Abstract

During speaking, humans perceive their own voices to control their speech production systems. This mechanism is reffered to as auditory feedback. Previous psychological and neurological studies related to auditory feedback have suggested the importance of acoustical monitoring during speech production. Those studies have mainly focused only on air-conducted (AC) speech transmission of one's own voice. However, there is another type of speech transmission: bone-conducted (BC) speech. It is still unclear whether/how BC speech perception during speaking affects one's speech production as auditory feedback, as well as AC speech. Although some studies have used acoustical stimuli such as a pink noise for masking one's BC speech, it has not been clarified whether the noise stimuli successfully mask one's BC speech enough or not. For further exploring the auditory information during speaking, the following things need to be understood: (1) acoustical characteristics of BC speech, (2) transmission pathways and transmission characteristics, (3) perceptual contribution of each transmission pathway.

Physiological studies related to BC hearing have argued that the glottal vibration and the sound inside the vocal tract are transmitted to the auditory system (i.e., the outer, middle and inner ear) as BC speech through multiple pathways. Since BC speech transmission to the middle and the inner ear cannot be observed directly, previous studies have not obtained consistent findings regarding the contribution of the middle/inner ear pathways to BC speech perception.

This study aims to clarify the process of the BC speech transmission from the vocal organ to the auditory system from the above aspects ((1) to (3)), hypothesizing that the spectral characteristics of the BC speech component reacing the middle/inner ear correspond to those of the vibration of regio temporalis (RT).

Firstly, the fundamental frequency (F0) and the spectral characteristics of BC speech were analyzed focusing on the RE vibration and the sound radiation in the ear canal (EC) during speaking. It was found that the RT vibration included the same information of the fundamental frequency (F0) and the lower-order formants below 2 kHz as AC speech, while the EC sound radiation includes the same information of F0, the first formant (F1) and the second formant (F2) as AC speech.

Secondly, transmission characteristics from the vocal organ to the RT and the EC were measured using transcutaneous excitation on the larynx and excitation from a sound source in the oral cavity. The measurement found that BC transmission from the vocal organ to the outer ear has an effect of band-pass filtering between 1 to 3 kHz, while BC transmission from the vocal organ to the middle/inner ear has an effect of low-pass filtering below 2 kHz.

Thirdly, voice timbre of one's own voice were subjectively evaluated using the transmission characteristics obtained above, to investigate perceptual contribution of AC and BC speech transmission. The evaluation found that the middle-/inner-ear part of transmission contributes BC speech perception almost as the same extent as the outer-ear part of transmission.

From the three findings above, this research indicated that not only the outer-ear but also the middle/inner-ear part of BC speech transmission may play a role mainly in monitoring the pitch information as auditory feedback. The transmission characteristics obtained in this study is helpful in designing the masking stimuli for the auditory feedback experiment. The role of BC speech perception in auditory feedback is expected to be revealed in future studies.

Keywords: Auditory feedback, bone-conducted speech, transmission characteristics, vocal organ, auditory system