Study of Complexity in Normal and Pathological Speech Signals*

M. E. Torres[1+], L. G. Gamero[1], H. L. Rufiner[2],
C. Martínez[2], D. H. Milone[2], G. Schlotthauer[1]
[1] Laboratorio de Dinámicas no Lineales y Señales, [2] Laboratorio de Cibernética,
Facultad de Ingeniería - Bioingeniería, Universidad Nacional de Entre Ríos,
CC 47, Suc. 3, (3100) Paraná, Entre Ríos, Argentina

Abstract— **The application of complexity measures to the analysis of different biological signals have contributed to give a better understanding of the dynamical systems involved in their generation. In this work we present a comparative study of the complexity of speech signals from subjects with normal phonation and patients with laryngeal pathologies of the vocal system. Different complexity measures were considered in this study. Our results suggest that some of them would allow to discriminate between normal and pathological voices. This result could give an indication in order to assist the diagnostic or the treatment in the clinical practice.**

Keywords— **Complexity measures; Entropy; Speech Disorders; Laryngeal pathologies**

I. Introduction

In the study of nonlinear dynamical systems we often deal with experimental data where the underlying dynamics is not well known. Most of them present a rich variety of self-oscillating regimes that involve either regular or complex behavior [1].

When a non linear dynamics can be represented by differential equations, there exist a variety of methods that provide a qualitative and quantitative characterization of its behavior. In the most frequent case, when explicit equations are not available, other approaches have been proposed in which the signal stationarity is taken for granted [1], [3], [4], [5].

Different notions of entropy have been used in order to characterize the complexity degree of differential and difference equations. The application of quantitative measures of complexity to the analysis of such signals have contributed to give a better understanding of the dynamical systems [6].

The algorithm complexity for sequences of finite length was suggested by Ziv and Lempel [7], and it is related with the number of distinct substrings and the rate of their recurrence along the given sequence. Ziv-Lempel (LZ) complexity can be a finer measure than the Lyapunov exponents for characterizing order [8]. Another measure of complexity (regularity) is the Approximate Entropy (ApEn) [9] that allows complex system classification. It is well known its ability to quantify complexity with a reduced amount of data, although it requires a relative high computational burden. LZ complexity and ApEn are not measures of chaos but they quantify the regularity embedded in the time series. The main disadvantage of these techniques, as most of the usual complexity measures (correlation dimension, Lyapunov exponents), is the large amount of data required for their estimation [1][3]. The renormalized entropy, opposite to other complexity measures, is defined relative to a fixed state [6] and has been used to indicate transitions from periodic to chaotic behavior as well as between different types of chaos. Basically it is the Kullback Information respect to a state with a given value of effective energy.

The classical Shannon entropy, which comes from information theory, describes the evolution of order. The more general Harvda-Charvat-Darovczy-Tsallis (q-entropies) [6], [10], [11] and their corresponding relative informations are considered as complexity measure estimators. In [12] we have presented a comparative analysis of different complexity measures and the involved algorithms, taking into account their computational cost and their robustness in the presence of noise. In [13] we have included complexity measures in order to improve the robustness of an automatic speech recognition system.

In this work we deal with pathological speech signals, which could present a nonlinear underlying dynamics, and our goal is to compare different complexity measures in order to characterize them. With this in mind we present and evaluate different complexity measures, derived from Shannon entropy, q-entropies, the corresponding relative information measures, and the LZ and ApEn. These preliminary results suggest that complexity measures would allow us to discriminate normal and pathological voices.

II. Materials and Methods

A. Complexity Measures

In this section we briefly review the complexity measures considered in this study. For more comprehensive discussions, see e.g. [12], [14].

1) Approximate Entropy

The ApEn can classify a system given at least 1000 data values in diverse settings including deterministic, chaotic and stochastic processes [9]. The capability to discern changing complexity from such relatively small amount of data holds promise for application of ApEn to a variety of contexts.

