Research Journal of Applied Sciences 2 (9): 984-987, 2007
ISSN: 1815-932X

Spectral Analysis of Voiced Sounds Tracheotomisian Case

H. Maalem
Department of Electronics, University of Skikda, Bp 26, Skikda 21000 Algeria,
A.I.L Guelma, Algeria

Abstract: This study presents the cepstral and spectral analysis of a speech signal produced by a woman affected by a tracheotomy, with and without phonation valve. The spectral analysis based on parametric method (Autoregressive AR) using the fourth order cumulant, these analyses are used to detect and compare the principals characteristics of the speech signal: pitches, the formants and the intensity (energy) of corresponding voiced sounds (example: vowel \i\ and \u\). The first results appear promising, since it seems, after several experiments there is no deformation of the spectrum as one could have supposed at the beginning, however this pathology influenced the pitch value, in addition the phonation valve influenced the energy value.

Key words: Spectrum, cepstrum, cumulant, tracheotomy, phonation valve

INTRODUCTION

Tracheotomy is a surgical procedure used to cut a hole in the trachea through which a small tube is inserted. Trachetomy is indicated for a variety causes of breathing difficulties in which the creation of a new breathing pathway is required, by-passing the nose, mouth and throat. A tracheotomy is usually considered when an Endotracheal (ET) tube (a tube that goes into the throat through the mouth) either will not be effective (in some emergency situations for example) or would be required for a long time. Sometimes, a tracheotomy is performed when an ET tube cannot be placed due to narrowing of the windpipe or blockage of the voice box (larynx).

The reasons for performing a tracheotomy, generally fall into three major categories (http://www.pediatric-ent.com/learning/surgeries/tracheotomy.htm):

- To by pass an obstruction in the airway (most common reason).
- To help with long term ventilation in patients who cannot do this on their own (patients with respiratory muscle problems or lung problems).
- To provide a temporary airway while reconstructive surgery is performed that may cause breathing problems.

Tracheotomy preserves the entirety of the phonation organs, the hole of tracheotomy would not remain empty, a tracheal cannula is placed without the anatomical structures modification, breathing is done by the air ways but arises also by the tracheal opening, this latter creates a difficulty in speaking; the air coming from the lungs arises directly by the opening without vibrating the vocal cords allowing thereby the speech. So, it is necessary to stop this air escape by using a special stopper called speaking stopper or phonation valve.

One purpose here is to integrate the spectral analysis based on fourth order cumulant and cepstral analysis of Arabic speech signal produced by a woman affected by a tracheotomy, with and without phonation valve; to show the advantage of phonation valve, to extract and compare the parameters which make it possible to describe its Principal characteristics.

DEFINITIONS

Fourth order statistics: If x(n), n=0, 1, 2, 3 ….. is a real stationary discrete time signal and its moments up to order p exist, then its p order moment function is given by (Fackrell and McLaughlin, 1994; Mendel, 1991).

$$m_p(\tau_1, \tau_2, \tau_3, ...\tau_{p-1})] = E\left\{\begin{matrix} x(n)x(n+\tau_1)x(n+\tau_2) \\ x(n+\tau_3).....x(n+\tau_{p-1}) \end{matrix}\right\} \quad (1)$$

and depends only the time differences τ_i for all i, here E{.} denotes statistical expectation, if in addition the signal has zero mean, then its 2nd and 4th order Cumulant function are given by:

2_{nd} order cumulant:

$$C_{2,x}(\tau) = m_{2,x}(\tau) \quad (2)$$

4_{th} order cumulant:

$$C_{4,x}(\tau_1,\tau_2,\tau_3) = E[x(n)x(n+\tau_1)x(n+\tau_2)x(n+\tau_3)] - C_{2,x}(\tau_1)C_{2,x}(\tau_2-\tau_3) - C_{2,x}(\tau_2)C_{2,x}(\tau_3-\tau_1) - C_{2,x}(\tau_3)C_{2,x}(\tau_1-\tau_2) \quad (3)$$

For non zero mean signal, it is enough to replace x(n) by x(n)-E[x(n)].

We give below some of the relevant properties of the 4th order cumulant, details can be found in Mendel (1991).

