

# Reduction of Speech Signal Deformation in Patients after Nasal Septum Surgery (Septoplasty)

M. KONIOR<sup>a,\*</sup>, M. KŁACZYŃSKI<sup>b</sup> AND W. WSZOŁEK<sup>b</sup>

<sup>a</sup>Chair and Clinic of Otolaryngology, Jagiellonian University Collegium Medicum  
Śniadeckich 2, 31-531 Kraków, Poland

<sup>b</sup>AGH Department of Mechanics and Vibroacoustics, University of Science and Technology  
al. A. Mickiewicza 30, 30-059 Kraków, Poland

Nasal blockage belongs to the most common symptoms of nasal diseases in vocal tract area. At the same frequency there appear acoustic symptoms, existing as the change of human voice color. Vocal and articulation disorders of the ear, nose and throat are usually observed in the form of closed rhinolalia and this observation can be performed both by patients and other listeners as well. Nasal polyps and nasal septum deviation are frequent reason of nasal blockage connected in consequence with decreased nasal ventilation. One of the main principles of the surgical treatment performed in mentioned situations is the restoration of nasal patency. The evaluation of the influence of nasal surgery on intensification of acoustic symptoms depends on verification of parameters of the human speech signal, so it was necessary to apply objective methods. That allowed to combine results of acoustic analysis with patient's subjective feeling and rhinomanometric evaluation of nasal patency. The main purpose of this research was to objectively evaluate the influence of surgical treatment improving nasal patency on deformation of the voice of operated patients.

PACS: 43.70.-h, 43.70+i

## 1. Introduction

Disorders of nasal patency are the most common symptoms of nasal pathology. It is observed that changes of the color of voice are manifested in the course of the said diseases. One of the most common etiologies of decreased nasal patency is the nasal septum deviation. The goal of the surgical therapy in this case is to improve the permeability of the nasal cavities.

According to acoustic impressions received by the human ear, nasal blockage can be present as a pathological deformation of recorded speech signal. Registration of changes in frequency and intensity (amplitude) of sounds in time is an objective procedure for verification of the patients' complaints.

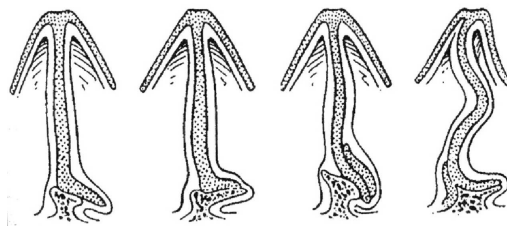


Fig. 1. Different types of nasal septum deviation.

In Poland skeletal abnormalities of nasal septum occur in approximately 60% of the population, but it is the de-

gree of stenosis of nasal canals caused by the deformation that determines intensity of the symptoms. Nasal septum deviation can take various forms: strip, spine, subluxation, dislocation, redouble or arcuate bend, as illustrated in Fig. 1 and is an indication for surgical treatment. The target of the surgery is not only anatomical correction, but above all to improve nasal ventilation.

## 2. Aim

The main goal of the study was to determine the effect of septoplasty improving the nasal patency on the degree of distortion of the voice of patients.

## 3. Material and methodology

The population in study consisted of 24 patients diagnosed with nasal septum deviation (17 men and 7 women). Mean age of patients eligible to participate in the research program was 31.5 years with the group range between 18 and 49 years. All patients underwent the following procedures:

- clinical evaluation,
- active anterior rhinomanometry,
- digital recording of speech samples.

Measurements were taken before treatment and twice after the surgical correction of nasal flow conditions. In all patients a septoplasty was performed.

Testing procedures included the following steps:

\* corresponding author; e-mail: marcin.konior@poczta.fm

- preoperative examination (on admission to the clinic),
- early follow up visit (28–32 days after surgery),
- late follow up (90 days after surgery).

Time of follow-up visits was established due to the post-operative course of healing, so as to register the early changes in the course of healing, and late changes occurring after the regeneration of tissue surrounding operated area. For rhinomanometry in the study rhinomanometer MES RHINOTEST 500 was used. The device was calibrated before each measurement [1].

