

# SIMULADOR DE SÍNTESIS DE VOZ EN EL IDIOMA ESPAÑOL

SEBASTIÁN HOYOS

**Abstract**-El simulador de síntesis de voz desarrollado en este proyecto está basado en el concepto de una fuente de excitación y un filtro autorregresivo que modela el tracto bucal. Los coeficientes del filtro son calculados por predicción lineal sobre la señal de voz original (corpus) y la fuente de excitación es la señal de error resultante al pasar el corpus por el filtro LPC *todo ceros*. Por lo tanto la síntesis propuesta es la basada en concatenación de unidades de sonido que en este caso corresponden a difonemas. El cambio de la curva melódica y de la duración de cada fonema es llevada a cabo con métodos nuevos que corresponden al aporte en esta área y que son evaluados comparándolos con resultados de otros sintetizadores.

## I. INTRODUCCIÓN

Un simulador de síntesis de voz es una herramienta de investigación que permite a los usuarios buscar la mejor descripción fonética y prosódica para obtener una buena calidad de la voz sintetizada. Entendiendo por calidad que la voz sea inteligible y natural.

En un sistema completo de texto a voz, es decir, un sistema que tiene como entrada un texto y cuya salida la reproducción en voz alta de éste texto, la síntesis constituye la última etapa. La etapa de síntesis recibe de las etapas anteriores toda la descripción fonética con la duración de cada fonema, la curva melódica y la amplitud que determinarán las pausas, acentos, indicación de pregunta o afirmación, etc. El sintetizador toma todas estas entradas y haciendo uso de

una base de datos de difonemas (síntesis por concatenación de difonemas)[2] realiza los procesos necesarios para darle a la voz las características indicadas. Es claro entonces que dependiendo de que tan correctos sean los datos entregados al sintetizador dependerá la calidad de la síntesis lograda. Por esta razón es a veces necesario realizar diversas modificaciones a este conjunto de entradas para lograr resultados óptimos y es precisamente allí donde es útil el simulador de síntesis de voz para escoger de una forma ágil el valor adecuado de estas variables.

El sintetizador descrito aquí además de ser una excelente herramienta para realizar el proceso descrito anteriormente, muestra las bondades y defectos de los métodos propuestos e investigados durante todo el desarrollo de este trabajo.

## II. MARCO TEÓRICO

### Modelo autorregresivo basado en LPC .

Este método está basado en el concepto de una fuente de excitación y un filtro que modela el tracto bucal tal como lo propuso Fant[1]. La fuente de excitación idealmente puede estar conformada por pulsos periódicos para los sonidos vocálicos y por ruido blanco para sonidos no vocálicos. Los coeficientes del filtro son extraídos directamente de la señal de voz haciendo predicción lineal. De esta forma cualquier señal de voz puede ser generada pasando la señal de excitación correcta (los pulsos periódicos o el ruido blanco) por el filtro autorregresivo correspondiente. La figura 1 muestra la síntesis de voz por medio del modelo autorregresivo basado en LPC .

Sin embargo, con este modelo ideal de la señal de excitación, no se ha logrado síntesis de alta calidad[1] y por eso se hace necesario reemplazarla por la señal de

excitación real obtenida de pasar la señal de voz por el filtro inverso al autorregresivo. El lector debe notar que este filtro hace las veces de un tracto anti-bucal imaginario al que le entra la señal de voz tal como la oímos y responde con la señal que proviene de los pulmones y cuerdas bucales.

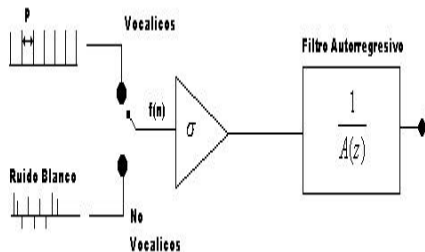


FIGURA 1  
Modelo autorregresivo basado en LPC

La principal ventaja que presenta utilizar el modelo de esta forma es que las operaciones de cambio de tono, de amplitud, y de duración se hacen sobre la señal de excitación tal como sucede en el caso natural y por lo tanto tiene más sentido que hacerlo sobre la señal de voz directamente.

