Comparison of equal-loudness-level contours between otologically normal young and older adults

Kenji Kurakata* and Tazu Mizunami

National Institute of Advanced Industrial Science and Technology (AIST), AIST Tsukuba Central 6, 1–1–1 Higashi, Tsukuba, 305–8566 Japan

(Received 20 September 2013, Accepted for publication 21 March 2014)

Abstract: To investigate the effects of aging on loudness perception, equal-loudness-level contours were derived from loudness estimates for pure tones with various combinations of frequency and sound pressure level. The listeners were young and older adults who had otologically normal hearing. Comparison of the contours between listener groups revealed a large difference at frequencies higher than 1,000 Hz, indicating that substantial effects of age and gender existed: older listeners were less sensitive to high-frequency tones than young listeners. In addition, older males were even less sensitive to those tones than older females. Recruitment of loudness was also observable for older listeners at high frequencies. These findings suggest that the use of frequency-weighting A, which is based on the hearing characteristics of young people, in noise evaluation is questionable when senior citizens are assumed to be the listeners. Another method that takes the effects of aging into account should be developed to estimate the loudness of sounds perceived by older listeners more accurately.

Keywords: Equal-loudness-level contours, Loudness perception, Aging effect, Gender effect, Magnitude estimation method, ISO 226

PACS number: 43.66.Cb, 43.66.Sr, 43.15.+s [doi:10.1250/ast.35.243]

1. INTRODUCTION

1.1. Loudness Perception of Older Adults

Equal-loudness-level contours (ELLCs) indicate combinations of the frequency and sound pressure level of pure tones that are perceived as equally loud to listeners under specified listening conditions. The contours have been used widely in acoustics and hearing sciences for a long time since they show fundamental characteristics of our auditory system, whose response to input sounds changes with their frequency and intensity.

Several attempts have been made to derive ELLCs under various measurement conditions (see [1] for a review). Among them, the contours of an ISO standard [2] have been accepted internationally as normative ones. The contours were obtained for pure tones of frontal incidence in a free field under binaural listening conditions. The listeners were otologically normal persons aged from 18 to 25 inclusive.

As humans become older, their sense of hearing changes progressively in various aspects. A well-known consequence of aging is the elevation of the hearing threshold [3]. People become less sensitive to sounds with age. The extent of degradation is larger for high-frequency tones than for low-frequency ones.

Aging can also affect the attributes of auditory sensation for sounds above threshold levels. Loudness is one such attribute. However, few studies have been made to examine loudness perception in presbycusis listeners [4], except for some studies motivated by clinical interest [5].

It is imaginable that older adults show different ELLCs than those obtained from young adults even if their hearing ability is normal for their age. The study of Robinson and Dadson [6] seems to so far be the only one that attempted to construct a set of ELLCs for older listeners [1].

One reason for the scarcity of loudness-perception research employing older listeners is attributable to larger individual differences in their hearing ability. A large number of listeners must be examined to derive a conclusion from measurement data. The measurement is thus more costly and time-consuming than that for young listeners.

Another reason is that the decline of hearing ability is caused not only by aging, but also by other factors such as infections, ototoxic drugs, noise exposure, and familial hearing disorders. Usually, these factors are not separable...
from the natural, biological ones. This makes it difficult to employ eligible listeners and to reveal the pure effects of aging.

1.2. Use of the Magnitude Estimation Method to Derive ELLCs

The method of magnitude estimation (ME) has been frequently used to measure sensory magnitudes including tone loudness. This method is efficient in terms of time for collecting a large amount of data. In most cases, only a few responses from each subject are necessary to estimate a value on a sensory magnitude continuum.

Although the ME method has been criticized as being subject to various response biases [7], the method might be useful for obtaining ELLCs for numerous listeners in a short time. In fact, Schneider et al. [8] examined its applicability to loudness-perception experiments and derived ELLCs for young listeners successfully.

The first author of the present paper has reported ELLCs measured for young and older adults using the ME method in his preliminary study [9]. He showed that the ME method could produce a set of ELLCs for young listeners and that the contours were comparable with those in the international standard [2] measured using other methods. Also, he demonstrated that ELLCs could be obtained from older listeners.

