Revista Árete: ISSN. 1657-2513 2012, vol. 12, No. 1, 24-32.

APLICACIÓN DE MÉTODOS ACÚSTICOS LINEALES Y NO LINEALES EN PARALISIS CORDALES UNILATERALES

Natalia Gabriela Elisei¹, Diego Alexis Evin², Francisco Goméz Fernández³, Jorge Alberto Gurlekian⁴

Fecha de Recepción: 29 de junio de 2012 Fecha de Aprobación: 15 de octubre de 2012

Resumen

Se presenta una nueva metodología que combina el análisis acústico y perceptual tradicional con algoritmos de la dinámica no lineal. Específicamente, coeficientes de Lyapunov aplicados a la medición de voces con alta perturbación. Se evaluaron emisiones de la vocal /a/ aislada y sostenida en 15 hablantes del español de Buenos Aires normales y 7 con parálisis unilaterales, de diferente sexo y edad. Las emisiones se registraron digitalmente en una cámara acústica con un micrófono dinámico vocal de respuesta conocida y a una intensidad de habla conversacional. Para la medición de los coeficientes se empleó un algoritmo Wolf (1985) modificado por Giovanni (1999). Se utilizaron índices tradicionales de perturbación de jitter (Horii, 1975 y Milenkovic 1987) y se evaluó perceptualmente el grado de severidad de la disfonía con la escala GRBAS (Hirano,1981). Se presentan para su comparación los valores de los coeficientes de Lyapunov y Jitter en voces normales y con parálisis. Las mayores alteraciones muestran exponentes de Lyapunov promedio de 0,83 versus 0,46 en voces normales. Los resultados y los gráficos obtenidos verifican el comportamiento caótico de la voz y demuestran la utilidad clínica de la medición. Los exponentes distinguen el grupo control respecto del grupo con parálisis.

Palabras clave: algoritmos, parálisis de los pliegues vocales, técnicas de análisis.

Doctora en Fonoaudiología, Universidad de Buenos Aires. Laboratorio de Investigaciones Sensoriales del Consejo Nacional de Investigaciones Científicas y Técnicas (LIS-CONICET). natalia.elisei@gmail.com

Doctor en Computación, Universidad de Buenos Aires. Laboratorio de Investigaciones Sensoriales del Consejo Nacional de Investigaciones Científicas y Técnicas (LIS-CONICET).

Licenciado en Computación, Universidad de Buenos Aires. Laboratorio de Investigaciones Sensoriales del Consejo Nacional de Investigaciones Científicas y Técnicas (LIS-CONICET).

Ingeniero Electrónico. Investigador Independiente. Laboratorio de Investigaciones Sensoriales del Consejo Nacional de Investigaciones Científicas y Técnicas (LIS-CONICET).

APPLICATION OF ACOUSTIC METHODS LINEAR AND NONLINEAR IN CHORDAL UNILATERAL PARALYSIS

Abstract

A new methodology that combines traditional acoustic analysis and perceptual algorithms for nonlinear dynamics is presented. Specifically, Lyapunov coefficients applied to measure high disturbance's voices. Vowel / a / isolated and sustained from 15 normal Buenos Aires Spanish speakers and 7 with unilateral paralysis of different sex and age were evaluated. The emissions were digitally recorded in an acoustic chamber with a dynamic vocal microphone and conversational speech intensity. For the measurement of the coefficients, Wolf algorithm (1985) modified by Giovanni (1999) was used. The traditional measures of perturbation jitter (Horii, 1975 and Milenkovic 1987) and perceptual scale GRBAS (Hirano, 1981) applied to evaluate the degree of severity of dysphonia. The values of Lyapunov coefficients and jitter in normal voices and paralysis were presented for comparison. The voices with highest alterations showed average Lyapunov exponents of 0.83 versus 0.46 in normal voices. The results and graphs obtained show the chaotic behavior of voice and demonstrate the clinical utility of measurement. The exponents differentiate control groups from group with paralysis.

Key words: algorithms, vocal cord paralysis, analytical methods.

INTRODUCCIÓN

En los últimos años se ha generado un interés creciente en desarrollar nuevas herramientas en el estudio de la voz humana dado que las mediciones acústicas y fisiológicas clásicas se correlacionan débilmente -sólo mediante la combinación de todas se llega a un diagnóstico de clasificación. El modelo hasta ahora más conocido y utilizado para caracterizar el sistema de producción de habla normal es el de fuente y filtro basado en teorías acústicas lineales.

