Acoustic Assessment of the Voices of Children Using Nonlinear Analysis: Proposal for Assessment and Vocal Monitoring

*Leonardo Wanderley Lopes, †Silvana Luciene do Nascimento Cunha Costa,
†Washington César de Almeida Costa, †Suzete Élida Nóbrega Correia, and †Vinícius Jefferson Dias Vieira,
*†*João Pessoa, Paraíba, Brazil*

Summary: Objective. To analyze the accuracy of recurrence measurements, both isolated and combined, to assess the intensity of vocal disorders in children.

Method. A total of 93 children of both sexes (48 girls and 45 boys), aged between 3 and 10 years, participated. The vocal-deviation intensity was evaluated by the consensus of three speech therapists from the pronunciation of vowel /ɛ/ using the visual analog scale. In the acoustic analysis, eight recurrence plot characteristics were evaluated and extracted with neighborhood radius values that maintained the recurrence rate at 1%, 2%, 3%, 4%, and 5%. The classification was performed using quadratic discriminant analysis applied for individual and combined measurements. The performance was evaluated by measuring the accuracy, which related the cases correctly classified to all the analyzed cases.

Results. In the classification cases concerning individual measure performance, the trapping time and maximum length of the diagonal lines showed the best classification potential to discriminate between healthy and disturbed voices, with accuracy rates above 80%. In the healthy and mild deviation cases, the trend (TREND) measure was also relevant. For the mild versus moderate deviation classification, the best performance was obtained by the TREND measure (85.00% ± 7.64%). A gain was obtained in the classification rate when the measures of recurrence were combined, reaching an accuracy of 95.00% ± 5.00%, for discriminating between healthy voices and those with mild deviation.

Conclusions. The measures of recurrence, either alone or combined, may be useful in detecting healthy and disturbed voices and in differentiating the intensity of vocal disorders in children.

Key Words: Voice analysis–Acoustic–Children–Nonlinear analysis.

INTRODUCTION

The high prevalence of dysphonia in children requires special attention in the assessment and diagnosis of their voices by developing objective measures that provide an understanding of the intensity of vocal deviation and its manifestation in different periods aged between 3 and 9 years.[1–4]

An acoustic signal is a complex product of the nonlinear interaction between the aerodynamic and biomechanical properties of the vocal production system. Therefore, a large correlation exists between laryngeal physiology and acoustic measures.[5]

Acoustic assessment is considered less subjective than perceptive analysis to provide quantification of the vocal deviation. However, findings are incipient to assert that a single instrumental assessment can be consistently and strongly correlated with perceptual analysis.[6]

An incessant search has been performed by clinicians and researchers to develop noninvasive measurements with high discriminatory power, enabling screening, assessment, diagnosis, and monitoring of voice disorders. In the literature, no consensus has been reached on the set of measures with greater accuracy for evaluating voice signals. Therefore, an urgent need exists for studies to investigate the discriminatory power of individual and/or combined acoustic measures that can be used in the classification of healthy and disturbed voices.[7]

In recent years, techniques based on nonlinear dynamic analysis and chaos theory have been used in acoustic analysis both in the classification of healthy and disturbed voices because of the different degrees of vocal disorder and the different types of voice quality.[8,9]

The methods of nonlinear dynamics can analyze irregular behavior and may be important in different studies of vocal production, including studies that evaluate the effectiveness of the treatment provided, classify voices in different degrees of modification, and differentiate between healthy and disturbed voices, and even contribute to the diagnosis of laryngeal diseases.[10–12]

Among the measures used in the nonlinear dynamic analysis are the recurrence plots (RPs) and their quantification measures, which are based on the Poincaré recurrence theorem and have the advantage of working on short and nonstationary series.[13–15]

RP is a square matrix that represents the evolution of a dynamic system, which can be visually analyzed qualitatively and subjectively by observing its structure of isolated points as well as the diagonal, vertical, and horizontal lines (Figure 1).

In the case of the voice production system, only one state variable is known, which is the digitized voice signal. For such representation, the temporal series under analysis needs to be

Accepted for publication February 25, 2014.
The research was conducted in the Department of Speech and Language Pathology in the Universidade Federal da Paraíba—UFPB—João Pessoa, Paraíba, Brazil.
Conflict of interests: None to report.
From the *Department of Speech and Language Pathology, Universidade Federal da Paraíba, João Pessoa, Paraíba, Brazil; and the †Electrical Engineering Department, Instituto Federal de Educação, Ciência e Tecnologia, João Pessoa, Paraíba, Brazil.
Address correspondence and reprint requests to Leonardo Wanderley Lopes, Departamento de Fonoaudiologia, Centro de Ciências da Saúde—Campus I, Universidade Federal da Paraíba, Cidade Universitária—Campus I, Castelo Branco, João Pessoa 58051-900, Paraíba, Brazil. E-mail: lwlopes@hotmail.com
Journal of Voice, Vol. 28, No. 5, pp. 565-573
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2014.02.013

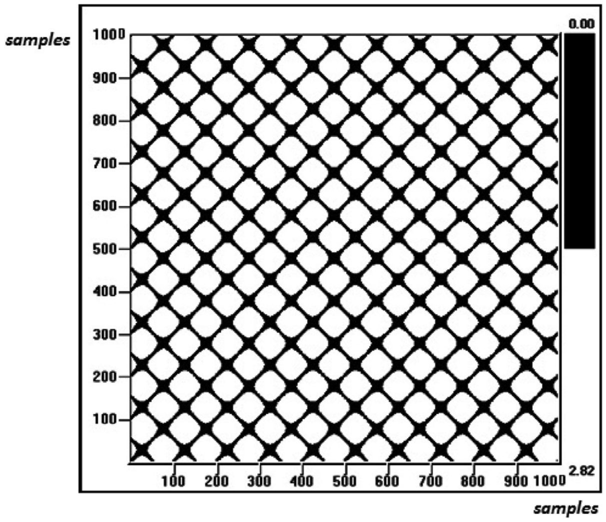

FIGURE 1. Recurrence plot of a pure tone (sinusoidal signal) with time delay $\tau = 2$ and embedded immersion dimension $m = 2$.

