Multidimensional Voice Program (MDVP) and amplitude variation parameters in euphonic adult subjects. Normative study

Multidimensional Voice Program e parametri di perturbazione d'ampiezza in soggetti adulti eufonici. Normativa

M. NICASTRI, G. CHIARELLA, L.V. GALLO, M. CATALANO, E. CASSANDRO
Department of Experimental and Clinical Medicine "G. Salvatore", Audiology Unit, "Magna Graecia" University, Catanzaro, Italy

Key words

Voice • Normal subjects • Voice analysis • Voice parameters

Parole chiave

Voce • Soggetti normali • Analisi della voce • Parametri vocali

Summary

The introduction, in the late 70s, of the first digital spectrograph (DSP Sonograph) by Kay Elemetrics has improved the possibilities of spectroacoustic voice analysis in the clinical field. Thanks to the marketing, in 1993, of the Multi Dimensional Voice Program (MDVP) advanced system, it is now possible to analyse 33 quantitative voice parameters which, in turn, allow evaluation of fundamental frequency, amplitude and spectral energy balance and the presence of any sonority gap and diplophony. Despite its potentials, the above-mentioned system is not widely used yet, partly on account of the lack of a standard procedure. Indeed, there are still only a few case reports in the literature taking into consideration prescriptive aspects related both to procedure and analysis. This study aims to provide the results of amplitude perturbation parameter analysis in euphonic adult patients. In our opinion, these are the most significant parameters in determining the severity of a phonation disorder. The study has been carried out on 35 patients (24 female, 11 male, mean age 31.6 years, range 19-59). The voice signal has been recorded using a 4300 B Kay Computer Speech Lab (CSL) supported by a personal computer including a SM48 Shure-Prolog microphone located at a distance of 15 cm and angled at 45°. Input microphone saturation has been adjusted to 6/9 of the CH1 channel. The voice sample consisted in a held /a/ and the analysis has been carried out on the central 3 seconds of the recording. The analysis has been carried out using a 5105 MDVP software version 2.3 and the signal digitalised at a 50 kHz sample rate. In order for the sample to be as free from intensity or frequency changes as possible, each patient underwent a training session (including at least 3 phonation tests) before the recording. The study included only emissions between 55 and 65 dB and with spectrum stability. Environmental noise has constantly been monitored and maintained below 30 dB. Data from the 24 female patients showed: absolute Shimmer (ShdB) = 0.203 dB, relative shimmer (Shim %) = 2.226, amplitude perturbation quotient (APQ) = 1.758, smoothed amplitude perturbation quotient = 3.309, peak amplitude variation (vAm) = 7.042. Data from the 11 males showed: absolute Shimmer (ShdB) = 0.269 dB, relative shimmer (Shim %)

Riassunto

Nella seconda metà degli anni '70, con l'introduzione da parte della Kay Elemetrics del primo spettrografo digitale (DSP Sonagraph), le possibilità di analisi spettroacustica della voce in ambito clinico si sono notevolmente ampliate e con la commercializzazione dal 1993 del sistema avanzato Multi Dimensional Voice Program (MDVP) è oggi possibile un'analisi dettagliata di ben 33 parametri quantitativi dell'emissione vocale, che consentono la valutazione sia della frequenza fondamentale, dell'ampiezza e del bilancio energetico spettrale, che della presenza di interruzioni di sonorità e diplofonia. Nonostante le potenzialità speculative di tale sistema, allo stato attuale non si è ancora avuta una larga diffusione dello stesso, anche per l'assenza di una metodologia procedurale standardizzata. Sono ancora poche, infatti, le ricerche in letteratura che prendono in considerazione aspetti normativi sia procedurali che di analisi. Scopo di questo lavoro è di fornire i risultati dell'analisi dei parametri di perturbazione d'ampiezza in soggetti adulti eufonici con età compresa tra 19 e 59 anni, da noi considerati i più significativi nella determinazione della gravità del disturbo di fonazione. Lo studio è stato condotto su un campione di 35 soggetti eufonici (età media 31,6, range 19-59 anni), di cui 24 di sesso femminile e 11 di sesso maschile. La registrazione del segnale vocale è stata effettuata tramite il Kay Computer Speech Lab (CSL) modello 4300B supportato da personal computer con microfono Shure-Prolog SM48 in dotazione, posizionato a 15 cm e angolato di 45°. La saturazione del microfono d'ingresso è stata regolata a 6/9 del canale CH1. L'analisi, effettuata sui 3 sec. centrali di un campione vocalico rappresentato da una /a/ tenuta, è stata realizzata mediante il software MDVP, modello 5105, versione 2,3, utilizzando, per la digitalizzazione del segnale, un sampling rate di 50 kHz. Al fine di ottenere un campione il più possibile senza variazioni d'intensità o frequenza, prima dell'acquisizione del campione, ogni soggetto è stato sottoposto ad allenamento, con almeno tre prove di fonazione. Sono state ritenute idonee le emissioni con intensità compresa tra 55 e 65 dB e stabilità dello spettro, scartando tutte le altre. Il rumore ambientale è stato costantemente monitorato e mantenuto al di sotto di 30 dB. Nei 24 soggetti di sesso femminile sono stati riscontrati valori di Shimmer assoluto (ShdB) di 0,203 dB, valori di Shimmer relativo (Shim%) di 2,226, valori di quoziente di perturbazione d'ampiezza (APQ) di 1,758, valori di quoziente mediato