En 1995 se clasificaron cualitativamente las señales de la voz en tipo 1, tipo 2 y tipo 3 (Titze, 1995). Las señales de tipo 1 son casi periódicas, el tipo 2 señales que contienen modulaciones o subármonicos fuertes y el tipo 3 señales irregulares y aperiódicas. Con esta clasificación surgió la recomendación de que las medidas acústicas de la perturbación tradicionales como el jitter y el

shimmer son apropiadas solamente para las señales casi periódicas del tipo 1 ya que se evalúa las irregularidades cuantitativamente más que temporalmente basándose en que las anormalidades en la señal se producen de manera aleatoria. Ambas medidas son de variación a corto plazo (de ciclo-a-ciclo) en la frecuencia fundamental (jitter) y en la amplitud (shimmer) lo que presupone, por definición, cierta periodicidad y por eso pueden ser poco confiables cuando se analizan voces con algún grado de desorden como los son los casos de señales tipo 2 y 3 (Titze, Horii, & Scherer, 1987; Bielamowicz, Kreiman, Gerratt, Dauer, & Berke, 1996).

Sin embargo, se ha demostrado que las anormalidades en las voces de pacientes disfónicos, en particular los que presentan parálisis unilateral, no están distribuidas arbitrariamente (Herzel, Berry, Titze, & Saleh, 1994). En algunos casos los patrones se pueden visualizar y en otros se re-

25

quiere de métodos específicos. Se concluyó que, generalmente, la complejidad de la señal afecta la aplicabilidad de los métodos de análisis tradicionales (Titze, Horii, & Scherer, 1987).

En los años 60' y 70' surgió en las teorías matemáticas y aplicadas el concepto de caos determinístico que actualmente se engloba en sistemas dinámicos no lineales. En principio este fenómeno fue pensado como único pero se fue desarrollando y refinando con el aporte de otras disciplinas como la meteorología, la biología molecular, la dinámica de fluidos, medicina, etc. (Hilborn, 2001). Actualmente se ha verificado en diferentes sistemas biológicos y contribuyen a diagnosticar desórdenes fisiológicos y a evaluar los efectos de tratamientos clínicos. En el área de la Otorrinolaringología y Fonoaudiología se conocen trabajos relacionados con el comportamiento caótico que exhibe el sistema auditivo (Gstoettner et al. 1996), se sugirieron medidas para cuantificar el patrón irregular del nistagmus (Aasen, 1993) y procedimientos aplicados al análisis de señales vocales, el análisis no lineal (Kumar & Mullick, 1996; Narayanan & Alwan, 1995).

Si bien las herramientas de análisis acústico tradicional ya son familiares y hasta esenciales en la clínica foniátrica, se ha establecido la existencia de fenómenos de caos en la voz humana como las variaciones en la frecuencia fundamental de la señal de voz y la presencia de subarmónicos (Berry, Herzel, Titze, & Story, 1996; Jiang, Zhang, & Ford, 2003; Zhang & Jiang, 2003). Por este motivo, se implementó la teoría de sistemas dinámicos no lineales (más popularmente llamada la teoría del caos) para describir este comportamiento pseudoaleatorio generado por un sistema que es determinista y no lineal (Peitgen, Jurgens, & Saupe, 1992).

Se define como determinista porque se puede encontrar una expresión para describir el estado de ese sistema dinámico para cada instante de tiempo y no lineal porque un cambio ínfimo de las condiciones iniciales puede llevar a un comportamiento tanto cuantitativo como cualitativo completamente diferente.

Existen varios métodos basados en la dinámica no lineal, incluyendo la dimensión general (dimensión de Hausdorff, dimensión de la información, dimensión de la correlación, etc.), entropía (entropía de Kolmogrov, entropía de segundo-orden, etc.), y exponentes de Lyapunov, que nos permiten describir cuantitativamente el comportamiento caótico. Estas mediciones consideran un factor no tenido en cuenta en los modelos fisiológicos clásicos de cuerdas vocales, la interacción de ellas.

Para cuantificar específicamente estos eventos se desarrolló una técnica que calcula el coeficiente de Lyapunov, el cual mide la inestabilidad vibratoria de las cuerdas vocales y se lo representa en un gráfico, el retrato de fase (Bergé & Pomeau y Vidal, 1984).

