Suprasegmental Analysis for Complex Quality Assessment in Pathological Voices

Grażyna Demenko

Institute of Linguistics
Adam Mickiewicz University, Poznań
lin@amu.edu.pl

Andrzej Obrębowski, Antoni Pruszewicz, Bożena Wiskirska-Woźnica, Piotr Świdziński, Waldemar Wojnowski

Chair and Department of Phoniatrics and Audiology
University of Medical Sciences, Poznań
bozena.woznica@wp.pl

Abstract

Phoniatric and acoustic examinations were carried out in a group of 30 patients with dysphonia, including 15 with organic and 15 with functional dysphonia. Another group of 18 patients with deep hypoacousis was divided into 3 categories: congenital, pre- and post-lingual. Phoniatric and stroboscopy examinations of the larynx was performed in all cases. The differences between functional and organic dysphonia were found after voice fatigue examination. The analysis of the global statistical features of fundamental frequency distribution estimated from the read text as well as the analysis of changes of Fo parameter in individual utterances was applied. The results of this study have been used in more detailed acoustical and perceptual pathological voice analysis.

In the other group of 100 patients with organic and functional voice disorders complex voice assessments including extended acoustic-perceptual voice estimation in GRBAS scale as well as classical spectrographic and complex MDVP analysis were performed on appropriate linguistic material including vowels, isolated words and a read text.

The importance of complex voice examination is stressed. Voice is a multidimensional phenomena and it cannot be assessed by simple acoustic or phoniatrics methods but only by more complex ones.

1. Introduction

In numerous publications the possibility of an objective assessment of the nature and degree of voice damage on the basis of acoustic information has been stressed ([3], [10]). A description of the degree and type of voice pathology determined by suprasegmental parameters was accepted as a very effective method in the diagnosis of evaluated pathological changes in the larynx. For the investigations of different voice disturbances, the analysis of suprasegmental speech features is very essential from a theoretical as well as from a practical point of view for numerous reasons: a) the prosodic structure is not affected by the frequency characteristics of the transmission systems, b) it is important in conveying nonlinguistic information c) it can be relatively direct compared with phoniatric examination. The diagnosis of larynx disorders cannot be determined only by acoustic examination. Until the present time, there has been no specific and characteristic change in acoustic voice analysis corresponding directly with larynx pathology. Voice and speech as a way of communication also requires non-instrumental perceptive evaluation. The aim of this research is to introduce and develop into phoniatric diagnostic procedures the complex methodology of vocal function assessment as a tool for voice quality multidimensional analysis.

2. Usefulness of the suprasegmentals speech features for the analysis of voice disorders

2.1 Melodic factors of speech and voice of individuals with hearing disturbances

In the evaluation of the communication process in individuals with different stages of hearing deficit special attention is usually paid to proper articulation, and at the same time the problems of the rehabilitation of the voice function, manifested in incorrect suprasegmental structures, is neglected ([5]). People with hearing disorders have problems with segmental features (vowel and consonant mistakes, incorrect articulation), as well as with suprasegmental ones (incorrectness in rhythm, stress and intonation). The proper realization of these features depends on the strength of articulation and the time of its duration (mutual temporal alignment). Suprasegmental mistakes especially indicate strength and time, e.g. the speed of speech production including the speed of articulation transients which depends on the time parameter, whereas the Fo changes are connected with vocal folds tensions and subglottic pressure. The most important problem for deaf speakers is not which articulation movements to use but what strength and duration of muscle tension and their mutual alignment should be engaged in the speech production. The examined patients (11 male, 7 female aged 11-59) with deep perceptive hypoacousis (above 70 dB) or with complete deafness were divided into 3 groups: those with congenital disturbances, those appeared in early childhood before speech development and those with deafness onset above the age of 15 years. The following analysis were performed: ENT examination, phoniatric examination including stroboscopy and a-p rentgenotomography of the larynx during phonation of the vowel ***a***, acoustic examination of isolated utterances and a read text. Parameters of acoustic examination are presented in Table 1. The speech disturbances in the range of periodicity are described by the following parameters: mean value of fundamental frequency, the instability of voice (described by maximum Fo change within 2 periods), the range of Fo, periodicity in vowels, jitter and its mean value.

