Temporal modulation perception of bone-conducted ultrasound

Takuya Hotehama[a,b*], Seiji Nakagawa[a]

[a] National Institute of Advanced Industrial Science and Technology (AIST), 1-8-31, Midorigaoka, Ikeda, Osaka 563-8577, Japan

[b] Japan Society for the Promotion of Science (JSPS)

*e-mail: t-hotehama@aist.go.jp

(Received 30 November 2009; accepted 7 December 2009)

Ultrasonic vibration generates a sensation of sound via bone-conduction. This phenomenon is called bone-conducted ultrasonic (BCU) hearing. In order to clarify perceptive characteristics of temporal modulation for BCU, the influence of modulation frequency on the sensitivity for detecting amplitude modulation of sinusoidal carriers 10, 20 and 30 kHz was examined. Temporal modulation transfer functions (TMTFs) obtained at each carrier frequency suggest that the auditory system has an ability to process timing information in the envelopes of amplitude modulated BCUs at lower modulation frequencies, as is the case with audible sounds. At higher frequencies, the possible contributions of peripheral filtering on the shape of the TMTF were examined.

Key words: bone-conducted ultrasound, amplitude modulation, temporal modulation transfer function

1. INTRODUCTION

Ultrasound is generally recognized as sound with a frequency higher than the upper limit of human hearing; this limit for air-borne sound is believed to be about 20 kHz. However, the audibility of ultrasonic vibrations above 20 kHz at least and up to 100 kHz via bone-conduction has been reported [1]. This audible bone-conducted sound in the ultrasonic range is called the bone-conducted ultrasound (BCU). The psychoacoustic characteristics for BCU represent some interesting features [e.g. 2-6]; however, the underlying mechanisms remain unclear. Severely sensorineural hearing-impaired persons as well as normal-hearing people can perceive BCUs. Complex information can be conveyed by an amplitude-modulated BCU (AM-BCU). Profoundly hearing-impaired people as well as normal-hearing people can recognize phonetic information from voice-modulated BCUs [7,8]. Thus, we have developed a new hearing-aid system for profoundly hearing-impaired that utilize the perceptive characteristics for BCUs; bone-conducted ultrasonic hearing aids (BCUHA) [9]. While the mechanisms by which the modulation signal of an AM-BCU can be perceived are still unclear, various possible mechanisms have been proposed [7,8,10]. Detection of audible distortion products is one of the mechanisms considered. Since the AM-BCUs are presented using a piezo-electric vibrator coupling with the head, combination tones, especially difference tones can be easily generated by the nonlinearity of reproduction and/or transmission pathway. In comparison, it is possible that the auditory system is equipped with a facility that processes the temporal fluctuation of an amplitude envelope in ultrasonic frequencies, as in the case in sonic frequency.

The temporal modulation transfer function (TMTF) which is the detection threshold curve of modulation depth as a function of modulation frequencies is a useful measure for presenting perceptive characteristics of the amplitude-modulated sound. In previous studies, detailed measurements of TMTFs using sinusoidal and broadband noise carriers were performed (e.g., [11,12]). For sinusoidal carriers between 2-10 kHz and broadband noise carriers, the TMTFs have a commonly observed feature: a low-pass characteristic at low modulation frequencies, where the modulation detection thresholds are relatively low and constant at low modulation frequencies and increase with increasing modulation frequency above a certain cutoff modulation frequency. This low-pass characteristic for these types of carriers is thought to reflect the limitation on temporal processing of fast amplitude-envelope variations in the central auditory pathway rather than the inner-ear filtering [12].

On the other hand, a significant difference in the shapes of the TMTFs in a range of high modulation frequencies were shown between these two types of carriers; a TMTF for sinusoidal carrier decrease with increasing modulation frequencies above a certain modulation frequency, whereas that for broadband noise carrier increase with increasing modulation frequencies. This difference is thought to be due to a discrepancy in spectral forms; a spectrum form of a SAM with sinusoidal carrier consists of a carrier and two sidebands with the spectral distance from the carrier that equals the modulation frequency, while a spectrum form of a SAM with broadband noise carrier is approximately identical with that of the unmodulated carrier signal. Therefore, the decrease of the TMTF with sinusoidal carrier at high modulation frequencies is considered as a result of detection of sidebands as separate spectral components. The detection of spectral sidebands causes difficulty in the TMTF with sinusoidal carrier to specify whether the cues of modulation detection are temporal or spectral at high modulation frequencies. In addition, considering this detection of spectral component, TMTFs obtained with sinusoidal carriers do not exactly represent “temporal” aspects of hearing especially at high modulation frequencies; however, we use the term TMTF as an amplitude-modulation detection threshold curve as a function of modulation frequency in a conventional manner.

