AMSE JOURNALS-AMSE IFRATH Publication –2016- Modelling C; Vol. 77; N° 2; pp 84-97
Submitted Sept. 2016; Accepted Oct. 31, 2016

Cochlear Implants : Influence of the Coding Strategy on Syllable Recognition in Noise

P.A. Cucis[1], C. Berger-Vachon[1,2,5], E. Truy[1,2], H. Thai Van[1,2], F. Millioz[1,3], S. Gallego[1,4]

[1] Université Claude-Bernard Lyon1, 43 bd du 11 Novembre, 69622 Villeurbanne-Cedex
[2] Centre de Recherche en Neurosciences de Lyon, (CRNL), INSERM U1028, 95 bd Pinel, 69675 Bron-Cedex
[3] Laboratoire CREATIS, INSA-Lyon, 7 Avenue Jean Capelle, 69621 Villeurbanne Cedex
[4] Centre Audition-Conseil, 34 avenue Lacassagne, 69003 Lyon
[5] IFSTTAR, 25 Avenue François Mitterrand, 69500 Bron
E-mail : cucis.pa@gmail.com

Abstract

Sound coding strategies for cochlear implant remain under development in order to improve patients' care. For cochlear implant recipient, understand speech in noisy environments is hard and an aging device makes it even harder. The present study investigated speech recognition in noise of cochlear implant recipients (n = 45) and speech recognition in noise of normal hearing subjects listening to a cochlear implant acoustic simulator (n = 20). We considered two groups, "CIS-like" strategy users (Continuous Interleaved Sampling) and n-of-m strategy users (Number of Maxima). Moreover we took into account the microphones' integrity ("Clean" or "Soiled"). For the cochlear implant simulator, soiling was simulated by a low pass filtering. Results showed that "CIS-like" users and n-of-m users have similar speech recognition scores even when the microphones were soiled. Nevertheless, "CIS-like" users performed a bit better. Both experiments, with cochlear implant users and normal hearing subjects, came to the same results.

Key Words

Cochlear Implant, "CIS-like" and n-of-m, Cochlear Implant Recipients, Normal Hearing Subjects, Microphone Soiling, Acoustic Cochlear Implant Simulator, Speech Recognition in Noise.

1. Introduction

Cochlear implant is now a common hearing device for profound hearing loss. Each year in the world about 30,000 people receive a cochlear implant including 1,000 in France. The domain of cochlear implant concerns both clinical and scientific aspects because in this case signal processing is applied to humans [2, 5].

Despite the existence of a lot of studies on the subject, the choice of a sound coding strategy is still an open question [3, 4, 8]. Two questions remain: how to manage speech recognition in noise [4] and which procedure apply to maintain aging devices in good function. The last point matters for biomedical instrumentation [1] particularly with the microphones' sensitivity loss [7] which is also linked to soiling.

A pilot study about the impact of microphones' sensitivity loss due to soiling was presented during the IFRATH's "Handicap" congress of Paris in 2014 [6]. After that, we received an accreditation from an ethic committee for conducting a study with a larger panel of patients.

We could have tested cochlear implant (CI) users only or normal hearing (NH) subjects only but there are two main objections:

- With cochlear implant users only, the sound coding strategy depends on the implant brand chosen by the surgeon. This choice could have an impact on the results,
- Normal hearing subjects listening to a simulator can face all different situations (in our case, the two coding strategies combined with the two states of the microphone). For the same subject, it allows a direct comparison of the results. But this approach is a bit far from reality and transposing the results to cochlear implant recipients may be non-accurate.

Subjects were tested in a noisy environment considering the two main coding schemes, "CIS-like" (Continuous Interleaved Sampling, which transmits sequentially all the information of the available channels to the cochlea) and n-of-m (Number of Maxima, which selects the "n" channels of high energy among the "m" available), and considering two states of the microphone ("Clean" and "Soiled").

After this introduction the methods section presents the principles of "CIS-like" and n-of-m coding schemes and the principles of vocoded speech simulating the hearing through a cochlear implant. Then we will talk about the subjects included in our study, the devices used to test speech understanding and the data analysis methods. In a third part we will present and discuss our results before a sum up of our main findings.