1.3. Aim of This Study

This paper is primarily intended to construct ELLCs for young and older listeners and examine the validity of the contours reported in the preliminary study [9]. Then, the contours are analyzed in terms of the effects of aging and gender.

The method and conditions of measurement were essentially identical to those in the previous study. However, the shortcomings of the previous study were eliminated or their adverse effects were reduced as much as possible, as described below.

First, the numbers of young and older listeners were increased to obtain more reliable contours; more than twice as many listeners as before were employed. Secondly, the age range of older listeners was restricted to 60–69 years, whereas the upper limit in the previous study was 86 years. As a result, the variation of responses among older listeners was substantially reduced. Thirdly, the effects of gender were investigated for older listeners. The number of listeners was too small in the previous study to investigate the effects. Lastly, the listeners were more rigorously screened for abnormalities in their hearing to minimize the effects of factors other than aging.

2. MEASUREMENT

2.1. Stimuli

Pure tones were used as stimuli. The frequency was 125, 250, 500, 1,000, 2,000, 4,000, or 8,000 Hz. The duration was 1 s with a rise/fall time of 50 ms.

The sound pressure level was varied from the sensation level (SL) of 5 dB to 80 dB in 5 dB steps to cover a wide range of loudness. The normative threshold in a free field [2] was taken as the reference of SL (i.e., 0 dB) at each frequency. Table 1 shows the sound pressure levels of the stimuli used.

The lowest-SL tones of 5 dB were presented to about half of the listeners while the highest ones of 65–80 dB (the level depended on the frequency) were presented to the other half of the listeners. For example, listeners in the former group heard 1,000 Hz tone stimuli at SLs of 5–70 dB and those in the latter group heard them at SLs of 10–75 dB. In this way, a wide range of sound levels was covered while reducing the measurement time per listener.

2.2. Apparatus

The pure tones were generated digitally using a personal computer at a sampling frequency of 44.1 kHz with 16-bit resolution and recorded on a compact disc (CD). They were played back by a CD player (DCD-755AE; Denon), fed into an equalizer (DEQ2496; Behringer), amplified by a power amplifier (AVC-1508, Denon), and presented to listeners in an anechoic room (W: 4.35 m, D: 6.00 m, H: 2.95 m) at AIST via a loudspeaker.
(i8; Tannoy). The loudspeaker was set on the room’s diagonal axis at a distance of 3 m from the reference point, which was the midpoint between the listener’s two ears. The characteristics of the output signals and the acoustic environment generally met the requirements of the preferred test conditions for ELLC determination [10].

2.3. Procedure

Listeners took part in the measurement individually after providing informed consent. They sat on a chair with a headrest and faced the loudspeaker directly. The position of the listener’s head was carefully adjusted to the reference point before the measurement. Listeners were asked to keep still during the measurement, with their head on the headrest.

The ME method without a modulus was used to measure the loudness of pure-tone stimuli, which were presented to listeners one by one in random order. After each presentation, the listeners assigned a positive number to the stimulus that they thought corresponded to the magnitude of loudness, which they communicated orally. After a short session for practice, every listener performed two trials for each stimulus. The measurement time per person was less than 1 h, including rest periods.

2.4. Participants

Two age groups, young and older adults, participated in the measurement. The young participants were 59 university students aged 18–25 years. The older participants were 71 men and women aged 60–69 years, who were introduced by a local employment agency.

All the participants underwent screening involving otoscopic examination, monaural audiometry with a pure-tone audiometer, tympanometry, and inquiry about difficulties in hearing [11].

3. RESULTS AND DISCUSSION

3.1. Screening of Participants

Participants were screened for the following otological abnormalities that were measured or reported: (1) a middle ear pressure out of the range of \( \pm 50 \text{ daPa} \), (2) extensive exposure to loud sounds at work and in daily life, (3) a history of severe ear disease or injury, (4) familial hearing loss, (5) self-reporting of unusual hearing on the day of measurement, or (6) an unbalanced threshold between left and right ears at any of the audiometric frequencies between 0.125–8 kHz, the criterion being a threshold level difference of 30 dB or more.

Data taken from the participants after screening were adopted for the analyses described below. Table 2 lists the ages and numbers of listeners.

