Abstract—This paper presents the cepstral and trispectral analysis of a speech signal produced by normal men, men with defective audition (deaf, deep deaf) and others affected by tracheotomy, the trispectral analysis based on parametric methods (Autoregressive AR) using the fourth order cumulant. These analyses are used to detect and compare the pitches and the formants of corresponding voiced sounds (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 it at the beginning, however these pathologies influenced the two characteristics: The defective audition influences to the formants contrary to the tracheotomy, which influences the fundamental frequency (pitch).

Keywords—Cepstrum, cumulant, defective audition, tracheotomisy, trispectrum.

I. INTRODUCTION

The performances of vocal signal decrease when corrupted by noise; in this direction, many studies were undertaken. The principal characteristic of the speech signal used is energy, but can be associated for example to the formants and the Fundamental frequency (pitch).

The observation that the vocal signal distribution is non-Gaussian led to various studies to consider the spectral analysis based on higher order statistics or Cumulant (H.O.S) for the speech signal detection. The idea of using H.O.S for speech processing is based on exploiting the Gaussian suppression and phase preservation properties of these statistics. In addition, the H.O.S of speech signals have distinctive features that may be exploited to lead to a better estimation and a more accurate discrimination between various speech types [1].

One purpose here, is to integrate the trispectral analysis based on the fourth order cumulant and cepstral analysis of Arabic speech signal produced by normal men, men with defective audition (deaf, deep deaf) and others affected by tracheotomy; to Extract the parameters which make it possible to describe its Principal characteristics.

II. DEFINITIONS

A. 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 th order moment function is given by [2]-[3].

\[ m_p(\tau_1, \tau_2, \tau_3, ..., \tau_{p-1}) = E[x(n) x(n+\tau_1) x(n+\tau_2) x(n+\tau_3) ... x(n+\tau_{p-1})] \] (1)

and depends only on the time differences \( \tau_i \) for all i, here \( E \) denotes statistical expectation, if in addition the signal has zero mean, then its 2 nd and 4 th order Cumulant function are given by:

2 nd order cumulant:
\[ C_{2,x}(\tau) = m_{2,x}(\tau) \] (2)

4 th order cumulant:
\[ C_{4,x}(\tau_1, \tau_2, \tau_3) = E [x(n) x(n+\tau_1) x(n+\tau_2) x(n+\tau_3)] - C_{2,x}(\tau_1) C_{2,x}(\tau_2) C_{2,x}(\tau_3) - C_{2,2}(\tau_1, \tau_2, \tau_3) \] (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 4 th order cumulant, details can be found in [3].

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

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

- If \( \tilde{y}(n) = x(n) + y(n) \), with y(n) is Gaussian then:

\[ C_{4,x}(i) = C_{2,x}(i) + C_{2,y}(i) \] (5)

\[ C_{4,y}(i, j, k) = C_{4,z}(i, j, k) \] (6)

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

B. Trispectrum

Like the PSD, the Trispectrum can be estimated using two methods: non-parametric methods (direct and indirect), and parametric methods.

- Non parametric methods
- Direct method
The measured signal, $x(n)$ $n=0,1,..,N$ is divided into $p$ segments, $x_i(n)$ $i=0,1,...,p-1$. Each segment can be first multiplied by a window and then its M-point discrete Fourier transform (DFT) $X_k(l)$, $k=0,1,...,M-1$, is computed. The trispectrum is, given by:

$$ T(k,l,m) = \frac{1}{p} \sum_{i=0}^{p-1} X_k(l)X_i(m)X_i^*(k+l+m) $$

(7)

- **Indirect method**

Estimation of the fourth order cumulants is first computed, and then 3D DFTs of this cumulant is computed after appropriate windowing [2]-[4].

$$ T(f_1,f_2,f_3) = \sum_{m,n=0}^{M-1} C_4(m,n,l)e^{-j2\pi(mf_1+nf_2+lf_3)} $$

(8)

- **Parametric method**

If $x(n)$ is a linear process, $x(n) = \sum_i h(i)u(n-i)$ and $u(n)$ is some white noise, then the trispectrum is, given by [3]-[5]:

$$ T(f_1,f_2,f_3) = c_{4}^{4}(0,0,0,0,0) \cdot H(f_1) \cdot H(f_2) $$

$$ H(f_1) \cdot H^*(f_1+f_2+f_3) $$

(9)

Where: $H(.)$ is the Fourier transform of the filter impulse response $h(n)$, $C_4$ represents the cumulant of order 4.