### Síntesis de voz basada en reglas vs basada en concatenación.

En la descripción anterior se plantearon dos métodos para llevar a cabo la síntesis de voz utilizando el modelo autorregresivo basado en LPC. El primero no utiliza ninguna señal de voz grabada previamente (corpus de voz) sino que está basado en modelos tanto para las fuentes de excitación como para el filtro que representa el tracto bucal. Por supuesto estos modelos son obtenidos después de estudios y observación de señales reales y por lo tanto en una primera fase sería necesario tener alguna base de datos de corpus grabada pero que no sería utilizada en el sintetizador final. Este método corresponde a la **síntesis basada en reglas**[2].

La segunda opción es utilizar el corpus de la señal como fuente de excitación tal como se describió anteriormente. Por requerir concatenación de las unidades básicas que conforman el corpus es llamada **síntesis basada en concatenación**[2].

## II. SÍNTESIS DE VOZ

A continuación se muestra el trabajo de investigación desarrollado y que corresponde a los aportes hechos en el área de síntesis de voz.

### LA FUENTE DE EXCITACIÓN

Para escoger la fuente de excitación entre una señal de error simulada y una señal de error real se hizo un estudio que se presenta a continuación en forma resumida.

Se evaluó la utilización del modelo autorregresivo basado en LPC de forma completamente ideal. El procedimiento fue extraer los coeficientes LPC de orden 18 de la señal de voz, desechar la señal de error resultante al pasar por el filtro *todo ceros* LPC y simular la señal de error  $E(n)$ . Finalmente la señal de error  $E(n)$  simulada fue pasada por el filtro autorregresivo para obtener la señal de voz sintetizada.

Para simular la señal de error  $E(n)$  se hicieron observaciones previas de la real obtenida al pasar la señal de voz por el filtro *todo ceros* LPC. Para el caso de sonidos vocálicos se ensayaron el modelo de impulsos periódicos y el de dobles impulsos periódicos cuyo periodo podía ser cambiado para obtener diferentes frecuencias fundamentales de la señal de tono. La amplitud de la señal cambiaba de acuerdo a la ganancia obtenida en el proceso de extracción de parámetros LPC. Para sonidos no vocálicos se simuló  $E(n)$  como ruido blanco. El análisis se hizo en ventanas de 10 ms.

La ventaja que tiene este modelo es que la base de datos se reduce considerablemente porque sólo se deben guardar los coeficientes del filtro LPC.

La desventaja evidente fue la difícil escogencia entre los pulsos periódicos y el ruido blanco para modelar la señal de error correcta en cada caso. Este problema se agravó para los sonidos nasales que tienen características periódicas y necesitan también de ruido blanco en la señal de excitación simulada.

Las diferencias de calidad entre la voz sintetizada y la voz original fueron demasiado grandes, por lo cual no se justificó el ahorro en espacio de memoria y rapidez de ejecución de la señal simulada, es decir, se decidió trabajar con la señal de error real que implica tener una base de datos con el corpus de voz grabado.

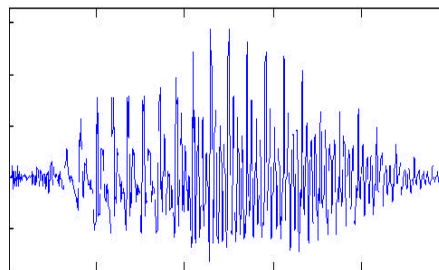
### **BASE DE DATOS Y CAMBIO DE LA DURACIÓN DE LOS FONEMAS.**

Uno de los pasos más importantes en este desarrollo fue escoger una base de datos con buenas características de grabación y segmentación del corpus. Por ser un sintetizador por concatenación de difonemas, estos debían estar segmentados de tal forma que el difonema sólo presentará el estado estable de los dos fonemas involucrados y el transitorio entre los mismos, pero no debía tener en sus bordes influencia de los difonemas que lo precedían y seguían en la palabra de la que era extraído. Por lo tanto en la segmentación de difonemas siempre se corre el riesgo de tomar parte de la señal del fonema anterior o del siguiente y esto puede traer consecuencias desastrosas en los resultados de la síntesis. De igual forma si el difonema era segmentado quitando información de su estado estable del principio o del final, seguramente todo el difonema o por los menos uno de sus dos fonemas no sería inteligible.

En los difonemas con estado estable definido como en los sonidos vocálicos es más fácil realizar la segmentación porque la señal es semiperiódica y normalmente existen periodos estables sin influencia de los fonemas vecinos y es precisamente allí donde debería realizarse la segmentación.

