CARLOS ARIEL FERRER RIESGO.
MARÍA ESPERANZA HERNADEZ-DIAZ HUICI.

Centro de Estudios de Electrónica y Tecnologías de la Información.
Universidad Central de Las Villas, Santa Clara, Cuba.

Una medida de imprecisión articulatoria en grabaciones de pacientes disártricos.

Introducción

Dada la importancia que el lenguaje hablado tiene para el hombre, en el mundo se le brinda una especial atención a las técnicas que permiten un mejor diagnóstico, seguimiento y/o rehabilitación de las personas con afectaciones en su comunicación oral. Una de las formas más comunes de alteraciones patológicas del habla la constituyen aquellos trastornos neurológicos que afectan en alguna medida los sistemas involucrados en la producción de la voz. Dentro de estos se encuentran los trastornos motores del lenguaje (TML) donde las disartrias constituyen un grupo principal.

A fines de la década de los 60's y principios de los años 70 Darley, Aronson y Brown (DAB) llevaron a cabo una serie de estudios ([1], [2], [3], [4], [5], [6], [7], [8]) en las Clínicas Mayo donde analizaron las características del habla de grupos de pacientes disártricos desde el punto de vista auditivo. A partir de estos estudios elaboraron una metodología de diagnóstico a partir de las valoraciones subjetivas de determinados rasgos perceptuales los cuales fueron evaluados en escala de 1 a 7 en nivel ascendente de severidad a partir de la realización de tres ejercicios: fonación de una vocal "a" sostenida durante un tiempo no menor de 8 segundos, repetición lo más rápido y uniformemente posible de los monosílabos con consonantes explosivas "Pa", "Ta" y "Ka" y la lectura de un segmento del párrafo del abuelo, utilizado con frecuencia en la literatura para el análisis de voces patológicas.

El uso de técnicas de procesamiento digital de señales en la caracterización automática de voz permite obtener medidas objetivas de algunos de los rasgos descritos en dicha metodología. El presente trabajo propone un método totalmente automático para la obtención de una medida para la valoración objetiva del rasgo denominado Consonantes Imprecisas (CI), contenido dentro del cluster de Imprecisión Articulatoria de la metodología DAB, a partir de grabaciones de un grupo de pacientes disártricos.

Desarrollo:

En el cluster de Imprecisión Articulatoria se agrupan defectos del habla como la articulación incorrecta, sustitución o repetición de uno o varios fonemas. Se percibe básicamente como errores involuntarios en la pronunciación. Las consonantes imprecisas son naturalmente una de la manifestaciones de este cluster. Para el desarrollo de un método automático de medición de severidad de este rasgo pueden usarse los ejercicios del párrafo y del PTK. En este trabajo se prefirió el segundo, por razones de factibilidad, pues ya existe un método automático de detección de sílabas en el PTK [11] y se había empleado con efectividad en la cuantificación automática de rasgos [12].

En el ejercicio del Pa-Ta-Ka las tres consonantes son explosivas sordas, por lo que sus características son similares y no es necesario determinar a qué consonante específica pertenece la sílaba si se refiere a falta de sonoridad previa a la vocal o a ausencia de sonido previa a la explosión de la consonante. Estas son las dos principales características de las consonantes de este ejercicio, y cuya alteración constituye la principal causa de perceptibilidad de Consonantes Imprecisas. El método descrito en este trabajo aborda la medición de CI a partir de sólo una de estas características: que busca la presencia de sonorización en estas consonantes, lo que auditivamente se refleja como la pronunciación de Ba-Da-Ga en lugar del Pa-Ta-Ka.

Se contó con 57 grabaciones del ejercicio del Pa-Ta-Ka, digitalitizadas a 22050 Hz, con 16 bits de resolución. Para cada una de ellas dos jueces emitieron valoraciones subjetivas en una escala ascendente del nivel de severidad de 0 a 6, dos valoraciones cada uno por grabación.

El nivel de CI del paciente se calcula como el promedio de los obtenidos para cada sílaba en particular, y este a su vez se halla como el valor del primer pico de la función de autocorrelación de los 40 milisegundos de señal previos al inicio de la sílaba.