2. Methods

A. *Sound coding strategies and cochlear implant simulator*

1) *Cochlear implant principles*

We will not expose the details of signal processing in cochlear implants here, but only the essential principles (figure 1). Cochlear implant can be divided in two main parts:

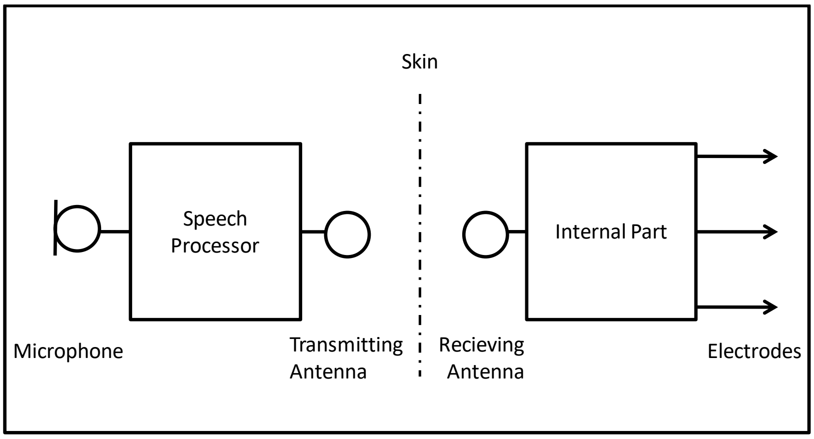
- The external block (signal processor) processes the signal and separates the sound in multiple frequency bands (channels).
- The internal block (the implant itself) receives the information from the processor and distributes the energy of the channels to the electrodes along the cochlea.

The information is transmitted from the processor to the internal part through the skin by a radiofrequency signal.

Sounds in the environment are captured by the microphones that are very impacted by soiling and aging. Then the processor fits the signal received from the microphones to the patient’s physiology and separates the signal in multiple frequency bands. The frequency bands correspond to the treatment channels scaled using a logarithmic scale. We chose to use the Bark scale for our CI simulator.

Most of the existing sound coding strategies are based on two main principles: there are “CIS-like” strategies and n-of-m strategies. It is important to notice that, for the n-of-m strategies, if “n = m” it is very similar to CIS. So the value of “n” is important to consider [8].

Fig 1: Shematic view of a cochlear implant.



2) On the effect of microphone soiling

In this study we tested speech recognition in noise of CI users, before and after a microphone cleaning process. For NH subjects we used a CI simulator integrating a low-pass filter mimicking microphones frequency drift due to soiling. The low-pass filter approximated a measured median attenuation. We based our work on a study about microphone frequency drift due to soiling on conventional hearing-aids. Knowles electronics (Itasca, Illinois, USA) is the manufacturer who leads the market of microphones for hearing-aids and cochlear implants. That is why we considered that this study matched for cochlear implants' microphones too.

The study of the microphones sensitivity loss had been conducted, during a routine maintenance appointment in a hearing care center, on regular hearing aids worn by hearing impaired patients (n = 129). 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 attenuation due to the microphone soiling.

Then the microphone transfer function was evaluated after a frequency warble sweep ranging from 200 Hz to 8,000 Hz (at 60 dB SPL). It permits to calculate a median attenuation corresponding to the attenuation measured on 50% of the microphones (percentile 50%).

The CI simulator used for testing NH subjects is a basic vocoder fitted for the purpose of the study (figure 2).

