Coding Strategy Behavior related to Microphone Integrity in Cochlear Implants using the Recognition of Syllables in a Noisy Environment

P.A. Cucis¹, E. Veuillet¹,², E. Truy¹,²,⁵, H. Thai Van¹,²,⁵, F. Millioz¹,³, S. Gallego¹,⁴

¹University Claude-Bernard Lyon1, 43 bvd du 11 Novembre, 69 622 Villeurbanne-Cedex, France
²Lyon Neurosciences Research Centre, (CRNL), INSERM U1028, 95 bvd Pinel 69 675 Bron-Cedex, France
³CREATIS Laboratory, Building Blaise Pascal, av J Capelle, INSA-Lyon 69621 Villeurbanne-Cedex, France
⁴Audition-Conseil Centre, 34 avenue Lacassagne, 69 003 Lyon, France
⁵CRIC, ORL Building, Edouard-Herriot Hospital, place d’Arsonval 69437 Lyon-Cedex 03, France
(E-mail: cucis.pa@gmail.com)

Abstract

Coding strategies in Cochlear Implants (CI) are still under investigation in order to improve the patient’s care.

Speech understanding in a noisy environment is an on-going problem. It is worsened by the aging of the components, mostly with the loss of sensitivity of the microphones.

In this work, disyllabic words have been presented to a group of 20 normal-hearing (NH) subjects. The signal was processed by a vocoder simulating the continuous interleaved sampling (CIS) and the number of maxima (NofM) classical strategies used in the implant. The loss of sensitivity of the microphone was simulated after a study conducted on regular hearing aids; in most cases the microphones are the same than those used in CI processors. Four degrees of sensitivity loss have been considered in simulation.

The evaluation of the subjects’ performances (linked to the experimental conditions) came from the Fournier’s lists mixed with a cocktail-party noise. Recognition percentages were fitted by a sigmoid regression curve; consequently, the 50% recognition level, the slope of the curve and the theoretical maximum could be evaluated with respect to the Signal to Noise Ratio (SNR).

Results indicated that the CIS coding was more resistant to the noise than NofM. The recognition of syllables ranged from 0% (for a SNR of -3 dB) to 100% (for a SNR of 9 dB).

CIS strategy was more efficient in the middle values of the SNR (0, 3, 6 dB). For the worse microphone sensitivity NofM was behind CIS.

The two strategies “CIS” and “NofM” are equivalent when N = M; in future the values of N and M in the NofM based strategies should be discussed according to the context (environmental sounds, medical history and electrophysiological measures).

Key words: Cochlear Implants, CIS and NofM Coding Strategies, Microphone Sensitivity, Simulation, Noisy Environment.
1 Introduction

Many classical failures of medical equipment come from aging. They affect “old” instruments [2, 3], and one of the goals of the engineers is to reduce their impact.

Concerning hearing rehabilitation systems, the input microphone is corruptible by aging and this effect cannot be avoided. The microphone integrity should be questioned. Obviously, damages or modifications on the input sensor will be amplified by the machine and they will affect the subjects’ performances. This aspect concerns classical hearing aids, and also cochlear implants [12, 13]. Several coding strategies are offered by the manufacturers [6, 7] and the impact of the microphone drift in front of aging is worthwhile to be studied.

Cleaning the hearing aids is a classical task done during the routine checks of hearing-impaired subjects; hearing aids analyzers are designed to perform this task. Cleaning the microphone is a routine maintenance of hearing aids. Moreover, hearing care professionals can check the results by using an analyzer designed to perform this task. Cleaning and checking are two of the bases of the maintenance.

In everyday life, hearing devices are used mainly in noisy environments. In order to understand the effect of noise it is necessary to conduct a study in reproducible experimental conditions. Thus the influence of the parameters driving the signal processing of the implants can be evaluated [7, 9]. The recognition of syllables and speakers is also a pending problem and this aspect deserve to be considered [1, 17], mostly with the tools offered by modeling and simulation.

In our work, processed signals were presented to NH subjects and the same subject was faced to different situations (mostly coding strategies and SNR) [11]. Furthermore, in the literature several authors have indicated that it is possible to extrapolate the results obtained in simulation to CI populations [7, 8, 10].

To be significant, the results should come from a large population. Consequently the rules of public health should be followed and an official approval was needed. Two stages have been taken. First, a pilot study with three people was done to evaluate the feasibility of the study [12]. Secondly we submitted an application to an Ethics Committee which gave the approval to a larger experiment.

Then the two main strategies used with cochlear implants (CIS and NoFIM) have been evaluated in noisy environments.

