

CARACTERIZACIÓN AUTOMÁTICA DE VOCES PATOLÓGICAS: MONOTONÍA DEL NIVEL

Eduardo Castillo Guerra, María E. Hernández-Días Huici

Centro de Estudios de Electrónica y Tecnologías de la Información (CEETI).
Universidad Central de Las Villas, Carretera a Camajuní km 5½ Santa Clara, Cuba.
Email: ecastillo@uclv.etcetsa.cu

RESUMEN

Este trabajo aborda el tema de la caracterización de voces patológicas utilizando algoritmos de procesamiento digital de señales y en específico procesamiento digital del Habla (PDH). Se enfoca el análisis en un rasgo particular, monotonía del período fundamental, de los rasgos definidos por Darley, Aronson y Brown en su metodología para el diagnóstico diferencial de Disartrias [3][4].

Se presenta un algoritmo para estimar la severidad con que se manifiesta la monotonía del período fundamental en voces patológicas y sanas. Se utiliza el algoritmo AMDF [7] para la estimación del período fundamental en un fragmento del párrafo "El abuelo" y se analiza la correlación de tres técnicas de PDH con la monotonía del período fundamental (Fo). Finalmente se propone una medida que se correlaciona mejor con la monotonía encontrada en las voces analizadas denominada *Alineación a la tendencia lineal* (ALT). Se realiza un análisis estadístico de los resultados con respecto a valoraciones subjetivas (VS) realizadas por tres especialistas en lenguaje patológico sobre una base de datos de voces con diferentes tipos de Disartrias. Se muestra la efectividad del algoritmo para cuantificar la ocurrencia de monotonía del nivel en voces patológicas y sanas.

Palabras clave: *Pitch monotony, Digital speech processing, Dysarthria.*

1. INTRODUCCIÓN

La utilización de los medios de cómputos en todas las esferas de la vida se ha hecho muy común en los tiempos actuales. La medicina moderna no ha escapado de esta tendencia y cada vez más se utilizan equipos y algoritmos que facilitan un diagnóstico y tratamiento más preciso de las enfermedades. El uso de técnicas de procesamiento digital de señales en la caracterización automática de voces patológicas permite poner a disposición de los especialistas medidas más objetivas de los diferentes rasgos que las caracterizan, reduciéndose así el factor subjetivo de los mecanismos de diagnóstico tradicionales. Las valoraciones subjetivas, como método tradicional de estudio de las voces patológicas, se apoya mucho en la experiencia adquirida por los especialistas a lo largo de su trabajo profesional e introducen diversidad de criterios en la clasificación y caracterización de algunas enfermedades, como por ejemplo, las Disartrias. Estas últimas, constituyen un grupo de enfermedades neurológicas que afectan específicamente el control de los músculos que intervienen en el proceso de producción de la voz y provocan trastornos del lenguaje que difieren ligeramente una de otras. La metodología descrita por Darley, Aronson y Brown para el diagnóstico diferencial de Disartrias [3][4] establece una relación entre 38 rasgos extraídos de la señal voz y la lesión

existente en el sistema nervioso. Los rasgos se agrupan en clusters y la combinación de clusters caracterizan el tipo de enfermedad. Este trabajo tiene como objetivo la caracterización automática de uno de esos 38 rasgos, *monotonía del nivel*, que se encuentra dentro del grupo relacionado con el período fundamental. La metodología agrupa los rasgos en siete categorías entre las que se encuentran: los rasgos relacionados con el período fundamental, los relativos al volumen o intensidad, los relativos a la respiración, los relativos a la calidad vocal, los relativos a la prosodia, los relativos a la articulación y los relativos a la impresión general [3][4]. Las medidas relacionadas con el período fundamental y monotonía del período en particular, son de especial importancia en la caracterización de voces patológicas disártricas.

El rasgo monotonía del nivel se define como el grado al cuál la voz no es capaz de seguir la inflexión y entonación apropiada que concuerde con el significado y el contexto de la alocución de acuerdo las características del lenguaje. En casos severos existe una ausencia de variaciones del período fundamental tendiendo a una voz monótona y carente de variaciones. De acuerdo a la definición dada anteriormente, el análisis de este rasgo debe realizarse en una muestra de voz de habla fluida donde se pueda apreciar la ausencia de variaciones del período fundamental donde el contexto y el idioma lo requiera.

