Performances Study of a New Speech Coding Strategy with Reduced Channels for Cochlear Implants

Bouafif Lamia*, Ouni Kais and Ellouze Noureddine

Signal Processing Laboratory, ENIT, BP 37 Tunis Belvedère, Tunisia

Abstract: A new cochlear implant speech coding strategy for representing acoustic information with reduced number of channels will be developed. After a brief description of the actual cochlear implant stimulation methods, the authors present a processing algorithm describing the new adaptive spectral analysis strategy (ASAS-GC). This technique is based on a Gammachirp perception model and an optimal stimulation rate selection. The excited electrodes signals and recognition performances are presented and compared with those of others strategies.

1. INTRODUCTION

A cochlear implant (CI) is an electronic device which is used to provide hearing aid to patients affected by severe hearing loss. The system consists of an external processor and an internal element that is implanted into the patient by means of a surgical operation. The implanted element has electrode array which is allocated into the cochlea, in order to provide stimulation of the auditory nerve by means of electrical stimuli [1].

Today there is enough experience with this treatment and we estimate that there are about 50.000 implanted patients around the world. The first successful multi-channel cochlear implant was implanted in 1978. Early signal processing devices extracted the second formant (F2) and pitch (F0) to control electrode stimulation. The frequency of F2 controlled the location of electrode stimulation, and F0 controlled the stimulation rate. Improvements were made by also extracting the first formant (F1) and adding a second stimulated electrode for each pitch period. The MULTIPEAK stimulation strategy added stimulation of a number of fixed electrodes to better represent high-frequency information. The next stages of development were the Spectral Maxima Sound Processor (SMSP) strategy, described in Australian Patent, and SPEAK strategy, described in U.S. Patent. Another fixed-rate strategy, called CIS, stimulated all of a small number of electrodes to represent the sound spectra.

More recently, the Multi-rate cochlear stimulation strategy has been developed that does not use fixed stimulation rates. The strategy determines rate of stimulation for each implant electrode by measuring average intervals between positive zero-crossings of the filtered signals for each band. Finally, in 1999, The Spike-based Temporal Auditory Representation (STAR) strategy improved speech perception ability in the presence of noise, including both pseudo-stationary random noise [2]. Adaptation effects improve detection of transient events, such as plosive onsets and formant transitions, and help to overcome pseudo-stationary

background noise by a neural adaptation process similar to spectral subtraction.

2. CIMPLANT FUNCTIONING

In a CI system, the external box is composed of a microphone, an antenna and a processor (Fig. **1**). The processor analyzes the sound and determines the stimulation level to be sent at each electrode. The stimulation pattern is sent into the internal part of the system by a radio transmission, and the internal part generates the electrical pulses, that are presented at each intra-cochlear electrode of the implant. The pulses at each electrode cause the activation of the auditory nerve providing a hearing sensation.
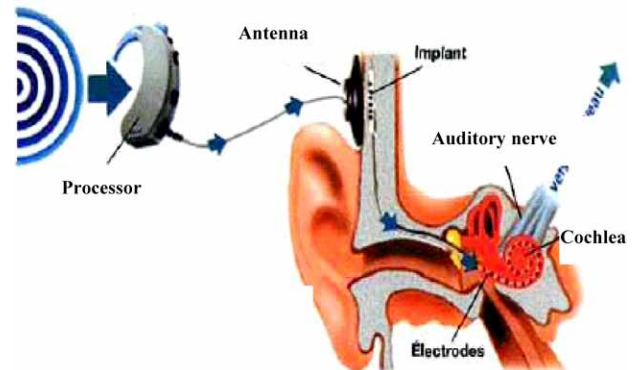

Fig. (1). The cochlea implant disposition.

This stimulation technique depends on:

- the filter bank frequencies
- the excited channel number and order
- the pulse rate and duration
- the dynamic of compression.

2.1. The Auditory Filter Bank

In fact, the filter bank must represent the auditory system with a non linear frequency distribution along the electrode array. These values can be computed by several models such

*Address correspondence to this author at the Signal Processing Laboratory, ENIT, BP 37 Tunis Belvedère, Tunisia; E-mail: b2lamia@yahoo.fr

as those of Moore [3], Patterson [4], Meddis [5] or Greenwood [6].

2.2. The Stimulation Rate

The pulse rate defines the number of pulses per second delivered to each electrode. Pulse rates can vary between 100 to 2500 pps. Wilson *et al.* [7] reported that some patients obtain a maximum recognition performance with a pulse rate of 833 pps and a pulse duration of 33 μs/phase. Other patients obtain small significant increases in performance as the pulse rate increases from 833 pps to 1365 pps, and 2525 pps. This can be interpreted and expected because high pulse-rate stimulation can better represent fine temporal variations (consonants and unvoiced speech).

2.3. The Stimulation Order

The stimulation order can be varied to minimize possible interaction between channels. The stimulation order refers to the order with which the electrodes are stimulated. One possibility is to stimulate the electrodes in an apex-to-base order (Fig. **2**). In this way, signals in the low frequencies (apex) are stimulated first, and signals in the high frequencies (base) are stimulated last.
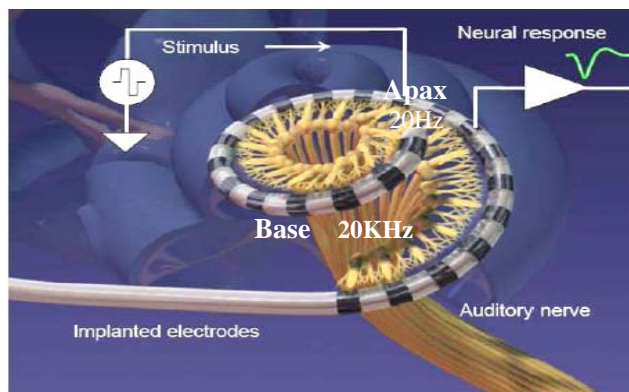

Fig. (2). The electrode array in the cochlea.

This apex-to-base order, however, does not minimize the spatial separation between sequentially stimulated electrodes. Alternatively, the electrodes can be stimulated in a so called "staggered" order, i.e., in the order of six-three-five-two-four-one, which maximizes the spatial separation between stimulated electrodes. Like the pulse rate, preference for stimulation order varies from patient to patient. Some patients prefer the apex-to-base stimulation, because as they say speech sounds more natural and more intelligible, while other patients prefer the staggered order stimulation [8].

2.4. Compression Function

The compression of envelopes of the filter bank outputs is an essential component of the CI processor because it transforms acoustical amplitudes into electrical amplitudes. This transformation is necessary because the range in acoustic amplitudes in conversational speech is considerably larger than the implant patient's dynamic range. Dynamic range is defined here as the range in electrical amplitudes between threshold (barely audible level) and loudness uncomfortable level (extremely loud). In conversational speech, the acoustic amplitudes may vary over a range of 30 dB. Implant listeners, however, may have a dynamic range as small as 5 dB. Logarithmic compression functions of the form:

$$Y = A \log(X) + B \qquad (1)$$

are typically used, where X is the acoustic amplitude (envelope output), A and B are constants, and Y is the compressed electrical amplitude.