Table 1: *Results of acoustical examination. Accordingly in columns are: (1) initials of patients, (2) mean Fo value (in bold are normal values for respective sex and age in Hz) (3) instability of voice (in Hz), (4) range of Fo (in Hz), (5) noise/lack of noise in vowels, (6) jitter and (7) type of intonation course within the utterance.*

	1	2	3	4	5	6	7
GROUP I	AW	282 256	400 276	93 350	noise	12	falling
	DB	240 256	300 218	120 308	noise	8	falling
	MB	315 256	445 325	100 448	noise	6	falling
	NJ	218 256	234 198	160 245	noise	9	falling
	KD	156 128	280 145	139 204	noise	4	falling
	WP	410 128	440 380	350 444	noise	3	falling
	AM	? 256	?	90 465	?	?	variability uncontrolled
GROUP II	SK	225 256	237 210	186 265	noise	4	falling
	NG	384 256	440 390	370 448	noise	3	falling
	WM	240 256	260 243	216 248	lack of noise	3	falling
	HB	321 256	382 318	289 440	lack of noise	6	equal rising
	NK	173 128	173 144	120 215	lack of noise	2	equal rising
	KP	160 128	181 169	129 190	noise	2	equal
	AP	215 256	223 201	160 296	lack of noise	2	falling
GROUP III	LJ	253 256	289 227	236 340	lack of noise	13	falling
	KR	116 128	104 95	90 156	lack of noise	1.5	equal
	SB	165 128	162 147	126 315	lack of noise	2	equal
	KF	245 256	286 256	219 285	lack of noise	1.5	falling

In first group a great variability of the Fo from period to period was observed and lack of control of amplitude. Jitter values were in the range 3-12, the instability of voice was observed (max: 400Hz/276Hz). Fundamental frequency courses were falling. The periodicity of consonants and vowels was disturbed and considerable noise was observed. In second group jitter values were smaller as in group I (in the range 2-6). The voices were characterized by better stability (237Hz/210Hz). Fundamental frequency courses were mainly falling (sometimes equal-rising). The periodicity of consonant was disturbed. In third group the value of jitter was mostly small (1,5-2; with one exception), a good correlation of amplitude and Fo parameter can be seen. Mostly rhythmical patterns were disturbed, some fragments were exaggerated and prolonged, intonation was monotonous (did not vary in statements or questions). In Group 1 and 2 the increase of Fo mean value was observed. In Group 3 the most important were disturbances of the melodic patterns. It seems that the first to be disturbed is prosody structure, next segmental structure.

2.2 Diagnosis of functional and organic voice disorders

The target of this investigation was an attempt to find a preliminary acoustic features of voice in groups of patients with organic and functional changes in order to create an algorithm of evaluation of the degree of voice disturbance. It is to a certain degree based on the biophysical assumption that the organic changes usually cause an asymmetric growth of vocal fold mass and that functional changes alter the degree of their tension and coordination ([6], [9]). In this pilot study the investigation included 2 groups of patients: 15 patients with organic changes confined to the glottis and 15 patients with functional voice disturbances in the form of dysphonia hyperfunctionalis. The phoniatrics examination consisted of: a) indirect laryngeoscopy, b) subjective evaluation of the voice, c) determination of voice frequency alone with voice range and voice intensity before and after voice fatigue, d) determination of maximum phonation time, e) stroboscopic examination f) larynx tomography. In the acoustic examinations, the following features were analyzed: the harmonic structure of the signal, statistics of F0, statistics of jitter, changes of Fo parameter within the utterance. In Group I essentially greater than in non-pathological voices jitter (1-8%) was found. In those patients the speech signal had a very poor prosody structure. In Group II with functional changes, smaller variability of the length of subsequent periods than in Group I was observed. The mean value of jitter in this group is considerably lower and amounts to 4 %. The melodic courses are more differentiated than in Group I.

