

Perception of the Speech Signal Against the Background of Noise in Normal, Hard of Hearing and Implanted Patients



Petrov SM*

PE, Saint Petersburg, Russia

Submission: February 11, 2026; **Published:** February 20, 2026

***Corresponding author:** Petrov SM, PE, Saint Petersburg, Russia

Abstract

One of the problems for most people with hearing loss is the difficulty of perceiving a speech signal in the presence of acoustic interference: extraneous speech, reverberation, music, various noises, i.e. they have a deterioration in the noise immunity of the auditory system. This article examines the issue of speech perception with simultaneous application of a masking signal – white noise - in normal subjects, patients with sensorineural hearing loss and implanted patients. Speech intelligibility against noise in normal and hearing-impaired patients can largely be explained by the high spectral redundancy of the speech signal. Taking into account the effect of noise on the result of audio signal processing in implanted patients, their noise immunity should be increased before electrical pulses are created, i.e. in the speech signal with a background of noise.

Keywords: Noise immunity; Signal-to-noise ratio; Peak factor; Sensorineural hearing loss; Cochlear implantation; Speech intelligibility; Spectral redundancy

Abbreviations: CI: Cochlear Implantation; SNHL: Sensorineural Hearing Loss; BM: Basilar Membrane; WN: White Noise

Introduction

One of the problems for most people with hearing loss is the difficulty of perceiving a speech signal in the presence (against the background) of acoustic interference: extraneous speech, reverberation, music, various noises, i.e. they have a deterioration in the noise immunity of the auditory system. A patient after cochlear implantation (CI) has auditory perception features, so the problem of speech perception against noise is more acute for them than in patients with hearing loss. This article examines the issue of speech perception while simultaneously applying a masking signal – white noise - in normal subjects, patients with sensorineural hearing loss (SNHL) and implanted patients. We have already touched upon this issue [1]. The Signal-to-Noise ratio (S/N) is used to characterize the difference in the levels of speech intensity and noise. If the S/N value is positive, then the useful signal is higher in intensity than the noise level. In accordance with the standard of the Russian Federation [2], the level of speech intelligibility of normal subjects is considered satisfactory if the sound pressure level of the speech signal exceeds the noise

level by 10 dB, i.e. the signal-to-noise ratio (S/N) is equal to or more than 10 db.

When considering the issue of speech perception when exposed to white noise (WN) at the same time, it is necessary to take into account the acoustic characteristics of these two signals- noise and speech. White noise has a wide spectrum, uniform spectral characteristics, and its spectrum is homogeneous in amplitude and random in phase. The spectral density of the WN is constant. The speech signal also has a wide spectrum, but the intensity level is variable over time, since speech has pauses between words, spectral maxima-peaks (for example, corresponding to formants), as well as minima, different intensity of phonemes [3]. The dynamic range of speech, that is, the ratio of the maximum and minimum power of the speech signal, ranges from 35 to 45 dB [4]. The average level of speech, that is, the value of the intensity level used in acoustic research, is measured within 15 seconds. Naturally, when measuring the average level of speech intensity, the contribution of all these irregularities-maxima and minima-to the measured value is evenly distributed.

To characterize deviations from the average level, that is, values exceeding the average level of speech, the peak factor is used, that is, the ratio of the instantaneous amplitude to the RMS value. On average, it is 12 dB, but in 2% of cases the peak values exceed even 20 dB [4]. Naturally, with constant exposure to masking noise, the spectral components of the speech signal are less intense than the level of the masker, they become noisy, and speech perception against the background of noise is determined by the remaining unmasked spectral regions of the speech signal, which exceed the noise level in intensity, i.e. the peak factor values of the speech signal. Next, the peak factors. To further discuss the issue of noise immunity, we will schematically present a picture of basilar membrane (BM) fluctuations in response to the peak factor of a speech signal in some frequency band in normal subjects (1a) and in patients with SNHL (1b). To illustrate the spectral redundancy of speech, the spectrum of the speech signal after comb filtering (1c) is schematically shown.