While a great deal of effort has been made on the perceptual characteristics for “sonic” SAM sounds using noise carriers and sinusoidal carrier, little is known about those for “ultrasonic” SAM sounds. The purpose of this study was to determine the influence of modulation frequency on modulation detection of amplitude modulation bone-conduction sounds above 10 kHz to ultrasonic range.

2. METHOD

2.1 Stimuli

In the detection task, a SAM and an unmodulated carrier were presented in the target and the standard intervals, respectively.

The formulas of the target stimuli are expressed as follows:

$$\text{SAM:}\quad (1+m\sin(2\pi f_m t))\sin(2\pi f_c t) \qquad (1)$$

Here, m is the modulation degree, and f_c and f_m are the carrier frequency of 10, 20 and 30 kHz and the modulation frequency, respectively. f_m ranged from 10–1600 Hz for the 10-kHz carrier and from 10–6400 Hz for the 20- and 30-kHz carriers. The spectrum of the SAM has three components which correspond to the carrier and one sideband above and one sideband below the carrier frequency with distances equivalent to the modulation frequency. The relative sideband levels of the SAM to the carrier level depend on m; the sideband level is $-6 + 20 \log m$ dB away from the carrier level.

The duration of the standard and target signals was 800 ms, including 150 ms rise/fall times shaped using a cosine ramp. The modulation was applied over the entire length of signals including the rise and fall. The inter-stimulus interval was 150 ms. Carrier levels of all stimulus intervals were kept constant. To achieve this condition, the intensity of the signal in the target interval was larger than that in standard intervals.

2.2 Psychophysical procedure

The modulation detection thresholds were measured in an adaptive three-interval, three-alternative forced choice (3-IFC) procedure using a transformed 2-down/1-up tracking method that estimates the modulation degree, m, required for 70.7% correct detection [13]. Subjects were required to identify the target interval out of three intervals in which the sinusoidal carrier was modulated. The modulation degree was varied in the target intervals. The step size of the modulation degree corresponded initially to 4 dB (in units of $20 \log m$) and was reduced to 2 dB after two reversals and finally to 1 dB after four reversals. The median of the modulation depth of the last 8 reversals in a block of 11 reversals was used as the detection threshold estimate for that block. Thresholds were based upon two estimates for each listener. The carrier levels of the stimuli were balanced in loudness with a 10-kHz sinusoid of 60 dB SPL presented via an earphone. The threshold of hearing for the carriers was also measured using an adaptive 3-IFC 2-down/1-up procedure similar to that for the modulation detection tasks to obtain the sensation levels of the carriers.

During the modulation detection tasks, masking noises, which were binaurally uncorrelated low-passed noises with certain cutoff frequencies, were continuously presented via earphones in order to mask the first-order combination tone at the modulation frequency possibly generated at the transducer and/or in the transmission pathway within the head.

2.3 Apparatus

Experiments were performed in an anechoic room at the AIST Kansai Centre. All stimulus waveforms were synthesized digitally by a PC (Dimension 8300, Dell) at a sampling frequency of 192 kHz, generated through a 24-bit D/A converter (Audiofire12, Echo Digital Audio), filtered by a low-cut filter (3625, NF), amplified by a piezo driver (M-2629B, MESS-TEK), and presented through a piezoelectric ceramic vibrator (MA40E7S, Murata) onto the mastoid using a fastening device. Masking noises were also synthesized digitally with 30-s duration on a notebook PC (VAIO, Sony) at a sampling frequency of 96 kHz, played back repeatedly with a sequence software (Audition, Adobe), generated through a 24-bit D/A converter (Layla24) attenuated by an attenuator (PA5, TDT) and presented through earphones (ER2, Etymotic Res.).

