Estimates of Cochlear Compression from Measurements of Loudness Growth

Stephen T. Neely, Kim S. Schairer, and Walt Jesteadt

Boys Town National Research Hospital, neely@boystown.org

1 Introduction

The growth of loudness can be measured by matching the loudness of a pure tone to that of a multi-tone complex. This method was originally used by Fletcher and Munson (1933) and has been used more recently by Buus et al. (1998). If the tones in the complex are equally loud and separated in frequency far enough that they do not mask each other, then the ratio of the loudness of the complex to the loudness of each component will be equal to the number of components. This empirical fact can be used to construct a relative loudness function without any additional assumptions about the form of this function. We have obtained loudness growth measurements using four-tone complexes with one-octave separation between equally-loud components. We assume that these four-tone complexes are four times as loud as any of their components. We consider loudness to be an intensity-like variable, so that a four-fold increase in loudness is equivalent to 6 dB. Therefore, we estimate cochlear compression by determining the number of dB (of physical intensity) required to achieve a four-fold increase in loudness and dividing that number by the corresponding 6-dB increase (in perceptual intensity).

2 Methods

Six adults with normal hearing served as subjects. Stimuli were 500-ms duration tones with cosine-squared envelopes (20-ms rise/fall) generated digitally (TDT AP2) at 0.5, 1, 2, and 4 kHz. Data were collected in three stages. (1) Quiet thresholds were obtained for the four individual tones. (2) The loudness of tones at 0.5, 2, and 4 kHz was matched to the loudness of a 1-kHz tone at 10 dB SL. (3) These four equally-loud tones were combined as a “chime” and the loudness of each component tone was matched to the loudness of this chime.

Threshold in quiet for each tone was estimated using a one-track paradigm, a two-interval forced-choice (2IFC) procedure and a 2-down, 1-up decision rule with 4-dB step size to estimate the 71% correct point on the psychometric function (PF).

The mean threshold across four repetitions was calculated for each frequency.

In the next stage, the 0.5-, 2-, and 4-kHz tones were individually matched in loudness to a 1-kHz tone at 10 dB SL using a 2IFC, six-track paradigm (two tracks for each of the three matching tones). The 1-kHz tone was presented randomly in one interval of each trial, and one of the other individual tones was presented in the other interval. The subject's task was to identify the interval that contained the louder sound, regardless of pitch. There were 300 trials in each repetition (50 trials per track), and two repetitions were completed for each subject. One of the two tracks per matching tone estimated the 71% point on the PF (descending track); the other track estimated the 29% point on the PF (ascending track). The matching tone level started at 30 dB SPL for the 71% track and at threshold for the 29% track. The step size was initially 8 dB and was reduced to 2 dB after the fourth reversal. The matched level for each track was calculated as the mean of reversal points after the fourth reversal. The level at which each matching tone was judged to be equally loud to the 1-kHz tone was estimated from the mean of the 29% and 71% tracks (Jesteadt et al. 1980). Next, the mean across the two repetitions was calculated. These matched levels for the 0.5-, 2-, and 4-kHz tones, along with the 10-dB SL 1-kHz tone, defined the first set of equally-loud tones for the final stage of data collection.

The four equally-loud tones were combined to create a fixed-level standard or "chime" that was four times as loud as any one of its components. The chime was presented randomly in one interval of each trial, and one of the individual tones was presented in the other interval. For this stage, a 2IFC, four-track paradigm was used. There were two stimulus blocks, with 0.5- and 1-kHz tones matched to the chime in one block, and 2- and 4-kHz tones matched to the chime in the other block. Each block had 200 trials, or 50 trials per track (two tracks per frequency as in the previous stage). In general, the starting stimulus level was about 16 dB above the chime level for the descending tracks and 16 dB below for the ascending tracks. The step size was initially 4 dB and reduced to 2 dB after the fourth reversal. The average matched level was calculated across two tracks for each frequency (i.e., across ascending and descending tracks) and across two repetitions. This resulted in a new set of four equally-loud tones that were combined to create a new, fixed-level chime. This process was repeated until the subject provided responses that would have required a matching tone to exceed 90 dB SPL. Thus, subjects were not presented with the same number of loudness matching conditions. The time required to complete data collection on each subject ranged from 12 to 15.5 hours.