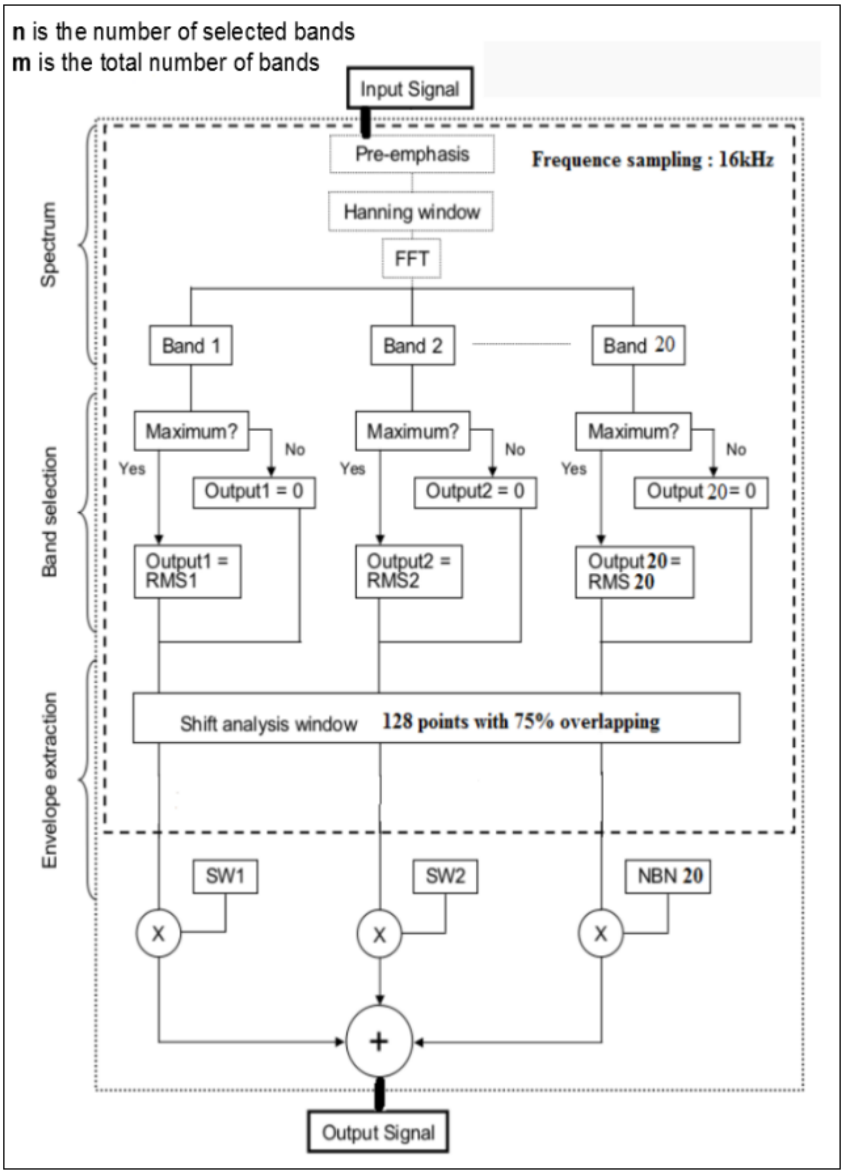
These are the different steps of signal processing applied by the CI simulator:

- The input signal goes through a pre-emphasis filter which is a high-pass filter (cutoff frequency 1.2 kHz and slope 3 dB/octave).
- 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.
- 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). 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-like” coding all the channels are taken. For the n-of-m coding implants. Then the signal is sampled (frequency 16 kHz) with an analysis frame of 8 milliseconds (128 samples are in a frame). There is a frame overlap of 6 milliseconds (75% overlap).

- Each channel is represented by a “narrow band” spectrum 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. The output signal is obtained by summing the selected channels (8 for n-of-m; 20 for CIS) and a headset delivered the acoustic signal to the participants.

Fig. 2: Block diagram representing the signal processing performed by the n-of-m simulator.



B. Participants

1) Cochlear implant users

The work presented in this paper follows a pilot study [6] and it was approved by the French Ethics Committee “Sud Est 2” (August, 27, 2014), under the supervision of the HCL (Hospitals of Lyon). Then all subjects filled and signed an agreement form before entering the study. Participants were tested at the CRIC (Fitting center for cochlear implants) in the Edouard Herriot Hospital in Lyon. Results were recorded by a certified audiologist before the annual fitting session.

Fifty-nine implanted patients were included in this study. Their age ranged from 18 to 60 years (average 37 years old). Nineteen subjects were fitted with “CIS-like” strategies (13 Medel® and 6 Advanced-Bionics®) and 26 had n-of-m strategies (16 Cochlear® and 10 Neurelec®). The numbers of channels were: Medel® (M # 12), Advanced-Bionics® (M # 16), Neurelec® (N # 8; M # 20), Cochlear® (N # 8; M # 22).