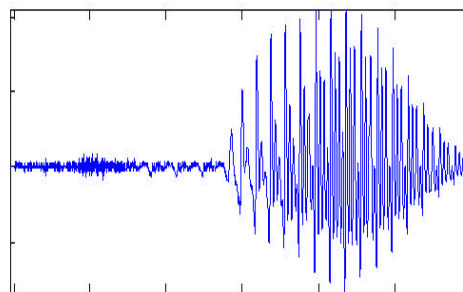
Para los difonemas no periódicos tales como los sonidos fricativos, el estado estable de la señal está claramente definido y también es fácil realizar la segmentación. Pero en los sonidos explosivos donde el transitorio corresponde a toda la información del fonema, es crítica la forma en que se realiza la segmentación. Si la segmentación es llevada a cabo ligeramente más allá del final del fonema explosivo, aparecerá inevitablemente el sonido indeseable del fonema siguiente, en el momento de la concatenación con otros difonemas. Si se

recorta la señal perdiendo así sea un poco de información del fonema, éste podría llegar a ser ininteligible. Para ilustrar esto, considere



**FIGURA 2**  
**DIFONEMA /ca DONDE SE DISTINGUEN EL**  
**TRANSITORIO DEL FONEMA /c**  
**Y EL ESTADO ESTABLE DEL FONEMA /ca**

la figura 2 que corresponde a la señal en tiempo del difonema /ca. Se ve claramente el transitorio del fonema /c sin estado estable por ser un sonido explosivo y también se ve el fonema /a que es un sonido vocálico y por lo tanto tiene estado estable semiperiódico definido. Debe notarse la rápida aparición del fonema /a una vez se acaba el fonema /c lo cual dificulta aún más saber cuál es la mitad del difonema.



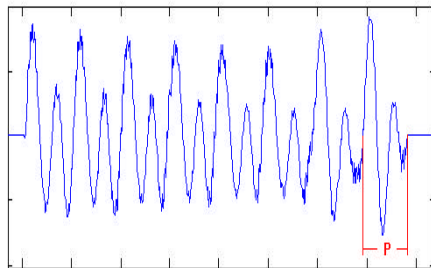
**FIGURA 3**  
**DIFONEMA /sa DONDE SE VEN LOS ESTADOS**  
**ESTABLES DE AMBOS FONEMAS**

En la figura 3 se muestra la señal en tiempo del difonema /sa. Ambos fonemas tienen estado estable definido, pero el fonema /s no es periódico por ser un sonido fricativo. En este difonema es más claro el punto de transición de un fonema a otro.

Aunque en este trabajo no se creó la base de datos de difonemas, sí se hizo un estudio del corpus para calificar su calidad de grabación y de segmentación porque es un punto clave en el éxito del sintetizador. Por esta razón son de gran importancia los comentarios hechos anteriormente.

La base de datos adquirida traía la información de inicio, mitad y final del difonema, por lo cual ese trabajo no fue necesario llevarlo a cabo. Sin embargo los métodos propuestos aquí para cambiar la duración de cada fonema también exigían el conocimiento del principio de su región más estable. Esta región es la última región periódica o no periódica del fonema, que al ser repetida y concatenada al final del fonema produciría un aumento de la duración. Para los sonidos vocálicos esta región está completamente determinada por el último periodo de la señal, mientras que para los sonidos fricativos corresponde a la región más estable del final del fonema, por lo tanto no es completamente objetiva al poder variar según lo que cada uno considere como estable. Sin embargo, en general no es crítico el punto exacto donde se escoja después que sea suficientemente estable como para no introducir los ruidos propios del transitorio.

Para sonidos explosivos se decidió no cambiar la duración tal como sucede en el caso natural.



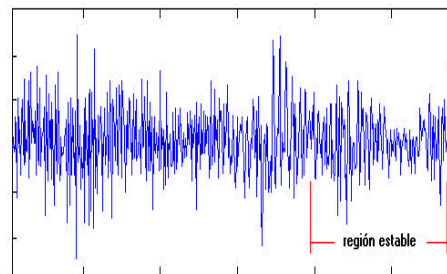
**FIGURA 4**  
**ÚLTIMO PERIODO DEL FONEMA /i**

Para ilustrar estas ideas considere la figura 4 que corresponde al fonema /i/. En la figura está marcado el principio del último periodo de la señal, que es a su vez el más estable de los periodos. Este sería el periodo que se repetiría al final del fonema para aumentar su duración.