En este trabajo se presenta una nueva metodología que combinan el análisis acústico tradicional y perceptual con algoritmos de la dinámica no lineal como los máximos coeficientes de Lyapunov (LE) aplicados a la medición de voces con alta perturbación como las parálisis unilaterales. El propósito de este artículo es presentar el desarrollo de nuevos métodos de análisis que puedan ser aplicados para cuantificar la actividad laríngea.

MATERIALES Y MÉTODOS

Materiales

Se evaluaron las emisiones de la vocal /a/ aislada y sostenida en 22 hablantes del español de Buenos Aires: 15 normales con y 7 con parálisis unilateral. Los pacientes fueron examinados y diagnosticados por medio de estroboscopía.

Métodos

Las emisiones se registraron digitalmente con un nivel de cuantización de 16 bits y una frecuencia de muestreo de 44100 muestras por segundo, en una cámara acústica con un micrófono dinámico vocal AKG de respuesta plana en el rango de frecuencias del habla (100-10.000 Hz) y a una intensidad conversacional. Se realizaron tres grabaciones de cada sujeto para seleccionar la más representativa.

Análisis perceptual

Las grabaciones fueron evaluadas perceptualmente por un panel de 7 jueces experimentados puntuando el grado G de la escala GRBAS (Hirano, 1981), la severidad de la disfonía, con rangos de valor entre 0 (normal) a 3 (alteración severa).

Análisis acústico lineal

Para el análisis objetivo tradicional se utilizaron índices de perturbación a través del cálculo de jitter –en adelante Jitter1– y la relación armónico ruido –SNR1– de (Horii, 1979). Se experimentó también con un método alternativo –en adelante Jitter2– para el cálculo del jitter basado en (Milenkovic, 1987).

El método considera a la onda como la suma de una componente periódica más una componente aperiódica o de error; cada nuevo ciclo de la frecuencia fundamental es detectado como el período que minimiza el valor del error cuadrático medio de la componente aperiódica, con respecto al ciclo anterior. De esta forma, se obtienen para cada ciclo valores del período de la fundamental, la amplitud relativa entre ciclos y parámetros de calidad, los cuales permiten calcular el jitter. Estas mediciones se realizan con un valor

inicial de F0. Para poder obtener buenas mediciones de los parámetros involucrados se utiliza la interpolación cuadrática entre componentes de error, pudiendo trabajar con señales muestreadas a baja resolución (a partir de 8.3 kHz).

Análisis acústico no lineal

En cuanto al análisis objetivo no lineal que estamos proponiendo, existen diversas alternativas para el cálculo del máximo exponente de Lyapunov (LE) a partir de una serie de temporal de datos.

Para la medición de los coeficientes se empleó el algoritmo (Wolf, Swift, Swinney, & Vastano, 1985) con las simplificaciones y parámetros sugeridos por (Giovanni, Ouaknine, & Triglia, 1999; Yu, Ouaknine, & Giovanni, 2000; Yu, Garrel, Nicollas, Ouaknine, & Giovanni, 2007), por lo cual nuestras señales fueron transformadas a una frecuencia de muestreo de 5 kHz para mantener condiciones similares. Este método consiste en reconstruir el estado de un sistema no linear disipativo en una representación espacial de los datos temporales (espacio embebido o retrato de fase) (Holzfuss & Lauterborn, 1989; Eckmann & Ruelle, 1992).

Los retratos de fase muestran las trayectorias 3 dimensionales reconstruidas usando un tiempo de retardo de 1.6 segundos. Si los parámetros externos se mantienen constantes se observara una órbita, figura geométrica denominada atractor. Ver Figura 1 A, B y C.

El concepto de exponente de Lyapunov formaliza la noción de divergencia en las trayectorias generadas a partir de la evolución de un sistema caótico con dos condiciones iniciales ligeramente diferentes. El principio general del algoritmo contempla en cálculo de la divergencia de dos trayectorias que se iniciaron como vecinas. En un sistema estable, ambas trayectorias continua-

ran cercanas y el LE será bajo. En un sistema inestable, las trayectorias que comenzaron siendo vecinas, luego divergirán rápidamente dando LE más altos. Cualquier sistema conteniendo al menos un exponente de Lyapunov positivo se define como caótico.

