



UNIVERSIDAD CENTRAL "MARTA ABREU" DE LAS VILLAS
VERITATE SOLA NOBIS IMPONETUR VIRILISTOGA. 1948

Facultad de Ingeniería Eléctrica

**Centro de Estudios de Electrónica y Tecnologías de la
Información (CEETI)**



TRABAJO DE DIPLOMA

Influencia de Modelos de *Jitter* en Medidas de Perturbación de la Periodicidad

Autor: Alexi José González Hernández

Tutor: Dr CT Carlos Ariel Ferrer Riesgo

Santa Clara

2011

"Año 53 de la Revolución"



Universidad Central “Marta Abreu” de Las Villas

Facultad de Ingeniería Eléctrica

**Centro de Estudios de Electrónica y Tecnologías de la
Información (CEETI)**



TRABAJO DE DIPLOMA

Influencia de Modelos de *Jitter* en Medidas de Perturbación de la Periodicidad

Autor: Alexi José González Hernández

alexijoseg@uclv.edu.cu

Tutor: Dr CT Carlos Ariel Ferrer Riesgo

cferrer@uclv.edu.cu

Consultante: MSc. Diana Torres Boza.

dtb@uclv.edu.cu

Santa Clara

2011

"Año 53 de la Revolución"

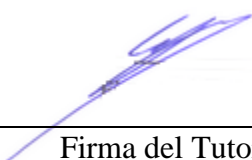


Hago constar que el presente trabajo de diploma fue realizado en la Universidad Central “Marta Abreu” de Las Villas como parte de la culminación de estudios de la especialidad de Ingeniería en Automática, autorizando a que el mismo sea utilizado por la Institución, para los fines que estime conveniente, tanto de forma parcial como total y que además no podrá ser presentado en eventos, ni publicados sin autorización de la Universidad.

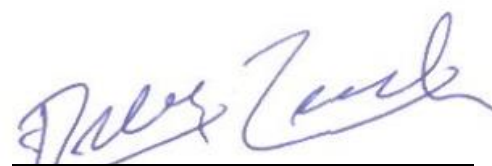


Firma del Autor


Los abajo firmantes certificamos que el presente trabajo ha sido realizado según acuerdo de la dirección de nuestro centro y el mismo cumple con los requisitos que debe tener un trabajo de esta envergadura referido a la temática señalada.



Firma del Tutor



Firma del Jefe de Departamento
donde se defiende el trabajo



Firma del Responsable de
Información Científico-Técnica

PENSAMIENTO

Cuida tus pensamientos porque se volverán palabras. Cuida tus palabras porque se volverán actos. Cuida tus actos porque se harán costumbres. Cuida tus costumbres porque forjarán tu carácter. Cuida tu carácter porque formará tu destino. Y cuida tu destino porque será tu vida.

Mahatma Ghandi

DEDICATORIA

A la memoria de mi padre que aunque ya no está siempre iluminó mi camino en ese sendero que todos transitamos que es la vida, a mi madre que supo llevar el peso de esta familia en su pecho, a mi hermano por su apoyo y transparencia, a mi abuela por sus consejos, a mi abuelo, a mi tía, en fin a toda la familia por estar siempre a mi lado y apoyarme en todo momento. A ellos les dedico este trabajo.

AGRADECIMIENTOS

A mi tutor Dr. CT. Carlos Ariel Ferrer Riesgo , porque me ha sido un faro en estos años de mi vida universitaria, por su ejemplo y por su ayuda incondicional, a usted le estoy eternamente agradecido.

A mí consultante la MSc Diana Torres Boza por sus consejos y ayuda.

Al Dr CT. Eduardo Gonzales Moreira por su actitud responsable, a la Ing. Beatriz Guerra Santana por estar a mi lado cuando lo necesité por muy difíciles que fuesen las circunstancias.

A todos los profesores y trabajadores de la UCLV en especial a los del CEETI por su profesionalidad, respeto y sabiduría.

A los padres de Beatriz que estuvieron dándome su apoyo incondicional como si fuesen los míos.

A mis compañeros de cuarto Alexey y Liam por estar siempre en las buenas y en las malas.

A mis compañeros de estudios que han compartido conmigo a lo largo de 5 años en todo momento, por comprenderme y brindarme su ayuda.

A todas las personas que han dedicado parte de su tiempo para ofrecer su ayuda y sugerencias y brindarme todo su apoyo en este trabajo.


... a todos, muchas gracias.

TAREA TÉCNICA

- Determinación de los modelos, medidas y pruebas a emplear.
- Síntesis de las señales con niveles de perturbación variable.
- Obtención de las medidas de perturbación.
- Evaluación de los resultados
- Escritura del Informe



Firma del Autor



Firma del Tutor

RESUMEN

La voz contiene información útil para el diagnóstico médico. Entre los parámetros que frecuentemente se utilizan para evaluar el estado del paciente se encuentran los de calidad vocal, relacionados con las perturbaciones de la periodicidad de la señal. Estas perturbaciones: duración del ciclo (*jitter*) y amplitud de pulso (*shimmer*) se presentan en voces patológicas como jadeo y aspereza.

Para obtener estas medidas de perturbación se requieren determinar las posiciones de los límites entre ciclos consecutivos de la señal (el período de los mismos). La presencia de *jitter* puede distorsionar el pulso glotal de diferentes maneras, y cada una de estas maneras afectar el cálculo de las medidas de perturbación en diferente magnitud. Este trabajo aborda esta problemática. Se evalúa el comportamiento de métodos reportados de estimación de duración y amplitud de los pulsos glotales, ante la presencia de niveles patológicos de *jitter*, considerando dos modelos de perturbación del pulso.

TABLA DE CONTENIDOS

PENSAMIENTO	i
DEDICATORIA	ii
AGRADECIMIENTOS	iii
TAREA TÉCNICA.....	iv
RESUMEN	v
INTRODUCCIÓN	1
CAPÍTULO 1. REVISIÓN BIBLIOGRÁFICA.....	5
1.1 Modelos de producción de la voz. Modelo Fuente-Filtro.....	5
1.1.1 Sonidos sonoros y sordos.....	7
1.1.2 El pulso glotal.	7
1.2 Perturbaciones de la Periodicidad. <i>Jitter</i> , <i>Shimmer</i> y Ruido.	8
1.3 Diferentes concepciones de <i>Jitter</i> . Implicaciones.....	9
1.3.1 Ejemplos de uso del modelo S/A (<i>jitter</i> por posición).....	10
1.3.2 Ejemplos de uso del modelo E/C (<i>jitter</i> por duración).	11
1.4 Conclusiones del capítulo	13
CAPÍTULO 2. MATERIALES Y MÉTODOS.....	14
2.1 Modelos de influencia del <i>jitter</i> en la forma de onda de los pulsos.....	14
2.2 Síntesis de señales.....	14

2.3	Medición de <i>jitter</i>	17
2.3.1	PDA's evaluados.....	19
2.4	Medición de <i>shimmer</i>	24
CAPÍTULO 3. RESULTADOS		25
3.1	PDA's.....	25
3.2	<i>Shimmer</i>	27
3.3	Conclusiones del capítulo	29
CONCLUSIONES Y RECOMENDACIONES		31
Conclusiones		31
Recomendaciones		31
REFERENCIAS BIBLIOGRÁFICAS		32
ANEXOS		36
Anexo I	PDA's.....	36
Anexo II	Evaluator	38
Anexo III	Generador de Onda	39
Anexo IV	Programa principal.....	44
Anexo V	Detector de <i>Shimmer</i>	47

INTRODUCCIÓN

El habla es de gran importancia por ser la principal vía de comunicación entre los seres humanos. En el proceso de producción del habla intervienen varios órganos y sistemas, por lo que la afección de estos repercute en la forma en que se genera y percibe el mismo. La influencia de dichas afecciones provoca que el habla porte no sólo el mensaje, estados emocionales o de ánimo y procedencia social de un individuo, sino también información útil desde el punto de vista clínico. Por ser tan variada la información contenida en la señal de voz esta ha sido utilizada como objeto de estudio en ramas como:

- Sistemas de comunicación, incluyendo el análisis, transmisión, síntesis, reconocimiento de voz, etc.
- Fonética y lingüística, comprendiendo el estudio de la entonación, los aspectos emocionales del habla, el estudio de los fonemas en diferentes idiomas, etc.
- Educación, abarcando tanto la ayuda en la enseñanza del habla a los sordos como el aprendizaje de lenguas extranjeras.
- Medicina, englobando tanto el diagnóstico de enfermedades a partir del habla como la obtención de información sobre el progreso de pacientes en proceso de rehabilitación.

El campo de las aplicaciones médicas del procesamiento de voz provoca gran interés en la comunidad científica internacional. En particular, resulta de gran utilidad la obtención de medidas capaces de indicar la severidad de las manifestaciones patológicas del habla. Contar con indicadores de este tipo permite efectuar diagnósticos más acertados, así como documentar de forma objetiva la evolución del paciente y la efectividad del tratamiento. Las diferentes medidas pueden clasificarse, en cuanto al nivel de las unidades del lenguaje involucradas, en tres categorías principales:

- Medidas de Calidad Vocal: Determinan perturbaciones en fonemas (sonidos individuales).
- Medidas de Articulación: Indican alteración de transiciones entre sonidos, a nivel de sílaba o palabra.
- Medidas de Prosodia: Evalúan el grado de alteración a nivel de frases, oraciones o párrafos.

A medida que se incrementa la complejidad de las unidades del lenguaje a analizar (desde los fonemas a las oraciones transitando de calidad vocal a prosodia) se incrementa también la complejidad del desarrollo y la obtención de mediciones correspondientes. Por este motivo, en las mediciones relacionadas con Calidad Vocal se han reportado más resultados que en Articulación y Prosodia, tendencia que se mantiene en la actualidad. Este trabajo se mantiene en la línea de desarrollo de mediciones de Calidad Vocal. Las medidas de este grupo se evalúan en segmentos sonoros donde la señal es aproximadamente periódica.

Entre las mediciones más frecuentes de Calidad Vocal se pueden mencionar las perturbaciones de la periodicidad dadas por la presencia de variaciones de frecuencia, amplitud y ruido aditivo en la señal. Estas mediciones de perturbación de la periodicidad dependen de la correcta determinación de los instantes de excitación glotal. Los métodos que extraen las posiciones de los instantes de excitación glotal de la señal de voz se denominan detectores de período fundamental (PDA's, del inglés *Pitch Detection Algorithms*). La detección de período fundamental no sólo es empleada con fines médicos, sino también en aplicaciones del procesamiento de voz como son la síntesis y codificación de voz y el reconocimiento del locutor, lo que la convierte en una tarea necesaria y frecuente. Esta amplia utilización de los algoritmos detectores en las diferentes ramas del procesamiento de voz ha traído como consecuencia el constante desarrollo y perfeccionamiento de infinidad de detectores. Por tal razón ha sido necesario el desarrollo de medidas de comparación entre los PDA's para la evaluación de los mismos y así garantizar su correcta selección para cada tarea específica. Sin embargo, en muchas ocasiones las medidas y procedimientos de comparación de los PDA's son inapropiados.

Una de las aplicaciones donde más fallan los PDA's es en el habla patológica, donde los niveles de perturbación presentes violan las suposiciones de periodicidad en que se basan

los PDA's. Las perturbaciones típicas incluyen alteraciones de la amplitud (*shimmer*) y duración (*jitter*) de los pulsos, además del ruido aditivo. La perturbación que impone mayores dificultades a las suposiciones de periodicidad es el *jitter* (Schoentgen 2003, Ferrer 2007), por lo que el estudio de su influencia y los modos de reducirla resultan de particular interés.

Existen varias formas de concebir la modificación de la forma del pulso glotal en presencia de *jitter* (variaciones en la duración total del mismo). Dos modelos empíricos muy empleados son por una parte la contracción/expansión de todas las fases del pulso de manera proporcional al cambio de la duración del pulso, y por otra la afectación sólo de la fase cerrada del pulso. Ambos modelos tienen implicaciones muy diferentes en la manera en que se afecta la señal, y presumiblemente en el rendimiento de los algoritmos diseñados para medir las perturbaciones de la periodicidad (*jitter*) y la amplitud (*shimmer*) de los pulsos. En los trabajos reportados en la literatura se emplean siempre sólo uno de los modelos, y por tanto los resultados no son comparables con trabajos que empleen otro modelo. Esta situación lleva al siguiente:

Problema Científico: Ausencia de referencias de cuánto afectan cada uno de estos modelos la estimación de las diferentes medidas de perturbación.

Objetivo General: Determinar cómo se afectan diferentes medidas de perturbación de la periodicidad ante los distintos modelos de modificación del pulso por causa del *jitter*.

Objetivos Específicos:

- Evaluar afectación de diferentes detectores de período fundamental ante ambos modelos de modificación del pulso.
- Evaluar afectación de diferentes medidas de *shimmer* ante ambos modelos de modificación del pulso.

Tareas Técnicas:

- Determinación de los modelos, medidas y pruebas a emplear.
- Síntesis de las señales con niveles de perturbación variable.
- Obtención de las medidas de perturbación.

- Evaluación de los resultados

Organización del informe

En el Capítulo 1 se presentan los principales conceptos que se tratan en este trabajo, introduciendo el tema de las mediciones del habla. Se brinda una panorámica general existente en torno al problema que se aborda y que motiva este trabajo.

En el Capítulo 2 se describen los modelos de producción que se consideraron en este trabajo. Se realiza la evaluación del desempeño de los Detectores de Período Fundamental y se explica la síntesis de las señales empleadas.

En el Capítulo 3 se muestran los resultados obtenidos en los experimentos y se efectúa su análisis, planteando las implicaciones de los mismos en cada caso.

CAPÍTULO 1. REVISIÓN BIBLIOGRÁFICA

El habla, como medio de comunicación humana, porta información en varios niveles: lingüístico, para-lingüístico y extra-lingüístico (Laver, 1991). En el primero se incluye la información de lo que se desea transmitir (el “texto” del mensaje). El segundo porta información sobre los estados de ánimo, emocionales u otros, de la persona, así como su procedencia social y otros factores culturalmente determinados. El tercer nivel comprende todos los factores físicos y fisiológicos (incluyendo todos los aspectos orgánicos involucrados en la producción del habla) característicos del locutor.

Dada la cantidad de información que porta el habla y la importancia que reviste para los seres humanos, por ser la vía fundamental de comunicación, se comprende la amplitud de las investigaciones que se dedican al estudio del habla en estos tres niveles. En este trabajo se abordan técnicas para analizar la información extra-lingüística del habla.

1.1 Modelos de producción de la voz. Modelo Fuente-Filtro.

El proceso de producción del habla pasa por varias etapas, desde la concepción del mensaje en el cerebro, la transmisión de los impulsos nerviosos necesarios a cada uno de los órganos involucrados, y la ejecución de los movimientos correspondientes en los mismos. En esta última parte, motora, radica el basamento físico de la generación de la voz, y sus características acústicas.

En la generación de la señal de voz intervienen varias estructuras del cuerpo humano, las más importantes las cuales se muestran en la Figura.1.1.

El funcionamiento simplificado del sistema desde el punto de vista físico puede explicarse (Liljencrants, 1991) de manera similar a un instrumento musical de viento (órgano, flauta).

En este caso, los pulmones contienen aire, que es comprimido por el diafragma de tal forma que esa presión produce sonidos al hacer pasar el aire por alguna constricción o estrechamiento en su trayecto entre las cuerdas vocales y los labios.