The abscissa axis is the frequency, and the ordinate axis is the amplitude. When considering the perception of a speech signal due to its peak factor values exceeding the noise level (Figure 1), it is interesting to look at the perception of a speech signal after comb filtering (CF) in normal subjects [5]. In an earlier study, a speech signal in the range of 200-8250 Hz was used. After processing it with a comb filter only one 50 Hz wide band was left in each subsequent 1000 Hz band,. Figure 1c schematically shows the spectrum of speech material with such parameters of comb filtering. We found that 5% (!) of the spectrum is enough for 100% understanding of words. As you can see, the spectral redundancy of the speech signal is large and it is thanks to it that speech can be understood due to peak factors, i.e. spectral bands exceeding the noise level. It should be noted that, unlike a masker that makes noise in low-intensity parts of the speech spectrum, with comb filtering, 95% of speech information is not noisy, but simply deleted. In our opinion, CF is a vivid demonstration of the spectral redundancy of a speech signal, which makes it easier to understand speech intelligibility against noise.

Sensorineural hearing loss

As is known, with the same S/N ratio as the norm, speech intelligibility in the presence of background noise worsens to a greater extent in patients with SNHL than in subjects with normal hearing [6], i.e., the noise immunity of the auditory system with SNHL, is lower than normal [7]. This is due to changes in the mechanisms of sound signal processing that occur due to the death of hair cells in patients with SNHL. The noise in the SNHL, of course, remains noise. Let's look at what happens to peak factors in patients with SNHL. As is known, after the death of the outer hair cells, the cochlear amplifier is destroyed [8], as a result of which the acute tuning of the auditory nerve fibers disappears. Figure 1 schematically shows fluctuations of basilar (BM) in response to the same peak factor in normal subjects (1a) and in patients with SNHL (1b). What do we see in patients with SNHL in

comparison with the norm? A decrease in the acuity of the tuning leads to a decrease in the amplitude of the vibrations of the BM and, consequently, in comparison with the norm, to a decrease in the S/N ratio during the transmission of nerve impulses to the central departments [9].

Due to the absence of a cochlear amplifier, the peak factor band is expanding and new spectral components appear that are not normally present. The appearance of new frequency components was previously noted by patients with unilateral SNHL in our study [10]. After setting the pitch equality of the tonal stimuli applied to both ears, they themselves indicate a difference in the timbres of equal-pitch signals in different ears. They describe the tone on the affected side as muffled, like going through cotton wool, through a bad phone, etc. In the same work, we discovered a shift in tone perception on the affected side towards high frequencies [10]. For example, when applying a tone with a frequency of 500 Hz to an ear with hearing loss, the patient perceived a pitch-equal tone signal with a frequency of 610 Hz with a normal ear. The temporal theory of frequency perception does not work here. A similar shift was found at all the frequencies examined. This result was discussed in more detail earlier [10]. The combined effect of the described factors-frequency selectivity, a decrease in the S/N ratio, a frequency shift-and not only them-reduces the redundancy of speech that is normally present, which complicates the understanding of speech by patients with SNHL under normal conditions, and even more so against the background of noise.

Cochlear implantation

Speech perception against a background of noise in normal subjects and in patients with SNHL occurs mainly due to the analysis of the amplitude-frequency characteristics of peak factors above the noise level, but as could be seen, to varying degrees in these two groups of subjects. To evaluate the perception of speech against noise by implanted patients, it is necessary to consider the processing of the frequency range of the input acoustic signal in a speech processor in silence and against noise. After dividing the speech spectrum of CI frequency range into certain frequency bands, energy is measured in each of them at a given time and, depending on its magnitude, an electric pulse of the appropriate amplitude is generated in each channel. A moment is a time interval during which the sound is analyzed and a series of pulses is created. For example, at a pulse rate of 1000 pulses/s, the moment is 1 ms. What happens when masking noise is applied at the same time? The white noise energy in the band of each channel added to the energy of the speech signal in the same band, and the amplitude of the pulse at the electrode is determined by the total energy of noise and speech. The energy of each band, due to the presence of noise, is greater than in silence, and therefore, with noise, the amplitude of the pulse will increase compared to its amplitude in the same channel in silence. There are no spectral changes, only the signals in all channels are getting louder. Since the intensity of the speech signal changes over time,

the energy in each channel also changes and the increment of the pulse amplitude at successive moments of speech processing in noise will occur by a different amount. The magnitude of the pulse amplitude increment depends on the S/N ratio at a given moment in each channel. The greater this ratio, the smaller the amplitude increment. Such unpredictable and varying pulse amplitude increments occur in all channels and, consequently, the “normal” (in silence) speech pattern changes significantly.