Given a finite time series $x(1)$, $x(2)$, ..., $x(N)$, a fix positive integer m and a positive real number r, we consider the embedding vectors $u(1),..., u(N-m+1)$ in R^m, where $u(j)=[x(j), x(j+1), ..., x(j+m-1)]^t$.

* This work is supported by U.N.E.R., under Projects Nº 6036, 6060 and 6062.

+ Corresponding author: metorres@ceride.gov.ar

For each i, $1 \le i \le N\text{-}m\text{-}1$,

C_i^m = number of $j \le N-m+1$ such that $d(u(i),u(j)) \le r$,

where $d(u(i),u(j))$ is the l_∞ norm. The ApEn is defined as:

$$ApEn(m,r) = lim_{N\to\infty}\left[\phi^m(r) - \phi^{m+1}(r)\right],$$

where

$$\phi^m = \sum_{i=1}^{N-m+1} \ln C_i^n(r)/(N-m+1).$$

2) Ziv-Lempel

As proposed in [7], the complexity of a finite sequence can be evaluated from the point of view of a simple self delimiting learning machine which, as it scans a given N-digits sequence $\mathbf{x} = x(1), ..., x(N)$ from left to right, adds a new word to its memory every time it discovers a substring of consecutive digits not previously encountered. The size of the compiled vocabulary and the rate at which new words are encountered along $\mathbf{x}$ serve as the basic ingredients in the LZ complexity evaluation.

B. Entropies

Given a signal $\mathbf{x}$, we can define its Shannon entropy as [16]:

$$H = -\sum_{i=1}^{M} p_i \ln(p_i),$$

where p_i is the probability that the signal belongs to a considered interval and with the understanding that $p.\ln(p) = 0$ if $p = 0$ and M is the partitions number. The entropy H is a measure of the information needed to locate a system in a certain state, meaning that H is a measure of our ignorance about the physic system.

The Harvda-Charvat-Daróvczy-Tsallis [6], [10], [11], q-entropy, that depends on a single real parameter $q \neq 1$, reads as:

$$H_q = \frac{1}{q-1}\sum_{i=1}^{M}\left(p_i - p_i^{\,q}\right)$$

C. Relative entropies

The relative entropy (or Kullback-Leiber distance) $K(f\,|\,g)$ between two probability densities f and g is defined by [17]:

$$K(f\,|\,g) = \int_X f(x)\ln\left[f(x)/g(x)\right]dx,$$

with the understanding that $y.\ln(y) = 0$ if $y = 0$.

In the q-entropies case for $q \neq 1$ and $q \in \Re$, the corresponding relative q-entropies are given by [14]:

$$D_q(f\,|\,g) = \frac{1}{1-q}\int_X f(x)\left[1-\left(\frac{f(x)}{g(x)}\right)^{q-1}\right]dx.$$

The Kullback divergence $K_D(f\,|\,g)$ between two probability densities f and g is defined by [17]:

$$K_D(f\,|\,g) = D(f\,|\,g) - D(g\,|\,f).$$

In our case, given two probabilities p_i and r_i, such that the signal belongs to different intervals, the corresponding discrete versions read as:

$$D(p\,|\,r) = \sum_{i=1}^{M} p_i \ln\left(p_i/r_i\right)$$

and

$$D_q(p\,|\,r) = \frac{1}{1-q}\sum_{i=1}^{M} p_i\left[1-\left(p_i/r_i\right)^{q-1}\right].$$

D. Speech Signals

The complexity measures were obtained from recordings of people's voices with normal phonation and patients with pathologies of the vocal system. Each signal is a recording of the sustained phonation of a vowel or a vocalic phoneme. The use of a vocal type stimulus has certain advantages. First, the isolated vowels are used in the routine of clinical practice for evaluation of the quality of pathological voices. Second, the objective measures are relatively direct, compared with the continuous speech. Also, they allow an easy and effective separation among normal and pathological voices [18]. The study of the continuous speech is a superior objective and an evident next step. However, first valid results are required based on stimuli of smaller complexity. In the upper part of figure 1, a segment of normal and pathological voices are shown, where the temporal differences of both waveforms can be appreciated, while in the bottom the spectra of the same segments are observed. A difference that is appreciated at first sight is the appearance of high frequency components in the pathological case.