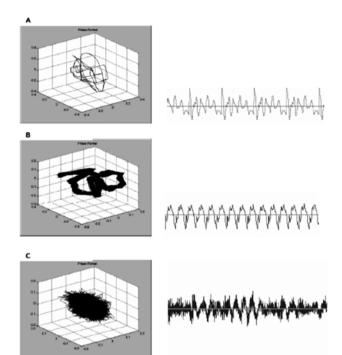


Figura 1: Retratos de fase (izquierdo) correspondientes a una voz sintética o artificial (A), de un sujeto con voz normal (B) y un sujeto con parálisis de cuerda vocal unilateral (C). Se puede observar que las trayectorias y atractores se hacen más complejos en las parálisis confirmando así el fenómeno no lineal.

RESULTADOS

Se presentan los valores de los máximos coeficientes de Lyapunov (LE), Jitter1 y SNR1, Jitter2 y SNR2 calculados según los métodos (Giovanni, Ouaknine, & Triglia, 1999; Horii, 1979; Milenkovic, 1987) respectivamente, para su comparación tanto en voces normales como en voces de pacientes con parálisis. Las mayores alteraciones muestran LE promedio de 0,83 comparado con las mediciones para las voces normales con exponentes promedio de 0,46.

Tabla 1. Valores obtenidos para ambas poblaciones, pacientes normales y con parálisis de cuerda vocal unilateral para cada índice estudiado.

	Normales		Con parálisis cordal unilateral		
	Media	DS	Media	DS	
LE	0,46	0,37	0,83	0,28	
Jitter 1	1,07	0,61	3,06	1,99	
Jitter 2	0,35	0,15	3,24	2,66	
SNR1	3,59	2,33	2,82	1,55	
SNR2	20,32	2,25	11,89	6,73	

Para la mejor visualización en el Figura 2 se presenta en diagramas de caja, la distribución de los máximos exponentes de Lyapunov en función de la percepción. El objetivo de este análisis es explorar la correlación entre los valores objetivos y las evaluaciones subjetivas realizadas con la escala GRBAS (Ver Tabla 2).

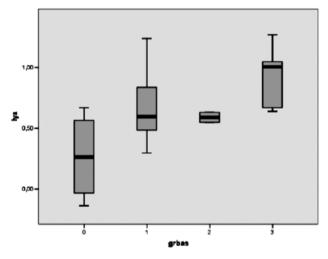


Figura 2. Distribución de LE según el grado general de severidad de la disfonía.

Se observa que LE diferencia tres grados: el 0, el 1-2 y el 3. Es decir no se alcanza a diferenciar el grado 1 del 2. Sin embargo a pesar de la poca cantidad de casos con parálisis es posible diferenciar dos grupos por lo que se espera que esta

tendencia se remarque al aplicarse a un número mayor de casos.

El análisis estadístico fue realizado con el programa SPSS 13.0. Se utilizó la prueba de Kruskal Wallis, dado que la distribución entre los grupos no es normal (Giovanni, Ouaknine, & Triglia, 1999). En la Tabla 2 se leen los valores de Chi cuadrado para cada uno de las mediciones. Se aprecia que el Jitter 2 presenta la mejor correlación seguido del LE, SNR2 y Jitter1. La medición SNR1 no es significativa para el análisis.

Se analizó mediante regresión lineal de LE y Jitter1. El valor de coeficiente de correlación de Pearson obtenido muestra que no es lo suficientemente significativo 0.13 por lo que interpretamos que la información de ambos métodos no resulta redundante, indicando que tal vez analicen fenómenos diferentes.

Tabla 2. Valores de correlación y significancia de los índices estudiados.

	Estadísticos de contraste				
	LE	Jitter 1	Jitter 2	SNR 1	SNR 2
Chi-cuadrado	10,497	7,464	16,093	1,301	8,919
Kruscal- Wallis p	0,015	0,058	0,001	0,729	0,030

Con el propósito de identificar las características (variables) que diferencian a dos o más grupos y determinar cuántas de ellas son necesarias para lograr la mejor clasificación posible, se realizó el análisis discriminante.

De esta forma podemos crear una función lineal capaz de distinguir a los miembros de cada grupo. En la Tabla 3 se brinda el detalle de los coeficientes estandarizados de la función canónica discriminante. En esa tabla se pueden comparar la importancia de cada variable a la hora de predecir el grupo de pertenencia de las clases perceptuales.