In a second work, cochlear implants recipients have been recruited. Speech recognition was recorded before and after cleaning the microphone. This study was done directly on the target population, but it was less powerful as CI users had only one strategy and only one loss of sensitivity on their microphone [4, 5].

The article is organized as follow. After the introduction the experimental design is depicted. The simulation of the coding strategies, the representation, the population which participated to the study, the acoustic material (signal & noise), the devices used in the experiment, the simulation of the microphone loss sensitivity, the session held with the subjects are described. In the third section the results are presented and discussed in the light of the objectives. Finally, the conclusion points out the main findings coming from this experiment.
2 Material & Methods

2.1 Implant simulation

The principles of cochlear implant are well known [6]. The cochlea of totally deaf patients is directly stimulated by electrical pulses delivered to the cochlea by the implanted electrodes. Several coding strategies are provided by the implant manufacturers and they can be simulated.

We used a vocoder to simulate the CIS and NofM coding strategies and to try to approximate on NH subjects the performances of CI users. The vocoder used is this study is represented on the figure 1.

The different steps are:

1. The input signal is taken by a microphone and it goes through a pre-emphasis filter which is a high-pass filter (cutoff frequency 1.2 kHz and slope 3 dB/octave),

2. Then the signal is sampled (frequency 16 kHz) and the analysis frames last 8 milliseconds (128 samples are in a frame). The overlap between two consecutive frames is 6 milliseconds (75% overlap),

3. Then amplitudes are corrected using a Hanning window (for continuity reasons),

4. After this correction a STFFT (Short Term Fast Fourier Transform) is applied to the samples, leading to 64 spectral beams (amplitude and phase), ranging from 0 to 8 kHz. The step between two beams is 125 Hz.

5. In the next step the spectral beams are grouped into frequency bands which are logarithmically distributed, according to the ear physiology (Bark scale). Considering the usual values taken in cochlear implant we used 20 bands (leading to 20 channels).

6. In each band, the energy is calculated using the Parseval’s formula (the squares of the amplitude of each beam are added). The final energy is represented by the beams RMS (Round Mean Square).

In the CIS strategy (Continuous Interleaved Sampling) all the channels are taken. For the NofM strategy only the 8 more energetic channels are kept. The value N=8 is classical in cochlear implants. If N = M the NofM coding becomes CIS (all channels are taken).

7. Each channel is represented by a spectrum “narrow band” coming from a white noise spectrum. The amplitude of the narrow band follows the energy detected in the corresponding channel.

The first two-channels, which are very narrow, were represented by sine waves.

8. Finally, the output signal is obtained by summing the selected channels (8 for NofM; 20 for CIS) and a headset delivered the acoustic signal to the participants.
Figure 1: Simulation of cochlear implant coding by a vocoder.

\[ n \text{ is the number of selected bands} \]
\[ m \text{ is the total number of bands} \]
2.2 Acoustic material

The acoustic material incorporates the Fournier’s lists mixed with a cocktail-party noise.

a) Fournier’s disyllabic lists

These lists were well adapted to the acoustic possibilities of the participants. They were written in French by Jean-Etienne Fournier and were uttered by a male speaker; 40 lists are available.

Each list is constituted by 10 two-syllable common French words (eg: le bouchon = the cork), leading to 20 syllables per list. The recognition step was one syllable (5%).

These lists are a French equivalent to the English spondees lists (eg: base-ball).

b) Noise

A cocktail party noise has been used. It came from a voice mix of 8 French speaking people, 4 males and 4 females.

This noise was sufficiently heterogeneous for the task and the masking was rather invariable throughout a session.

c) Input signal

The signal delivered to the vocoder was a mix of noise and voice. The volume and the mix were controlled by an audiometer (Madsen orbiter 922 ®). This signal was delivered on the right ear of the NH subjects by a TDH 39 headset. The acoustic level of the Fourier’s lists was 60 dB SPL and the noise level ranged from 51 to 63 dB SPL. Consequently, the maximum level delivered to the subjects was below 65 dB SPL. According to the conditions requested by the ethics committee it did not exceed the 80 dB SPL limit recommended for professional noises exposure.

With these values the SNR ratios were: -3 dB, 0 dB, 3 dB, 6 dB and 9 dB. Recognition was 100% at 9 dB, and 0% at -3 dB.

d) Experimental conditions

Eight conditions have been considered:

-2 coding strategies (CIS and NoFm)

-4 degrees of microphone sensitivity loss (none, medium, strong, very strong).