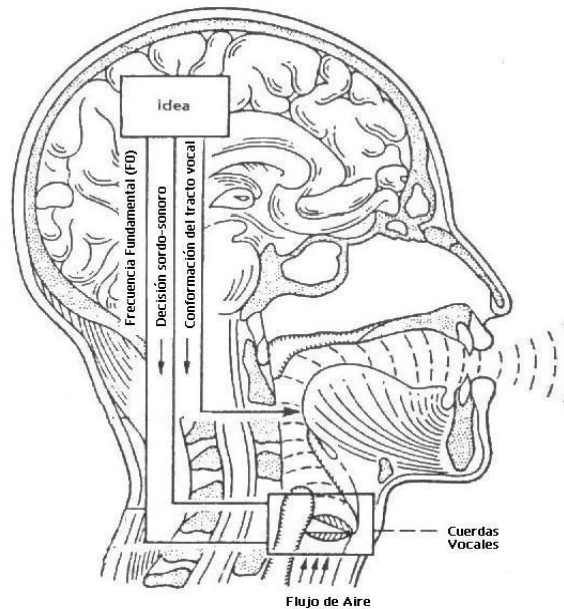


Figura.1.1 Modelo de Producción de Voz

Modelo Fuente Filtro.

El modelo más comúnmente empleado en síntesis y análisis de voz es el modelo Fuente-Filtro (Fant 1960, Titze 1995, Stevens 1997, Marasek 1997, Veldhuis 1998), donde se asume que los pulsos de aire $g(t)$ generados en la glotis son conformados espectralmente por la función de transferencia del tracto vocal $H(f)$ y posteriormente por la radiación en los labios $R(f)$ (ver Figura.1.2). La principal ventaja de este modelo es su sencillez, lo que no le resta aplicabilidad, y los resultados de su uso son la base de los adelantos en la codificación y transmisión de voz. Su limitación principal radica, simultáneamente, en su sencillez (Steinecke, 1995, Titze, 1980). Otro aspecto a considerar es la suposición de estacionariedad del tracto vocal, o sea, que el tracto no varía su posición en el intervalo de análisis, pero aquí también se imponen consideraciones prácticas para poder aplicar las técnicas más comunes de análisis, como la correlación, la Transformada de Fourier y otras.

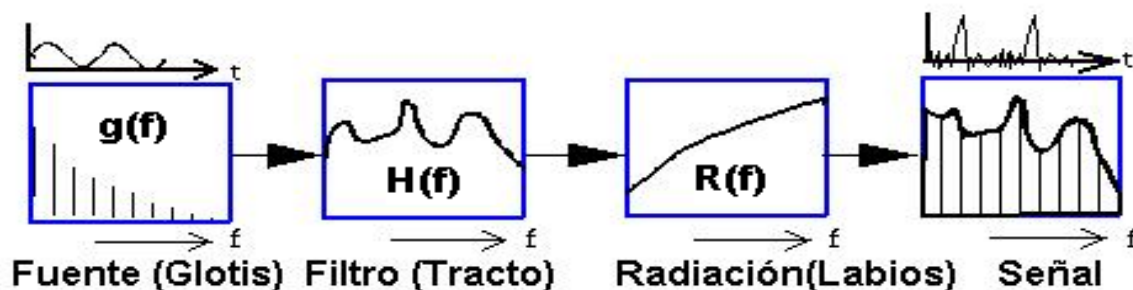


Figura.1.2 Modelo Fuente-Filtro

1.1.1 Sonidos sonoros y sordos.

Los sonidos generados se clasifican en sonoros o sordos en dependencia de si existe periodicidad o no en la señal emitida. Los sonidos sonoros (periódicos) se generan cuando la constricción se produce en las cuerdas vocales. Los sonidos sonoros son típicos de las vocales (y algunas consonantes como “m”, “n”, “b”, etc.). Los sordos (aperiódicos), son únicos de las consonantes, y se generan al hacer pasar el aire por constricciones en el tracto vocal (consonantes fricativas: “f”, “s”, “j” y otras) o al liberar repentinamente una oclusión en el mismo (consonantes explosivas sordas: “p”, “t” y “k”).

1.1.2 El pulso glotal.

La obtención de la forma de onda de $g(t)$ a emplear no es un problema sencillo. Algunas formas de medición directa en sujetos vivos han sido altamente invasivas debido a la colocación de transductores. Otras basadas en el cálculo del área de las cuerdas vocales como la electroglotografía (Fourcin, 2000) y laringoboscopia (Cranen & de Jong, 2000) han fallado debido a la conductividad eléctrica transglotal y los requisitos muy estrictos con la configuración instrumental para obtener una secuencia de imágenes adecuada para medir el área, respectivamente.

Ante las dificultades que se reportan en la estimación de la señal de excitación (ya sea experimental directa o mediante modelos físicos) ha sido frecuente la búsqueda de alternativas con menor basamento físico y mayor abstracción matemática. El método más usado es el empleo de la señal residual del filtrado inverso para estimar la forma de onda de $g(t)$, aunque el filtrado inverso tiene aparejadas una serie de limitaciones que se tornan

severas en voces patológicas. Las formas de onda obtenidas han sido representadas mediante modelos paramétricos donde se le asignan determinadas funciones matemáticas a las fases de apertura y cierre de la forma de onda de $g(t)$ (ver Figura 1.3). Entre estos modelos los más empleados son los polinomiales y trigonométricos de Rosenberg (Rosenberg, 1971) y el modelo de Liljencrants-Fant (Fant et al., 1985).

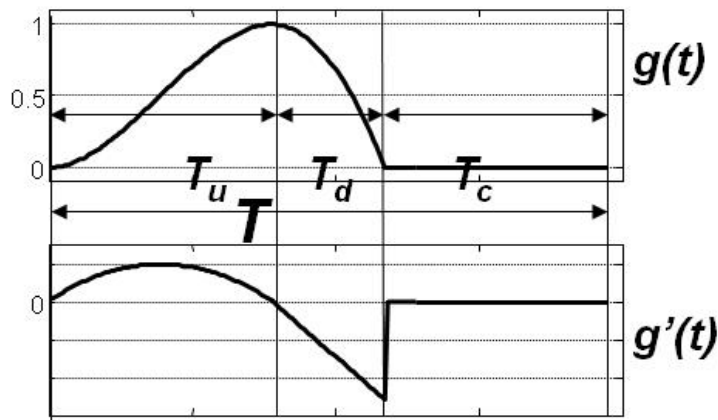


Figura.1.3 Formas de onda y fases de la señal de excitación glotal $g(t)$ y su derivada $g'(t)$. T_u es la duración de la fase de apertura, T_d la del cierre, T_c la fase cerrada, y T la duración total del pulso.

1.2 Perturbaciones de la Periodicidad. *Jitter*, *Shimmer* y Ruido.

Existen varios tipos de perturbaciones de la periodicidad como son:

1. Perturbación de la duración de los pulsos (conocida como *jitter*), consiste en la separación de los pulsos en intervalos de tiempo no exactamente iguales a T .
2. Perturbación de la amplitud de los pulsos (conocida como *shimmer*), consiste en que los pulsos puedan expresarse uno en función de otro a través de un factor de proporcionalidad diferente de la unidad.
3. Presencia de ruido aditivo: Cada pulso se diferencia del otro en una magnitud aleatoria en cada instante de muestreo, de valor medio cero.
4. Perturbación de forma de onda: La $g(t)$ varía de pulso a pulso de manera que el cambio no se reduce a un factor de proporcionalidad.

Una expresión (Ferrer, 2005) que incluye todos estos factores es:

$$s(t) = h(t) * r(t) * \left(\sum_{i=1}^N g_i(t) * \delta(t - (T_{i-1} + \bar{T}_0 + \Delta T_{i-1})) \right) + e(t) \quad (1)$$

con $r(t)$ la respuesta al impulso del radiador, $g_i(t)$ la forma de onda de la fuente en el pulso i -ésimo, incluyendo perturbaciones de amplitud (factores 2 y 4), ΔT_i la perturbación de duración con respecto a T (factor 1), y $e(t)$ representando el ruido aditivo (factor 3).

1.3 Diferentes concepciones de *Jitter*. Implicaciones

Para este estudio se analizaron dos modelos de *jitter*, uno que concibe las perturbaciones de duración como contracciones o expansiones de los pulsos y otro que las concibe como acercamientos y separaciones de los mismos (ver Figura 1.4).

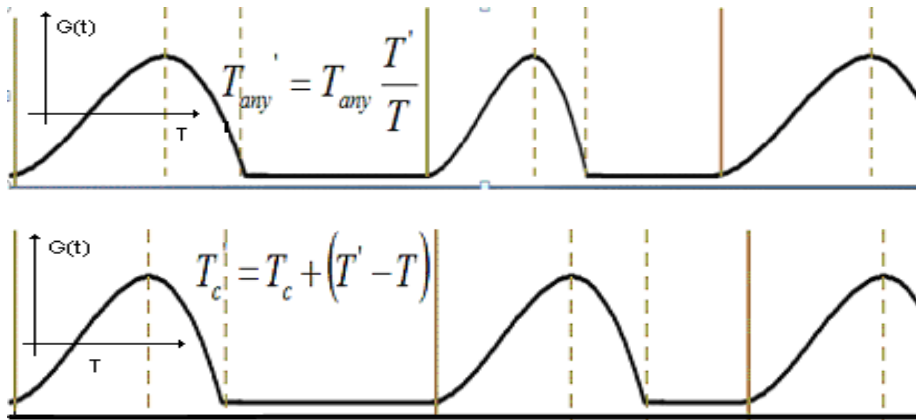


Figura.1.4 Diferentes modelos de *Jitter*: Arriba, Contracción/Expansión de los pulsos, Abajo, Acercamiento/Separación de los pulsos.

- Implicaciones de estas diferencias
 1. Posibilidad de *shimmer* causado por *jitter*.

Dado el efecto derivativo sobre $g(t)$, en el modelo Expansión/Contracción (E/C) aparece *shimmer*, y en el Separación/Acercamiento (S/A) no aparece debido a que no hay variación en la primera derivada.

2. Invariancia de medidas glotales

Medidas glotales ampliamente reportadas en la literatura como el *Open Quotient (OQ)* y el *Speed Ratio (SR)*, consistentes en relaciones entre duraciones de las fases del pulso glotal, se mantienen todas constantes en el modelo E/C, no siendo así en el modelo S/A.

3. Definición de las fronteras del pulso.

La secuencia (contorno) de posiciones o duraciones de los pulsos glotales es empleada en la evaluación del rendimiento de detectores de $T0$. Los detectores de período fundamental (PDA's) funcionan generalmente detectando los instantes de cierre, mientras el ciclo glotal se representa generalmente a partir del instante de apertura. En el modelo de E/C las separaciones entre dos pulsos adyacentes no son las mismas si se miden entre cierres que entre aperturas, mientras que en el modelo de S/A sí son iguales. Cuando se evalúa un detector de período fundamental hay que tener en cuenta esto, para evaluar su capacidad para seguir el contorno de los instantes que intenta detectar, y no el contorno de duraciones con que se sintetizaron los pulsos.

1.3.1 Ejemplos de uso del modelo S/A (*jitter* por posición).

1.3.1.1 Software para sintetizar formantes en cascada / paralelo. Klatt (1980).

En este trabajo se sintetizan formantes tanto en cascada como en paralelo para la síntesis digital del habla empleando computadoras, y se muestran las diferencias entre ambos modos y sus implicaciones. El artículo explica el empleo de algoritmos para simular el tracto vocal, la radiación de los labios, y la generación del estímulo (la onda glotal) y sus diferentes configuraciones y características. El modelo que se utiliza es el de S/A ya que su síntesis contempla el paso de un tren de impulsos por un filtro paso bajo, esto provoca que aparezca un pulso en las posiciones de los impulsos de Dirac, la forma del pulso (fase abierta) siempre es constante (Klatt, 1980).

1.3.1.2 Comparación de PDA's de alta precisión en vocales sostenidas. Parsa-Jamieson (1999).

Este artículo hace una breve descripción del funcionamiento de los PDA's y realiza una serie de experimentos con señales sintetizadas contaminadas con *jitter*, *shimmer* y ruido y con señales reales tanto normales como de voces patológicas. En el artículo se examinaron siete técnicas de estimación de pith de las cuales cinco fueron en el dominio del tiempo

(AMDF, AC, ACC, FIC, WM), una en el dominio de la frecuencia (HPS) y por último una combinación de los dos dominios (CEP). Se empleó como modelo de síntesis el pasar un tren de impulsos a través de un filtro representativo del tracto vocal, por lo que sigue el modelo de S/A, con comienzo del pulso en el instante de apertura (Parsa & Jamieson, 1999).

1.3.2 Ejemplos de uso del modelo E/C (*jitter* por duración).

1.3.2.1 Efecto de la forma del pulso glotal en la calidad de voces naturales. Rosenberg (1971).

El artículo evalúa la influencia de diferentes formas de onda para modelar el pulso glotal en la naturalidad con que se escucha el habla sintetizada. Es un artículo prolífico, donde se presenta por primera vez el modelo polinomial de Rosenberg, ampliamente utilizado con posterioridad. Las variaciones de *jitter* se modelan mediante el modelo de E/C, pues las relaciones de duración entre las fases se asumen constantes. El inicio del pulso se asume en el instante de cierre (Rosenberg, 1971).

1.3.2.2 Modelos espectrales de ruido aditivo y de modulación en señales de excitación y de habla. Schoentgen (2003).

El objetivo del artículo es modelar el efecto del modelo espectral de ruido aditivo y multiplicativo (o de modulación) así como las perturbaciones de la amplitud y duración del pulso glotal. El propósito del modelado es establecer criterios acerca de las causas del ruido en el espectro normal y las voces con desórdenes para intentar, con las propiedades del espectro de las perturbaciones de la señal de excitación fonatoria, deducir las propiedades espectrales de la señal del habla.

El análisis que se desarrolla en este artículo se centra en las perturbaciones de todo tipo de ruido. El término ‘ruido’ designa un término no deseado o señales extrañas, y con frecuencia, implica un componente estocástico. El análisis, sin embargo, es determinista, así como el patrón de las perturbaciones estocásticas a considerar. El modelo de excitación glotal que se usa es el de E/C ya que se asume una modulación de frecuencia provocando que la onda se expanda o se contraiga de acuerdo con la modulación. No se trabaja con una

forma de onda en específico, sino con las representaciones espectrales, por lo que no se declara de manera explícita el punto de inicio o final del pulso (Schoentgen, 2003).

1.3.2.3 Medidas de HNR libres de perturbación empleando análisis espectral. Murphy (1999).

En este artículo se trata el tema de la estimación de la Relación Armónico-Ruido (HNR por sus siglas en inglés *Harmonics to Noise Ratio*). El trabajo expone las posibles implicaciones que esta relación conlleva para el análisis de calidad vocal, expone una serie de concepciones como aperiodicidad, ruido, define cuales son las principales perturbaciones y hace un análisis detallado de estas. El objetivo fundamental es el de automatizar el procedimiento de detección en señales de habla patológica, donde coexisten niveles elevados de todo tipo de perturbación. Con este propósito se sintetizan señales donde se le introducen diferentes niveles de *jitter*, *shimmer* y ruido y se evalúa un nuevo método para determinar el HNR ante estas perturbaciones. El modelo glotal que utiliza es el polinomial de Rosemberg y la síntesis con *jitter* se efectúa escalando el pulso en correspondencia con un factor dado, por lo que se trata de un modelo de E/C. El pulso se asume de apertura a apertura. En correspondencia con este modelo, se declara que siempre el *jitter* induce *shimmer* (Murphy, 1999).

1.3.2.4 Declaración final del taller de análisis acústico del habla. Titze. (1995).

En esta declaración se explican las concepciones de cada perturbación y cómo modelarlas, así como su influencia en los modelos. Se describen matemáticamente las funciones de las perturbaciones y sus medidas. Se analizan diferentes tipos de modelos de señal para ver las implicaciones de estas en los algoritmos de detección de pitch. El modelo que se asume es el de E/C, lo cual se aprecia en las ilustraciones de las formas de onda, que muestran una modulación de frecuencia, provocando la expansión o contracción de los pulsos (Titze, 1995).

1.3.2.5 Análisis, síntesis y percepción de variaciones de la calidad vocal en hablantes femeninos y masculinos. Klatt&Klatt (1990).

