Lección 1: Modelo del Aparato Vocal

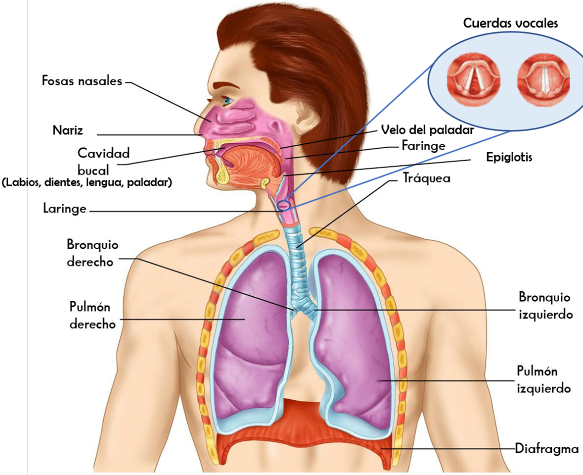
Dr. Jesús Savage
Dr. Carlos Rivera

6 de septiembre de 2021

Índice

Modelo del Aparato Vocal

Un modelo lineal de producción de voz para sonidos vocales fue propuesto por Fant en 1959.



Modelo del Aparato Vocal

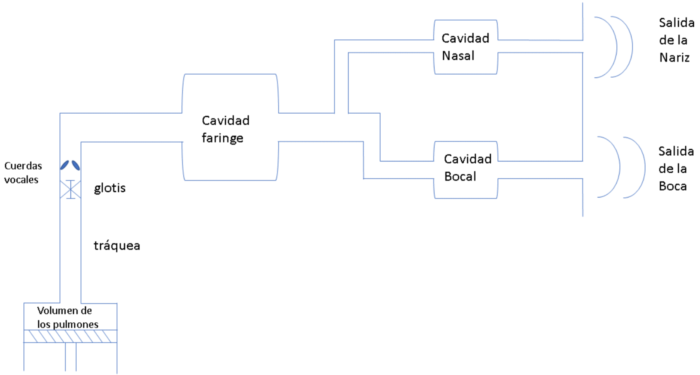


Figura: Modelo del Aparato Vocal.

Modelo del Aparato Vocal

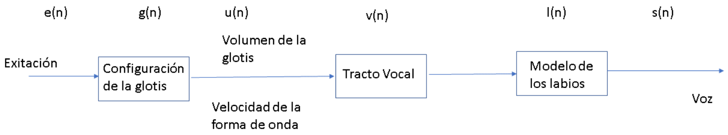
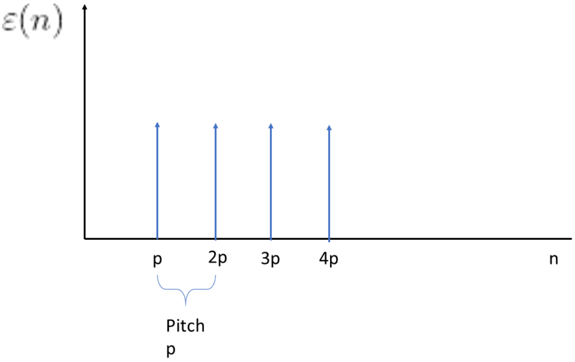


Figura: Diagrama de bloques del aparato vocal.

Modelo de la Excitación $e(n)$

La excitación $e(n)$ es modelada como una serie de impulsos unitarios escalados y espaciados con un período que corresponde al tono (pitch) para sonidos vocales.



Modelo de la Excitación $e(n)$

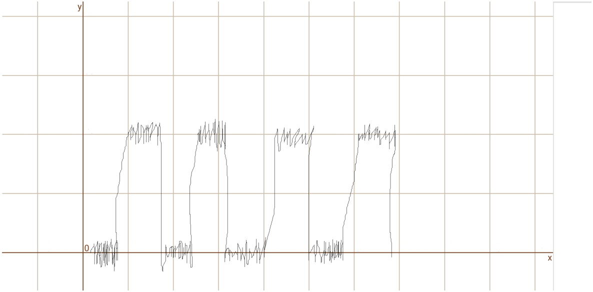


Figura: Gráfica de la vocal A, se observa que es quasiperiodica

Modelo de la Excitación $e(n)$

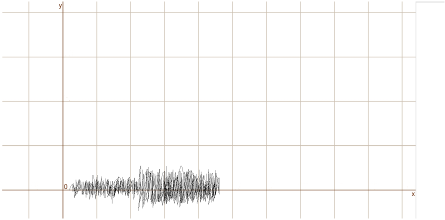


Figura: Gráfica de la consonante S, se observa que es parecido a una señal aleatoria.