Five levels of SNR were tested for each condition leading to 40 (2*4*5) combinations. Each combination was assigned to a Fournier’s list so that the lists were not repeated.

Each session started with a short training period to help the listener to understand the instructions. Then the 40 lists were randomly presented to each subject. The session lasted about 45 minutes.

2.3 Analysis of the microphones

The aim of this work was to study the resistivity of the coding strategies to the microphone partial occlusion, mostly by dirt, and its impact on speech understanding in noise.

a) Microphone testing

The manufacturer Knowles Electronics (Itaska, Illinois) is the leader on the microphone market and its products are used for classical hearing aids and for cochlear implants. The microphones used in this work were omnidirectional.
A study of the microphone sensitivity had been conducted on regular hearing aids worn by hearing impaired patients when they came at the audiologist’s laboratory for routine maintenance.

Two steps were necessary in order to check the microphone frequency drift:
- First the earphone and the earmold were deeply cleaned leading to the state “clean earphone and dirty microphone”,
- Secondly, the microphone was carefully cleaned, leading to the state “clean earphone and clean microphone”.

The subtraction of the two states gives the loss due to the microphone occlusion.

Then the microphone transfer function was evaluated after a frequency warble sweep ranging from 200 Hz to 8,000 Hz (at 60 dB SPL). This procedure had been conducted with several hearing aids trademarks, mostly Oticon®, Phonak®, Starkey® and Siemens®. The frequency check at the audiologist’s laboratory was 3 or 4 times per year.

b) Microphone sensitivity curves
The microphone response curves have been recorded, before and after cleaning, with an hearing aid analyzer.

The contribution of the cleaning (given in dB) has been measured.

Four degrees of attenuation (“cleaning efficiency”) have been defined:
- case 1: clean microphone (no loss); the response was not modified,
- case 2: medium loss corresponding to the attenuation seen on 50% of the microphones (percentile 50%)
- case 3: strong loss, corresponding to the 20% worse microphones (percentile 80%),
- case 4: very strong loss, corresponding to the 10% worse microphones (percentiles 90%).

The clean signal, coming from the input microphone, was modified by the different attenuation curves and the passed through the vocoder.

2.4 Participants
The study has been approved by the French “Sud Est II” Ethics Committee, under the supervision of the Hospitals of Lyon (August 27, 2014).

All the participants filled an agreement form prior entering the study.
All the recordings were made by an audiologist.

Twenty NH subjects participated to this experiment. Their age ranged from 18 to 33 years (average 25 years).
All the subjects had an ORL examination before entering the study, in order to eliminate previous pathologies or deafness which may corrupt the study.
The auditory thresholds were measured; they were always below 20 dB HL for all the frequencies between 250 and 8,000 Hz. According to the BIAP (International Bureau for Audio-Phonology) all the subjects were considered as NH listeners.
2.5 Mathematical analysis of the data

a) Comparison of the percentages
For each SNR and for each subject the recognition percentages were compared for:
- the strategies CIS and NofM,
- the microphone’s frequency drift (4 degrees).
The paired Wilcoxon test was used with a 5% significance threshold

b) Curve fitting with a sigmoid function
The recognition percentages versus the SNR can be represented by a sigmoid regression curve (figure 2).

In audiology, three values are considered:
- the SNR corresponding to 50% of the maximum recognition (maximum was usually 100%) denoted $x_{50\%}$; it shows the word recognition ability in noise,
- the “slope” (SNR range, given in dB, between 25% and 75% of the maximum recognition) and denoted $\Delta_{75\%-25\%}$. It shows the speed of acquisition of the syllables when the SNR increases (the smaller $\Delta_{75\%-25\%}$, the steeper the slope)
- the top asymptote shows the maximal recognition score: $y_{\text{max}}$

These values are presented on the sigmoid curve.

When the minimum recognition is 0% (seen for SNR = -3dB) the sigmoid equation is:

$$y = \frac{a}{1 + e^{-b(x-c)}}$$

Where:
- $y$ is the recognition percentage
- $x$ is the SNR
- $a$ is $y_{\text{max}}$
- $c$ is $x_{50\%}$
- $b$ is linked to the slope: $b = 2.2/\Delta_{75\%-25\%} \Rightarrow \Delta_{75\%-25\%} = 2.2/b$. 


3 Results and discussion

3.1 Microphone sensitivity

The study of the microphone sensitivity has been done on 129 microphones. The range 250 Hz-8,000 Hz was swept by third of octave [16].