En este artículo se emplean tres formas de excitación: una denominada KLSYN que es la empleada en (Klatt, 1980), correspondiente al modelo de S/A, y dos formas nuevas que responden al modelo E/C. Una es una modificación del modelo LF de Liljencrants-Fant

(Fant *et al.*, 1985) y la otra es la polinomial de (Rosemberg, 1971), con algunas características adicionales como la posibilidad de cierre incompleto de la glotis. El pulso comienza en el instante de apertura (Klatt&Klatt, 1990).

1.4 Conclusiones del capítulo

La manera en que se modifica la señal de excitación en presencia de *jitter* ha sido abordada de diferentes maneras en la literatura, y no existen reportes donde se compare la influencia de uno u otro modelo en el rendimiento de las medidas de perturbación de la periodicidad. El empleo de uno u otro modelo provoca diferencias significativas en aspectos importantes y diversos como la evaluación de PDAs, la aparición de *shimmer*, y las medidas de las fases del pulso glotal.

CAPÍTULO 2. MATERIALES Y MÉTODOS

En este capítulo se describen los modelos a utilizar en la síntesis de las señales a evaluar. Se describen además los niveles de perturbación introducidos y las medidas que se emplearán en la estimación de estos niveles. En el caso de la medición de *jitter*, se describen los diferentes PDA's evaluados.

2.1 Modelos de influencia del *jitter* en la forma de onda de los pulsos.

Los modelos considerados son los modelos de S/A y E/C expuestos en la sección 1.3. Es de destacar que no existen trabajos previos donde se evalúen simultáneamente ambos modelos. En los artículos disponibles ni siquiera se considera la posibilidad de la existencia de un modelo alternativo al que se asume en el trabajo en cuestión.

2.2 Síntesis de señales.

2.2.1 Consideraciones para seleccionar el tipo de señal:

En cuanto a los tipos de señales, se pueden mencionar tres disyuntivas principales: reales/simuladas, normales/patológicas y vocales/habla fluida y la decisión está en dependencia del propósito del detector de período fundamental.

Si el detector se desea emplear para aplicaciones de codificación y síntesis generalmente se usan señales reales, de habla fluida de personas normales, mientras que para medir perturbaciones de la periodicidad se emplean vocales aisladas normales y patológicas, ya sean reales o simuladas. En la decisión reales/simuladas están a favor de las últimas el conocer y poder cambiar los parámetros de simulación y la repetitividad de los resultados, y en contra tienen las simplificaciones en que se incurre al elegir un modelo de síntesis, y la

dudosa representatividad de las señales generadas. En contra de las reales está la ausencia de bases de datos estándar contra las cuales validar los algoritmos.

El uso de señales normales o patológicas depende de la magnitud de las perturbaciones ante las cuales se desea evaluar el detector. De acuerdo a un sistema de clasificación ampliamente aceptado (Titze 1995, Vieira *et al* 2002) las señales normales son de Tipo 1, cuasi-periódicas, con las energías de las perturbaciones como mínimo por debajo de un orden de magnitud con respecto a la de la componente periódica. Las voces patológicas van desde Tipo 2 (energías de las perturbaciones cercanas a la de la periódica) hasta Tipo 3, sin estructura periódica aparente (también denominadas señales caóticas). En los estudios reportados en la literatura no se emplean perturbaciones superiores al 15% de la energía de la periódica.

Por último, la selección de vocales aisladas o habla fluida depende totalmente del propósito. Para medir perturbaciones de la voz se deben emplear vocales sostenidas, donde cualquier desviación de la periodicidad pueda considerarse perturbación, sin la influencia de los cambios en la configuración del tracto vocal que ocurren en el habla fluida. Para aplicaciones de codificación y síntesis es preferible el uso de grabaciones de habla fluida.

Atendiendo a estas consideraciones, en este trabajo se evalúan las medidas con señales sintéticas, de una vocal “a” sostenida, con niveles patológicos de perturbación.

2.2.2 Síntesis implementada:

Para sintetizar las señales empleadas se usó el MatLab 7.1, donde se diseñó un algoritmo que nos permite sintetizar ambas señales. El programa genera la forma de onda glotal de acuerdo al modelo utilizado la cual es convolucionada con la función del tracto vocal que es fijo para cada modelo y como resultado se obtienen las señales sintéticas, produciéndose una vocal sostenida “a”.

El modelo para la síntesis de la forma de onda glotal que se utilizó fue el modelo polinomial de Rosemberg “B” (Rosemberg 1971) que responde a la ecuación 2 (donde α es la amplitud y Tp es la duración de la fase de apertura) ya que en estudios realizados por (Rosemberg 1971, Veldhuis 1998, O’Leidhin & Murphy 2003) utilizando criterios de naturalidad de la voz sintetizada y de minimización del error medio cuadrático con respecto a la señal residual del filtrado inverso (Fant *et al* 1985, O’Leidhin & Murphy 2003) el

modelo de Rosemberg dio buenos resultados debido a que las señales sintetizadas se percibieron como más naturales que las de las grabaciones, mientras que el modelo de Liljencrants-Fant (Fant *et al* 1985), con el que se reportan los menores valores de error cuadrático medio en su ajuste a las señales residuales, ha producido “resultados decepcionantes” (O’Leidhin & Murphy 2003) al sintetizar nuevamente la voz empleando la $g(t)$ estimada.

$$F_b = \alpha \left[3 \left(\frac{t}{T_p} \right)^2 - 2 \left(\frac{t}{T_p} \right)^3 \right] \quad (2)$$

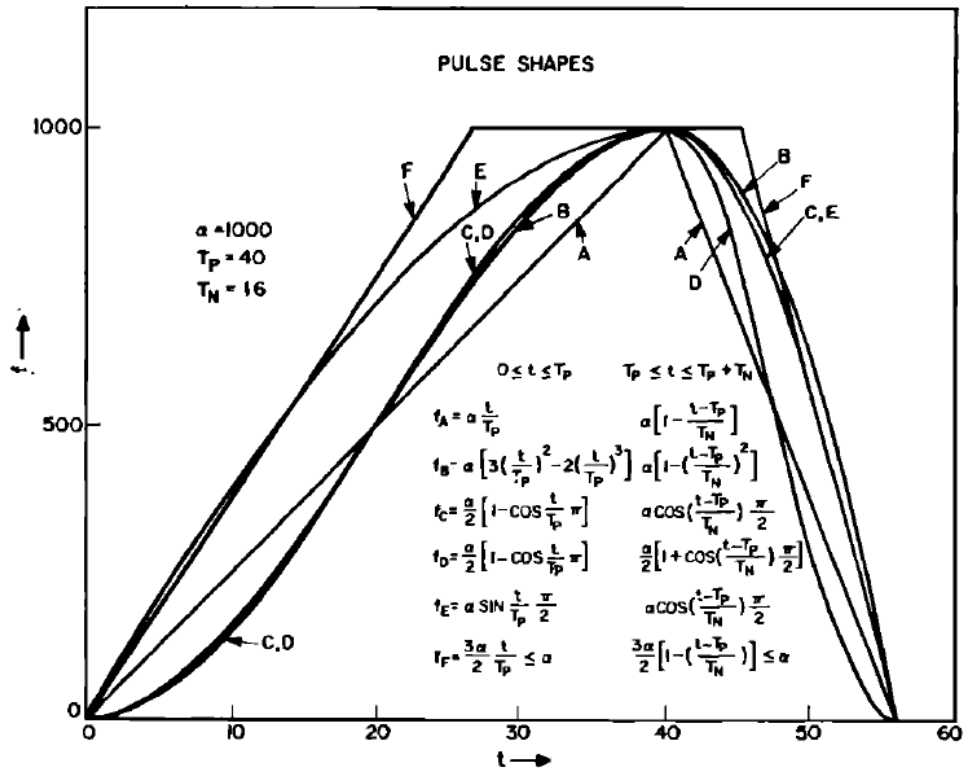


Figura 2.1. Formas de onda evaluadas por (Rosemberg 1971). En este trabajo se emplea la B.

Se utilizan dos funciones para generar los pulsos, para el modelo E/C la función acepta como parámetros el nivel (n) de *jitter*, la duración de cada vocal en segundos ($TmScs$), período fundamental ($T0$), Frecuencia de muestreo (Fs) y un vector con los instantes de tiempo donde se coloca la onda glotal. Esta función devuelve un vector con la señal resultante de convolucionar el vector de pulsos de Rosemberg con *jitter* con la función del tracto vocal $h(t)$. Para el modelo S/A la función necesita los mismos parámetros que para la

de E/C sin el parámetro ``n`` devolviendo un vector con el resultado de la convolución del tracto con el vector de ondas glotales con un ruido que produce 30 dB de SNR en la señal.

En la síntesis se empleó una frecuencia de muestreo de 22050 Hz, y una F_0 promedio de 150 Hz. La duración seleccionada de las señales fue de 2 segundos, por lo que cada una contiene como promedio 300 pulsos glotales. La función de transferencia del tracto vocal empleada fue la misma que en (Rabiner *et al.*, 1976, Parsa & Jamieson, 1999 y Ferrer, 2006).

Cada medida de perturbación se evalúa para 100 realizaciones de señales en cada nivel de *jitter* (3.4%, 6.8%, 10.2%, 13.6%, 17%, 20.4%, 23.8%).

2.3 Medición de *jitter*.

Las diferentes medidas de *jitter* se basan en cálculos efectuados sobre el contorno de T_0 , por lo que la exactitud de las mismas depende de la efectividad en la detección de dicho contorno. De esta manera, la evaluación de la precisión de las medidas de *jitter* se convierte en la evaluación de la precisión del PDA.

La comparación de los rendimientos de los PDA's se efectúa midiendo la similitud entre los valores de T_0 reales y los detectados, usados como referencia. Varias medidas representativas de las reportadas en la literatura se muestran en la Tabla 2.1.

Tabla 2.1. Algunas de las medidas de rendimiento reportadas en la comparación de los PDA's.

Medidas		Descripción	Ref.
1	V/U (sonoro a sordo)	% de segmentos sonoros clasificados como sordos, y sordos como sonoros. Orientadas principalmente a PDA's por segmentos	1,2
2	U/V (sordo a sonoro)		
3	PI (inserción de pulsos)	% de inserciones por detección submúltiplos de T_0 y supresiones por la detección de múltiplos de T_0 . Evitables con correcto margen de búsqueda de T_0 .	2
4	PD (supresión de pulsos)		
5	GE (errores gruesos)	Errores mayores que un determinado umbral. (%).	1*, 2
6	VFE (var. de errores finos)	Varianza de errores menores que un umbral.	1
7	α	Medidas de perturbación del período (<i>jitter</i>)	3,4,5
8	β	Diferencias relativas promedio con los T_0 reales	5,6

Los nombres dados a las medidas no necesariamente coinciden con los empleados en las referencias, y puede haber ligeras diferencias en las fórmulas empleadas en cada referencia para alguna medida. Las referencias se han denotado como: [1] (Rabiner *et al.*, 1976), [2] (Veprek and Scordilis, 2002), [3] (Deem *et al.*, 1989), [4] (Titze and Liang, 1993), [5] (Parsa and Jamieson, 1999), [6] (Medan *et al.*, 1991). * En [1], PI y PD se incluyen en GE. En [5], β es un híbrido con VFE, empleando el cuadrado de los errores.

Las primeras cuatro medidas reportan errores de una magnitud tal que no involucran la precisión del PDA, determinando errores en la clasificación sordo/sonoro y de múltiplos y submúltiplos. Las cuatro medidas (V/U, U/V, PI y PD) se expresan como porcentaje del número total de pulsos.

La precisión de los PDA es abarcada por las otras cuatro medidas (GE, VFE, α y β), entendiendo que precisión es la exactitud en seguir los valores reales de T_0 dado que se haya determinado que el segmento sea sonoro y que no se esté cometiendo error de múltiplos o submúltiplos. La exclusión de los errores gruesos (GE) de la VFE tiene el objetivo de reducir la influencia de los mismos, ya que la magnitud de estos es muy variable y afectaría la relevancia de VFE. Ambas medidas son complementarias, representando VFE la precisión en el seguimiento de la curva de T_0 y GE la frecuencia con que se pierde dicho seguimiento. Los umbrales empleados para separar errores gruesos y finos se fijan de forma heurística.

Las múltiples medidas de *jitter* se han agrupado en la Tabla 2.1 bajo el término α (Ferrer, 2005) y una de las expresiones más comúnmente empleadas se muestra en (3), siendo $T_d(n)$ el período (duración) del enésimo pulso tal como lo detecta el PDA, y N el número total de pulsos detectados.

$$\alpha = \frac{1}{N-1} \sum_{n=1}^{N-1} \frac{|T_d(n+1) - T_d(n)|}{0.5 * (T_d(n+1) + T_d(n))} * 100 \quad (3)$$

La razón de emplear medidas de *jitter* como esta en la comparación de rendimientos de PDA's consiste en que dada una curva de T_0 de referencia (T_r) la diferencia $\Delta = |\alpha_d - \alpha_r|$ entre los valores detectados y reales de α (calculados empleando las curvas T_d y T_r respectivamente) sirve de indicador de fallo del PDA. A pesar de ser muy usadas en la

comparación de PDA's (ver Tabla 2.1) las medidas de *jitter* no son apropiadas para estos fines. Existen infinitos contornos de T_d que resultan en el mismo valor de α , y un contorno de T_d completamente erróneo puede conducir por casualidad a un mismo valor de α_d que α_r , el *jitter* del contorno real. Esto hace que las otras medidas (GE, VFE y β) sean más apropiadas para la comparación de los PDA's, pues emplean T_d y T_r en sus expresiones de cálculo.

Algunas medidas muy relacionadas con VFE pero preferidas por varias razones (la normalización de los errores, evitar la raíz cuadrada y la influencia no lineal de los *outliers*, por ejemplo) se agrupan en la Tabla 2.1 bajo el nombre de β y una fórmula típica se muestra en (4) (Ferrer, 2005)

$$\beta = \frac{1}{N} \sum_{n=1}^N \frac{|T_d(n) - T_r(n)|}{T_r(n)} * 100 \quad (4)$$

En el caso de estas medidas que trabajan tanto con T_d como con T_r existen varios aspectos que requieren de un análisis más detallado para evaluar cuán apropiadas resultan. Dicho análisis se reporta en (Ferrer, 2010), donde se propone un algoritmo de alineamiento temporal de contornos que produce correcciones a las medidas de α y β , así como a los valores de GE, PI y PD.

En este trabajo se emplearon las medidas de β_c y GE para evaluar el rendimiento de los PDA's.

2.3.1 PDA's evaluados

Debido a la gran cantidad de métodos de detección de período fundamental existentes (la más amplia compilación existente se halla en Hess, 1983 con una actualización posterior en Hess, 1992), la tarea de comparar los rendimientos de los mismos, en aras de establecer sus ventajas y desventajas, ha sido abordada en múltiples trabajos (Rabiner *et al.*, 1976, Veprek and Scordilis, 2002, Deem *et al.*, 1989, Titze and Liang, 1993, Parsa and Jamieson, 1999, Medan *et al.*, 1991, Ma *et al.* 1994, de Cheveigné & Kawahara 2003).

Los PDA's pueden clasificarse en cuanto al resultado que brinda en detectores por segmento y detectores ciclo a ciclo. Los detectores por segmento obtienen un valor de

período fundamental para todo el segmento de análisis. Dicho segmento puede contener varios pulsos glotales, por lo que el valor de T_0 obtenido no representa exactamente la separación entre ninguno de los pares de pulsos contiguos, sino un valor que aproxima de manera general a esas separaciones. Los detectores ciclo a ciclo obtienen las posiciones (o separaciones) específicas de cada pulso. En las mediciones de alteraciones de la periodicidad de los pulsos es imprescindible el uso de detectores ciclo a ciclo.