3 Results

3.1 Loudness

Our most basic loudness result is the determination of equally-loud tones at the four stimulus frequencies. We will describe individual results for the subject with the least intra-subject variability, known as our "best" subject, and describe average results for the entire group. Iso-loudness contours for our best subject are shown in

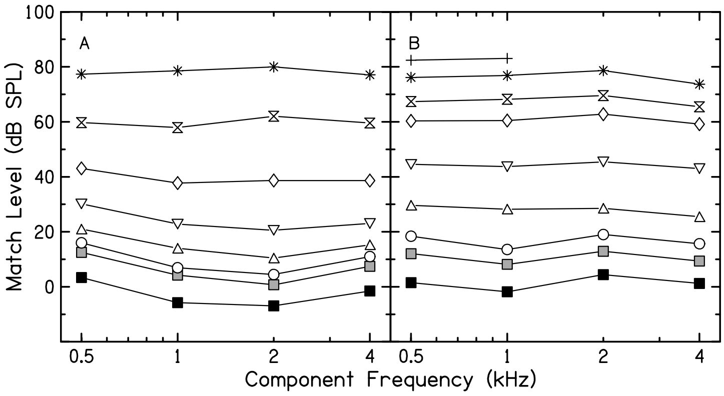

Fig. 1. Equal loudness contours. Black squares represent threshold for each component in quiet. Shaded squares represent the condition in which tones were matched to a 10-dB SL 1-kHz tone. A: Best subject. B: Average across subjects.

Fig. 1A and for the group average in Fig. 1B. The solid-black symbols at the lowest levels represent quiet thresholds. The gray symbol at 1 kHz is at 10 dB SL, which is our reference stimulus for all loudness values. The gray symbols at other frequencies have the same loudness as the reference tone. The open symbols above the gray symbols all have the same loudness as the combined loudness of the first four equally-loud tones. Assuming that the loudness of these four tones is additive (Fletcher and Munson 1933), the loudness of each subsequent iso-loudness contour is four times the loudness of the iso-loudness contour below it. Repeated measurements were performed in two subject to provide an estimate of intra-subject variability. Inter-subject standard deviations (SD) always exceeded intra-subject for matched levels in the 40 to 60 dB SPL range, by as much as 12 dB.

Using the sum of four, equally-loud tones as the standard for loudness matching allows us to determine the number of dB required to quadruple the loudness of each

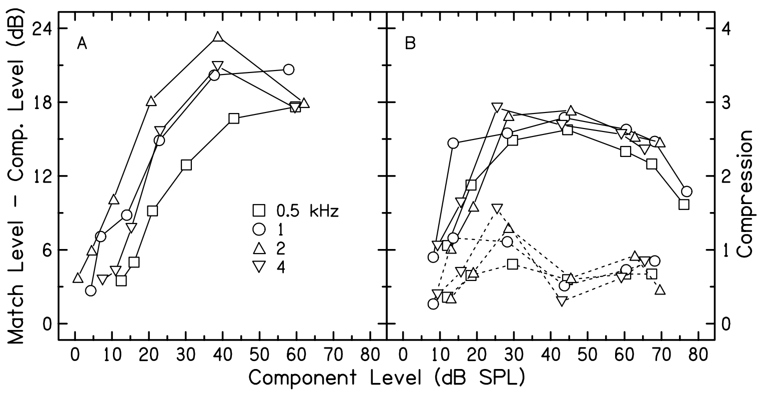

Fig. 2. Number of dB required to quadruple loudness of a tone as a function of its level. Compression is obtained by dividing this number by 6 dB. A: Best subject. B: Average across subjects. Dashed lines represent the corresponding standard deviations.

of the component tones. This result is shown in Fig. 2. As in Fig. 1, panel A represents our best subject and panel B represents group data. In Fig. 2B, the symbols connected by solid lines indicate the group average and the symbols connected by dashed lines the standard deviation across subjects. The number of dB to quadruple loudness is equal to the vertical separation between symbols in Fig. 1. In general, about a 6-dB increase in tone level is required to quadruple the loudness of the lowest-level tones. The number of dB increases rapidly as level increases reaching 15 to 18 dB at moderate levels. At the highest levels, the number of dB often decreases.

Another way to plot the data shown in Fig. 1 is to take vertical slices through iso-loudness contours and plot the relative loudness of each contour as a function of tone level separately for each frequency. This representation of the data is shown in Fig. 3A for our best subject and Fig. 3B for the group average. The loudness values plotted in Fig. 3 are relative to the loudness of a 1-kHz tone at 10 dB SL; hence the units sones/reference on the ordinate. No assumption beyond loudness additivity was required to construct these loudness functions. For comparison, the loudness function (for a 1-kHz tone) of Fletcher and Munson (1933) is superimposed in Fig. 3B. Only loudness values that could be achieved in all subjects are plotted in Fig. 3B. Higher loudness values were not achievable in some subjects because they would have required tone levels exceeding 90 dB SPL. Note that although the highest mean tone level was around 60 dB SPL, the actual range of tone levels across subjects in that chime-stimulus condition was 33.8 at 500 Hz, 44.5 at 1000 Hz, 45.7 at 2000 Hz, and 43.6 at 4000 Hz. We cannot rule out the possibility that the four tones in the chime might partially mask each other at high levels, even though they are an octave apart. Any such masking would be contrary to our assumption of loudness additivity.

