PROCEEDINGS of the
23rd International Congress on Acoustics

9 to 13 September 2019 in Aachen, Germany

# Loudness of ramped and damped sounds that are temporally shifted across ears

Josef SCHLITTENLACHER[1]; Robbie ZHAO; Brian C. J. MOORE[2]

University of Cambridge, UK

## ABSTRACT

In a previous study we have shown that amplitude-modulated sounds are louder when their modulation is out of phase across the two ears than when it is in phase. The level difference required for equal loudness (LDEL) of sounds with diotic presentation and an interaural modulation phase difference of 180° was about 2 dB. This could be explained by a loudness model where binaural summation lags behind binaural inhibition. The present study investigated the binaural loudness of ramped and damped sounds in a similar manner. Stimuli consisted of trains of 1000-Hz tone pulses with linear rise and fall times with ratios of 1:10 (damped sounds) or 10:1 (ramped sounds). Stimuli contained 28 55-ms pulses, 14 110-ms pulses or 7 220-ms pulses, resulting in a stimulus duration of 1540 ms plus half the pulse duration for the interaurally shifted stimuli. The LDEL between diotic and interaurally shifted stimuli was close to 0 dB for all of these conditions. For a single 220-ms pulse, the LDEL was 1.4 dB for damped sounds, and 3.0 dB for ramped sounds, the diotic sounds being louder. The difference between a single pulse and a pulse train suggests differences between short-term and long-term loudness judgments.

Keywords: loudness, time-varying, binaural

## 1. INTRODUCTION

The loudness of a sound that is presented to both ears is greater than when it is presented to one ear. This effect is called binaural loudness summation. If this summation was perfect in terms of loudness, the level difference required for equal loudness (LDEL) between a tone that is presented monaurally and a tone that is presented diotically would be about 10 dB. However, recent findings suggest that this LDEL is about 5-6 dB (e.g. (1, 2), see (3, 4) for reviews). The effect of non-perfect binaural summation has been explained and modelled by binaural inhibition, based on the assumption that a sound in one ear reduces the internal response to a sound in the other ear (3, 5).

The perceived loudness of a time-varying sound depends on the duration over which it is judged. In one loudness model (6), three aspects of loudness are calculated: [A] Instantaneous loudness, which is calculated from a single very short sample (frame) of the sound and is assumed not to be available to conscious perception; [B] Short-term loudness, which is the loudness of a short segment of sound, such as a syllable; and [C] long-term loudness, which is the loudness judged over longer durations, such as a sentence. Imaging studies have provided evidence for [A] and [B] at different stages of the auditory pathway (7, 8). However, it is not known at which stage binaural inhibition and binaural summation take place. Real-world sounds can differ in both spectrum and time pattern across ears, and predictions of loudness depend on the sequence of implementation of binaural inhibition and binaural summation.

Figure 1 shows a block diagram of the Cambridge loudness model for binaural time-varying sounds (9, 10). The model implements binaural inhibition after the calculation of short-term loudness for each ear. In contrast, long-term loudness is calculated from the inhibited short-term loudness for each ear, and this involves an additional stage of temporal smoothing after binaural inhibition has been applied. With this implementation, the model correctly predicts that amplitude-modulated (AM) sounds whose modulation phase is shifted by 180° across the two ears have a greater long-term loudness than sounds

---

[1] js2251@cam.ac.uk

[2] bcjm@cam.ac.uk

with in-phase AM. The LDEL for such sounds depends on the modulation frequency and is up to 2 dB.