Múltiples métodos han sido propuestos para extraer las posiciones y/o separaciones entre pulsos individuales (Milenkovick, 1987; Medan *et al.*, 1991; Kadambe & Boudreaux-Bartels, 1992) y existe un consenso general en cuanto a que las técnicas que buscan similitud de formas de onda (WM, del inglés *Waveform Matching*) son las más exactas y confiables en esta tarea (Titze and Liang, 1993; Titze, 1995; Parsa and Jamieson, 1999; Veprek and Scordilis, 2002). Los WMPDA's más abordados en la literatura son los reportados por Milenkovick (Milenkovick, 1987) y Medan (Medan *et al.*, 1991). Varias fuentes los reportan como los PDA's de mejor rendimiento en la medición de perturbaciones de la voz (Titze and Liang, 1993; Titze, 1994; Parsa and Jamieson, 1999; Veprek and Scordilis, 2002).

Por este motivo se seleccionaron, para la evaluación de la influencia del *jitter* en la estimación de los contornos de T_0 , los algoritmos de (Milenkovick, 1987) y (Medan *et al.*, 1991), así como dos variantes posteriores que han producido buenos resultados (Ferrer, 2006).

Método de Milenkovick:

Este método se deriva de un modelo de la forma de onda $s(t)$ dado por su descomposición en una parte periódica $p(t)$ con periodicidad en $t=T_i$ y una aleatoria (ruido) $e(t)$:

$$\begin{aligned} s(t) &= p(t) + e(t) \\ p(t) &= K_i p(t - T_i) \end{aligned} \tag{5}$$

donde K_i representa el cambio de amplitud en el período i -ésimo (Milenkovick, 1987). La inclusión de K_i en el modelo mejora el de (Wise & di Caprio 1976) al tomar en consideración el *shimmer*.

Asumiendo que $e(t)$ es estacionario en la media y la varianza, con media cero y con $e(t)$ no correlacionado con $e(t-T_i)$ se llega a una expresión para el error cuadrático medio entre dos segmentos de señal separados en T_i :

$$MSE = E \{e(t)^2\} = \frac{1}{1+K^2} (S * S' - 2K * S * S'_{T_i} + K^2 S_{T_i} * S'_{T_i}) \quad (6)$$

donde S y S_{T_i} son las versiones muestreadas de los dos segmentos, descritas por (7), y S' y S'_{T_i} sus traspuestas.

$$S = [s(nT_s)] \quad (7)$$

$$S_{T_i} = [s(nT_s - T_i)] \quad n_0 - N < n \leq n_0$$

En (7) T_s es el período de muestreo, N la longitud de los segmentos en muestras y n_0 la muestra de inicio del primer segmento. El valor de N debe ser menor que T_i en muestras, por limitaciones que impone el modelo (Milenkovick, 1987). Se asume que MSE es mínimo para los valores de K_i y T_i que representan las verdaderas amplitudes y separaciones de los pulsos, por lo que estos pueden en principio determinarse minimizando MSE. De la minimización de MSE con respecto a K_i resulta:

$$K_i = R + \sqrt{1 + R^2} \quad R = \frac{(S * S' - S_{T_i} * S'_{T_i})}{2S * S'_{T_i}} \quad (8)$$

La obtención de K_i y T_i se efectúa en dos pasos, primero se obtiene un valor de K_i según (8) empleando un estimado de T_i , y luego, con el valor de K_i obtenido se efectúa una búsqueda del mínimo valor de MSE para valores de T_i alrededor de un 10% del estimado. El valor de T_i que minimiza MSE en ese intervalo es considerado el correcto (óptimo).

El procedimiento se repite posicionando n_0 en $n_0 + T_i$, para efectuar la búsqueda del siguiente valor de T_i y así obtener el contorno de T_d .

Método de Medán

El método de Medán (Medan *et al.*, 1991) parte de definir los segmentos de análisis x_τ e y_τ adyacentes y no solapados para un instante de referencia t_0 como:

$$\begin{aligned}
 x_\tau(t, t_0) &= s(t)\omega_\tau(t - t_0) & \omega_\tau(t) &= \begin{cases} 1 & 0 < t < \tau \\ 0 & t < 0 \text{ y } t > \tau \end{cases} \\
 y_\tau(t, t_0) &= s(t + \tau)\omega_\tau(t - t_0)
 \end{aligned} \tag{9}$$

con $s(t)$ la señal de voz, de esta forma los segmentos tienen en realidad longitud τ . Considerando el caso de dos pulsos contiguos de duración $\tau = T_i$, se asume un modelo de similitud donde:

$$x_{T_i}(t, t_0) = a_i y_{T_i}(t, t_0) + e(t, t_0) \tag{10}$$

siendo a_i un factor positivo de amplitud que representa el *shimmer* y $e(t, t_0)$ un término de error que agrupa el resto de las diferencias entre las formas de onda de los dos pulsos. Se define el período fundamental T_0 como el valor de τ que minimiza una función de costo dada por:

$$J(\tau) = \frac{\int_0^{\tau} [x_\tau(t, t_0) - a_i y_\tau(t, t_0)]^2 dt}{\int_0^{\tau} [x_\tau(t, t_0)]^2 dt} \tag{11}$$

donde el numerador es representativo de la energía del error en el intervalo y el denominador se introduce para normalizar el efecto de la longitud variable del intervalo.

En (Medan *et al.*, 1991) se demuestra fácilmente que la maximización de $J(\tau)$ equivale a maximizar el coeficiente de cross-correlación entre los segmentos x e y :

$$\rho_\tau = \frac{(x, y)_\tau}{\sqrt{(x, x)_\tau (y, y)_\tau}} \tag{12}$$

donde entre paréntesis se expresa la operación de producto vectorial entre las versiones muestreadas de los mismos. Se destaca una diferencia fundamental con las funciones de correlación y es la dependencia no sólo del desplazamiento sino de la duración de los segmentos de análisis. El rango de búsqueda también se limita a un intervalo (25%) alrededor de un estimado del valor de período fundamental.

Limitaciones ante *jitter*. Variantes 1 y 2

En (Ferrer, 2006) se desarrollaron dos variantes de PDA's para superar algunas limitaciones observadas en los WMPDA's de Milenkovich y Medan. Estas limitaciones, por estar estrechamente vinculadas al *jitter*, se repiten a continuación.

En el método de Milenkovich se obtiene el valor de K_i a partir de un estimado de T_i , pero un estimado de T_i individual para cada pulso no resulta práctico (es precisamente lo que se desea obtener) y un estimado “global” puede no ser adecuado en voces patológicas. Con valores de *jitter* de hasta un 25% del T_0 estimado el valor de K_i obtenido puede ser excesivamente alterado, al obtenerse para posiciones muy desplazadas con respecto a la correcta. Un valor de K_i muy desviado del real puede a su vez alterar excesivamente los resultados de la búsqueda subsiguiente de T_i . En (Titze and Liang, 1993) se reporta que el asumir $K_i=1$ en el método de Milenkovich provocó que esa variante sufriera un deterioro considerable de su rendimiento para señales con *shimmer* por encima del 5%, lo que es equivalente a haber cometido errores del 5% en la estimación de K_i al asumirlo unitario. El método de Milenkovich es entonces propenso a errar en presencia de *jitter*, al fallar la estimación de K_i , mientras el de Titze permuta esta sensibilidad hacia el *shimmer*, al asumir K_i constante y unitaria, ignorando las variaciones de amplitud. Ambos métodos deben fallar con frecuencia ante las perturbaciones que se pueden encontrar en voces patológicas.

El método de Medan es inmune a las diferencias de amplitud, pues en la derivación de (12) se emplea directamente el valor de a_i óptimo para cualquier τ . La sensibilidad del mismo está dada ante el *jitter*, por asumir el modelo que los dos pulsos son de igual duración. El efecto del *jitter* en esta suposición se puede explicar gráficamente. Sean N_1 la duración del primer pulso, igual a la separación del mismo al segundo, y N_2 la duración del segundo igual a su separación del tercero. Existen dos casos en presencia de *jitter*: a) $N_1 < N_2$ ó b) $N_2 < N_1$. En la Figura 2.3 se muestra una ilustración de ambos casos.

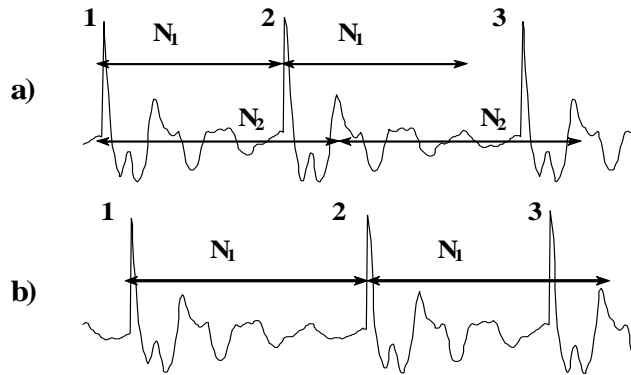


Figura. 2.2 Ejemplo del efecto del *jitter* en el método de Medan. a) $N_1 < N_2$, b) $N_2 < N_1$.

Cuando $N_1 < N_2$ la mayor similitud (menor J_τ) se halla para $\tau = N_1$, pero otra alta similitud se halla para $\tau = N_2$ debido al parecido de los picos 2 y 3 al final de los segmentos de análisis. En el caso $N_2 < N_1$ la presencia del pico 3 dentro del segundo segmento de análisis (y_τ) cuando $\tau = N_1$ introduce una diferencia considerable que puede hacer que J_τ no sea mínima para ese valor de τ (ver numerador de (11)). Los efectos de estas interferencias deben ser mayores al combinarse con *shimmer* y ruido, al bajar el valor de los picos correctos de autocorrelación por el ruido y poder ser mayores las correlaciones en las posiciones incorrectas por las relaciones de amplitud entre los picos involucrados.

2.4 Medición de *shimmer*.

En el caso del *shimmer* se consideraron también cuatro métodos, los cuales se evalúan igualmente para cada nivel de perturbación de *jitter*. Los métodos empleados son:

1. Pico máximo de la señal en el pulso.
2. RMS de la señal en el pulso.
3. Relación de amplitud estimada por el PDA de Milenkovick (K_i).
4. Relación de amplitud estimada por el PDA de Medan (a_i).

2.5 Conclusiones del capítulo.

Se programaron experimentos para evaluar la influencia de dos modelos de variación del pulso glotal en función del *jitter* en la efectividad de medidas de perturbación de la periodicidad. Las medidas consideradas fueron el mismo *jitter*, a través de la evaluación de los PDAs que se requieren para su cálculo, y un conjunto de estimadores de *shimmer*. Los experimentos se efectuaron con señales de voz sintéticas, con niveles de *jitter* de hasta un ~25%, y fueron programados en MatLab.

CAPÍTULO 3. RESULTADOS

En este capítulo se muestran los resultados de los experimentos realizados para validar las contribuciones propuestas en este trabajo y se comentan las implicaciones que pueden extraerse de estos resultados.

3.1 PDA's

Los resultados de los experimentos “con *Jitter*” se muestran en la Figura 3.1. Los métodos se han denominado “Milenk” (Milenkovick, 1987), “Medan” (Medan *et al.*, 1991), “V1” y “V2” (Ferrer, 2006). Para cada PDA se obtuvieron las medidas de rendimiento β_c y GE en los experimentos con señales simuladas.

A partir de estos resultados se puede apreciar que “Milenk” tiene los peores resultados, no es confiable en presencia de *jitter*, El método “Medan” no presenta dificultades al tratar con señales con *jitter*, aparentemente debido a la ausencia de las otras perturbaciones. Sin embargo, estos métodos son superados por Variante 1 y Variante 2.

La Figura 3.1 muestra una gráfica con los valores de β_c para los métodos estudiados.

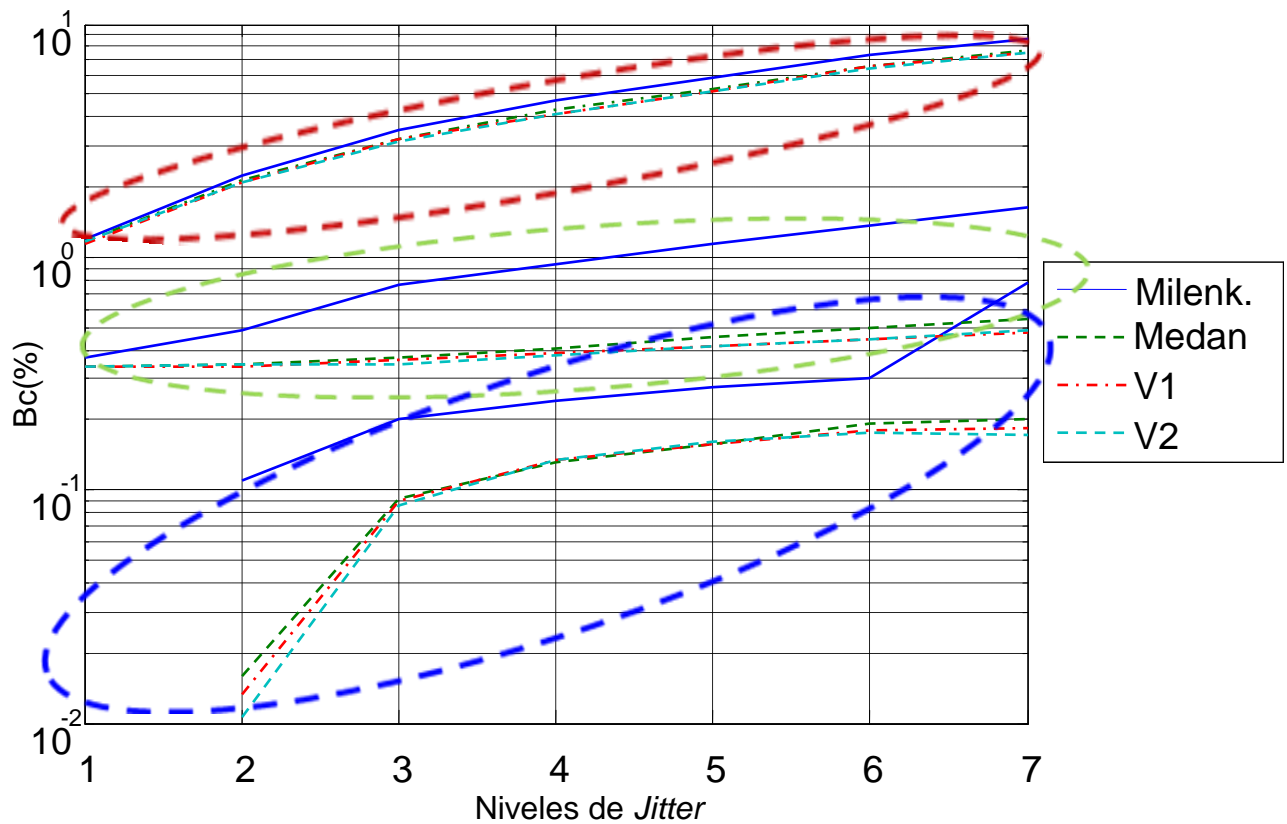


Figura. 3.1 Resultados de β_c en %. Arriba (elipse roja), modelo E/C siguiendo las aperturas. Al centro (verde), modelo de E/C siguiendo cierres Abajo (azul), modelo S/A (no hay diferencia entre aperturas y cierres).

Puede apreciarse que el error cometido (β_c) es menor ante el modelo S/A, y mayor en el caso del E/C siguiendo las aperturas. Los resultados en cuanto a Errores Gruesos (con umbral del 3% de $T0$), promediados para 100 corridas aleatorias de los experimentos, pueden observarse en la Tabla 3.1.

Tabla.3.1 Resultado Promedio de los Errores Gruesos.