Recording of speech samples for later analysis of the speech signal was the last part of the preoperative examination. The study, oriented towards determination of usefulness of selected methods of atypical speech signal parameterization in diagnosis of pathological speech, has been carried out using time domain samples of both pathological and normal speech signals, the latter being treated as a separate, individually pre-adjusted reference

set. The registration of time-dependent acoustic speech signals has been carried out in the anechoic chamber, located in Department of Mechanics and Vibroacoustics of AGH University of Science and Technology in Kraków. The applied in research phonetic test (presented in Table I) was arranged and proposed by the group of otolaryngologists and phoniatriest experts. The same test was used in earlier authors' research [2, 3]. The selection of sounds and groups of words uttered by the examined persons has been based on morphological and functional analysis of the expected (for a given pathology type) dysfunction of speech organs. The phonetic material available for the study consisted of groups of words, selected with respect to their phonetic features, in order to provide maximum contribution of information concerning the examined pathology. The recording system employed a professional digital magnetic recorder (HHB type PDR 1000) allowing for recording in the frequency range from 20 Hz to 20 kHz and dynamic range exceeding 80 dB.

Phonetic test.

TABLE I

A, A, A	Ala, Ala, Ala	konie, konie, konie	mądry, mądry, mądry	płatczyk,
E, E, E	Ela, Ela, Ela	siano, siano, siano	ręka, ręka, ręka	płatczyk, płatczyk
U, U, U	Ula, Ula, Ula	kino, kino, kino	wąs, wąs, kęs	móżdżek, móżdżek,
I, I, I	Dziś jest ładna pogoda, Dziś jest ładna pogoda, Dziś jest ładna pogoda	miękki, miękki, miękki	kęs, kęs, kęs	móżdżek

In this study nasal sound “n” from words “konie” and “siano” was analyzed to assess an improved voice quality (reduction of nasality) for the voice after nasal septum surgery treatment. Because septoplasty does not change the characteristics of vowels in speech signal, we used the oral sound “a” from word “trawa” as a reference (our expectations: the value of metric — distance of the nasal sounds before the surgical treatment from the external standard of speech signal is greater than the respective value in the post-treatment and final evaluations and simultaneously and objectively the distance for oral sound is unchanged).

Recorded data was transferred to a computer (as wave file type) and then fragmented and segregated. Using Matlab environment “n” and “a” sounds were analyzed in time domain by the zero crossing analysis (ZCA), in the frequency domain by the short-term Fourier transform (STFT) and with the use of a mel scale — mel frequency cepstral coefficients (MFCC) [4]. As a result of the STFT, the dynamic spectrum  $W(t, f)$  was obtained. This spectrum contained 96 frequency bands with fre-

quency step  $\Delta f = 125$  Hz, time step  $\Delta t = 9$  ms and signal level step equal to  $\Delta L = 0.2$  dB.

The following parameters of voice signals digital conversion were applied in the MFCC calculations:

- Hamming's time-weighting window (length of  $N = 384$  samples),
- The 384-point digital Fourier transform (DFT) calculated at every 8 ms (signal sampling frequency of 48 kHz),
- Frequency axis was rescaled to the mel scale according to the equation given in [4],
- Mel filtering — meaning the spectral conversion using triangular band pass filters corresponding to mel scale (by summing the weighted spectral lines). The number of filters in the set is equal to  $N = 12$ ,
- Calculation of cepstral coefficients in the mel-scale bands, MFCC, as a discrete conversion of logarithm cosines of the parameters of filter data,

according to the formula [4]:

$$C_n = \sqrt{\frac{2}{N}} \sum_{i=1}^N \log(s_i) \cos\left(\frac{\pi n}{N} \left(i - \frac{1}{2}\right)\right), \quad (1)$$

where  $C_n$  —  $n$ -th cepstral coefficient,

- $S_i$  —  $i$ -th coefficient obtained from signal conversion by the set of filters,  $N$  — number of filters in the set,  $N = 12$ .