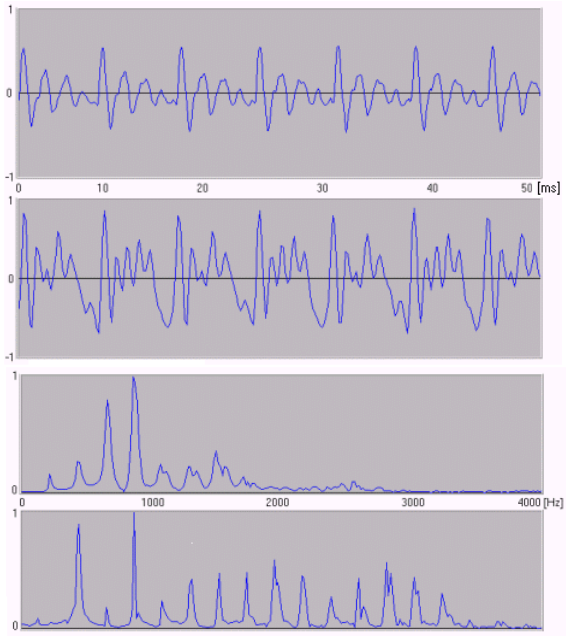

Fig. 4. Normal and pathological speech signals in time and frequency domains.

1) Data of normal voice

The speech signals of normal voice were obtained from the corpus of continuous voice TIMIT [19]. From the

sentence SA1.WAV, of the original group of training sentences, signal portions that contain the phoneme /aa/ were extracted. When creating the different sets of patterns, the speakers were selected at random among the dialectical regions DR1 to DR8, in such a way of having represented a wide variety of dialects and not to repeat the patterns.

2) Data of pathological voice

Signals of pathological voice were obtained from a library of recordings of voices taken in VA Hospital (West L.A.) by investigators of the Speech Processing and Auditory Perception Laboratory (SPAPL), UCLA. The signals were recorded with a miniature microphone mounted on the head AKG C410, placed to 4 cm of the patient's lips. The signals were gone by a lowpass filter of 8 Khz, digitized directly to 20 Khz and sampled to 10 Khz. A segment of 1 second was extracted of the half portion of each recording [20]. For the purpose of this work, the signals were resampled to 16 Khz, to obtain the same temporal reference that the signals of TIMIT.

The classification of the signals was carried out by the mentioned investigation team, being contained in the following categories: rough and rough-breathy: 11 files, bicyclic (also well-known as diplophony): 8, rough-bicyclic: 1, strained-breathy: 2, and strained-rough: 2.

III. Results and Discussion

Tables I and II present the results evaluating the complexity measures in records of speaking feminine and masculine respectively on both normal and pathological conditions. They are expressed as mean values obtained through the evaluation of sliding windows of 500 samples from records of 2000 points. This is equivalent to use an analysis window of 33 ms approximately. In order to avoid the dependency of the complexity measures with the signal energy, all the records were normalized to have unit standard deviation previous to the processing. The presence of significant differences in the estimated parameters was evaluated using the unpaired Student's *t*-test. As can be appreciated, some complexity measures do not present meaningful differences in the mean values corresponding to normal and pathological signals, as for example the Shannon entropy. The LZ presents a high variance in women and in the case of the men the mean value is similar. The Kullback relative entropy K offers a good discrimination in the male cases, but not in the case of the feminine voices. H_q, D_q, and K_q presented significative differences on both groups.