		1	2	3	4	5	6	7
Sep/Acerc	Mil	0	0.55	0.73	0.57	0.62	0.66	2.88
	Med	0	0	0	0	0	0	0
	V1	0	0	0	0	0	0	0
	V2	0	0	0	0	0	0	0
Exp/Cont Cierres	Mil	0	1.03	3.50	4.89	6.09	6.76	7.36
	Med	0	0	0.01	0.02	0.08	0.12	0.21
	V1	0	0	0	0	0	0	0.02
	V2	0	0	0	0	0	0	0
Exp/Cont Aperturas	Mil	1.61	39.96	66.52	83.59	96.08	104.04	109.31
	Med	1.24	40.39	67.04	84.30	95.63	103.89	108.99
	V1	1.00	38.52	66.18	83.37	94.82	103.01	108.65
	V2	1.04	38.68	65.28	82.81	94.57	102.82	108.38

Puede apreciarse nuevamente la superioridad de los métodos desarrollados por los autores (Ferrer, 2006), la mayor dificultad que representa el modelo de E/C, y la posibilidad de obtener resultados aceptables dentro de este último si se siguen los instantes de cierre.

3.2 *Shimmer.*

El error de estimación promedio para 100 corridas del experimento se muestra en la Figura 3.2, para el caso del modelo E/C de *jitter*. Para el caso del modelo de S/A existen resultados reportados en la literatura (Ferrer *et al*, 2007) que se muestran en la Figura 3.3.

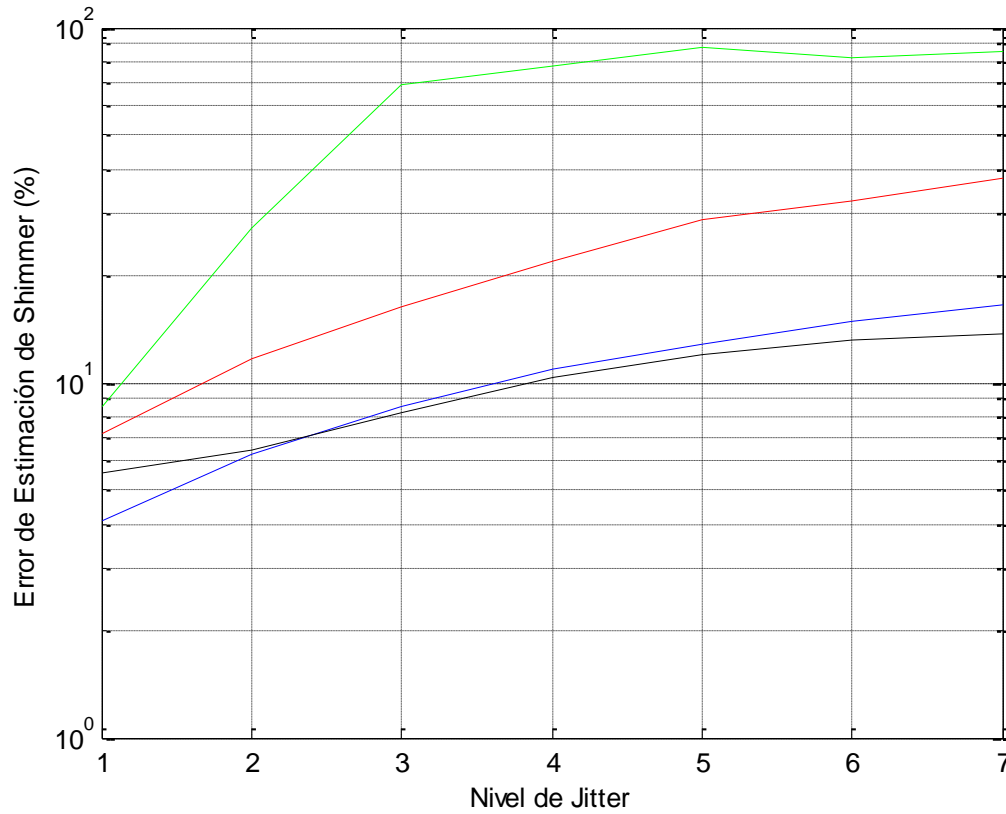


Figura. 3.2 Error promedio de estimación de *Shimmer*, en %. Rojo: Pico máximo, Azul: RMS del pulso, Verde: Milenkovich, Negro: Medán.

Aquí debe señalarse que las curvas de los métodos basados en Picos (roja) y RMS (azul) se obtuvieron para los pulsos correctamente localizados a partir del conocimiento de las fronteras de los pulsos en las señales sintéticas, mientras que los de Milenkovich y Medán (verde y negro, respectivamente) parten de los pulsos detectados por el PDA del mismo nombre. De esta manera, la estimación de Milenkovich y Medán se ve afectada además por los posibles errores de detección de las posiciones de los pulsos que cometen los PDA's. Es de destacar la enorme degradación que sufre el método de Milenkovich como estimador de *shimmer*, (en correspondencia con su ineffectividad como PDA) y la robustez que muestra el método de Medán, con los mejores resultados a partir del 2^{do} nivel de perturbación.

Para el caso del modelo de S/A, la figura siguiente muestra los resultados obtenidos en (Ferrer *et al*, 2007) para los detectores de Pico, RMS y Milenkovich.

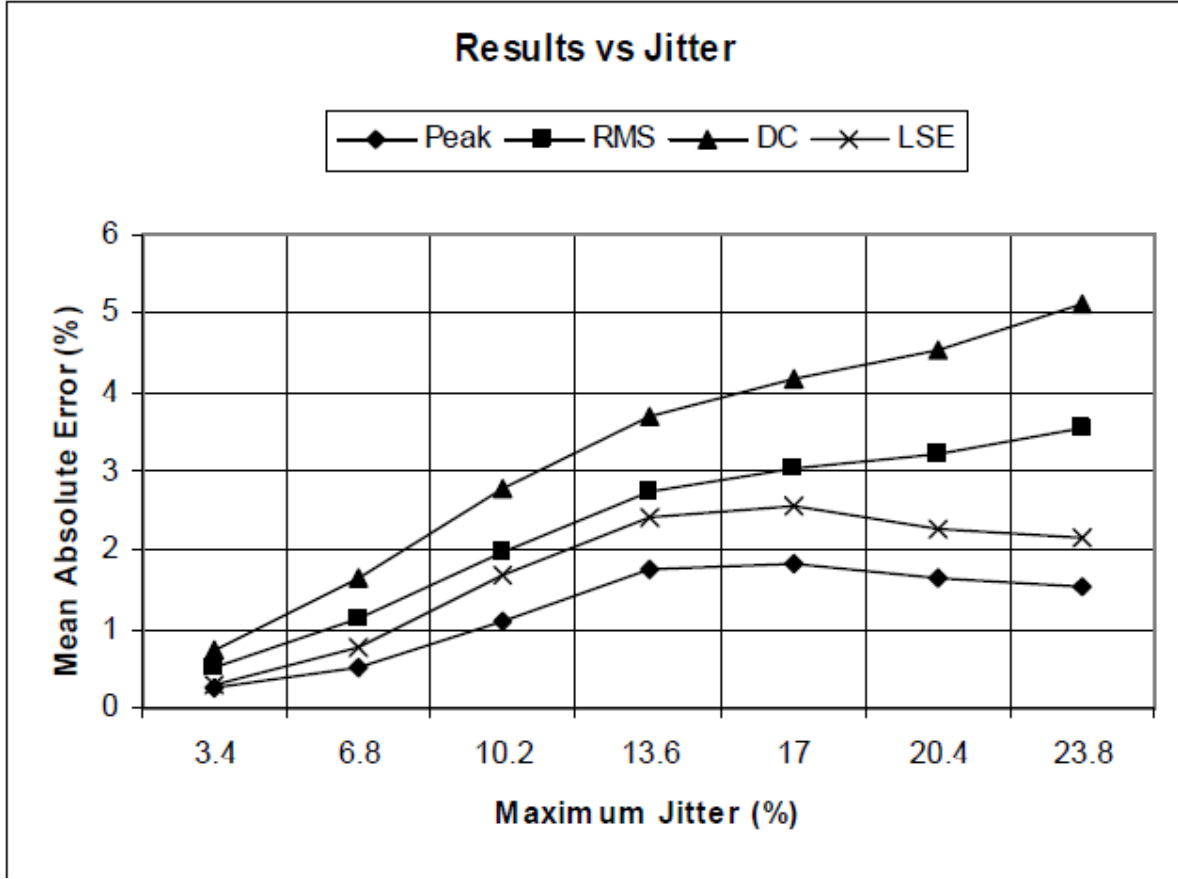


Figura. 3.3 Error promedio de estimación de *Shimmer* para tres de los métodos evaluados en este trabajo, empleando el modelo S/A de *jitter*. Picos (♦), RMS (■), y Milenkovich (x). Tomado de (Ferrer *et al*, 2007) con permiso de los autores.

En la estimación de *shimmer* ante *jitter* que sigue el modelo S/A, existen resultados reportados en (Ferrer *et al*, 2007) para tres de los métodos comparados en este trabajo, y se muestran en la Figura 3.3. Aquí es notable el buen comportamiento de todos los métodos, lo cual corrobora que ante el modelo S/A los resultados de las mediciones tienden a ser mejores.

3.3 Conclusiones del capítulo

Cuando se introduce *jitter* según el modelo E/C los PDA's se ven mucho más afectados que ante el modelo S/A. Sin embargo, la precisión en el seguimiento de los instantes de cierre en el modelo de E/C por parte de los PDA's se acerca al logrado en el modelo de S/A.

Entre los PDA's, las variantes desarrolladas en el CEETI son las de mejores resultados en todos los casos.

Con respecto a los estimadores de perturbaciones de amplitud (*shimmer*), se ratifica la mayor complejidad del modelo E/C, donde el mejor estimador va de un 5 a un 15% de error, mientras en el modelo S/A el error siempre queda por debajo de un 5%.

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

Jitter: El modelo E/C es mucho más difícil de seguir por los PDA's, aunque la precisión en el seguimiento de los instantes de cierre en el modelo de E/C por parte de los PDA's se acerca al logrado en el modelo de S/A. Entre los PDA's, las variantes desarrolladas en el CEETI son las de mejores resultados en todos los casos.

Shimmer: Con respecto a los estimadores de perturbaciones de amplitud, se ratifica la mayor complejidad del modelo E/C, donde el mejor estimador va de un 5 a un 15% de error, mientras en el modelo S/A el error siempre queda por debajo de un 5%.

En general, los resultados de los estimadores difieren significativamente en dependencia del modelo de *Jitter* con que se genera la señal glotal. Estudios posteriores deben definir qué modelo se corresponde en mayor medida con las señales reales, para una mejor evaluación de la efectividad de los algoritmos de estimación.

Recomendaciones

Se recomienda extender estos análisis de la influencia del modelo de deformación del pulso causado por el *jitter* a las mediciones de ruido, para evaluar las diferencias en la estimación de SNR que provoque cada modelo.

Debe explorarse además la posibilidad de estimar, mediante modelos matemáticos dada la ausencia de señales EGG, el modelo que siguen las señales de excitación en voces reales.

REFERENCIAS BIBLIOGRÁFICAS

- COX, N. B., ITO, M. R., & MORRISON, M. D. (1989a) "Technical considerations in computation of harmonics to noise ratios for sustained vowels" *JS&HR*. 32, 203-218.
- COX, N. B., ITO, M. R., & MORRISON, M. D. (1989b) "Data labeling and sampling effects in harmonics-to-noise ratios" *JASA*. 85, 2165-217.
- CRANEN, B. & DE JONG, F. (2000) "Laryngostroboscopy" en R. D. Kent & M. J. Ball (eds) *Voice Quality Measurement*. Singular. San Diego
- DE CHEVEIGNÉ, A. & KAWAHARA, H. (2001) Comparative evaluation of F0 estimation algorithms. *EuroSpeech 2001*. Scandinavia
- DEEM, J.F., MANNING, W.H., KNACK, J.V. & MATESICH, J.S. (1989) "The automatic extraction of pitch perturbation using microcomputers: Some methodological considerations" *Journal of Speech and Hearing Research*, 32, 689-697
- FANT, G. (1960) *Acoustic Theory of Speech Production*. Mouton. The Hague.
- FANT, G., LILJENCANTS, J. & LIN, Q. (1985) A four parameter model of the glottal flow. *Speech Transmission Laboratory Quarterly Progress Status Report No 4*. Suecia, Royal Institute of Technology
- FERRER, C. A. (2005) *Cuantificación de parámetros de la voz para aplicaciones médicas. Centro de Estudios de Electronica y Tecnologias de la Informacion*. Santa Clara, Universidad Central "Marta Abreu" de las Villas.
- FERRER, C. A. (2006) *Cuantificación de parámetros de la voz para aplicaciones médicas. Tesis de PhD*.

- FERRER, C. A., HERNANDEZ-DIAZ, M. E. & GONZALEZ, E. (2007) Using Waveform-Matching techniques in the measurement of *shimmer* in voiced signals. *INTERSPEECH 2007*. Amherst.
- FERRER, C., TORRES, D., & HERNÁNDEZ-DÍAZ, M. E. (2010) "Using dynamic time warping of T0 contours in the evaluation of cycle-to-cycle Pitch Detection Algorithms", *Pattern Recognition Letters*. ISSN 0167-8655. Vol. 31 (6), pp 517-522.
- FOURCIN, A. (2000) "Voice Quality and Electrolaryngography" en R. D. Kent & M. J. Ball (eds) *Voice Quality Measurement*. Singular. San Diego.
- HERTZ, D. (1986) "Time delay estimation by combining efficient algorithms and generalized cross-correlation methods". *IEEE Trans. Acoust., Speech, Signal Processing*. ASSP 34 pp. 1-7.
- HESS, W. (1983) *Pitch Determination of Speech Signal: Algorithms and Devices*. Springer Verlag. Berlin.
- HESS, W.J. (1992) "Pitch and Voicing Determination" en S. Furui and M.M. Sondhi (eds) *Advances in Speech Signal Processing*. Marcel Dekker, New York.
- KADAMBE, S., BOUDREAUX-BARTELS, G. F., (1992). "Application of the wavelet transform for pitch detection of speech signals". *IEEE Trans. on Information Theory*. 38 (2). 917-924
- KLATT D.H. (1980). "Software for a cascade/parallel formant synthesizer" *J. Acoust Soc. Am.*, Vol. 67, No. 3. pp. 979-990
- KLATT D.H. & KLATT L.C. (1990). "Analysis, synthesis and perception of voice quality variations among female and male talkers" *JASA*, 87. pp. 820-857
- LAVER, J. (1991) *The gift of speech*, Edinburgh, UK.
- LILJENCRAFTS, J. (1991) *Analogies in the production of speech and music*, Macmillan Academic and Professional Ltd.
- MA, C., KAMP, Y. & WILLEMS, L. F. (1994) "A Frobenius norm approach to glottal closure detection from the speech signal" 2, 258-265

- MARASEK, K. (1997) "EGG & Voice Quality" disponible en <http://www.ims.uni-stuttgart.de/phonetik/EGG/frmst1.htm>.
- MEDAN, Y. YAIR, E. & CHAZAN, D. (1991) "Super resolution pitch determination of the speech signals". IEEE Transactions on Signal Processing. 39. pp 40-48.
- MILENKOVICK, P. (1987) "Least mean squares measures of voice perturbation" Journal of Speech & Hearing Research. 30. 529-538
- MURPHY, P. J. (1999) "Perturbation-free measurement of the harmonics-to-noise ratio in voice signals using pitch synchronous harmonic analysis" JASA. 105, 2866-2881
- O'LEIDHIN, E. & MURPHY, P. J. (2003) "Preliminary Glottal Source modeling for pathologic voices", *Proc. Maveba Conf.*, Florence, pp. 237-240
- PARSA, V & JAMIESON, D. G. (1999) "A comparison of high precision Fo extraction algorithms for sustained vowels". JSLHR. Vol. 42, pp. 112-126.
- RABINER, L.R., CHENG, M.J., ROSEMBERG, A.E., MCGONEGAL, C.A., (1976). "A comparative study of several pitch-detection algorithms" IEEE Trans. Acoust. Speech Signal Process. ASSP. 24. pp 399-417
- ROSEMBERG, A. E. (1971) Effect of glottal pulse shape on the quality of natural vowels. Journal of the Acoustical Society of America, 84, 583-588.
- SCHOENTGEN, J. (2003) "Spectral models of additive and modulation noise in speech and phonatory excitation signals". Journal of the Acoustical Society of America. Vol. 113 (1) pp. 553-562.
- STEINECKE, I. & HERZEL, H. (1995) Bifurcations in an asymmetric vocal fold model. Journal of the Acoustical Society of America, 97, 1874-1884.
- STEVENS K. N. (1997) "Articulatory-Acoustic-Auditory Relationship" en Hardcastle W. J., Laver, J. (eds.) *The Handbook of Phonetic Sciences*, Blackwell Publishers Ltd., Oxford.
- TITZE, I. R. (1980) Comments on the myoelastic-aerodynamic theory of phonation. Journal of Speech and hearing Research, 23, 495-510.
- TITZE I.R. (1994), *Principles of Voice Production*, Prentice Hall, Englewood Cliffs.