En la distribución de los grados perceptuales en función de dos componentes principales (Función 1 y 2) se observa la distribución de los datos correspondientes a cada clase perceptual, el centroide para cada una de ellas obtenido luego del análisis discriminante, todos estos datos representados en las coordenadas de las funciones discriminantes. Se puede observar cuatro grupos con centroides bastantes separados y escasos solapamientos entre clases.

Tabla 3. Coeficientes de tres funciones discriminantes para los índices estudiados

Coeficientes estandarizados de las funciones discriminantes canónicas				
	Función			
	1	2	3	
LE	0,165	0,169	0,652	
Jitter 1	-1,081	0,773	-0,119	
Jitter 2	1,106	0,310	-0,054	
SNR 1	-0,099	0,251	-0,275	
SNR 2	-0,042	-0,361	0,747	

Tabla 4. Matriz de confusión de los cuatro grados de evaluación perceptual donde se indican los porcentajes de pertenencia a cada grupo.

Grupo de pertenencia original	Grupo de pertenencia pronosticado			
	0	1	2	3
0	87,5	12,5	0	0
1	28,6	71,4	0	0
2	100	0	0	0
3	0	0	0	100

En la Tabla 4 se muestran los resultados del clasificador por análisis discriminante. Se observa que permite clasificar con claridad las clases 3,

0 y 1. La clase 2 no puede clasificarse correctamente por la cantidad insuficiente de datos.

DISCUSIÓN

Los métodos de análisis acústico basados en el modelo de producción de habla normal, es decir, el enfoque fuente-filtro se fundamanta en la idea de que los pliegues vocales son la fuente y el tracto vocal es un conjunto de filtros acústicos, pero esto es claramente algo idealizado ya que existe considerable evidencia de que el comportamiento de los pliegues vocales no se puede separar de la acción del tracto vocal. Es por esto que se postula que una representación verdaderamente realista debería basarse en un modelo no lineal.

Desde la teoría no lineal, las vibraciones de un sistema dinámico como los pliegues vocales pueden ser vistas como trayectoria en espacio de una fase en la evolución del tiempo. Si bien los exponentes de Lyapunov requieren algoritmos numéricos complejos, han demostrado ser uno de los descriptores más eficaces del caos.

Los exponentes de Lyapunov se relacionan con el índice medio de la divergencia exponencial o la convergencia de órbitas vecinas en espacio de la fase. Generalmente, un sistema que contiene por lo menos un exponente de Lyapunov positivo se define como caótico, mientras que un sistema sin exponente positivo es regular. Un sistema con un exponente de Lyapunov positivo demuestra inestabilidades locales y sensibilidad extrema a las condiciones iniciales. Es decir, cualquier cambio leve en el estado se amplificará exponencialmente, dando lugar a una salida totalmente distinta.

El objetivo de este trabajo fue presentar una nueva metodología que combina el análisis acústico tradicional y perceptual con algoritmos de la dinámica no lineal como los coeficientes de Lyapunov aplicados a la medición de voces con alta perturbación como las parálisis unilaterales. Se implementó un algoritmo para diferencia los sujetos con voz normal de pacientes con parálisis unilateral de cuerda vocal

En las parálisis existe una situación de asimetría tal entre las cuerdas que hacen claros los fenómenos no lineales. Existen muchas fuentes de no linealidad en la señal vocal, no linealidad en la relación fuerza y elongación de las cuerdas vocales, en el flujo aéreo y el tejido cordal, en la interacción de las cuerdas vocales introducidas por la vibración. En los casos de inestabilidades organizadas como en las parálisis solo los LE serían teóricamente adaptables y podrían proporcionar más información a los clínicos. Ambos tipos de medición (Jitter/LE) evalúan la inestabilidad de la señal pero no exploran la misma dimensión de la señal, por lo que podemos suponer que las mediciones son complementarias.

Esta alternativa basada en una definición aún más amplia y no tradicional de evaluación de la voz, incorpora información relacionada con el comportamiento global del sistema vocal. Esta nueva medición posee la ventaja de tener una representación gráfica (los retratos de fase) cuya interpretación es relativamente intuitiva. A mayor desorden vocal, mayor es el grado de complejidad de la imagen que se muestra, más desordenada es la misma.