Figures 2 and 3 present the results of the temporary evolution of the mean values of all the speakers by window and their standard deviation in the case of the feminine voices. In red line the pathological cases are presented and in blue the normal ones. In the same way, Figures 4 and 5 show the results corresponding to the masculine voices. As can be seen, the time evolution of the considered complexity measures presents a slowly time varying pattern. They allow us to appreciate the conformity with the results described in the corresponding tables. In terms of discrimination, the best results were obtained with the q-divergence H_q for women and the Kullback information K for the men group, with $p<0.005$.

TABLE I
COMPLEXITY MEASURES FOR WOMEN

Normal	H	H_q	D_q	K	K_D	LZ
	0.52	6.13	0.08	0.016	0.105	14.67
	0.57	6.18	0.07	0.007	0.076	18.87
	0.57	5.93	0.07	0.010	0.074	16.87
	0.52	5.98	0.04	0.006	0.056	16.27
	0.54	5.97	0.05	0.005	0.052	17.13
Mean	**0.54**	**6.04**	**0.06**	**0.009**	**0.073**	**16.76**
SD	**0.03**	**0.11**	**0.02**	**0.004**	**0.021**	**1.52**

Pathologic	H	H_q	D_q	K	K_D	LZ
	0.53	6.40	0.02	0.003	0.030	17.13
	0.54	6.41	0.02	0.003	0.023	15.13
	0.56	6.29	0.03	0.006	0.047	20.73
	0.57	6.46	0.02	0.004	0.031	21.60
	0.56	6.11	0.03	0.005	0.038	22.80
Mean	**0.55**	**6.33**	**0.02**	**0.004**	**0.034**	**19.48**
SD	**0.02**	**0.14**	**0.01**	**0.001**	**0.009**	**3.22**
* $p<0.05$		*	*		*	

TABLE II
COMPLEXITY MEASURES FOR MEN

Normal	H	H_q	D_q	K	K_D	LZ
	0.56	6.10	0.07	0.014	0.096	17.87
	0.59	6.27	0.03	0.008	0.042	17.13
	0.51	5.37	0.06	0.007	0.044	19.33
	0.57	6.07	0.06	0.015	0.103	17.73
	0.45	5.07	0.05	0.011	0.054	19.93
Mean	**0.53**	**5.77**	**0.05**	**0.011**	**0.068**	**18.40**
SD	**0.05**	**0.52**	**0.02**	**0.003**	**0.029**	**1.18**

Pathologic	H	H_q	D_q	K	K_D	LZ
	0.54	6.66	0.02	0.002	0.026	19.33
	0.55	6.71	0.02	0.003	0.033	20.07
	0.54	6.69	0.04	0.005	0.045	19.07
	0.53	6.50	0.03	0.004	0.028	19.53
	0.55	6.58	0.02	0.003	0.025	18.60
Mean	**0.54**	**6.63**	**0.03**	**0.003**	**0.031**	**19.32**
SD	**0.01**	**0.09**	**0.01**	**0.001**	**0.008**	**0.54**
* $p<0.05$		*	*	*	*	

IV. Conclusion

In this study, complexity of voice signals was assessed using ApEn, LZ and information theory derived complexity measures. Preliminary results were presented comparing normal and pathological cases stratified by gender. Significative differences were found between groups, suggesting that these measures can be successfully applied to discriminate normal and pathological speech signals. Further studies should consider a larger amount of data and the analysis of each pathology individually.

sinc(*i*) Laboratory for Signals and Computational Intelligence (http://fich.unl.edu.ar/sinc)
M. E. Torres, L. G. Gamero, H. L. Rufiner, C. Martinez, D. H. Milone & G. Schlotthauer; "Study of complexity in normal and pathological speech signals"
Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. pp. 2339--2342, 2003.

