Advances in Applied Physiology
2020; 5(2): 58-69
http://www.sciencepublishinggroup.com/j/aap
doi: 10.11648/j.aap.20200502.18
ISSN: 2471-9692 (Print); ISSN: 2471-9714 (Online)

A Possible Way to Improve the Quality of Speech Perception by Increasing the Number of Electrodes in a Cochlear Implant from 8 to 22

Margarita Stefanovich

Independent Researcher, Saint-Petersburg, Russia

Email address:
marg.stefanovich@yandex.ru

To cite this article:
Margarita Stefanovich. A Possible Way to Improve the Quality of Speech Perception by Increasing the Number of Electrodes in a Cochlear Implant from 8 to 22. *Advances in Applied Physiology*. Vol. 5, No. 2, 2020, pp. 58-69. doi: 10.11648/j.aap.20200502.18

Received: November 21, 2020; **Accepted**: December 8, 2020; **Published**: December 31, 2020

Abstract: A cochlear implant (CI) helps a person with auditory receptor pathology restore the ability to hear and understand speech. The partial loss of the information about the frequency and loudness of the audio signal is the main difference from the natural sound perception. CIs of various types differ in the number of stimulated electrodes: 12, 16, 20, or 22. The quality of speech perception is significantly improved by increasing the number of stimulated electrodes from 4 to 7 and almost does not change with a further increase in the number of electrodes from 8 to 22. When speech is perceived in all types of CI, 3 to 6 adjacent electrodes are usually stimulated simultaneously. When a single electrode is stimulated, the CI user hears a sound with a single pitch. When two adjacent electrodes E(n) and E(n+1) are simultaneously stimulated, a virtual frequency channel (VFC) is formed, and the CI user hears a sound with an intermediate pitch {Z(n)+Z(n+1)}/2, provided that the adjacent electrodes on the right and left are not stimulated. When three or more adjacent electrodes are simultaneously stimulated, the CI user hears a sound with the same pitch, which depends on the distance from the electrode with the largest amplitude of stimulating pulses to the apex of the cochlea. There was no suggestion as to why the quality of speech perception does not improve when the number of electrodes increases from 8 to 22 in the scientific literature for the period from 1997 to 2019. This article proposes a method for adjusting the mode electrodes stimulation, which creates conditions for the formation of virtual frequency channels. The CI user will be able to hear a greater number of sounds of various pitch. A method for correcting the algorithm for controlling the stimulation of electrodes, which can be used in any type of CI, is proposed.

Keywords: Cochlear Implantation, Virtual Frequency Channel, Speech Perception

1. Introduction

Within the human auditory system, information about the frequency and intensity of sound is converted into an increase in the receptor potential (RP) in a group of inner hair cells (IHC) located in the cortical organ along the basilar membrane of the cochlea (BM). Each IHC is innervated by several spiral ganglion cells (SGC) with different excitation thresholds. Axons of SGC form auditory-nerve fibers (ANFs). SG cells generate pulses when the RP of innervated IHC increases. Pulses are transmitted to higher parts of the auditory system via the AN fibers. In the case of IHC pathology, hearing significantly deteriorates, up to complete deafness. Auditory system experts assumed that the CI user could hear sound signals if the SGCs retain the ability to generate pulses when excitatory signals arrive at the synapses. In 1961, William Fouts House (1923-2012), Los Angeles, performed the first intracochlear implantation with a cochlear implant he created himself. The main part of the cochlear implant (CI) is a thin, elastic tube containing platinum electrodes. The results of experiments show that the quality of speech perception by CI users significantly improves with an increase in the number of stimulated electrodes from 4 to 8, and then almost does not change with an increase in the number of electrodes from 8 to 22 [1-4]. In scientific literature for 20 years from 1997 to 2019, there are no assumptions about the cause of this phenomenon and about the possibility of improving the quality of speech perception with an increase in the number of electrodes. As a result, there is no competition between manufacturers of cochlear implants (CI) for the main parameter: the quality of speech perception. The article discusses the features of perception of audio signals through CI in comparison with normal hearing and suggests a possible way to correct the program for controlling the stimulation of electrodes to improve the quality of speech perception.

2. Factors That Affect the Quality of Speech Perception

The study of the potential improvement of the quality of speech perception by CI users is based on the analysis of the frequency spectrum of speech signals. When audio signals are perceived through the CI, information about the frequency and intensity of the audio signal is converted into a sequence of pulse signals to stimulate the preserved cells of the spiral ganglion. In this case, some of the information is lost irrevocably. The results of experiments on the quality of speech perception by CI users with different numbers of stimulated electrodes are known [1-4]. Experiments were conducted to compare the quality of speech perception when using different types of speech signal conversion strategies in the same CI: CIS, ACE, MDE2, PDT, and MEM [5]. However, according to the results of all experiments, the quality of speech perception depends mainly on the individual characteristics of CI users and does not depend on increasing the number of electrodes from 8 to 22. A comprehensive analysis of the results of known experiments suggests the reason for the inability to improve the quality of speech perception in the standard mode of stimulation of the electrodes.

2.1. Block Diagram of a Cochlear Implant (CI)

For electrical stimulation of preserved cells of the spiral ganglion, a thin flexible tube is surgically inserted into the inner ear, within which platinum electrodes (an electrode array) are linearly located. Figure 1 shows a typical block diagram of a cochlear implant (CI) [6]. When each individual electrode is stimulated by electrical pulses, the preserved SGC located near the electrode at a distance of up to 3 mm generates pulses with exactly the same parameters as to when natural SGC stimulation occurs in normal hearing [7]. The audio signal from the microphone output (1) is transmitted to the inputs of the group of parallel-connected band-pass filters (BF) (2). The frequency spectrum is divided into the number of frequency bands corresponding to the number of stimulated electrodes. For example, the CI of type Med-El (Pulsar/Sonata/Concerto implant) has 12 active electrodes and 12 band-pass filters, respectively. The CI of type Cochlear CI24RE has 22 active electrodes and 22 band-pass filters. Each electrode is stimulated by the perception of an audio signal with a frequency within the bandwidth of the corresponding BF. At the output of the BF, the amplitude detector (3) is enabled. The rectified voltage at the detector output determines the mode of stimulation of the corresponding electrode. It should be noted that at the output of the amplitude detector, information about the change in the frequency of the audio signal within the BF bandwidth is lost irrevocably.

