



II CONGRESO CUBANO DE LOGOPEDIA Y FONIATRÍA Y II ENCUENTRO INTERNACIONAL DE TRASTORNOS DEL LENGUAJE, HABLA Y VOZ.

**TÍTULO:** ALGORITMOS Y PROGRAMAS PARA EL ESTUDIO DE VOCES PATOLÓGICAS DISARTRICAS.

**AUTORES:** Eduardo Castillo Guerra

**CENTRO DE PROCEDENCIA:** Centro de Estudios de Electrónica y Tecnologías de la Información (CEETI), UCLV. Carr. A Camajuaní km 5 ½, Santa Clara, Villa Clara, Cuba.,54830. Email: eduardo.castillo@unb.ca

## RESUMEN

El trabajo aborda el tema de la clasificación de enfermedades neuronales conocidas como disartrias. Se estudian diferentes herramientas para el estudio de voces patológicas, haciendo énfasis en un novedoso enfoque basado en rasgos extraídos de voces patológicas mediante valoraciones objetivas y subjetivas. La metodología para la clasificación se basa en el uso de algoritmos de procesamiento digital de señales para apreciar la severidad de las perturbaciones menos confiablemente percibidas por los especialistas, mientras que las otras son tomadas directamente de las valoraciones subjetivas. Se desarrolla un sistema experto para la clasificación de Disartrias que no solo proporciona un elevado porcentaje de clasificación correcta, sino que suministra información precisa acerca de la contribución de los rasgos extraídos a la decisión del clasificador. El clasificador utilizado es no lineal, basado en mapas auto-organizados de Kohonen, que proporciona una mayor efectividad en la clasificación que otros clasificadores reportados. Los análisis derivados de este trabajo permiten correlacionar los rasgos de la voz con la localización de la lesión dentro del sistema nervioso central o periférico. El sistema experto que incluye la metodología y el clasificador no lineal proporciona a los especialistas una herramienta potente para el estudio y el diagnóstico de las Disartrias.

Palabras claves: Procesamiento de voz, Disartrias, voces patológicas.

## I. INTRODUCCIÓN

El desarrollo de las nuevas tecnologías de la información ha impactado directamente en la mayoría de los aspectos de la vida moderna. En la esfera de la salud, el impacto ha sido de gran relevancia dando lugar al desarrollo de innumerables herramientas y equipos para el diagnóstico, tratamiento y rehabilitación. Especialmente el estudio de la voz ha sido objeto de innumerables investigaciones, destacándose el nacimiento de varias herramientas para el estudio de diferentes aspectos del proceso de producción de la voz que no son fácilmente percibidos de forma perceptual o con la ayuda de dispositivos mecánicos sencillos. Una de las herramientas más utilizadas y

comercializadas para el análisis de voces patológicas es el CSpeech, desarrollado por Kay Elemetrics Inc. [25] Dicha herramienta es de propósito general y permite obtener varios parámetros de la voz que reflejan muchas de las perturbaciones escuchadas. Sin embargo, las medidas obtenidas mediante esta herramienta no están relacionadas directamente con patologías específicas, sino que corresponde a los especialistas en lenguaje patológico correlacionar las perturbaciones obtenidas con las patologías estudiadas. Varias investigaciones reportadas han demostrado correlaciones entre dichas mediciones y varias patologías [26].

Otra herramienta de propósito mucho más general que el CSpeech es el Praat [27] ([www.praat.org](http://www.praat.org)). Esta herramienta es muy popular debido a que no es comercializada, sino que se brinda gratis por los desarrolladores para el estudio de voces. El software Praat está diseñado para manipular objetos de software que se relacionan con la fonética y el análisis de voz mediante algoritmos computacionales. El mismo pone a disposición de los especialistas la mayoría de los algoritmos de procesamiento digital de señales disponibles en la literatura para el estudio de la voz, incluyendo análisis estadísticos y herramientas de clasificación. Por su complejidad, el Praat es una herramienta potente, pero requiere de un conocimiento previo sostenido de procesamiento de señales y fonética por parte de los especialistas en lenguaje patológico para su explotación adecuada. Sin embargo, esta herramienta tampoco vincula las mediciones obtenidas con las características de enfermedades en particular.