The microphones cleaning process consisted in a brushing phase and a dry air blow phase. Which is a common procedure used in hearing-care centers.

2) Normal hearing subjects

Twenty NH subjects participated to this experiment. Their age ranged from 18 to 33 years (average 25 years). The subjects had an ORL check before entering the study. It was to eliminate previous pathologies or deafness which may corrupt the study.

The auditory thresholds were measured; they were below 20 dB HL for all the frequencies between 250 and 8,000 Hz. According to the BIAP (International Board for Audiophonology), the subjects were considered as normal.

C. Acoustic devices

1) Fournier’s lists

These lists are well adapted to the patients’ recognition ability in this environment. They were uttered by a male voice and were the vocal part of the signal. They are similar to the spondees lists used in English. Each list contains 10 disyllabic words (for instance “le bouchon” = the cork). Forty lists are available, and the recognition unit was the syllable. So the recognition step was 5%.

2) Noise

A cocktail-party noise has been used in this study. This noise was a mix of voices from 8 speakers, 4 males and 4 females.

3) Delivered signal

The signal delivered to the subjects was made of the Fournier's lists mixed with the cocktail-party noise. The mixing was managed by a Madsen Orbiter 922 audiometer and the signal to noise ratio (SNR) was perfectly adjusted.

For CI users, tests were made in free field in an audiometric booth. In all cases only one ear was stimulated. When a subject was fitted with two implants, only the best ear was kept.

For NH subjects, the Fournier's lists and the cocktail-party noise was first processed by the CI simulator and then passed through the audiometer and finally delivered on the right ear by a TDH 39 headset.

The level of the Fournier's list was 60 dB SPL and the noise level was increased by a 3 dB step between each lists. Noise and words combined, the output levels were below 65 dB SPL (below 80 dB SPL, the maximum allowed for professional noise exposure). This was a requirement of the Ethics Committee.

There are the SNRs presented to the subjects:

- CI users: -3 dB, 0 dB, 3 dB, 6 dB, 9 dB, 12 dB, 15 dB, 18 dB,
- NH subjects: -3 dB, 0 dB, 3 dB, 6 dB, 9 dB (Speech recognition was 100% at 9 dB, and 0% at -3 dB).

4) Experimental conditions

Eight conditions have been considered:

- 2 coding strategies ("CIS-like" and n-of-m)
- 2 microphone states ("Clean" and "Soiled")

Each session started with a short training to help the listener to understand the instructions. With CI users we tested their speech recognition in noise before and after a microphone cleaning process. Each session lasted less than 30 minutes.

With NH subjects, 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. More, the 40 lists were randomly presented to the subject. The session lasted about 45 minutes.

D. *Statistical evaluation of the data*

1) *Speech recognition scores comparison*

We included n = 19 CI recipients using "CIS-like" strategies and n = 26 CI recipients using n-of-m strategies. To compare the speech recognition of the CI users using "CIS-like" and n-of-m coding strategies we used a Student's T-Test and a Mann-Whitney's test (non-parametric).

Paired data lower the number of participants required to have accurate statistical analysis. That was the case for NH subjects because each subject could have been tested with the two different coding strategies and the two microphone states. In order to compare the speech recognition with "CIS-like" strategy and n-of-m strategy we used Wilcoxon's hypothesis test for paired data (non-parametric).

2) *Fitting with a sigmoid curve*

The recognition percentages, versus the SNR, were fitted by a sigmoid curve (figure 3). We considered two main features: the SNR corresponding to 50% of the maximal recognition score ($x_{50\%}$) and the slope of the curve; which was evaluated by the interval (in dB) needed to go from 25% to 75% of the maximum recognition score ($\Delta_{75\text{-}25\%}$). The feature $x_{50\%}$ is a good indicator of the speech intelligibility and the feature $\Delta_{75\text{-}25\%}$ shows the speed of acquisition of the syllables with the SNR.