Collected research material analysed by means of ZCA, STFT and MFCC techniques allowed for a creation of the 29-element feature vector (2):

$$X_n = \langle M_0, M_1, M_2, F_1, F_2, F_3, F_4, AF_1, AF_2, AF_3, AF_4, W_1, W_2, W_3, C_1, C_2, C_3, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}, C_{11}, C_{12}, f_0, J, S \rangle \quad (2)$$

where  $n$  — identification number — according to the data base of patients,  $M_0$ – $M_2$  — spectrum moments of the order: 0, 1, 2, [2, 5],  $F_1$ – $F_4$  — formant frequencies [2, 4, 5],  $AF_1$ – $AF_4$  — amplitude of formant frequencies [2, 4, 5],  $W_1$ – $W_3$  — relative power coefficients [2],  $C_1$ – $C_{12}$  —  $n$ th cepstral coefficient,  $f_0$  — mean value of the basic frequency of the larynx tone vibration [2, 4, 5],  $J$  — jitter, denotes a frequency deviation of the basic tone — in consecutive periods [2, 4],  $S$  — shimmer, denotes an amplitude deviation of the basic tone — in consecutive periods [2, 4].

In the measurement of speech signals distances from the signal pattern Euclidean, weighted Euclidean, Hamming and weighted Hamming metrics were used. Differences in distances from the speech signal pattern identified in this way were a measure of deformation of human voice [6, 7]. In this paper, the choice of metrics has been made taking into account that the most meaningful are those whose values show greater differences. Polish speech pattern (control group) was created by recording and analyzing samples of normal speech collected from non-vocally trained healthy men and women without any pathology which may affect the voice color. Described pattern has already been used as a reference for other studies [2, 3].

In the study StatSoft Statistica software was used basing on the analysis of variance ANOVA, and verifying tests: Duncan and the least significant difference (NIR). In the calculations, the following conditions:  $N = 120$  (24 patients, 96 persons in control group),  $k = 2$ ,  $p = 0.001 \div 0.01$ , where  $N$  — number of observations,  $k$  — number of groups,  $p$  — level of significance were established.

#### 4. Results

The results of measurements of the distance of nasal sounds from the signal pattern presented a reduction in all eight measurements (100% of the surveyed population) in early control and a further reduction in the next four from eight measurements (50%) in the late control.

In this study, there were no statistically significant difference in the distance of oral sounds from the signal pattern, neither in the early nor in the late control. Aggregated qualitative summary of the results, taking into account the variability of measured parameters is shown in Tables II and III.

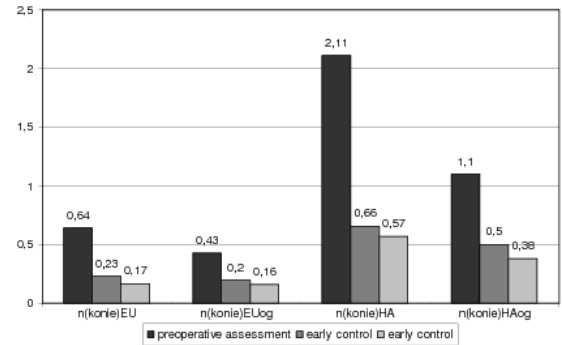


Fig. 2. Distance of the sound “n” speech signal from the pattern “konie”.

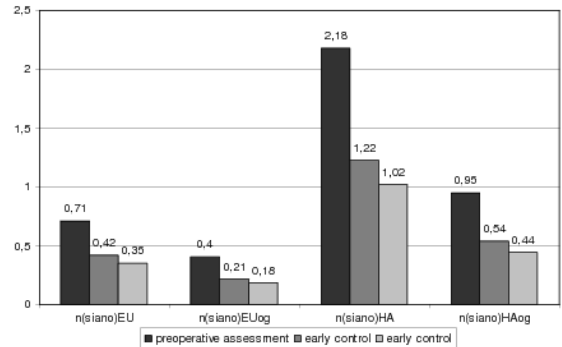


Fig. 3. Distance of the sound “n” speech signal from the pattern “siano”.

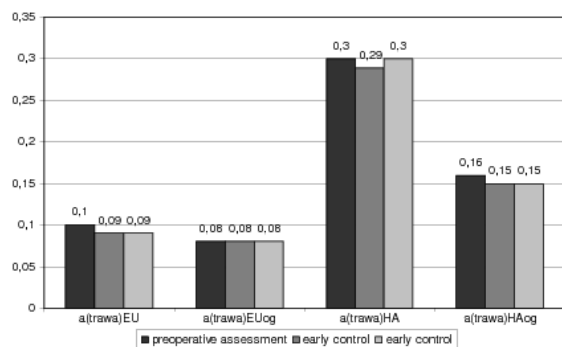


Fig. 4. Distance of the sound “a” speech signal from the pattern “trawa”.

