Speech quality evaluation of subcutaneously implanted microphone using *in vivo* experiment

Seong Tak Woo[a], Gihyoun Lee[b], Eui Sung Jung[a], Hyung-Gyu Lim[a], Ki Woong Seong[c], Jyung Hyun Lee[c], Myoung Nam Kim[c,d] and Jin-Ho Cho[e,*]

[a]*Graduate School of Electronics Engineering, Kyungpook National University, Sangyuk-dong, Buk-Gu, Daegu, 702-701, S. Korea*
[b]*Department of Medical and Biological Engineering, Graduate School, Kyungpook National University, Dongin2-Ga, Jung-Gu, Daegu, 700-422, S. Korea*
[c]*Department of Biomedical Engineering, Kyungpook National University Hospital, Dongin2-Ga, Jung-Gu, Daegu, 700-422, S. Korea*
[d]*Graduate School of Medicine, Kyungpook National University, Dongin2-Ga, Jung-Gu, Daegu, 700-422, S. Korea*
[e]*School of Electronics Engineering, College of IT Engineering, Kyungpook National University, Sangyuk-dong, Buk-Gu, Daegu, 702-701, S. Korea*

Abstract. The microphone in a fully implantable hearing device (FIHD) is generally implanted under the skin covering the temporal bone. However, the implanted microphone can be affected by the skin, which causes both sound attenuation and distortion, particularly at high frequencies. As the degree of attenuation and distortion through the skin is severe, speech quality evaluation parameters are needed for the received signal when designing an implantable microphone. However, the performance of most implantable microphones is only assessed based on the sensitivity and frequency response. Thus, practical indicators based on human auditory characteristics are needed for an objective evaluation of the performance of implantable microphones. In this study, a subcutaneously implantable microphone was designed, and its frequency response investigated using an *in vivo* experiment. Plus, to evaluate the objective indicators, the speech quality of the signals measured by the implanted microphone was calculated using a MATLAB program, and the indicators compared before and after implantation.

Keywords: Implantable microphone, *in vivo* experiment, speech quality score, fully implantable hearing devices

1. Introduction

Various types of fully-implantable hearing devices (FIHDs) have recently been developed to mitigate the auditory inconveniences experienced by hearing-impaired people. In particular, FIHDs have at-

*Corresponding author: Jin-Ho Cho, School of Electronics Engineering, College of IT Engineering, Kyungpook National University, Sangyuk-dong, Buk-Gu, Daegu, 702-701, S. Korea. Tel.: 82-53-427-5538; Fax: 82-53-427-5539; E-mail: jh-cho@ee.knu.ac.kr.
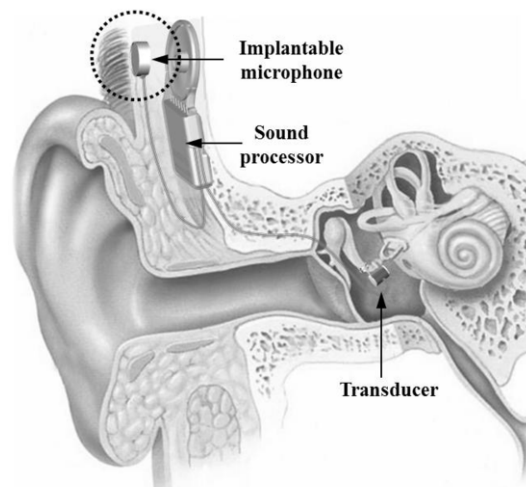
Fig. 1. Schematic of fully implantable hearing device.

tracted attention as they can overcome acoustic feedback and cosmetic problems [1–5]. Figure 1 shows a schematic of an FIHD. Generally, the FIHD microphone is implanted under the skin of the temporal bone [6]. However, the microphone characteristics vary after implantation, as the microphone is surrounded by skin or muscle tissue that can come into contact with the microphone membrane or casing. When the microphone is implanted under the skin, the acoustic signal collected through the implanted microphone is attenuated through scattering and absorption by the skin [7]. Furthermore, a subcutaneously implanted microphone can be significantly affected by mastication noise when the patient is chewing food. Thus far, the performance of most implantable microphones has only been assessed based on the sensitivity and frequency response. However, the sound attenuation with average 7mm skin is 10 dB to 20 dB, particularly in a high frequency region [8,9]. Plus, the sound degradation due to frequency and phase distortions is also severe. As a result, these deteriorated conditions significantly reduce the sound perception ability of an implanted microphone [10]. Therefore, various additional speech quality evaluation parameters with practical indicators based on human auditory characteristics need to be included in the design of implantable microphones.