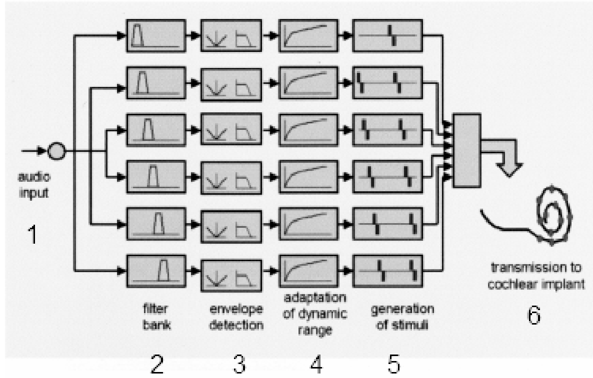

Figure 1. *Block diagram of the CI [6].*

The frequency range of the audio signal is divided by the number of subbands in accordance with the number of electrodes in the CI. When a speech signal is perceived, the change in sound intensity from the threshold of audibility to the maximum comfortable level can be more than 50 dB. When the electrode is stimulated, the change in the amplitude of the stimulating pulses from the threshold of audibility to the maximum comfortable level of loudness is on average no more than 30 dB [2]. To coordinate the dynamic range of speech signals and the possible range of changes in the amplitude of the pulses, a special unit of the dynamic range converter (DRC) (4) is switched on at the detector output. The quietest signals get additional amplification, and the loudest ones get additional attenuation. This way the CI user can hear both the quietest and loudest sounds. The generator of stimulating pulses (5) converts the rectified voltage at the output of the adaptation of dynamic range (ADR) (4) into pulses with a certain amplitude, duration, and frequency of repetition. The parameters of stimulating pulses depend on the strategy of converting the envelope of the speech signal at the BF output into stimulation pulses. For example, the CI of firm Med-El uses pulses with a duration of 24-40 microseconds (mcsec) and a frequency of 1000-6000 Hz to stimulate the auditory nerve. The maximum frequency of stimulation in modern cochlear implantation systems can be more than 50000 Hz [4]. Through the induction coil (6), stimulating pulses are transmitted to the inner ear to the electrode array. The amplitude of the stimulating pulses for each electrode depends on the value of the rectified voltage at the output of the corresponding BF. When a single electrode is stimulated, the CI user hears a sound with one specific pitch, which depends on the type of CI and the distance of the electrode from the apex of the cochlea.

2.2. Features of the Auditory Sensations of CI Users in the Perception of a Tonal Audio Signal

Let's assume that a person with normal hearing hears a tone signal. The audible pitch is measured in conventional units "mel". The audible pitch is directly proportional to the frequency of the signal in Hz in the frequency range up to 500 Hz and is proportional to the logarithm of the frequency change in the range from 1 kHz to 8 kHz. A normally hearing person remembers the audible pitch by the ordinal number of the IHC located in the middle of the general group of receptor cells with increased receptor potential (RP) [8]. What will the CI user actually hear when receiving a tone signal with a frequency of 1 kHz? Table 1 shows, for example, the BF frequency ranges for a CI with 12 electrodes. When a tone signal is received at a frequency of 1 kHz, the E7 electrode will be stimulated, which is located near the point of the basilar membrane of the cochlea at a distance of about 10,5 mm from the apex of the cochlea.

Table 1. *Frequency range of the BF for stimulation of 12 electrodes in the Med-El CI [9].*

E	**E12**	**E11**	**E10**	**E9**	**E8**	**E7**
F, Hz	150–240	240–357	357–510	510–710	710–973	973–1319
E	E6	E5	E4	E3	E2	E1
F, Hz	1319–1776	1776–2380	2380–3180	3180–4238	4238–5641	5641–7500

The audible pitch will mainly depend on the distance from the E7 electrode to the apex of the cochlea. It is important to note that when the electrode is stimulated from the pulse generator (5), pulses with exactly the same parameters are transmitted to the 1st projection of the cochlear frequency scale (PCFS) in the cochlear nucleus complex (CCN), as in the natural perception of sound. The auditory system analyzes the pulse intensity of auditory neurons linearly located on the 1st PCFS. The algorithm for the perception of information about the frequency and (loudness) of an audio signal in the human auditory system is the same for normal hearing and perception through CI. The audible pitch in the natural perception of the sound signal depends on the distance from the point of the basilar membrane of the cochlea with the greatest amplitude of transverse vibrations to the apex of the cochlea. This point is called the "excitation point of the cochlea membrane (PMVM)". The auditory system selects 146 neurons from the general group of auditory neurons with an increased pulse intensity, which are located in the middle of the general group of neurons with an increased pulse density [8].

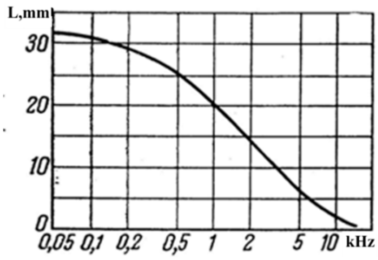

Figure 2. *The correlation between the maximum transverse oscillations of the BM of the cochlea and the frequency of the tonal signal. The abscissa axis is the tone frequency in kHz, the ordinate axis is the distance from the L base in mm [10].*

In normal hearing, when receiving a tone signal with a frequency of 1 kHz, the PMVM is located at a distance of 12 mm from the apex of the cochlea. In CI of firm Med-El, the E7 electrode is located at a distance of approximately 18 mm from the apex of the cochlea. For this reason, Cochlear Implant user hears a sound with a different pitch when receiving a tone signal with a frequency of 1 kHz than he would have heard a signal with a frequency of 1 kHz with normal hearing.

When an audio signal is perceived through the CI, the audible pitch depends on the location of a group of spiral ganglion cells, whose synapses receive the largest number of signal pulses from the stimulated electrode. When a single electrode is stimulated, stimulating pulses act on the SGC located at a distance of up to 3 mm from the electrode [7]. The closer to the stimulated electrode the SGC is located, the greater is the signal pulse density transmitted along with the auditory nerve fiber. When a tone signal is perceived in normal hearing, the sound wave propagates along the basilar membrane from the base to the apex of the cochlea. The complex of elements of OHC-BM-IHC is a chain of sequentially connected parallel resonant circuits according to biophysical characteristics [11]. At the BM point, where the frequency of the sound wave most closely coincides with the frequency of the resonant circuit, the greatest amplitude of transverse vibrations will be observed. When an audio signal is perceived through the CI, stimulating pulses are sent to SGC synapses, if there is an audio signal at the input of the corresponding BF. However, due to the inclusion of an amplitude detector at the output of the BF, information about the frequency of the audio signal at the input of the BF is lost irrevocably.

For the CI user, the audible pitch will only depend on the distance from the stimulated electrode to the apex of the cochlea. The auditory system registers an increase in the pulse density in a group of auditory neurons on the 1st projection of the Cochlea Frequency scale (PCFS), which innervate SGC located near the stimulated electrode. By the standard algorithm for the perception of information about the frequency of audible sound, the auditory system selects 146 central auditory neurons from the general group and transmits information about the ordinal number N0 of the auditory neuron located in the middle of the group of neurons with an increased pulse intensity to the long-term memory block [8]. The CI user hears a sound with the pitch that he heard before hearing loss when the same group of auditory neurons is excited on the 1st PCFS.