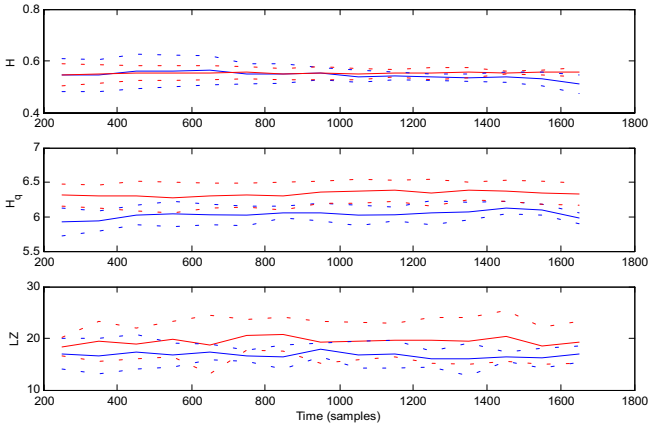

Fig. 2. Feminine voices – In red the pathological and in blue the normal ones.

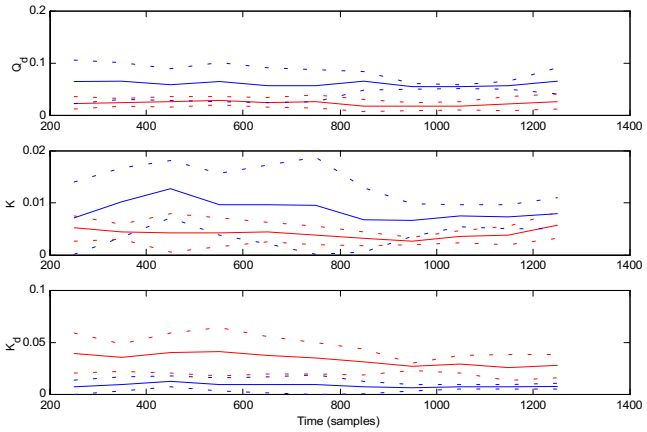

Fig. 3. Feminine voices – In red the pathological and in blue the normal ones.

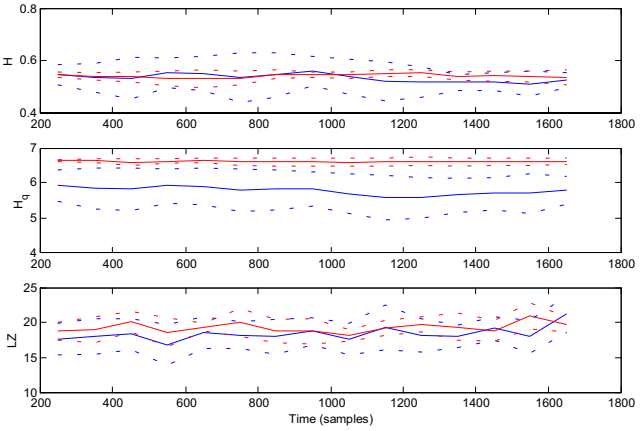

Fig. 4. Masculine voices – In red the pathological and in blue the normal ones.

REFERENCES

[1] R. C. Conant, "Detecting subsystems of a complex system," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 2, pp. 550-553, Sept 1972.

[2] L. Iasemidis, J. Sackellares, H. Zaveri and W. Williams, "Phase space topography and the Lyapunov exponent of electrocorticograms in partial seizures," *Brain Topogr.*, vol. 2, no. 3, pp. 187-201, 1990.

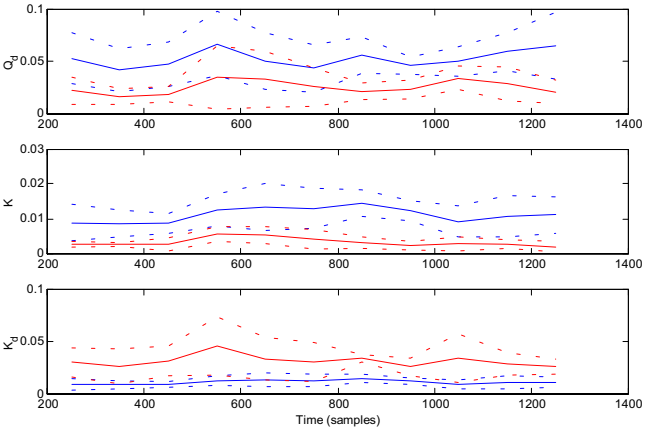

Fig. 5. Masculine voices – In red the pathological and in blue the normal ones.