- TITZE, I. R. (1995). Summary Statement. *Workshop-on acoustic voice analysis*. National Center of Voice and Speech. Iowa
- TITZE, I.R. & LIANG, H. (1993). "Comparison of Fo extraction methods for high-precision voice perturbation measurements" *JSHR*, 36, pp. 1120-1133.
- VELDHUIS, R (1998) "A computationally efficient alternative for the Liljencrants-Fant model and its perceptual evaluation" *JASA*, 103. 566-571.
- VEPREK, P. & SCORDILIS. M. S. (2002) "Analysis, enhancement and evaluation of five pitch determination techniques" *Speech Communication*. 37. pp 249-270
- VIEIRA, M. N., MCINNES F. R. & JACK, M. A. (2002) "On the influence of laryngeal pathologies on acoustic and electroglottographic *jitter* measures" *JASA*. 111. 1045-1055.
- WISE, J. D., CAPRIO, J. R. & PARKS, T. W. (1976) "Maximum likelihood pitch estimation". *IEEE Trans. Acoust., Speech, Signal Processing*. ASSP 24 pp. 418-423.

ANEXOS**Anexo I PDA's****1. Variante 1**

```
function [Pts]= PtFa (y,Fm,T0Est,SR,PIni)
WL = T0Est-SR;
Pts = PIni;
while (length(y) > (PIni+T0Est+SR+WL+1) )
S=y(PIni:PIni+WL);
CC=[];
SD=y(PIni+T0Est:PIni+T0Est+WL);
for i=-SR:SR
SD=y(PIni+T0Est+1+i:PIni+T0Est+WL+i+1);
CC(SR+i+1)=sum(S.*SD)/sqrt(sum(SD.^2));
end
[a,PosM]=max(CC);
PP=PIni+T0Est-SR+PosM;
PIni=PP;
Pts = [Pts PIni];
end.
```

2. Variante 2

```
function [Pts]= PtFa (y,Fm,T0Est,SR,PIni)
WL = T0Est-SR;
Pts = PIni;
while (length(y) > (PIni+T0Est+SR+WL+1) )
S=y(PIni:PIni+WL);
CC=[];
SD=y(PIni+T0Est:PIni+T0Est+WL);
for i=-SR:SR
SD=y(PIni+T0Est+1+i:PIni+T0Est+WL+i+1);
CC(SR+i+1)=sum(S.*SD);
end
[a,PosM]=max(CC);
PP=PIni+T0Est-SR+PosM;
PIni=PP;
Pts = [Pts PIni];
end
```

3. Milencovick

```
function [Pts,K, K1]= PtFa (y,Fm,T0Est,SR,PIni)
WL = T0Est;
Pts = PIni;
K=[];
K1=[];
%v=0;
%pinta=0;
while ( length(y) > (PIni+T0Est+SR+WL+1) )
%v=v+1;
S=y(PIni:PIni+WL);
MSE=[];
SD=y(PIni+T0Est+1:PIni+T0Est+1+WL);
R=(sum(S.^2)/sum(S.*SD) - sum(SD.^2)/sum(S.*SD))/2;
Temp = R+sqrt(1+R^2);
RMS= sqrt(sum(SD.^2)/sqrt(sum(S.^2)) ;
K1=[K1 RMS];
K=[K Temp];
for i=-SR:SR
%if v == 1
```

```
% pinta=1;
%end
```

```
% SD=y(PIni+T0Est+1+i:PIni+T0Est+WL+i+1);
% if pinta==1
% figure(1)
% subplot(4,1,3)
% plot(R,'go')
% grid on
% subplot(4,1,4)
% plot(K,'ro')
% grid on
% hold off
% subplot(4,1,1)
% plot(S)
% hold on
% plot(SD,'r')
% hold off
% hold off
% plot(S)
% hold on
% plot(SD,'r')
% grid on
% pause
%MSE(SR+i+1)=1/(1+K^2)*(-2*K*sum(S.*SD) +
sum(S.^2) + K^2*sum(SD.^2));
%subplot(4,1,2)
%plot(MSE)
%grid on
%else
SD=y(PIni+T0Est+1+i:PIni+T0Est+WL+i+1);
MSE(SR+i+1)=1/(1+Temp^2)*(-2*Temp*sum(S.*SD) +
sum(S.^2) + Temp^2*sum(SD.^2));
%end
```

```
end
[a,PosM]=min(MSE);
PP=PIni+T0Est-SR+PosM;
PIni=PP;
Pts = [Pts PIni];
%figure(2)
%hold off
%plot(y)
%hold on
%plot(Pts,y(Pts),'o')
end
```

4. Medan

```
function [Pts,K2,K3]= PtFa(y,Fm,T0Est,SR,PIni)
Pts = PIni;
K2=[];
K3=[];
while (length(y)>(PIni+2*T0Est+2*SR+1))
CCC=[];
R=[];
RMS=[];
for i=-SR:SR
```

```
S=y(PIni:PIni+T0Est+i);
SD=y(PIni+T0Est+i+1:PIni+2*T0Est+2*i+1);
CCC(SR+i+1)=sum(S.*SD)/sqrt(sum(S.^2)*sum(SD.^2));
R(SR+i+1) = sum(S.*SD)/sqrt(sum(S.^2)) ;
RMS(SR+i+1)= sqrt(sum(SD.^2))/sqrt(sum(S.^2)) ;
end
[a,PosM]=max(CCC);
a=R(PosM);
b=RMS(PosM);
PP=PIni+T0Est-SR+PosM;
PIni=PP;
Pts = [Pts PIni];
K2=[K2 a];
K3=[K3 b];
end
```

Anexo II Evaluador

```
function
[A,B,GE,SR,SL,PI,PD,TdAl,Ch]=PerturbMeasures(Torig,Text
r);
%To find the perturbation measures A,B, and N.
%Inputs: The original T instants of glottal excitation and
%the detected ones.
%Outputs: Alpha (vector of real alpha, detected alpha, and
corrected alpha),
%          Beta, (vector of original and corrected)
%          GE: Number of Gross Error greater than 3% of
mean Textr,
%          before and after
%          SR, SL, corrected shifts to the right and left
%          PI, PD, Number of actually inserted pulses,
%          TdAl: Aligned contour,
%          and correction history (Ch).

% *****
% ****
%Important Notice: There is a nomenclature issue with the
pitch contours,
%being Torig and Textr the absolute positions, while Pr and Pd
are the
%relative positions

%A=[mean(abs(diff(diff(Torig))))/mean(diff(Torig))*100
mean(abs(diff(diff(Textr))))/mean(diff(Textr))*100];
if length(Textr)>1
Dlys=diff(Textr);
InDlys=diff(Torig);
minlength=min(length(Dlys),length(InDlys));

B=abs(Dlys(1:minlength)-InDlys(1:minlength));
InDlys=InDlys(1:minlength);%los comparables
possup=find(abs((InDlys-mean(InDlys))./InDlys) > 1);% Ver
outliers Problema!!! Debio ser como en Beta!!!
%possup=find(Temp == 0);
B(possup)=[];% Eliminarlos
InDlys(possup)=[];% Aqui tambien
ndx=2:length(InDlys);
A=mean(abs((InDlys(ndx)-InDlys(ndx-
1)))./(InDlys(ndx)+InDlys(ndx-1)))*100;
ndx=2:length(Dlys);
A=[A mean(abs((Dlys(ndx)-Dlys(ndx-
1)))./(Dlys(ndx)+Dlys(ndx-1)))*100];
B=mean(B./InDlys)*100;
%% B=sum(abs((Dlys(1:minlength)-InDlys(1:minlength))./...
% InDlys(1:minlength)))/minlength*100;
Tr=diff(Torig);
Td=diff(Textr);
TdAl=zeros(size(Tr));
TdAl(1)=Td(1);
nr=2;
nd=2;
```

```
GE=0;
SR=0;
SL=0;
PI=0;
PD=0;
Ch=[];
Thr=mean(Tr)*3/100;
%GEAntes=sum(abs(Tr(1:minlength)-Td(1:minlength))>Thr);
while ((nr<length(Tr)) & (nd<length(Td))),
while (Textr(nd-1)>=Torig(nr+1))
%PDCond=(Textr(nd)>=Torig(nr))
PD=PD+1;
nr=nr+1;
if ~((nr<length(Tr)) & (nd<length(Td)))
break
end
end
if ~((nr<length(Tr)) & (nd<length(Td)))
break
end
while (Textr(nd+1)<=Torig(nr-1))
%PICond=(Textr(nd+1)<=Torig(nr))
PI=PI+1;
nd=nd+1;
if ~((nr<length(Tr)) & (nd<length(Td)))
break
end
end
if ~((nr<length(Tr)) & (nd<length(Td)))
break
end
if (abs(Td(nd)-Tr(nr)) > Thr)%GECond= (abs(Td(nd)-
Tr(nr)) > Thr)
%SRCond= ( ((Td(nd)-Tr(nr+1)) < Thr) & ((Td(nd-1)-
Tr(nr)) < Thr) & (Torig(nr) > Textr(nd)) )
SRCond=( (abs(Td(nd)-Tr(nr+1)) < Thr) & (abs(Td(nd-
1)-Tr(nr)) < Thr) & ((Torig(nr+1)+Thr) < Textr(nd+1)) );
SLCond=( (abs(Td(nd)-Tr(nr-1)) < Thr) & (abs(Td(nd+1)-
Tr(nr)) < Thr) & (Torig(nr+1) > (Textr(nd+1)+Thr)) );
if (SRCond) & ~(SLCond)
SR=SR+1;
nd=nd-1;
Ch=[Ch; nr 1];
end
if (SLCond) & ~(SRCond)
SL=SL+1;
%TdAl(nr-1)=Td(nd);
%nd=nd+1;
nr=nr-1;
Ch=[Ch; nr -1];
end
GE=GE+1;
if ~((nr<length(Tr)) & (nd<length(Td)))
break
end
end
```



```

TdAl(nr)=Td(nd);
nr=nr+1;
nd=nd+1;
end
GE=[GE];
TdAl(nr)=Td(nd);
Temp=TdAl;
% Temp(find(abs(Temp-mean(Temp))./Temp > 0.25))=[];
% borrar esto si...
possup=find(Temp == 0);
Temp(possup)=[];
Tr(possup)=[];
possup=find(abs(Tr-mean(Tr)) > 0.25*mean(Tr) );
Temp(possup)=[];
Tr(possup)=[];
% hasta aqui
ndx=2:length(Temp);
A=[A mean(abs((Temp(ndx)-Temp(ndx-1)))./(Temp(ndx)+Temp(ndx-1))*100);
% ndx=find(abs(TdAl-Tr)./Tr < 0.2);
% B=[B mean(abs(TdAl(ndx)-Tr(ndx))./Tr(ndx))*100];
B=[B mean(abs(Temp-Tr)./Tr)*100];
else
A=[NaN NaN NaN];
B=[NaN NaN];
GE=[NaN];
SR=NaN;
SL=NaN;
PI=NaN;
PD=NaN;
TdAl=NaN;
Ch=NaN;

```

end

Anexo III Generador de Onda

1. Modelo E/C

```

function
[POPMio,randsett,pulse]=genpulseglottaleat(n,TmScs,T0,Fs,Tinst);
% TmScs=1;
% F0=150;
% Fs=22050;
F0=1/T0;
% n=1

JittAm=[5*(1:7)];
MeanT0=round(Fs/F0);
randjitt=[round((2*JittAm(1)+1)*rand(1,length(Tinst)))-
(2*JittAm(1)+1)/2);
round((2*JittAm(2)+1)*rand(1,length(Tinst)))-
(2*JittAm(2)+1)/2);
round((2*JittAm(3)+1)*rand(1,length(Tinst)))-
(2*JittAm(3)+1)/2);
round((2*JittAm(4)+1)*rand(1,length(Tinst)))-
(2*JittAm(4)+1)/2);

```

```

round((2*JittAm(5)+1)*rand(1,length(Tinst)))-
(2*JittAm(5)+1)/2);
round((2*JittAm(6)+1)*rand(1,length(Tinst)))-
(2*JittAm(6)+1)/2);
round((2*JittAm(7)+1)*rand(1,length(Tinst)))-
(2*JittAm(7)+1)/2);
randsett=randjitt+MeanT0;
p1=[];
% a=1+0.5*rand(1,TmScs*F0);
for l=1:TmScs*F0,
T=randsett(n,l);
[pulse,pdiff,pos]=glottalpulseanother(T/Fs,0.33*T/Fs,0.09*T/Fs,Fs);
p1=[p1 pdiff];
l;
end
HtSize=300; % size of vocal tract impulse response
h(t)
SN=[22 18 15 12 8 5 2]; % SNRs used
SNR = 30;
FFrq=[520 800 1190 1840 2390]; % Frequencies
FAmp=[250 60 215 25 40]; % Amplitudes
FBwt=[320 720 520 770 350]; % BandWidths
i=0:HtSize-1;
Ht=[];
FsHt=Fs;
for j=1:5,
Ht(j,:)=FAmp(j)*exp(-
FBwt(j)*i/FsHt).*cos(2*pi*i*FFrq(j)/FsHt);
end;
Ht=sum(Ht);
POPMio=conv(Ht,-p1); % POP = PO 'Pure'
end

```

2. Modelo S/A

```

function [POJ,TJitt]=POGenGlt(TmScs,T0,Fs,Tinst)
% Constants Used. (if using POGen alone, remove comments

% TmScs=1; % Duration of every vowel in
Secs
F0=1/T0; % Mean Fundamental
Frequency
% Fs=22050; % Sampling Frequency
% T0=1/F0; % Fundamental Period

PrsD=0.0001; % Presicion Desired
HtSize=300; % size of vocal tract impulse response
h(t)
SN=[22 18 15 12 8 5 2]; % SNRs used
SNR = 30;

% Vowel Formant Data
FFrq=[520 800 1190 1840 2390]; % Frequencies
FAmp=[250 60 215 25 40]; % Amplitudes
FBwt=[320 720 520 770 350]; % BandWidths