With sequential stimulation of single electrodes, the CI user can hear as many different sounds in pitch, as the frequency range is divided into bandpass filters (BF). Through CI with 12 electrodes, the user can hear 12 different sounds in pitch. Using the Advanced Bionics CI with 16 electrodes, the CI user can hear 16 different sounds in pitch. Through the CI24RE CI with 22 electrodes – 22 different in pitch sounds. When two adjacent electrodes are simultaneously stimulated, the CI user can hear a sound with an intermediate pitch. For example, when stimulating the electrode E1, the CI user hears a sound with a pitch Z1, when stimulating the electrode E2, the user hears a sound with a pitch Z2. When simultaneously stimulating the electrodes E1 and E2, the CI user can hear a sound with an intermediate pitch Z1/Z2. When two adjacent electrodes are simultaneously stimulated, a "virtual frequency channel" is formed [12, 13]. In this case, most CI users hear a sound with an intermediate pitch (Z1+Z2) / 2. Sequential stimulation of pairs of adjacent electrodes by pulses with approximately the same amplitude the CI user will hear an additional 11 sounds with different pitch via CI with 12 electrodes, more 15 sounds of the different pitch via CI with 16 electrodes, and 21 sound with a different pitch by using CI with 22 electrodes. When three or more neighboring electrodes are simultaneously stimulated, the VFC is not formed, since the middle electrode is simultaneously affected by the stimulation of two neighboring electrodes on the left and right. CI user will hear a sound with a pitch that depends on the stimulation of the "loudest" electrode.

In normal hearing, when a tonal audio signal is perceived, the audible loudness is proportional to the total pulse density of the auditory nerve fibers (AN), which innervate 146 inner hair cells located in the middle of the general group of IHC with increased RP.

In normal hearing, SGC synapses with different excitation thresholds connect to different points on the innervated inner hair cell (IHC). On average, each IHC is innervated by 6 SG cells. The threshold for the excitation of SGC depends on which point of the IHC the SGC synapse connects to. For example, if the intensity of the audio signal is not more than 40 dB above the threshold of audibility, the pulse density will increase only in those fibers of the auditory nerve that have an excitation threshold of 20 dB and 40 dB. When an audio signal is perceived through the CI and a single electrode is stimulated, stimulating pulses arrive at the synapses of spiral ganglion cells with different excitation thresholds simultaneously. As a result, the total pulse intensity of several SGC, which were united at the synapses of one summing auditory neuron (SAN) before hearing loss, increases avalanche-like to the pain threshold. There is a need to artificially reduce the amplitude of stimulating pulses. The pulse density of the SGC increases artificially when quiet audio signals are perceived. The pulse density is artificially reduced when loud audio signals are perceived. The dynamic range of changes in the pulse density of summing auditory neurons on the 1st PCFS is approximately 30 dB, which is significantly less than the dynamic range of changes in loudness of the signal at the BF inputs.

3. Stimulation of Electrodes in a CI with 22 Electrodes in the Perception of (I) Sound

The results of experiments on the structure of control signals for the stimulation of electrodes in the perception of vowel sounds and syllables are known. Figure 4 shows the amplitudes of stimulatory pulses in the perception of (I) sound in relative units, with the frequency of the main tone of 98 Hz and 740 Hz. For two variants of the (I) sound, several adjacent electrodes are simultaneously stimulated: {(E22 – E19) + (E12 – E7)} and {(E19 – E17) + (E14 – E9)}.

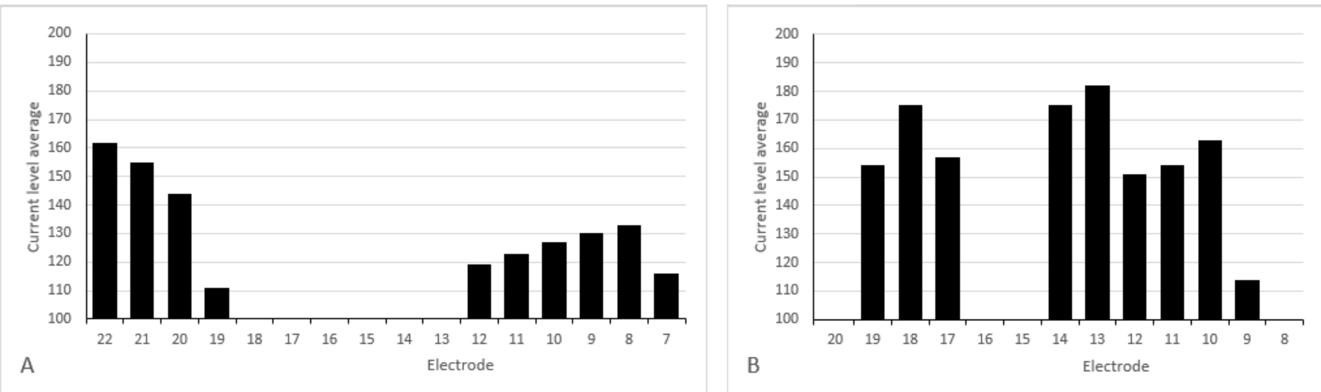

Figure 3. Mode of stimulation of electrodes in the perception of (I) sound [15]. The abscissa shows the number of stimulated electrodes, and the "y" axis shows the amplitude of the stimulating pulses. According to Figure 4A, the main frequency of the tone is 98 Hz, according to Figure 4B – 740 Hz [14].

Figure 4 shows the spectrum of the vowel (I) sound with the main frequency of approximately 250 Hz [13]. The location of electrodes that are stimulated when a vowel sound is perceived depends on the division of the total frequency range by bandpass filters (BF) of individual channels that are set when configuring the CI. In the frequency spectrum of the vowel (I) sound, there are 3 formants: (100-500) Hz, (2500-2800) Hz, and (3000-4300) Hz. Table 2 shows a possible distribution of the frequency spectrum of (I) sound along the channels of stimulation of electrodes in the CI (22).

Table 2. Distribution of the frequency spectrum of sound (I) along the channels of electrodes stimulation in CI (22) [7].

Electrode	E22	E21	E20	E10	E9	E8	E7
F, Hz	120–280	280–440	440–600	2460–2860	2860–3350	3350–3920	3920–4600

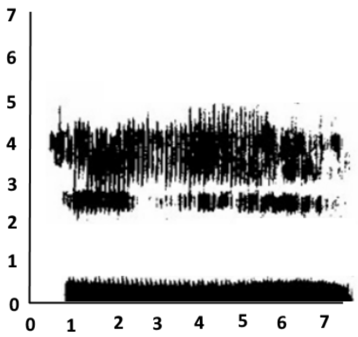

Figure 4. *The frequency spectrum of a vowel (I) sound [15].*

When the (I) sound is perceived (according to Figure 5), 2 groups of neighboring electrodes (E22 – E20) and (E10 – E7) are stimulated simultaneously in the CI (22) with the frequency spectrum distribution according to Table 2.

Figure 4 shows the ratio of the amplitudes of stimulating pulses in the perception of (I) sounds with the frequency of the main tone of 98 Hz and 740 Hz. However, there is no information about the frequency range distribution between bandpass filters. Table 2 shows stimulated electrodes in the perception of (I) sound with the main frequency of approximately 250 Hz but do not indicate the ratio of the amplitude values of the stimulating pulses for individual electrodes. In this case, to understand what the CI user actually hears, more information can be obtained from the diagram in figure 4.