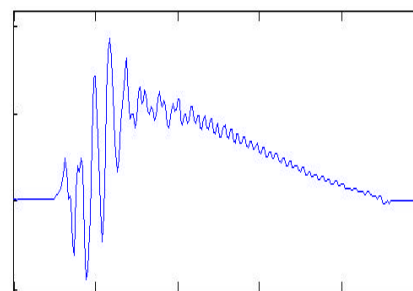
En la figura 5 aparece el fonema /s que es un fricativo y por lo tanto no es periódico, pero sí presenta una región estable que serviría para aumentar su duración.

En la figura 6 se muestra el fonema /k que es un explosivo y por lo tanto no presenta

estado estable, además no tiene sentido cambiar su duración.



**FIGURA 5**  
**FONEMA /s Y SU REGIÓN ESTABLE**



**FIGURA 6**  
**FONEMA /k**

Con los criterios expuestos arriba se terminó de segmentar la base de datos. Si era el primer fonema del difonema y se respetaba la región transitoria entre los dos fonemas los resultados eran buenos; sin embargo aumentar la duración del segundo fonema, no produjo un efecto natural. La razón de este problema era que en el caso natural, cuando se alarga por ejemplo una vocal, se disminuye al mismo tiempo su amplitud y esto inicialmente no lo estaba haciendo el sintetizador. Esta conclusión fue obtenida después de múltiples observaciones de señales de voz no sintetizadas con diferentes duraciones en vocales y sonidos fricativos. Para disminuir la amplitud del fonema se multiplicó por una envolvente con caída dada por una ventana Hamming. La decisión de utilizar una ventana Hamming fue arbitraria porque no se encontró una relación consistente en las caídas de la amplitud de varios ejemplos estudiados. El resultado de la envolvente fue muy bueno mejorando tanto la inteligibilidad como la naturalidad de la síntesis.

## CAMBIO DE LA CURVA MELÓDICA Y DE LA AMPLITUD

Para marcar los acentos e indicaciones de pregunta, afirmación, exclamación etc., se pueden combinar cambios en la frecuencia del tono fundamental y de la amplitud de cada uno de los fonemas.

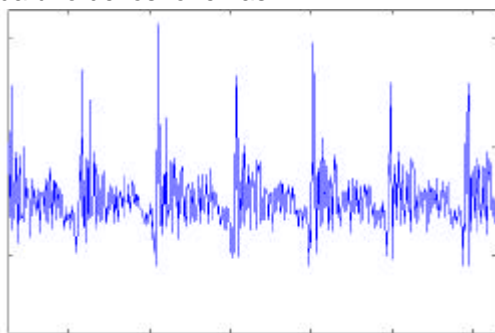


FIGURA 7  
SEÑAL DE ERROR CON UN FILTRO LPC DE ORDEN 18

El cambio de la frecuencia del tono fundamental se realizó sobre la señal de error obtenida de pasar el corpus de voz por el filtro *todo ceros* LPC. La señal error estaría conformada idealmente por pulsos y ruido blanco si los coeficientes del filtro fuesen infinitos, pero como se utilizó un filtro de orden 18 por adaptarse lo suficientemente bien al modelo propuesto[4], la señal de error tiene características como las mostradas en la figura 7.

La frecuencia fundamental de la señal de voz es el inverso del tiempo medido entre dos pulsos consecutivos de la figura 7. Por lo tanto cambiando este tiempo pueden hacerse variaciones de la curva melódica. El método utilizado para el cambio de este periodo, fue el de interpolación ó diezmo de orden cúbico de la señal. Es importante que fuese de orden cúbico porque se debían respetar los bordes de los pulsos que modelan el sonido vocálico de la señal. Si estos pulsos se degeneraran por causa de la interpolación de igual forma lo haría la señal vocálica resultante de pasar la señal de error por el filtro autorregresivo.

El cambio de la amplitud puede ser un recurso adicional para cambiar los acentos de algunas palabras, incluso acentuando sólo

con la amplitud sin cambiar la curva melódica en algunos casos ha dado buenos resultados.