3. Complex voice evaluation in organic and functional disorders

The basic assumption of this research is to prove the correlative existence of vocal folds vibration quality in videostroboscopy, parameters describing acoustic wave generated by vibration system in the larynx (MDVP) and perceptual subjective voice assessment (GRBAS scale) in organic and functional voice disorders ([3], [4]). Research was conducted on 100 subjects, aged 7-74 years randomly chosen from population of patients treated at the Department of Phoniatrics and Audiology, University School of Medical Sciences in Poznań and 60 subjects as a control group without any voice disorders according to phoniatric examination. The first group was composed of 65 patients with organic etiology of hoarseness (chronic laryngitis, chronic laryngitis, hypertrophic laryngitis, Reincke edema, vocal folds polyps, posttraumatic dysplastic changes in the larynx, benign neoplastic organic pathology on vocal folds, long time persistent vocal fold paralysis, granuloma, papilloma and cysts of vocal folds). Out of this group 15 subjects were selected, to whom surgical treatment was performed and who were examined before and after operation. The next group consisted of 35 patients with functional voice disorders (hyperfunctional and hypofunctional dysphonia, unilateral vocal fold paralysis, infant dysphonia, phonasthenia, psychogenic dysphonia). The research methodology included perceptual voice evaluation based on the Japanese ([3]) GRBAS scale adapted to Polish ([7], [8]), where G was grade of hoarseness in scale 0 to 3 (0- normal voice, 1 – slight hoarseness, 2 – moderate hoarseness, 3 – extreme hoarseness), R – rough voice in binary scale 1 or 2 (yes or no), B – breathy 1 or 2 (yes or no), A- asthenic 1 or 2 (yes or no), S – strain 1 or 2 (yes or no). The vibrations of vocal folds in videostroboscopy have been described with attempt at quantification (scale 1 to 3) of selected measures: symmetry and regularity **N** (1-yes, 2-no), amplitude **A** (1-normal, 2- reduced or increased), mucosal wave **P** (1-normal, 2 – reduced, 3- none) and glottic closure **Z** (1-complete, 2- incomplete). Vowels recorded in isolation and the Polish sentence " *Ten dzielny żołnierz był z nim razem* " (eng. *The brave soldier was with him*) were used for acoustic voice analysis for the Multi Dimensional Voice Program (MDVP) and the software analyzer Kay Elemetrics Model CSL 4305.

The results in the group of organic voice disorders showed the greatest number of significant dependences in R Spearman and Tau Kendall tests, for all parameters describing vocal folds vibrations in videostroboscopy in relation to parameters of both remaining methods (acoustic and perceptive). Out of subjective descriptions in perceptive voice evaluation (GRBAS scale) the grade of hoarseeness G showed the most significant correlation with videostroboscopic and acoustic analysis parameters (Fig.1, 2, 3). The MDVP analysis in this group presented the correlation for frequency perturbation measures (Jita, Jitt, RAP, PPQ, vFo) in relation to parameters of GRBAS scale and videostroboscopy (Fig.4). The evaluation of differences between parameters before and after surgery (improvement after operation) in perceptive GRBAS scale revealed significant differences for grade of hoarseness G and for rougheness R. In MDVP analysis differences after surgery comparing to those before it were noticed for two frequency parameters RAP and PPQ at a level of $p<0,05$ according to Wilcoxon test and signs test. All the videostroboscopic measures (**N, A, P, Z)** were statistically different after operations ($p<0,05$) from those before it, particularly amplitude at a level of $p< 0,001$. The evaluation of the voice in functional disorders showed significant correlation between perceptive parameter- breathy voice **B** in GRBAS scale and noise parameter NHR (noise to harmonic ratio) in acoustic analysis MDVP (Fig.5). In this group a relation was found also between the grade of hoarseness **G** (GRBAS scale) and two videostroboscopic parameters – amplitude **A** and glottal closure **Z** ($p<0,05$)**.** The multiparameter assessment of voice disorders in comparison to the normal voices showed essential correlation for almost all measures of perceptive, acoustic and videostroboscopic examination, according to U Mann- Whitney and Kołgomorow- Smirnow tests at a level of $p<0,05$. The exception was functional voice disorders, where correlation was significant only for amplitude and glottal closure in videostroboscopy. For both organic and functional changes the most significant correlation in GRBAS scale was found in grade of hoarseness **G**, but in MDVP in parameters measured frequency perturbation. In differentiating between organic and functional voice disorders, only videostroboscopic parameters were significant.