The sigmoid curve's equation used in our study was:

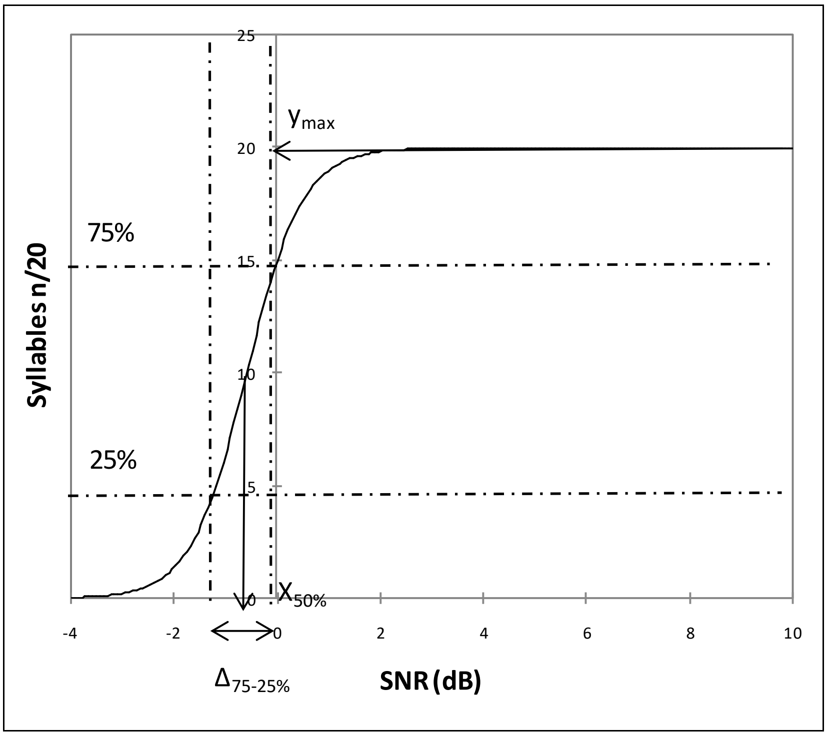
$$y = \frac{a}{1 + e^{-b(x-c)}}$$

The parameters a, b and c were given by the XLSTAT software.

The following denotations are:

- -y: recognition percentage,
- -x: the SNR
- -a: y_{max} (the top asymptote)
- -c: $x_{50\%}$
- -b: represents the slope: $b = 2.2/\Delta_{75\text{-}25\%}$

Fig 3: Fitting of the recognition percentages by a sigmoid curve.



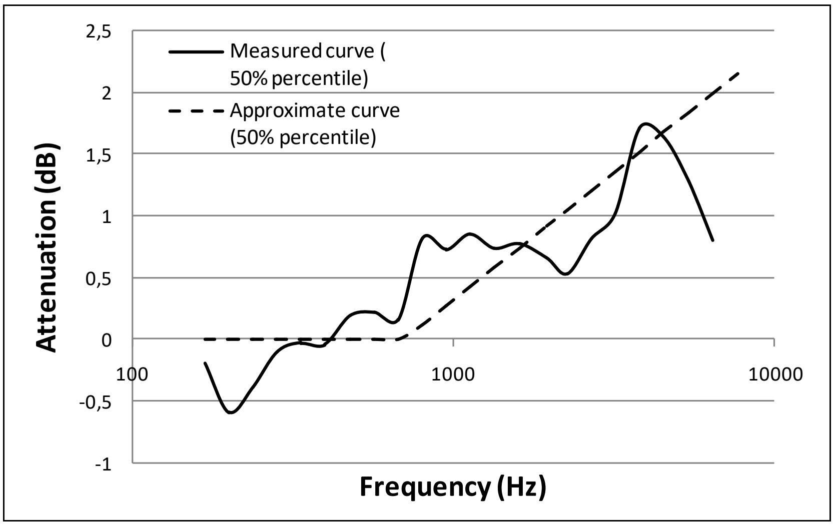
3. Results and discussion

A. Microphone sensitivity

The study of the microphone sensitivity has been done on 129 microphones. The range 250 Hz-8 kHz was swept by third of octave [9].