On figure 3, the three degrees of loss of sensitivity, corresponding to the 50%, 80% and 90% percentiles, are indicated. Each degree was represented by two straight lines for simplification. The signal attenuation started at 800 Hz and then three slopes were introduced, respectively, 2 dB, 7 dB and 9 dB per decade. No attenuation was applied between 0 and 8000 Hz.
The curves were subtracted to the clean spectrum.

![Graph showing Microphone’s Frequency Drift due to partial inlet occlusion](image)

**Figure 3**: Simulation of the microphone loss of sensitivity.

### 3.2 Syllable recognition

The syllable recognition, versus the SNR, is shown on figure 4. The strategies CIS and NofM are indicated. The attenuation degree (4 losses) is also indicated.

Comparisons were done with paired Wilcoxon tests (on the 20 subjects who participated to the experiment). In the whole experiment, we had 40 paired series (one for each strategy, each degree of attenuation and each SNR); each paired series had 20 couples of values (one per subject).

a) **Clean microphone**

Results are indicated on figure 4a (for the clean microphone simulation, the spectrum was not modified).

For the extreme points (-3 dB and 9 dB) the recognition percentages were not significantly different, between CIS and NofM. Significant differences were found for the SNRs 0 dB, 3 dB and 9 dB (the middle of the SNR interval).

b) **Micro with medium loss (Percentile 50%)**

This degree is the average loss of the microphones. There was a significant difference for 3 and 9 dB. For 9 dB the difference was significant; the difference was small but the standard deviation was also small (figure 4b).
4a: Average syllable recognition in noise for CIS and NofM simulated strategies. Case 1: Clean microphone

Intelligibility (%) vs. SNR (dB)

4b: Average syllable recognition in noise for CIS and NofM simulated strategies. Case 2: Medium sensitivity loss

Intelligibility (%) vs. SNR (dB)

4c: Average syllable recognition in noise for CIS and NofM simulated strategies. Case 3: Strong sensitivity loss

Intelligibility (%) vs. SNR (dB)
c) Micro with a strong loss (percentile 80%)

Now, we are with the microphones having a strong sensitivity loss (degree 3). Results obtained with the two strategies were equivalent. The only significant difference found between CIS and NofM was for the SNR 0 dB (figure 4c).

d) Microphone with a very strong loss (percentile 90%)

This is the worse category. Significant differences have been seen for high SNRs (figure 4d). It is suggested that an incomplete spectrum (NofM coding) may lead to a lower representation of the signal and it has a disadvantage.

Also it is seen that, in this case, the results for the NofM strategy did not reach 100% for a high SNR. On the contrary the 100% recognition was reached with CIS. This aspect should be seen investigated more deeply in future works.

e) Effect of the microphone sensitivity

Figure 5 shows the effect of the microphone robustness to noise according to the strategy.

When the SNR was small (3 and 0 dB), the CIS strategy was more resistant to the noise; the recognition scores decreased more slowly than for the NofM. The impact of the microphone sensitivity loss on speech recognition was not the same for the two strategies.

Also, as expected, the higher the loss, the lower the recognition.

The differences between the strategies were mainly observed for the SNR values 0 and 3 dB (figure 4). The differences were 15 to 20% corresponding to 3 or 4 syllables in a Fournier’s list.
Figure 5: Effect of the microphone loss of sensitivity on the recognition of the syllables, according to the microphone loss of sensitivity, for the NofM and the CIS strategies.

3.3 Audiological values
Results coming from the sigmoid curve regression are represented on table 1. The four degrees of sensitivity loss are studied.
It can be reminded that:
- $x_{50\%}$ represents the recognition ability in noise,
- $\Delta_{75-25\%}$ is linked to the sigmoid slope with respect to the SNR variation.
It is clear that fitting the data by a sigmoid regression curve did not change the results but it gives a new insight into them.

a) Clean microphone (case 1)
For $x_{50\%}$ the results were significantly better for CIS than for NofM. For $\Delta_{75-25\%}$ the results were equivalent (case 1).
b) Median loss (case 2)
Columns 1b indicate that the average recognition ($x_{50\%}$) was better for the CIS strategy; the slopes were not significantly different.

c) Strong loss (case 3)
In this situation (columns 1c) the $x_{50\%}$ and $\Delta_{75-25\%}$ values were not significantly different.

d) Very strong loss (case 4)
Columns 1d in table 1 show that CIS and NofM were different. Also the asymptotes $y_{\text{max}}$ were different (figure 4d).

$x_{50\%}$ and the slope were better for CIS.

As indicated above, the incomplete spectrum transmitted by the NofM strategy seemed to be a disadvantage in simulation.