Accordingly, this paper presents a design for a subcutaneously implanted microphone, and evaluates its performance using an animal experiment. To evaluate the speech quality of the implanted microphone, 7 objective indicators, including the perceptual evaluation of speech quality (PESQ) and total harmonic distortion (THD), are computed using MATLAB and each parameter value is compared before and after implantation.

2. *In vivo* experiment and speech quality indicators

2.1 Method

Microphones for FIHD must satisfy many requirements, such as biocompatibility, a high sensitivity, low distortion, and high signal-to-noise ratio. The casing and diaphragm of the implantable microphone used in this study were fabricated using biocompatible titanium. Figure 2 shows a photograph of the microphone implanted in a guinea pig and the frequency response. The electret con-

denser microphone (ECM) with a 4mm diameter was inserted into the titanium case. The overall dimensions of the fabricated implantable microphone were a 9 mm diameter, 3.8 mm height, and 0.02 mm diaphragm thickness. The connections of the fabricated microphone were sealed using a biocompatible epoxy (Medicure-222, Dymax).

The fabricated microphone was implanted at the left lumbus of the guinea pig, where the skin thickness was approximately 2.5 mm. To ensure sufficient contact between the skin and the diaphragm, the experiment was performed two weeks after the microphone was implanted. Moreover, for the animal experiment, the fabricated microphone was sterilized using an ethylene oxide gas.

The sound experiment was conducted in an acoustic chamber with a background noise of approximately 30 dB SPL. Speech at approximately 80 dB SPL and a pure tone at 94 dB SPL produced by a standard speaker (FX120, Fostex) were applied to the microphone implanted in the guinea pig. An audio-band measurement and analysis device (Fast Ultra Track 8R, M-Audio) were used to measure the output signal of the microphone. For the sensitivity calibration, a reference microphone (ER-10B, Etymotic) was used to measure the sound level of the speaker.

The mean sensitivity of the ECM was - 30 dB in a free field (re 0 dB = 1V/Pa). As shown by the solid lines in Figure 2, the sensitivity of the fabricated implantable microphone was measured in the flat region from 0.1 to 3 kHz at about – 37 dB, while resonance occurred at about 6 to 8 kHz. Meanwhile, as shown by the dotted line, the sensitivity of the microphone implanted in the guinea pig was measured in the flat region from 0.1 to 1 kHz at about – 40 dB, resonance occurred at about 1.5 kHz, and the sensitivity decreased by 20 dB/decade after the first resonance. The resonant frequency shifted from 6 ~ 8 kHz to 1.5 kHz due to the loading effect of the skin, which affects the diaphragm of the implantable microphone. Typically, the resonant frequency of circular diaphragm is decreased according to the increases of diaphragm mass [11]. In this study, the resonant frequency of microphone was extremely shifted by added mass of diaphragm and animal skin.

2.2 Speech quality

To evaluate the quality of the speech signals measured by the implanted microphone, various objective analysis methods were investigated using MATLAB. These analyses included perceptual, spectral, composite, distortion, and signal-to-noise–ratio analyses.
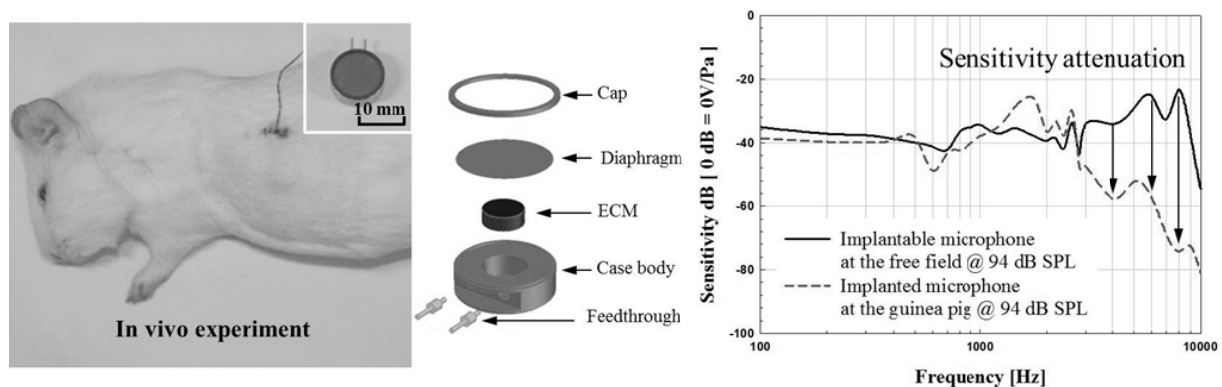

Fig. 2. Photograph of microphone implanted in guinea pig and frequency response.