2. METODOLOGÍA

En este estudio se utilizaron las voces patológicas de pacientes disártricos tomados de una base de datos obtenida por Aronson (1993) [1] con diferentes afecciones de la voz y pacientes sanos (grabados por los autores) que sirven como control. En todos los casos se digitalizaron las señales de voz con una frecuencia de muestreo de 22050Hz y una resolución de 16 bits.

El análisis de este rasgo se realizó teniendo en cuenta la lectura de un fragmento del párrafo del abuelo [8] en 82 pacientes, 6 pacientes sanos y 76 pacientes enfermos con diferentes patologías entre las que se encuentran Disartria Flácida, Espástica, Atáxica, Hipocinética (Parkinson), Hipercinética (Corea, Distonía), Temblor Orgánico de la voz, Myoclonus, Esclerosis Lateral Amiotrófica, Enfermedad de Wilson, Esclerosis Múltiple y Disfonía Flácida. En cada caso se solicitó la lectura del título del pasaje para comenzar el análisis después de este y eliminar la parte inestable del comienzo. En todos los casos les fue solicitado a los pacientes que leyeran a un ritmo natural y trataran de pronunciar todos los fonemas correctamente [1].

El idioma utilizado fue el inglés. Se utilizaron las valoraciones subjetivas realizadas por tres especialistas en lenguaje patológicos sobre las 82 voces analizadas para el rasgo monotonía del nivel. Se realizaron las valoraciones en

una escala de 7 niveles como las descritas en [3], distribuidas de la siguiente forma:

0. Variaciones normales del Fo.
1. Ligeramente notable falta de variaciones de Fo.
2. Notable falta de variaciones de Fo.
3. Moderada falta de variaciones de Fo.
4. Acentuada falta de variaciones de Fo.
5. Severa falta de variaciones de Fo.
6. No hay variaciones de Fo (Monotonía total).

En todos los casos se utilizó el algoritmo de detección grueso del período fundamental AMDF [7] para estimar los valores instantáneos del período durante el segmento analizado, dado su buen comportamiento ante voces patológicas [5]. Una vez obtenido el vector de los estimados del período fundamental se le aplicaron algoritmos como *perturbación relativa promedio modificada (MRAP)*, *Factor de Jitter modificado (MJF)* y *alineación a la tendencia lineal (ALT)* para obtener una medida la ausencia de variaciones relacionadas con la monotonía del nivel. Las medidas MRAP y MJF descritas a continuación difieren ligeramente de su definición original [2] ya que el período fundamental se calcula con respecto a un segmento de la señal de voz y no respecto a la diferencia entre pulsos glotales.

MRAP

Es una medida que elimina los efectos producidos por cambios bruscos en la señal mediante un promediado de las muestras adyacentes a la analizada [6]:

$$MRAP = \frac{\frac{1}{\sqrt{n-2j}} \sum_{i=j+1}^{n-j-1} \left| \frac{F_{i-j} + F_i + F_{j+1}}{2j+1} - F_i \right|}{\frac{1}{\sqrt{n}} \sum_{i=1}^n |F_i|} \quad (1)$$

en este caso $j=30$ para promediar las 60 muestras adyacentes a la analizada con el objetivo de eliminar las transiciones abruptas causadas por otras características patológicas de la voz y solo obtener las variaciones suaves que están relacionadas cambios del período de acuerdo a lo requerido por el lenguaje.

MJF

El factor de jitter modificado se define como la razón entre la perturbación media que esta por debajo de un umbral de 10Hz: y la duración media de la onda[2]:

$$MJF = \frac{\frac{1}{\sqrt{n-1}} \sum_{i=1}^{n-1} |F_i - F_{i-1}| < 0.1}{\frac{1}{\sqrt{n}} \sum_{i=1}^n |F_i|} \quad (2)$$

esta medida solo incluye aquellos valores que su variación esta por debajo de 10 Hz.

ALT

El factor alineación a la tendencia lineal se define como la razón entre la perturbación media y la duración media de la onda:

$$ALT = \sum_{i=1}^N \frac{MnP(i)}{N} \quad (4)$$

$$MnP = \begin{cases} 1 & \frac{1}{n-10} \sum_{i=5}^{n-5} \left| \frac{F_{i-1.5} + F_i + F_{i+1.5}}{10} - F_i \right| < 0.05 * \overline{F_0} \\ 0 & \frac{1}{n-10} \sum_{i=5}^{n-5} \left| \frac{F_{i-1.5} + F_i + F_{i+1.5}}{10} - F_i \right| > 0.05 * \overline{F_0} \end{cases} \quad (3)$$

Donde N es la longitud del vector MnP. El umbral de 0.05 en la ecuación 3 representa las variaciones por debajo del 5% del valor medio de las variaciones de Fo.

