Acoustic analysis on the speech of a cochlear implant patient: A case study of post-lingual deafness
Risa Shimizu, Yuko Sakaguchi, Rion Iwasaki, Takayuki Arai, Moriyuki Mano, Atsushi Kawano and Kyoko Shirai

1Sophia University, 7–1 Kioi-cho, Chiyoda-ku, Tokyo, 102–8554 Japan
2Speech, Language, Hearing Sciences Program, Graduate Center, City University of New York, 365 5th Ave, New York, NY 10016, United States
3Jinkō-chōkaku Jōhō-gakkai, 5–17–9 Hatanodai, Shinagawa-ku, Tokyo, 142–0064 Japan
4Tokyo Medical University Hospital, 6–7–1 Nishishinjuku, Shinjuku-ku, Tokyo, 106–0023 Japan

(Received 6 February 2019, Accepted for publication 7 May 2019)

Keywords: Cochlear implant, Post-lingual deafness, Hearing impairment, Intelligibility of speech

PACS number: 43.70.Dn [doi:10.1250/ast.40.399]

1. Introduction
Hearing-impaired people have problems with auditory feedback. In the case of post-lingual deafness, the longer the duration of the deafness the more the speech intelligibility deteriorates [1]. Moreover, speech sounds uttered by hearing-impaired people tends to become nasalized [2]. Therefore, some of the cochlear implant patients easily feel nervous about their speech communication.

In this study, an investigation was conducted to understand the acoustical characteristics of the utterance of a cochlear implant patient by observing speech waveforms, sound spectrograms, and spectra.

2. Participant
A male adult (Japanese native speaker) with cochlear implants on both his ears, cooperated with our experiments where he underwent an operation in which the cochlear implants were embedded, after suffering from a sudden severe hearing loss in 1988. He was a patient with high performance among cochlear implant patients.

3. Method of data collection
We conducted four different types of evaluations as follows:
- 100 syllable intelligibility test [3],
- Word intelligibility test [4] (list A),
- Sentence test [5] (textbook is "cherry"), and
- Accent and intonation [5].

In all of these tests, utterances by the participant were recorded for acoustic analysis. In addition, we recorded the same speech samples of a normal-hearing listener for comparison. The recordings were done with a digital recorder (Marantz, PMD670), where the sampling frequency and the quantization were set to 44.1 kHz and 32 bits, respectively. A unidirectional microphone (Sony, EMC-23F5) with a stand was placed approximately 15 cm in front of the target’s mouth.

4. Analysis
We analyzed the data by using an acoustic analysis software, ‘Praat’ [6], as described below.

4.1. Transcription analysis
Three of the listeners commented that syllables, words and sentences were frequently nasalized, in general. While in an actual conversation, it is possible to predict the content of the conversation from the context, hence it is considered that the intelligibility becomes much higher than the tests in the current study. From the auditory impression, accents and intonations had no problem regarding the fluency in the sentences.

4.2. Formant of the vowels
In the 100 syllable intelligibility test, we measured the first formant frequency (F1) and the second formant frequency (F2) of the isolated Japanese five vowels [8]. Table 1 shows the formant frequencies of the vowels. From the results, it was found that the values of the vowel formants in the target utterance were close to the average by normal-hearing data [9].

4.3. Acoustic analysis of the consonants
We focused on the following four points from the error analysis.

4.3.1. Voicing of consonants
The voice onset time (VOT) was measured regarding all 120 stop consonants in the word intelligibility test [10]. Table 2 shows the VOT values of consonants. No matter whether the listeners perceived an utterance correctly, the parameter values of all target sounds were within the average range of each phoneme [11].
4.3.2. Fricatives — affricates
It was observed that the fricative energy is completely weak by measuring the waveforms and spectra. Figures 1 and 2 show the waveform and spectrogram of three-mora word “musume” in the word intelligibility test. A burst is observed at the starting point of /s/. In addition to the sound being extremely weak in frication, a closure section which is usually unseen in fricative sounds is formed. Considering the “su” sound at the end of the sentence, it can be assumed that the sound was not configured correctly because there were cases where the sound “su” is either heard as “tsu” or was even barely heard. In the case where “su” was heard as “tsu” it was obviously observed that there were a closure section and a burst. As the case was much in the medial position of words, it could be effective to conduct an articulation training that eliminates bursts before fricative sound in words.