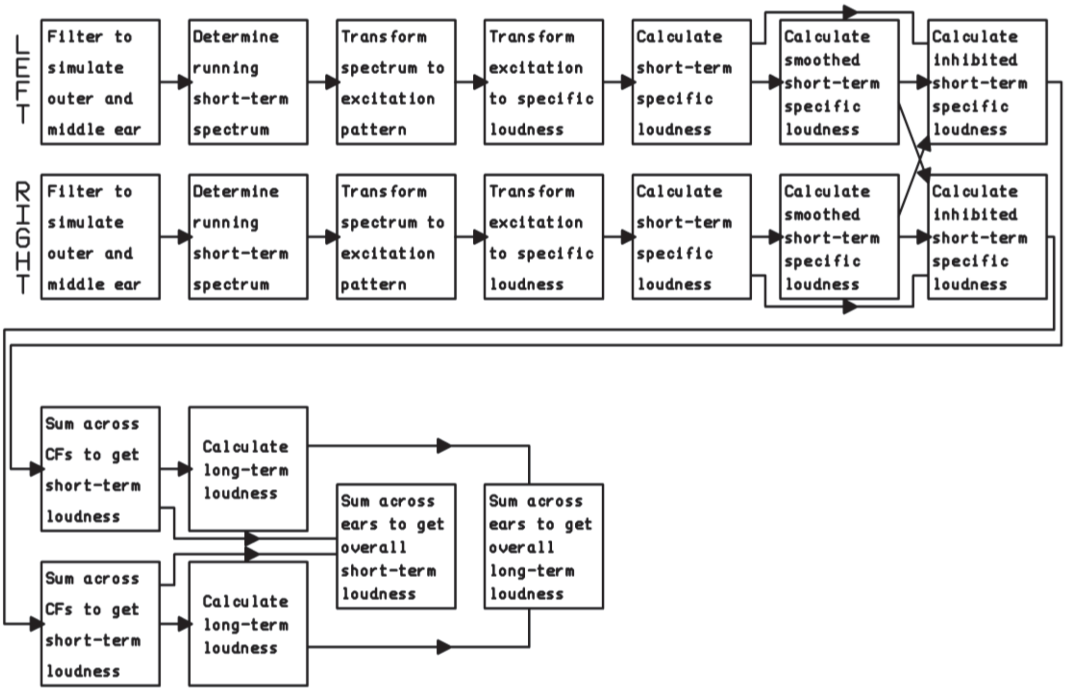

Figure 1 – Block diagram of the Cambridge loudness model

It has been shown previously that ramped sound (with a slow rise and a fast fall) are perceived as louder than damped sounds (with a fast rise and a slow fall) when the overall levels of the sounds are matched (11). However, we are not aware of any studies of the loudness of ramped and damped sounds when the sounds are not identical at the two ears. In the present study we assessed the loudness of trains of tone pulses with different pulse durations that were presented diotically or shifted by half a pulse duration across ears. The pulses were either ramped or damped. The duration of the whole pulse train was slightly longer than 1.5 s (see section 2.3 for details), probably leading to a decision based on long-term loudness. Loudness was also assessed for a single pulse with a duration of 220 ms. For this duration the decision is more likely to be based on short-term loudness.

## 2. METHOD

### 2.1 Subjects

Sixteen subjects, aged 18 to 22 years, ten of them male and six female, participated in the experiment. All of them had normal hearing, with audiometric thresholds not exceeding 20 dB HL between 125 and 8000 Hz in both ears.

### 2.2 Apparatus

Stimuli were generated digitally at a sampling rate of 44100 Hz, D/A-converted by an M-Audio Delta 44 audio interface (Cumberland, RI) and passed through manual attenuators (Hatfield, 2125, Hatfield, UK) to a Sennheiser HD580 headset (Wedemark, Germany).

### 2.3 Stimuli

Stimuli consisted of tone pulses with durations of 55, 110 or 220 ms. The carrier was a 1-kHz pure tone. The tone pulses had linear rise and fall times with a ratio of 1:10 (damped) or 10:1 (ramped), i.e. 5 ms rise and 50 ms fall, 50 ms rise and 5 ms fall, 10 ms rise and 100 ms fall, 100 ms rise and 10 ms fall, 20 ms rise and 200 ms fall, and 200 ms rise and 20 ms fall. The tone pulses were concatenated to form a train with a duration of 1540 ms, formed of 28 55-ms pulses, 14 110-ms pulses or seven

220-ms pulses. A single 220-ms pulse was also used. All stimuli were presented diotically or shifted by half a pulse duration across ears. In the latter case, the total stimulus duration was half a pulse duration longer.