ANALISIS ESTADÍSTICO

Se realizó un análisis estadístico para observar el grado de correlación de los resultados de cada algoritmo con el valor medio de las valoraciones subjetivas realizadas por tres especialistas en lenguaje patológico para cada paciente analizado. Se obtuvo un índice de correlación que muestra la relación que existe entre los resultados de los algoritmos y las valoraciones subjetivas hechas por los especialistas para diferenciar cuál de ellos se asemeja más a las valoraciones subjetivas. Se realizó la prueba T para saber si el valor de correlación obtenida esta en el limite del 95% del intervalo de confianza y se obtuvieron los coeficientes de un modelo de regresión lineal para indicar el grado de confiabilidad de que ese índice sea verdadero de forma similar a lo descrito por Fukasawa[6].

3. RESULTADOS

La aplicación de algoritmos descritos anteriormente a las alocuciones de los pacientes analizados permitió obtener índices de correlación que se muestran en la tabla I.

Tabla I

Tabla de correlaciones de los algoritmos utilizados con respecto al valor medio de las valoraciones subjetivas

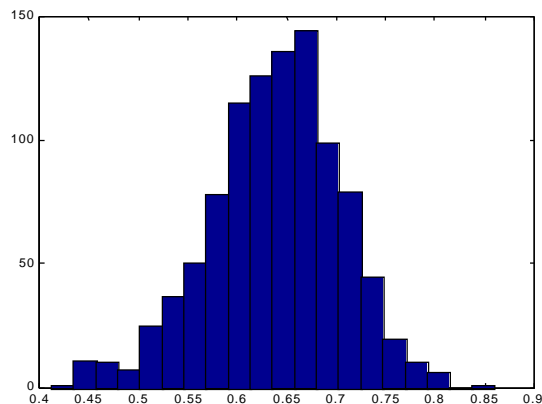
Algoritmo	Correlación con VS (R)	H	Intervalo de confianza del 95%
MRAP	0.5735	0	0.5707-0.5789
MJF	0.4246	0	0.4269 0.4393
ALT	0.6054	0	0.6012-0.6095

H representa que la hipótesis nula de la prueba T es aceptable a un nivel de significación 0.05. La hipótesis nula es asumir que las muestras del vector de entrada tienen una distribución normal y su valor medio es R.

Se puede apreciar que el algoritmo ALT es el que obtiene una mayor correlación con respecto a las valoraciones subjetivas hechas por los especialistas, siguiéndole MRAP y por último MJF. Es importante destacar que en el análisis estadístico, la prueba T, para el caso de ALT se acepta la hipótesis nula para un intervalo de confianza de 99%, mientras que para los otros algoritmos se acepta para un 95%. Las figuras 1, 2 y 3 muestran los histogramas que exhiben el comportamiento del índice de correlación entre VS y VO para 1000 muestras insertadas de forma aleatoria con los resultados de cada uno de los algoritmos utilizados.

En todos los casos la curva de contorno presenta una forma similar a la de una distribución normal corroborando la hipótesis asumida en el análisis estadístico (prueba t). El índice obtenido con el algoritmo ALT y MRAP es similar a los reportados y aceptados como indicativos de interdependencia por Fukasawa [6]. Para caso del algoritmo MJF se obtiene un índice de correlación relativamente bajo. Se utilizó el paquete estadístico del Matlab para obtener los resultados[9].

La figuras 4, 5 y 6 muestran los resultados de ALT, MJF y MRAP para todos los pacientes contemplados en el trabajo (EL- ELA, FD- Disartria Flácida, FF- Disfonía Flácida, HC- Hipercinética Corea, HD- Hipercinética Distonía, HP- Hipocinética Parkinson, SD- Disartria Espástica, WD- Enfermedad de Wilson, AD- Disartria Atáxica y PS- Pacientes sanos). Se muestra con una línea roja el umbral a partir del cuál se considera la existencia de cierto grado de severidad del rasgo caracterizado.



g. 1 Histograma del coeficiente de correlación VS-VO para ALT.