## CONCATENACION

La concatenación es sin duda un punto clave en el resultado final de la síntesis de voz. Una vez se han hecho los cambios de duración, curva melódica y amplitud a cada uno de los difonemas, se debe proceder a concatenarlos.

Para evitar ruidos indeseables debido a los cambios de fase, se tuvo en cuenta mantener la fase en la región de concatenación. La regla seguida fue la de empezar el difonema con concavidad negativa (segunda derivada negativa) y terminarlo con concavidad positiva (segunda derivada positiva). Ambos extremos deben quedar con primera derivada positiva tal como lo muestra la figura 8.

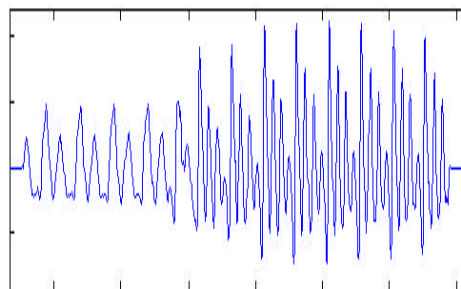


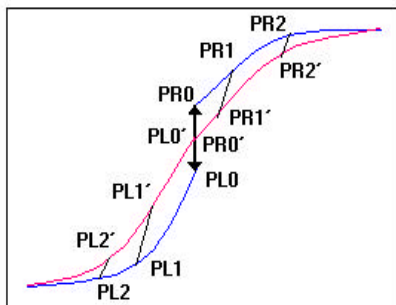
FIGURA 8  
DIFONEMA QUE MUESTRA LA REGLA SEGUIDA PARA MANTENER LA FASE

Adicionalmente para evitar ruidos debido a saltos de amplitud en el punto de concatenación se empezó y terminó el difonema en la región de más baja energía que permitiera la señal digital. Los saltos que pudiesen quedar en el punto de concatenación fueron suavizados repartiendo el error en la región de concatenación siguiendo las siguientes formulas de interpolación:

$$PL_i' = PL_i + (PL_0 - PR_0) * (ML - i) / (2ML)$$
$$PR_i' = PR_i + (PR_0 - PL_0) * (MR - i) / (2MR)$$

Para  $i=0 \dots ML-1$  y  $ML=MR$ , donde  $MR$  y  $ML$  son el número de muestras interpoladas a derecha e izquierda respectivamente.

La figura 9 ilustra el resultado de la interpolación realizada alrededor del punto de concatenación.



**FIGURA 9**  
**REPARTICIÓN DEL SALTO EN LA REGIÓN DE**  
**CONCATENACIÓN**

### OTRAS TÉCNICAS DE SÍNTESIS DE VOZ

Debe ser claro en este momento, que el método de síntesis de voz presentado aquí consiste en la utilización de un filtro conformado por coeficientes calculados por predicción lineal para generar la voz sintetizada, pasando una señal de excitación (la señal de error) conformada idealmente por un tren de pulsos (para los segmentos vocálicos) o ruido (para los segmentos no vocálicos). Es decir el método corresponde al modelo clásico de filtro y fuente de excitación. Sin embargo también existe otro método llamado PSOLA (*Pitch Synchronous Overlap and Add*)[3] que no utiliza este principio clásico. En este modelo se multiplica la señal de voz original (corpus) por ventanas Hamming de longitud igual a dos veces el periodo de la señal de tono, obteniendo una señal que es cero excepto en las regiones localizadas por las ventanas. La concatenación se lleva a cabo suavizando las regiones de concatenación. Para cambiar la frecuencia fundamental del tono se mueven entre sí las señales obtenidas del muestreo con ventana, de tal forma que su espaciamiento refleje el periodo deseado para luego adicionarlas. Para cambiar la duración se suman el número de señales necesarias.

Existen otras técnicas que son variantes del modelo LPC ya que rempazan la excitación de pulsos y ruido por una señal con forma de onda explícita, o por una descomposición espectral de las partes periódicas y no

periódicas de la señal. Existen otras variantes de PSOLA como un híbrido del LPC y PSOLA (LP-PSOLA)[3] donde la operación de suma de las ventanas es realizada con la señal de error obtenida del filtro *todo ceros* LPC.