Resultados:

El parámetro descrito debe ser validado de forma tal que se demuestre la correspondencia entre la magnitud objetiva, obtenida mediante su uso, y la evaluación perceptual de la característica subjetiva que se pretende medir con él (Consonantes Imprecisas). En el trabajo de Fukazawa et al [9] se realiza una tarea similar, con un índice para medir jadeo. Fukazawa encontró una correlación entre dos especialistas de 0.714 para jadeo y de 0.589 para aspereza de la voz. Su índice tuvo una correlación de 0.734 con el promedio de las valoraciones de los especialistas para el jadeo y de 0.211 con las valoraciones de aspereza. De ahí concluyen que el índice puede ser empleado en la misma medida en que se usan las valoraciones subjetivas de jadeo. Este método también se utilizó en ([12] y [13]) en los cuales se obtuvieron índices de correlación entre las valoraciones subjetivas y las objetivas de 0.6724 para Razon Lenta, 0.8415 para Ruptura del Periodo y 0.6054 para Monotonía del Periodo.

En este trabajo se sigue en lo fundamental ese método de validación, aunque se brindan también otras medidas que dan una idea de la correspondencia entre las valoraciones subjetivas (Vsubj) y el índice CI. Además del coeficiente de correlación se muestra un gráfico del índice CI vs. Vsubj (Figura 1), junto a la recta de mejor ajuste. También se muestra un intervalo de confianza (***ci*** en la Tabla 1) para el valor del coeficiente de correlación obtenido y los coeficientes b_1 y b_2 (***b*** en la Tabla 1) de la recta de mejor ajuste $CI = b_2 * Vsubj + b_1$ junto a sus respectivos intervalos de confianza (***bint*** en la Tabla 1). En la Tabla 1 se muestran además (en ***stats***) tres valores correspondientes a la regresión lineal [10] de CI-Vsubj, el primero de ellos el porcentaje de varianza que el modelo de regresión lineal justifica, el segundo es el valor de F en la distribución de probabilidad del mismo nombre y el último indica la probabilidad de haber obtenido una muestra como la analizada para el caso de que las dos variables CI y Vsubj no estuvieran correlacionadas. Como indicador adicional se brinda el histograma del coeficiente de correlación entre CI y Vsubj. Los resultados del índice CI se obtuvieron sobre la muestra de 57 pacientes, y se muestran en la Figura 1 , la Tabla 1 y la Figura 2.

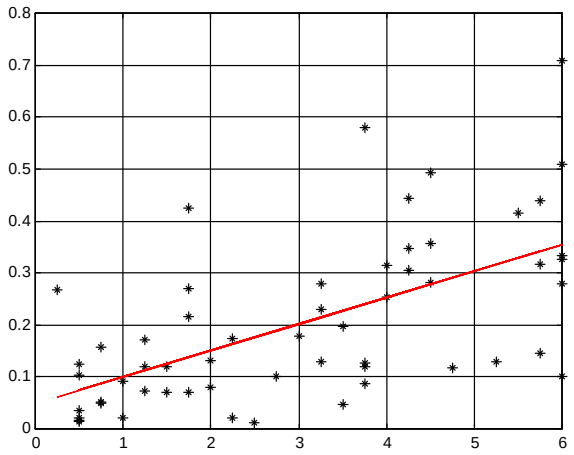

Figura 1 : Índice CI vs. percepción subjetiva de Consonantes Imprecisas del Nivel. Recta de mejor ajuste.

Corr.	T-test		Regresión Lineal					
R	ci		b	Bint		Stats		
0.6046	0.601	0.6117	0.051	0.0329	0.0692	0.3656	31.69	6.5e-012
			0.0487	-0.016	0.1134			

Tabla 1 : Resultados del índice CI.

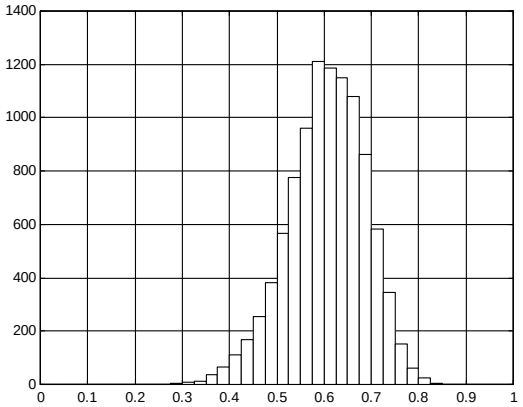

Figura 2 : Histograma del coeficiente de correlación entre el índice CI y la percepción subjetiva de Consonantes Imprecisas.

Con relación a las valoraciones subjetivas, fueron efectuadas por dos jueces, emitiendo dos valoraciones por paciente cada uno, escuchando las grabaciones en orden aleatorio. La correlación entre las dos rondas de valoraciones para cada juez fue de 0.88 y 0.79, mientras que la correlación entre los promedios de ambos fue de 0.75.

