Sensory unpleasantness of high-frequency sounds

Kenji Kurakata¹,*, Tazu Mizunami¹ and Kazuma Matsushita²

¹National Institute of Advanced Industrial Science and Technology (AIST), AIST Central 6, 1–1–1 Higashi, Tsukuba, 305–8566 Japan
²National Institute of Technology and Evaluation (NITE), 2–49–10, Nishihara, Shibuya-ku, Tokyo, 151–0066 Japan

(Received 5 March 2012, Accepted for publication 2 August 2012)

Abstract: The sensory unpleasantness of high-frequency sounds of 1 kHz and higher was investigated in psychoacoustic experiments in which young listeners with normal hearing participated. Sensory unpleasantness was defined as a perceptual impression of sounds and was differentiated from annoyance, which implies a subjective relation to the sound source. Listeners evaluated the degree of unpleasantness of high-frequency pure tones and narrow-band noise (NBN) by the magnitude estimation method. Estimates were analyzed in terms of the relationship with sharpness and loudness. Results of analyses revealed that the sensory unpleasantness of pure tones was a different auditory impression from sharpness; the unpleasantness was more level dependent but less frequency dependent than sharpness. Furthermore, the unpleasantness increased at a higher rate than loudness did as the sound pressure level (SPL) became higher. Equal-unpleasantness-level contours, which define the combinations of SPL and frequency of tone having the same degree of unpleasantness, were drawn to display the frequency dependence of unpleasantness more clearly. Unpleasantness of NBN was weaker than that of pure tones, although those sounds were expected to have the same loudness as pure tones. These findings can serve as a basis for evaluating the sound quality of machinery noise that includes strong discrete components at high frequencies.

Keywords: Sensory unpleasantness, High-frequency sound, Loudness, Sharpness, Equal-unpleasantness-level contour, Sound quality

PACS number: 43.66.Cb, 43.66.Jh, 43.66.Lj

1. INTRODUCTION

1.1. High-Frequency Sounds and Evaluation of Their Perceived Quality

Some electric and electronic products, such as information technology and telecommunications (ITT) equipment, emit sounds including salient high-frequency (HF) components of several thousand hertz. Other products reported to emit HF noise are electric home appliances, medical devices [1], automobiles [2], and railway trains including the Shinkansen [3]. Those sounds are often generated from built-in inverters or mechanical devices that rotate at a very high speed. Sources for some other HF sounds were unidentifiable [4].

 Those HF noises can elicit an unpleasant impression and are annoying for users of the product or equipment, or nearby people. Although such products generating HF sounds have become popular in daily life, methods for evaluating psychoacoustic effects of HF sounds have not yet been established.

For evaluation, the hearing threshold can be the first criterion [5,6]. If the sound pressure level (SPL) of a tone is higher than the threshold value, then the tone is judged to be audible for the population from which the threshold was obtained. However, comparison with the hearing threshold provides no useful information for evaluating the auditory impression of sounds above the threshold.

For tones accompanied by broadband noise, indices such as the tone-to-noise ratio (TNR) and prominence ratio (PR) are used. Measurement methods for these two indices are specified for ITT equipment in an international standard [5]. However, TNR and PR are not applicable properly, by definition, to tones without surrounding noise or when the noise is at a subthreshold level. The latter case can often occur at high frequencies because the hearing threshold increases as the frequency becomes higher than about 4 kHz. The level of noise is usually lower than the threshold. Furthermore, even when the
prominence of two tones is equal, their effects of sensory unpleasantness might differ depending on their frequencies. The relationship between those indices and sensory unpleasantness of discrete tones has not been fully investigated.

Loudness and sharpness [7] are sensory attributes of sounds that have been investigated widely for sound-quality evaluation and which might be related to the auditory impression of unpleasantness. It has been argued that loudness and sharpness are inversely related to sensory pleasantness [7]. In addition, the perceived quality of some machinery noise is degraded when sharpness is high [8]. However, having high loudness and sharpness is not always the cause of an unpleasant impression of the sound. For example, an appropriate increase in the amount of sharpness can give a sound the character of powerfulness [9], which is not a negative attribute of sounds. Consequently, acoustic correlates of sensory unpleasantness of HF sounds having a high loudness or sharpness value must be investigated further.

1.2. Objectives of This Study

This study was conducted to measure the sensory unpleasantness of HF sounds. The frequency, SPL, and bandwidth of stimulus sounds were varied systematically and their respective relationship with the calculated sharpness, loudness and other sound-quality metrics of those sounds were analyzed.