```

```
% *****
% *****
% *****
% Generation of the PO waveforms
%First, h(t)
i=0:HtSize-1;
Ht=[];
FsHt=Fs;
for j=1:5,
    Ht(j,:)=FAmp(j)*exp(-
    FBwt(j)*i/FsHt).*cos(2*pi*i*FFrq(j)/FsHt);
end;
Ht=sum(Ht);

% *****
% *****
% *****
% Pulse Train, No jitter, no shimmer
i=1:TmScs*F0; %Quantity of Glottal Pulses,
Fs/F0 samples appart
Tinst=i*Fs/F0; %By good luck,
Fs/F0=147=integer!
PlsTrn=zeros(1,TmScs*Fs); %Zeros everywhere
PlsTrn(Tinst)=1; %Except in Tinst, Glottal inst. of
exc.

% Glottal function
[gp,pd,p]=glottalpulsesanother(T0,0.33*T0,0.09*T0,Fs);

% POP
T=conv(Ht,PlsTrn); %POP = PO 'Pure'
POP=conv(T,-pd);
% TNorm=Tinst;

% *****
% *****
% ONLY JITTERED SIGNALS PLUS 30db
% *****
% *****
% Jittered Pulse Trains
%round((2*x+1)*rand(1,N)-(2*x+1)/2) generates a random
%uniformly distributed jitter of length N in the interval
%of +/- x samples, producing a mean jitter of x(x+1)/(2x+1)
%multiplied by 1/L, the influence of one sample in the L of F0
%Pulse Train, with increments of 3 samples of jitter, no
shimmer

i=1:TmScs*F0; %Quantity of Glottal Pulses,
Fs/F0 samples appart
Tinst=i*Fs/F0; %By good luck,
Fs/F0=147=integer!
```

```
JittAm=[5*(1:7)];
```

```
randjitt=[round((2*JittAm(1)+1)*rand(1,length(Tinst))-
(2*JittAm(1)+1)/2);
round((2*JittAm(2)+1)*rand(1,length(Tinst))-
(2*JittAm(2)+1)/2);
round((2*JittAm(3)+1)*rand(1,length(Tinst))-
(2*JittAm(3)+1)/2);
round((2*JittAm(4)+1)*rand(1,length(Tinst))-
(2*JittAm(4)+1)/2);
round((2*JittAm(5)+1)*rand(1,length(Tinst))-
(2*JittAm(5)+1)/2);
round((2*JittAm(6)+1)*rand(1,length(Tinst))-
(2*JittAm(6)+1)/2);
round((2*JittAm(7)+1)*rand(1,length(Tinst))-
(2*JittAm(7)+1)/2)];
```

```
MeanT0=round(Fs/F0);
TJitt=zeros(7,length(Tinst));
TJitt(:,1)=MeanT0;
```

```
for i=2:length(Tinst),
    TJitt(1,i)=TJitt(1,i-1)+ MeanT0 + randjitt(1,i);
    TJitt(2,i)=TJitt(2,i-1)+ MeanT0 + randjitt(2,i);
    TJitt(3,i)=TJitt(3,i-1)+ MeanT0 + randjitt(3,i);
    TJitt(4,i)=TJitt(4,i-1)+ MeanT0 + randjitt(4,i);
    TJitt(5,i)=TJitt(5,i-1)+ MeanT0 + randjitt(5,i);
    TJitt(6,i)=TJitt(6,i-1)+ MeanT0 + randjitt(6,i);
    TJitt(7,i)=TJitt(7,i-1)+ MeanT0 + randjitt(7,i);
end;
```

```
PlsTrn=zeros(size(TJitt,1),TmScs*Fs); %Zeros everywhere
```

```
PlsTrn(1,TJitt(1,:))=1;
PlsTrn(2,TJitt(2,:))=1;
PlsTrn(3,TJitt(3,:))=1;
PlsTrn(4,TJitt(4,:))=1;
PlsTrn(5,TJitt(5,:))=1;
PlsTrn(6,TJitt(6,:))=1;
PlsTrn(7,TJitt(7,:))=1; %Except in TJitt, Glottal
instants
```

```
%POJ= PO 'Jittered'
```

```
PO=[conv(Ht,PlsTrn(1,:));
conv(Ht,PlsTrn(2,:));
conv(Ht,PlsTrn(3,:));
conv(Ht,PlsTrn(4,:));
conv(Ht,PlsTrn(5,:));
conv(Ht,PlsTrn(6,:));
conv(Ht,PlsTrn(7,:));];
```

```
POJ=[conv(-pd,PO(1,:));
```

```

conv(-pd,PO(2,:));
conv(-pd,PO(3,:));
conv(-pd,PO(4,:));
conv(-pd,PO(5,:));
conv(-pd,PO(6,:));
conv(-pd,PO(7,:));];

fnoise = randn(1,length(POJ));
vfn = var(fnoise);

POJ=[POJ(1,:)+sqrt(var(POJ(1,:))/(vfn*10^(SNR/10)))*fnoise;
      POJ(2,:)+sqrt(var(POJ(2,:))/(vfn*10^(SNR/10)))*fnoise;
      POJ(3,:)+sqrt(var(POJ(3,:))/(vfn*10^(SNR/10)))*fnoise;
      POJ(4,:)+sqrt(var(POJ(4,:))/(vfn*10^(SNR/10)))*fnoise;
      POJ(5,:)+sqrt(var(POJ(5,:))/(vfn*10^(SNR/10)))*fnoise;
      POJ(6,:)+sqrt(var(POJ(6,:))/(vfn*10^(SNR/10)))*fnoise;
      POJ(7,:)+sqrt(var(POJ(7,:))/(vfn*10^(SNR/10)))*fnoise;
      ];

POJ(:,(TmScs*Fs+HtSize-1):end)=[];

% *****
% *****
% ONLY SHIMMERED PLUS 30db
% *****
% *****
%Shimmered pulse trains
%Pulse Train, No jitter, Variable shimmer
%Shim=50/100/2;
%Shim=[10 20 30 40 50 60 70]*2/147;

randshim=ones(7,TmScs*F0);
randshim(:,2:TmScs*F0)=1+2*randjitt(:,2:TmScs*F0)/(Fs/F0);
%shimmer is proportional to jitter
i=1:TmScs*F0;
%Quantity of Glottal Pulses, Fs/F0 samples appart
Tinst=i*Fs/F0; %By
good luck, Fs/F0=147=integer!
PlsTrn=zeros(7,TmScs*Fs); %Zeros
everywhere
PlsTrn(:,Tinst)=randshim; %Except in
Tinst, Glottal inst. of exc.

P=[conv(Ht,PlsTrn(1,:));
   conv(Ht,PlsTrn(2,:));
   conv(Ht,PlsTrn(3,:));
   conv(Ht,PlsTrn(4,:));
   conv(Ht,PlsTrn(5,:));
   conv(Ht,PlsTrn(6,:));
   conv(Ht,PlsTrn(7,:));]; %POS= PO 'Shimmered'

POS=[conv(P(1,:),-pd);
      conv(P(2,:),-pd);
      conv(P(3,:),-pd);

```

```

conv(P(4,:),-pd);
conv(P(5,:),-pd);
conv(P(6,:),-pd);
conv(P(7,:),-pd);];

fnoise = randn(1,length(POS));
vfn = var(fnoise);

POS = [ POS(1,:)+
sqrt(var(POS(1,:))/(vfn*10^(SNR/10)))*fnoise;
      POS(2,:)+ sqrt(var(POS(2,:))/(vfn*10^(SNR/10)))*fnoise;
      POS(3,:)+ sqrt(var(POS(3,:))/(vfn*10^(SNR/10)))*fnoise;
      POS(4,:)+ sqrt(var(POS(4,:))/(vfn*10^(SNR/10)))*fnoise;
      POS(5,:)+ sqrt(var(POS(5,:))/(vfn*10^(SNR/10)))*fnoise;
      POS(6,:)+ sqrt(var(POS(6,:))/(vfn*10^(SNR/10)))*fnoise;
      POS(7,:)+
sqrt(var(POS(7,:))/(vfn*10^(SNR/10)))*fnoise];

% *****
% *****
% GAUSSIAN NOISE *
% *****
% *****
%Oral Pressure (PO), No jitter, no shimmer, variable dB S/N
%SN=[-5 -2 0 2 5 10 15 20 25];%PO

i=1:length(SN); %We use the pure
signal (POP) plus noise
NoiseSTD(i)=sqrt(var(POP)/10.^(SN(i)/10));
POG=[POP+ NoiseSTD(1)*randn(1,length(POP));
      POP+ NoiseSTD(2)*randn(1,length(POP));
      POP+ NoiseSTD(3)*randn(1,length(POP));
      POP+ NoiseSTD(4)*randn(1,length(POP));
      POP+ NoiseSTD(5)*randn(1,length(POP));
      POP+ NoiseSTD(6)*randn(1,length(POP));
      POP+ NoiseSTD(7)*randn(1,length(POP));]; %PON= PO
'Noised'

POG(:,(TmScs*Fs+HtSize-1):end)=[];

% *****
% *****
% FILTERED NOISE *
% *****
% *****
%POP will be filtered noise

fnoise = randn(1,length(POP));
fnoise = conv(Ht,fnoise);
fnoise =
fnoise(round(length(Ht)/2)+1:round(length(Ht)/2)+length(POP
));
vfn = var(fnoise);
POCF = zeros(7,length(POP));

```

```
POF = [POP + sqrt(var(POP)/(vfn*10^(SN(1)/10)))*fnoise;
      POP + sqrt(var(POP)/(vfn*10^(SN(2)/10)))*fnoise;
      POP + sqrt(var(POP)/(vfn*10^(SN(3)/10)))*fnoise;
      POP + sqrt(var(POP)/(vfn*10^(SN(4)/10)))*fnoise;
      POP + sqrt(var(POP)/(vfn*10^(SN(5)/10)))*fnoise;
      POP + sqrt(var(POP)/(vfn*10^(SN(6)/10)))*fnoise;
      POP + sqrt(var(POP)/(vfn*10^(SN(7)/10)))*fnoise; ];
```

```
POF(:,(TmScs*Fs+HtSize-1):end)=[];
```

```
% *****
```

```
% JITTERED_SHIMMERED PLUS 30 DB *
```

```
% *****
```

```
% Compound distortions, shimmer(%)=2*jitter(%), decreasing  
dB S/N,
```

```
% Distortions in shimmer are based on the "jittered"  
experiment,
```

```
% the amplitude of the pulse N is found based on the jitter  
distortion
```

```
% of the N+1 pulse. jitter(%) is the randjitt matrix times  
1/(Fs/F0)*100
```

```
i=1:TmScs*F0; %Quantity of Glottal Pulses,  
Fs/F0 samples appart  
Tinst=i*Fs/F0; %By good luck,  
Fs/F0=147=integer!
```

```
PlsTrn=zeros(7,TmScs*Fs); %Zeros everywhere  
PlsTrn(:,147)=1; %Except in TJitt, Glottal inst. of exc.
```

```
% Notice that shimmer perturbations are identical (same values)  
than if we
```

```
% used randshim, it is just a matter of code.
```

```
PlsTrn(1,TJitt(1,1:length(TJitt)-  
1))=1+2*randjitt(1,2:length(TJitt))/(Fs/F0);  
PlsTrn(2,TJitt(2,1:length(TJitt)-  
1))=1+2*randjitt(2,2:length(TJitt))/(Fs/F0);  
PlsTrn(3,TJitt(3,1:length(TJitt)-  
1))=1+2*randjitt(3,2:length(TJitt))/(Fs/F0);  
PlsTrn(4,TJitt(4,1:length(TJitt)-  
1))=1+2*randjitt(4,2:length(TJitt))/(Fs/F0);  
PlsTrn(5,TJitt(5,1:length(TJitt)-  
1))=1+2*randjitt(5,2:length(TJitt))/(Fs/F0);  
PlsTrn(6,TJitt(6,1:length(TJitt)-  
1))=1+2*randjitt(6,2:length(TJitt))/(Fs/F0);  
PlsTrn(7,TJitt(7,1:length(TJitt)-  
1))=1+2*randjitt(7,2:length(TJitt))/(Fs/F0);
```

```
% PO shimmered and jittered
```

```
PO=[conv(Ht,PlsTrn(1,:));  
    conv(Ht,PlsTrn(2,:));  
    conv(Ht,PlsTrn(3,:));  
    conv(Ht,PlsTrn(4,:));  
    conv(Ht,PlsTrn(5,:));  
    conv(Ht,PlsTrn(6,:));
```

```
conv(Ht,PlsTrn(7,:));  
POSJ=[conv(-pd,PO(1,:));  
      conv(-pd,PO(2,:));  
      conv(-pd,PO(3,:));  
      conv(-pd,PO(4,:));  
      conv(-pd,PO(5,:));  
      conv(-pd,PO(6,:));  
      conv(-pd,PO(7,:));];
```

```
fnoise = randn(1,length(POSJ));
```

```
vfn = var(fnoise);
```

```
POSJ=[POSJ(1,:)+  
sqrt(var(POSJ(1,:))/(vfn*10^(SNR/10)))*fnoise;  
      POSJ(2,:)+  
sqrt(var(POSJ(2,:))/(vfn*10^(SNR/10)))*fnoise;  
      POSJ(3,:)+  
sqrt(var(POSJ(3,:))/(vfn*10^(SNR/10)))*fnoise;  
      POSJ(4,:)+  
sqrt(var(POSJ(4,:))/(vfn*10^(SNR/10)))*fnoise;  
      POSJ(5,:)+  
sqrt(var(POSJ(5,:))/(vfn*10^(SNR/10)))*fnoise;  
      POSJ(6,:)+  
sqrt(var(POSJ(6,:))/(vfn*10^(SNR/10)))*fnoise;  
      POSJ(7,:)+  
sqrt(var(POSJ(7,:))/(vfn*10^(SNR/10)))*fnoise;  
      ];
```

```
POSJ(:,(TmScs*Fs+HtSize-1):end)=[];
```

```
% *****
```

```
% JITTERED SIGNALS PLUS [22 18 15 12 8 5 2]DB OF  
GAUSSIAN NOISE
```

```
% *****
```

```
% Jittered-Noised-Gaussian
```

```
i=1:length(SN); %We use the pure  
signal (POP) plus noise  
NoiseSTD(i)=sqrt(var(POJ(i,:))/10.^(SN(i)/10));
```

```
POJG=[ POJ(1,:)+ NoiseSTD(1)*randn(1,length(POJ));  
      POJ(2,:)+ NoiseSTD(2)*randn(1,length(POJ));  
      POJ(3,:)+ NoiseSTD(3)*randn(1,length(POJ));  
      POJ(4,:)+ NoiseSTD(4)*randn(1,length(POJ));  
      POJ(5,:)+ NoiseSTD(5)*randn(1,length(POJ));  
      POJ(6,:)+ NoiseSTD(6)*randn(1,length(POJ));  
      POJ(7,:)+ NoiseSTD(7)*randn(1,length(POJ));  
      ];  
POJG(:,(TmScs*Fs+HtSize-1):end)=[];
```

```
% *****
```

```
% JITTERED SIGNALS PLUS [22 18 15 12 8 5 2]DB OF  
FILTERED NOISE
```

```
% *****
```

```
% Jittered-Noised-Filtered
```

```
fnoise = randn(1,length(POJ));
fnoise = conv(Ht,fnoise);
fnoise
fnoise(round(length(Ht)/2)+1:round(length(Ht)/2)+length(POJ)
);
vfn = var(fnoise);
POJFN = zeros(7,length(POJ));
```

```
POJF=[POJ(1,:)+
sqrt(var(POJ(1,:))/(vfn*10^(SN(1)/10)))*fnoise;
POJ(2,:)+ sqrt(var(POJ(2,:))/(vfn*10^(SN(2)/10)))*fnoise;
POJ(3,:)+ sqrt(var(POJ(3,:))/(vfn*10^(SN(3)/10)))*fnoise;
POJ(4,:)+ sqrt(var(POJ(4,:))/(vfn*10^(SN(4)/10)))*fnoise;
POJ(5,:)+ sqrt(var(POJ(5,:))/(vfn*10^(SN(5)/10)))*fnoise;
POJ(6,:)+ sqrt(var(POJ(6,:))/(vfn*10^(SN(6)/10)))*fnoise;
POJ(7,:)+ sqrt(var(POJ(7,:))/(vfn*10^(SN(7)/10)))*fnoise;
];
```

```
POJF(:,(TmScs*Fs+HtSize-1):end)=[];
```

```
% *****
% SHIMMERED SIGNALS PLUS [22 18 15 12 8 5 2]DB OF
GAUSSIAN NOISE
% *****
% Shimmered-noised-Gaussian signal
j=1:size(POS,1);
NoiseSTD(j)=sqrt(var(POS(j,:))./10.^(SN(j)/10));
```

```
POSG=[POS(1,:)+ NoiseSTD(1) *randn(1,length(POS));
POS(2,:)+ NoiseSTD(2) *randn(1,length(POS));
POS(3,:)+ NoiseSTD(3) *randn(1,length(POS));
POS(4,:)+ NoiseSTD(4) *randn(1,length(POS));
POS(5,:)+ NoiseSTD(5) *randn(1,length(POS));
POS(6,:)+ NoiseSTD(6) *randn(1,length(POS));
POS(7,:)+ NoiseSTD(7) *randn(1,length(POS));];
```

```
POSG(:,(TmScs*Fs+HtSize-1):end)=[];
```

```
% *****
% SHIMMERED SIGNALS PLUS [22 18 15 12 8 5 2]DB OF
FILTERED NOISE
% *****
% Shimmered-noised-Filtered signal
```

```
fnoise = randn(1,length(POS));
fnoise = conv(Ht,fnoise);
fnoise
fnoise(round(length(Ht)/2)+1:round(length(Ht)/2)+length(POS
));
vfn = var(fnoise);
POSF = zeros(7,length(POS));
```

```
POSF=[POS(1,:)+
sqrt(var(POS(1,:))/(vfn*10^(SN(1)/10)))*fnoise;
POS(2,:)+
sqrt(var(POS(2,:))/(vfn*10^(SN(2)/10)))*fnoise;
POS(3,:)+
sqrt(var(POS(3,:))/(vfn*10^(SN(3)/10)))*fnoise;
POS(4,:)+
sqrt(var(POS(4,:))/(vfn*10^(SN(4)/10)))*fnoise;
POS(5,:)+
sqrt(var(POS(5,:))/(vfn*10^(SN(5)/10)))*fnoise;
POS(6,:)+
sqrt(var(POS(6,:))/(vfn*10^(SN(6)/10)))*fnoise;
POS(7,:)+
sqrt(var(POS(7,:))/(vfn*10^(SN(7)/10)))*fnoise;
];
```

```
POSF(:,(TmScs*Fs+HtSize-1):end)=[];
```

```
% *****
% GAUSSIAN COMPOUND *
% *****
%POC will be PO 'Compound', by adding noise to POSJ
```

```
j=1:size(POSJ,1);
NoiseSTD(j)=sqrt(var(POSJ(j,:))./10.^(SN(j)/10));
```

```
POCG=[POSJ(1,:)+ NoiseSTD(1)*randn(1,size(POSJ,2));
POSJ(2,:)+ NoiseSTD(2)*randn(1,size(POSJ,2));
POSJ(3,:)+ NoiseSTD(3)*randn(1,size(POSJ,2));
POSJ(4,:)+ NoiseSTD(4)*randn(1,size(POSJ,2));
POSJ(5,:)+ NoiseSTD(5)*randn(1,size(POSJ,2));
POSJ(6,:)+ NoiseSTD(6)*randn(1,size(POSJ,2));
POSJ(7,:)+ NoiseSTD(7)*randn(1,size(POSJ,2));];
```

```
POCG(:,(TmScs*Fs+HtSize-1):end)=[];
```

```
% *****
% FILTERED COMPOSITE *
% *****
%POPC will be POP 'Compound', by adding filtered noise to
POSJ
```

```
fnoise = randn(1,length(POSJ));
fnoise = conv(Ht,fnoise);
fnoise
fnoise(round(length(Ht)/2)+1:round(length(Ht)/2)+length(POS
J));
vfn = var(fnoise);
POCF = zeros(7,length(POSJ));
```

```
POCF = [POSJ(1,:) +
sqrt(var(POSJ(1,:))/(vfn*10^(SN(1)/10)))*fnoise;
POSJ(2,:) +
sqrt(var(POSJ(2,:))/(vfn*10^(SN(2)/10)))*fnoise;
```

```

    POSJ(3,:);
    sqrt(var(POSJ(3,:))/(vfn*10^(SN(3)/10)))*fnoise;
    POSJ(4,:);
    sqrt(var(POSJ(4,:))/(vfn*10^(SN(4)/10)))*fnoise;
    POSJ(5,:);
    sqrt(var(POSJ(5,:))/(vfn*10^(SN(5)/10)))*fnoise;
    POSJ(6,:);
    sqrt(var(POSJ(6,:))/(vfn*10^(SN(6)/10)))*fnoise;
    POSJ(7,:);
    sqrt(var(POSJ(7,:))/(vfn*10^(SN(7)/10)))*fnoise; ];

    POCF(:,(TmScs*Fs+HtSize-1):end)=[];