When neighboring electrodes E22–E21–E20–E19 are simultaneously stimulated, the user hears a single sound with a pitch that depends mainly on the stimulation of the E22 electrode. The stimulation of E21, E20, and E19 electrodes has little effect on the audible pitch. When a group of neighboring electrodes (E12—E7) is stimulated, the user hears a single sound with a pitch that depends on the mode of E8 and E9 electrodes stimulation. When (I) sound is perceived with a pitch frequency of 740 Hz (Figure 4B), the CI user hears 3 different sounds with a pitch that depends on the stimulation of the E18, E13, and E10 electrodes. Figure 6 shows the location on the 1st projection of the cochlea frequency scale (PCFS) of a group of auditory neurons with an increased pulse intensity in the perception of (I) sound with the main frequency of 740 Hz. Auditory neurons that innervate groups of SGC located at a distance of about 15 mm, 20,5 mm and 23 mm from the apex of the cochlea have the highest density of pulses. The CI user simultaneously hears three audio signals with a pitch of Z18, Z13, and Z10.

The envelope of the general group of auditory neurons with an increased pulse density has 3 inflection points at a distance of 15, 20,5, and 23 mm from the apex of the cochlea. The human auditory system works in such a way that information about the frequency of the audible tone is stored in a linear coordinate relative to the apex of the cochlea of the auditory neuron, which has the highest intensity of pulses. Information about the audible loudness is stored by the pulse density of 146 auditory neurons, that are located symmetrically *relative to the neuron with the highest pulse intensity [8]*.

Notably, when a multi-frequency audio signal is perceived through the CI, stimulating pulses arrive simultaneously to the synapses of a large number of SGC. As can be seen in Figure 6, most SG cells located along a 13 mm OM segment generate pulses. At the 1st PCFS, approximately 1000 auditory neurons are in a state of excitation. When a large number of auditory neurons are simultaneously excited on the 1st PCFS, the auditory system analyzes the pulse density of only such groups of auditory neurons in which the envelope has an inflection point [8].

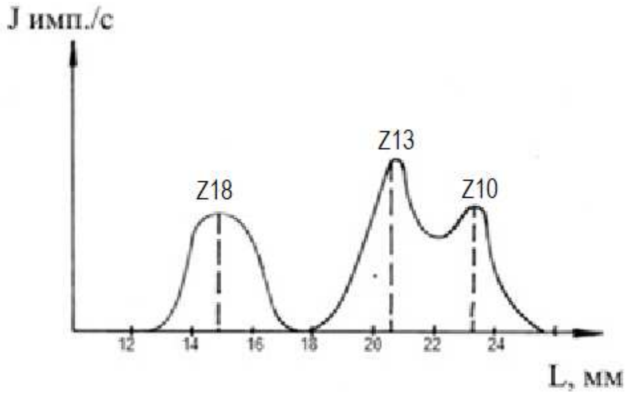

Figure 5. *The envelope of the general group of excited auditory neurons on the 1st projection of the FCS in the CCN at the perception of a vowel (I) sound (Figure 4B). The abscissa shows the distance in mm from the apex of (the) cochlea; the ordinate shows the relative density of summing pulses in auditory neurons (SAN) on 1st PCFS.*

The envelope of a large group of auditory neurons with an increased pulse intensity does not have an inflection point if the pulse density increases or decreases monotonously. This situation is observed when a noise signal is perceived. When hearing noise with a band less than two octaves in normal hearing, the audible pitch depends on the central frequency. If the width of the white noise spectrum is greater than two octaves, then two pitches equal to the boundary frequencies are heard. If the noise passes through the LPF or HPF, the audible pitch corresponds to the filter cutoff frequency [16]. The algorithm of the human hearing system when hearing sound through the CI is the same as in normal hearing. From the general group of auditory neurons with increased pulse density (see Figure 6) the auditory system allocates 3 separate groups of 146 neurons located symmetrically relative to the envelope points with the designations Z18, Z13, Z10. A possible algorithm for isolating 146 central neurons with increased pulse density is shown in [8].

4. Features of Perception of Syllables in the Standard Mode of Stimulation of Electrodes

The results of an experiment to assess the quality of perception of syllables in the V–CV format (vowel-consonant/vowel) by CI users with 22 active electrodes are known [17]. The experiment involved 30 subjects (adults with experience in speech communication).

Each subject listened to 19 syllables; on average, each syllable has been listened to 28 times, in random order. In total, each syllable has been listened to 870 times by all the subjects. Table 3 shows the number of correctly heard consonants for several syllables, and the most common variants of incorrectly heard consonants, and the number of consonants that the subjects could not identify or did not hear at all. To fully understand speech, all vowels and consonants must be heard as different in sound. In this case, the CI user remembers the new sound of previously familiar words as foreign words.

Table 3. *Perception of syllables in the format (V-CV) [17].*

№	Syllable	Number of correct responses	Incorrect answer options
1	A-PA	530	148 (T); 106 (K)
2	A-TA	545	199 (K); 85 (P)
3	A-KA	571	167 (T); 97 (P)
4	A-BA	490	105 (D); 60 (G)
5	A-DA	389	360 (G)
6	A-GA	639	107 (D)
7	A-FA	417	262 (S)
8	A-SA	438	151 (F)
9	A-VA	288	116 (Z)
10	A-MA	532	23 (B)
11	A-NA	19	67 (M); 21 (D)
12	A-LA	–	97 (M)

Note: The variant of the heard consonant sound is indicated in brackets.

To assess the ability of a CI user to hear syllables as different in sound, it is necessary to know the frequency spectrum of each syllable. Figure 7 shows for example the dynamic spectra of syllables (a-TA) and (e-TE). The dynamic syllable spectrogram shows that information about the consonant (T) sound is contained in a fragment of fewer than 0,2 seconds before pronouncing the 2nd vowel sound. The frequency of the formant For. 2 increases to about 4 kHz in the syllable (a-TA) and about 6 kHz in the syllable (e-TE). As can be seen in Table 3, the most difficult syllables for the subjects were (a-NA) and (a-LA). In these syllables, in a short fragment before pronouncing the 2nd vowel, the frequency of the For. 2 formant not! changes, but changes the loudness of For. 2 [15].

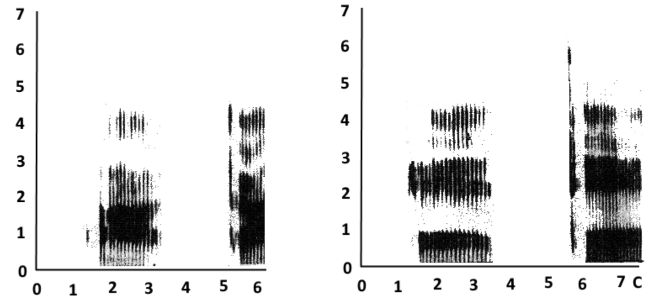

Figure 6. *Dynamic spectrograms of the consonant (T) sound in symmetrical syllables (a-TA) and (e-TE) [14].*

Figure 8 shows the dynamic spectrogram of the syllable (a-NA). The darker color of formant For. 2 after a pause in the noise band in the range from 1500 Hz to 2000 Hz, indicates that the frequency of the formant For. 2 does not change in the sound spectrum (N), but the loudness of formant For. 2 does. Since the dynamic range of changes in audible loudness when hearing sounds through the CI can not be more than 30 dB, the CI user does not hear a loudness widening by leaps when pronouncing a consonant sound (N). A similar pattern is observed in the perception of the syllable (a-LA). For this reason, users of CI almost do not hear consonants in the syllables (a-NA) and (a-LA). The quality of speech perception by CI users of the same type, depending on individual characteristics, can be different: from 20% to 85% of the correct perception of speech stimuli [17].