4.3.3. Intensity of consonant
In the unvoiced consonants of the fricative and the affricate, we compared the difference of intensity between the consonant and the following vowel for the participant with a normal-hearing listener. The value was measured at the most stable intensity point around the center of each consonant. In case the syllable containing a consonant was devoiced, the differences were calculated between the intensities of the target consonant and the preceding vowel.

As a result, in some sounds including “musume” and “kitsune” with hearing errors, the participant’s intensity difference was 10 dB larger than that of the normal-hearing listener. Furthermore, among 40 target sounds, the median was calculated to 15 sounds which is an exceptionally large value obtained by subtracting the difference between the intensity of the normal-hearing listener and the intensity of the hearing-impaired participant. Compared to them, the intensity difference of the participant was 10 dB larger than the normal-hearing listener. From this result, it is confirmed that the frication of the unvoiced fricatives is weak in the participant’s utterance.

4.3.4. Nasalization
In addition, the spectrum of vowel /e/ was observed. Figures 3 and 4 show a spectrum of /e/ in isolated “e,” and Figs. 5 and 6 show a spectrum of /e/ in “musume.” It can be confirmed that the F1 and F2 were clearly observed in Figs. 3, 4, and 5; however, there were anti-formants near the F1 and nasal formants between the F1 and the F2 in Fig. 6. This

Table 1 Formant frequencies of the vowels (Hz).

<table>
<thead>
<tr>
<th></th>
<th>/a/</th>
<th>/i/</th>
<th>/u/</th>
<th>/e/</th>
<th>/o/</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH</td>
<td>F1</td>
<td>790</td>
<td>250</td>
<td>340</td>
<td>460</td>
</tr>
<tr>
<td></td>
<td>F2</td>
<td>1,180</td>
<td>2,300</td>
<td>1,180</td>
<td>2,060</td>
</tr>
<tr>
<td>CI</td>
<td>F1</td>
<td>906</td>
<td>323</td>
<td>380</td>
<td>490</td>
</tr>
<tr>
<td></td>
<td>F2</td>
<td>1,351</td>
<td>2,008</td>
<td>1,332</td>
<td>1,811</td>
</tr>
</tbody>
</table>

NH: normal-hearing listeners [9], CI: the participant in the present study.

Table 2 VOT of consonants (ms).

<table>
<thead>
<tr>
<th>phoneme</th>
<th>/p/</th>
<th>/b/</th>
<th>/t/</th>
<th>/d/</th>
<th>/k/</th>
<th>/g/</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH</td>
<td>41</td>
<td>30</td>
<td>30</td>
<td>75</td>
<td>66</td>
<td>75</td>
</tr>
<tr>
<td>CI</td>
<td>31</td>
<td>28</td>
<td>20</td>
<td>43</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

NH: normal-hearing listeners [11], CI: the participant in the present study.
difference comes from a change of the spectra with nasalization of “musume.” Therefore, it can be assumed that the nasalization occurred in the utterance of /e/ in “musume.”

4.4. Issues of accents and intonation

In accents and intonation, there are no large differences among the participant and the normal-hearing because of the participant’s proper accents and intonation.

5. Summary

In this study, we investigated speech sounds of a cochlear-implant patient from the perceptual and acoustical viewpoints by analyzing the waveforms, the spectrograms, and the spectra. Speech sounds of the participant in the present study were generally clear, and there were only few points which differed from that of normal-hearing.

Regarding the fricatives and the affricates, it was confirmed that the frication energy is weak after observing the speech waveforms and the spectrogram. Also, it was confirmed that the intensity difference of the participant was larger than the normal-hearing. In the case where “su” is heard as “tsu,” the frication was extremely weak and there were a closure section and a burst. Moreover, some sounds became nasalized.

This study should be a starting point of our future studies with more varieties of other cochlear implant patients.

Acknowledgements

The authors would like to extend their gratitude to Dr. Eri Osawa of the Arai lab, and Dr. Setsuko Imatomi of Mejiro University for their constant support during this study.

References