- The 4th order cumulant of a Gaussian process are identically zero.
- If Z(n)=x(n)+y(n), where x(n) is statistically independent of y(n), then the 4th order cumulant of z(n) is the sum of the 4th order cumulant of x(n) and y(n).

$$C_{4,z}(i,j,k) = C_{4,x}(i,j,k) + C_{4,y}(i,j,k) \quad (4)$$

- If Z(n) = x(n)+y(n), with y(n) is Gaussian then:

$$C_{2,z}(i) = C_{2,x}(i) + C_{2,y}(i) \quad (5)$$

$$C_{4,z}(i,j,k) = C_{4,x}(i,j,k) \quad (6)$$

Note that the effects of the Gaussian noise have been suppressed when the fourth order cumulant is used.

Cepstrum: The cepstrum is defined as the inverse transform of the logarithm of the absolute value of the Fourier transform of the signal (Kadri, 2000)

$$C_R = \int_{-\infty}^{\infty} \mathrm{Ln}|X(f)| e^{-j2\pi fn} df \quad (7)$$

Linear speech production model: Many different models have been postulated for quantitatively describing certain factors involved in the speech process, one of the most used models of acoustical speech behaviour is the linear speech production model developed by Fant (Marker and Gray, 1976). The speech production model (AR) is show in Fig. 1.

The transfer function H (z) is then:

$$H(z) = \frac{\sigma}{A(z)} = \frac{\sigma}{1 + \sum_{i=1}^{p} a(i) z^{-i}} \quad (8)$$

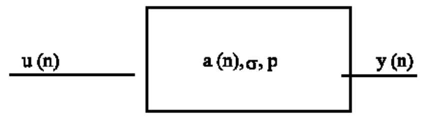

Fig. 1: Linear speech production model

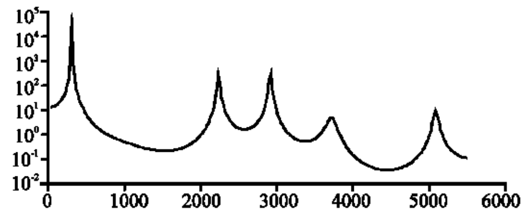

Fig. 2: Spstrum of vowel \i\ of a Tracheotomysan woman with phonation valve

Where, σ represents the system gain and a (i) the AR coefficients.

Many ways exist for determining the AR coefficients, one of which is based on the following normal equation

$$\sum_{k=1}^{p} a(k) r_y(k) = r_y(0) \quad (9)$$

Where, r_y is the autocorrelation function.

It has been shown (Mendel, 1991) that the AR coefficient can also be determined using 4th order cumulant, giving the following equation:

$$\sum_{k=1}^{p} a(k) C_{4,y}(\tau - k, k_0, 0) = C_{4,y}(\tau, k_0, 0) \quad (10)$$

With:

$$\tau = 1,\ 2, \ldots,\ p+n,\ n \geq 0 \text{ and } k_0 = -p,\ -p+1, \ldots,\ 0$$

RESULTS

The signals used in the application consist of sevral voiced sounds of 512 samples' sections for each one (example: The Arabic vowels \i\ and \u\) pronounced by a women of 43 years old affected by a tracheotomy, firstly with phonation valve and secondly without phonation valve.

First, estimation computes the spectrum of these vowels using the Autoregressive (AR) of order 12 based on the fourth order cumulant (Fig. 2-5). Contemplating these different figures, one notices that these spectra are the same from the formantic distribution and the number of the formants point of view.