Graphic presentation of the change of the distance of the sound “n” and “a” speech signal from the pattern is

shown in Figs. 2–4. The changes of nasal resistance has been shown in graphic form in Fig. 5.

TABLE II  
Variability of measured parameters.

Parameter	Statistical difference		<i>P</i>
	Early follow-up	Late follow-up	
n(konie)EU	+	+	0.01
n(konie)EUog	+	–	0.01
n(konie)HA	+	–	0.01
n(konie)HAog	+	+	0.01
n(siano)EU	+	+	0.01
n(siano)EUog	+	–	0.01
n(siano)HA	+	–	0.01
n(siano)HAog	+	+	0.01
a(trawa)EU	–	–	0.001
a(trawa)EUog	–	–	0.001
a(trawa)HA	–	–	0.001
a(trawa)HAog	–	–	0.001
RNSCex	+	+	0.001
RNSCin	+	+	0.001
RNBCex	+	+	0.001
RNBCin	+	+	0.001

where: + statistically significant difference according to previous examination; – statistically non-significant difference according to previous examination

TABLE III  
Variability of measured parameters — average values.

Parameter	Unit	Assessment		
		Preop	Early	Late
n(konie)EU	scalar	0.64	0.23	0.17
n(konie)EUog	scalar	0.43	0.20	0.16
n(konie)HA	scalar	2.11	0.66	0.57
n(konie)HAog	scalar	1.10	0.50	0.38
n(siano)EU	scalar	0.71	0.42	0.35
n(siano)EUog	scalar	0.40	0.21	0.18
n(siano)HA	scalar	2.17	1.22	1.02
n(siano)HAog	scalar	0.95	0.54	0.44
a(trawa)EU	scalar	0.10	0.09	0.09
a(trawa)EUog	scalar	0.08	0.08	0.08
a(trawa)HA	scalar	0.30	0.29	0.30
a(trawa)HAog	scalar	0.16	0.15	0.15
RNSCex	kPa/l/s	0.29	0.19	0.09
RNSCin	kPa/l/s	0.25	0.17	0.08
RNBCex	kPa/l/s	0.34	0.20	0.08
RNBCin	kPa/l/s	0.34	0.20	0.09

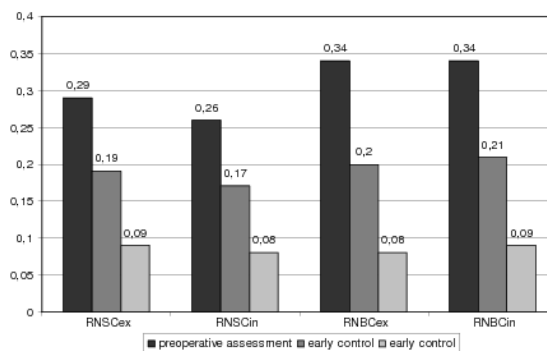


Fig. 5. Changes of nasal resistance.

## 5. Discussion

Value of the nasal resonance depends mostly on the number of nasal vowels in the speech and it varies depending on the language [8, 9]. Acoustic disorders associated with pathologies of the nasal cavity in which the leading symptom is impairment of the permeability are not well known [10].

Procedures such as septoplasty are aimed primarily at nasal ventilation improvement, but also influences a change of patients' "voice color" [11, 12]. Nowadays, according to rapid development of information technology acoustic analysis is assessed as very useful diagnostic instrument in vocal studies. Many authors indicate the need to objectify patients' symptoms because of the large discrepancy between subjective assessments of listeners [13]. Results of measurements of the distance of nasal vowel signal from the pattern signal clearly show the improved acoustics of voice and speech during the articulation of these sounds in patients undergoing septoplasty. The improvement is mainly manifested in the spectrographic picture of sound "n" by the better selection of formant frequencies, in particular for first and second control and the increase in the amplitude of the sound wave. This phenomenon can be explained by better availability of resonance space of the nasal cavity in the way of improvement of the air ventilation. Surgical intervention is the only explanation for the reported acoustic differences. Chen and Hosemann obtained similar results in their study in patients that underwent endoscopic sinus surgery [9, 14].