3. CODING STRATEGIES

Several speech coding strategies are used in references and industrial implantation such as:

3.1. Fo/F1/F2 Strategy

The F0/F1/F2 strategy was the first strategy developed for the Nucleus device in the early 1980 [9]. In this strategy, the fundamental frequency (Fo) and the second formant (F2) are extracted from the speech signal using zero crossing detectors (Fig. **3**). One zero-crossing detector is used to estimate F0 from the output of a 300 Hz low-pass filter, and another zero-crossing detector is used to estimate F2 from the output of a 800-2300 Hz band pass filter. The amplitude of F2 is estimated with an envelope detector by rectifying and low-pass filtering (at 35 Hz) the band passed signal. The F0/F2 processor conveys F2 frequency information by stimulating the appropriate electrode in the 22-electrode array. Voicing information is conveyed with F0 by stimulating the selected electrode at a rate of Fo pps. The amplitude of the pulses is set in proportion to the amplitude of F2. During unvoiced segments, the selected electrode is stimulated at quasi-random intervals with an average rate of 100 pulses/sec.
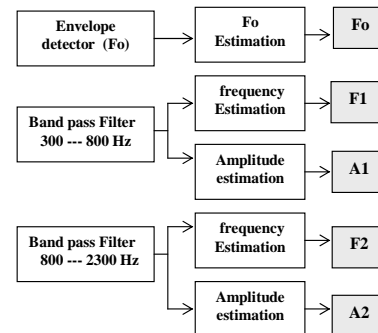

Fig. (3). The Fo/F1/F2 strategy principle.

3.2. The MPEAK Strategy

The block diagram of the MPEAK strategy is shown in Fig. (**4**). The extraction of the formant F1 and F2 was performed using zero crossing detectors, and the amplitudes of F1 and F2 were computed using envelope detectors. The frequency range for F2 was refined in the MPEAK strategy to 800-4000 Hz. Additional high-frequency information was extracted, using envelope detectors, from the frequency bands 2000-2800 Hz, 2800-4000 Hz and 4000-6000 Hz. The motivation for using the three additional band pass filters (> 2 kHz) was twofold: (1) to enhance the representation of the second formant (F2), and (2) to include high-frequency information which is important for the perception of consonants. The estimated envelope amplitudes of the three band pass filters were delivered to fixed electrodes. Electrodes 7,

4, and 1 were allocated to the outputs of the filters 2-2.8 kHz, 2.8-4 kHz and 4-6 kHz respectively. The MPEAK strategy stimulates four electrodes at a rate of F0 pulses/sec for voiced sounds, and at quasi-random intervals with an average rate of 250 pulses/sec for unvoiced sounds. For voiced sounds, stimulation occurs on the F1 and F2 electrodes and on the high-frequency electrodes 4 (2000-2800 Hz) and 7 (2800-4000 Hz).
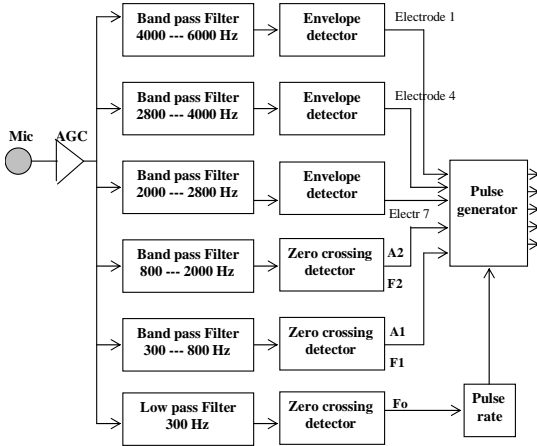

Fig. (4). The MPEAK strategy principle.

3.3. CIS Strategy

The Continuous Interleaved Sampling (CIS) strategy is unique to Nucleus Freedom. It emphasises timing information of speech.
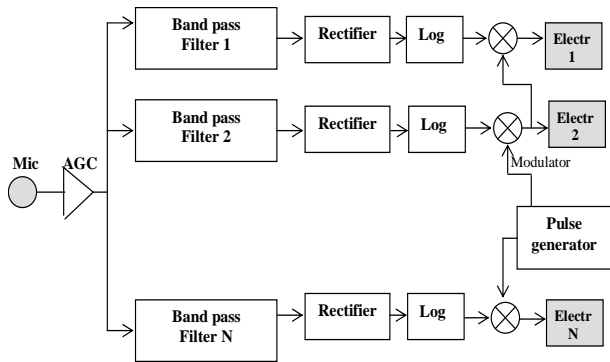
Fig. (5). The CIS strategy principle.

This strategy uses a fixed set of electrodes and offers new stimulation rates. Researchers at the Research Triangle Institute (RTI) developed CIS approach by using non-simultaneous, interleaved pulses. Trains of biphasic pulses are delivered to the electrodes in a non-overlapping fashion, that is, in a way such that only one electrode is stimulated at a time to avois interactions. The pulse amplitudes are derived by extracting the envelopes of band passed waveforms. The CIS approach is shown in more detail in Fig. (**5**).

3.4. N-of-M Strategy

The fundamental of n-of-m strategies consists on the activation, at each stimulation cycle, of only n channels (the n channels with more energy at this cycle) of the m available

channels. The objective of n-of-m strategies is to provide an increment in the stimulation rate. This is possible because the reduction of the number of channels activated at each cycle makes the duration of the whole cycle to be shorter. The increment of the stimulation rate using n-of-m has as consequence a reduction of the quality because the information corresponding to those channels with lower energy is lost. In order to simulate the effect of n-of-m strategies, at each stimulation cycle the envelopes corresponding to the different channels are compared. The n 20channels with the highest energy are selected and the rest of the channels are set to a null value. This way, the information corresponding to the non selected channels is removed from the synthesized signal.

3.5. CA Strategy

The CA approach was implanted by Symbion Inc. *via* Inraid processor. This strategy uses analogue stimulation that delivers four continuous analogue waveforms to four electrodes simultaneously at 0.5 KHz, 1 KHz, 2 KHz, and 3.4 KHz (Fig. **6**). A major concern associated with simultaneous stimulation is the interaction between channels caused by the summation of electrical fields from individual electrodes [10]. Neural responses to stimuli from one electrode may be significantly distorted by stimuli from other electrodes. These interactions may distort speech spectrum information and therefore degrade speech understanding.
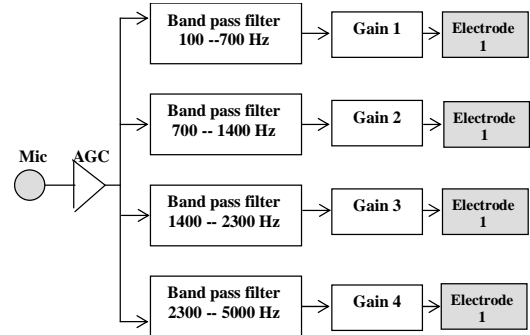

Fig. (6). The CA strategy principle.

