Perception of fundamental frequency in cochlear implant patients

Angel de la Torre[(1)], *Cristina Roldán*[(2)], *Manuel Sainz*[(2,3)]

[(1)] Dpto. Teoría de la Señal, Telemática y Comunicaciones, University of Granada, Spain
[(2)] ENT Service, Hospital Universitario San Cecilio, Granada, Spain
[(3)] Dpto. Cirugía y sus Especialidades, University of Granada, Spain

Abstract

Fundamental frequency is important in speech perception, since it provides prosodic information to emphasize key words and contributes to speaker identification. Additionally, fundamental frequency contains phonetic value in tonal languages. In the case of hearing impaired patients who have received a cochlear implant, the perception of the fundamental frequency is affected by the technical limitations associated to the procedure of auditory nerve stimulation. In this paper we analyze the possibilities and limitations of cochlear implant systems with respect to perception of the fundamental frequency.

Index Terms: fundamental frequency, cochlear implant, frequency resolution, speech perception.

1. Introduction

Cochlear implants have significantly improved the treatment of severe and profound hearing loss [1]. A cochlear implant system consists of an internal element (implanted by means of surgery) and an external element. The internal element includes an electrode array (allocated inside the cochlea), a ground electrode (usually allocated below the temporal muscle), and a receptor-stimulator. The external part includes the batteries, a microphone, a processor and a transmitter to send information and energy to the receptor-stimulator.The system receives the sound, processes it, and sends electrical pulses into the cochlea, which stimulates the auditory nerve providing a hearing sensation to the patient. This way, for those patients affected by sensorineural hearing loss, the transduction mechanism is substituted by the cochlear implant [1, 2]. Currently, the effectiveness of cochlear implants is widely accepted. In most cases, the cochlear implant provides a perception of the sound with enough quality for speech understanding, and a high percentage of implanted patients are able to maintain fluent conversations without any visual aid (i.e. without lipreading) [3, 4].

However, hearing perception provided by cochlear implants presents some limitations due to the procedure of stimulation. Normal hearing involves around 6.000 inner hair cells, between 15.000 and 20.000 outer hair cells, and about 40.000 neural ends, while cochlear implants include a very reduced number of channels (currently between 8 and 32, depending on the model). On the other hand, there is a synaptic connection between hair cells and neural ends in normal hearing, while in cochlear implants, the electrodes spread current to a relatively wide region. Finally, the active mechanisms of tuning in normal hearing (associated to efferent innervation) are not possible with cochlear implants. These facts cause limitations in perception with cochlear implants (compared with normal hearing), like worse intelligibility of speech under noise conditions, or poor quality in perception of complex sounds (like music). These limitations are connected to the spectral resolution that can be obtained from the cochlear implant [5, 6].

In this paper we study the capability of perception of frequency by cochlear implant patients. Frequency resolution plays an important role in perception of fundamental frequency of speech signals. Perception of tone is important since it affects tone control when a cochlear implant patient speaks. It also affects, perception of supra-segmental information associated to fundamental tone, or speech recognition in the case of tonal languages. We analyze the mechanisms involved in the perception of tone and the limitations associated to stimulation with cochlear implants. We also present some results of tests designed to measure frequency resolution in patients wearing a cochlear implant. For this tests, pure tones and synthetic speech-like sounds (periodic signals generated using a speech production model) have been presented to the patients. The results show that current cochlear implant systems provide an acceptable degree of perception of the fundamental frequency for speech-like signals.

2. Perception of frequency in a normal ear and with a cochlear implant

2.1. Mechanisms of frequency perception

When the ear is stimulated with a pure tone, frequency perception is a complex process, based on the place of maximum displacement of the basal membrane and on the active mechanism of the hair cells in the organ of Corti. The point of maximum stimulation in the cochlea depends on the frequency of the tone and each point of the cochlea presents a characteristic frequency, according to the tonotopic theory. Spectral resolution in a normal ear depends on the number of inner hair cells and neural ends and it is close to 1/10 of tone (i.e, $\Delta f/f \approx 1.2\%$) for stationary pure tones in the case of a trained listener.

Audio signals are usually non stationary, and their properties (spectral distribution of the energy) change in time. The ear is able to perceive these variations due to the dynamic response of the hair cells and the auditory nerve. This way, the stimulation pattern transmitted by the auditory nerve changes in time according to the evolution in time of the audio signal. Due to the processes involved in generation of action potentials, after a firing of the hair cells or neurons they need some time before being ready for the next firing. This interval (the refractory period) is of about 2 ms, and limits the firing rate to about 500 firings per second, which limits the time resolution of the normal ear.