Two psychoacoustic experiments measured the sensory unpleasantness of HF sounds. In the first experiment, pure tones with various frequencies were used as stimuli. The effects of frequency and SPL on perceived unpleasantness were investigated. In the second, narrow-band noise (NBN) with various center frequencies and bandwidths was used. The unpleasantness of NBN was compared to that of pure tones to ascertain the effects of frequency and bandwidth.

In both experiments, young people aged about 20 years with normal hearing, who were expected to be most sensitive to HF sounds, participated as listeners.

Sensory unpleasantness in this study was defined as the sensation giving rise to the sentiment that “I would not want to hear the sound if I had to hear it continuously for a long period.” Any loud sound might be annoying if presented continuously. However, the degree of unpleasantness might differ depending on the acoustic characteristics of the sound.

Whether a sound is perceived as unpleasant or not depends on the subjective relation of the listener to the sound [7]. An unpleasant sound might not be as annoying as it could be if the listener is indifferent to the sound source (e.g., a person or a machine), and vice versa. To differentiate sensory unpleasantness from annoyance in this respect, the sound source was not specified in the experiments; the listeners were asked to concentrate on a judgment based on their perceptual impression.

2. EXPERIMENT I: SENSORY UNPLEASANTNESS OF PURE TONES

2.1. Measurement Method

Stimuli Pure tones with a frequency of 1, 2, 4, 5, 6.3, 8, 10, 12.5, 16, or 18 kHz were used. Table 1 shows that the SPL was changed by 5 or 10 dB in several steps. The duration of tones was 10 s with a rise/fall time of 1 s so that the transients at the beginning and end of tones would not elicit an unwanted impression of timbre change.

Apparatus The signals were calculated digitally on a personal computer with 24-bit resolution and 48 kHz sampling and were generated from a D/A converter (UA-5; Edirol, Roland Corp.). Then, they were fed into a control amplifier (SU-C1010; Technics, Matsushita Electric Industrial Co., Ltd.) and a power amplifier (SE-A1010; Technics, Matsushita Electric Industrial Co., Ltd.), and
were presented to the listener via a two-way concentric loudspeaker (i8; Tannoy Ltd.).

Measurement was conducted in an anechoic room of 4.35 m (W) × 6.00 m (D) × 2.95 m (H) at AIST. The loudspeaker and the listener’s chair were set 3.0 m apart from each other on the room’s diagonal axis. The listener sat on the chair with a headrest and faced the speaker directly. The listener’s head movement was monitored with two stick-type sensors attached to the headrest: movement during measurement was restricted to about 1 cm.

Procedure The magnitude estimation (ME) method without a modulus [10] was used to measure the unpleasantness that listeners perceived. The pure tone stimuli were presented to individual listeners one by one in random order. After listening to each tone, they estimated the magnitude of sensory unpleasantness and verbally reported a positive number that they assigned to the degree of impression. No limit was imposed to the range of numbers that listeners might use.

Listeners were given an explanation about the meaning of “unpleasantness” as described in Sect. 1.2. They were also instructed not to confuse unpleasantness with loudness because some sounds might be soft but unpleasant, although others might be loud but pleasant.

About 30 trials were run for practice before the main trials started. Each listener gave two estimations for every stimulus tone in the main session.

In addition to unpleasantness, listeners were asked to estimate the loudness of 1 kHz tones in a separate session using the same apparatus and procedure. Thereby, the relationship of unpleasantness with loudness was investigated.

Participants Study participants were 46 university students aged 19–24 years. Otoscopic examination, monaural audiometry using a pure-tone audiometer, tympanometry, and a questionnaire about difficulties in hearing [11] were administered to screen participants for otological abnormalities. Consequently, 38 students (18 men, 20 women) were selected as listeners. Their average age was 21 years. The average hearing threshold level (HTL) at 1, 2, 4, and 8 kHz of their better ear was −0.3 dB. Those listeners were regarded as being sampled from a young, otologically normal population.

2.2. Results and Discussion

2.2.1. Initial data processing

The correlation coefficient of the values assigned to the same stimuli in two trials was calculated to examine the reliability of listeners’ responses. One listener showed a markedly low coefficient (r = 0.24). Therefore, her responses were regarded as unreliable. After deleting her data, those of the remaining 37 listeners were adopted for the following analyses.

