1P5-15

Speech perception by distantly-presented bone-conducted ultrasound: Effects of AM-method and speaker gender

Naoya Takahashi[1‡], Sho Otsuka[1,2,3], and Seiji Nakagawa[1,2,3*] ([1] Grad. School Sci. and Eng., Chiba Univ.; [2] Center for Frontier Med. Eng., Chiba Univ.; [3] Med-Tech Link Center, Chiba Univ. Hospital)

1. Introduction

Ultrasound, high-frequency sound above 20 kHz, cannot be perceived by air-conduction, however, it can be perceived clearly by bone-conduction (BC). This "bone-conducted ultrasound (BCU)" can be perceived even by the profoundly sensorineural hearing impaired, and can transmit speech by using amplitude-modulation (AM). [1)]

Two pitches are perceived by AM-BCU; one is due to carrier and another came from the modulator. It is thought that hearing impaired people perceive the pitch corresponding to the envelope of modulated ultrasound as temporal information. [2)] On the other hand, in normal-hearing people, it is likely to be even perceived as a self-demodulation compornent is caused by non-linearities of the biological tissues that exist in propagation paths. [3)]

BC stimuli are typically presented to the mastoid process of the temporal bone, however, BCU can be perceived even when presented to body parts distant from the head, e.g., the neck, trunk, and upper arm. [4)] In the distantly-presented BCU, the distance between the stimuli location and the auditory receptors (cochleae) is longer and more complex than ordinally BCU. Therefore, it is likely that more self-demodulation components will be generated. Previous research has indicated that the self-demodulation component could enhance hearing of disatntly-presented BCU. [5)]

The envelope of the modulated wave is determined by a AM-method and a modulator. Previous studies have investigated the effect of AM-method on monosyllable articulation and word intelligibility of AM-BCU presented to the mastoid process. [5, 6)] However, the effects of AM-method and characteristics of modulator have not been systematically investigated in the distal presentation.

In this study, a monosyllable articulation test was conducted to assess how the AM-method and the speaker gender affect the hearing of distantly-presented AM-BCU.

2. Methods

The experiments were conducted on 7 normal hearing participants (5 males and 2 females, 21–23 years). 30-kHz sinusoid amplitude-modulated by Japanese monosyllables were used as stimuli. Following previous studies, [5)] we used four AM-methods; DSB-TC, Transposed, DSB-SC, and SSB. The amplitude of each monosyllable was adjusted to avoid over modulation. **Fig. 1** shows examples of temporal waveforms and frequency spectra for each AM-method. We used 100 Japanese-monosyllables uttered by four speakers (2 males and 2 females) from the commercially available database (FW03, NTT-AT).

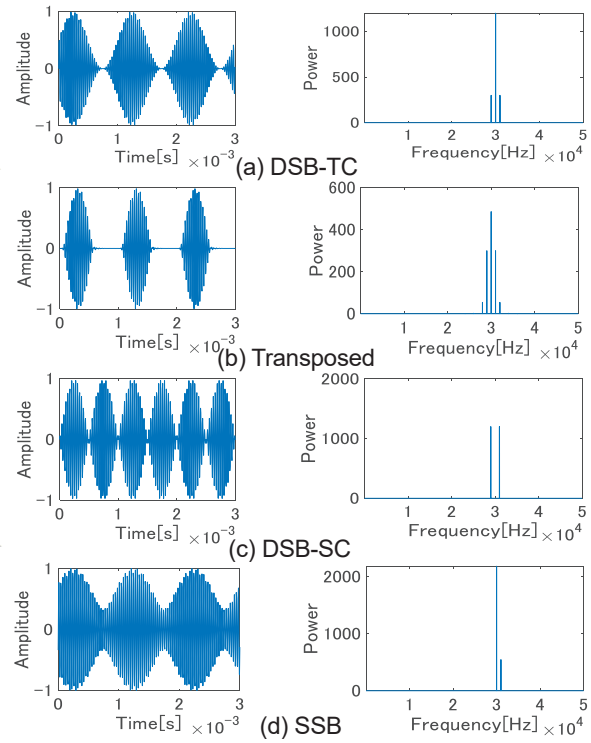

Fig. 1 Temporal waveforms (left column) and frequency spectra (right column) of AM-BCU using each AM-method. The carrier and modulator are 30-kHz and 1-kHz sinusoid.

AM-BCU stimuli were presented to the mastoid, sternocleidomastoid, and clavicle as shown in **Fig. 2**. For all body parts, stimuli were presented in random order and with 5-s interval. Subjects were instructed to enter the monosyllables they heard into a PC. Intensities of the stimuli were adjusted to the most clearly perceiving level for each participant.

E-mail: ‡n-takashi@chiba-u.jp, *s-nakagawa@chiba-u.jp

Fig. 2 Presentation of AM-BCU stimulus in the experiment.