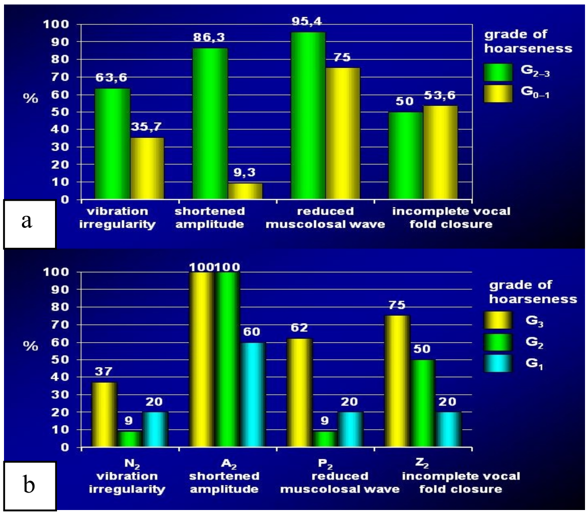

Figure 1: *Correlation between stroboscopic examination and grade of hoarseness* ***(G)*** *in subjective GRBAS scale*
a) in organic changes
b) in functional voice disorders

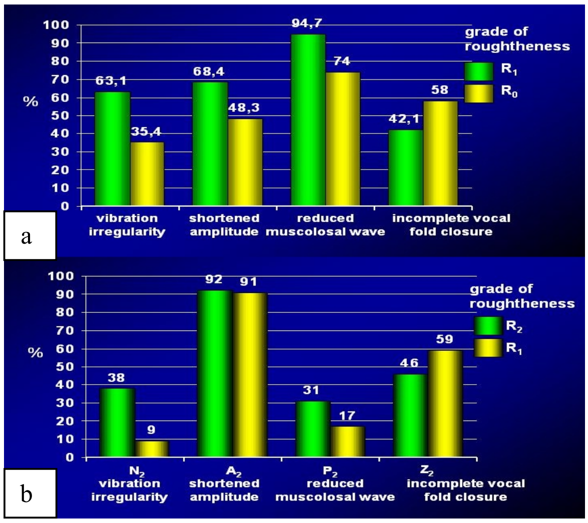

Figure 2: *Correlation between stroboscopic examination and grade of roughteness* ***(R)*** *in subjective GRBAS scale*
a) in organic changes
b) in functional voice disorders

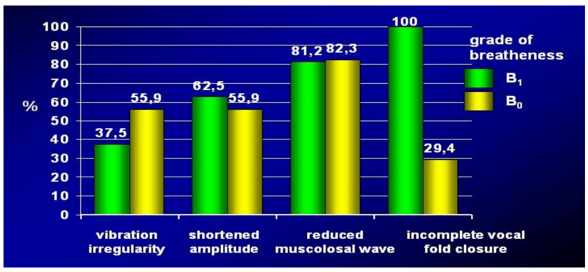

Figure 3: *Correlation between stroboscopic examination and grade of breatheness* ***(B)*** *of the voice in organic changes in subjective GRBAS scale*

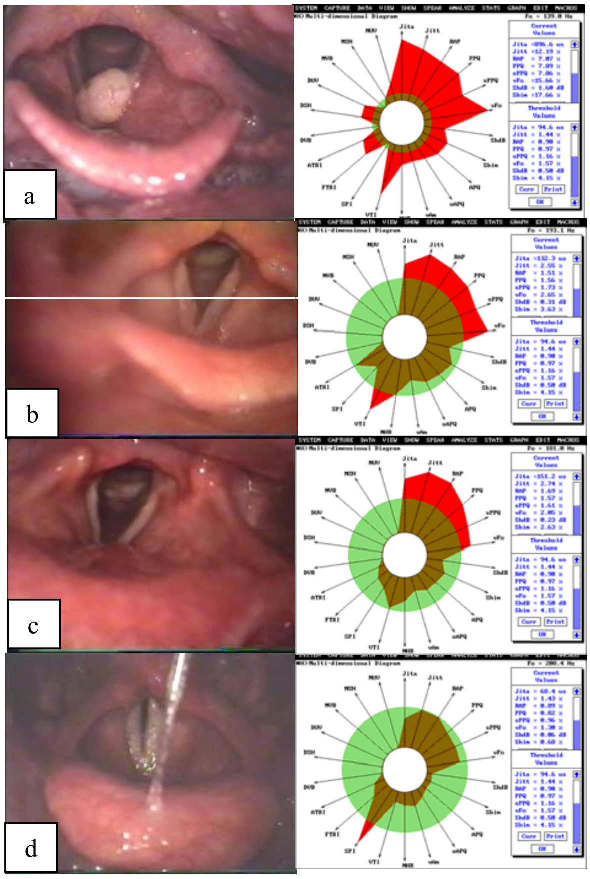

Figure.4*: Videostroboscopic picture of the vocal folds and acoustic analysis in MDVP*

a) organic-Abrikosow tumor of the larynx
b) organic-vocal nodules
c) functional hyperfunctional voice disorders
d) functional-hypofunctional voice disorders