An audiogram of the listeners after screening is presented in Fig. 1. Although the listeners were not screened on the basis of their hearing threshold levels per se, they were regarded as having normal hearing ability for their age, as indicated by their low median thresholds.

3.2. Derivation of ELLCs

The geometric mean of the two numbers assigned to each stimulus was first calculated to obtain a magnitude estimate for individual listeners. Then, group medians for young and older listeners were calculated for each stimulus. Data of males and females in the older listener group were treated separately to examine the effects of gender.

Medians, not means, were used in the calculation because stimuli at the lowest SLs were not audible for...
some listeners and, therefore, means were not calculable for those tones. A median was calculable for stimuli to which more than half of the listeners could respond.

A loudness growth function [12] was fitted to the thus-calculated magnitude estimates of listener groups for each frequency tone, expressed as

$$\Psi = k(P_2^n - P_0^n).$$  \hspace{1cm} (1)

This function has the same form as that used for deriving the ELLCs in ISO 226 [2]. In this equation, $\Psi$ is the group median of the loudness estimates, $P$ is the sound pressure of the stimulus tone, and $k$ and $n$ determine the form of the function. $P_0$ is the sound pressure of the stimulus tone at the threshold, which was estimated using the audiometric data for each age and gender group.

The number of listeners was taken into account when the fitting was carried out (MATLAB R2013a; The MathWorks, Inc.) since the number differed with the stimulus level.

Figure 2 depicts the loudness functions of selected frequencies for three listener groups. The power law function, Eq. (1), approximates the loudness estimates of both young and older listeners well on the whole, as has been repeatedly observed for young adults [4]. However, a closer look at the results reveals that the function deviated from the data points at very low SLs in some cases. It might have been difficult even for young listeners to judge the loudness of tones at levels near the threshold. Therefore, the ELLCs at an SL of about 10 dB or below should be treated carefully in the following sections since the contours are subject to a large error.

### 3.3. ELLCs of Young Listeners

An assumption was made to derive ELLCs from the magnitude estimates of loudness: tones to which the same number was assigned by listeners had the same loudness, irrespective of their frequency. Once the loudness growth functions at 1,000 Hz and other frequencies are obtained, as shown in Fig. 2, ELLCs can be drawn by connecting the sound pressure levels of tones that have a different frequency but the same magnitude of loudness.

According to the two-stage theory of ME [7], in which the first stage is sensory and the second is cognitive and involves processes of judgment, the loudness functions obtained in the present experiment can be affected by the second stage. Therefore, they may not express the sensory magnitude directly.

However, the cognitive factors associated with the ME method can be canceled out if we assume that these factors affected the listeners’ judgments in the same manner regardless of the stimulus level and frequency. The contour represents a relative, not absolute, magnitude of loudness. This assumption produces an advantage of drawing ELLCs over investigating loudness functions, such as those shown in Fig. 2, as they are.

Figure 3 shows ELLCs of the young listener group obtained in the measurement. Also drawn for comparison are two sets of ELLCs, one specified in ISO 226 and the other derived using a loudness calculation procedure in an ANSI standard [13]. The contours in the current study generally show good agreement with the standardized ones. At moderate frequencies and loudness levels, the data points obtained from our measurement lie between the standard contours or close to either of them. The size of the discrepancy was only a few decibels at maximum.

Relatively large deviations can be found at the highest loudness level of 70 phons. The deviation might have been caused by response biases because the judgments could be variable for these stimuli under extreme conditions. It is noteworthy here that the variation among the study results adopted to establish the ISO standard was as large as 10 dB even under moderate stimulus conditions [2]. Therefore, the size of the deviation in Fig. 3 does not invalidate the overall results at all.

Thus, it was confirmed that the ME method can produce ELLCs with comparable precision to that of other psycho-physical methods such as the method of constant stimuli and the bracketing method, which were used to construct the
ISO ELLCs. In addition, the method requires much less time than the other methods. The reproducibility of ELLCs by the ME method will be discussed later in Sect. 4.