= 2.892, amplitude perturbation quotient (APQ) = 2.611, smoothed amplitude perturbation quotient (SAPQ) = 3.433, peak amplitude variation (vAm) = 6.385. Statistical analysis (t test) showed no statistically significant difference between sexes, thus a single set of rules was used for both sexes.

di perturbazione d'ampiezza (SAPQ) di 3,309 e valori di variazione di ampiezza di picco (vAm) di 7,042. Negli 11 soggetti di sesso maschile sono stati riscontrati valori di Shimmer assoluto (ShdB) di 0,269 dB, valori di Shimmer relativo (Shim%) di 2,892, valori di quoziente di perturbazione d'ampiezza (APQ) di 2,611, valori di quoziente mediato di perturbazione d'ampiezza (SAPQ) di 3,433 e valori di variazione di ampiezza di picco (vAm) di 6,385. II confronto statistico (T test) tra i due sessi non ha messo in evidenza differenze statisticamente significative, consentendo quindi di considerare un'unica normativa per i due sessi.

Introduction

The first voice spectrographs were introduced soon after the Second World War for the medico-legal identification of individuals. However, it was only in 1961 that Lieberman [1] carried out the first prescriptive study on the correlation between laryngeal phonatory behaviour, in terms of vocal fold vibratory function, and voice electroacoustic measures, as its objective evaluation, thus defining the Pitch Perturbation Factor (PPF). This factor was later used in the clinical field in order to visualize changes in the spectral pattern of the vocal signal over time. In the late 70s, the introduction by Kay Elemetrics of the first digital spectrograph (DSP Sonograph) significantly improved the analysis possibilities. Despite this, the first studies focused mainly on the qualitative description of the easiest vocal parameters (fundamental frequency, time, amplitude), providing just a few of those quantitative elements which are definitely more significant when describing a disease [2-11]. In the attempt to identify dysphonic electroacoustic features, many Authors focused on the search for new parameters which could be objectively quantified. Their first studies, therefore, led to the identification of temporal variation of fundamental frequency, as the jitter [12], of amplitude, as the shimmer [13], and of other parameters relating to the share of noise present in the vocal signal: H/N ratio [14], SNL [15], S/N ratio [16], NNE [17] and RNL [18]. At the end of 1989, a new processing card, the Voice Analysis Program, was marketed allowing more specific real- and non-real-time analysis, leading to the identification of parameters such as the "breath" quotient relating to the energy content at high frequencies and the H/N ratio. At present, the software update with the MDVP [19] advanced system enables the detailed analysis of 33 vocal parameters which allow the evaluation of both fundamental frequency/amplitude/spectral energy balance and the presence of any breakdown in sonority and diplophony. Despite the speculative potentials of the above-mentioned system [20-22], it is not yet widely used, probably due also to the lack of a standard procedure [23-26] and the reference to provisions relating to extremely limited samples [23 26].

In this context, it would be useful and worthwhile carrying out an extensive study for standardization of methodology and to establish a database with normative values for children, as well as for young and older adults. A normative database is the first step towards the analysis of possible correlations between the single voice disorders and the degree of deviation from normal values; in other words, the analysis of deviant patterns, if existent, able to significantly characterize the different pathological patterns for use in differential diagnosis.

