CAPÍTULO 2:

MODELO MATEMÁTICO DEL APARATO FONADOR.

En el capítulo anterior se ha explicado cómo es el aparato fonador y cómo funciona desde el punto de vista anatómico y fisiológico, tanto para el adulto como para el neonato. Pero para llegar a diseñar un programa que analice el llanto del recién nacido es necesario tener un modelo matemático de todo lo explicado anteriormente. En este capítulo se procederá el desarrollo de este modelo, llegando a describir un sistema discreto con una entrada y una salida. En la figura 2.1 se puede apreciar el modelo del aparato fonador.

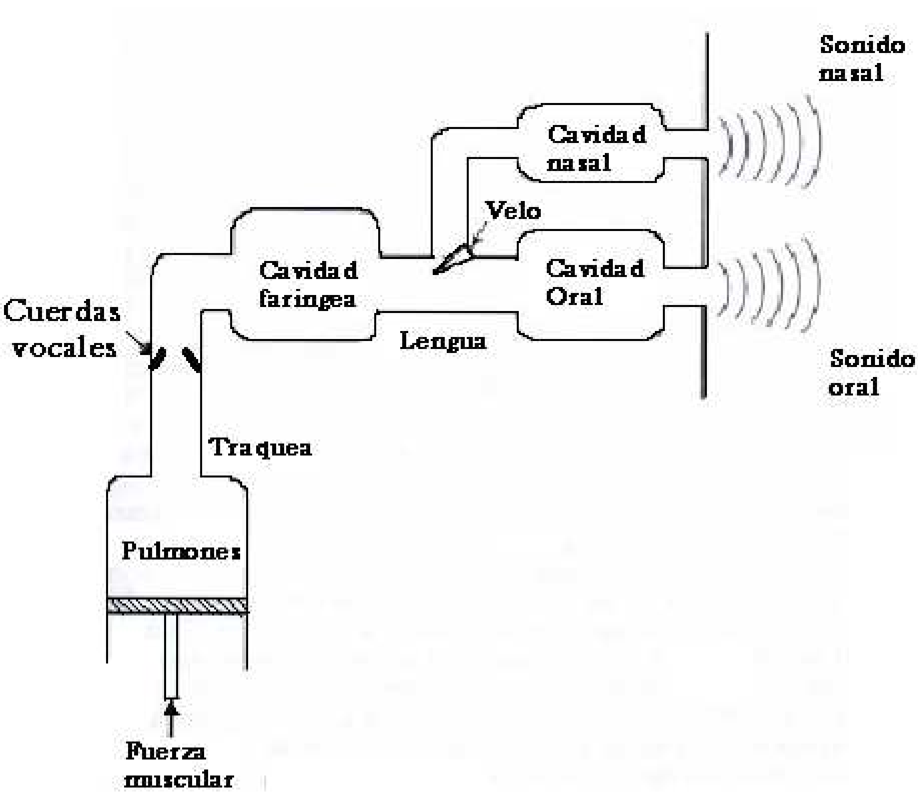

Figura 2.1: Esquema del aparato fonador.

La producción del habla en el ser humano puede ser representado como tres subsistemas independientes: la fuente de excitación $U(z)$, el tracto vocal $H(z)$ y el efecto de la radiación acústica producida por los labios $R(z)$. Por tanto, la salida $S(z)$ de los tres subsistemas se puede definir como [6]:

$$S(z) = U(z)H(z)R(z)$$

A continuación se explicarán cada uno de los subsistemas.

2.1.- Modelo de la excitación.

La fuente de excitación es de dos tipos: las cuerdas vocales, en el caso de un sonido voiced; o el aire al pasar por el tracto vocal sin hacer vibrar las cuerdas vocales, sonido unvoiced. Por tanto, se deberá buscar un modelo matemático para cada uno de los dos casos.