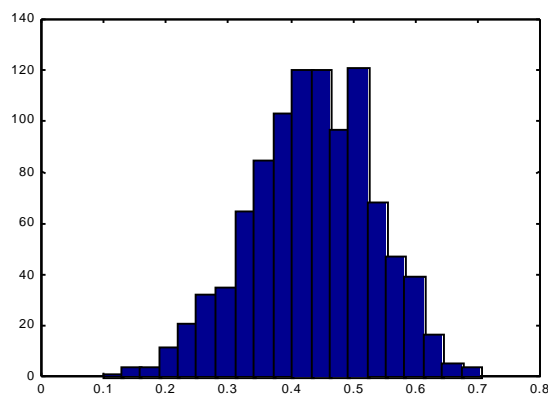


Fig. 2 Histograma del coeficiente de correlación VS-VO para MJF.

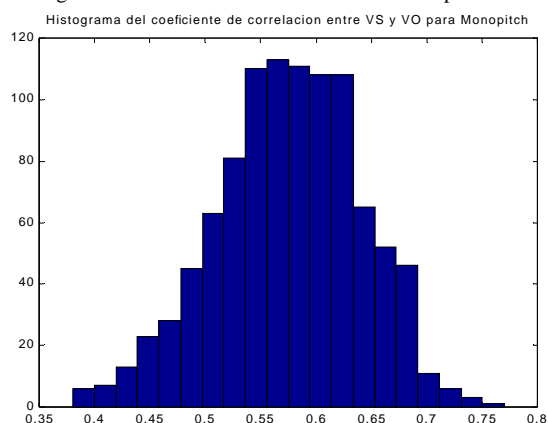
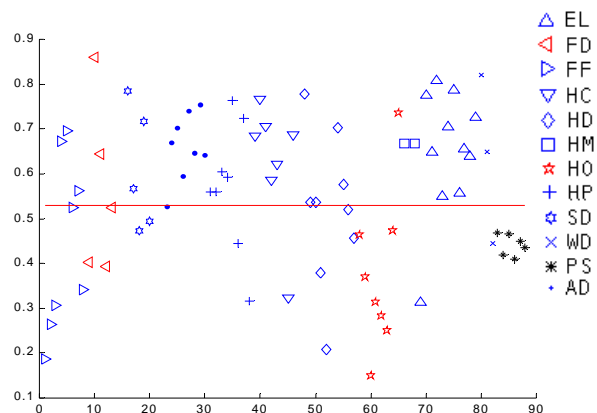


Fig. 2 Histograma del coeficiente de correlación VS-VO para MRAP.

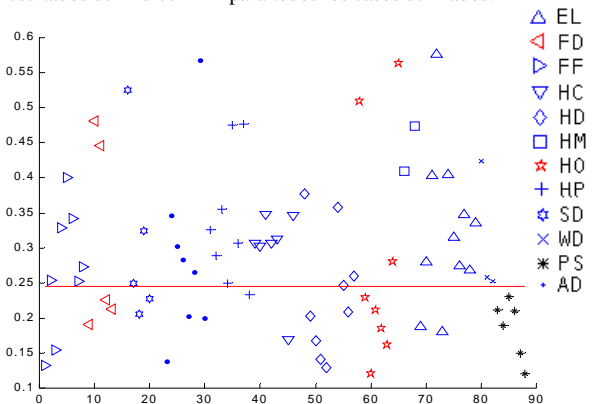
Se puede apreciar que los pacientes sanos introducidos como control se ubican próximos al umbral como era de esperar al existir voces con variaciones del período debidas a otras causas. Otras enfermedades como Disartria Flácida, Espástica, ELA y Corea, se clasifican adecuadamente por encima del umbral según lo descrito en [3][4]. Es importante resaltar que no todos los pacientes de la base de datos

utilizada exhiben todas las características de su grupo ya que el grado de severidad de las enfermedades es medio. Se puede apreciar que los pacientes sanos, introducidos como control, están ubicados próximos al umbral como era de esperar, al existir voces con variaciones del período debidas a otras causas. Otras enfermedades como Disartria Flácida, Espástica, ELA y Corea, se clasifican adecuadamente por encima del umbral según lo descrito en [3][4]. Es importante resaltar que no todos los pacientes de la base de datos utilizada exhiben todas las características de su grupo ya que el grado de severidad de las enfermedades es medio.



Fi

g. 4 Resultados del índice ALT para todos los casos utilizados.