<table>
<thead>
<tr>
<th>Case 1 (1a)</th>
<th>Case 2 (1b)</th>
<th>Case 3 (1c)</th>
<th>Case 4 (1d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta_{75-25%}$</td>
<td>$x_{50%}$</td>
<td>$\Delta_{75-25%}$</td>
<td>$x_{50%}$</td>
</tr>
<tr>
<td>NofM</td>
<td>2,54</td>
<td>0,39</td>
<td>2,04</td>
</tr>
<tr>
<td>CIS</td>
<td>2,19</td>
<td>-0,28</td>
<td>1,87</td>
</tr>
<tr>
<td>$\Delta = \text{CIS - NofM}$</td>
<td>-0,35</td>
<td>-0,67</td>
<td>-0,17</td>
</tr>
<tr>
<td>p (Wilcoxon)</td>
<td>0,287</td>
<td><strong>0,038</strong></td>
<td>0,837</td>
</tr>
</tbody>
</table>

Table 1 Slopes and mean values seen on the sigmoid curves, according to:
- the degree of sensitivity of the microphone
- the coding strategy (CIS and NofM).

Significant differences are in a green background

3.4 Results analysis
An overlook of the percentages differences seen between the strategies CIS and NofM shows that the percentages observed with CIS were better. One value did not follow this trend: “clean microphone, 9 dB SNR”. Consequently the CIS strategy appeared to be more robust than NofM for syllable recognition when N=8 and M =20. In order to attest these results, further studies should be conducted with others values of N and M [15].

If we consider the extreme percentages (the asymptotes on the sigmoid curve) the recognition was equivalent for both strategies. It was 0% for a SNR of -3dB and 100% when the SNR was +9 dB.

This was not true when the attenuation on the microphone was large; the NofM strategy did not reach the 100% (but CIS did). It may come from the incomplete spectrum. This situation has to be investigated further.

In the middle of the SNR range (0, 3 and 6 dB) the CIS strategy led to better results; the difference was about 1 to 2 syllables per list (5 to 10%). When the loss of sensitivity was high (figure 4), the difference between the two strategies could reach 4 syllables (20%). When the loss was small, the difference observed between the strategies was about 10% (2 syllables) in favor of CIS; this is the situation when the machines are regularly cleaned and checked, every three months for instance.
Also, on figure 5 it could be seen that speech intelligibility was less affected by the decrease of the SNR when the microphone was clean, for the two strategies. The maintenance periodicity needs further investigations.

These results, achieved with NH subjects, have to be confirmed with CI users. But for implanted patients we cannot modify so easily the strategy and the microphone corruption.

This information could be useful for CI manufacturers and more experiments are necessary in order to see again our results. More than ever the collaboration between physicians, audiologists, scientists, engineers and the manufacturers is necessary.

4 Conclusions

This study measured the syllable recognition in noise (cocktail party noise) of NH subjects listening to simulated CIS and NofM coding strategies, according to microphone frequency drift due to aging. The microphone drift was also simulated. Several combinations were tested on the same person and the number of recognized syllables in the disyllabic Fournier’s lists represented the performance.

Results indicated that:
- CIS led to better scores than NofM, mostly for the middle SNR (signal to noise ratio) values, such as 0, 3 and 6 dB. This result stood for clean and dirty microphones. The difference was around 10%,

- for the extreme values of the SNR (-3 dB and 9 dB) results were rather similar whatever the strategy, 0% for -3 dB and 100% for 9 dB. When the microphone was very corrupted (the 10% worst microphones) the results with the NofM strategy did not reach the 100% (contrarily to CIS).

In another companion study, done with CI Users, we found similar patterns of results. However the measured recognition percentages were collected with less degrees of liberty than in NH participants.

Several extensions of this work are suggested, such as a focus on the extreme SNR conditions (-3 dB and 9 dB), the number of open channels N and M for both strategies and the periodicity of the check in hospital specialized departments.

Similar conclusions were obtained after fitting recognition curves with a sigmoid regression curve. The loss in the recognition ability, due to the microphone corruption, was about 1 or 2 dB whatever the coding strategy.

Acknowledgements

The authors are grateful to the persons and the bodies who participated to the study; M. Kevin Perreaut who initiated the work, Dr Fabien Seldran for the scientific contribution, the members of the CRIC Lyon from the Edouard-Herriot hospital of Lyon for their collaboration, the subjects who listened to the Fourier’s lists, the Hospitals of Lyon and the Polytechnic School of the University of Lyon. Also, we thank the AMSE referee for the modifications suggested in the manuscript.
References