El presente trabajo aborda la problemática del desarrollo de herramientas de estudio de perturbaciones de la voz relacionadas directamente con un grupo de enfermedades neurológicas conocidas como disártrias. Los objetivos del trabajo consisten en desarrollar algoritmos computacionales que permitan apreciar automáticamente ciertas patologías de la voz y asociarlas con cada grupo de disartrias estudiadas. El trabajo fue concebido para mejorar la efectividad del diagnóstico de estas enfermedades y desarrollar un sistema de ayuda al diagnóstico que permita obtener medidas objetivas del estado de salud de los pacientes para su tratamiento y rehabilitación. Se desarrolla una nueva metodología por la clasificación de disartrias que combina valoraciones objetivas y subjetivas para obtener diagnósticos acertados mediante clasificadores implementados con redes neuronales artificiales.

Las disartrias constituyen una causa de padecimiento de un porcentaje representativo de la población mundial, especialmente en países subdesarrollados donde la población es más numerosa y existe un número limitado de especialistas entrenados para diagnosticar eficientemente este tipo de enfermedad. Las mismas pueden manifestar síntomas muy similares entre sí, siendo originadas por lesiones en diferentes lugares del sistema nervioso central o periférico (SNCP). Consecuentemente, el desarrollo de nuevas herramientas alternativas para el estudio de estas enfermedades es necesario y bien aceptado por especialistas y pacientes.

## **II. MATERIALES Y MÉTODOS**

En esta sección se describen brevemente los principales métodos desarrollados en el marco de este trabajo para el estudio de los diferentes tipos de disártrias.

### ***A. Descripción de las disártrias***

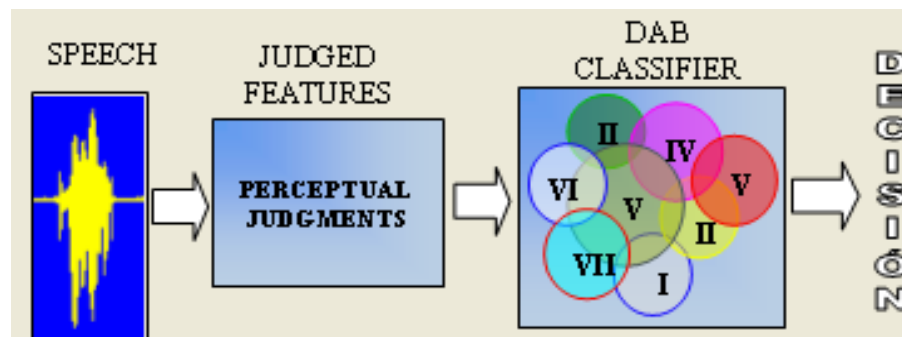
Disartria es un término relacionado con un grupo de enfermedades neuronales que se diferencian entre sí por la localización de la lesión en el SNCP. Estudios clínicos realizados sobre disártrias han revelado que existe una relación causa-efecto entre la localización de la lesión en el SNCP y el tipo de disartria exhibida. La localización del daño en el SNCP determina las manifestaciones y los síntomas producidos por la enfermedad y por consiguiente, las

perturbaciones de la voz que se observan. Consecuentemente, se puede establecer una correspondencia entre las perturbaciones de la voz y la localización de la lesión en el SNCP. Sin embargo, las diferencias entre estas enfermedades pueden ser muy sutiles, produciendo frecuentemente errores en la clasificación, fundamentalmente en etapas tempranas. La voz es uno de los mecanismos más susceptibles a lesiones en el SNCP y por consiguiente es un medio muy deseado para el diagnóstico y estudio de este tipo de enfermedad .

Entre los tipos de disártrias más comunes se encuentran: disartria flácida, espástica, atáxica, hipocinética (enfermedad de Parkinson), hipercinética (corea, dystonia, temblor orgánico de la voz) y disartria mixta (esclerosis lateral amyotrófica (ALS), esclerosis múltiple y enfermedad de Wilson).

### ***B. Método tradicional de diagnóstico diferencial de disartrias***

El clasificador tradicional para el diagnóstico diferencial de las disártrias esta basado en 38 rasgos apreciados subjetivamente por los especialistas en lenguaje patológico. Estos rasgos son agrupados en clusters y la combinación de estos cluster determina el tipo de enfermedad. La secuencia de la realización de este método se muestra en la figura 1.