Because of this dynamic response of the ear, for signals with a fundamental frequency significantly lower than the maximum firing rate, the firings tend to be synchronized with the peaks of energy in time. This way, the fundamental frequency can be encoded in the pattern of activity in the auditory nerve. For signals with a fundamental frequency below 400 Hz, the tone can be perceived from this mechanism of time-based encoding.

Therefore, frequency can be perceived either by tonotopic mechanism or by time encoding mechanism. The last one is more important for low frequencies.

2.2. Limitations associated to electrical stimulation

There are important differences in the generation of action potentials between normal hearing and cochlear implant-based hearing. The synaptic connection between the hair cells and the neural ends cause that vibration in a given point of the basal membrane produces action potentials only in the neural ends connected to this point of the cochlea. However, the electrical stimulation provided by the electrodes produces a field of electric currents in a non confined region, and it generates action potentials in a wider region of the cochlea (compared to the case of normal hearing) [7].

On the other hand, in normal hearing, the generation of action potentials in a given hair cell does not affect to the hair cells in the vicinity, and therefore activity in one area of the cochlea does not interfere the activity in a different area. In the case of cochlear implants, if two channels are simultaneously activated, there will be an interference between both stimulations because: (a) the activation of the first channel is produced by creating a difference of potential between the first active electrode and the ground electrode $(V_1 - V_{GND})$; (b) the activation of the second channel is obtained by a difference of potential between the second and the ground electrodes $(V_2 - V_{GND})$; and (c) in general, there will be an uncontrolled difference of potential between first and second active electrodes $(V_1 - V_2)$ causing an uncontrolled stimulation of the neural ends between both active electrodes. For this reason, simultaneous stimulation is avoided in most cochlear implant systems, and electrical pulses are sequentially presented at the different channels, with only one channel activated at each moment.

The previously described limitations of cochlear implants are common to the different designs, and are a consequence of the fact that cochlear implants use electrodes that spread electric current inside the cochlea in order to activate the auditory nerve. Current technology does not allow a better connection between electrodes and neural ends for a more selective excitation of the neural ends or providing simultaneous stimulation without interference. For this reason, all the cochlear implant systems present a very small number of channels (in comparison with the number of inner hair cells), and this causes a very reduced tonotopic spectral resolution. Taking into account the interaction among channels (due to the spread of electric current field around the electrodes), obtaining tonotopic spectral resolutions better than 12 channels per decade (or better than $\Delta f/f \approx 21\%$) would be very difficult [7].

However, perception of tone by cochlear implant patients usually improves this tonotopic limit. There are evidences of it, since patients are able to control their own tone quite well. There are also implanted patients who are able to sing in tune, which means that their spectral resolutions are better than 1/4 tone ($\Delta f/f < 3\%$). In such cases, the spectral resolution is associated to the time pattern of stimulation provided by the cochlear implant. Cochlear implants with a high stimulation rate (providing, at each electrode, pulses at a rate significantly higher than the maximum firing rate of the auditory nerve) allow the perception of the fundamental frequency from the time pattern of stimulation, because the highest stimulation levels (and therefore most of the neural activity) are synchronized with the peaks of energy of the signal, in a similar way as normal hearing.

Current cochlear implant systems are able to provide high stimulation rates (some models can reach several thousands of pulses per second at each electrode). In these cases, the temporal resolution and the spectral resolution associated to the time encoding mechanism would be close to those obtained with normal hearing.

2.3. Spectral resolution based on tonotopic and time encoding mechanisms

Perception of sound based on tonotopic and time encoding mechanisms can be illustrated by analyzing signals with spectrograms. Narrow-band spectrograms would show which aspects of the signal can be perceived with a good tonotopic spectral resolution, where frequency tuning dominates over time synchronization. Wide-band spectrograms would show the aspects that would be perceived from the time pattern of neural activity, where synchronization is more important.

In the case of pure tones, frequency resolution requires a good capability for tuning, and due to the stationarity of the signal, it only depends on the tonotopic perception. In the case of speech-like signals (for example, synthetic signals generated with a filtered periodic train of impulses), the narrow-band spectrogram show spectral peaks corresponding to the harmonic series, while wide-band spectrogram show the peaks of energy in time (which would corresponds to the glotal pulses in the case of a speech signal). In that case, the fundamental frequency could be obtained from the tonotopic mechanism (narrow-band spectrogram) or from the time encoding mechanism (wide-band spectrogram). This concept can be applied to other harmonic periodic sounds.