PESQ is an important parameter for performing a perceptual analysis, which is the most complex to compute. The PESQ indicator is expressed in Eq. (1) as a linear combination of the average symmetric disturbance value D_{ind} and average asymmetrical disturbance value A_{ind} [12], where coefficient value a_0 is 4.5, a_1 is -0.1, and a_2 is -0.0309. In this paper, the PESQ diameter was calculated using the equation developed by Rix et al. and their coefficient values [13], which were computed using an auditory transform, followed by optimization of the compensation coefficients of D_{ind} and A_{ind}.

$$PESQ = a_0 + a_1 D_{ind} + a_2 A_{ind} \tag{1}$$

Normal subjective test values fall between 1.0 for bad and 4.5 for no distortion, thus a higher PESQ value indicates a better speech quality.

The weighted spectral slope (WSS) measures the direct spectral distance, which is based on a comparison of smoothed spectra from clean and distorted speech samples. The WSS indicator is expressed in Eq. (2), where $W(j, m)$ is the computed weight, $S_c(j, m)$ and $S_d(j, m)$ are the spectral slopes for the enhanced speech signal frames in the frequency band, K is the total number of critical bands, and M is the number of data segments[13,14].

$$WSS = \frac{1}{M} \sum_{k=0}^{M} \frac{\sum_{j=1}^{10} W(j,m)(S_c(j,m) - S_d(j,m))^2}{\sum_{j=1}^{k} W(j,m)} \tag{2}$$

The composite objective measures are a linear combination of four basic objective measures, including the PESQ, WSS, segmental SNR (SNR_{seg}), and log likelihood ratio (LLR), indicating the correlation between speech and noise distortions and the overall quality [15]. In this paper, the composite measure for signal distortion (C_{sig}) and composite measure for the overall speech quality (C_{ovl}) were calculated using Eqs. (3) and (4). The parameter and weighting factor for Eqs. (3) and (4) were computed based on the method of Hu [15], where higher values for C_{sig} and C_{ovl} indicate a better speech quality.

$$C_{sig} = 3.093 - 1.029 LLP + 0.603 PESQ - 0.009 WSS \tag{3}$$

$$C_{ovl} = 1.594 + 0.805 PESQ - 0.512 LLR - 0.007 WSS \tag{4}$$

The total harmonic distortion (THD) is determined from the fundamental frequency and first to fifth harmonics using a modified periodogram of the same length as the input signal [16]. In this paper, the THD was computed using Eq. (5), where I_1 is the magnitude of the fundamental frequency and $I_3, I_4,$ and I_5 are the magnitudes of the harmonics, respectively. For microphones, the reference distortion limit is generally established as the sound pressure level at 1 kHz that produces a THD of less than 0.5% at the microphone output voltage.

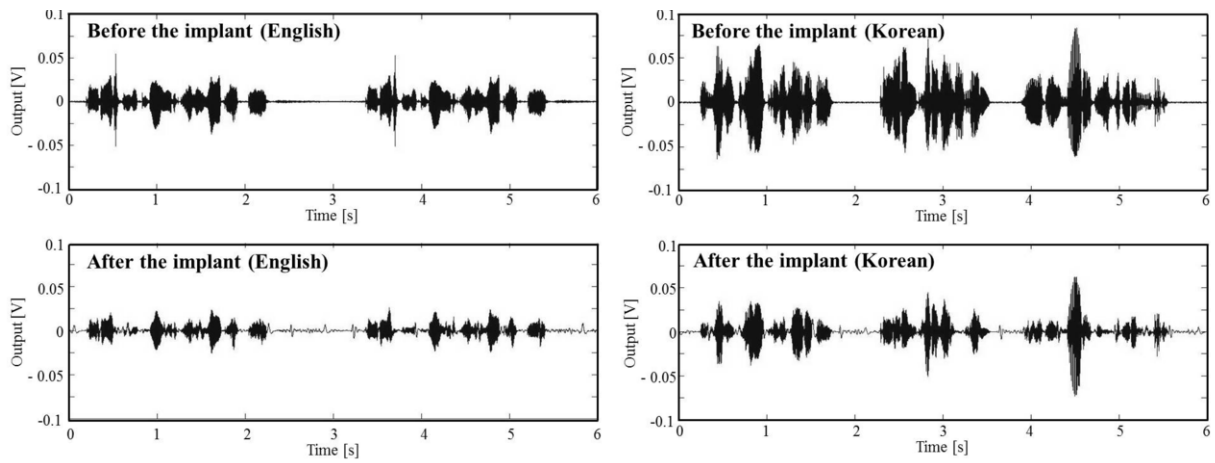$$THD = \frac{\sqrt{I_2{}^2 + I_3{}^2 + I_4{}^2 + I_5{}^2}}{I_1} \ (\%) \tag{5}$$