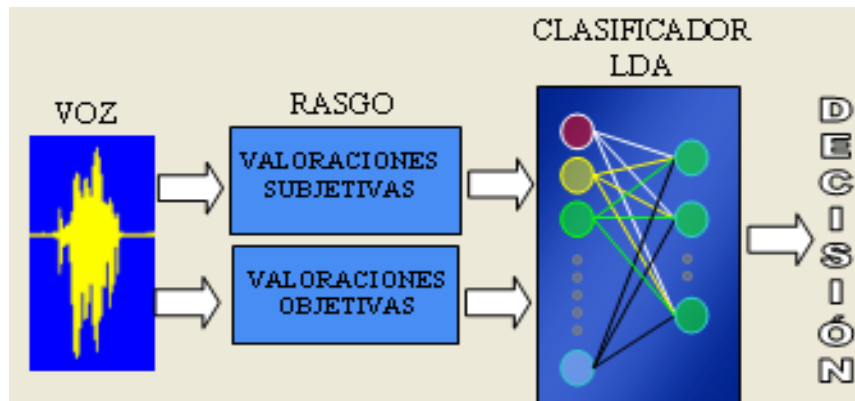


**Figura 1** Diagrama en bloques del clasificador tradicional.

Los métodos tradicionales para la clasificación de disártrias, como es descrito anteriormente, están basados en las perturbaciones de la voz y se apoyan fundamentalmente en las valoraciones subjetivas realizadas por especialistas en lenguaje patológicos como descriptores fundamentales. Las valoraciones subjetivas suelen ser imprecisas e inconsistentes en algunos de los rasgos de la voz, especialmente cuando son realizadas por diferentes especialistas que frecuentemente tienen puntos de referencia diferentes debido a que vienen de diferentes escuelas. Esto hace que la confiabilidad y repetibilidad del proceso de descripción de las perturbaciones sea baja, lográndose en ocasiones niveles bajos de clasificación y diagnósticos muy difíciles de estandarizar.

### ***C. Método del discriminante lineal***

El clasificador mediante el algoritmo del discriminante lineal (LDA) se basa en un análisis de los grupos o clases a través de combinaciones lineales de los rasgos de entrada. La decisión se realiza sobre la ecuación que produce una máxima magnitud, lo que equivale a una mínima distancia del vector de entrada con la clase seleccionada. La implementación puede ser apreciada en la figura 2.

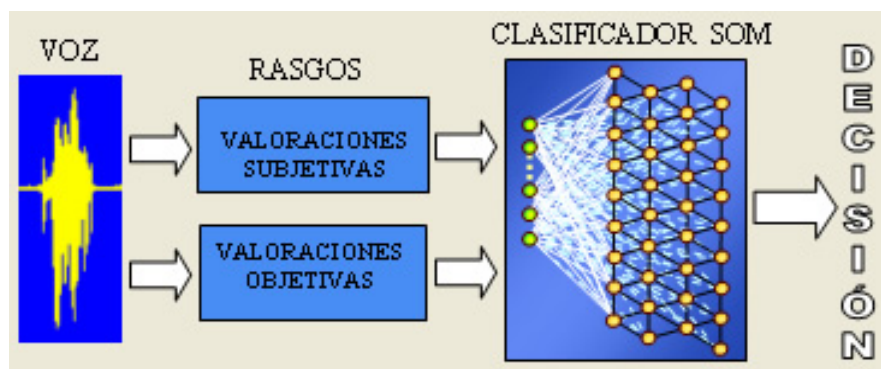


**Figura 1** Diagrama en bloques del clasificador mediante discriminante lineal.

Este método separa los diferentes grupos de enfermedades mediante hiperplanos lineales y es incapaz de diferenciar eficientemente las clases cuando los descriptores de los rasgos no obedecen a comportamiento descritos por ecuaciones lineales. Esto trae como consecuencia que el análisis de relevancia de los rasgos que más intervienen en el proceso de clasificación puede ser poco preciso.

#### ***D. Método de clasificación mediante Mapas autoorganizados de Kohonen***

El método de clasificación mediante mapas auto-organizados de Kohonen (SOM) está basado en el uso de redes neuronales artificiales que realiza un análisis no lineal del conjunto de datos de entrada. Este tipo de red neuronal aprende a reconocer las regularidades y correlaciones del conjunto de entrada y adapta las respuestas futuras de acuerdo a las entradas. Este tipo de red no solo aprende a reconocer grupos de vectores de entrada similares sino que también las neuronas vecinas aprenden a reconocer la vecindad de las secciones delimitadas por el espacio de entrada. Por consiguiente, las SOMs aprenden la distribución y la topología de los vectores de entrada que son usados para su entrenamiento y producen un vector de pesos asociados a cada neurona que sirve para identificar cuáles de los rasgos de entrada contribuyen más a la decisión final del clasificador. Esto determina su uso en este tipo de aplicación donde el análisis de la contribución de los rasgos es decisivo para el estudio de este grupo de enfermedades [2]. La figura 3 muestra el diagrama en bloques del clasificador diseñado para la aplicación de este tipo de red neuronal a la clasificación de disartrias.