As previously discussed, tonotopic spectral resolution is very limited with cochlear implants, due to the low number of channels. However, time resolution is high, particularly for high stimulation rate systems. Therefore, in this case, the spectral resolution obtained from the time encoding mechanism is adequate for perception of fundamental frequency in speech-like sounds.

Figure 1 illustrates this with some examples. For three signals, we represent a narrow-band spectrogram (resolution: 16 Hz), a wide band spectrogram (resolution: 200 Hz), the amplitude of the signal and the "stimulogram" or representation of the pattern of activity at each channel of the cochlear implant. The stimulogram was obtained for the default configuration of a MED-EL Combi40+ cochlear implant (using 12 channels and a stimulation rate of 1527 pulses per second at each channel). Three different signals were analyzed: a 404 Hz pure tone, a synthesized speech-like signal with a fundamental frequency of 150 Hz, and a speech signal corresponding to the syllable /pa/ (a Hamming window has also been applied to these signals).

In the case of pure tones, the narrow-band spectrogram shows a spectral line. This line is significantly wider in the wide-band spectrogram, due to its lower spectral resolution. The signal is quasi-stationary (the signal was modulated with a 0.4 seconds Hamming window). The stimulogram provides a pattern of activity similar to that of the wide-band spectrogram, where most of the energy is concentrated on channel 2 in this case, and adjacent channels also provide some stimulation due to the filter bank design. The spectral resolution provided by the stimulogram is very poor and an accurate estimation of the frequency would be very difficult for a patient with this stimulation pattern.

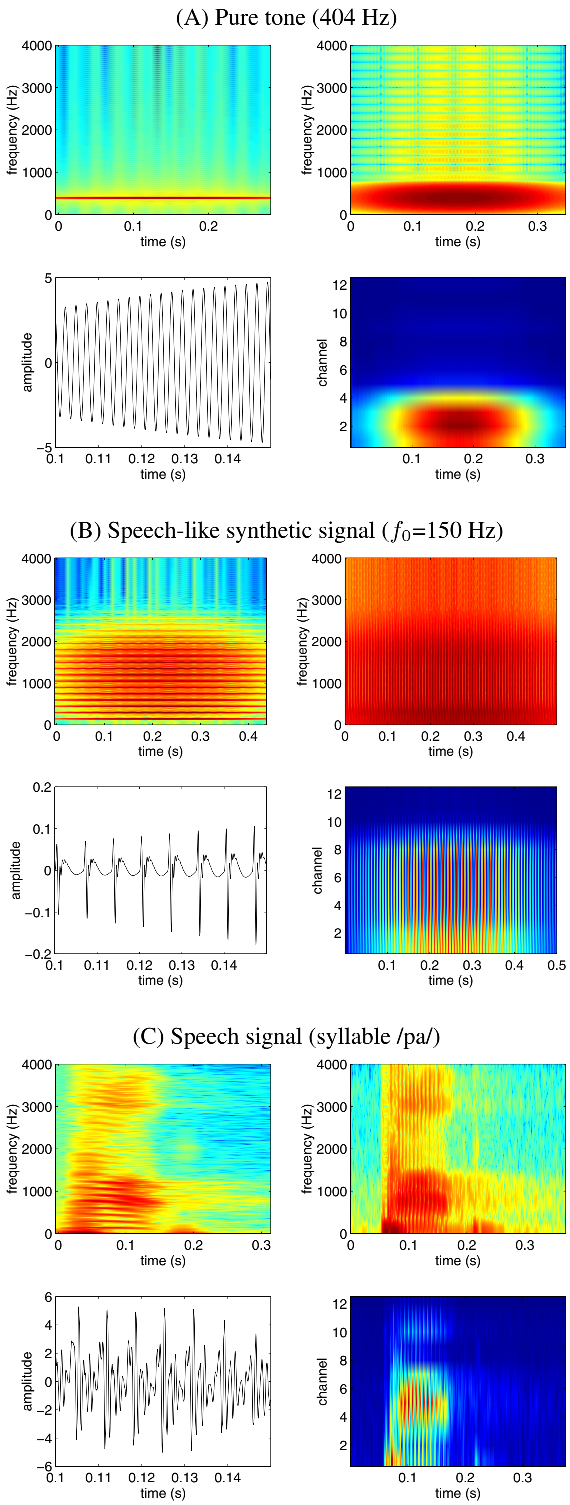

Figure 1: Perception of the frequency for three signals: (A) pure tone; (B) speech-like signal synthesized according to a speech production model; (C) speech signal corresponding to the syllable /pa/. For each signal we represent: (1) narrow-band spectrogram (resolution: 16 Hz); (2) wide band spectrogram (resolution: 200 Hz); (3) the amplitude of the signal in time; and (4) the "stimulogram" or pattern of activity for each channel (for a MED-EL Combi40+ cochlear implant).