3.4. Effects of Gender on ELLCs of Older Listeners

It is widely known that the gender has a large effect on the hearing threshold of older adults; men have a higher threshold on average than women at high frequencies [3]. Therefore, they are also expected to show a large difference in loudness perception and the effects of gender might be observable in their ELLCs. Robinson and Dadson [6], who measured the ELLCs of older listeners, did not report the effects of gender.

Figure 4 displays a comparison of the ELLCs obtained from the older male and female groups. As expected from the effects of gender on the hearing threshold, the ELLCs of male listeners showed a steeper rise at high frequencies than those of female listeners. This result indicates that males were also less sensitive to tones at high frequencies at supra-threshold levels than females.

For frequencies lower than 1,000 Hz, the contours for female listeners are flatter than those for male listeners. However, the difference was relatively small.

3.5. Comparison of ELLCs of Young and Older Listeners

Although large effects of gender were observed in the previous section, comparing the ELLCs of young and older listeners regardless of the gender is of interest for investigating the effects of aging.

The two sets of ELLCs obtained from male and female listeners separately were averaged to obtain the general characteristics of their loudness perception. Figure 5 illustrates the thus-obtained ELLCs of older listeners and those of young listeners in Fig. 2.

Some characteristics of aging are notable in the figure. First, the contours of the older listener group were higher than those of the young listener group at high frequencies of above 1,000 Hz. This result indicates that older listeners are less sensitive to tones at higher frequencies than young listeners, as is expected from the threshold elevation of the former listeners.

Second, the difference in ELLCs between the two age groups became smaller at 4,000 and 8,000 Hz as the loudness level increased. This result supports the argument that loudness recruitment occurs in a sizable proportion of people with presbycusis [6,14]. The existence of recruitment is also evidenced by the steeper loudness functions of older listener groups in Fig. 2(b).

Third, the difference in ELLCs between the two age groups was relatively small at frequencies lower than 1,000 Hz, indicating little effect of aging at lower frequencies.

4. GENERAL DISCUSSION

4.1. Reproducibility of ME Measurement

Although the ME method is efficient for collecting responses from a large group of listeners, the responses
of an individual listener can vary from trial to trial. This makes it difficult to construct ELLCs for individual listeners with high accuracy. Therefore, in this study, only a single set of ELLCs was derived from the median responses of each listener group.

To evaluate the accuracy of the measurement results, the ELLCs of young listeners obtained in the present study were compared with those in the author’s previous study [9]. Since the listeners in the two studies shared common attributes such as the age range and hearing ability, they can be regarded as samples from the same population. They were expected to produce essentially the same contours. Figure 6 shows the results of the comparison.

The two sets of contours agreed well in general except for the loudness level of 10 phons, at which the tone produced only a faint sensation and the judgment of loudness could be ambiguous. For moderate frequencies and loudness levels, the size of the discrepancy was a few decibels. This can be considered as a rough estimate of the size of variation when the ELLC measurement is repeated under the same conditions.

If this estimate of variation is taken into account, the measurement results in Sect. 3 can be reevaluated as follows:

1. The deviation of ELLCs of young listeners from those based on standards is within the range of variation of repeated measurements (Fig. 3).
2. The effects of gender observed in the ELLCs for older listeners are large at high frequencies. Those at low frequencies are close to the size of variation (Fig. 4).
3. The effects of aging observed in the ELLCs are large at high frequencies. Those at low frequencies are within the range of variation (Fig. 5).

### 4.2. Comparison of ELLCs of Older Listeners with Those in Other Studies

The listeners in the preliminary study [9] are regarded as samples from a different population than that in the present study in terms of their age range and hearing ability, as described in Sect. 1. Robinson and Dadson [6] estimated ELLCs for people aged up to 60 years. The participants were screened for abnormalities in hearing and are thus comparable with those in the present study.

Although there are some differences in the listeners’ attributes and the experimental procedure, it is interesting to compare these three sets of ELLCs to examine the effects of aging from a wider perspective. Figure 7 shows a comparison of these three measurements.

The listeners in the preliminary study were older on average (median, 69.5 years; age range, 61–86 years) than those in the present study. The effects of aging are manifested in the higher threshold of the former group. On the other hand, the threshold of Robinson and Dadson’s study was lower than those in the other two studies. The curves were estimated for listeners up to an age of 60.