The present study aims to provide a first prescriptive analysis of amplitude variation parameters in euphonic adult patients aged 19-59 years. Some Authors [27-28] consider amplitude variations as the most significant factors in determining the severity of the phonation disorder. The intensity, indeed, reflects the changes in subglottic pressure, adjusted, in turn, by both laryngeal and respiratory muscles and elastic forces and, also, by their coordination.

In pathological conditions, the vibratory cycle can take place without complete closure of the glottis, e.g. due to a functional failure of interarytenoidal muscles (which can cause a permanent rear triangular gap due to the lack of contact with the intercartilagineous area), a failure of the adductors (in which case the glottis appears as oval) or, finally, to all those productive diseases which, due to interference of a neoformation (nodular, polypoidal etc.), make full contact between the vocal folds impossible. This lack of complete closure creates a breath which compromises the ability to produce a constant sound emission and leads to changes which can affect the vibratory amplitude without affecting the frequency and which can already be highlighted in the peak-to-peak analysis. The above-mentioned inability can also be caused by extreme tension accompanied by folds hyperadduction. It can be highlighted in the short-term control analysis of the vocal emission amplitude and is more and more frequently associated with frequency perturbations. The importance of long-term evaluation appears clear in the case of a phonatory disorder (with laryngeal negative objectivity) due mainly to inadequate expiratory air flow, thus causing changes which can be detected only if examining many consecutive periods rather than single periods.

Materials and methods

The present study has been carried out on 35 euphonic patients (mean age 31.6 years, range 19-59), of whom 24 female (mean age 33.5 years, range 19-47) and 11 male (mean age 42 years, range 28-59). In order to exclude any vocal disorder, patients had to meet the following criteria: dysphonia negative history, negative physical and videolaryngoscopic pattern, a zero overall value in acoustic-perceptive evaluation with GIRBAS scale [29]. The vocal signal has been recorded with a Kay Computer Speech Lab (CSL) 4300B supported by a personal computer including a Shure-Prolog SM48 microphone, located at a distance of 15 cm (Kay) and at an angle of 45°. Background noise (< 30 dB) has been constantly monitored. Analysis of a voice sample, directly recorded using digital technology and with a sample frequency of 50 kHz, has been carried out using the 2.3 version of the MDVP 5105 software. In accordance with the vocal material used in the system standardization and with almost all the studies carried out using the same material, we have chosen the vowel /a/ as sound sample. The vowel has been kept at a "conversation" intensity (55-65 dB acceptability range) after the patient had previously been trained for an emission, as constant as possible, without intensity or frequency changes (three training tests). For further standardization of the analysis procedure, the saturation point has been constantly set on 6/9 of the acquisition channel (CH1). Each patient has been asked to carry out a phonation lasting at least 6 seconds, out of which the central 3 seconds have been extracted under default provisions being the most regular and those affected the least by signal perturbations due to the onset and end of the vocal signal [30]. As already pointed out, of all the parameters calculated by the system, only those related to amplitude perturbations were evaluated:

- absolute shimmer (Sh dB): mean absolute variability, peak-to-peak, of amplitude in every single period. This value provides information concerning the short-term amplitude that shows irregularity (cycle-to-cycle);
- relative shimmer (Shim): mean relative variability, peak-to-peak, of amplitude in each single period;
- peak amplitude variation (vAm): mean relative variability of amplitude, 11 periods out of 11, with a 1-step peak-to-peak;
- amplitude perturbation quotient (APQ): mean relative variability of amplitude, s periods out of s (default s = 55), with a 1-step peak-to-peak;
- smoothed amplitude perturbation quotient (SAPQ): relative variability of the peak-to-peak amplitude standard deviation compared to the average peak-to-peak amplitude.

Of the data obtained, we calculated the mean values, standard deviations and any statistic correlation. Results have then been compared to those obtained in different centres using the same voice evaluation system.

Table I. Values detected in females.

	Min value	Max value	Mean value	SD
ShdB	0.155	0.279	0.203	0.046
Shim %	1.274	3.176	2.226	0.647
APQ	1.443	2.005	1.758	0.223
SAPQ	2.264	5.189	3.309	0.619
vAm	5.909	8.099	7.042	0.767

SD = standard deviation

Table II. Values detected in males.

	Min value	Max value	Mean value	SD
ShdB	0.227	0.354	0.269	0.049
Shim %	2.532	3.011	2.892	0.175
APQ	2.084	3.101	2.611	0.378
SAPQ	2.495	5.002	3.433	0.756
vAm	5.892	7.262	6.385	0.536

SD = standard deviation

Table III. Overall prescriptive values.