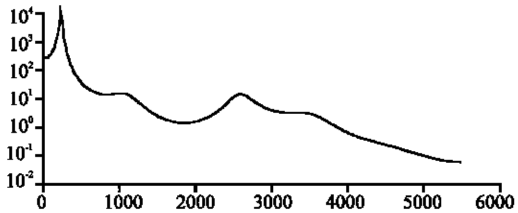

Fig. 3: Spstrum of vowel \i\ of a Tracheotomysan woman without phonation valve

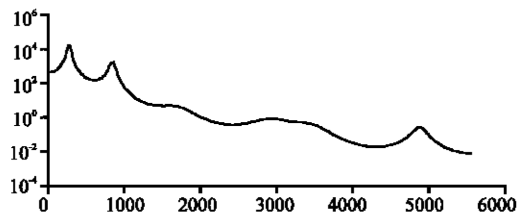

Fig. 4: Spstrum of vowel \u\ of a Tracheotomysan woman without phonation valve

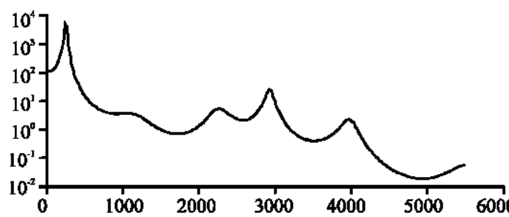

Fig. 5: Spstrum of vowel \u\ of a Tracheotomysan woman with phonation valve

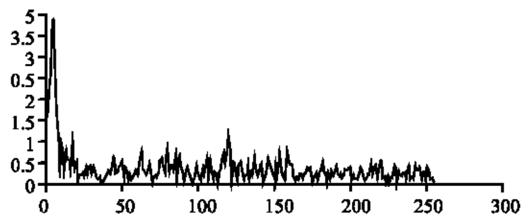

Fig. 6: Cepstrum of vowel \i\ of a Tracheotomysan woman without phonation valve

The Arabic vowels are voiced sounds then, they are characterized by their fundamental frequency (pitch). The frequential method (cepstrum) is used for pitch detection (Fig. 6-9).

The values of the first three formants, pitch value and the maximum energy value corresponding to the maximum spectrum magnitude for each vowel, produced by the woman in the two cases, are given respectively in Table 1 and 2.

It appears, after several experiments that:

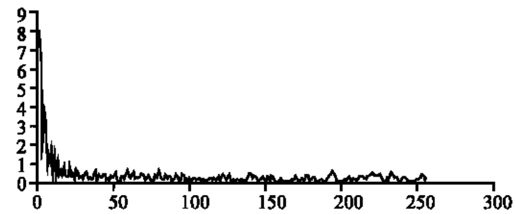

Fig. 7: Cepstrum of vowel \i\ of a Tracheotomysan woman with phonation valve

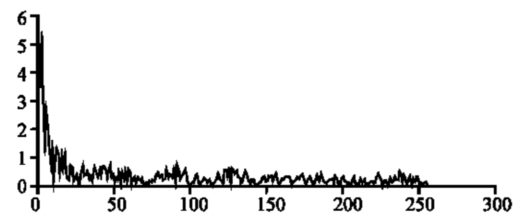

Fig. 8: Cepstrum of vowel \u\ of a Tracheotomysan woman with phonation valve

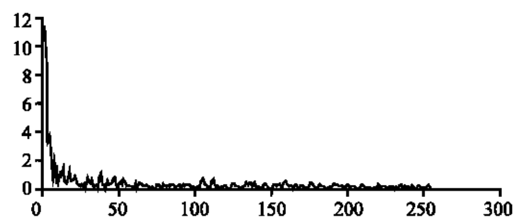

Fig. 9: Cepstrum of vowel \u\ of a Tracheotomysan woman without phonation valve

Table 1: The first three formants (Hz) values, the pitch (Hz) and the maximum energy (dB) of the vowel \i\ of a Tracheotomysan woman with and without phonation valve

	F1	F2	F3	Pitch	Max energy
With phonation valve	301	2239	2928	649	44860
Without phonation valve	236	2605	3359	525	8233

Table 2: The first three formants (Hz) values, the pitch (Hz) and the maximum energy (dB) of the vowel \u\ of a Tracheotomysan woman with and without phonation valve

	F1	F2	F3	Pitch	Max energy
With phonation valve	258	841	2891	649	14535
Without phonation valve	237	1057	2934	649	4596

- Tracheotomy does not alter the values of the formants.
- The pitch value of the affected woman by a Tracheotomy does not correspond to that of a normal woman that locates in the interval of [180Hz at 450 Hz] (faculties/arts/linguistics/russell/138/sec5/phonatio.htm).

- Phonation valve alter the energy value, the latter decreases in the phonation valve absence case.

CONCLUSION

It has been shown, although many more tests have to be made, that the pitch value of woman affected by a Tracheotomy is different from that of a normal woman. The first three formants values, however, are located in the same interval of the normal women.

Thus, in addition to the healthy advantages of the phonation valve, it increases indeed the necessary energy for phonation.

REFERENCES

Pediatric Otolaryngology: Learning Center-Tracheotomy URL: http://www.pediatric-ent.com/learning/surgeries/tracheotomy.htm

Fackrell, J.W.A. and S. McLaughlin, 1994. H.O.S. of speech signals. IEEE. Coll. Technology in Speech Signal Proceeding, London.

Mendel, J., 1991. Tutorial on higher-order statistics. Proc. IEEE., 79: 277-303.

Kadri, L., 2000. Spectral Analysis of arabic speech application of child with defective audition. Memory of master, University of Constantine, Algeria.

Markel, J. and A. Gray, 1976. Linear prediction of speech. Springer-Verlag.

<< Phonation >> URL: http://www.umanitoba.ca/faculties/arts/linguistics/russell/138/sec5/ phonatio.htm