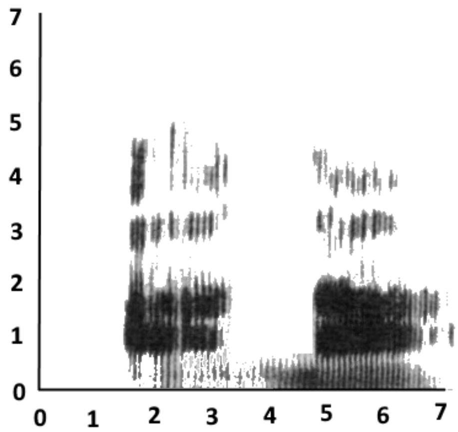

Figure 7. *Dynamic syllable spectrogram (a-NA) [14].*

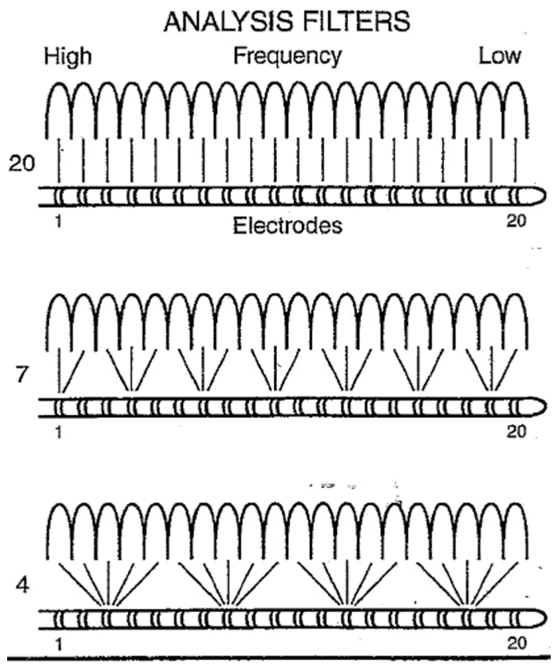

Figure 8. *Method for reducing the number of stimulated electrodes. Single BF output is connected to one electrode out of 20, three BF outputs are connected to one electrode out of 7, and five BF outputs are connected to one electrode out of 4 [1].*

5. Dependence of the Quality of Speech Perception from the Mode Stimulation of the Electrodes

The results of several experiments on the dependence of the quality of speech perception on the number of active electrodes in CI are known [1-4]. CI of various types differ in two main characteristics: the number of active electrodes and the type of strategy for encoding speech signals (SESS). In [1, 4] there is a detailed description of an experiment to test the quality of speech perception by CI users when the number of stimulated electrodes is reduced from 20 to 4. The Kim and Fishman experiment [1] involved 11 users of the Nucleus 22 (SPEAK processing strategy) CI. The quality of speech perception was determined by stimulating 20 electrodes (one BF output was connected to each active electrode), 7 electrodes (three BF outputs were connected to each active electrode), and 4 electrodes (five BF outputs were connected to each active electrode). The quality of speech perception significantly improved with an increase in the number of active electrodes from 4 to 7, and practically did not change with a further increase in the number of electrodes from 10 to 20.

Table 4. *Percentage of correct recognition of speech stimuli with diffrent number of active electrodes [1].*

Number of electrodes	1	2	4	7	10	20
Consonant	30%	50%	65%	70%	70%	70%
Vowel	30%	45%	55%	75%	70%	80%
Monosyllabic words	3%	15%	25%	40%	42%	50%
Sentences	15%	38%	42%	65%	67%	70%

Different speech signal encoding strategies can be used to convert the audio signal from the microphone output into a pulse pattern for stimulating the electrodes. Experiments were conducted in which CI users compared the quality of speech perception when enabling different strategies. When changing the strategy, there were no significant differences in the quality of speech perception by one CI user.

Table 5. *Comparison of the recognition quality of test speech signals when using different speech encoding strategies by the same CI user [5].*

Test signal	Words	Vowel	Consonant	Phonemes	Sentences
ACE strategy	32%	57%	57%	57%	72%
CIS strategy	30%	55%	56%	56%	65%
MDE3 strategy	30%	62%	56%	58%	66%
MEM strategy	32%	58%	58%	57%	72%

Table 5 shows the average results of comparing the quality of speech perception in an experiment involving 5 subjects, where different speech signal encoding strategies (speech coding strategy/SCS) were included in a CI with 22 electrodes for each user sequentially with an interval of 7 days: ACE, CIS, MDE3, and MEM. Taking into account the fact that the spread of individual indicators of speech perception quality in a group of CI users of the same type and with the same speech signal encoding strategy can be in the range of 40% to 70% of correct responses [5], the deviation of the average results when comparing different strategies in the range of 2% to 7% can not be the basis for ranking the SCS by the degree of their preferential use. The accuracy of word recognition, vowel and consonant sounds, phonemes, and sentences was virtually unchanged with the change of encoding strategy.

6. Dependence of the Quality of Speech Perception on the Number of Electrodes in the CI

Cells of the spiral ganglion, which are located along with the electrode array, generate pulses with exactly the same parameters when the electrodes are stimulated, as in the natural perception of an audio signal. When a single electrode is stimulated, stimulating pulses propagate to the SGC synapses, which are located up to 3 mm away from the stimulated electrode [7]. In a CI with 22 electrodes, the length of the working part of the electrode array is 19 mm (the distance between neighboring electrodes is approximately 0.9 mm). Figure 6 shows the envelope of a group of auditory neurons on the 1st PCFS during simultaneous stimulation of the electrodes (E19+E18+E17) and (E14+E13+E12+E11+E10+E9) when the (I) sound is perceived with the main frequency of 740 Hz. All SGC located along with the electrode array at a distance of 13 mm to 16 mm and 19 mm to 24 mm from the apex of the cochlea has an increased pulse intensity.

In a CI with 10 active electrodes, as shown in Figure 9, the electrodes (E19, E17, E14, E11, E8) are simultaneously stimulated. The distances between the stimulated electrodes are approximately 2 mm, 3 mm, 3 mm, and 3 mm. This means that practically stimulating pulses will affect all SGC located along a segment of the electrode array of approximately the same length as in the perception of (I) sound through the CI (22). When speech is perceived through CI (4) with 4 active electrodes, the distance between the simultaneously stimulated E18 and E13 electrodes is approximately 5 mm. In this case, there is no group of SGC, the excitation of which is simultaneously affected by the stimulation of two different electrodes. The CI user (4) can hear 4 different pitch sounds. However, in CI (4), due to the decrease in the number of bandpass filters, information about the change in the frequency of the audible signal is lost to a much greater extent than in CI with 22 and 10 electrodes.

***Table 6.** Stimulation of electrodes in the perception of (I) sound through CI with different numbers of electrodes.*

Electrode	E19	E18	E17	E14	E13	E12	E11	E10	E9
BF (22), Hz	600–760	760–920	920–1080	1414–1624	1624–1866	1866–2144	2144–2463	2463–2856	2856–3347
E aud.		+			+			+	
BF (10), Hz	E19		E17	E14			E11		E8
	440–760		760–1240	1240–1870			1870–2860		2860–4595
E aud.			+	+			+		
BF (4), Hz		E18			E13		E8		E3
		600–1414			1414–2856		2856–6308		
		+			+				

Note: The (+) sign indicates the audible electrodes "E aud."