3.6. ACE Strategy

The advanced combinational encoder (ACE) strategy used with the Nucleus implant is an "N of M" speech coding strategy. The signal from the microphone is first pre-emphasized by a FIR filter that amplifies the high-frequency components. Adaptive-gain control (AGC) is then used to limit distortion of loud sounds by reducing the amplification at the right time. Afterwards, the signal is digitized and sent through a filter bank. An estimation of the envelope is calculated for each spectral band of the audio signal. The envelopes are obtained by computing the magnitude of the complex output. Each band pass filter is allocated to one electrode and represents one channel (Fig. **7**).

3.7. Speak Strategy

The Spectral Peak strategy uses a filter bank with 22 band-pass filters. Each filter corresponds to one intra cochlear electrodes. The *energy calculation* block computes energy in each processed band. The *band selection* block selects *M* bands (between 6 and 10) with maximal energy.
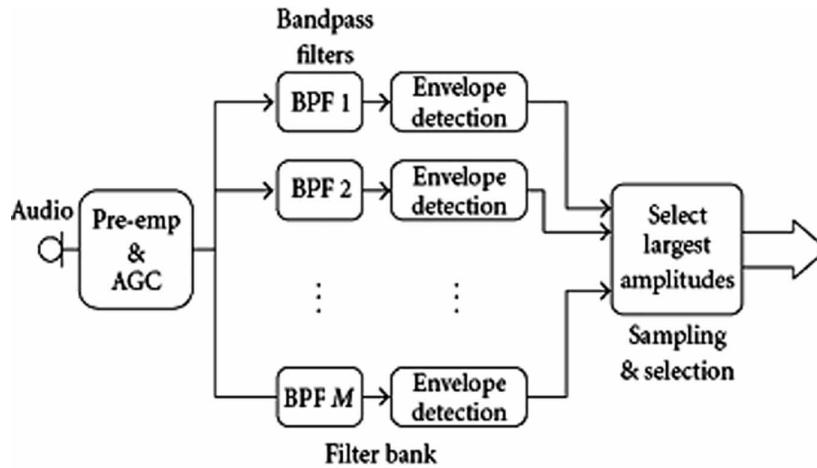
Fig. (7). The ACE strategy principle.

The Logarithmic Gain Function block implements the compression relationship between current amplitude and loudness of auditory perception. A moderate stimulation rate is used to increase the richness of important pitch information by stimulating electrodes across the entire electrode array.

3.8. SMSP Strategy

The Spectral Maxima Sound Processor (SMSP), was developed in 1990 for the University of Melbourne. It analyzed the speech signal using a bank of 16 band pass filters and a spectral maxima detector. Fig. (**8**) shows the block diagram of the SMSP processor. The signal from the microphone is first preamplified and then sent through a bank of 16 band pass filters with center frequencies ranging from 250 to 5400 Hz. The output of each filter is rectified and low-pass filtered with a cut off frequency of 200 Hz. After computing all 16 filter outputs, the SMSP processor selects, at 4 msec intervals, the six most significant (in amplitudes) filter outputs. The six amplitudes of the spectral maxima are

finally logarithmically compressed, to fit the patient's electrical dynamic range, and transmitted to the six selected electrodes through a radio-frequency link.

Only the 16 most-apical electrodes are activated, the remaining basal electrodes in the 22-electrode implant are left inactive. Six biphasic pulses are delivered to the selected electrodes in an interleaved fashion at a rate of 250 pps.

3.9. Electro-Acoustic Stimulation

As illustrated in Fig. (**9**), Electro-Acoustic Stimulation (EAS) combines the electrical stimulation (through the cochlear implant) with the acoustic stimulation (through a hearing aid). This technique is useful for those patients who keep residual hearing for low frequencies [11]. In this case, the audio spectral range is split into two parts: lower frequencies corresponding to acoustic stimulation and upper frequencies corresponding to electrical stimulation. The synthesized signal is the sum of the two kinds of stimulation.
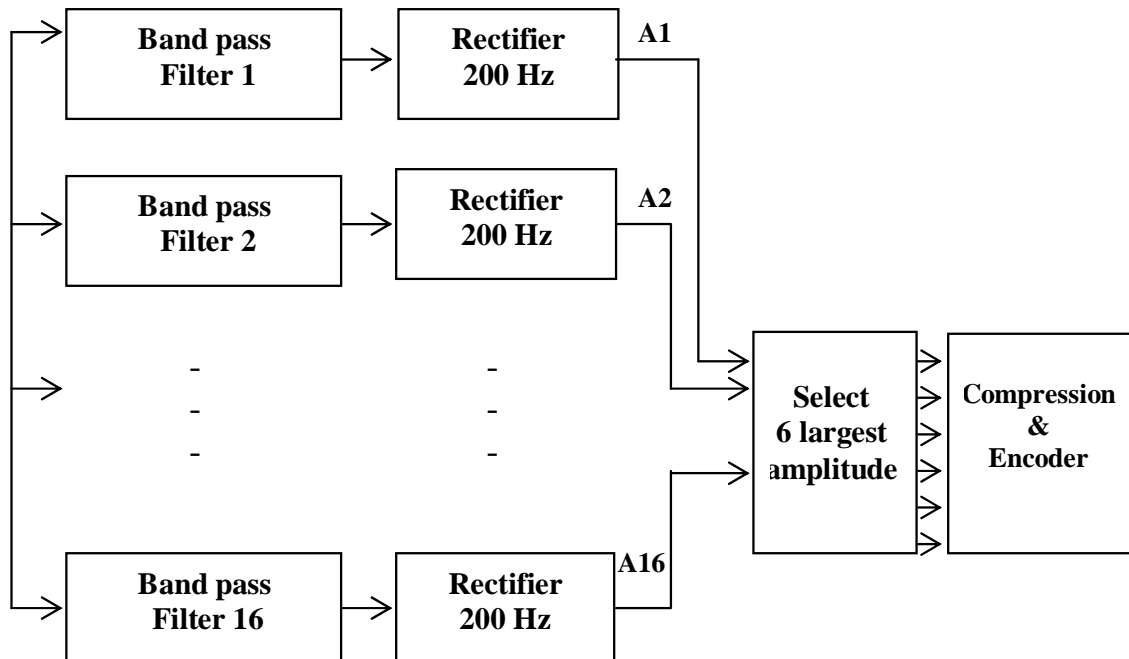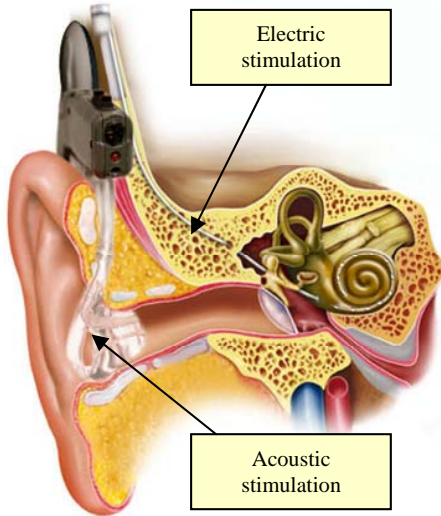

Fig. (8). The SMSP strategy principle.