In the case of the synthetic speech-like signal, the narrow-band spectrogram provides a series of harmonics associated to the periodicity of the signal. Fundamental frequency can be obtained from the first spectral line, or by the difference in frequency between two consecutive spectral lines. The wide-band spectrogram does not provide the series of harmonics due to its poor spectral resolution (compared to the fundamental frequency of 150 Hz). However, the periodicity of the signal can be observed in time, because the analysis window is short enough to detect peaks of energy in time. This way, the wide-band spectrogram shows that the signal is periodic with a fundamental period of 6.6 ms, which corresponds to a fundamental frequency of 150 Hz. This time-based encoding of the frequency is also observed in the stimulogram, and therefore, the stimulation provided by cochlear implants allows resolving the fundamental frequency for such speech-like signal. For a real speech signal, a similar behavior is observed. Cochlear implants allow, on one hand the perception of formants (which allows identification of phonemes) and on the other hand the perception of the fundamental frequency from the time pattern of stimulation.

This analysis illustrates that cochlear implants allow perception of fundamental frequency based on the time encoding mechanism. In order to perceive fundamental frequency, a high stimulation rate in the cochlear implant system is necessary (in order to allow an appropriate time representation), and in any case, the temporal resolution will be limited by the maximum firing rate of the neural ends of the auditory nerve. In addition to the technical and physiological aspects, the ability for perception of the fundamental frequency will also depend on how the patient is able to extract information from the hearing provided by the cochlear implant, and therefore, it requires a development, learning or training of certain hearing abilities.

3. Evaluation of the frequency resolution

In order to study frequency resolution (both, tonotopic and time-encoding based), we have prepared subjective tests with cochlear implant patients. Normal hearing subjects have also been tested for comparisons. We have included results from 10 normal hearing subjects and 10 patients implanted with the MED-EL Combi40+ system.

For different frequencies, spectral resolution has been measured for pure tones (in order to evaluate tonotopic spectral resolution), and for synthetic speech-like signals (in order to evaluate the spectral resolution based on time-encoding mechanism). Signals were presented to the subjects in pairs, with frequencies f_0 and $f_1 = f + \Delta f$, and they were asked whether the stimuli were perceived similar or different. In order to verify the answer, pairs of signals with the same frequency ((f_0, f_0) or (f_1, f_1)) or with different frequencies ($(f, f_1$ or (f_1, f)) were randomly presented. The frequency resolution has been measured in percentage, taking into account the ratio $\Delta f / f_0$, where Δf is the difference of frequency between two signals that can be systematically distinguished by the subject.

Figure 2 shows the results of the frequency resolution tests (mean and standard deviation as a function of the frequency). In the case of normal hearing subjects, tonotopic frequency resolution reaches the maximum value (mean value close to 0.8%) around 1 kHz. Implanted patients present a significantly worse tonotopic frequency resolution, with mean values reaching 8% at lower frequencies and with a faster degradation for high frequencies (around 30% at 4 kHz). A higher dispersion is also observed

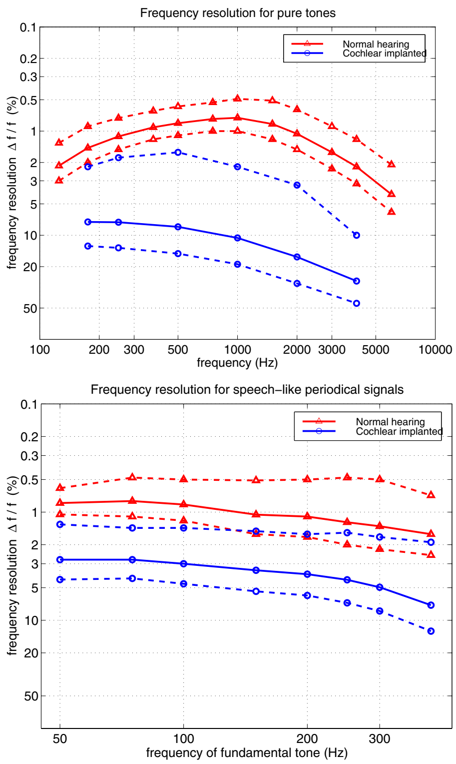

Figure 2: Evaluation of the frequency resolution for normal hearing and cochlear implant subjects. Dashed lines represent mean ± standard deviation.

in the results of cochlear implant subjects.