A significant difference between auditory sensations when a signal is perceived through CI (4) and sensations through CI (22) and CI (10) is that in CI (4), when two, three or four electrodes are simultaneously stimulated, isolated groups of SGC are excited and each group of excited auditory neurons on the 1st PCFS will have one inflection point. Table 6 shows the frequency ranges of BF for the electrodes that are stimulated by the perception of (I) sound with the main frequency of 740 Hz. The sign (+) indicates the electrodes that are stimulated by pulses with the highest amplitude. In CI (22), when three adjacent electrodes (E19–17) are simultaneously stimulated, all SGC located in parallel to a segment of the electrode array approximately 6 mm long are excited. When six adjacent electrodes (E14–E9) are simultaneously stimulated, all SGCs located in parallel to a segment of the electrode array approximately 10 mm long are stimulated.

The conversion operation of the output signal BF in the amplitude of stimulating pulses to the respective electrode does not take into account the frequency of the sound signals at the input of BF, but only the magnitude of the rectified voltage at the output of the amplitude detector. Information about the frequency of the audio signal at the BF input is lost almost completely. Information about the frequency of the audio signal at the BF input is converted to the excitation of a group of SGC located near the stimulated electrode. In other words, the frequency information is converted at the 1st PCFS to the linear coordinate of the summing auditory neuron (SAN) located in the middle of the overall SAN group with the highest pulse density. In a CI with any number of electrodes, a significant part of the information about the frequency of the audio signal is lost at the output of the amplitude detectors. For this reason, the CI user hears different frequency signals at the BF input as the same pitch at the BF output.

The frequency spectrum of vowel sounds has three main formants and a significant noise component. The presence of a noise component is the reason for the fact that several neighboring electrodes are simultaneously stimulated. When simultaneously stimulating pairs of neighboring electrodes and single electrodes, the envelope of the general group of all auditory neurons on the 1st PCFS in the CCN has a greater number of inflection points. The CI user hears a greater number of different sounds at the same time.

When perceiving syllables in the V–CV format, for most syllables, the main distinguishing feature of a consonant is a change in the frequency of the 2nd formant in a segment lasting up to 0,2 seconds before the 2nd vowel (Figure 7). For syllables (a-LA), (a-MA), and (a-NA), the main feature for distinguishing a consonantal sound is the "leap" in intensity when switching from a consonantal sound to a vowel [15]. For syllables (a-LA) and (a-NA), the leap in intensity before the 2nd vowel is less than 10 dB. For a syllable (a-MA), the leap in intensity is approximately 22 dB. According to the results of the experiment shown in Table 3, no one could correctly identify the "L" sound. The "N" sound was only correctly identified 19 times. The "M" sound was correctly identified 532 times. The CI user does not hear an abrupt change in the intensity of the 1st formant by less than 10 dB.

7. Virtual Frequency Channel

The CI has a reserve that is not yet used to improve the quality of speech perception.

When two adjacent electrodes E(n) and E(n+1) are simultaneously stimulated, a virtual frequency channel (VFC) is formed. The CI user hears an audio signal with an intermediate pitch [Z(n) +Z(n+1)]/2 [12, 13]. At the same time, the number of different sounds that the user of the CI (22) can hear, increases from 22 to 43. When an audio signal is perceived through a CI with 16 electrodes, the number of sounds of different pitches increases to 31. A user of CI with 12 electrodes can hear 23 different pitch sounds. Electrodes that are located to the left and right of the pair of electrodes E(n) and E(n+1) should not be stimulated. The ability of the CI user to hear a sound with an intermediate pitch while simultaneously stimulating two adjacent electrodes was tested experimentally. The experiment involved 8 subjects. The average percentage of correct estimation of the change in audible pitch was 80% when the apical and medial electrodes were stimulated with signals lasting from 300 ms to 1000 ms. When stimulating pairs of basal electrodes, the average percentage of correct evaluation of the audible pitch change was 70%.

The frequency spectrum of (I) sound has a significant noise component (see Figure 5).

Several neighboring electrodes are stimulated simultaneously (see Fiure 4) due to the presence of a large noise component. There are no isolated pairs of electrodes for the possibility of implementing a virtual FC. The CI user can only hear the sound of the electrodes with the highest amplitude of stimulating pulses.

Figure 9 shows the proposed electrode stimulation mode, where there are favorable conditions for the implementation of VFC. Table 8 shows audible electrodes for (I) sound perception with a main frequency of 98 Hz and 740 Hz for the standard and proposed stimulation modes. The number of audible electrodes depends on the ratio between the frequency spectrum of the vowel sound and the distribution of the frequency range of the general spectrum at the input of the speech processor between 22 bandpass filters (BF). The ability to hear a larger number of simultaneously stimulated electrodes allows the CI user to hear different speech sounds as different auditory sensations. When implementing the possibility of creating a VFC, the user of the CI (22) will hear a larger number of simultaneously stimulated electrodes.

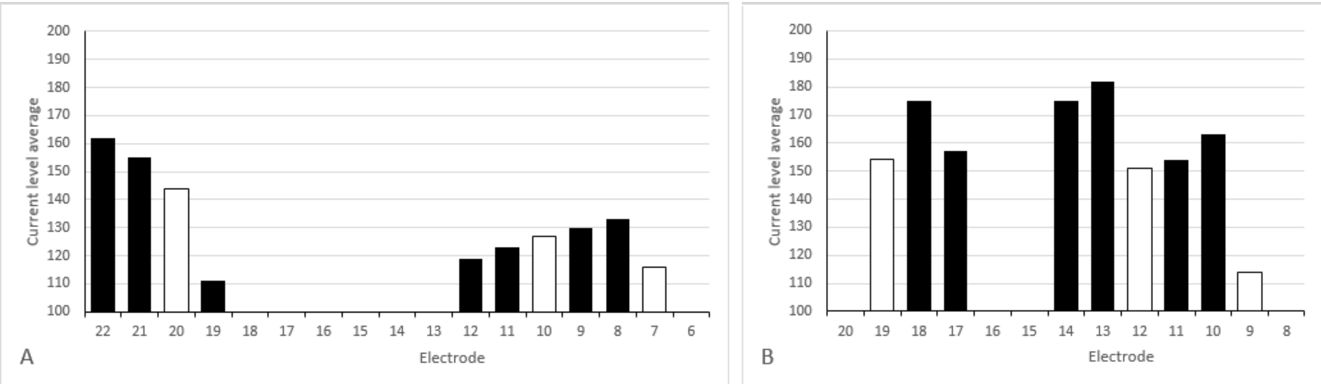

Figure 9. *A possible way to correct the mode of stimulation of the electrodes for the perception of (I) sound with a different frequency of the main tone.*

Table 7. *Audible electrodes in the standard and proposed modes of stimulation in the perception of (I) sound.*

Stimulation mode	Figure 9A	Figure 9B	Audible electrodes
Standard	E22, E21, E8, E7	E18, E14, E13	4 (Figure 9A), 3 (Figure 9B)
Proposed	E22, E21, E19, E12, E11, E9, E8	E18, E17, E14, E13, E11, E10	7 (Figure 9A), 6 (Figure 9B)