In the analyses, the median ME values of the listeners were used. The values were obtained in the following way. (1) The geometrical mean of the two estimates that the listeners gave for each stimulus tone was calculated for individual listeners. (2) The group median of the logarithm of the geometrical means was calculated for each stimulus level and frequency. The median was taken because some listeners failed to respond to the stimulus tones at the lowest levels and highest frequencies; the mean was not calculable for those tones. (3) For smoothing purposes, an INEX function [12] was fitted to the median as a function of SPL. The INEX function has been applied successfully to the loudness growth function of the 1 kHz tone, taking the place of the conventional power function [13]. Fitting was done separately for each tone frequency. (4) The 1 kHz, 40 dB tone was chosen as a reference tone and the ME value for the tone was adjusted to one. The INEX functions and ME values of other tones were shifted in proportion to it.

According to the two-stage theory of magnitude estimation, in which the first stage is sensory and the second is cognitive and involves processes of judgment, the unpleasantness growth function obtained in the present experiment can be affected by the second stage and may not express the sensory function directly. However, in the case of loudness growth functions, the effect of the second stage is known to be negligibly small (for review, see [10]). Therefore, the ME values given by the listeners in this study were assumed to reflect their sensory magnitude properly and were adopted as they were.

2.2.2. Relationship with sharpness

The unpleasantness for HF tones was regarded as resembling the auditory sensation of sharpness. To examine the relationship between those two auditory impressions, the sharpness of the stimulus was calculated using a software package (Type 7698 ver. 15.1.0; Brüel & Kjær). Among several methods for calculating sharpness, Aures’ method [14] was employed, which corrected Zwicker’s method [7] to give improved level dependence, so that the calculated sharpness can be compared with the present dataset measured for the wide SPL range.

Figure 1 shows median ME values of unpleasantness and calculated sharpness respectively, as a function of the SPL of stimulus tone, for stimulus frequencies of 1 kHz to 12.5 kHz. Data for 16 kHz and 18 kHz tones are not presented in the figure because the software package used did not provide reliable sharpness calculations for those extremely high tones.

The curves in the three panels illustrate clearly that, although unpleasantness and sharpness increased as the SPL became higher, they are two different sensory attributes. Unpleasantness increased more rapidly than sharpness did as the SPL of tone became higher. When SPL
increased from 20 dB to 80 dB, unpleasantness increased by a factor of a few tens, or one hundred or more for some frequencies, although sharpness increased only by a factor less than 2. In contrast to that, the effect of tone frequency on unpleasantness was much smaller than on sharpness. For example, unpleasantness of the 12.5 kHz tone was only slightly larger than that of the 1 kHz tone at 40 dB, whereas sharpness of those two tones differed by a factor of 10. Consequently, sensory unpleasantness is not determined solely by sharpness; unpleasantness is more level dependent, but less frequency dependent than sharpness.

2.2.3. Relationship with loudness

Although sharpness is not closely related to unpleasantness, loudness might be because loud sounds are annoying in general. The relationship between sensory unpleasantness and loudness was investigated by estimating the loudness as described below. (1) An INEX function [12] was fitted to the group medians of ME values that the listeners assigned to the stimuli in the separate session for loudness measurement. (2) The MEs of loudness at frequencies other than 1 kHz were obtained using the equal-loudness-level relation described in an ISO standard [15]. Tones having a different frequency but the same loudness level as a 1 kHz tone were given the same ME value as assigned to the 1 kHz tone. Data for 16 kHz and 18 kHz tones were not adopted here either because the international standard does not cover those frequencies.

Figure 2 presents the unpleasantness ME’s as a function of thus-estimated loudness ME. For the ME’s for each frequency, the power function expressed below was separately fitted:
or

$$\log U = \log a + b \cdot \log L,$$

where $U$ is the unpleasantness ME, $L$ is the loudness ME, and $a$ and $b$ are, respectively, a constant and the exponent of the power function that varies depending on the tone frequency.

It can be seen that the functions ran close to the diagonal line overall, indicating that unpleasantness increased proportionally with increasing loudness. Closer observation reveals that the increase of unpleasantness was greater than that of loudness, as is evident from the slope of the function. Furthermore, the slope differed depending on the tone frequency; higher-frequency tones generally showed a steeper slope. The 95% confidence interval of $b$, the power exponent or the slope of the function, was from 0.94 to 1.09 for 1 kHz tones. On the other hand, that for 12.5 kHz tones was from 1.09 to 1.38, which was significantly larger than one. The tones with those frequencies elicited greater unpleasantness than those with other frequencies even when their loudness was equivalent.