**C. Cepstrum**

The cepstrum is defined as the inverse transform of the logarithm of the absolute value of the Fourier transform of the signal.

$$ C_R = \int_{-\infty}^{\infty} \ln |X(f)| e^{-j2\pi f n} df $$

(10)

**D. 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 [6]. The speech production model (AR) is shown in Fig. 1.

![Fig. 1 Linear speech production model](image)

The transfer function $H(z)$ is then:

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

(11)

Where $\sigma$ 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 [3]:

$$ \sum_{k=0}^{p} a(k)r_{y}(k) = r_{y}(0) $$

(12)

Where, $r_y$ is the autocorrelation function.

It has been shown that the AR coefficient can also be determined using $4^{th}$ order cumulant, giving the following equation [3]:

$$ \sum_{k=0}^{p} a(k)C_{a_{y}}(\tau - k,0) = C_{a_{y}}(\tau,0) $$

(13)

With: $\tau = 1,2,...,p+n$, $n \geq 0$ and $k_0 = -p,-p+1,...,0$

**E. Tracheotomy**

The larynx is essential for breathing and vocalizing. When damaged, an emergency tracheotomy may be necessary to ensure survival. Vocal cord paralysis is the second most common congenital irregularity of the larynx, The most common treatment for laryngeal paralysis is a tracheotomy [7].

- **What is a tracheotomy?**

The TRACHEA is the part of the AIRWAY (or breathing passage) commonly known as the windpipe! A tracheotomy is a surgical procedure that creates a temporary opening in the trachea. The hole itself is called a tracheotomy. The tube that is placed through this hole is called a tracheotomy tube [7].

- **What are the indications for a tracheotomy?**

A tracheotomy is a temporary or permanent treatment for a variety of 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 in the throat through the mouth) either will not be effective (in some emergencies 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 [7]:

- 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

**III. EXPERIMENTAL RESULTS**

The signals used in the application consist of 512 samples sections of several voiced sounds for each one (example: the Arabic vowels \a ,\i \,\u ) pronounced, firstly by a normal man, a man with defective audition (deaf, deep deaf) and another affected by tracheotomy. First, estimation computes the trispectrum of these vowels using the Autoregressive (AR) based on the fourth order cumulant. The spectra obtained by the second horizontal slices of the trispectra are shown in figures: 2, 3, 4, 5, 10, 11, 12, 13, 18, 19, 20 and 21. Contemplating these different figures, one notices that these
spectra are the same formantic distribution and the number of
the formants from point of view.

The Arabic vowels are voiced sounds then, they are
characterized by its fundamental frequency (pitch). The
frequential method (cepstrum) is used for pitch detection. See
figures (6, 7, 8, 14, 15, 16, 17, 22, 23, 24 and 25).

The values of the first three formants and the pitch value of
the vowels produced by the four men are given in tables I, II,
III, IV, V and VI respectively.

It appears, after several experiments that:
- In spite of the audition deficiency of the deaf man or deaf
depth the fundamental frequency (pitch) of its voice is normal
as that of a tending man who locates in the interval of [60Hz
at 250 Hz].
- The value of the first formant of the vowel \( \bar{a}, \bar{i}, \text{ and } \bar{u} \)
of a deaf man does not correspond to that obtained by
the normal man who is located between \([600 \text{ and } 750 \text{ Hz}]\), \([250 \text{ and } 400 \text{ Hz}]\) and \([200 \text{ and } 450 \text{ Hz}]\) respectively.
- The value of the second formant of the vowel \( \bar{a}, \bar{i}, \text{ and } \bar{u} \)
of a deaf man does not correspond to that obtained by
the normal man who is located between \([1150 \text{ and } 1400 \text{ Hz}]\), \([2100 \text{ and } 2700 \text{ Hz}]\) and \([1000 \text{ and } 1400 \text{ Hz}]\) respectively.
- Tracheotomy does not alter the values of the formants.
- The value of the pitch of the man affected by
Tracheotomy does not correspond to that of a normal man.