Results obtained in this study will contribute to better understanding and explanation of phenomena occurring in the space of the nasal cavity having an impact on the final shape of the patient's voice. That allows, in practice, to understand better the pathology that cause abnormal nasal patency, the course of surgical treatment and its anticipated consequences.

## 6. Conclusions

1. After septoplasty there is a decrease in the distance of speech signal from the signal pattern of nasal sounds "n".

2. Better conditions of nasal flow improve the nasal resonance expressed as a reduction of deformation of nasal sounds.

### Abbreviations

a(trawa)EU — distance of the speech signal of the sound “a” from the pattern isolated from the word “trawa” measured by Euclidean metric,

a(trawa)EUog — distance of the speech signal of the sound “a” from the pattern isolated from the word “trawa” measured by weighted Euclidean metric,

a(trawa)HA — distance of the speech signal of the sound “a” from the pattern isolated from the word “trawa” measured by Hamming metric,

a(trawa)HAog — distance of the speech signal of the sound “a” from the pattern isolated from the word “trawa” measured by weighted Hamming metric,

n(konie)EU — distance of the speech signal of the sound “n” from the pattern isolated from the word “konie” measured by Euclidean metric,

n(konie)EUog — distance of the speech signal of the sound “n” from the pattern isolated from the word “konie” measured by weighted Euclidean metric,

n(konie)HA — distance of the speech signal of the sound “n” from the pattern isolated from the word “konie” measured by Hamming metric,

n(konie)HAog — distance of the speech signal of the sound “n” from the pattern isolated from the word “konie” measured by weighted Hamming metric,

n(siano)EU — distance of the speech signal of the sound “n” from the pattern isolated from the word “siano” measured by Euclidean metric,

n(siano)EUog — distance of the speech signal of the sound “n” from the pattern isolated from the word “siano” measured by weighted Euclidean metric,

n(siano)HA — distance of the speech signal of the sound “n” from the pattern isolated from the word “siano” measured by Hamming metric,

n(siano)HAog — distance of the speech signal of the sound “n” from the pattern isolated from the word “siano” measured by weighted Hamming metric,

RNBCex — total expiratory nasal resistance (Broms),

RNBCin — total inspiratory nasal resistance (Broms),

RNSCex — total expiratory nasal resistance (standard),

RNSCin — total inspiratory nasal resistance (standard).

### References

- [1] M. Sarzynska, K. Dudziec, M. Wieclawska, A. Kukwa, *Otolaryngol. Pol.* **51**, 290 (1997).
- [2] Z. Engel, M. Kłaczyński, W. Wszolek, *Int. J. Occupat. Safety Ergon.* **13**, 369 (2007).
- [3] W. Wszolek, M. Kłaczyński, *Arch. Acoust.* **30** (Supp.), 193 (2005).
- [4] J.R. Deller, Jr., J.G. Proakis, J.H.L. Hansen, *Discrete-Time Processing of Speech Signals*, Macmillan Publishing Company, New York 1993, p. 380.
- [5] R. Tadeusiewicz, *Speech Signals*, WKiL, Warsaw 1988, p. 159 (in Polish).
- [6] M. Haapanen, L. Liu, T. Hiltunen, L. Leinonen, J. Karhunen, *Folia Phoniatr. Logop.* **48**, 35 (1996).
- [7] R. Tadeusiewicz, M. Flasiński, *Pattern Recognition*, PWN, Warszawa 1991, p. 43 (in Polish).
- [8] R. Anderson, *Cleft Palate Craniofac. J.* **33**, 333 (1996).
- [9] M. Chen, R. Metson, *Arch. Otolaryngol. Head Neck Surg.* **123**, 845 (1997).
- [10] H. Bauer, *Hals-Nasen Ohrenheilkd.* **28**, 171 (1980).
- [11] M. Constantian, R. Clardy, *Plast. Reconstr. Surg.* **98**, 38 (1996).
- [12] G.-S. Lee, C.C.H. Yang, C.-P. Wang, T.B.J. Kuo, *J. Voice* **19**, 71 (2005).
- [13] R. Dalston, G. Neiman, G. Gonzalez-Landa, *Cleft Palate Craniofac. J.* **30**, 285 (1993).
- [14] W. Hosemann, U. Gode, J.E. Dunker, U. Eysholdt, *Eur. Arch. Otorhinolaryngol.* **255**, 499 (1998).