**Figura 2** Diagrama en bloques del clasificador mediante mapas autoorganizados.

## ***E. Metodología***

En este estudio se utilizaron grabaciones de voces patológicas de pacientes que exhiben 8 tipos de disártrias entre los que se encuentran disartria flácida, espástica, atáxica, enfermedad de Parkinson, corea, dystonia, temblor orgánico de la voz y Esclerosis Lateral Amyotrófica [22].

Las mediciones de los rasgos fueron realizadas sobre la alocución de tres ejercicios que muestran más eficientemente las perturbaciones en la voz. Las grabaciones utilizadas fueron tomadas de dos bases de datos, una grabada por Aronson, 1993 [23] y otra de pacientes cubanos grabada por los autores de este trabajo [16]. Se analizaron un total de 127 pacientes teniendo en cuentas un conjunto de 38 rasgos de la voz reportados entre los más importantes para describir los tipos de disartria utilizados.

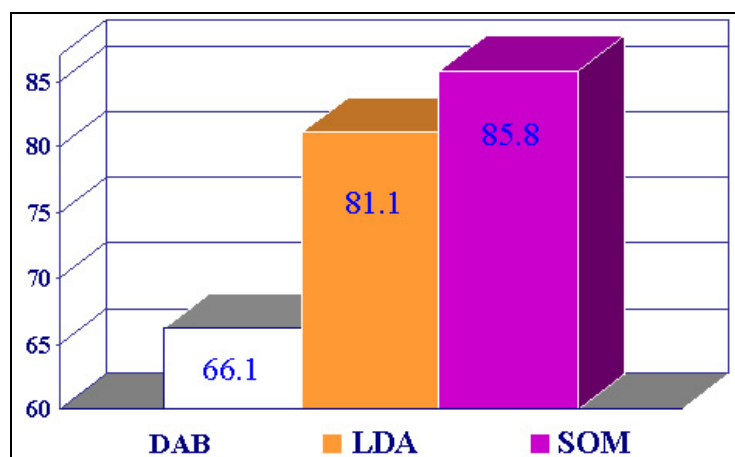
Como resultado de un análisis relacionado con la correlación entre especialistas y criterios de síntesis del tipo de patología se determinó que 20 rasgos de la voz eran más importantes para describir las enfermedades analizadas[22]. Once rasgos fueron apreciados más confiablemente mediante el uso de algoritmos de procesamiento digital del habla (PDH), mientras 9 rasgos son tomados directamente del promedio de las valoraciones subjetivas hechas por tres especialistas en lenguaje patológico.

Entre los rasgos más confiablemente apreciados mediante el uso de algoritmos computacionales están: Nivel del Período, Rupturas del período, temblor de la voz, excesivas variaciones del nivel, voz áspera, voz jadeante, interrupciones de la voz, inspiraciones audibles, razón del habla, frases cortas y ráfagas cortas del habla. En varios de estos se desarrollaron algoritmos novedosos que proporcionan mejores resultados que algoritmos similares reportados en la literatura especializada [1][3][5][6][7][9][13]. Entre los rasgos mejor apreciados de forma subjetiva por parte de los especialistas están: monotonía del nivel, hipernasalidad, estrés reducido, razón variable, intervalos prolongados, silencios inapropiados, estrés igual o excesivo, rupturas irregulares de la articulación y vocales distorsionadas.

El conjunto final de rasgos fueron extraídos en todos los pacientes de la base de datos para evaluar la efectividad de tres técnicas fundamentales de clasificación. La primera es el clasificador tradicional propuesto por Darley, Aronson y Brown mediante su diagnóstico diferencial de disártrias [22], la segunda esta basada en el análisis mediante técnicas de discriminante lineal (LDA) y la tercera esta basada en algoritmos no lineales usando mapas auto-organizados de Kohonen (SOM) [24]. Los clasificadores no solo tienen como objetivo clasificar correctamente los diferentes grupos de disártrias sino que deben permitir el estudio de la contribución que realiza cada rasgo de entrada en la selección de un grupo dado y permitir así correlacionar los rasgos de la voz con la lesión en el SNCP.