We can define *compression* as the number of dB increase in physical intensity for each dB increase in perceptual intensity. If we treat loudness as proportional to

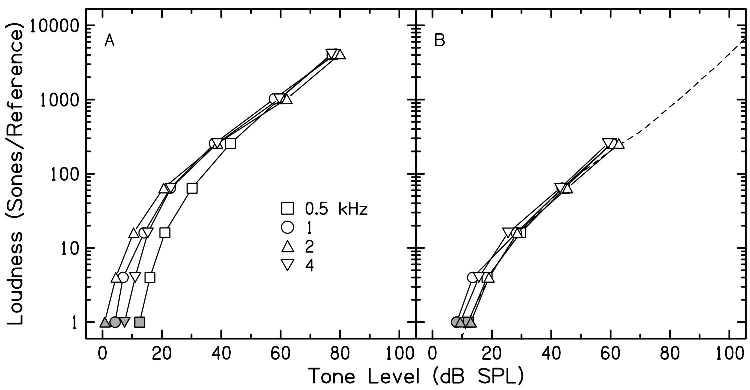

Fig. 3. Loudness functions. These functions are derived from the same data shown in Fig. 1 assuming that the equal-loudness contours are separated by a factor of four. The shaded symbols represent tones that have the same loudness as 1 kHz tone at 10 dB SL. A: Best subject. B: Average across subjects. The dashed line is from Fletcher and Munson (1933), but divided by 20 to account for the difference in reference value

the perceptual intensity, then a four-fold increase in loudness is equivalent to 6 dB increase in perceptual intensity. We can estimate compression conveniently from our loudness measurements by taking the number of dB required to quadruple loudness (as shown in Fig. 2) and dividing that value by 6 dB. The right-hand axis in Fig. 2 shows the compression values that correspond to the dB values on the left-hand axis. Our compression estimates start around 1 for the lowest level tones and increase to between 2.5 and 3 at moderate levels. Some of the highest-level group-average compression values in Fig 2B represent only a subset of subjects, because tone levels exceeding 90 dB SPL would have been required for some subjects to achieve loudness matches. The abrupt decrease in compression at the highest level in Fig. 2B is because the "average" includes only one subject. None of the individual subjects' data exhibited such an abrupt decrease in compression. The compression values in Fig. 2 are equal to the reciprocal of the slope of the loudness functions in Fig. 3.

3.2 Weber fraction

In addition to our primary results for estimates of loudness, we consider Weber fraction estimates derived from the same data.

If we assume that a just noticeable difference (JND) in loudness is determined by the variability imposed by Poisson internal noise, then the loudness JND will be proportional to the square root of the average loudness (McGill and Goldberg 1968, Hellman and Hellman 1990). Using the variable N to represent loudness, this can be written as follows (e.g., Allen and Neely 1997).

$$\Delta N = h\sqrt{N} \quad (1)$$

The factor h is a constant that depends on the reference selected for the loudness scale. If we let α denote compression and assume that JNDs are small, our definition of compression allows us to relate the perceptual JND ΔN to the corresponding physical JND ΔI (e.g., Allen and Neely 1997).

$$\frac{\Delta I}{I} = \alpha \frac{\Delta N}{N} \quad (2)$$

Combining these two equations gives us a way to estimate the Weber fraction from loudness matching data.

$$\frac{\Delta I}{I} = \frac{h\alpha}{\sqrt{N}} \quad (3)$$

This estimate is plotted in Fig. 4A for our best subject using the compression values in Fig. 2 and the loudness values in Fig. 3. The unknown scale factor h has arbitrarily been set to 1 for this figure. Any other value of h would shift the plotted values vertically.