## 2.4 Procedure

Subjects compared the loudness of the diotic and shifted sounds for each pulse duration and pulse shape, resulting in eight runs per subject. A two-alternative-forced-choice task was used. The task was to decide whether the first or second stimulus was louder. A run consisted of four interleaved tracks, each using a 1-up/1-down adaptive procedure (12). For two of the tracks, the diotic sound was the standard stimulus and the shifted sound was the comparison stimulus whose level was varied from trial to trial. For the other two tracks, the shifted sound was the standard and the diotic sound was the comparison. The peak level of the standard was fixed at 70 dB SPL. For each type of standard (diotic or shifted), one track started with a comparison level 10 dB above that of the standard and the other started with a comparison level 10 dB below that of the standard. The order of the standard and comparison stimuli was random. For each trial, the level of both stimuli was changed by the same amount, drawn from a uniform distribution between -5 and +5 dB, to encourage subjects to compare the stimuli within a trial rather than using their long-term memory across trials.

On a given trial, if the comparison was judged as louder than the standard, its level was decreased; otherwise its level was increased. For each track, the change in level was 5 dB until a reversal occurred, 3 dB until a second reversal occurred, and 1 dB thereafter. A run terminated when all tracks reached the eighth reversal. The level of the comparison was averaged across the last four reversals for each track and then averaged across the four tracks (with the sign adjusted so that the mean reflected the level of the shifted sound minus the level of the comparison sound) to yield the LDEL.

# 3. RESULTS

The results are shown in Figure 2. Circles, connected by a solid line for the pulse-train conditions, show results for damped pulses. Squares, connected by a dashed line for the pulse-train conditions, show results for ramped pulses. The LDELs were between -0.7 and 0.7 dB for the six pulse-train conditions. For the single 220-ms pulses, the LDEL was positive at 1.4 dB for the damped pulse and 3.0 dB for the ramped pulse, meaning that at equal levels the diotic sound would be louder.

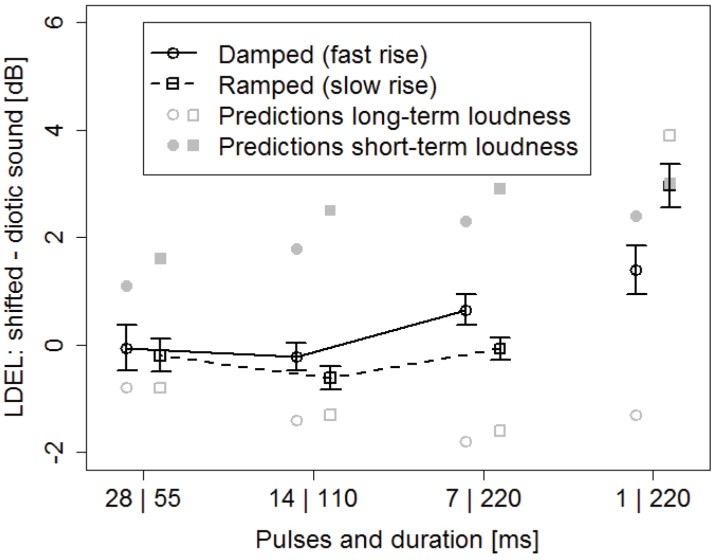

Figure 2 – Level difference required for equal loudness (LDEL) of sounds that were temporally shifted between the two ears by half a pulse duration and diotic sounds. Circles show means for damped pulse shapes, squares for ramped pulse shapes, and error bars the standard error of the mean. Open and filled grey symbols show predictions of the Cambridge loudness model based on long-term and short-term loudness, respectively. Negative values indicate that the shifted sounds would be louder at equal root-mean-square levels.

The results for the single pulses are close to the predictions based on short-term loudness. The results for the pulse trains are between the predictions for short-term loudness and long-term loudness, with a tendency to be closer to long-term loudness with decreasing pulse duration.