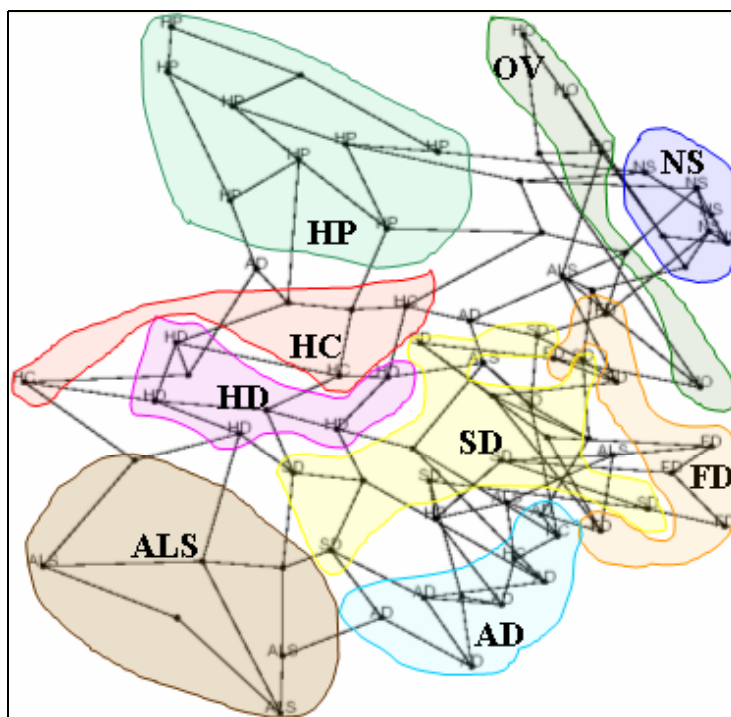
## **III. RESULTADOS Y DISCUSIÓN**

La evaluación realizada a los diferentes tipos de clasificadores analizados en cuanto a porcentaje de clasificación correcta (PPC) y la efectividad del análisis de relevancia de los rasgos de entrada en la decisión del clasificador, evidenció un mejor comportamiento del clasificador no lineal. La figura 4 muestra los resultados del proceso de clasificación.



**Figura 4** Porcentaje de clasificación correcta de los clasificadores evaluados.

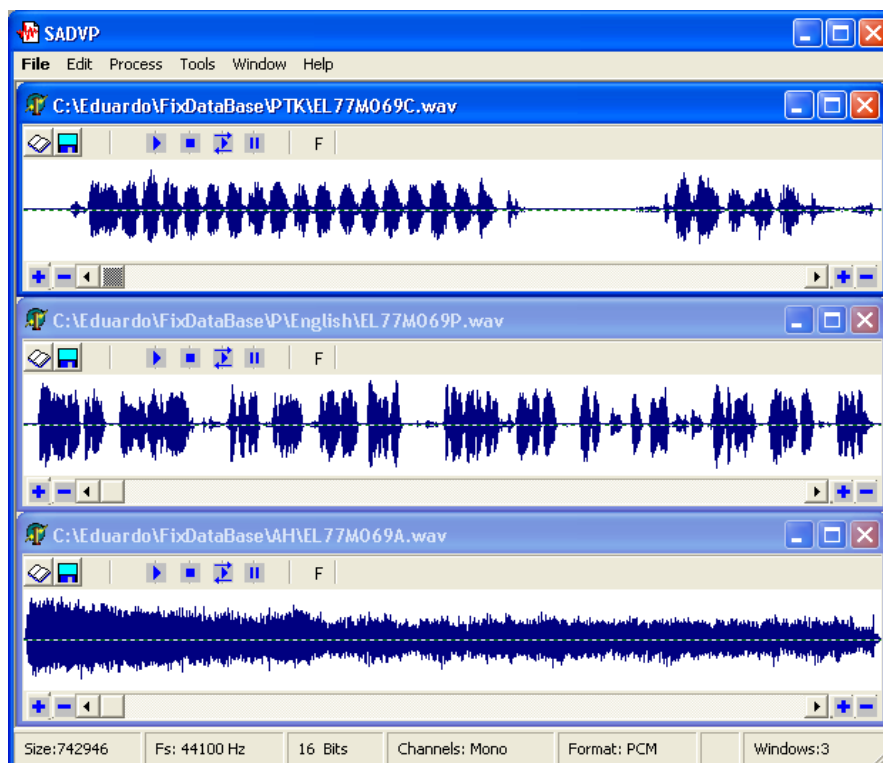
Se puede apreciar que el clasificador LDA supera la efectividad mostrada por el clasificador tradicional en un 15%, mientras que el clasificador SOM supera a ambos por cerca de 20% y 5% respectivamente. El clasificador SOM no solo proporciona un mayor PPC, sino que permite un análisis de relevancia más preciso y proporciona un mapa bidimensional de la distribución espacial de los diferentes grupos de disártrias. La figura 5 muestra una asociación evidente entre las neuronas que responden a los diferentes grupos de disártrias sombreados. Las áreas correspondientes a la disartria flácida (FD) y espástica (SD) están más solapadas entre sí debido a que tienen un gran grupo de rasgos comunes y la combinación de ambas conforman el grupo ALS [2].