Fig. (9). Electric and Acoustic stimulation.

4. THE ASAS-GC STRATEGY

4.1. ASAS-GF Algorithm

The principle of our strategy is given by Fig. (**10**). After a pre-emphasis and segmentation step of the speech signal into overlapped hamming windows, our developed algorithm uses decomposition through 22 channels gammachirp filter bank GFB. This auditory filter is chosen because it simulates with high accuracy the ear psychoacoustic model and the non linear aspects of the hearing system. Each signal of the 22 GFB outputs is analysed in order to compute its energy. The most significant bands constituting 90% of the signal Energy (generally between 3 to 6) are selected to be coded according to an adaptive coding strategy and then this parameter α_i is chosen by experimental tests with optimal performances (SNR, MQE, RR).

4.2. ASAS-GF Coding Strategy

The GFB filter bank outputs are spectrally analysed in order to compute the energy of each channel. Only the most significant bands constituting 90% of the input signal are selected. Our experience and measurements demonstrate that

this number varies between 3 to 6 depending on the speech information. Then, we compute the Hilbert transform of each selected channel in order to extract the envelope (for the amplitudes of spikes) and the analytic signal (for the instantaneous frequency of stimulation f_i).

Transmitted to the basilar membrane electrodes. This new method is based on two enhanced factors:

- The first one is the reduced number of excited channels (3 to 6 instead of 8, 16 or 22 as others strategies). This parameter is significant because it allows us to increase the stimulation speed especially for transitory information (consonants, fricatives,…)

- The second innovation and modification is the variable stimulation rate Ri which depends on the instantaneous channel frequency fi and the signal variance of the concerned channel Vi. The variation law of Ri can be expressed as:

$$R_i = f_i + \alpha_i V_i \tag{2}$$

α_i : is a balanced factor.

i : indices referring to the excited channel i.

V_i : signal variance.

f_i : stimulation frequency

The information is delivered to the electrodes using a pulsate mode (a set of narrow pulses). The amplitudes of these pulses are extracted from the envelopes. The advantage of this type of stimulation is that the pulses can be delivered in a non-overlapping fashion, thereby minimizing channel interactions and using high pulse rates tend to yield better performances.

In fact, the stimulation frequency $f_i(t)$ can be deduced from the phase $\phi(t)$ of the analytic signal $s(t)$ as :

$$\phi(t) = \arg(s(t)) \tag{3}$$

with : $f_i(t) = \dfrac{1}{2\pi} \dfrac{d\phi(t)}{dt}$ \hspace{1cm} (4)

This method is more accurate than the strategy which uses the central frequencies as speed stimulation.
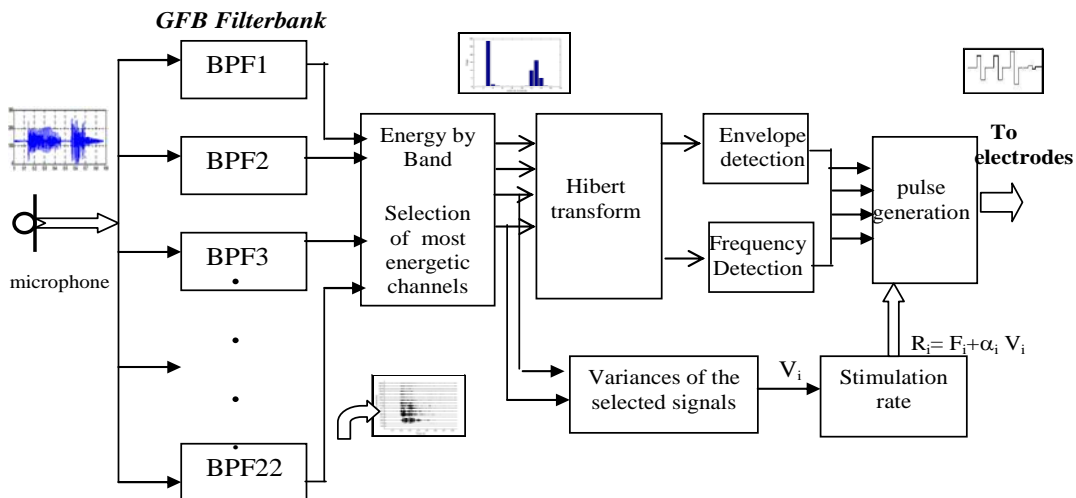

Fig. (10). The ASAS-GC method.

Finally, our last modification and contribution is the possibility of optimizing the stimulation rate in function of the three factors:

- the patient

- the environment

- the kind of information

To resolve this problem, we adopted a balanced and experimental factor which uses an adaptive law taking into account the signal dynamics (variance) and the instantaneous channel frequency. The relation is given by expression (2).

5. IMPLEMENTATION

We have integrated the algorithm of Fig. (**1**) in a Matlab speech processing program for evaluation and simulation of our coding strategy. Then, the script is compiled with C language in a programmable DSP 21065L of Analog Devices.

5.1. The Gammachirp GBF Implementation

Reference [12] proposed a GFB model of the auditory filter called gammachirp. The impulse response of the gammachirp filter is given by the following expression [13]:

$$g_c(t) = at^{n-1}\exp(-2\pi bERB(f_r)t) \times \exp(j2\pi f_r t + jc\ln t + j\phi) \quad (5)$$

with :

n   :   filter order,

f_r :   is the modulation frequency of the gamma function,

a   :   is the carrier normalization parameter,

c   :   is the asymmetry coefficient of the filter,

ϕ :   is the initial phase,

BERB :   is the filter envelope,

ERB   :   represents the equivalent rectangular band given by [6, 14]:

$$ERB (fr) = 24.7 + 0.108\ fr \quad (6)$$

The GFB frequency response of Fig. (**11**). is [15]:

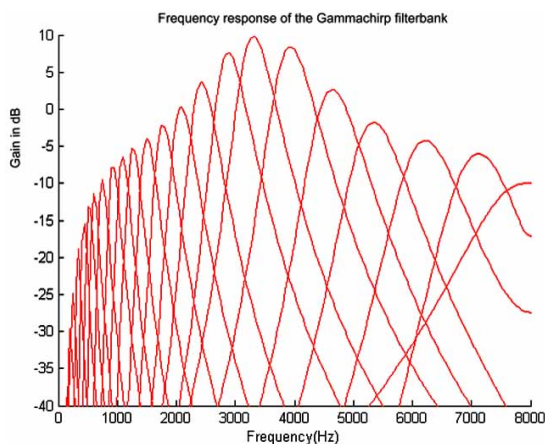$$G_c(f) = \frac{a\Gamma(n + jc)e^{j\phi}}{\left\{2\pi bERB(f_r) + j2\pi(f - f_r)\right\}^{n+jc}}$$

Fig. (11). Frequency response of the Gammachirp filter bank GFB filtred by the external ear function.