MEAN	Hoarseness			Rougheness		Breatheness		NORM
	G3	G2	G1	R2	R1	B2	B1	
JITT%	6,31	5,53	2,49	6,41	2,98	4,82	5,55	1,44
SHIM%	9,87	7,62	3,08	9,19	3,79	6,90	9,08	4,15
VFO %	7,31	6,46	3,26	7,78	3,11	5,76	6,38	1,57
NHR	0,23	0,19	0,12	0,15	0,11	0,27	0,21	0,15

Figure 5: *The correlations between MDVP acoustic parameters and subjective voice estimations in GRBAS scale (Jitt-Jitter percent, Shim-Shimmer percent, vFo-Fundamental Frequency variation, NHR-Noise- to- HarmonicRatio).*

4. Discussion

Complex voice assessment requires phoniatric as well as perceptual examination (eg. GRBAS) and objective (e.g. MDVP) scales. Suprasegmental speech features were mostly useful for the diagnosis of different types of pathology. The most effective parameters in the MDVP analysis were: jitter, shimmer and fundamental frequency variation. Probably more precise suprasegmental analysis would make it possible to recover hidden information coming from different types of pathological speech ([2]). Caution should be taken in using normative values in MDVP because these values may not be appropriate for various age-sex populations. The concept of an integral, representative database is important ([1]). The set of complex voice evaluation ought to include perceptual examination of voice quality with a quantitative GRBAS scale as a subjective method, vocal folds vibration in videostroboscopy as a quasi-objective method and acoustic voice estimation in Multi Dimensional Voice Program (MDVP) as an objective method.

5. Conclusions

Methodology and scale based on application of three methods of voice examination could be an objective rate in estimation the degree of voice disorder in differential diagnosis of organic and functional changes in the larynx. It could also be useful in the estimation of effectiveness of treatment and rehabilitation methods, in monitoring benign organic changes on the vocal folds, in post-treatment voice assessment after conservative and surgical therapy as well as in the selection process for professions requiring efficient vocal function.

6. References

[1] Dejonckere P.H., Crevier L., Elbaz E., Marraco M., Remacle M., Woisard V., 2000, Clinical implementation of a multidimensional basic protocol for assessing functional results of voice therapy, *Proc.of 4th EUFOS Congress,* Berlin, 561-565.

[2] Demenko G., 1999, Analysis of Polish suprasegmentals for Speech Technology, *ed. UAM,University,* Poznań.

[3] Hirano M., 1989, Objective evaluation of the human voice clinical aspects. *Folia Phoniatr*., 1989,41, 89-114.

[4] Hirose H., 1988, High-speed digital imaging of vocal fold vibration, *Acta Otolaryngol. Stockh*., 458, 151-153.

[5] Pruszewicz A., Demenko G., Wika T., 1993 Variability Analysis of F_0 Parameter in the Voice of Individuals with Hearing Disturbances, *Acta Otolaryngol (Stockholm*) 113, 450- 454

[6] Pruszewicz A., Obrębowski A., Świdzinski P., Demenko G., Wika T., Wojciechowska A., 1991, Usefulness of Acoustic Studies on the Differential Diagnostics of Organic and Functional Dysphonia, *Acta Otolaryngol, Stockholm,* 111, 414- 419.

[7] Woźnica B., Pruszewicz A., Obrębowski A., Swidziński P., 2001, Complex voice assessment after microsurgery of vocal folds, *Aktuelle phoniatrisch-pädaudiologische Aspekte 2000/2001,* band 8, ed. M.Gross, E.Kruse, 65-69.

[8] Woźnica B., 2002, The Complex Voice Evaluation in Organic and Functional Larynx Disorders*, ed. AM University*, Poznań.

[9] Świdzinski P., 2000, The usefullness of objective acoustic technics in organic and functional voice disorders, *ed. AM University,* Poznań, 77-83.

[10] Kent R., Vorperian H., Duffy J., 1999, Reliability of the Multi-Dimensional Program for the Analysis of Voice Samples of Subjects with Dysarthria*, American Journal of Speech- Language Pathology,* vol.8, 129-136.