	Min value	Max value	Mean value	SD
ShdB	0.155	0.354	0.233	0.059
Shim %	1.274	3.176	2.538	0.615
APQ	1.443	3.101	2.101	0.483
SAPQ	2.264	5.189	3.212	0.620
vAm	5.909	8.099	6.800	0.658

SD = standard deviation

Results

The 24 adult females examined showed the following results: ShdB = 0.203 dB; Shim% = 2.226; APQ = 1.758; SAPQ = 3.309; vAm = 7.042 (Table I).
The 11 males examined showed the following results: ShdB = 0.269 dB; Shim% = 2.892; APQ = 2.611; SAPQ = 3.433; vAm = 6.385 (Tab. II).
The statistical analysis (t test) did not reveal any significant difference, thus the same provisions were taken into consideration for both sexes (Table III).
Comparison with data obtained in other centres did not show any statistically significant difference (Tables IV, V).

Conclusions

Acoustic digital voice analysis basically requires the application and evaluation of a correct set of rules. It is, therefore, necessary to find a common procedure in order to enable different centres to obtain comparable and superimposable values [23-25 31]. The present study shows how the variability detected in previous studies is already being overcome by simply reducing the variability factors while carrying out the test and by following a common protocol. Ursino et al. [23] did not detect any significant difference between males and females. Indeed, the results obtained are comparable to those obtained in the present investigation and to those obtained by De Colle [20]. Conversely, significant differences were detected when comparing data concerning the system default values. Ursino et al. [23] and De Colle [20] used an analysis procedure which is similar to that used in the present study (same distance from the microphone, acquisition of at least 6 seconds and analysis of the 3 central seconds of the recorded signal, same sample frequency, etc.). The one used by default, instead, is a little different from ours (the microphone is located at only 5 cm, the choice of the segment to analyse is randomised, the recording only lasted 4

Table IV. Statistical comparison of values detected in different centres using MDVP system.

	Catanzaro	De Colle, 2001	Ursino, 1999
ShdB	0.233	0.209	0.298
Shim %	2.538	2.399	/
APQ	2.101	1.853	/
SAPQ	3.212	3.014	2.73
vAm	6.800	6.916	8.41

Table V. Statistical variance.

	Catanzaro/De Colle	Catanzaro/Ursino	Total
ShdB	0.000288	0.002112	0.00212
Shim %	0.009661	/	/
APQ	0.030752	/	/
SAPQ	0.019602	0.116162	0.058697
vAm	0.006728	1.29605	0.806265

seconds). Although our results are comparable, in part to those of the other centres, univocal provisions appear to be evident and in order to standardise the procedure should not be delayed. Once again, we emphasize the need for a more extensive evaluation of new acoustic parameters, particularly in normal subjects, according to sex and age, in order to identify all the correct correlations.

References

1 Lieberman P. *Perturbations in vocal pitch*. J Acoust Soc Am 1961;33:597-603.

2 Biondi S, Zappalà M, Amato G. *La spettrografia della voce*. In: Piragine F, Ursino F, editors. *La semeiotica foniatrica strumentale nella patologia vocale non neoplastica*. Relazione Ufficiale XXIV Congresso SIFEL, Acta Phon Lat 1990;12:199-236.

3 Eskenazi L, Childers DG. *Acoustic correlates of vocal quality*. J Speech Hearing Res 1990;33:298-306.

4 Sanderson RJ, Maran AGD. *The quantitative analysis of dysphonia*. Clin Otolaryngol 1992;17:440-3.

5 Di Nicola V. *Perturbazioni del suono vocale: jitter e shimmer*. Acta Phon Lat 1995;17:243-52.

6 Wolfe V, Fitch J, Cornell R. *Acoustic prediction of severity in commonly occurring voice problems*. J Speech Hearing Res 1995;38:273-9.

7 Di Nicola V, D'Onofrio A, Mangiatordi F, Fiorella ML. *Analisi acustica della voce normale: correlazioni tra jitter, shimmer e H/N ratio*. Acta Phon Lat 1998;2:313-23.

8 Di Nicola V, D'Onofrio A, Mangiatordi F, Fiorella ML. *La variazione dei parametri acustici della voce (H/N ratio, jitter e shimmer) dopo fonochirurgia*. Acta Phon Lat 1998;2:324-30.