The main reason why the improvement of the speech perception quality by increasing the number of electrodes in standard stimulation mode is the impossible reduction of the distance between the electrodes. If 7 electrodes out of 22 are stimulated, the distance between neighboring electrodes is approximately 3 mm. When 8 or more electrodes out of 22 are stimulated, the distance between adjacent electrodes decreases from 2,4 mm to 1,0 mm. Each SGC group is simultaneously affected by the stimulation of 3 to 5 adjacent electrodes. The CI user hears the sound of one electrode with the largest amplitude of stimulating pulses from each group of several simultaneously stimulated neighboring electrodes. At a distance of less than 2 mm between adjacent electrodes, the total number of SGC that generate pulses and transmit them along the auditory nerve fibers in the CCN to the 1st projection of the frequency scale of the cochlea practically does not change. The shape of the envelope group of auditory neurons on the 1st PCFS does not change. Under these conditions, there are also no grounds for improving the quality of speech perception. The CI user will be able to hear more sounds of different pitches if conditions are created for the formation of virtual frequency channels.

8. A possible Way to Optimize the Electrode Stimulation Mode

For practical implementation of the proposed method of improving the quality of speech perception, it is enough to add one logical operation to the algorithm for controlling the stimulation of electrodes: the exclusion of simultaneous stimulation of three neighboring electrodes. Experimental verification of the proposed method of creating conditions for improving the quality of speech perception was not carried out since the author was not able to change the program for controlling the stimulation of electrodes. However, given the fact that from 1997 to 2019 [1, 2, 3, 4] in the scientific literature there were no assumptions about the impossibility of improving speech understanding by increasing the number of electrodes in a CI, it would be useful to test the proposed method.

When an audio signal is perceived in a speech processor, the following operations are performed sequentially: fixing the voltage value at the outputs of amplitude detectors (see Figure 1), regulating the amplitude of stimulating pulses for each active electrode, correcting the ratio between the maximum and minimum permissible values of the pulse amplitude in the range from the audibility threshold to the pain threshold, and stimulating the corresponding electrode. To optimize the electrode stimulation mode, it is enough to add one more logical operation: excluding the stimulation of the 3rd neighboring electrode for the possibility of organizing a VFC. For the stimulation option according to Figure 4A, the electrodes E20, E10, and E7 are turned off; according to Figure 4B, the electrodes E19, E12, and E9 are turned off. The quality of speech perception should be improved by optimizing the conditions for creating "virtual frequency channels". In the pattern of stimulating pulses, the electrodes that should be stimulated by the pulses with the highest amplitude are selected. For example, in the mode of stimulation of the electrodes in Figure 4B, this is (E18 + E13). Then the amplitude of stimulating pulses for neighboring electrodes to the right and left of E18 and E13 is estimated. The neighboring electrode that is stimulated by pulses with a higher amplitude is selected for stimulation (E18/E17 + E13/14) according to Figure 4B.

If amplitudes of stimulating pulses for electrodes on the right and left are the same, then the electrode located closer to the apex of the cochlea should be stimulated. From the group of remaining electrodes, the electrode with the largest amplitude of stimulating pulses is selected and the 2nd pair of active electrodes (E10/E11) is formed according to Figure 4B. As a result, three virtual FC function. The CI user will hear three different pitch sounds both in the standard electrode stimulation mode and in the mode of stimulation of pairs of neighboring electrodes. However, when the frequency spectrum of the speech signal changes, the ratio between pulse amplitudes for stimulating neighboring electrodes change. When two adjacent electrodes are simultaneously stimulated, the E18/E19 or E17/E16 pairs can be stimulated instead of the E18/E17 pair. The CI user will hear a sound with a different pitch. When six electrodes are simultaneously stimulated, the CI user will hear three sounds with pitches of Z18a, Z13a, and Z10a. Exception of 3rd neighboring electrode leads to following: Z18a < Z18, Z13a > Z13, and Z10a > Z10. Depending on the ratio of amplitudes of the stimulating pulses for E19, E18, or E17 electrodes, the CI user can hear three different pitch sounds when stimulating E19/E18, E18/E17, or E18 electrodes. When three adjacent electrodes are simultaneously stimulated, the user is most likely to hear a sound with a single Z18 pitch.

9. Discussion

9.1. Features of Auditory Sensations in the Perception of a Tonal Sound Signal in Normal Hearing and When Stimulating Single Electrodes

In normal hearing, when receiving a tonal audio signal with an intensity of up to 60 dB, the receptor potential increases in a group of about 500 inner hair cells (IHC). On the 1st projection of the FCS, 3520 summing auditory neurons (SAN) with an increased pulse density are linearly located. The auditory system remembers the ordinal number of the SAN, which is located in the middle of the general group with an increased pulse intensity. The auditory system does not know why the pulse intensity increases in a group of ordered auditory neurons on the 1st PCFS. The algorithm of functioning of the auditory system does not change depending on the cause of SGC excitation. When an audio signal is perceived through the CI and a single electrode is stimulated by signal pulses with relatively high amplitude in the mode of the most comfortable level of perception, the pulse intensity increases in the group of SGC, which are located along with the BM at a distance of up to 3 mm from the electrode. From this group, a noticeable increase in the pulse intensity will be observed in the group of SGC, located along the BM segment with a length of approximately 6 mm. In normal hearing, 660 IHC and approximately 3000 spiral ganglion cells are located along a 6 mm-long segment of BM. At the 1st PCFS in the CCN, one group of summing auditory neurons will have an increased pulse intensity. From this group, the auditory system allocates 146 auditory neurons located symmetrically relative to the central SAN. The audible pitch will depend not on the frequency of the audio signal at the BF input, but, in fact, on the distance of the stimulated electrode from the apex of the cochlea.

9.2. Features of Speech Perception Through a Cochlear Implant (CI)

Figure 7 and Figure 8 show the dynamic spectrograms of consonantal sound (T) in a syllable (a-TA) and (N) in a syllable (a-NA) [13]. The main feature of the frequency spectrum of syllables is the presence of 3 formants with a significant noise component. As a result, even when a single vowel (I) is perceived, several electrodes are simultaneously stimulated (see Figure 5). CIs of different manufacturers have 12, 16, 20, or 22 electrodes. The distance between adjacent electrodes in all types of CI is less than 3 mm. This means that when three or more adjacent electrodes are simultaneously stimulated, the stimulating pulses excite almost all the preserved SGC, which are located along the corresponding segment of the BM plus 2 mm to the right and left of the extreme electrodes.

If only 10 active electrodes are stimulated from the total group of 22 electrodes (as shown in Table 6), then when the (I) sound is perceived, the electrodes E19 and E17 (1st formant) and E14, E11, E8 (2nd and 3rd formants) are simultaneously stimulated. The distance between the) stimulated electrodes is less than 3 mm. For this reason, the total number of preserved SGC located along the BM segment from the E19 electrode to the E8 electrode will be exactly the same as when 22 electrodes are stimulated. The pulse intensity of auditory neurons on the 1st projection of the FCS in the CCN will be the same. The auditory sensations of the CI user are the same for both 22 active electrodes and 10 active electrodes.