Fi

Las figuras 7, 8 y 9 muestran la relación entre valoraciones subjetivas y las objetivas (VO) de los algoritmos ALT, MJF y MRAP. Se muestra en todos los casos la recta de mejor ajuste que representa la tendencia lineal de la relación VS-VO. El paradigma de esta relación es una recta de pendiente positiva con 45 grados que representa una relación lineal directa entre VS y VO. Se aprecia como ALT presenta la característica más lineal con una pendiente más cercana a la ideal. Se mostraron, con círculos, los valores medios por cada clase para analizar la tendencia lineal de la relación descrita. En todas las figuras se aprecia que para los niveles de severidad bajos (0 y 1) ocurren los errores más significativo, aunque en la figura 8 se aprecia cierta dispersión en los valores medios.

4. DISCUSIÓN

El algoritmo ALT fue el que obtuvo un comportamiento superior respecto a los algoritmos MJF y MRAP en la cuantificación del rasgo monotonía del período descrito en la metodología de Darley, Aronson y Brown.

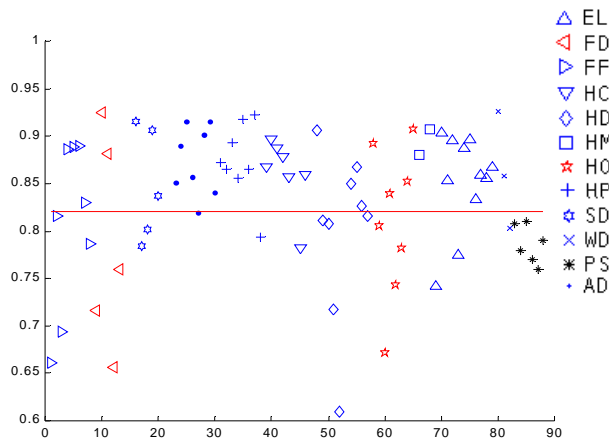


Fig. 6 Resultados del índice MRAP para todos los casos utilizados.

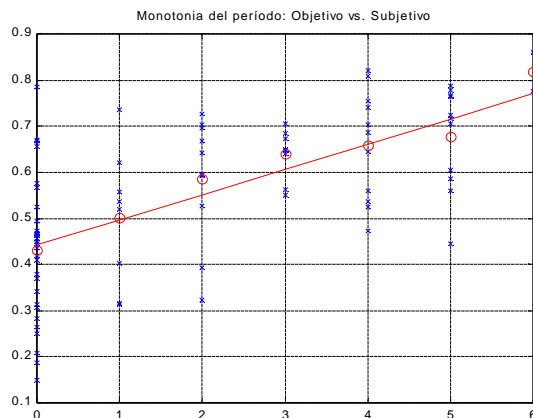


Fig. 7 Índice ALT vs. Valoración subjetiva de rupturas del período. Se muestra la recta de mejor ajuste. Los círculos muestran los valores medios de RTP por cada valor valoración subjetiva.

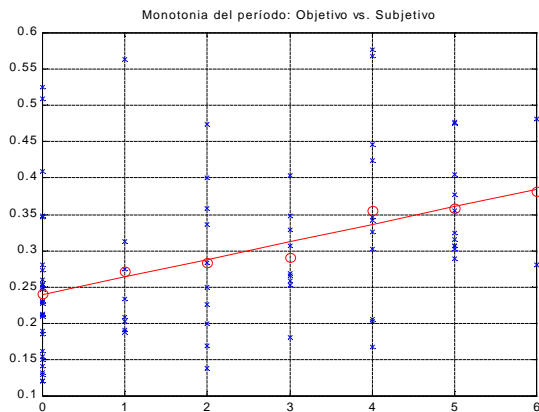


Fig. 8 Índice MJF vs. Valoración subjetiva de rupturas del período. Se muestra la recta de mejor ajuste. Los círculos muestran los valores medios de RTP por cada valor valoración subjetiva.

El algoritmo MRAP supera a MJF en la realización de esta tarea. La mayor parte de las diferencias entre VS y VO, en general para todos los algoritmos se encuentran en los valores más pequeños de la escala utilizada, mientras que en los extremos se obtienen mejores resultados. Para los casos ALT y MRAP se obtuvieron valores de correlación similares a los reportados y aceptados como indicativos de interdependencia por Fukasawa[6]. El MJF no es apropiado para caracterizar la monotonía del nivel dado que el valor de correlación con las VS es muy bajo. Se puede apreciar que las valoraciones subjetivas son, en ocasiones, inconsistentes con el criterio de validación seguido. Este rasgo requiere de gran

entrenamiento por parte de los especialistas para ser valorado correctamente ya que se aprecia una correlación de 0.67 entre los especialistas.