Modelo de la Excitación $e(n)$

Para señales tipo vocal:

$$e(n) = \varepsilon_0 \sum_{j=0}^{\infty} \delta(n - j\rho)$$

donde ρ es el tono del parlante o el periodo del tren de pulsos y la delta de Dirac es :

$$\delta(n) = \begin{Bmatrix} 1 & n = 0 \\ 0 & n \neq 0 \end{Bmatrix}$$

La transformada Z de $e(n)$

$$Z\{e(n)\} = \sum_{n=0}^{\infty} e(n) z^{-n}$$

Modelo de la Excitación $e(n)$

$$= \sum_{n=0}^{\infty} \varepsilon_0 (\sum_{j=0}^{\infty} \delta(n - jp)) z^{-n}$$

$$= \varepsilon_0 \sum_{j=0}^{\infty} \sum_{n=0}^{\infty} \delta(n - j\rho) z^{-n}$$

$$= \varepsilon_0 \sum_{j=0}^{\infty} z^{-jp} = \varepsilon_0 \sum_{j=0}^{\infty} (z^{-p})^j$$

Haciendo:

$$A = \sum_{j=0}^{\infty} (z^{-p})^j \qquad (1)$$

Modelo de la Excitación $e(n)$

Multiplicando (1) por z^{-p}:

$$z^{-p} \cdot A = \sum_{j=0}^{\infty} (z^{-p})^{j+1}$$

Haciendo $j + 1 = k$, entonces para $j = 0 \Rightarrow k = 1$

$$= \sum_{k=1}^{\infty} (z^{-p})^{k} = \sum_{j=1}^{\infty} (z^{-p})^{j} \qquad (2)$$

$$(1) - (2) = A - z^{-p} A = \sum_{j=0}^{\infty} (z^{-p})^{j} - \sum_{j=1}^{\infty} (z^{-p})^{j} = 1$$

$$\Rightarrow A = \frac{1}{1 - z^{-p}}$$

Modelo de la Excitación $e(n)$

Entonces la transformada Z de la excitación $e(n)$ es:

$$Z\{e(n)\} = Z\{\varepsilon_0 \sum_{j=0}^{\infty} \delta(n - j\rho)\} = \frac{\varepsilon_0}{1 - z^{-p}}$$

Modelo Cavidad Glotal

La cavidad glotal se modela con la siguiente expresión:

$$g(n) = G_0(n+1)e^{-cnT} = G_0 ne^{-cnT} + G_0 e^{-cnT} \tag{3}$$

La transformada Z de $g(n)$ es:

$$Z\{g(n)\} = Z\{G_0 ne^{-cnT} + G_0 e^{-cnT}\} \tag{4}$$

Separando esta transformada en dos partes, para el segundo termino:

$$Z\{G_0 e^{-cnT}\} = \sum_{n=0}^{\infty} G_0 e^{-cnT} z^{-n} = G_0 \sum_{n=0}^{\infty} (e^{-cT} z^{-1})^n \tag{5}$$

Recordar que

$$\sum_{n=0}^{\infty} A^n = \frac{1}{1-A} \tag{6}$$

Modelo Cavidad Glotal

Por lo tanto:

$$G_0 \sum_{n=0}^{\infty} (e^{-cT} z^{-1})^n = \frac{G_0}{1 - e^{-cT} z^{-1}} \tag{7}$$

Por otra parte la transformada Z para el primer termino de $g(n)$

$$Z\{G_0 n e^{-cnT}\} = G_0 \sum_{n=0}^{\infty} n \cdot e^{-cnT} z^{-n} = A \tag{8}$$

$$= 0 + G_0 \sum_{n=1}^{\infty} n \cdot (e^{-cT} z^{-1})^n$$

Modelo Cavidad Glotal

Haciendo

$$n = k + 1; n = 1 \Rightarrow k = 0$$

$$= G_0 \sum_{k=0}^{\infty} ((k+1) \cdot (e^{-cT} z^{-1})^{k+1}$$

$$= G_0 \sum_{k=0}^{\infty} k \cdot (e^{-cT} z^{-1})^{k+1} + G_0 \sum_{k=0}^{\infty} (e^{-cT} z^{-1})^{k+1}$$

$$= G_0 (e^{-cT} z^{-1}) \cdot \sum_{k=0}^{\infty} k \cdot (e^{-cT} z^{-1})^{k} + G_0 (e^{-cT} z^{-1}) \cdot \sum_{k=0}^{\infty} (e^{-cT} z^{-1})^{k} \quad (9)$$