The quality of speech perception by (the) CI user depends on what percentage of words and syllables the CI user hears as different in sound. The main difference between the auditory sensations of a CI user is that after cochlear implantation, many words are heard differently than they were before hearing loss. Auditory sensations depend on which groups of auditory neurons on the 1st projection of the FCS have an increased pulse intensity when each word is perceived. If a CI user hears all the words as different in sound, then they remember the new sound like foreign words, and the quality of speech perception can be very high. There may be a situation when the user hears some different syllables and words as identical in sound. In this case, the quality of speech perception is significantly reduced. Table 8 shows the number of audible electrodes in the standard mode of stimulation of electrodes and if it is possible to implement the conditions for the formation of VFC in the perception of (I) sound. Increasing the number of audible electrodes from 4 to 7 (Figure 4A) and from 3 to 6 (Figure 4B) ensures that the CI user can hear a greater number of different auditory sensations of syllables and words.

9.3. The main Factors That Negatively Affect the Quality of Speech Perception by CI Users

The quality of speech perception in the standard mode of stimulation of electrodes is negatively affected by two factors: the noise component in the spectrum of the audio signal and the small distance between neighboring electrodes. Due to the presence of a significant noise component in the speech signal spectrum, not only electrodes that transmit information about the frequency and intensity of the main formants are stimulated but also electrodes located nearby that contain information about the intensity of the noise component. As a result, out of three adjacent electrodes, the CI user hears the sound of only one electrode with the highest amplitude of stimulating pulses. However, this increases the pulse intensity in all preserved SG cells that are located along the BM segment with a length of at least 6 mm. The auditory system perceives information about a significant increase in the total pulse intensity in a large group of SGC as an avalanche increase in the intensity of the audio signal and adjusts the dynamic range of changes in volume.

Figure 5 shows the frequency spectrum of the vowel (I) sound [15]. Average formant frequency values: F1=300 Hz, F2=2650 Hz, F3=4000 Hz.

Table 8. *Frequency ranges of BF for CI (22) [7].*

Electrode	E20	E19	E18	E8	E7	E6	E5
F, Hz	120–280	280–440	440–600	2460–2860	2860–3350	3350–3900	3900–4600

If it is possible to turn off the electrodes, the stimulation of which depends only on the noise component, the CI user (22) hears the sound of only three electrodes: E19 (280-440) Hz, E8 (2460-2860) Hz, and E5 (3900-4600) Hz. The distance between the electrodes E19 and E8 is 11 mm, between the electrodes E8 and E5 there are 3 mm. At the 1st projection of the CFS in the CCN, 3 groups of auditory neurons will have an increased pulse intensity. The CI user will hear 3 sounds at the same time with different pitches.

10. Conclusion

During the process of converting information about the frequency and intensity of the audio signal into a pattern of pulses for stimulating the electrodes in a cochlear implant, some part of the information is lost irrevocably. The main condition for understanding speech is to hear syllables and words as different in sound. The more different pitch sounds a CI user can hear, the better they understand speech.

A comprehensive analysis of well-known experiment results to assess the quality of speech perception by different types of CI users shows that the accuracy of speech stimulus recognition improves only in one case: when increasing the number of stimulated electrodes from 4 to 8. The CI user (22) can theoretically hear 43 different sound pitches, while the CI user (12) can hear 23 different by pitch sounds. But this is only possible if the stimulation of the electrodes has the conditions for the formation of a virtual frequency channel.

If only pairs of neighboring electrodes and isolated single electrodes are stimulated in the CI, the CI user hears more sounds of different pitches, which should help to improve the quality of speech perception. It is recommended to add an operation to the program for converting information about the frequency and intensity of the audio signal to exclude the stimulation of electrodes located between pairs of stimulated electrodes. When isolated pairs of adjacent electrodes and isolated electrodes are stimulated, the advantages of CI (22) over CI (12) will be implemented.

References

[1] Kim E., Fishman U. A. Speech Recognition as a Function of the Number of Electrodes Used in the Speak Cochlear Implant Speech Processor // JSLHR. 1997. Vol. 40. № 5. P. 1201-1215.

[2] B. Laback et al. Coding of vowellike signals in cochlear implant listeners // JASA. 2004. Vol. 116. № 2. P. 1208-1223.

[3] Swanson B. A. Pitch Perception with Cochlear Implant. Melbourn. 2008. 290 p.

[4] Berg K. A., Noble J. H., Dawant B. M., Labadie R. F., Gifford R. H. “Speech recognition as a function of the number of channels in perimodiolar electrode recipients”. JASA 2019 Vol. 145 (1), 1556 p.

[5] Vandali A. E. et al. Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies // JASA. 2005. Vol. 117. № 5. P. 3126–3138.

[6] Cochlear Implant Simulation version 2. 0. University of Granada. Granada, December 2004.

[7] McKay C. M., Henshall K. R. Frequency-to-electrode allocation and speech perception with cochlear implants // JASA. 2002. Vol. 111. № 2. P. 1036-1044.

[8] Stefanovich M. (2020). Possible Method of memorizing tonal sound signal frequency in the human auditory system. Advances in applied Physiology, 2019: 19-27.

[9] M. Stefanovich, V. Pudov (2013). Peculiarities of auditory sensations in electrode hearing aids. Neurophysiological mechanisms of the formation of auditory sensations in the perception of a sound signal. Saarbrucken: LAMBERT. 120 p.

[10] Zwicker E., Feldkeller R. (1967). Das Ohr als Nachrichten Empfanger. Stuttgart: Hierzel-Verlag. 250 p.

[11] M. Mauermann (2004). Fine structure of hearing threshold and loudness perception. JASA, 116: 1066-1088.

[12] Patent No: US 7,251,530 B1 31.07.2007. "Optimizing pitch and other Speech stimuli allocation in a Cochlear implant".

[13] Luo Xin. Encoding pitch contours using current steering // JASA. 2010. Vol. 128. № 3. P. 1215-1223.

[14] Looi V. The effect of cochlear Implantation on music perception. A review, 2008. (Tsang D. J. Musical pitch perception with cochlear implants: a comparison of strategies. // Honours dissertation. Melbourne).

[15] Derkach M. F. "Dynamic spectra of speech signals" Lyvov, 1983, 168 p.

[16] Fastl, H., Zwicker, E. "Psychoacoustics: Facts and Models" 2007, 471p.

[17] Munson B., Donaldson G. S. and Allen Shanna L., Collison Elizabeth A., Nelson David A. Patterns of phoneme perception errors by listeners with cochlear implants as a of overall speech perception ability // JASA. 2003. Vol. 113. № 2. P. 925–935.

Biography

Margarita Stefanovich, I have worked for 34 years as an engineer specializing in radio measuring instruments. I have several patents for inventions in the field of measuring technology. Now I am a pensionary and an independent researcher. Since 2000 I do my own research on hearing physiology and neurophysiology. Between 2012 and 2018, 9 of my articles and one book were published: M. Stefanovich, V. Pudov "Peculiarities of auditory sensations in electrode hearing aids. Neurophysiological mechanisms of the formation of auditory sensations in the perception of a sound signal (2013)". Of particular interest to me is Cochlear Implantation.