5.2. Spectral Analysis

Figs. (**12**, **14**) illustrate temporal responses of the G.F.B filter bank outputs respectively of the vowel /a/ and the consonant /sh/ pronounced by a female speaker. We can deduce that for the analysed vowel /a/, there are 3 most significant channels which are the 3rd, 5th and 7th corresponding to central frequencies: *263Hz, 507 Hz, 871 Hz*. According to our coding strategy, the stimulation rates of the 3 excited channels are computed by their couple of variances and instantaneous frequencies (F_3,V_3, F_5,V_5 and F_7,V_7). The simulation results give the following values: 344 pps for the 3th electrode, 467 pps for the 5th electrode and 785 pps for the 7th electrode.
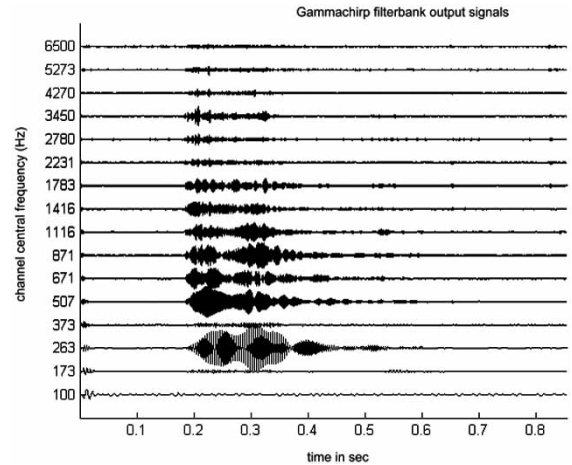

Fig. (12). GFB output signals of the vowel /a/ pronounced by a female speaker.

Yet, according to Figs. (**13**, **15**), we can deduce that for the analysed consonant */sh/*, there are four most significant channels (Energy>90%) which are the 12th, 13rd , 14th and 15th corresponding to channels with central frequencies: 2780 Hz, *3450Hz, 4270 Hz and 5273 Hz*. The 4 stimulation rates computed by the ASAP-GC (Vi,fi) are 1205 pps for the 12th electrode, 1281 pps for the 13th electrode, 1429 pps for the 14th electrode and 1532 pps for the 15th electrode.
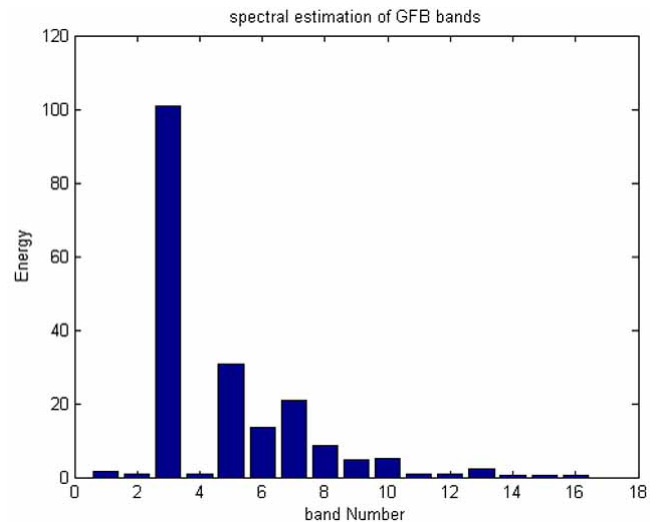

Fig. (13). Spectral analysis: l energy distribution by channel of the same vowel /a/.
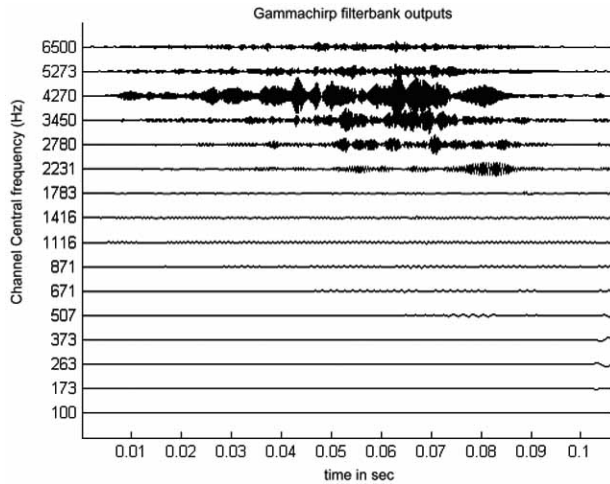
Fig. (14). GFB output signals of the consonant /sh/ pronounced by a female speaker.
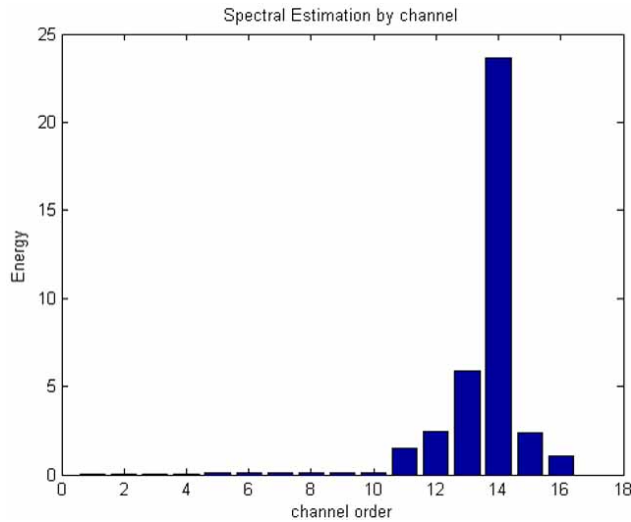

Fig. (15). Spectral analysis: l energy distribution by channel of the same consonant /sh/.

5.3. Envelope Detection

Envelope detection can be implemented either by a rectifier followed by a low-pass filter or with the Hilbert transform. We used the HT because it provides an optimal envelope of the signal temporal evolution and a good estimation of the frequency stimulation. Fig. (**16**) illustrates the first four envelope bands of a vowel /a/.

5.4. Stimulation Spikes

The stimulation rate limits the temporal resolution of the implant and its capability for the perception of fast changes. As the stimulation rate is lower, the quality of the sound perceived is worse. Besides, the temporal resolution for a cochlear implant patient is limited by the refractory period of the neurons in the auditory nerve. In fact, the time required for the repolarisation of the neurons after a neural firing is about 2 ms. For this reason, a stimulation rate above 500 pulses per second is convenient. In Figs. (**17** and **18**), we illustrate the temporal evolution of the speech signal and the electrode spikes for the word "bienvenu" pronounced by a female speaker.