**Figura 5** Representación bidimensional del mapa auto-organizado de la red neuronal artificial.

Este clasificador aprende a reconocer las regularidades y correlaciones del conjunto de entrada y adapta las respuestas futuras de acuerdo a las entradas. Este tipo de red neuronal no solo aprende a reconocer grupos de vectores de entrada similares sino que también las neuronas vecinas aprenden a reconocer la vecindad de las secciones delimitadas por el espacio de entrada. Por consiguiente, las SOMs aprenden la distribución y la topología de los vectores de entrada que son usados para su entrenamiento y producen un vector de pesos asociados a cada neurona. Dicho vector sirve para identificar cuales de los rasgos de entrada contribuyen más a la decisión final del clasificador. Esto determina que su uso en este tipo de aplicación sea muy satisfactoria donde el análisis de la contribución de los rasgos es decisivo para el estudio de este grupo de enfermedades.

Los resultados de la realización de la metodología muestran una superioridad de los métodos de clasificación que usan combinaciones de parámetros valorados objetiva y subjetivamente, corroborando la efectividad de esta nueva metodología desarrollada. El sistema experto desarrollado agrupa en un programa la metodología, los algoritmos de caracterización de voces patológicas y el clasificador para proporcionar una herramienta de ayuda al diagnóstico, el tratamiento y la investigación. La figura 6 y 7 muestran la interfase desarrollada para el sistema experto y la base de datos para el análisis subjetivo de los rasgos de la voz respectivamente.



**Figura 6** Interfase del sistema experto para la clasificación de disartrias.

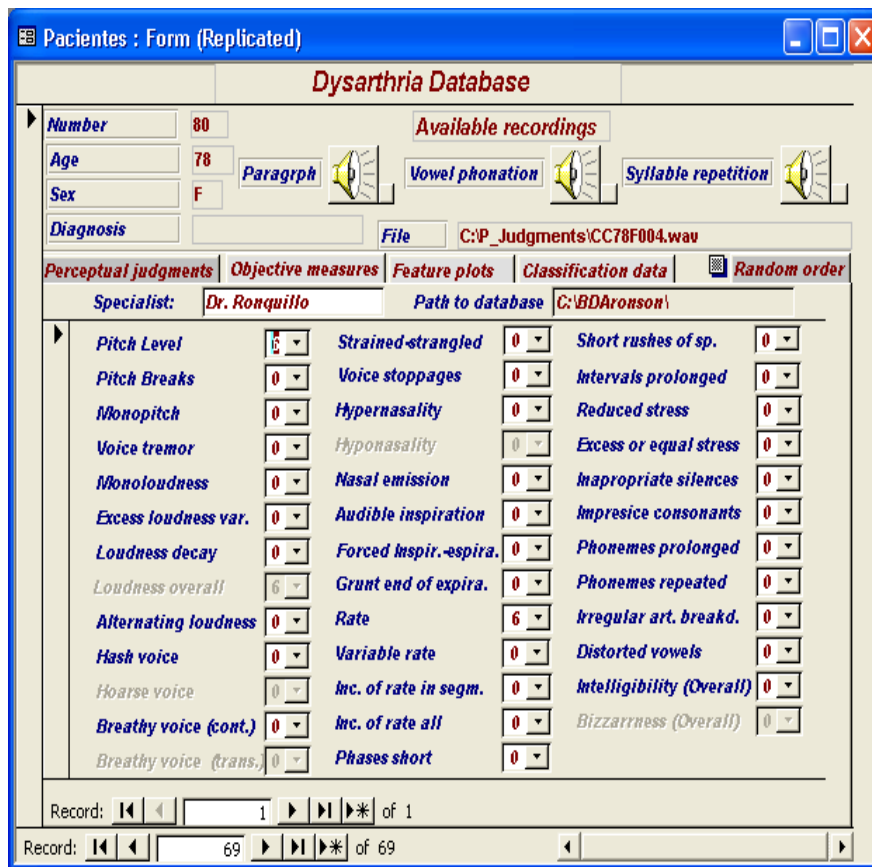


Figura 7 Interfase de la base de datos para el estudio subjetivo de los rasgos de la voz.

