-28.
11. Deliyski DD, Evans MK, Shaw HS. Influence of data acquisition environment on accuracy of acoustic voice quality measurements. J Voice 2005;19(2):176-86.