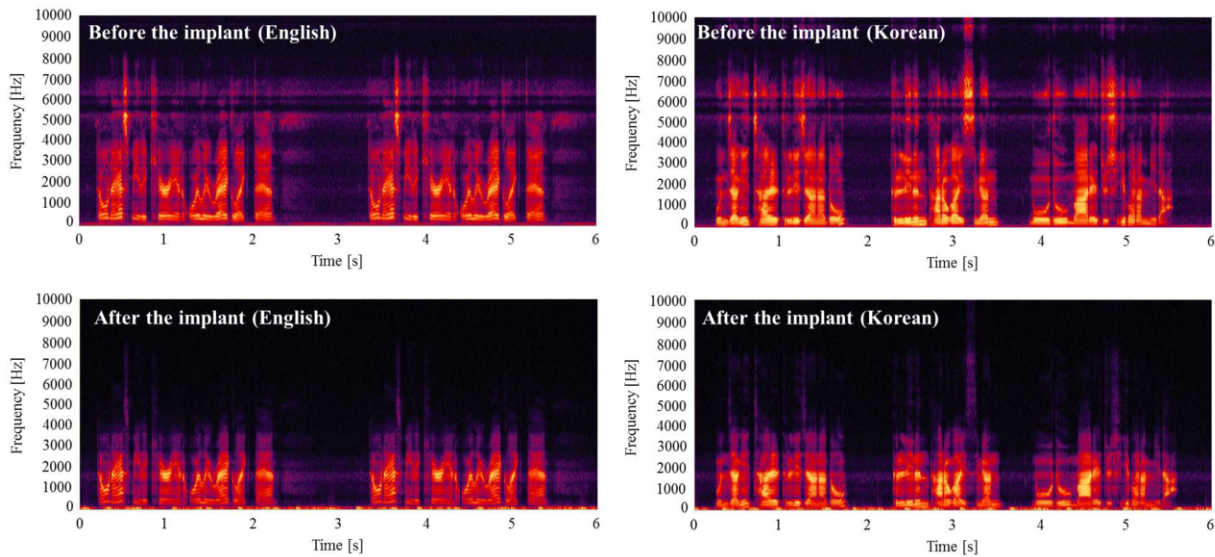Fig. 3. Speech signals measured by implanted microphone.

Fig. 4. Spectrum analysis of measured speech signals.

3. Results and discussion

Figure 3 shows the speech signals measured by the implanted microphone. After implantation, the speech signal values measured by the microphone were low, and the signals were distorted. Figure 4 shows the spectral analysis of the measured speech signals. The intensity of the speech signals after implantation was weaker above 3 kHz. However, the intensity of the speech signals below 2 kHz was similar before and after the implantation.

Tables 1 and 2 show the speech quality indicators for the signals measured by the implanted microphone and the THD for the implanted microphone. According to Table 1, after implanting the micro-

phone in the guinea pig, the PESQ, C_{sig}, C_{ovl}, SNR_{mean}, and mean SNR (SNR_{seg}) generally decreased by 0.495, 1.978, 3.265, 2.508, and 30.062, respectively, in the case of English words, and by 0.219, 1.917, 3.604, 0.534, and 32.43, respectively, in the case of Korean words. A higher PESQ, C_{sig}, C_{ovl}, SNR_{mean}, and SNR_{seg} indicates a better speech quality. Meanwhile, the WSS value increased to 41.13 and 41.009 in the case of English and Korean words, respectively. This means the WSS was degraded, as a lower WSS value indicates a better speech quality. According to Table 2, the THD measured in air was less than 0.5 and 1% at 109 and 115 dB SPL, respectively, while the THD measured under the skin was less than 0.5 and 1% at 101 and 107 dB SPL, respectively. Plus, the maximum sound pressure was about 8 dB lower after the microphone was implanted. The THD is a form of nonlinearity that causes unwanted signal to be added to the fundamental signal that is harmonically related to it. Typically, the THD and maximum sound pressure level of microphone are expressed the dB SPL at either 0.5 % or 1 %, and where the higher level of sound pressure indicate a better microphone performance.