La colección de rasgos extraídos mediante el programa permite dar seguimiento a la evolución de las enfermedades, estudiar la efectividad de las medicinas recomendadas, compartir información con otros especialistas en lugares remotos para obtener criterios alternativos e incrementar la información disponible para desarrollar investigaciones en el tema; además de constituir una evidencia objetiva del estado del paciente.

El sistema experto incrementa también las opciones de los pacientes de países en desarrollo que pueden ser examinados por personal más entrenado mediante colaboración remota con especialistas más especializados usando las tecnologías de la información disponible actualmente [4].

#### IV. CONCLUSIONES

Se ha demostrado mediante el análisis realizado que los ocho grupos de disartrias estudiados son más precisamente descritos mediante rasgos obtenidos a través de combinaciones de valoraciones objetivas y subjetivas. El clasificador no lineal que usa SOM ha mostrado resultados superiores a los restantes y constituye una excelente herramienta el estudio de este tipo de enfermedad mediante su integración a un sistema experto diseñado para estos fines.



El desarrollo del presente trabajo conllevó a la producción de contribuciones originales relacionadas con la clasificación de disártrias. Las contribuciones pueden ser separadas en dos formas principales: una relacionada con la caracterización de las voces patológicas y la otra con la clasificación y estudio de las disártrias. Las contribuciones relacionadas con la caracterización de las voces patológicas se basan en el desarrollo de novedosos algoritmos para la apreciación automática de diferentes perturbaciones de la voz entre los que se encuentran:

1. Nivel del Período fundamental
2. Rupturas del período fundamental
3. Temblor de la voz
4. Excesivas variaciones del nivel
5. Voz jadeante
6. Inspiraciones audibles

Las contribuciones relacionadas con la clasificación y estudio de las disártrias están definidas por:

1. Desarrollo de la nueva metodología donde se combinan medidas objetivas y subjetivas de la voz y se seleccionan 20 parámetros más relevantes para el diagnóstico.
2. Desarrollo del Clasificador SOM que no solo proporciona un mejor PPC sino que ofrece un análisis más preciso de la relevancia de los rasgos y proporciona el mapa bidimensional de las enfermedades.
3. La base de datos de pacientes disártricos que posibilita el desarrollo de investigaciones futuras en el tema.
4. Desarrollo del sistema experto que produce un conjunto de parámetros que proporciona medidas objetivas acerca del estado de salud de los pacientes, además de permitir la colaboración remota entre especialistas.

## V. REFERENCIAS BIBLIOGRÁFICAS

- [1] **Eduardo Castillo Guerra:** A modern approach to dysarthria classification, Tesis Doctoral, Universidad de New Brunswick, Canada, Mayo, 2003.
- [2] **Eduardo Castillo Guerra & Dennis F Lovely:** Suboptimal classifier for dysarthria assessment. Lecture Notes on Computer Science. Aceptado para publicar.
- [3] **Eduardo Castillo Guerra y Dennis F Lovely:** Contributions to a modern approach to dysarthria classification ", Memorias del XXV Congreso Mundial de la Sociedad de Biología y Medicina de la IEEE, Cancún, México, Agosto de 2003, ISBN 0-7803-7790-7, Catalogo número 03CH37439C.
- [4] **Eduardo Castillo Guerra:** State of the Art and Trends in Dysarthria Classification", Editorial Feijoo, UCLV, Junio 2003, ISBN: 959-250-094-0.
- [5] **Eduardo Castillo Guerra y Carlos A. Ferrer Riesgo:** Monotonía del período fundamental en lenguaje patológico", Ingeniería Electrónica, Automática y Computación, vol. XXII, No.2, p.p. 75-78, 2001.