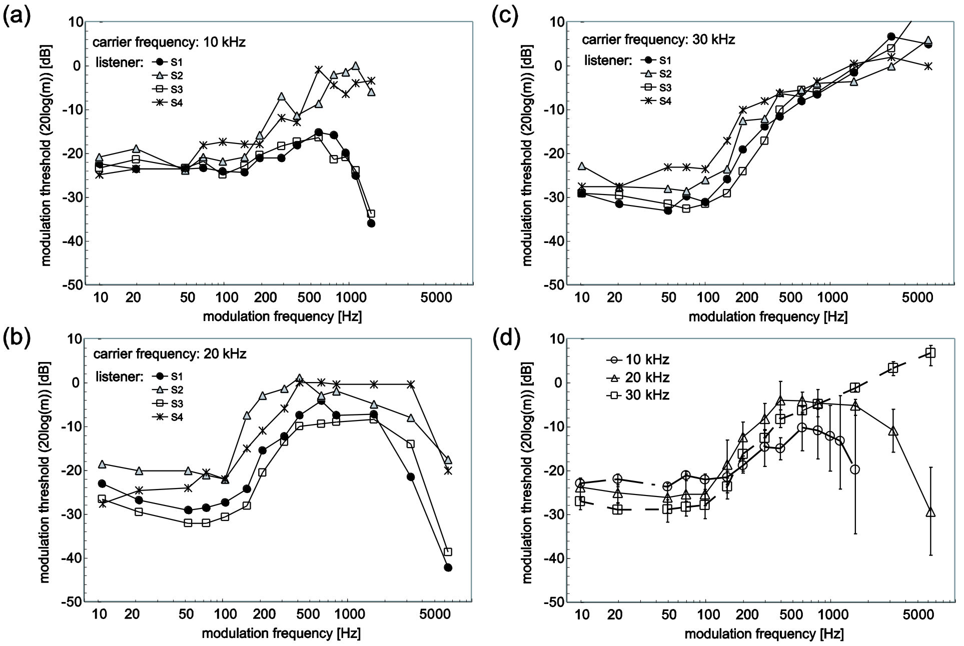

Fig. 1. Amplitude-modulation detection thresholds as a function of modulation frequency at a carrier frequency of (a) 10 kHz, (b) 20 kHz and (c) 30 kHz at different subjects; (d) averaged data across subjects for different carrier frequencies.

2.4 Subjects

One of the authors and three adult volunteers (30–45 years old) with normal hearing participated in the experiments. None of the subjects had any neurological or otological diseases in the ears that were stimulated.

3. RESULTS

The results of the SAM detection experiments at carrier frequencies of 10, 20, and 30 kHz for four subjects and their average are shown in Fig. 1. Detection thresholds are plotted as the just-detectable modulation index in units of 20 log m as a function of modulation frequency. Individual data are represented as means of two measurements. Averages across subjects are shown as means ± 1 standard deviation across subjects’ means.

For all carriers, the detection threshold curves clearly indicate the low-pass characteristics at low modulation frequencies. The detection thresholds were almost constant at modulation frequencies from 10 to about 100 Hz and then began to increase at about 100–150 Hz with increasing modulation frequencies. The threshold curves above the modulation frequency where this initial increase occurred varied between carrier frequencies.

At a carrier frequency of 10 kHz, the thresholds increased up to about 600 Hz with a slope of 8 dB/oct and then decreased with a slope of 18 dB/oct. In addition, the thresholds curve for the 20-kHz carrier had the initial increase at about 100–150 Hz up to 400 Hz with a slope of 10 dB/oct and did not depend on the modulation frequency in the middle-modulation frequency region between 400–1600 Hz. Above 1600 Hz, the threshold curve declined. For a 30-kHz carrier, the threshold curve showed a monotonic increase after the initial increase at about 100–150 Hz with slopes of 8 dB/oct at modulation frequencies between 100–400 Hz and 6 dB/oct above 400 Hz, respectively.

The shapes of the threshold curves for each individual roughly denoted the same tendency as that of the averaged data among all subjects with the exception of the 10-kHz carrier at high modulation frequency. Although the threshold curves for all subjects indicated the initial increase at about 100–150 Hz, different shapes were observed between subjects in the middle and high modulation frequencies ($150 < f_m < 3200$ Hz): the threshold curves of S1 and S3 had a peak at the modulation frequency of 600 Hz and those of the other subjects (S2 and S4) showed a continual increase with increasing modulation frequencies. Therefore, the deviations of the averaged data of the 10-kHz carrier are large at the high modulation frequencies.

4. DISCUSSION

4.1 Detectability of temporal modulation

Our results of the TMTFs obtained from sinusoidal carriers in the range from sonic high-frequency to ultrasonic frequency agree with the low-pass characteristics of the TMTFs in the region of low modulation frequency [11,12]. The cutoff frequency of the low-pass characteristics also agree with a previous study obtained with sinusoidal carriers [12].

The factors for having low-pass characteristics in the TMTFs at low modulation frequencies, where the modulation detection cue was the envelope fluctuation, have been thought to be a combination of the limitation of amplitude resolution and temporal resolution [12]. While the flat portion is explained as being the result of the limitation of resolution for the amplitude envelope, the initial increases in the threshold are due to the limitation of the temporal resolution, that is, limitation in resolving fast fluctuations of the amplitude envelope, rather than the result of peripheral filtering accounting for the agreement of cut-off modulation frequencies of low-pass characteristics of the tonal TMTF between carrier frequencies. In the current experiments, the cutoff frequencies of low-pass portions in the TMTF with 10-, 20- and 30-kHz carriers corresponded to data reported in the literature. Our results suggest the detectability of temporal modulation for the frequency region above the upper limit of the audible range and a common limitation between the temporal resolutions in processing amplitude envelopes in the sonic and the ultrasonic frequency range.