Perception of a new implanted, significantly impoverished language [11] in silence presents certain difficulties for CI patients, and with constant and unpredictable changes in pulse amplitudes in all channels with noise, it will undoubtedly make it even more difficult to understand speech in noise compared to the perception of speech in silence. Apparently, it should be borne in mind that, given such a significantly transformed speech signal, the redundancy of speech perceived by such patients initially reduced and therefore the slightest destructive effect (in our case, white noise) negatively affects speech intelligibility. In addition to the uninformative increment of the pulse amplitude, there is another disadvantage from the presence of noise-in channels where there is currently no speech signal, the patient hears an uninformative (!) sound. This is our explanation of noise immunity in CI patients. It should be noted that the lower the S/N ratio the more the uninformative increment of the pulse amplitude, which further distorts the “normal” picture of speech in silence. This effect is confirmed in a model experiment [12].

With a signal-to-noise ratio of +2 dB, maximum sentence intelligibility (plateau) was achieved using 12 stimulation channels. With a signal-to-noise ratio of -2 dB, maximum intelligibility was achieved using 20 stimulation channels. It follows that with an increase in noise intensity, an increase in the number of channels is required to achieve a plateau of intelligibility [12]. It should be noted that with an increase in the number of channels, their bands narrow, therefore, the channel selectivity of stimulation increases, which improves intelligibility. And in the noise too [13]. It's possible to assume. that with a narrowing of the frequency range, noise immunity increases in CI patients. Obviously, noise immunity in CI patients' needs to be increased before creating a noise signal pulse, i.e. noise reduction should be carried out in the audio signal.

Conclusions

a) Speech intelligibility against noise in normal and hearing-impaired patients can largely be explained by the high spectral redundancy of the speech signal.

b) Noise immunity in CI patients should be increased in the input audio signal (noise reduction).

References

1. Petrov SM (1999) Redundancy of the speech signal and noise immunity of the auditory system. *News of otorhinolaryngology and speech pathology* 4: 20-22.
2. (2013) National standard of the Russian Federation. *Ergonomics. Evaluation of speech communication.*
3. Rulenkova LI, Smirnova OI (2003) *Audiology and hearing prosthetics.* Moscow: Publishing center Academy pp. 208.
4. Efimov AP (1989) *Acoustics. Handbook.* Moscow : Radio and Communications p. 336.
5. Petrov SM (2003) Comfortable loudness level and signal intelligibility of spectrally transformed speech. *Fiziol Cheloveka* 29(4): 45-48.
6. Millman RE, Sven L, Mattys SL, Gouws AD, Prendergast G (2017) Magnified Neural Envelope Coding Predicts Deficits in Speech Perception in Noise. *J Neurosci* 9(32): 7727-7736.
7. Salvi RJ, Henderson D, Hamernik R, Ahroon WA (1983) Neural correlates of sensorineural hearing loss. *Ear Hear* 4(3): 115-129.
8. Ricci A (2003) Active hair bundle movements and the cochlear amplifier. *J Am Acad Audiol* 14 (6): 325-338.
9. Leek MR, Summers V (1996) Reduced frequency selectivity and the preservation of spectral contrast in noise. *J Acoust Soc Am* 100(3): 1796-1806.
10. Petrov SM (1998) The perception of the frequency of tonal stimuli in a lesion of the sound-perceiving apparatus. *Fiziol Cheloveka* 24(6): 38-41.
11. Petrov SM (2025) Amplitude-Frequency-Temporal picture of a New Implanted Language. *Glob J Otolaryngol* 27(3): 1-3
12. Dorman MF, Loizou PC, Fitzke J, Tu Z (1998) The recognition of sentences in noise by normal-hearing listeners using simulations of cochlear-implant signal processors with 6-20 channels. *J Acoust Soc Am* 104(6): 3583-3585.
13. Petrov SM (2022) What frequency range of the cochlear implant is the best for speech perception? *Scholarly J Otolaryngology* 7(5): 811-813.



This work is licensed under Creative Commons Attribution 4.0 License
DOI: [10.19080/GJO.2026.28.556239](https://doi.org/10.19080/GJO.2026.28.556239)

**Your next submission with Juniper Publishers
will reach you the below assets**

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats
(Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission
<https://juniperpublishers.com/online-submission.php>