Mbrola hace uso del método PSOLA pero realizando un re-síntesis del corpus de voz, que consiste en dejar todo el corpus con la misma frecuencia fundamental del tono y evitar desfases de un segmento a otro para que cuando se apliquen las ventanas, no haya corrimientos entre las unidades básicas de sonido que se van a concatenar. Este método es llamado MBR-PSOLA (*Multi-Band Resynthesis PSOLA*)[3].

### III. EVALUACIÓN Y CONCLUSIONES

La evaluación fue realizada con 50 personas utilizando un conjunto de evaluación que cubre todos los dominios de inteligibilidad necesarios para este tipo de pruebas. Adicionalmente se hicieron las mismas pruebas con la voz sintetizada por el programa Mbrola® con el mismo conjunto de parámetros de entrada y la misma base de datos con el fin de poner a ambos sistemas en igualdad de condiciones.

Ninguna de las personas fue entrenada antes de hacer la evaluación y por eso los resultados pueden ser ligeramente inferiores a los que entregaría un grupo entrenado. El entrenamiento es un factor muy importante en inteligibilidad de voces sintetizadas y voces naturales. Por ejemplo cuando se oye en la televisión un programa de España, Argentina u otro país extranjero, es necesario que pasen algunos segundos o incluso minutos para que el oído se adapte a ese acento particular, y en algunos casos después de esto es muy probable que no se entiendan algunas palabras.

Es importante anotar que el proyecto Mbrola es uno de los más destacados en el área de síntesis de voz y que tanto el tiempo como el número de personas a escala internacional que han colaborado con dicho proyecto superan ampliamente a los recursos utilizados en el trabajo aquí realizado.

A continuación se muestran una tabla con los resultados de la prueba de inteligibilidad en la

que se distinguen los dominios de evaluación.

TABLA 1: Resultado de prueba de inteligibilidad

TIPO DE INTELIGIBILIDAD	MBROLA	ESTE TRABAJO
OCCLUSIVAS	94% (750)	90% (750)
VOCALICAS	100% (750)	100% (750)
FRICATIVAS	90% (750)	86% (750)
NASALES	90% (750)	90% (750)
LIQUIDAS	96% (750)	84% (750)
SITIO DEL ACENTO	100% (800)	100% (800)
TOTAL	95%	92%

También se hicieron las pruebas de naturalidad MOS (Mean Opinión Score) cuyos resultados se muestran en la siguiente gráfica.

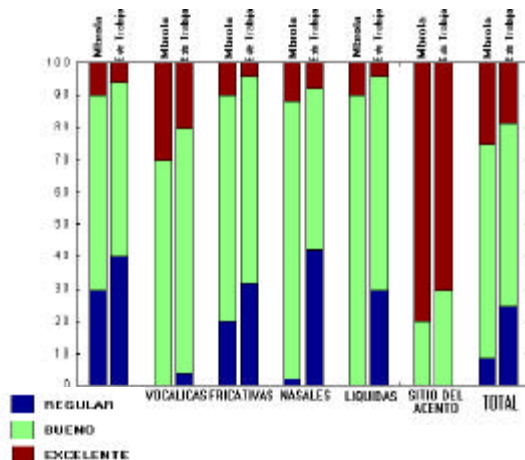


FIGURA 10 Resultados de la prueba de naturalidad

Aunque todas las ideas llevadas a cabo llenaron las expectativas que se tenían al principio del trabajo, también mostraron sus debilidades que abren nuevos horizontes de investigación. Algunas de estas fueron:

Todos los esfuerzos realizados para lograr concatenar los difonemas evitando ruidos indeseables surtieron el efecto esperado pero no contribuyeron lo suficiente en la naturalidad de la voz sintetizada porque

introdujeron algunas veces silencios que no deberían estar presentes. Sería importante en futuros trabajos buscar algoritmos de concatenación que complementen los realizados aquí.

El sintetizador es muy sensible a la calidad del corpus de la base de datos porque utiliza el método de concatenación de difonemas y por eso se hicieron varios contactos para conseguir una base de datos que tuviera buenas características. La que se consiguió es de características aceptables sin embargo existen otras bases de datos excelentes que mejorarían la calidad de la síntesis. Infortunadamente no se contó con la colaboración de sus autores para hacer uso de ellas. Se recomienda a los futuros investigadores en síntesis de voz buscar una base de datos de mejores características que la utilizada aquí, aunque esta base puede ser un buen punto de partida para empezar a investigar.