6. PERFORMANCES EVALUATION

In order to evaluate our ASAS approach, we have computed its performances especially with noisy environments. Besides we succeed to set and to find the optimal setting parameters such as the energy threshold, the number of stimulated channels N_{max}, and the balanced factor. These results show that it is not advantageous to increase the number of excited electrodes above an optimal number (No). Fig. (**19**) demonstrate that this number is 4 if M=16 (Total number of electrodes) and it becomes 5 if M=22. According to Fig. (**20**), this number can be chosen between 16 and 22 to obtain optimal values of the Mean Quadratic Error (MQE) which is defined as:

$$MQE = \frac{1}{N}\sqrt{\sum_{k=1}^{N}(S_r(k)-S_s(k))^2} \tag{7}$$

S_r　:　　is the original speech signal
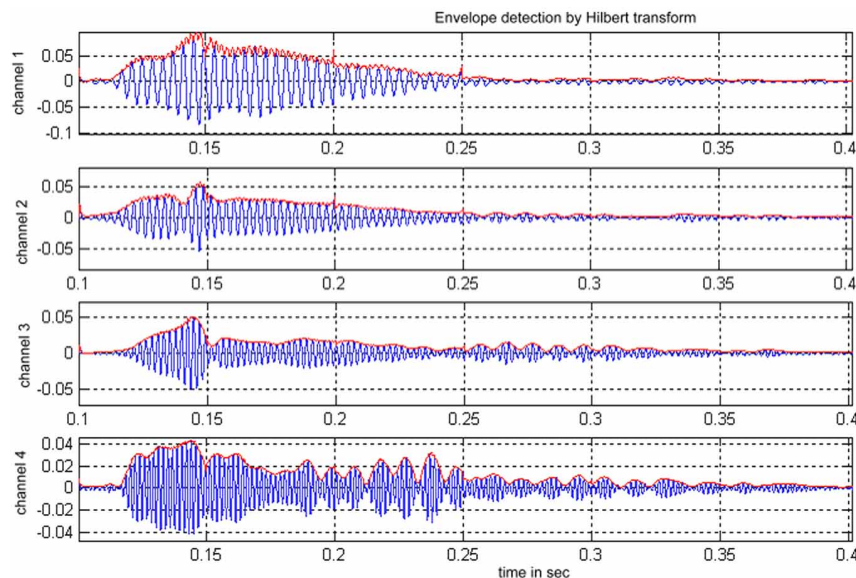
S_s　:　　is the stimulated speech signal

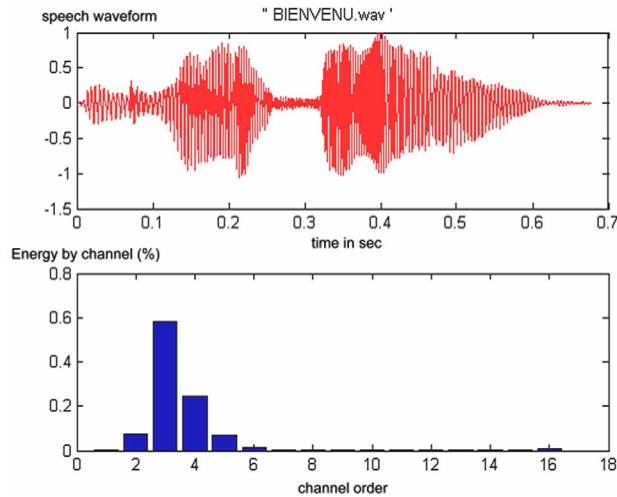**Fig. (16).** Envelope detection with Hilbert Transform.

Fig. (17). Speech waveform and bands energy of the speech "bienvenu" : female sound.

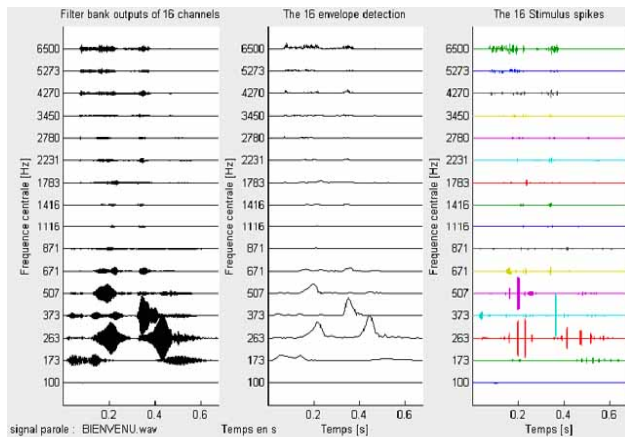## 6.1. Effect of the Number of Channels

Fig. (18). Basilar membrane signals, envelopes and stimulation spikes of the 16 CI channels of the same speech "bienvenu" : female sound.
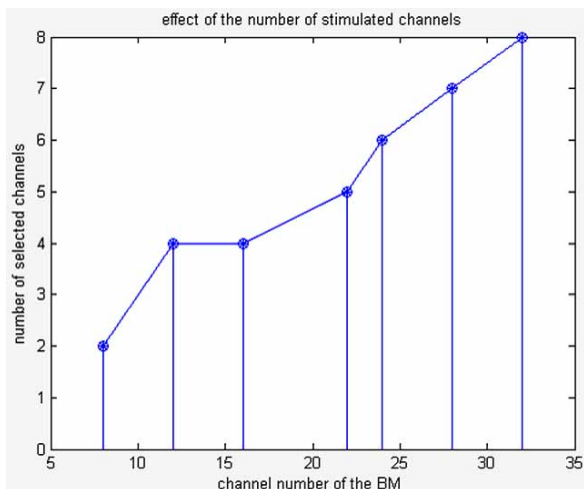

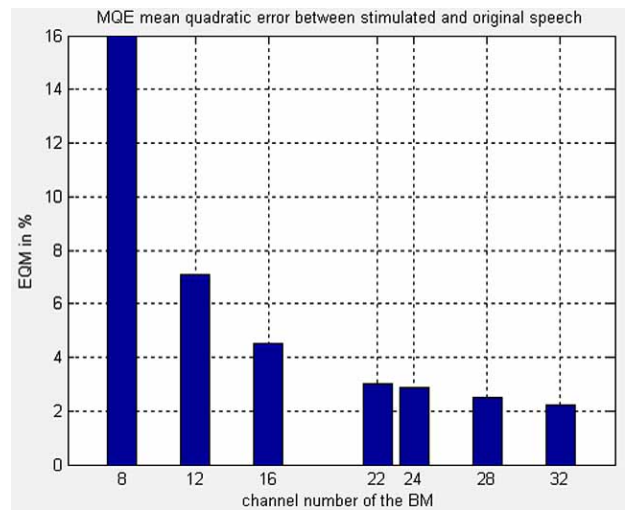**Fig. (19).** Optimal number of selected electrodes.

Fig. (20). Mean Quadratic error between stimulated and original signal in function of the number of electrodes M.

6.2. Effect of the Noisy Environment

To study the robustness of the CI coding strategy, we computed the same parameters with noisy environments. For example, Table **1** and Fig. (**21**) represent the MQE in function of the SNR values and the number of excited channels (No). We can easily observe that the ASAS strategy became vulnerable for SNR< 5dB. This problem is resolved but in Fig. (**22**) by increasing the number of excited electrodes (No).