4.2 Detectability of sideband components

Generally, a TMTF with tonal carrier has decreases with increasing modulation frequency at high modulation frequencies. The modulation frequency at which the TMTF turns downward depends on the carrier frequency. Our results obtained for 10- and 20-kHz carriers also have these characteristics. This decrease in modulation detection thresholds is thought to be due to the detectability of the spectral component of sidebands [12]. In the following, we will discuss the detectability of the spectral components of sidebands at each carrier frequency used in our experiments.

For the 30-kHz carrier, the decrease in TMTF was not observed at high modulation frequencies up to 6400 Hz. One potential explanation for this phenomenon is that there are no other auditory filters in the range of ± 6400 Hz centered at 30 kHz adjacent to the auditory filter tuned to the 30-kHz carrier. Therefore, the sidebands could not be □heard out□ Regarding the auditory filters in the ultrasonic range, Buus et al. proposed three assumptions for explaining the sharp increase in hearing threshold of air-conducted sound above 14 kHz: (i) inefficient transmission of acoustic energy to the inner ear, (ii) decreasing sensitivity of auditory channel tuned to high frequencies, and (iii) running out of channels or the end of the cochlea [14]. Buus et al. concluded that the increase in threshold reflects the end of the auditory channel. Ashihara et al. also interpreted their results of a sharp increase between 14 to 20 kHz and a plateau above 20 kHz in the hearing threshold of air-conducted sound as representation of slope and shallower skirt, respectively, of the upper side of the end auditory channel□s tuning curve whose characteristic frequency is between 14□18 kHz. Our results with the 30-kHz carrier seem to be in line with □the end channel□ assumption: all components of SAM with the 30-kHz carrier were processed within the last auditory channel [15]. Moreover, based on this assumption, one of the specific characteristics of tonal BCU perceptions as a high-pitch tone independent of frequency [1-6] can be reasonably interpreted.

5. CONCLUSIONS

In this study, sinusoidally amplitude-modulation detection thresholds for sinusoidal carriers of 10, 20 and 30 kHz were measured. The results can be summarized as follows:

(i) For carriers of 10, 20 and 30 kHz, the TMTFs commonly low-pass characteristic in modulation frequency domain. These results are in line with common characteristics in temporal processing over the sonic range.

(ii) The TMTF with a 30-kHz carrier seem to reflect that sideband components for a 30-kHz carrier could not be detected. These results can be reasonably interpreted on the bases of an assumption of □the end of the auditory channels□

REFERENCES

[1] Gavreau, V. (1948). Compt. Rendu. 226, 2053□2054.

[2] Pumphrey, R. J. (1950). Nature 166, 571.

[3] Detherage, B. H., Jeffress, L. A., and Blodgett, H. C. (1954). J. Acoust. Soc. Am. 26, 582.

[4] Haeff, A. V. and Knox, C. (1963). Science 139, 590□592.

[5] Corso, J. F. (1963). J. Acoust. Soc. Am. 35, 1738□1743.

[6] Dieroff, H. G. and Ertel, H. (1975). Arch. Otorhinolaryngol. 209, 277□290.

[7] Lenhardt, M. L., Skellett, R., Wang, P., and Clarke, A. M. (1991). Science 253, 82□85.

[8] Hosoi, H., Imaizumi, S., Sakaguchi, T., Tonoike, M., and Murata, K. (1998). Lancet 351, 496□497.

[9] Nakagawa, S., Okamoto, Y., Fujisaka, Y. (2006). Trans. Jpn. Soc. Med. Bio. Eng., 44, 184-189.

[10] Dobie, R. A. and Wiederhold, M. L. (1992). Science 255, 1584□1585.

[11] Bacon, S. P. and Viemeister, N. F. (1985). Audiology 24, 117□134.

[12] Kohlrausch, A., Fassel, R., and Dau, T. (2000). J. Acoust. Soc. Am. 108, 723□734.

[13] Levitt, H. (1971). J. Acoust. Soc. Am. 49, 467□477.

[14] Buus, S., Florentine, M., and Mason, C. R. (1986). in Auditory Frequency Selectivity, 341□350. (Plenum, New York).

[15] Ashihara, K., Kurakata, K., Mizunami, T., and Matsushita, K. (2006). Acoust. Sci. & Tech. 27, 12□19.