1P5-12

Speech perception by distantly-presented bone-conducted ultrasound: Assessment by phonetic-feature transmission analysis

Seiji Nakagawa[1, 2, 3†], Koichiro Doi[4], and Sho Otsuka[1, 2, 3] ([1]Ctr. for Frontier Medical Eng., Chiba Univ.; [2]Grad. School of Eng., Chiba Univ., [3]Med-Tech Link Ctr., Chiba Univ. Hospital, [4]Dept. of Medical Eng., Grad. School of Sci. & Eng., Chiba Univ.)

1. Introduction

Bone-conducted ultrasound (BCU) can be experienced as sound by the severely hearing-impaired as well as the normal-hearing.[1-3] We have been working on elucidation of perception mechanisms of BCU[4-6] and the application to a novel hearing-aid (BCU hearing aid).[7, 8]

In the BCU hearing aid, ultrasonic sinusoids with a frequency of about 30 kHz are amplitude-modulated by speech and presented to the mastoid process of the temporal bone, which is located behind the pinna. However, BCU can be heard not only on the mastoid process, but also on wider area of the body; for example, the forehead, the neck, the trunk, the superior limb, and the inferior limb. In the previous study, we measured thresholds of 30-kHz tone bursts presented to the neck and the upper and lower arms in normal hearing participants. The results showed that BCUs presented to the distal parts, including the lower arm, can be perceived whereas threshold increased depending on the distance from the head.[9)]

The air-conducted components radiated from the vibrators cannot be perceived, and ther is no vibratory sensation caused by the BCU. Therefore, this "distantly-presented BCU" can be applied to a novel audio interface that can transmit the sound information selectively to users who touch the device.[10)] However, no previous study has examined distantly-presented BCUs' ability to transmit speech information.

In this study, monosyllable articulation tests were conducted, to assess transmission of speech information by distantly-presented BCU. Additionally, sequential information transfer analyses (SINFA)[11, 12)] were conducted to determine the type of articulatory features that were well transmitted.

2. Methods

2.1 Subjects

10 adults (males, 22-26 years old), with normal-hearing, whose mother tongue was Japanese, participated in the experiment.

2.2 Speech materials

100 Japanese monosyllables recorded in a female voice were taken from a commercially available database (NTT-AT FW03). A mono-syllable is a component of mora.

†s-nakagawa@chiba-u.jp

2.3 Stimulus and procedure

Speech sounds were presented under both bone-conducted (BC) and air-conducted (AC) conditions. In the BC condition, ultrasounds were amplitude-modulated by speech sounds using two modulation methods: double side band-transmitted carrier (DSB-TC) and transposed modulation7). The carrier frequency was set at 30 kHz. BCU stimuli were presented against the following body parts on the left side of the body: (a) Mastoid process of the temporal bone, (b) Sternocleidomastoid muscle (muscle of the neck), (c) Clavicle. The presentation pressure was set at 2 N for the sternocleidomastoid muscle and 5 N for the mastoid and clavicle. In the AC condition, words and monosyllables recorded in FW03 were presented as-is for comparison to BC condition.

In both AC and BC conditions, 100 monosyllables were presented randomly. The stimulus interval was 5.0 s. Subjects were required to write down what they heard after each word or monosyllable was presented. The intensity of BC stimuli was set at most clearly perceiving level, and the average AC sound level was set at 20 dB above the threshold.

2.4 SIFNA

To analyze what types of phonetic features were transmitted, SIFNA was conducted separately for the vowels and consonants. SINFA determines the ratio of phonetic features transferred using the confusion matrices of the monosyllable articulation tests and defined phonetic features matrices.

3. Results and Discussions

Words and monosyllables were recognized at all stimulus placements. The AC condition score was higher than that of BC condition ($p < .01$), except for between the AC and the sternocleidomastoid DSB-TC. In the BC condition, the main effect of the modulation method was significant ($p < .01$). In the sternocleidomastoid and clavicle, the score for DSB-TC was higher than that for Transposed.

Confusion matrices for each stimulus type and stimulus placement are shown in Fig. 1. The SINFA results for vowels and consonants are shown in Table 1 and 2, respectively. For vowels, the transfer ratios of "openness" in the AC condition were higher than those in the BC condition. On the other hand, several BC condition showed a higher transfer ratio than the AC conditions for "frontness."

Additionally, at the mastoid, the transfer ratios for "frontness" were higher than those for "openness". By contrast, at the distal body parts