More importantly, these three sets of ELLCs exhibit a systematic change in shape, as shown by the vertical arrows in the figure. As the age of the listeners increases, the contours generally become higher at high frequencies and lower at low frequencies. (In Fig. 7, the curves of
Robinson and Dadson’s study are drawn for frequencies of above 1,000 Hz only because “the effects of age were found to be negligible below and including 1,000 c/s” [6] and the validity of curves at lower frequencies has been questioned [1].

The lowering of contours below 1,000 Hz does not mean that the older group could hear the tones of those frequencies better than the younger group since an ELLC represents relative, not absolute, sensitivity to a 1,000 Hz tone. Rather, the curves in Fig. 7 indicate that the relative sensitivity to low-, middle-, and high-frequency tones changes with increasing age. The change in the balance of sensitivity is larger for older people.

### 4.3. Implications of the Effects of Aging for Noise Evaluation

The frequency-weighting A is used when the level of sound is measured with a sound level meter [15]. The weighting curve was determined on the basis of loudness-level contours by Fletcher and Munson [16]. Although this weighting does not aim to accurately simulate the frequency characteristic of our auditory system, the time-averaged, A-weighted sound level ($L_{Aeq}$) of a sound corresponds well with the perceived loudness [17].

Thus, $L_{Aeq}$ could be a good measure to predict the loudness of sound for young people but not for older people. Consequently, the loudness of a sound perceived by older ears will be overestimated when the sound has relatively large energy at high frequencies and is evaluated in terms of the $L_{Aeq}$ value.

The same argument can be applied to loudness calculation methods [13,18] that have been developed on the basis of the hearing characteristics of young people without taking the effects of aging into consideration. These methods should be modified appropriately if they are to be used to estimate the loudness of sounds perceived by older people. This modification should be carried out as future work.

### 5. CONCLUSION

In this study, we measured the loudness of tones with various frequencies and sound pressure levels as perceived by otologically normal young and older adults. On the basis of their estimates of loudness, a set of ELLCs was derived for each age and gender group.

The results of the measurement can be summarized as follows:

1. The ME method could produce ELLCs comparable with the normative ones in standards obtained using other measurement methods.
2. The ELLCs of older adults showed an effect of gender. The difference in contours was large at high
frequencies since older males had decreased sensitivity to tones with high frequencies.

3) The effects of aging were also large at high frequencies since older listeners were less sensitive to tones with high frequencies than young listeners. The effects became larger with increasing age.

4) As the loudness level increased, the difference in the ELLCs between the two age groups became smaller at 4,000 and 8,000 Hz. This result evidences the existence of loudness recruitment in presbycusis.

5) The deviation of the 40-phon ELLC from the A-weighting curve was larger for older listeners than for young listeners at high frequencies. A new measure other than the A-weighted sound pressure level should be developed to evaluate noise for senior citizens.

This study employed only otologically normal persons to investigate age-related changes in loudness perception. However, for the purpose of revealing the loudness perception of the general population, it would be interesting to conduct a similar measurement, employing people who represent the population regardless of their hearing ability.

ACKNOWLEDGMENTS

The authors are grateful to the editor and the two anonymous reviewers for their helpful comments on an earlier version of the manuscript. This work was supported in part by JSPS KAKENHI Grant Number 24653219.

REFERENCES


Kenji Kurakata received his Ph.D. degree from Osaka University in psychoacoustics in 1994 and is currently working for the National Institute of Advanced Industrial Science and Technology (AIST), Japan, as Leader of the Accessible-Design Research Group. His research interests include the effects of aging on hearing and human-interface design for older persons and persons with disabilities. He has been engaged in international standardization work in ISO/TC 43 “Acoustics,” ISO/TC 159 “Ergonomics,” and other committees on acoustics and accessibility in ISO and IEC. http://staff.aist.go.jp/kurakata-k/index-e.html

Tazu Mizunami received her Ph.D. degree from Osaka University in psychoacoustics in 2001 and is currently working for AIST, Japan, as a technical staff member of the Accessible-Design Research Group. She has been involved in research of psychoacoustics, the effects of noise on humans, music psychology, and human-interface design using sound signals and voices. She is a member of the Acoustical Society of Japan and other academic societies on acoustics and psychology.