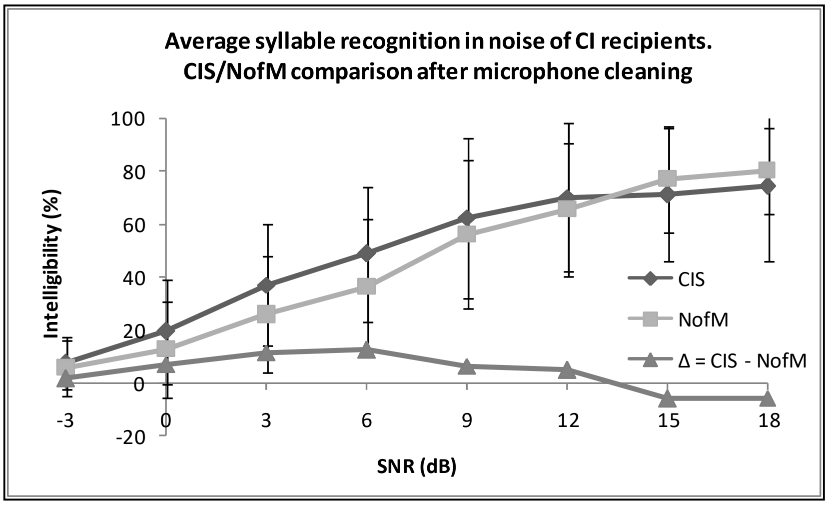
On figure 4 we indicate the attenuation due to soiling, corresponding to the 50% percentiles. The measured average attenuation was simplified by two straight lines. The signal attenuation started at 800 Hz at 2 dB per decade. No attenuation was applied between 0 and 800 Hz.

Fig 4: Frequency attenuation due to microphone soiling.



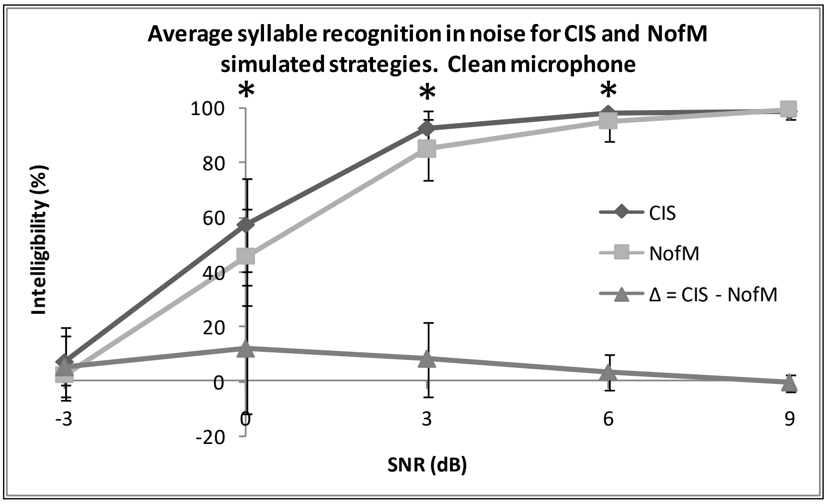
B. Syllable recognition scores in noise

Fig 5: Average syllable recognition function of the SNR, for each coding strategy("CIS-like" and n-of-m) and state of the microphone (Clean and Soiled). Bellow each graph we indicate the p-value relatively to the SNR and the p-value for the comparison of $x_{50\%}$ and $\Delta_{75\text{-}25\%}$.