CONCLUSIONES

Los resultados verifican el comportamiento caótico de las parálisis y la utilidad clínica de la medición, dado que los exponentes varían proporcionalmente con el grado de severidad. Estos estudios son prometedores, porque muestran la complementariedad de distintas medidas cuantitativas que explicaría la categorización perceptual. Se prevé extender estas mediciones a otros grupos poblaciones con diagnósticos diversos para estudiar su desempeño.

Agradecimientos

Consejo Nacional de Investigaciones Científicas y Técnicas.

REFERENCIAS

- Aasen, T. (1993). Chaos theory applied to the caloric response of the vestibular system. Computers and Biomedical Research (26): 556-567.
- Bergé, P., & Pomeau y Vidal, C. (1984). Order within chaos: towards a deterministic approach to turbulence. New York: U.S.A., Wiley.
- Berry, D. A., Herzel, H., Titze, I. R., & Story, B. H. (1996). Bifurcations in excised larynx experiments. Journal of Voice (10): 29-138.
- Bielamowicz, S., Kreiman, J., Gerratt, B. R., Dauer, M. S., & Berke, G. S. (1996). Comparison of voice analysis systems for perturbation measurements. J. Speech Hear. Res. (39): 126–134.
- Eckmann, J. P., & Ruelle, D. (1992). Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems. Physica D (56), 185–187.
- Giovanni, A., Ouaknine, M., & Triglia, J. M. (1999). Determination of largest Lyapunov exponents of vocal signal: application to

- unilateral laryngeal paralysis. J. Voice (13: 341-354.
- Gstoettner, W., Baumgartner, W., Hamzavi, J., Felix, D., Svozil, K., Meyer, R., y otros (1996). Auditory fractal random signals: Experimental data and clinical application. Acta Otolaryngol (116): 222-223.
- Herzel, H., Berry, D., Titze, I. R., & Saleh, M. (1994). Analysis of vocal disorders with methods from nonlinear dynamics. J. Speach Hear. Res. (37): 1008-1019.
- Hilborn, R. (2001). Chaos and Nonlinear Dynamics. An Introduction for Scientists and Engineers (2nd ed.). Oxford Univerity Press.
- Hirano, M. (1981). Clinical Examination of the voice. New York: U.S.A. Springer Verlag.
- Holzfuss, J., & Lauterborn, W. (1989). Lyapunov exponents from a time series of acoustic chaos. Phys. Rev. A (39): 2146–2152.
- Horii, Y. (1979). Fundamental Frequency Perturbation Observed in Sustained Phonation. Journal of Speech and Hearing Research (22): 5-19.
- Jiang, J., Zhang, Y., & Ford, C. N. (2003). Nonlinear dynamics of phonations in excised larynx experiments. J. Acoust. Soc. Am 114 (4 pt1): 2198-205
- Kumar, A., & Mullick, S. K. (1996). Nonlinear dynamical analysis of speech. J.Acoust. Soc. Am. (100): 615–629.
- Milenkovic, P. (1987). Least mean square measures of voice perturbation. J. Speech Hear. Res. (30): 529–538.

- Narayanan, S. S., & Alwan, A. A. (1995). A non-linear dynamical systems analysis of fricative consonants. J. Acoust. Soc. Am. (97): 2511.
- Peitgen, H. O., Jurgens, H., & Saupe, D. (1992). Chaos and Fractals, New Frontiers of Science. New York: U.S.A. Springer-Verlag.
- Titze, I. R. (1995). Workshop on acoustic voice analysis; Summary Statement. National Center for Voice and Speech, Denver.
- Titze, I. R., Horii, I., & Scherer, R. C. (1987). Some technical considerations in voice perturbation measurements. J Speech Hearing Research (30): 252-260.
- Wolf, A., Swift, J. B., Swinney, H. L., & Vastano, J. A. (1985). Determining Lyapunov exponents from a time series. Physica D (16): 285–317.
- Yu, P., Garrel, R., Nicollas, R., Ouaknine, M., & Giovanni, A. (2007). Objective Voice Analysis in Dysphonic Patients: New Data Including Nonlinear Measurements. Folia Phoniatr Logop, 1 (59): 20-30.
- Yu, P., Ouaknine, M., & Giovanni, A. (2000). Intérêt clinique du calcul des coefficients de Lyapunov pour l'analyse objective des dysphonies. Rev. laryngol. otol. Rhinol, 5 (121): 301-305.
- Zhang, Y., & Jiang, J. (2003). Nonlinear dynamic analysis of signal typing of pathological human voices. Electron. Lett. (39): 1021–1023.