- [6] **Eduardo Castillo Guerra y María E. Hernández-Díaz Huici:** Caracterización automática de voces patológicas: monotonía del nivel”, Memorias del II Congreso Latinoamericano de Ingeniería Biomédica Habana’2001, Ciudad de la Habana, mayo de 2001.
- [7] **Eduardo Castillo Guerra y Carlos A. Ferrer riesgo:** Caracterización de rupturas del período en voces patológicas”, Memorias del II Congreso Latinoamericano de Ingeniería Biomédica Habana’2001, Ciudad de la Habana, mayo de 2001.
- [8] **Eduardo Castillo Guerra y Norma L. Méndez Rodríguez:** Metodología para la obtención de una base de datos de voces patológicas”, Memorias del VII Simposio Internacional sobre Comunicación Social, Santiago de Cuba, enero de 2001.
- [9] **Eduardo Castillo Guerra y Carlos A. Ferrer Riesgo:** Estudios de parámetros relativos al período fundamental en voces patológicas”, Memorias del VII Simposio Internacional sobre Comunicación Social, Santiago de Cuba, enero de 2001.
- [10] **Eduardo Castillo Guerra:** Alteraciones a la prosodia en pacientes con Parkinson”, Memorias del V Simposio Internacional sobre Comunicación Social, Santiago de Cuba, enero de 1999.
- [11] **Eduardo Castillo Guerra y Dennis F Lovely:** Contributions to a modern approach to dysarthria classification, XXV Congreso Mundial de la Sociedad de Biología y Medicina de la IEEE, Cancún, México, Agosto de 2003.
- [12] **Eduardo Castillo Guerra y Dennis F Lovely:** Sistema experto para la clasificación de disártrias”, V Congreso de la Sociedad Cubana de Bioingeniería, Ciudad de la Habana, junio de 2003.
- [13] **Eduardo Castillo Guerra y Juan V. Lorenzo Ginori:** Experiencias en la cuantificación de temblor en voces patológicas, X Simposio de Ingeniería Eléctrica, Universidad Central "Marta Abreu" de Las Villas, Junio de 2001.
- [14] **Eduardo Castillo Guerra y María E. Hernández-Díaz Huici:** Caracterización automática de voces patológicas: monotonía del nivel, II Congreso Latinoamericano de Ingeniería Biomédica Habana’2001, Ciudad de la Habana, mayo de 2001.
- [15] **Eduardo Castillo Guerra y Carlos A. Ferrer riesgo:** Caracterización de rupturas del período en voces patológicas, II Congreso Latinoamericano de Ingeniería Biomédica Habana’2001, Ciudad de la Habana, mayo de 2001.
- [16] **Eduardo Castillo Guerra y Norma L. Méndez Rodríguez:** Metodología para la obtención de una base de datos de voces patológicas, VII Simposio Internacional sobre Comunicación Social, Santiago de Cuba, enero de 2001.
- [17] **Eduardo Castillo Guerra y Carlos A. Ferrer Riesgo:** Estudios de parámetros relativos al período fundamental en voces patológicas, VII Simposio Internacional sobre Comunicación Social, Santiago de Cuba, enero de 2001.
- [18] **Eduardo Castillo Guerra:** Alteraciones a la prosodia en pacientes con Parkinson”, V Simposio Internacional sobre Comunicación Social, Santiago de Cuba, enero de 1999.
- [19] **Eduardo Castillo Guerra y Dennis F Lovely:** Procedimiento para caracterizar automáticamente inspiraciones audibles en voces patológicas”, patente solicitada al OCPI # 2-2003-01-05, marzo del 2003.

- [20] **Eduardo Castillo Guerra y Dennis F Lovely:** Procedimiento para apreciar automáticamente la perturbación producida por el jadeo en voces patológicas, patente solicitada al OCPI # 3-2003-00-73, febrero del 2003.
- [21] **Eduardo Castillo Guerra y Dennis F Lovely:** Procedimiento para mejorar la efectividad en la clasificación de disártrias, patente solicitada al OCPI # 1-2003-01-06, febrero del 2003.
- [22] **Darley F.L, Aronson A.E., Brown J.R.:** Differential diagnostic patterns of dysarthrias, JSHR, vol. 12, (1969a) 462-496.
- [23] **Aronson, A.E., Dysarthrias:** differential diagnosis (audio tape), Rochester, MN: Mentor Seminars (1993)
- [24] **Kohonen, T.,** "Self-organized formation of topologically correct features", Biological Cybernetics, vol. 43, (1982) 59-69.
- [25] **Kay Elemetrics:** Multi-dimensional voice program (MDVP™), model 4305, Lincoln park, NJ: Kay Elemetric Corporation, 1993.
- [26] **Kent, R.D. & Read C:** The acoustic analysis of speech, Singular publishing group Inc, First Edition, 1992.
- [27] **Boersma, P. & Weenink D.:** Praats, a system for doing phonetics by computers, Ver 4.0, Summer Institute of Linguistic, Amsterdam, , 2001.