[3] A. Wolf, J. Swift, H. Swinney and J. Vastano, "Determining Lyapunov exponents from a time series," *Physica D*, vol. 16, pp. 285-317, 1985.

[4] H. Abarbanel, R. Brown, J. Sidorowich, and L. Tsimring "The analysis of observed chaotic data in physical systems," *Rev. of Modern Physics*, vol. 65, no. 4, pp. 1331-1392, 1993.

[5] P. Grassberger, T. Schreiber, and C. Schaffrath "Nonlinear time sequence analysis," *Int. J. of Bifurc. and Chaos*, vol. 1, no. 3, pp. 521-547, 1995.

[6] P. Saparin, A. Witt, J. Kurths and B. Anishchenko, "The renormalized entropy - an appropriate complexity measure?," *J. Chaos, Solitons and Fractals*, vol. 4, no. 10, pp. 1907-1916, 1994.

[7] A. Lempel and J. Ziv, "On the complexity of finite sequences," *IEEE Transactions on Information Theory*, vol. IT-22, no. 1, pp. 75-81, 1976.

[8] K. Kaspar and H. Schuster, "Easily calculable complexity of spatio-temporal patterns," *Phys. Rev. A*, vol. 36, pp. 843-848, 1987.

[9] S. Pincus, "Approximate entropy (ApEn) as a complexity measure," *Chaos*, vol. 5, pp. 110-117, 1995.

[10] Z. Daróvczy, "Generalized information functions," *Inf. Control*, vol. 16, pp. 36-51, 1970.

[11] C. Tsallis, "Somme comments on Boltzmann-Gibbs statistical mechanics," *Chaos, Solitons and Fractals*, vol. 6, pp. 539, 1995. and references therein.

[12] M.E. Torres and L.G. Gamero, "Relative Complexity Changes in Time Series using Information Measures," Physica A, vol. 286, no. 3-4, pp 457-473, 2000.

[13] H. Rufiner, M. Torres, L. Gamero, and D. Milone, "Introducing Complexity Measures in Robust Speech Recognition," submitted to Physica A, May 2003.

[14] M. E. Torres, *El procesamiento de señales ligadas a problemas no lineales.* PhD thesis, Universidad Nacional de Rosario - Argentine, 1999. (Math. D. Thesis).

[15] J. Havrda and F. Charvat, "Quantification method of classification process: Concept of structural α-entropy," *Kybernetica*, vol. 3, pp. 30-35, 1967.

[16] C. Shannon, "A mathematical theory of comunication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379-423, 1948.

[17] T. M. Cover and J. A. Thomas, *Information Theory*. NY: John Wiley and Sons, 1991.

[18] B. Boyanov and S. Hadjitodorov, "Acoustic analysis of pathological voices and screening of laryngeal diseases," *IEEE Engineering in Medicine and Biology Magazine*, vol.16, No 4, pp.74-82, 1997.

[19] Garofolo, Lamel, Fisher, Fiscus, Pallett and Dahlgren, DARPA TIMIT Acoustic-Phonetic Continuous Speech Corpus Documentation, National Institute of Standards and Technology, February 1993.

[20] A. Alwan, P. Bangayan, J. Kreiman, and C. Long, "Time and Frequency Synthesis Parameters for Severe Pathological Voice Qualities," Proc. of ICPhS, Stockholm, Sweden, vol. 2, pp. 250-253, August 1995.