Algunos aciertos importantes que se demuestran en los resultados del trabajo de investigación son:

El algoritmo encargado de cambiar la duración de los fonemas, adicionando periodos estables a la señal, funcionó muy bien para los sonidos vocálicos. Para los sonidos no vocálicos como los sonidos explosivos (/p/, /b/, /t/, /d/, /k/ y /g/) se les dejó intacta la duración porque solamente es importante la parte transitoria del fonema, lo cual funcionó muy bien. Para ilustrar esto el lector debe tratar de cambiar la duración de la letra *p* en el difonema *pe* y se dará cuenta que no lo va a lograr, sin embargo hacer más larga la letra *e* no es difícil. En el difonema *sa* ambos fonemas pueden tener diferentes duraciones y por lo tanto para los sonidos no vocálicos y no explosivos como la /s/, /f/, /j/ y /ch/ también se repitieron las regiones estables (pero no periódicas por no ser señales con periodicidad) para alargar las señales. En definitiva este método demostró buenos resultados.

El cambio de la curva melódica interpolando la señal de error dio muy buenos resultados. La interpolación fue cúbica con el fin de respetar los bordes y se realizó sobre la señal de error para no afectar las

características impuestas por el filtro formado por el tracto bucal.

Los resultados de la evaluación mostraron que el trabajo realizado aquí, está muy cercano en inteligibilidad al desarrollado en el proyecto Mbrola® (95% vs 92%), lo cual es un verdadero éxito teniendo en cuenta la diferencia entre los recursos utilizados en ellos y el reconocimiento de Mbrola® en el ámbito internacional [4]. Esto motiva mucho a continuar investigando con el fin de perfeccionar la inteligibilidad y mejorar la naturalidad del sistema.

El simulador de síntesis de voz que arrojó este trabajo de investigación es una buena plataforma para seguir investigando en el área porque permite que se aplique con facilidad cualquier nueva idea que pretenda mejorar la calidad de la síntesis de voz, y es sin duda una herramienta versátil porque permite escoger de una forma rápida el mejor conjunto de variables para una síntesis determinada.

## BIBLIOGRAFÍA

- [1] ALLEN JONATHAN. A Perspective on Man-Machine Communication by Speech. Proceedings of the IEEE, VOL. 73, NO. 11, Noviembre 1985.
- [2] CORTES A. RICARDO, ONOFRE MAURICIO. Síntesis de voz por medio digitales. Trabajo de Grado Pontificia Universidad Javeriana, Departamento de Electrónica. 1982.
- [3] DUTOIT THIERRY. An Introduction to Text to Speech Synthesis. Editorial Kluwer Academic Publishers, 1997.
- [4] DUTOIT THIERRY, LEICH HENRI. A comparison of Four Candidate Algorithms in context of High Quality Text to Speech Synthesis (Resumen). Facultad Politécnica de Mons, TCTS-Multitel, 1994
- [5] LEVINSON STEPHEN E., OLIVE JOSEPH P., TSCHIRGI JUDITH S. Speech Synthesis in Telecommunications. IEEE Communication Magazine, Noviembre 1993.
- [6] PARRA CARLOS ALBERTO, INFANTE CARLOS FERNANDO. Conversión semiautomática de texto a descripción fonética en el idioma Español. Anteproyecto

Pontificia Universidad Javeriana, Departamento de Ingeniería Electrónica. 1999.

[7] RANGEL MAX GUILLERMO, RICO TITO MAURICIO. Interfaz datos parlante con capacidad de unión de fonemas en el español. Trabajo de Grado Pontificia Universidad Javeriana, Departamento de Electrónica. 1999.

[8] SPROAT RICHARD. Multilingual Text to Speech Synthesis: The Bell Labs Approach. Editorial Kluwer Academic Publishers, 1997.

[9] VARGA ANDREW, FALLSIDE FRANK. A Technique for Using Multipulse Linear Predictive Speech Synthesis in Text to Speech type Systems. IEEE Transactions on acoustics, Speech, and Signal Processing, Vol. ASSP-35, No. 4, Abril 1987.

[10] VIOLARO FÁBIO, BOEFFARD OLIVIER. A Hybrid Model for Text-to-Speech Synthesis. IEEE Transaction on Speech and Audio Processing, VOL. 6 NO. 5, Septiembre 1998.