9 Barillari U, Previtero G. *Valutazione dell'indice spettrografico H/N*. Acta Phon Lat 1998;20:216-23.

10 Ferrero FE, Accordi M, Accordi D, De Filippis C, Staffieri A. *Semeiologia avanzata della funzione vocale*. Acta Phon Lat 1998;20:33-58.

11 Case JL. *Technology in the assessment of voice disorders*. Semin Speech Lang 1999;20:169-84.

12 Ricci Maccarini A, De Colle W, Lucchini E, Casolino D. *L'esame spettroacustico della voce*. Relazione Ufficiale LXXXIX Congresso Nazionale SIO 2002:150-82.

13 Jacob L. *Normative study of laryngeal jitter*. Unpublished master's thesis, University of Kansas, Lawrence 1968.

14 Koike Y. *Vowel amplitude modulations in patients with laryngeal diseases*. JASA 1969;45:839-44.

15 Yumoto E, Sasaski Y, Okamura H. *Harmonics-to-noise ratio and psychophysical measurement of the degree of hoarseness*. J Speech Hearing Res 1982;27:2-6.

16 Deal R, Emanuel F. *Some waveform and spectral features of vowel roughness*. J Speech Hearing Res 1978;2:250-64.

17 Kojima H, Gould W, Lambiase A, Isshiki N. *Computer analysis of hoarseness*. Acta Oto-Laryngologica 1980;89:547-54.

18 Hirano M, Hibi S, Yoshida T, Hirada Y, Kasuya H. *Acoustic analysis of pathological voice*. Acta Otolaryngol 1988; 105:432-44.

19 Fujiu M, Hibi S, Hirano M. *An improved technique for measurement of the Relative Noise Level using sound spectrograph*. Folia Phoniat 1988;40:53-67.

20 De Colle W. *Voce & computer: analisi acustica digitale del segnale verbale*. Torino: Ed Omega 2001.

21 Ferrero FE, Lanni R, De Colle W. *Primi risultati di uno studio per la validazione del sistema MDVP come strumento per la caratterizzazione multiparametrica della voce*. Acta Phon Lat 1995;17:161-80.

22 Ferrero FE, Lanni R. *Valutazione del sistema CSL-MDVP con campioni di voce normale, patologica, digitale e analogica, microfonica ed elettroglottografica*. Quaderni del CSRF-CNR. Padova; Ed. CSRF CNR 1995;13:257-89.

23 Ursino F, Matteucci F, Trianni V, Della Rossa S, Piragine F. *Il ruolo clinico del Multi Dimensional Voice Program nelle disfonie disfunzionali*. Acta Phon Lat 1999;21:306-12.

24 Titze I, Horii Y, Scherer R. *Some technical considerations in voice perturbation measurements*. J Speech Hearing Res 1987;30:252-60.

25 Karnell MP, Scherer RS, Fisher LB. *Comparison of acoustic voice perturbation measures among three independent voice laboratories*. J Speech Hearing Res 1991;34:781-90.

26 Titze I, Winholtz W. *Effect of microphone type and placement on voice perturbation measurements*. J Speech Hearing Res 1993;36:1177-90.

27 Baken RJ, Orlikoff RF. *Changes in vocal fundamental frequency at the segmental level: control during voiced fricatives*. J Speech Hear Res 1988;31:207-11.

28 Peterson GE, Barney HL. *Control methods used in a study of the vowels*. J Acoust Soc Am 1952;24:175-84.

29 Dejonckere PH, Remacle M, Fresnel-Elbaz E, Woisard V, Crevier-Buchman L, Millet B. *Differentiated perceptual evaluation of pathological voice quality: reliability and correlations with acoustic measurements.* Rev Laryngol Otol Rhinol 1996;117:219-24.

30 Gelfer MP. *Fundamental frequency, intensity, and vowel selection: effects on measures of phonatory stability*. J Speech Hear Res 1995;38:1189-98.

31 Di Nicola V, Fiorella ML, Luperto P, Fiorella R. *L'analisi digitale acustica della voce: metodologia procedurale*. Acta Phon Lat 2000;22:401-7.

■ Riceived: July 15, 2003
Accepted: September 16, 2004

■ Correspondence: Prof. Ettore Cassandro, Cattedra di Audiologia dell'Università "Magna Graecia" di Catanzaro, c/o Clinica Villa Del Sole, viale Pio X 202, 88100 Catanzaro, Italy - Fax +39 0961 742021 - E-mail: cassandro@unicz.it