4. Conclusion

This study objectively measured the speech quality and frequency characteristics of a subcutaneously implanted microphone based on an *in vivo* experiment. To evaluate the speech quality of the microphone, the PESQ, C_{sig}, C_{ovl}, SNR_{seg}, SNR_{mean}, THD, and WSS as parameter values were calculated using MATLAB before and after implantation. As a result, the sensitivity of the microphone implanted under the skin decreased due to the effect of skin attenuation, particularly in the high frequency range. Furthermore, the PESQ, C_{sig}, C_{ovl}, SNR_{seg}, SNR_{mean}, THD, and WSS all decreased after implantation. Accordingly, it was confirmed that when the proposed parameter values negatively deviated from the original values (before implant), this also represented a decrease in the discrimination ability between voice and noise. Thus, the proposed quality evaluation method can be a useful tool for developing high-performance implantable microphones for future implantable hearing devices.

Table 1

Sound quality before and after implantation

Index / Data set	In air (English)	Under skin (English)	In air (Korean)	Under skin (Korean)
PESQ	4.5	4.005	4.5	4.281
WSS	0	41.130	0	41.009
C_{ovl}	5.216	3.238	5.126	3.209
C_{sig}	5.806	2.541	5.806	2.202
SNR_{mean}	Infinite	-2.508	Infinite	-0.534
SNR_{seg}	35	-4.938	35	-2.57

Table 2

Total harmonic distortion before and after implantation

Index / Data set	In air (Sound pressure level)	Under skin (Sound pressure level)
THD 5-order less than 0.5%, 1kHz	109 dB SPL	101 dB SPL
THD 5-order less than 1%, 1kHz	115 dB SPL	107 dB SPL

Acknowledgement

References

[1]   S. Kochkin and MarkeTrak VIII, Consumer satisfaction with hearing aids is slowly increasing, Hearing Journal **63** (2010), 19–32.

[2]   J.H. Spindel, Middle ear implantable hearing devices, American Journal of Audiology **11** (2002), 104–113.

[3]   S. Stenfelt, N. Hato and R.L. Goode, Fluid volume displacement at the oval and round windows with air and bone conduction stimulation, Journal of Acoustical Society of America **115** (2004), 797–812.

[4]   D.W. Kim, K.W. Seong, J.H. Cho et al., A 1-channel 3-band wide dynamic range compression chip for vibration transducer of implantable hearing aids, Bio-Medical Materials and Engineering **24** (2014), 1009–1017.

[5]   P. Counter, Implantable hearing aids, Journal of Engineering in Medicine **222** (2008), 837–852.

[6]   J.H. Han, M.W. Kim, D.W. Kim et al., Design and fabrication of an implantable microphone for reduction of skin damping effect through FEA simulation, Korea Journal of Biomedical Engineering **29** (2008), 59–65.

[7]   I.P. Herman, Physics of the Human Body, Springer, Berlin-Heideler, New York, 2007, pp. 555–567.

[8]   E.S. Jung, H.G. Lim, K.W. Seong et al., Implantable microphone with acoustic tube for fully implantable hearing devices, IEICE Transaction on Information and Systems **8** (2011), 215−219.

[9]   A.E. Deddens, E.P. Wilson, J.M. Fredrickson et al., Totally implantable hearing aids: The effects of skin thickness on microphone function, American Journal of Otolaryngology **11** (1990), 1–4.

[10]   S.T. Woo, E.S. Jung, J.H. Cho et al., *In vivo* evaluation of mastication noise reduction for dual channel implantable microphone, Bio-Medical Materials and Engineering **24** (2014), 439–444.

[11]   J. Zhao and S. Yu, Effects of residual stress on the hydro-elastic vibration of circular diaphragm, World Journal of Mechanics **2** (2012), 361–368.

[12]   S. Moller and R. Heusdens, Objective estimation of speech quality for communication systems, Proceedings of the IEEE **101** (2013), 1955–1967.

[13]   A. Rix, J. Beerends, M. Hollier and A. Hekstra, Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs, Proc. IEEE Int. Conf. Acoust., Speech **2** (2001), 749–752.

[14]   K. Kondo, Subjective quality measurement of speech, Signal and Communication Technology **14** (2012), pp. 32–40.

[15]   Y. Hu and P. Loizon, Evaluation of objective quality measures for speech enhancement, IEEE Transaction on Audio, Speech, and Language processing **16** (2008), 229–238.

[16]   J. Arrillaga and N.R. Watson, Power System Harmonics, John Wiley and Sons, Chichester, 2003, pp. 11–22.