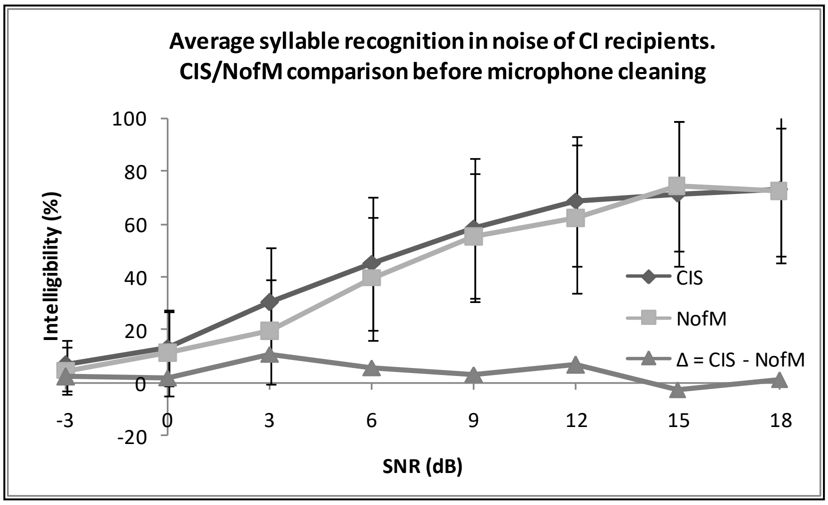
SNR (dB)	-3	0	3	6	9	12	15	18
p (Student)	0,61	0,24	0,10	0,12	0,48	0,56	0,37	0,38

	$x_{50\%}$ (dB)	$\Delta_{75\text{-}25\%}$ (dB)
CIS-like	3,95	7,52
n-of-m	6,16	6,09
p (Mann-Whitney)	0,0418	0,18



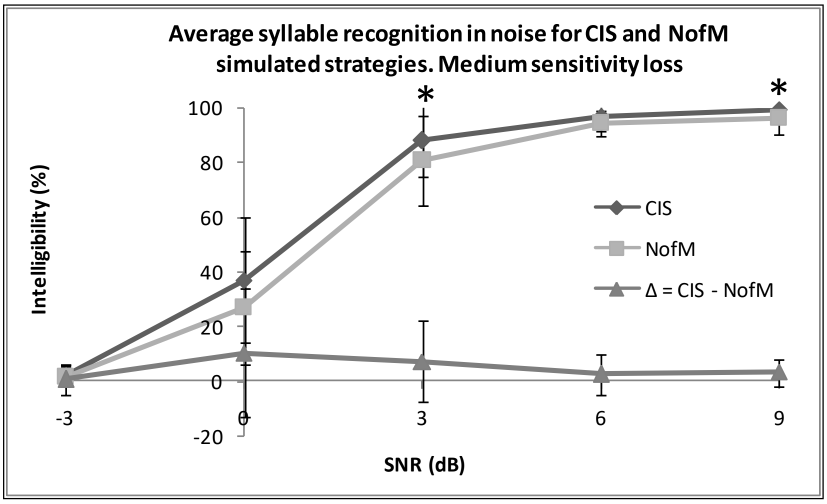
SNR (dB)	-3	0	3	6	9
p (Wilcoxon)	0,065	0,020	0,020	0,046	0,374

	$x_{50\%}$ (dB)	$\Delta_{75\text{-}25\%}$ (dB)
CIS-like	-0,28	2,18
n-of-m	0,38	2,54
p (Wilcoxon)	0,038	0,28



SNR (dB)	-3	0	3	6	9	12	15	18
p (Student)	0,44	0,67	0,08	0,43	0,67	0,42	0,73	0,89

	$x_{50\%}$ (dB)	$\Delta_{75\text{-}25\%}$ (dB)
CIS-like	4,48	6,45
n-of-m	6,44	6,35
p (Mann-Whitney)	0,0217	0,8512



SNR (dB)	-3	0	3	6	9
p (Wilcoxon)	0,615	0,089	0,040	0,156	0,020

	$x_{50\%}$ (dB)	$\Delta_{75\text{-}25\%}$ (dB)
CIS-like	0,65	1,87
n-of-m	1,31	2,04
p (Wilcoxon)	0,0419	0,83

C. *Discussion*

1) *Clean microphone*

Curves recorded with CI users show that people using "CIS-like" coding strategies reach higher speech recognition scores for the SNRs between 0 and 12 dB than people using n-of-m coding strategies. We cannot conclude here because no statistical differences were found. But, the feature $x_{50\%}$ is statistically better for "CIS-like" than for n-of-m. Moreover, for the extreme values of SNR (-3 dB, high noise level, and 18 dB, low noise level) scores are very similar for both coding strategies but we observed an advantage for n-of-m strategies at 18 dB of SNR.