Time-encoding based frequency discrimination is better for low frequencies than for high frequencies

For low frequencies, time encoding based frequency discrimination is better than tonotopic frequency discrimination for both groups, normal hearing and cochlear implant subjects. This is consistent with the fact that both, synchronization capability of the neural ends and capability of analysis of the time pattern of activity are better for lower frequencies. In this test, frequency resolution is close to 1% for normal hearing subjects and close to 4% for implanted subjects, and the differences between both groups are smaller than in the case of tonotopic spectral resolution.

Tonotopic spectral resolution provided by cochlear implants is very poor (compared to normal hearing subjects). However, cochlear implants provide, from the time pattern of stimulation, enough frequency resolution for perceiving variations of the fundamental frequency in a speech signal. Variations of tone in a sentence, which typically vary in the range between $\Delta f/f$=20% and $\Delta f/f$=50% (depending on the speaker) would be clearly perceived by an implanted subject. Frequency resolution is also adequate for perceiving the musical tone for a high percentage of implanted subjects (a difference of 1/4 of tone corresponds to $\Delta f/f$=3%). This capability would be verified for signals like singed speech and also for other sounds (or musical instruments) with low fundamental frequencies, slow attack and decay periods and harmonically rich spectrums.

The differences observed among the different implanted patients regarding frequency resolution seem to be connected to several factors such as the aetiology of the hearing-loss, the state of the neural ends, the experience with the use of the cochlear implant, and the hearing training. Also a better frequency resolution has been observed for patients with higher stimulation rates.

4. Conclusions

Cochlear implants provide a poor tonotopic spectral resolution, due to the technical limitations associated to the electrical stimulation. However, they provide an acceptable degree of frequency resolution for speech-like sounds. This resolution is obtained from the time pattern of stimulation according to the time-encoding mechanism. For this reason, coding strategies oriented to preserving time resolution (with high stimulation rate) provide better perception of the fundamental frequency.

Main factors involved on perception of the fundamental frequency can be grouped into technical (stimulation rate or coding strategy), and physiological (state of the neural ends, and capability for synchronization of the neural response with the electrical stimuli). Hearing training and the experience with the cochlear implant system also affect the perception of the fundamental frequency. Cochlear implants with a high stimulation rate provide an adequate frequency resolution for fundamental frequency of speech signals. In some cases, they also allow a perception of fundamental frequency appropriate for music. In order to develop such abilities, hearing training programmes for implanted patients should include specific training for the perception of fundamental frequency as well as exercises oriented to musical education.

5. References

[1] B. Wilson, C. Finley, D. Lawson, R. Wolford, D. Eddington and W. Rabinowitz. “Better speech recognition with cochlear implants”. *Nature*, (352):236–238, 1991.

[2] P.C. Loizou. “Mimicking the human ear”. *IEEE Signal Processing Magazine*, pages 101–130, 1998.

[3] P. Blamey, P. Arndt, F. Bergeron, G. Bredberg and J. Brimacombe. “Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants”. *Audioogy and Neuro-otology*, (1):293–306, 1996.

[4] R.S. Tyler, H. Fryauf-Bertschy, D.M. Kelsay, B.J. Gantz, G.P. Wood-worth and A. Parkinson. “Speech perception by prelingually deaf children using cochlear implant”. *Otolaryngol. Head Neck Surg.*, (117):180–187, 1997.

[5] M. Sainz, A. de la Torre and C. Roldán. “Frequency resolution in cochlear implant patients”. *6th Europ. Symp. on Paediatric Cochlear Implantation*, Feb 2002.

[6] A. de la Torre, M. Sainz and C. Roldán. “Perception of the tone with cochlear implants. Influence of the coding strategy”. *7th Intern. Cochlear Implant Conference*, Sep 2002.

[7] A. de la Torre, M. Sainz and C. Roldán. “Influence of placement and configuration of the electrodes over the perception using cochlear implants”. *6th Europ. Symp. on Paediatric Cochlear Implantation*, Feb 2002.