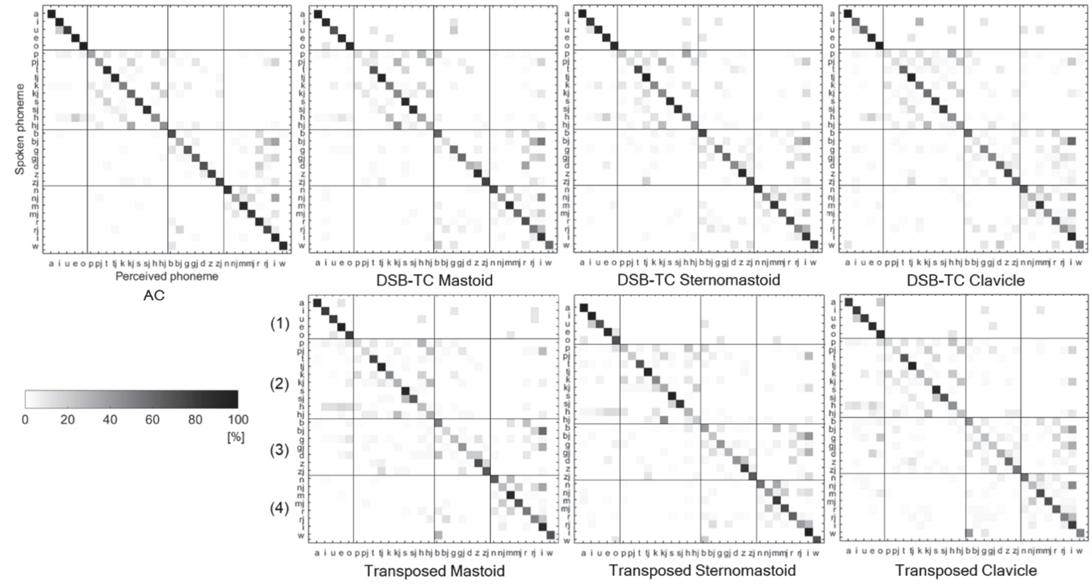

Fig. 1 Confusion matrices for each stimulus type and stimulus placement. Each monosyllable was compressed into 30- elements and classified into four categories: (1) vowels, (2) unvoiced, (3) voiced, and (4) others.

Table 1 The transfer ratio (%) (Transmitted information (bits)) for vowels obtained by SIFNA.

Phonetic feature	AC	Mastoid		Sternomastoid		Clavicle	
		DSB-TC	Transposed	DSB-TC	Transposed	DSB-TC	Transposed
openness	100.00 (1.52)	92.07 (1.42)	86.98 (0.67)	100.00 (1.54)	93.29 (1.43)	100.00 (1.53)	86.74 (1.33)
frontness	80.50 (0.64)	97.23 (0.75)	91.85 (1.42)	87.36 (0.68)	63.66 (0.50)	87.02 (0.66)	79.86 (0.62)

Table 2 The transfer ratio (%) (Transmitted information (bits)) for consonants obtained by SIFNA.

Phonetic feature	AC	Mastoid		Sternomastoid		Clavicle	
		DSB-TC	Transposed	DSB-TC	Transposed	DSB-TC	Transposed
Voicing	67.50 (0.55)	65.57 (0.64)	50.41 (0.41)	59.44 (0.49)	55.86 (0.46)	60.91 (0.60)	57.35 (0.56)
Nasality	72.40 (0.48)	61.48 (0.31)	58.67 (0.40)	67.16 (0.44)	62.73 (0.42)	54.97 (0.28)	43.89 (0.23)
Manner	58.67 (0.29)	65.57 (0.64)	35.14 (0.39)	43.00 (0.48)	38.69 (0.43)	47.23 (0.23)	45.25 (0.23)
Position	49.05 (0.81)	40.43 (0.67)	29.33 (0.30)	39.94 (0.41)	30.68 (0.32)	33.92 (0.56)	31.28 (0.51)
C. S.	70.23 (0.66)	57.46 (0.54)	47.1 (0.44)	60.19 (0.57)	51.36 (0.48)	51.70 (0.49)	45.14 (0.43)

(sternocleidomastoid and clavicle), the transfer ratios for "openness" were higher than those for "frontness." The transfer ratio for "openness" was the highest at the sternocleidomastoid, whereas the transfer ratio for "frontness" was the highest at the mastoid. For consonants, the transfer ratios were generally lower than those of vowels. The transfer ratio of "position" was the lowest of all the phonetic features. Additionally, the sternocleidomastoid showed the highest transfer ratio for the "nasality" and "contracted sounds."

These results indicate the phonetic features transfer ratios of vowels varied depending on the stimulus placements, and suggest that the feature information transmission may depend on the characteristics in the biological tissues such as attenuation rate and frequency characteristics. The detailed elucidations of these characteristics are necessary to optimize distantly-presented BCUs for speech transmission.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Numbers JP19K22950 and JP20H04497.

References

1) R. J. Pumphrey: Nature **166** (1950) 571.
2) R. J. Bellucci and D. E. Schneider: Ann. Otol. Rhinol. Laryngol. **71** (1962) 719.
3) M. L. Lenhardt et al.: Science 253 (1991) 82.
4) S. Nakagawa and A. Nakagawa: J. Acoust. Soc. Am. **120** (2006) 3123.
5) S. Nakagawa: Brain Topography and Multimodal Imaging (2009) 95.
6) S. Nakagawa and M. Tonoike: Unveiling the Mystery of the Brain, **ICS1278** (2005) 333.
7) S. Nakagawa et al.: Jpn. J. Appli. Phys. **51** (2012) 07GF22.
8) S. Nakagawa et al.: Jpn. J. Appli. Phys. **52** (2013) 07HF06.
9) S. Nakagawa et al., Jpn. J. Appl. Phys. **57** (2018) 07LD22.
10) R. Ogino et al., Jpn. J. Appl. Phys. **58** (2019) SGGE12.
11) M. D. Wang and R. C. Biliger, J. Acoust. Soc. Am., **54** (1973) 1248.
12) D. van Leeuwen, http://davidvanleeuwen.ruhosting.nl/resources/software/sinfa/ (2020).