```

Anexo IV Programa principal

```

% -----Matrices Modelo2-----

ShM2Bor=zeros(4,7,100);
ShM2Bcorr=zeros(4,7,100);
ShM2GE=zeros(4,7,100);
ShM2SR=zeros(4,7,100);
ShM2SL=zeros(4,7,100);
ShM2PI=zeros(4,7,100);
ShM2PD=zeros(4,7,100);

%-----Matrices Modelo1-----

ShM1Bor=zeros(4,7,100);
ShM1Bcorr=zeros(4,7,100);
ShM1GE=zeros(4,7,100);
ShM1SR=zeros(4,7,100);
ShM1SL=zeros(4,7,100);
ShM1PI=zeros(4,7,100);
ShM1PD=zeros(4,7,100);

% -----3-----

ShM3Bor=zeros(4,7,100);
ShM3Bcorr=zeros(4,7,100);
ShM3GE=zeros(4,7,100);
ShM3SR=zeros(4,7,100);
ShM3SL=zeros(4,7,100);
ShM3PI=zeros(4,7,100);
ShM3PD=zeros(4,7,100);

% -----Cabecera-----

RMS=zeros(1,7,100);
PEAK=zeros(1,7,100);

```

```

+   Milenk=zeros(1,7,100);
+   Medam=zeros(1,7,100);
+   TmScs=1;
+   F0=150;
+   Fs=22050;
+   T0=1/F0;
+   SrchR=round(.25*Fs/F0);
+   TpEst=round(Fs/F0);
+   i=1:TmScs*F0;
+   Tinst=i*Fs/F0;

% -----Ciclo de 100 repeticiones----

for c=1:100

% ----Ciclo para cada Nivel (1:7)-----

for n=1:7

% ----Prueba con Milenkovick-----

% -----Modelo2 VS Milenkovick-----

[POPMio,randsett,pulse]=genpulseglottaleat(n,TmScs,T0,Fs,Ti
nst);

% TR2=Milenkovick(POPMio,Tinst,F0,Fs,SrchR,TpEst);

[TR2,K,
K1]=Mlk(POPMio,Fs,TpEst,SrchR,round(.42*randsett(n,1)));

% KMil(1,n,c)=K;

Torig2=cumsum(randsett(n,:))-round(.58*randsett(n,:));

Textr2=TR2;

[A2,B2,GE2,SR2,SL2,PI2,PD2]=PerturbMeasures(Torig2,Text
r2);

ShM2Bor(1,n,c)=B2(1);
ShM2Bcorr(1,n,c)=B2(2);
ShM2GE(1,n,c)=GE2;
ShM2SR(1,n,c)=SR2;
ShM2SL(1,n,c)=SL2;
ShM2PI(1,n,c)=PI2;
ShM2PD(1,n,c)=PD2;

% -----

Torig2=cumsum(randsett(n,:));

```

```

[A2,B2,GE2,SR2,SL2,PI2,PD2]=PerturbMeasures(Torig2,Text
r2);
    ShM3Bor(1,n,c)=B2(1);
    ShM3Bcorr(1,n,c)=B2(2);
    ShM3GE(1,n,c)=GE2;
    ShM3SR(1,n,c)=SR2;
    ShM3SL(1,n,c)=SL2;
    ShM3PI(1,n,c)=PI2;
    ShM3PD(1,n,c)=PD2;
    % las otras...
% -----Modelo1 VS Milenkovich----
[POJ,TJitt]=POGenGlt(TmScs,T0,Fs,Tinst);
% TR=Milenkovich(POJ,Tinst,F0,Fs,SrchR,TpEst);
[TR]=Mlk(POJ(n,:),Fs,TpEst,SrchR,TJitt(n,1));

Torig=(TJitt(n,:));
Textr=TR;
[A,B,GE,SR,SL,PI,PD]=PerturbMeasures(Torig,Textr);
    ShM1Bor(1,n,c)=B(1);
    ShM1Bcorr(1,n,c)=B(2);
    ShM1GE(1,n,c)=GE;
    ShM1SR(1,n,c)=SR;
    ShM1SL(1,n,c)=SL;
    ShM1PI(1,n,c)=PI;
    ShM1PD(1,n,c)=PD;
% -----Prueba con Meda-----
-----Modelo2 VS Meda-----
[TR2]=Medan(POPMio,Tinst,F0,Fs,SrchR,TpEst);
[TR2,K2,          K3]=          Med
(POPMio,Fs,TpEst,SrchR,round(.42*randsett(n,1)));
Textr2=TR2;
    Torig2=cumsum(randsett(n,:))-round(.58*randsett(n,:));

[A2,B2,GE2,SR2,SL2,PI2,PD2]=PerturbMeasures(Torig2,Text
r2);
    ShM2Bor(2,n,c)=B2(1);
    ShM2Bcorr(2,n,c)=B2(2);
    ShM2GE(2,n,c)=GE2;
    ShM2SR(2,n,c)=SR2;
    ShM2SL(2,n,c)=SL2;
    ShM2PI(2,n,c)=PI2;
    ShM2PD(2,n,c)=PD2;
% -----
Torig2=cumsum(randsett(n,:));

[A2,B2,GE2,SR2,SL2,PI2,PD2]=PerturbMeasures(Torig2,Text
r2);
    ShM3Bor(2,n,c)=B2(1);
    ShM3Bcorr(2,n,c)=B2(2);
    ShM3GE(2,n,c)=GE2;
    ShM3SR(2,n,c)=SR2;
    ShM3SL(2,n,c)=SL2;
    ShM3PI(2,n,c)=PI2;
    ShM3PD(2,n,c)=PD2;
-----Modelo1 VS Meda-----
% [TR]=Medan(POJ,Tinst,F0,Fs,SrchR,TpEst);
[TR]= Med (POJ(n,:),Fs,TpEst,SrchR,TJitt(n,1));
    Textr=TR;
[A,B,GE,SR,SL,PI,PD]=PerturbMeasures(Torig,Textr);
    ShM1Bor(2,n,c)=B(1);
    ShM1Bcorr(2,n,c)=B(2);
    ShM1GE(2,n,c)=GE;
    ShM1SR(2,n,c)=SR;
    ShM1SL(2,n,c)=SL;
    ShM1PI(2,n,c)=PI;
    ShM1PD(2,n,c)=PD;
% ---Modelo2 VS MedaModificado-----
TR2=CAI(POPMio,Fs,TpEst,SrchR,round(.42*randsett(n,1)));
    Textr2=TR2;
    Torig2=cumsum(randsett(n,:))-round(.58*randsett(n,:));

[A2,B2,GE2,SR2,SL2,PI2,PD2]=PerturbMeasures(Torig2,Text
r2);
    ShM2Bor(3,n,c)=B2(1);
    ShM2Bcorr(3,n,c)=B2(2);
    ShM2GE(3,n,c)=GE2;

```

```

ShM2SR(3,n,c)=SR2;
ShM2SL(3,n,c)=SL2;
ShM2PI(3,n,c)=PI2;
ShM2PD(3,n,c)=PD2;
% -----
Torig2=cumsum(randsett(n,:));

[A2,B2,GE2,SR2,SL2,PI2,PD2]=PerturbMeasures(Torig2,Text
r2);

ShM3Bor(3,n,c)=B2(1);
ShM3Bcorr(3,n,c)=B2(2);
ShM3GE(3,n,c)=GE2;
ShM3SR(3,n,c)=SR2;
ShM3SL(3,n,c)=SL2;
ShM3PI(3,n,c)=PI2;
ShM3PD(3,n,c)=PD2;
%--Modelo1 VS MedanModificado-----
% TR=MedanMod(POJ,Tinst,F0,Fs,SrchR,TpEst);
TR= CAI (POJ(n,:),Fs,TpEst,SrchR,TJitt(n,1));
Textr=TR;
[A,B,GE,SR,SL,PI,PD]=PerturbMeasures(Torig,Textr);
ShM1Bor(3,n,c)=B(1);
ShM1Bcorr(3,n,c)=B(2);
ShM1GE(3,n,c)=GE;
ShM1SR(3,n,c)=SR;
ShM1SL(3,n,c)=SL;
ShM1PI(3,n,c)=PI;
ShM1PD(3,n,c)=PD;
% ----Modelo2 VS MyCC-----

TR2=CAII(POPMio,Fs,TpEst,SrchR,round(.42*randsett(n,1)));
Textr2=TR2;
Torig2=cumsum(randsett(n,:))-round(.58*randsett(n,:));

[A2,B2,GE2,SR2,SL2,PI2,PD2]=PerturbMeasures(Torig2,Text
r2);

ShM2Bor(4,n,c)=B2(1);
ShM2Bcorr(4,n,c)=B2(2);
ShM2GE(4,n,c)=GE2;
ShM2SR(4,n,c)=SR2;
ShM2SL(4,n,c)=SL2;
ShM2PI(4,n,c)=PI2;
ShM2PD(4,n,c)=PD2;
% -----
Torig2=cumsum(randsett(n,:));

[A2,B2,GE2,SR2,SL2,PI2,PD2]=PerturbMeasures(Torig2,Text
r2);

ShM3Bor(4,n,c)=B2(1);
ShM3Bcorr(4,n,c)=B2(2);
ShM3GE(4,n,c)=GE2;
ShM3SR(4,n,c)=SR2;
ShM3SL(4,n,c)=SL2;
ShM3PI(4,n,c)=PI2;
ShM3PD(4,n,c)=PD2;
%----Modelo1 VS MyCC-----
TR=CAII (POJ(n,:),Fs,TpEst,SrchR,TJitt(n,1));
Textr=TR;
[A,B,GE,SR,SL,PI,PD]=PerturbMeasures(Torig,Textr);
ShM1Bor(4,n,c)=B(1);
ShM1Bcorr(4,n,c)=B(2);
ShM1GE(4,n,c)=GE;
ShM1SR(4,n,c)=SR;
ShM1SL(4,n,c)=SL;
ShM1PI(4,n,c)=PI;
ShM1PD(4,n,c)=PD;
Shimmer
ERROR=mean(( abs(AmRMS/mean(AmRMS) - 1) ))*100;
RMS(1,n,c)=ERROR;
ERROR=mean(( abs(AmPEAK/mean(AmPEAK) - 1) )
)*100;
PEAK(1,n,c)=ERROR;
ERROR=mean(( abs(K/mean(K) - 1) ))*100;
Milenk(1,n,c)=ERROR;
ERROR=mean(( abs(K2/mean(K2) - 1) ))*100;
Medam(1,n,c)=ERROR;
display(n)

```



```

end
display(c)
end
a=mean(RMS,3);
b=mean(PEAK,3);
d=mean(Milenk,3);
e=mean(Medam,3);

figure(1)
plot(a)
hold on
plot(b)
hold on
plot(d)
hold on
plot(e)
hold on
plot(e,'k')

f=mean(ShM1Bcorr,3);
g=mean(ShM2Bcorr,3);
h=mean(ShM3Bcorr,3);
figure(2)
plot(f)
hold on
plot(g)
hold on
plot(h)

```

Anexo V Detector de *Shimmer*

```

AmRMS=[];
AmPEAK = [];
rs=[ones(7,1) randsett];
rs=cumsum(rs,2);
for i = 1:size(rs-2)
    S=POPMio(rs(1,i):rs(1,i+1));
    SD=POPMio((rs(1,i+1)+1):(rs(1,i+2)));
AmRMS=[AmRMS sqrt(sum(SD.^2))/sqrt(sum(S.^2))];

```