Modelo Cavidad Glotal

De la definición de A en (9) entonces (10):

$$\Rightarrow A = (e^{-cT}z^{-1}) \cdot A + \frac{G_0(e^{-cT}z^{-1})}{1 - e^{-cT}z^{-1}}$$

$$A(1 - (e^{-cT}z^{-1})) = \frac{G_0(e^{-cT}z^{-1})}{1 - e^{-cT}z^{-1}}$$

$$A = \frac{G_0(e^{-cT}z^{-1})}{(1 - (e^{-cT}z^{-1}))^2}$$

Entonces la transformada Z en la expresión (5)

$$G(z) = A + \frac{G_0}{1 - e^{-cT}z^{-1}}$$

Modelo Cavidad Glotal

$$= \frac{G_0(e^{-cT}z^{-1})}{(1-(e^{-cT}z^{-1}))^2} + \frac{G_0}{1-e^{-cT}z^{-1}}$$

$$= \frac{G_0(e^{-cT}z^{-1}) + G_0(1-e^{-cT}z^{-1})}{(1-e^{-cT}z^{-1})^2}$$

$$= \frac{G_0}{(1-e^{-cT}z^{-1})^2}$$

Tracto Vocal

La transformada Z del tracto vocal:

$$V(z) = \prod_{i=1}^{k} \frac{1}{(1 - z^{-1} z_i)(1 - z^{-1} z_i^*)}$$

Efecto de los Labios (Lip Radiateur):

$$L(z) = L_0(1 - z^{-1})$$

Modelo total

$$S(z) = L(z)V(z)G(Z)e(Z) \tag{10}$$

$$= \frac{L_0 G_0 E_0 [1 - z^{-1}][1 - e^{-cT} z^{-1}]^{-2}[1 - z^{-P}]^{-1}}{\prod_{i=1}^{k}(1 - z^{-1} z_i)(1 - z^{-1} z_i^*)} \tag{11}$$

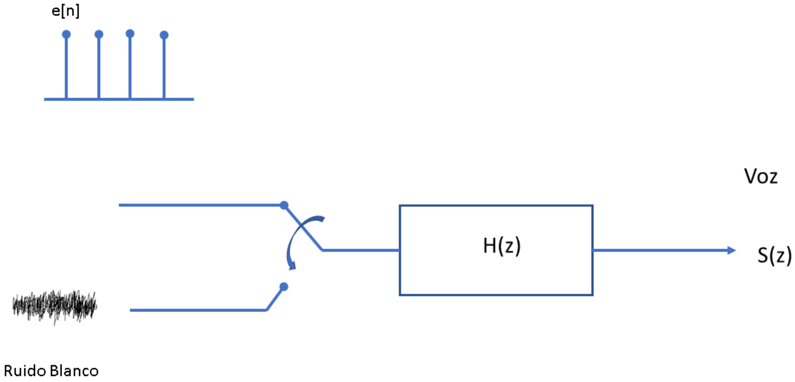
Tracto Vocal

$[1 - z^{-1}]$ se cancela con $[1 - e^{-cT} z^{-1}]^{-1}$
$[1 - e^{-cT} z^{-1}]^{-1}$ se incluye en el producto del denominador

$$S(z) = \frac{L_0 G_0 E_0 [1 - z^{-P}]^{-1}}{\prod_{i=1}^{k} (1 - z^{-1} z_i)(1 - z^{-1} z_i^*)}$$

$$= \frac{\sigma \cdot E(z)}{\sum_{i=0}^{M} a_i \cdot z^{-i}} = E(z) H(z)$$

Tracto Vocal



Tracto Vocal

$$H(z) = \frac{\sigma}{\sum_{i=0}^{M} a_i \cdot z^{-i}}$$

$$S(z) = \frac{\sigma \cdot E(z)}{\sum_{i=0}^{M} a_i \cdot z^{-i}}$$

$$S(z)\sum_{i=0}^{M} a_i \cdot z^{-i} = S(z)a_0 + S(z)a_1 z^{-1} + ... + S(z)a_m z^{-m} = \sigma \cdot E(z)$$

Si $a_0 = 1$

$$S(z) = \sigma \cdot E(z) - S(z)a_1 z^{-1} - ... - S(z)a_m z^{-m}$$

Tracto Vocal

Obteniendo la transformada Z inversa de $S(z)$:

$$Z^{-1}\{S(z)\} = s[n] = \sigma e[n] - a_1 s[n-1] - a_2 s[n-2] \cdots - a_m s[n-m]$$

Dada una señal de voz el objetivo es encontrar los siguientes componentes del modelo para cada bloque de la señal j que representan ésta:
$\underline{a}^j, \sigma_j, e^j[n]$
En donde el vector para el bloque j esta definido como:

$$\underline{a}^j = [1 \; a_1^j \; a_2^j \; ... \; a_m^j \;]$$