2.2.4. Derivation of equal-unpleasantness-level contours

The frequency dependence and level dependence of sensory unpleasantness and the relation to loudness can be made even clearer by drawing equal-unpleasantness-level contours (EUPLCs). Drawing EUPLCs has the advantage that the cognitive factors associated with the ME method that were described in Sect. 2.2.1 can be canceled out if we assume that those factors affected the listeners’ judgments in a similar manner regardless of stimulus frequencies; the contours represent the relative, not absolute, degree of unpleasantness of tones as a function of frequency.

In deriving the contours, another assumption that tones to which the same number had been assigned by listeners in the experiment had the same unpleasantness irrespective of the tone frequency was made. Combinations of frequency and SPL were identified using the INEX functions of unpleasantness and loudness presented in the previous sections.

Figure 3 shows some EUPLCs derived in this manner. The equal-loudness-level contours (ELLCs) [15] were superimposed to compare the two families of contours. The EUPLCs are lower than the ELLCs in general, suggesting that pure tones with a frequency higher than
1 kHz can elicit a stronger than expected sensation of unpleasantness from the equal-loudness relation. That is, HF tones can be more unpleasant even when their loudness is equal to that of lower frequency tones.

The differences between EUPLCs and ELLCs are largest at 4 kHz and around 10 kHz, as was demonstrated by the analysis of power functions described in Sect. 2.2.3. The tones at and around those two frequencies can yield a high degree of unpleasantness even when their loudness is low.

ISO 226 does not specify equal loudness levels at frequencies higher than 12.5 kHz, but two studies [17,18] that were used for deriving the ISO contours measured those of 20 phons and 40 phons at 16 kHz. Average equal-loudness-level values of those studies are presented in Fig. 3. Apparently, equal-unpleasantness levels and equal-loudness levels do not differ much at that frequency.

3. EXPERIMENT II: SENSORY UNPLEASANTNESS OF NARROW-BAND NOISE

3.1. Measurement Method

Stimuli NBN with a center frequency of 1, 4, 8, 10, 12.5, or 16 kHz was used. The bandwidth was 1/3, 1/6, 1/12, or 1/24 octave. The SPL was kept constant at 40 dB, except for the 16 kHz noise, whose level was 60 dB. In addition to the noises, pure tones with the same frequency and SPL were also used as reference sounds. The sound duration was 10 s with a rise/fall time of 1 s so that the transients at the beginning and end of sounds would not elicit an unwanted impression of timbre change.

Apparatus The signals were calculated digitally on a personal computer with 24-bit resolution and 96 kHz sampling. They were presented to listeners using apparatus identical to that used in Experiment I.

Procedure The procedure for the NBN measurement was fundamentally identical to that in Experiment I for pure tones: the ME method without a modulus was used; listeners estimated the magnitude of unpleasantness of noises and pure tones presented one by one in random order; then, they reported a positive number that they assigned to the degree of impression; every stimulus sound was presented twice for each listener.

The listeners were given the same instruction about the meaning of “unpleasantness” and its difference from loudness, as in Experiment I.

After a short training session, each listener gave two estimations for every stimulus sound in the main session.

Participants Study participants were the 29 university students (13 men, 16 women) who participated in Experiment I and who were screened for hearing abnormalities. Their average age was 21 years. The average HTL at 1, 2, 4, and 8 kHz of their better ear was −0.8 dB, which was normal for their age.

3.2. Results and Discussion

3.2.1. Initial data processing

One listener showed a markedly low correlation coefficient (r = 0.26) between the two trials for the same stimulus. Therefore, his responses were regarded as unreliable. After deleting his data, those obtained from the remaining 28 listeners were adopted for the following analyses.

Median ME values of the listener group were calculated similarly, as conducted in Experiment I. (1) The geometric mean of the two estimates that the listeners gave for each stimulus was calculated for each listener; listeners estimated the magnitude of unpleasantness of noises and pure tones presented one by one in random order; then, they reported a positive number that they assigned to the degree of impression; every stimulus sound was presented twice for each listener.

The listeners were given the same instruction about the meaning of “unpleasantness” and its difference from loudness, as in Experiment I.

After a short training session, each listener gave two estimations for every stimulus sound in the main session.

Participants Study participants were the 29 university students (13 men, 16 women) who participated in Experiment I and who were screened for hearing abnormalities. Their average age was 21 years. The average HTL at 1, 2, 4, and 8 kHz of their better ear was −0.8 dB, which was normal for their age.
statistically significant (the $\chi^2$ statistic varied from 24.8 for 12.5kHz sounds to 51.1 for 4kHz sounds; $df = 4$, $p < 0.01$). The difference was not significant for the 1kHz sounds ($\chi^2 = 3.31$, $df = 4$, $p > 0.50$) and 16kHz sounds ($\chi^2 = 3.14$, $df = 4$, $p > 0.50$).