Table 1.   Effect of Noise on MQE Error

SNR in dB	Number of Stimulated channels	MQE	Total Number of channels
15	4	0,0021	22
10	6	0,0037	22
5	11	0,0078	22
0	13	0,0095	22
-5	16	0,0110	22

By using the results of Fig. (**22**) we can deduce the next expression:

$$No = N_{max} - 0.6 \, (SNR+5) \qquad (8)$$

This relation demonstrate that with high noisy signals (SNR=-5dB) all channels (N_{max}) must be stimulated to obtain acceptable speech intelligibility and recognition ratio. Generally, every decreasing of the SNR of 5 dB must be compensated by and increasing of the number of stimulated channels of 3.

This result is very important because it can also follow the optimal No values under several conditions and environments.

6.3. Comparison with others Strategies

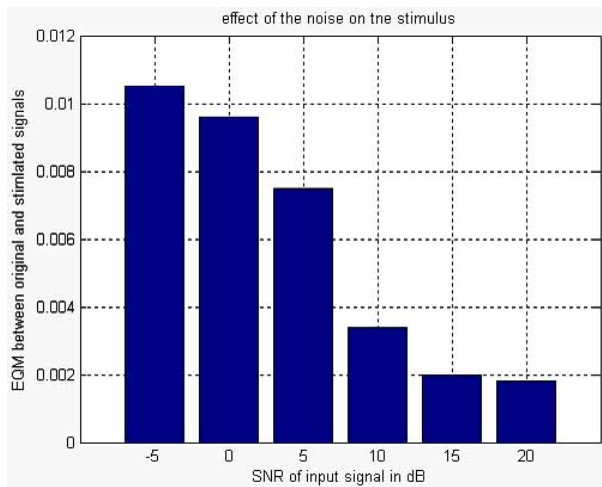Another degree of performances evaluation of our ASAS-GF technique is the comparison of the recognition

Fig. (21). Mean Quadratic error between stimulated and original signal in function of SNR.
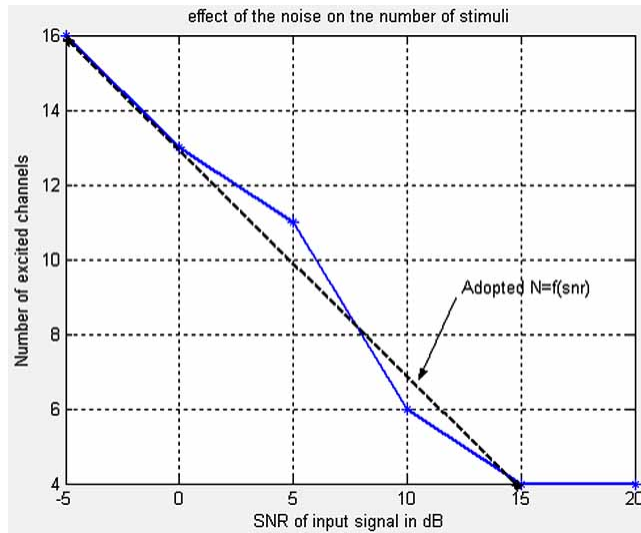

Fig. (22). Optimal number of excited channels (N_o) evolution in function of SNR.

ratios with those obtained by classic CIS and "N of M" strategies. Figs. (**23**, **24** and **25**) represent respectively the
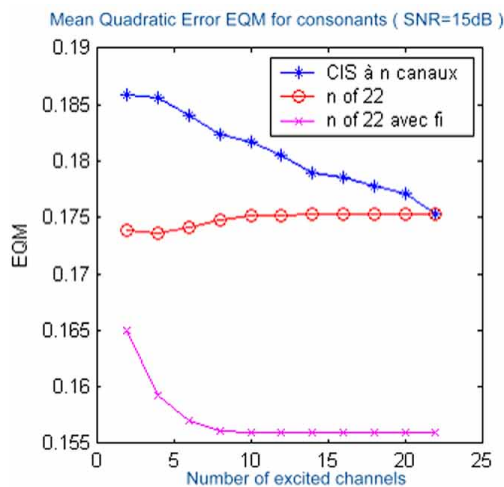
mean quadratic error, the recognition ratio for vowels analysis and the recognition ratio for consonants analysis under different conditions especially in function of the the number (No) and the SNR and the kind of information (vowels or consonants).

By analysing the obtained results by objective and subjective judgements, we can deduce that the developed ASAS-GF strategy (N-off-22-with-fi and Vi) offers two advantages. The first one is the reduced number of channels necessary to reach an appreciated performance especially under noise environment. This number varies between 3 to 6 with respectively recognition ratios 60 to 85% and a mean quadratic error MQE around 17%. In addition, when the SNR decreases, this strategy seems to be the least vulnerable when compared with the n-of-m and CIS strategies. Indeed, we observed that the stimulation density can be reduced without deteriorating or degrading the speech intelligibility. Moreover, the selecting of a reduced number of channels makes that the wide band noise is filtered. Consequently, the noise power to stimulation is much weakened.

7. CONCLUSION

In this paper, we presented a new ASAS-GF speech processing and coding strategy which is intended for cochlear implants. This approach is based on an adaptive parameters extraction of the speech signal. These parameters are the number of the stimulation channels (generally optimized and varies between 3 to 6), the stimulation order and finally the stimulation rate (number of pulses per second). The first and second parameters are chosen after a spectral analysis and energy computing of the filter bank output signals. However, the last parameter is chosen in function of the signal envelope and the signal variance. The simulation results demonstrated that, for vocal, vowels and voiced speech, the stimulation rate is less than 600 pps, hence for fricatives, plosives and consonants speech frames, the stimulation rate varies between 1200 and 1800 pps. The tests in noisy environments showed that the system is affected with lower SNR(SNR< 5dB) but this problem was resolved by increasing the number of excited electrodes (3 additional channels for a decreasing of 5 dB). The stimulation results of the excited channels and their interferences for different words, demonstrated a good discrimination between these informa-