Excitación Voiced

En este caso el aire sale de los pulmones, hace vibrar las cuerdas vocales y pasa luego por el glotis que hace de filtro. Este filtro se puede definir como una función todo-polo con dos polos [17]:

$$G(z) = \frac{1}{(1 - e^{-cP} z^{-1})^2} \qquad (2.1)$$

Esta función se ha comprobado experimentalmente viendo que los resultados obtenidos son muy próximos a los de la realidad. Representada en el tiempo sería:

$$g(n) = [\alpha^n - \beta^n]u(n) \quad , \beta < \alpha < 1, \alpha \approx 1$$

donde $u(n)$ es el escalón unitario. La salida de la glotis se puede representar como la convolución entre esta señal $g(n)$ y un tren de pulsos periódico, con periodo P igual al periodo de las cuerdas vocales:

$$u(n) = e(n) \otimes g(n) = \sum_{i=-\infty}^{\infty} e(n) g(n - iP)$$

Aplicando esto a la ecuación 2.1 se obtiene:

$$S(z) = E(z)G(z)H(z)R(z)$$

La forma de la onda obtenida es periódica, con periodo par a la frecuencia fundamental.

Excitación Unvoiced

En este caso la señal de entrada es muy parecida al ruido y, por tanto, puede ser modulada como ruido blanco $e(n)$. Esta señal de excitación no afecta al espectro de potencia, siendo la densidad espectral de potencia constante. En este caso la ecuación 2.1 sería:

$$S(z) = E(z)H(z)R(z)$$

2.2.- Modelo del tracto vocal.

El modelo que se presenta para el tracto vocal es tan sólo válido cuando la entrada es un tren de pulsos periódicos y para un número de frecuencias y fundamentales dados. Es decir, sólo se puede usar para las partes voiced.

En este modelo sólo se tendrán en cuenta las cavidades faríngea y bucal. No se tienen en cuenta las cavidades laríngea, que está incluida en el modelo de la glotis; y la nasal. De esta manera el modelo será válido para sonidos no nasales.

Las emisiones sonoras producidas por el ser humano son señales no estacionarias, pero se pueden considerar como tal en breves intervalos de tiempo de unos 5 a 20ms. De esta manera se puede considerar el tracto vocal como un filtro variable en el tiempo modificando su respuesta en frecuencia y su espectro. Este filtro puede ser modelado como:

$$H(z) = k\,\frac{z^{-\frac{N}{2}}}{1 - \sum_{i=1}^{N} a_i z^{-i}}$$

donde los $N/2$ ceros en el origen representan el retardo entre salida y la entrada, los N polos representan las frecuencias de resonancia y k es la ganancia que consideraremos unitaria.

Esta forma de representar el tracto vocal es definido como un modelo ARMA (Autoregressive Moving Average) que tiene tanto polos como ceros. Sin embargo en trabajos precedentes a este [18] se ha demostrado que una forma de representación basada sobre un modelo AR (Autoregressive) con un número finito de polos no mete grandes diferencias. Así el tracto vocal podrá ser representado mediante el filtro todo-polo:

$$H(z) = \frac{1}{1 - \sum_{i=1}^{N} a_i z^{-i}}$$

Ahora basta hallar el orden del modelo. Este paso es de gran importancia ya que elegir un modelo con un orden muy alto puede causar división en los picos correspondientes a las formantes; si se toma un valor muy bajo estos picos pueden desaparecer. En los dos casos se tiene pérdida de información. En la siguiente parte del trabajo, donde se habla de la realización del programa, se explica cómo se ha solucionado la elección del orden del modelo.

2.3.-Modelo de los labios.

El modelo para los labios es un filtro digital simple que funciona como un filtro paso alto:

$$L(z) = 1 - z_0 z^{-1} \qquad , z_0 < 1$$

Al meter este filtro se pierde la característica todo-polo explicada anteriormente. Una de las maneras de solucionar esto es decir que este cero en z_0 es anulado con uno de los polos de $G(z)$. En la tesis realizada por D'Aniello [18] se demuestra que la eliminación de este filtro no afecta de manera considerable a los resultados finales, por tanto la ecuación 2.1 quedaría como:

$$S(z) = E(z)G(z)H(z)$$