Conclusiones:

El coeficiente de correlación obtenido entre el índice propuesto y las valoraciones subjetivas (0.60) está en el orden de los reportados en la literatura entre mediciones objetivas y subjetivas (0.60 [13], 0.67 [12], 0.73 [9] y 0.84 [13]), y de correlación entre especialistas para valoraciones subjetivas de características de la voz (0.59 y 0.71 [9]) aunque es ligeramente inferior a la reportada en este mismo trabajo (0.75) entre especialistas. Este resultado es muy alentador teniendo en cuenta que el método sólo aborda la distorsión de las consonantes por sonorización, mientras las valoraciones subjetivas se basan en distorsión por cualquier causa. También atenta contra el desempeño del método la tasa de efectividad que pueda tener el detector automático de sílabas. A pesar de esto, en el intervalo de confianza de b_2 (pendiente de la recta) el valor superior apenas duplica el inferior, por lo que se puede apreciar que la relación DN - Vsubj es bastante definida. Por otra parte, el tercer valor de stats indica que la probabilidad de que DN y Vsubj no estén correlacionadas es menor que 10^{-11}. El histograma también muestra que la correlación entre ambas variables es notable, con casi la totalidad de los valores por encima de 0.5.

De acuerdo a estos resultados podemos afirmar que el ejercicio de la repetición de sílabas puede ser empleado para la determinación objetiva de la Imprecisión en las Consonantes, y el índice propuesto puede ser empleado para la valoración automática del rasgo del mismo nombre con relativamente buena efectividad. Estos resultados deben ser superiores introduciendo mediciones de sustituciones por otras consonantes como nasales y fricativas.

Referencias:

[1]. Aronson, A.E. “Dysarthria: Differential Diagnosis”. Mentor Seminars. Rochester. Michigan. (4 cassettes). 1993.

[2]. Berry, W.R.; Aronson, A.E.; Darley, F.L. & Goldstein, N.P. “Effects of penicillamine therapy and low-copper diet on dysarthria in Wilson’s disease (hepatolenticular degenerative)”. Mayo Clinic Proceedings. 49, pp 405-408. 1974.

[3]. Berry, W.R.; Darley, F.L.; Aronson, A.E. & Goldstein, N.P. “Dysarthria in Wilson’s disease”. Journal of Speech & Hearing Research. 17, pp 169-183. 1974.

[4]. Darley, F.L.; Aronson, A.E. & Brown, J.R. “Clusters of deviant speech dimensions in the dysarthria”. Journal of Speech & Hearing Research. 12, pp 462-496, 1969.

[5]. Darley, F.L.; Aronson, A.E. & Brown, J.R. “Differential diagnostic patterns of dysarthria”. Journal of Speech & Hearing Research. 12, pp 246-269, 1969.

[6]. Darley, F.L.; Aronson, A.E. & Brown, J.R. “Motor Speech Disorders”. Philadelphia. Saunders. 1975.

[7]. Darley, F.L.; Aronson, A.E. & Brown, J.R. “Motor Speech Signs in Neurologic Disease”. Medical Clinics of North America, 52, 835-844. 1968.

[8]. Darley, F.L.; Brown, J.R. & Goldstein, N. “Dysarthria in multiple sclerosis”. Journal of Speech & Hearing Research. 15, pp 229-245, 1972.

[9]. Fukazawa, T.; El-Assuooty, A. & Honjo, I. “A new index for evaluation of the turbulent noise in pathological voice”. Journal of the Acoustical Society of America. Vol. 83, No 3. pp 1189-1193. March. 1988.

[10]. Statistic Toobox User’s Guide. MathWorks. January. 1998.

[11]. Ferrer, C. A., Hernandez, M. E. “Detector de Posición de Sílabas Aisladas en Grabaciones de la Señal de Voz de Pacientes con Trastornos Motores del Lenguaje mediante Técnicas de Procesamiento Digital de Señales”. Memorias de la Conferencia Internacional TELEC’02, Santiago de Cuba. En CD ISBN 84-8138-506-9. Julio 2002.

[12]. Ferrer, C. A., Hernandez, M. E. “Obtención de un indice objetivo de razon lenta”. Actas del VII Simposio Internacional de Comunicación Social Santiago de Cuba. Editorial Academia. ISBN 959-02-0314-0. Vol. II. pp 390-394. Enero 2001.

[13]. Castillo, E., Ferrer, C. A. “Estudio de parámetros relativos al periodo Fundamental en Voces patologicas”. Actas del VII Simposio Internacional de Comunicación Social Santiago de Cuba. Editorial Academia. ISBN 959-02-0314-0. Vol. II. pp 395-400. Enero 2001.