3. Results and Discussions

Fig. 3 shows percent corrects obtained under each condition. Both DSB-TC and Transposed showed higher articulation than DSB-SC and SSB. A similar pattern was observed in a previous study examining in the mastoid process. [5] This indicates that the increase of the self-demodulation component due to distal presentation is not large enough to overcome the differences in hearing among the AM-methods.

As the presentation location moved away from the head, the articulation decreased. This result is similar to previous studies, [4] indicated that there might be some missing or distorted information in the transmission paths. However, the sternocleidomastoid showed a higher articulation compared to the mastoid when the female voice was modulated with DSB-TC and Transposed. These results suggest that improvements in hearing performance due to the self-demodulation component may have exceed the energy loss in the transmission process.

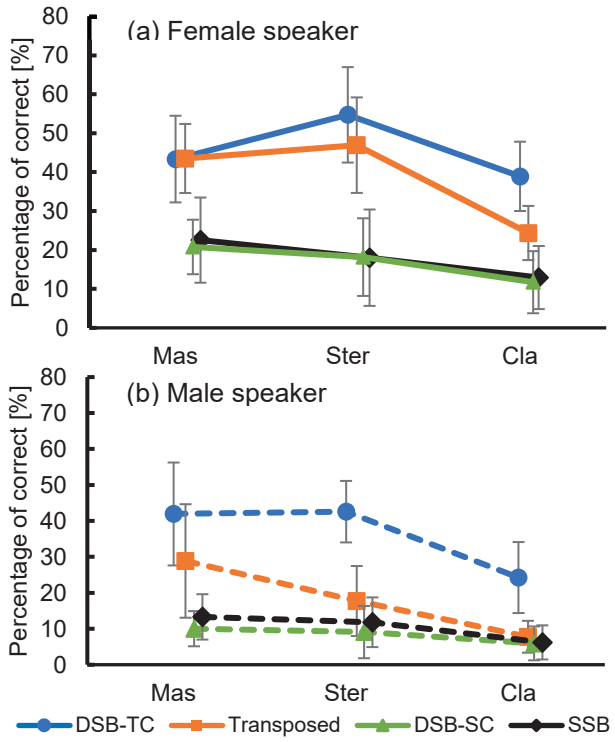

Fig. 3 The score of the monosyllable articulation tests for each condition. “Mas”, “Ster” and “Cla” stand for the mastoid, sternocleidomastoid and clavicle, respectively.

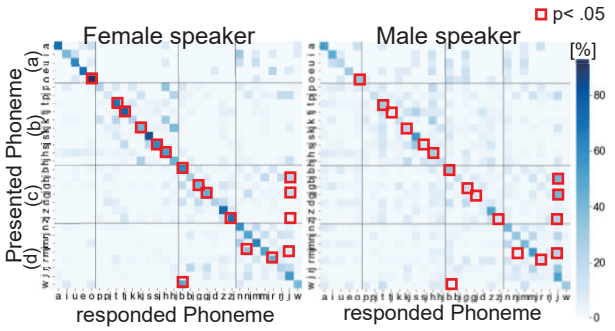

Fig. 4 Confusion matrices obtained in the monosyllable articulation tests when Transposed stimuli was presented to the sternocleidomastoid.

In addition, female voices were more intelligible than male voices in most conditions. Previous studies of the mastoid have reported no clear differences in intelligibility or articulation based on speaker gender. [6] The discrepancy may arise from self-demodulation component generated in transmission processes.

Fig. 4 shows examples of confusion matrices, i,e, the probability that a presented phoneme is perceived as a different phoneme, when Transposed stimuli was presented to the sternocleidomastoid. In this condition, speaker-gender differences were very apparent. For male voices, the probability of mishearing /bj/, /gj/, /zj/ for /j/ is significantly higher than for female voices. It has been previously reported that the mishearing rate for the contracted sound /j/ is higher in BCU sounds. In the current experiment, this tendency was more pronounced in male voices, possibly because distortion by Transposed modulation made it difficult to capture the formant frequency change in the fricative and wandering parts of the contracted sound.

Moreover, DSB-SC and SSB also had a higher probability of hearing the presented vowel differently. It is possible that formant information for vowel identification is lost during modulation.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Numbers 19K22950 and 20H04497.

References

1) ML Lenhardt, R. Skellett, P. Wang, and AM Clarke, Science, **253** (1991) 82-85.
2) S. Nakagawa, C. Fujiyuki, and T. Kagomiya, Jpn. J. Appl. Phys., 51 (2012) 07GF22.
3) K. Ito and S. Nakagawa, Jpn. J. Appl. Phys., **54** (2015) 07HF07.
4) S. Nakagawa, R. Ogino, and S. Otsuka, Jpn. J. Appl. Phys., 57 (2018) 07LD22.
5) K. Doi, R. Ogino, S. Otsuka, and S. Nakagawa, Jpn. J. Appl. Phys., **59** (2020) SKKE26.
6) T. Kagomiya, and S. Nakagawa, Proc. Speech Prosody 2014, 467-471, (2014).