### TABLE I

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal man</td>
<td>689</td>
<td>1270.5</td>
<td>2734</td>
</tr>
<tr>
<td>deaf man</td>
<td>366.1</td>
<td>1227.4</td>
<td>2519.4</td>
</tr>
<tr>
<td>deep deaf</td>
<td>689.1</td>
<td>1808</td>
<td>2476.3</td>
</tr>
<tr>
<td>Tracheotomysan man</td>
<td>624</td>
<td>1227</td>
<td>2562</td>
</tr>
</tbody>
</table>

### TABLE II

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal man</td>
<td>344</td>
<td>2670</td>
<td>3380</td>
</tr>
<tr>
<td>deaf man</td>
<td>559</td>
<td>2261</td>
<td>2691</td>
</tr>
<tr>
<td>deep deaf</td>
<td>279</td>
<td>2885</td>
<td>3337</td>
</tr>
<tr>
<td>Tracheotomysan man</td>
<td>344</td>
<td>2477</td>
<td>3100</td>
</tr>
</tbody>
</table>

### TABLE III

<table>
<thead>
<tr>
<th></th>
<th>F1</th>
<th>F2</th>
<th>F3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal man</td>
<td>430</td>
<td>1012</td>
<td>22433</td>
</tr>
<tr>
<td>deaf man</td>
<td>516</td>
<td>1378</td>
<td>2670</td>
</tr>
<tr>
<td>deep deaf</td>
<td>323</td>
<td>1722</td>
<td>2325</td>
</tr>
<tr>
<td>Tracheotomysan man</td>
<td>215</td>
<td>1227</td>
<td>2476</td>
</tr>
</tbody>
</table>

### TABLE IV

<table>
<thead>
<tr>
<th></th>
<th>Pitch (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal man</td>
<td>200.25</td>
</tr>
<tr>
<td>deaf man</td>
<td>200</td>
</tr>
<tr>
<td>deep deaf</td>
<td>239</td>
</tr>
<tr>
<td>Tracheotomysan man</td>
<td>482</td>
</tr>
</tbody>
</table>

### TABLE V

<table>
<thead>
<tr>
<th></th>
<th>Pitch (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal man</td>
<td>189</td>
</tr>
<tr>
<td>deaf man</td>
<td>225</td>
</tr>
<tr>
<td>deep deaf</td>
<td>157</td>
</tr>
<tr>
<td>Tracheotomysan man</td>
<td>367</td>
</tr>
</tbody>
</table>

### TABLE VI

<table>
<thead>
<tr>
<th></th>
<th>Pitch (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal man</td>
<td>143</td>
</tr>
<tr>
<td>deaf man</td>
<td>225</td>
</tr>
<tr>
<td>deep deaf</td>
<td>167</td>
</tr>
<tr>
<td>Tracheotomysan man</td>
<td>380</td>
</tr>
</tbody>
</table>

### IV. CONCLUSION

It has been shown that the trispectral analysis can be used
for the detection and the recognition of the speech signal.

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

as well as the values of the first and the second formants of
men with defective audition (deaf, deep deaf) respectively is
different from that of men with normal audition contrary to
the value of the pitch which is the same.

### REFERENCES

Fig. 2  Spectrum of vowel \( \text{a} \) of a normal man

Fig. 3  Spectrum of vowel \( \text{a} \) of a deaf man

Fig. 4  Spectrum of vowel \( \text{a} \) of a deep deaf man

Fig. 5  Spectrum of vowel \( \text{a} \) of a Tracheotomysan man

Fig. 6  Cepstrum of vowel \( \text{a} \) of a normal man

Fig. 7  Cepstrum of vowel \( \text{a} \) of a deaf man

Fig. 8  Cepstrum of vowel \( \text{a} \) of a deep deaf man

Fig. 9  Cepstrum of vowel \( \text{a} \) of a Tracheotomysan man

Fig. 10  Spectrum of vowel \( \text{i} \) of a normal man

Fig. 11 Spectrum of vowel \( \text{i} \) of a deaf man

Fig. 12 Spectrum of vowel \( \text{i} \) of a deep deaf man

Fig. 13  Spectrum of vowel \( \text{i} \) of a Tracheotomysan man
Fig. 14 Cepstrum of vowel \( \hat{a} \) of a normal man

Fig. 15 Cepstrum of vowel \( \hat{a} \) of a deaf man

Fig. 16 Cepstrum of vowel \( \hat{a} \) of a deep deaf man

Fig. 17 Cepstrum of vowel \( \hat{a} \) of a Tracheotomysan man

Fig. 18 Spectrum of vowel \( \hat{a} \) of a normal man

Fig. 19 Spectrum of vowel \( \hat{a} \) of a deaf man

Fig. 20 Spectrum of vowel \( \hat{a} \) of a deep deaf man

Fig. 21 Spectrum of vowel \( \hat{a} \) of a Tracheotomysan man

Fig. 22 Cepstrum of vowel \( \hat{a} \) of a normal man

Fig. 23 Cepstrum of vowel \( \hat{a} \) of a deaf man

Fig. 24 Cepstrum of vowel \( \hat{a} \) of a deep deaf man

Fig. 25 Cepstrum of vowel \( \hat{a} \) of a Tracheotomysan man