Because our procedure for obtaining loudness matches involved two separate tracks for each condition that converged to the 29% and 71% points on the PF for

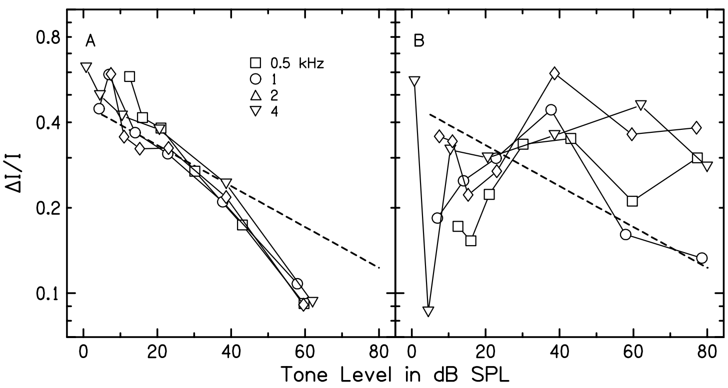

Fig. 4. Weber fraction estimates for best subject. A: Estimated directly from the loudness data in Fig. 3A. B: Estimated from the 29% and 71% points on psychometric function for loudness matching task. The dashed line is the Weber fraction estimate from Jesteadt et al. (1977) as a function of sensation level.

comparison of the loudness of a tone relative to the loudness of a fixed-level chime, we have sufficient data to provide an estimate of the slope of this PF. Relative to a tone at the midpoint of the PF, we can consider the 29% and 71% points on the PF as JNDs in loudness. If the intensity of the lower-level point is I_1 and the higher-level point is I_2, then we can estimate the Weber fraction by the following formula.

$$\frac{\Delta I}{I} = \frac{I_2 - I_1}{I_2 + I_1} \qquad (4)$$

This Weber fraction estimate for our best subject is shown in Fig. 4B. For comparison, the fit by Jesteadt et al. (1977) to their intensity discrimination data, which was independent of frequency, is superimposed as a dashed line in both panels of Fig. 4.

4 Discussion

The method of "loudness matching to the sum of equally-loud tones" was first described as a way to quantify loudness by Fletcher and Munson (1933). One advantage of this method over cross-modality scaling methods (e.g., Hellman, 1999) is its ability to more accurately characterize the rate of loudness growth. Buus et al. (1998) used a similar loudness-matching method to measure loudness; however, they combined equal-SL tones for their loudness matches instead of combining equally-loud tones. One advantage of using equally-loud tones is being able to determine the loudness growth rate directly without additional modeling assumptions.

The variability in the rate of loudness growth across subjects combined with our stimulus-level limit of 90 dB SPL made it difficult to obtain group-average loudness functions at high levels. Within the range of allowable stimulus levels, only four subjects at each frequency perceived a loudness 1024-times the reference and only three subjects perceived a loudness 4096-times the reference. Using only these subsets for the group average biases the result by including only subjects with the largest loudness growth rate. To avoid presenting biased results, we show only group averages in Fig. 3B for loudness values that were perceived by all six subjects.

The group average compression functions in Fig. 2B appear to be representative of the individual compression functions. At moderate levels, compression appears to have a relatively constant value of a little less than three. This is consistent with Steven's Law. At the lowest levels, the compression falls to about one, consistent with energy detection, where perception growth is proportional to sound intensity. At high levels, compression falls to about 2, consistent with a displacement detector, where perception growth is proportional to sound pressure.

We had hoped to see better agreement between the two estimates of the Weber fraction shown in Fig. 4, since both estimates were based on the same loudness measurements. Whether the disagreement indicates a failure of the Poisson internal noise model or reflects measurement errors remains to be determined. Note that the agreement between our Weber fraction estimates and the Jesteadt et al. estimates in Fig 4A could be improved by assuming that the internal noise grows more rapidly at high levels (Allen and Neely, 1997).

Acknowledgements

This work was funded by grants R01 DC02251, R01 DC00136, T32 DC00013, and P30 DC04662 from the National Institutes of Health.

References

Allen and Neely (1997) The relation between the intensity JND and loudness for pure tones and wide-band noise. J. Acoust. Soc. Am. 102, 3628-3646.

Buus, S., Musch, H., Florentine, M. (1998) On loudness at threshold. J Acoust Soc Am. 104, 399-410

Fletcher and Munson (1993) Loudness, its definition, measurement, and calculation. J. Acoust. Soc. Am. 5, 82-108.

Hellman, R. P. (1999) Cross-modality matching: a tool for measuring loudness in sensorineural impairment. Ear Hear. 20, 193-213.

Hellman, W. and Hellman, R. (1990) Intensity discrimination as the driving force of loudness: application to pure tones in quiet. J. Acoust. Soc. Am. 87, 1255-1271.

Jesteadt, W. (1980) An adaptive procedure for subjective judgments. Percept. Psychophys. 28, 85-88.

Jesteadt, W., Wier, C., and Green, D. (1977) Intensity discrimination as a function of frequency and sensation level. J. Acoust. Soc. Am. 61, 169-177.

McGill, W. and Goldberg J. (1968) Pure-tone intensity discrimination as energy detection. J. Acoust. Soc. Am. 44, 576-581.