The insensitivity of 16kHz sounds to the bandwidth change is explainable as follows. At around 16kHz, loudness decreases rapidly as the frequency increases, as expected from the rapid increase of the hearing threshold (see Fig. 3). Consequently, the energy at the lower-frequency edge of the noise band is dominant, whereas that at higher frequencies becomes less effective in perception, which yields a pure-tone-like impression of sound; the NBN might have been perceived as a pure tone by listeners.

3.2.3. Relationship with other sound-quality metrics

Although not portrayed in Fig. 4, Aures’ sharpness of NBN changed very little when the bandwidth was widened from 1/24 octave to 1/3 octave and could not explain the experimental results. The values for NBN were almost equal to that for the pure tone of the same frequency, according to the calculations using a software package (Type 7698 ver. 15.1.0; Brüel & Kjaer), which was expected from the derivation procedure of sharpness [7].

Other popular sound-quality metrics, roughness and fluctuation strength, are related to the magnitude of the temporal variation of sounds. NBN has a time-varying envelope and that variation can elicit the sensory impressions of roughness and fluctuation strength [7]. On the other hand, pure tones, whose temporal envelope is constant over time, do not yield those sensations.

Those two sensory impressions can cause the degradation of sound quality, as was demonstrated by the studies on the prediction of perceived annoyance on the basis of the composed metrics including roughness and fluctuation strength [9]. If the unpleasantness of HF sounds were associated with roughness and fluctuation strength, NBN would produce stronger unpleasantness than would pure tones. However, the experimentally obtained results presented in Fig. 4(a) were quite opposite to that inference.

The reduction of unpleasantness of NBN might be explainable as follows: the NBN did elicit the sensation of roughness and fluctuation strength to some extent, which could lead to the increase of unpleasantness. However, their broader spectrum than that of pure tones had an effect of lessening the penetrating impression that HF sounds could evoke, and decreased the unpleasantness. The latter effect was greater than the former one of roughness and, consequently, the overall impression of sensory unpleasantness of NBN became weaker than that for pure tones.

To summarize, the variation of sensory unpleasantness among pure tones and NBN with a different bandwidth were not related to the difference in loudness and sharpness. Furthermore, producing the sensation of roughness or fluctuation strength did not lead to the increase of unpleasantness. The metrics that have often been used in sound-quality analyses cannot be used to evaluate HF noise unpleasantness.

Sensory unpleasantness might be more closely related to pitch strength, which can be labeled as faint pitch or strong pitch on a scale. The pitch strength of NBN has been shown to be smaller than that of pure tones [19]. It is imaginable that the weaker pitch sensation could lessen the penetrating impression that HF sounds could evoke, which resulted in mitigating the unpleasantness of NBN.

Unfortunately, models for predicting the pitch strength of sounds have not yet been established [20]. Neither has the pitch strength of HF sounds been fully investigated. Therefore, it is worth investigating the relation of unpleasantness with pitch strength so that the effect of spectral contents on unpleasantness will be clarified.

4. SUMMARY

Measurement results of sensory unpleasantness of high-frequency sounds in the two experiments can be summarized as described below.

(1) Sharpness (Aures’ sharpness) was not closely related to the unpleasantness of pure tones. The unpleasantness was more level dependent and less frequency dependent than sharpness.

(2) Pure tone unpleasantness increased as loudness increased. The growth rate of unpleasantness was higher for high-frequency tones.

(3) Comparison between equal-unpleasantness-level contours and equal-loudness-level contours revealed that pure tones at 1kHz and higher elicit a stronger impression of unpleasantness than expected from their loudness. The two families of contours were most different at 4kHz and around 10kHz.

(4) Unpleasantness of narrow-band noise was less than that of pure tones having the same center frequency and sound pressure level as the noise had; nevertheless, those sounds did not differ in loudness.

(5) The unpleasantness of narrow-band noise did not change greatly when their bandwidth varied from 1/24 octave to 1/3 octave.

(6) The difference in unpleasantness between pure tones and narrow-band noise was not explainable in terms of loudness or sharpness. Furthermore, other sound-quality metrics, roughness and fluctuation strength, were not directly related to the unpleasantness. Instead, pitch strength must be investigated further to ascertain a perceptual correlate with the effect of bandwidth on unpleasantness.
ACKNOWLEDGEMENTS

The authors are grateful to the editor and two anonymous reviewers for their helpful comments on an earlier version of the manuscript.

REFERENCES