Fig. (23). MQE errors in function of excited channels (No) for the 3 coding strategies.
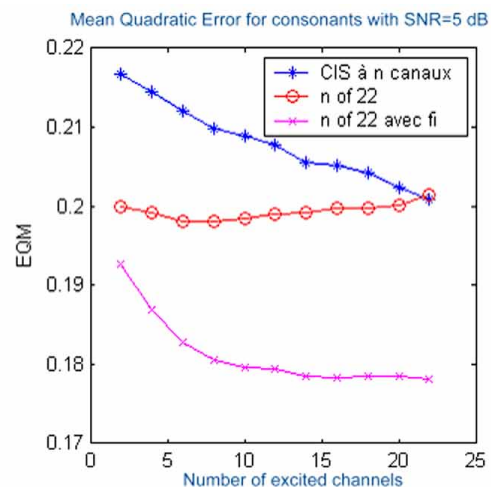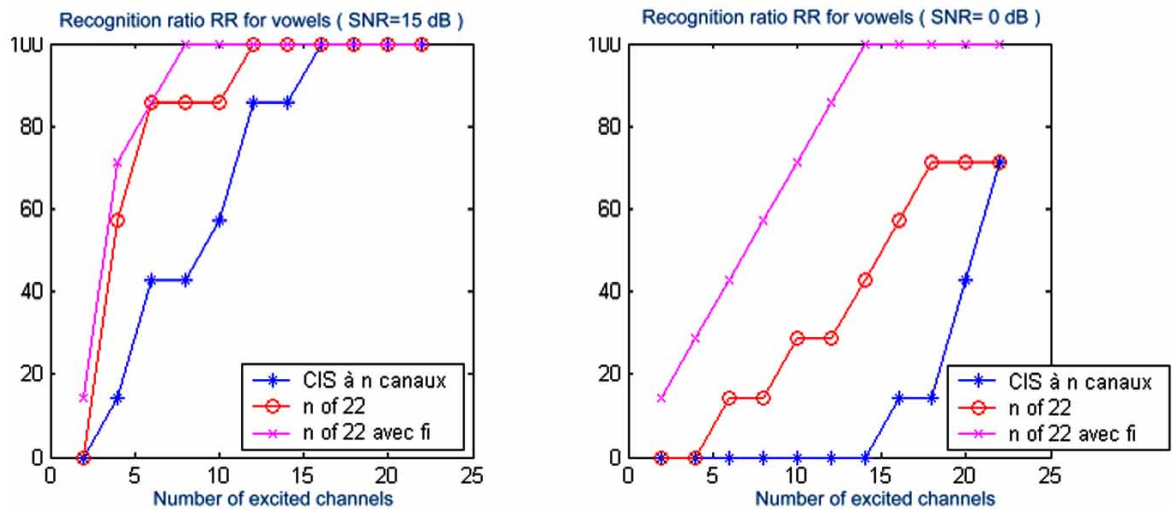
Fig. (24). Recognition ratios (RR) in function of excited channels (No) for the 3 coding strategies (signals with SNR= 15dB and 0dB): for vowels analysis.
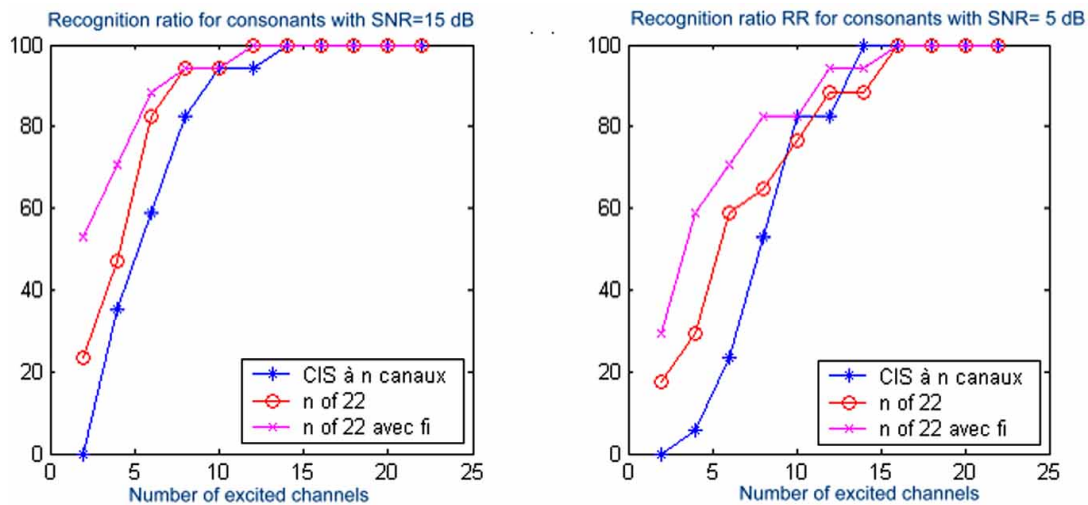
Fig. (25). Recognition ratios (RR) in function of excited channels (No) for the 3 coding strategies (signals with SNR= 15dB and 5dB): for consonants analysis.

tions especially for vowels, consonants, voiced and unvoiced speech frames. The obtained recognition ratio varies between 85% and 99% (with optimized parameters). The other advantage of our adaptive strategy is the reduced number of electrodes and the variable stimulation speed with an automatic detection of the transitory components and stationary information in the processed word. This result is important because it minimises the electrodes interferences and conduct to best signal intelligibility.

REFERENCES

[1] B.H. Ahmed, M. Samet, M. Benmessaoud, H.N. Lakhoua, and J. Mouin, "A speech processing algorithm based on a programmable filter bank for cochlear prosthesis", *In Proceedings of the International Electronic Conference JTEA*, Tunis 1998.

[2] D.B. Grayden, A.N. Kenny, O.P. Clarey, and J.C. Clark, "A cochlear implant speech processing strategy based on an auditory model", *Proceedings of the Sensor Networks and Information Processing Conference*, 2004, pp. 491- 496.

[3] B.R. Glasberg, and B.C. Moore, "Derivation of auditory filter shapes from notched-noise data", *Hearing Research*, vol. 47, 1990.

[4] R.D. Patterson, M.H. Allerhand, and C. Giguere, "Time-domain modelling of peripheral auditory processing: A modular architec-

ture and a software platform", *J. Acoust. Soc. Am.*, vol. 98, pp. 1890-1894, 1995.

[5] R. Meddis, "Simulation of mechanical to neural transduction in the auditory receiver", *J. Acoust. Soc. Am.*, vol. 79(3), pp. 702-711, 1986.

[6] B.C.J. Moore, and B.R. Glasberg, "Suggested formula for calculating auditory-filter bandwidths and excitation patterns", *J. Acoust. Soc. Am.*, vol. 74(3), pp. 750-753, 1983.

[7] J.O. Smith, and J.S. Abel, "Bark and ERB bilinear transforms", *IEEE Transactions Speech and Audio Processing"* vol. 7(6), pp. 697-708, 1999.

[8] R.F. Lyon, and C. Mead, "An analogue electronic cochlea", *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol.36(7) , pp. 1119-1134, 1988.

[9] P.C. Loizou, "Introduction to cochlear implants" *IEEE Signal Processing Magazine*, pp. 101-130, September 1998.

[10] Y. Hu, and P.C. Loizou. "A new sound coding strategy for suppressing noise in cochlear implants", *J. Acoust. Soc. Am.*, vol. 124(1), pp. 498-509. 2008.

[11] E. Zwicker, and R. Feldkeller, "*Psychoacoustics*", NY, Masson Press, 1981.

[12] T. Irino, and R.D. Patterson, "A time-domain, level dependent auditory filter: the gammachirp", *J. Acoust. Soc. Am.,* vol. 101(1), pp. 412- 419, 1997.

[13] T. Irino, and R.D. Patterson, "Modelling temporal asymmetry in the auditory system", *J. Acoust. Soc. Am.*, vol. 104(5), pp. 2967-79, November 1998.

[14]   A. G. Katsiamis, E. M. Drakakis and R.F. Lyon, "Practical gamma-tone like filters for auditory processing". *EURASIP J. Audio, Speech, Music Process.*, vol. 10, pp. 1-14, 2007.

[15]   M. Unoki, and T. Irino, "Analysis auditory filter bank based on an IIR implementation of the gammachirp" *J. Acoust. Soc. Jpn.,* vol. 20(5), pp. 397-406, 1999.