immersed in a phase space from the Takens' immersion theorem,[16] given by

$$\vec{x}_i = \{x(t_i), x(t_i + \tau), x(t_i + 2\tau), \ldots, x(t_i + (m-1)\tau)\}, \quad (1)$$

which represents the ith vector constructed from the original signal and its m lagged versions in time, where τ is the delay applied to the voice signal and m is the immersion dimension. Using the Takens' theorem, variables that attempt to reproduce the dynamics of speech production system are created.

RP is formed according to the following equation[17]:

$$\vec{R}_{i,j}^{m,\varepsilon} = \theta\left(\varepsilon - \left\|\vec{x}_i - \vec{x}_j\right\|\right), \quad \vec{x}_i \in \Re^m, \quad i,j = 1, \ldots, N. \quad (2)$$

where

- N is the number of $\vec{x}_i$ states considered
- ε is the neighborhood radius (*threshold*) at point $\vec{x}_i$
- $\|.\|$ is the neighborhood measure, usually the Euclidean measure
- $\theta(.)$ is the unit step function
- m is the immersion dimension.

If $\vec{R}_{i,j} = 1$, the state is considered recurrent; as a result, a black dot is marked on the RP. If $\vec{R}_{i,j} = 0$, the state is "nonrecurrent," and a white dot is marked on the RP.[7,15,18,19]

Figures 1 and 2 show examples of the RPs obtained from classic signals such as a sine wave (a pure tone, a deterministic signal), without noise addition, and white noise (a random signal that contains equal power within any frequency band width). The diagonal lines of the RPs in the periodic signals are completely filled (Figure 1), whereas in the nonstationary random signals such as the white noise, the plot shows the predominance of isolated points (Figure 2).

The RPs obtained from signals of infant voices are shown in Figures 3–5, which represent the voice signals of a healthy child, with mild and moderate degree of deviation, respectively.

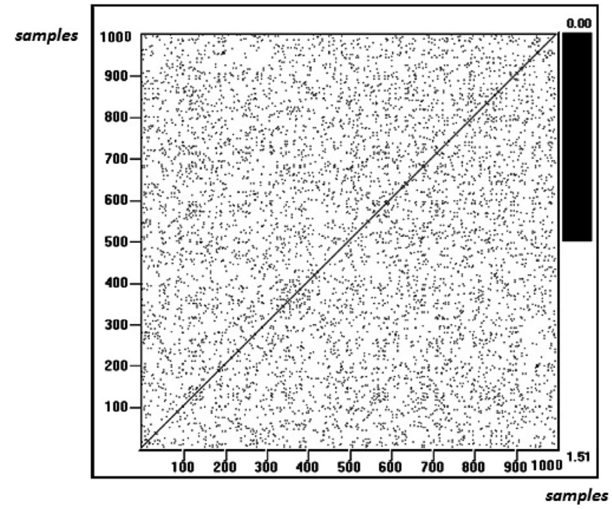

FIGURE 2. Recurrence plot obtained from a white noise signal (time delay $\tau = 1$ and embedded immersion dimension $m = 3$).

Because of the near-periodicity characteristic of the voice signals considered as normal, they tend to exhibit increased formation of diagonal structures. The higher the degree of voice disorder, the more the voice signal loses this characteristic, showing a noisier aspect with the appearance of isolated points on the graph.

The visual analysis of the RPs is subjective and can lead to different interpretations. The recurrence quantification analysis (RQA) uses pattern recognition algorithms to quantify the recurrence features depicted in RPs, which is more objective than the visual inspection of the graphs. To add more robustness to the analysis of the quantifying structures present in the RPs, the recurrence quantification measures[14,15] were created. Such measures include the following: recurrence rate (REC), determinism (DET), maximum length of the diagonal lines (L_{max}), Shannon entropy (ENTR), trend (TREND), laminarity (LAM), mean length of vertical structures (TT), and maximum length of vertical structures (V_{max}), which are described in the following.

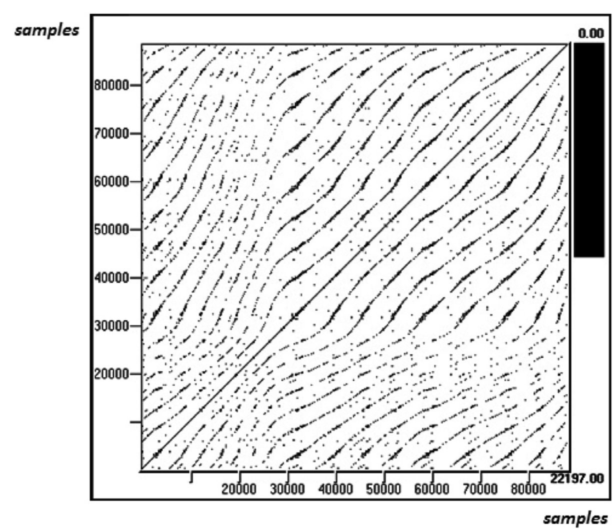

FIGURE 3. Recurrence plot obtained from a healthy child's voice (time delay $\tau = 14$ and embedded immersion dimension $m = 5$).