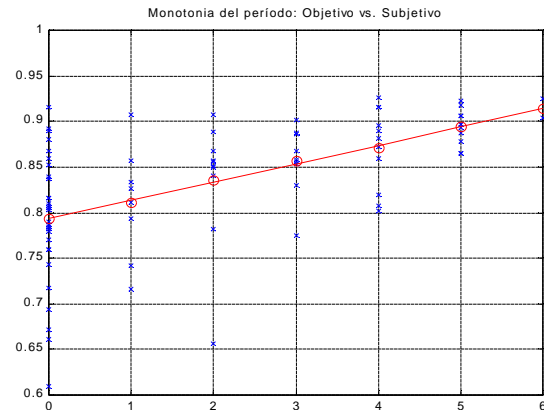


Fig. 9 Índice MRAP vs. Valoración subjetiva de rupturas del período. Se muestra la recta de mejor ajuste. Los círculos muestran los valores medios de RTP por cada valor valoración subjetiva.

5. CONCLUSIONES

Se demuestra la efectividad de ALT para diferenciar voces sanas y con monotonía del período, así como para cuantificar la severidad con que aparece dicha perturbación. El ALT supera los demás algoritmos de sus resultados con las VS. Se obtienen resultados que pueden contribuir a la introducción de un sistema de caracterización automática de voces patológicas donde se recomienda el uso de ALT como alternativa para medir la monotonía del período.

REFERENCIAS

- [1] Aronson A. E., "Dysarthria Differential Diagnosis" . Mentor Seminars. Rochester. Michigan. (4 Cassettes). 1993.
- [2] Baken R. J., "Clinical measures of speech and voice". Singular Publishing Group, Inc, 1996.
- [3] Darley F.L, Aronson A.E., Brown J.R.," Differential diagnostic Patterns of Disarthria.JSHR, vol:12 pp 246-269, 1969a.
- [4] Darley F.L, Aronson A.E., Brown J.R.," Differential diagnostic Patterns of Disarthria.JSHR, vol:12, pp. 462-496, 1969b.
- [5] Ferrer C., "Cuantificación de parametros subjetivos de la voz para el diagnostico de Disartrias" . Tesis de Master, UCLV, 2000.
- [6] Fukasawa T., El-Assuooty A.& Honjo I., " A new index for evaluation of the turbulent noise in pathological voice". J.A.S.A. vol. 83, No 3. pp 1189-1193. March 1988.
- [7] Hernández-Díaz M.E., "Algoritmos para la extracción del período fundamental de la voz. Desarrollo y evaluación". Tesis doctoral. UCLV.Cuba 1996.
- [8] Cannito Michael P., Yorkston K.M., Beukelman D.R., "Neuromotor speech disorders. Nature, Assessment, and Management". © Paul H. Brookes Publishing Co, 1998.
- [9] The MathWorks, Inc.; " Statistics Toolbox User's Guide"; Third printing 1995 - 1999.

AUTOMATIC CHARACTERIZATION OF PATHOLOGICAL VOICES MONOTONY OF LEVEL

ABSTRACT

This work deals with the topic of automatic characterization of pathological voices using digital signal processing techniques (PDH), specially using speech-processing algorithms. The analysis is focused in a specific voice feature defined by Darley, Aronson & Brown in their methodology for differential diagnosis of Dysarthrias **¡Error!No se encuentra el origen de la referencia.¡Error!No se encuentra el origen de la referencia..** An algorithm for estimation of pitch monotony severity in pathological and normal voices is presented. AMDF **¡Error!No se encuentra el origen de la referencia.** algorithm is used to estimate the fundamental frequency in a fragment of the standard paragraph “The GrandFather” and the correlation between some PDH techniques and anomalous Pitch monotony is analyzed. Finally, a measure –aligning with linear trend (ALT)- that best correlates with monotony is proposed. A statistical analysis of the results is performed with respect to the subjective appreciation done by 3 speech language pathologists over a database of different kind of dysarthria. The efficacy of the algorithm proposed to measure the occurrence of monotony in dysarthric and normal voices.