Concerning the scores obtained with NH subjects we found approximately the same conclusions. For the intermediates SNRs (from 0 to 6 dB, scores are statistically different), when subjects are listening to the "CIS-like" simulation they reach higher recognition scores than when they were listening to the n-of-m simulation. Here, the feature $x_{50\%}$ is also better for "CIS-like" and the difference is statistically significant. Moreover, scores are also very similar for extreme SNRs (-3 dB and 9 dB).

2) *Soiled microphone*

Again, for CI users, "CIS-like" curve shows higher speech recognition scores than the n-of-m curve but no significant differences were found. Here $x_{50\%}$ is also better for "CIS-like" and the difference is statistically significant.

3) *General discussion*

Overall, we can say that "CIS-like" coding strategies give better $x_{50\%}$ than n-of-m coding strategies for CI users as well as for NH subjects. Conclusions are the same whatever if the microphone is soiled or not.

In accordance with the results of the pilot study [8], for the extreme values of SNRs recognition scores are the same for the two coding strategies. For recall we chose to select the N = 8 more energetic channels among the M = 20 for our n-of-m simulator. May be the criteria for selecting the stimulation channels must be studied further, knowing that the n-of-m strategies are used to reduce channel interaction and give to the cochlea only the "useful" information.

This work gives additional clues about sound coding strategies influence on speech recognition in noise but raise questions also, mainly with n-of-m strategies.

4. Conclusion

In this study we worked on the influence of the coding strategy used in cochlear implants on syllable recognition in noise.

We considered two states for the microphone: "Clean" and "Soiled". The syllable recognition in noise was tested on 65 subjects, 45 cochlear implant users and 20 normal hearing subjects.

Results showed that the "CIS-like" strategies seems to be more robust in front of a noisy environment than the n-of-m strategy and that the recognition scores are similar, for both strategies, for the extreme values of SNRs (-3 dB and 9/18 dB).

May be a homogenization of the populations is to consider and the criterion for selecting the stimulation channels should be studied further.

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, Ms Evelyne Veuillet for the links with the ethic committee, the members of the CRIC Lyon from the Edouard-Herriot hospital of Lyon for their collaboration, the subjects who participated in the study, 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

[1] R.Batmer, G.O'Donoghue, T.Lenartz "A multicenter study of device failure in European cochlear implant centers", Ear Hear, 28 (2 suppl), 95S-99S, 2007.

[2] G.Clark "Cochlear implants: fundamental and applications" Springer Verlag, 2003

[3] M.Dorman, P.Loizou, A.Spahr, E.Maloff "A comparison of the speech understanding provided by acoustic models of fixed-channel and channel-picking signal processors for cochlear implants. J Speech Lang Hear Res, 45(4), 783-8, 2002.

[4] T.Green, A.Faulkner, S.Rosen "Spectral and temporal cues to pitch in noise-excited vocoder simulation of continuous-interleaved sampling cochlear implants" J Acoust Soc Am, 112, 2155-64, 2002.

[5] Q.Nam Tran, H.Arabnia "Emerging trends in computational Biology, Bioinformatics and System Biology", Moran Kauffmann publishers, Elsevier, 2015.

[6] K.Perreaut, S.Gallego, C.Berger-Vachon, F.Millioz “Influence de l’encrassement du microphone sur les performances de l’implant cochléaire. Etude préliminaire en simulation des codages CIS et N of M“, Congress Handicap 2014, IFRATH, Paris, 11 – 13 June, 169-173, 2014.

[7] S.Razza, S.Burdo “An underestimated issue: unsuspected decrease of sound processor microphone sensitivity, technical and clinical evaluation” Cochlear Implants Int, 12(2), 114-23, 2011

[8] F.Seldran, S.Gallego, H.Thai-Van, C.Berger-Vachon “Influence of coding strategies in electric-acoustic hearing: a simulation dedicated to EAS cochlear implant in the presence of noise”, Applied Acoustics, 76, 300-30, 2014.

[9] P.O.Serra “Effet de l’entretien des aides-auditives sur leur performances”, Audiology dissertation, University of Montpellier, France, 2015.