A two-way 3x2 (duration x pulse shape) within-subjects analysis of variance on the LDEL values for the pulse trains yielded a significant effect of duration, $F(2,30) = 5.12$, $p < 0.05$, and a significant effect of pulse shape, $F(1,15) = 5.20$, $p < 0.05$. The LDEL was larger for the longest duration and for the damped than for the ramped sounds. The interaction was not significant, $F(2,30) = 1.16$, $p = 0.33$. A two-way 2x2 (number of pulses x pulse shape) within-subjects analysis of variance for the conditions with a pulse duration of 220 ms yielded a significant effect of the number of pulses, $F(1,15) = 69.1$, $p < 0.001$. The LDEL was larger for the single pulse than for seven pulses. The effect of pulse shape was not significant, $F(1,15) = 1.85$, $p = 0.19$. The interaction was significant, $F(1,15) = 11.9$, $p < 0.01$. For one pulse, the LDEL was larger for the ramped than for the damped sound, while for seven pulses the reverse was the case.

## 4. DISCUSSION

The loudness of ramped and damped sounds had been found previously to be predicted well by an early version of the Cambridge loudness model (13). These authors used stimuli with durations between 10 ms and 500 ms and found short-term loudness to predict the results better than long-term loudness (open symbols without error bars in figure 2), although they argued that long-term loudness might be better for longer sounds. In the present experiment the model predicts negative LDELs if the subjects based their estimates on long-term loudness, because binaural loudness summation is implemented after binaural inhibition, which involves an additional stage of temporal smoothing of the inhibited short-term loudness for each ear. In contrast, the short-term loudness for each ear is calculated prior to the application of binaural inhibition, and this leads to the prediction of a positive LDEL if loudness estimates are based on short-term loudness (grey symbols without error bars in figure 2). The clearly positive LDEL values for the single 220-ms pulse suggest that subjects used short-term loudness for this stimulus. For the pulse train with a total duration of about 1.5 s, the LDEL values are not clearly positive or negative. A possible reason is that for sequences of tone pulses, subjects sometimes based their judgments on the loudness of the individual pulses and sometimes on the overall loudness of the whole sounds, leading to a very small net effect.

Just as ramped and damped sounds differ in loudness in a manner that depends on the duration of the sounds (13, 14), the LDEL between sounds that were temporally shifted between the two ears and diotic sounds also depended significantly on pulse shape; the LDEL values were slightly larger for the damped than for the ramped sounds.

The most important findings of the present study are:

(a) For a single tone pulse, the LDELs values were positive, indicating that at equal levels the diotic sounds would be louder than the shifted sounds. This result can be explained assuming that loudness judgments for the single pulse were based on short-term loudness, and that binaural inhibition is applied after the determination of short-term loudness, as assumed in the Cambridge loudness model (9, 10).

(b) LDEL values for the trains of pulses were close to zero, but increased slightly for the longest pulse duration. It is possible that for these sounds subjects sometimes based their judgments on the loudness of the individual pulses and sometimes on the overall loudness of the whole sounds.

## REFERENCES

1. Edmonds BA, Culling JF. Interaural correlation and the binaural summation of loudness. J Acoust Soc Am. 2009;125:3865-70.

2. Schlittenlacher J, Ellermeier W, Arseneau J. Binaural loudness gain measured by simple reaction time. Atten Percept Psychophys. 2014;76:1465-72.

3. Moore BCJ, Glasberg BR. Modeling binaural loudness. J Acoust Soc Am. 2007;121:1604-12.

4. Sivonen VP, Ellermeier W. Binaural Loudness. In: Florentine M, Popper AN, Fay RR, editors. Springer

Handbook of Auditory Research. 37: Loudness. New York: Springer; 2011. p. 169-97.

5. Gigerenzer G, Strube G. Are there limits to binaural additivity of loudness? J Exp Psychol. 1983;9:126-36.

6. Glasberg BR, Moore BCJ. A model of loudness applicable to time-varying sounds. J Audio Eng Soc. 2002;50:331-42.

7. Thwaites A, Glasberg BR, Nimmo-Smith I, Marslen-Wilson WD, Moore BCJ. Representation of instantaneous and short-term loudness in the human cortex. Front Neurosci. 2016;10:183.

8. Thwaites A, Schlittenlacher J, Nimmo-Smith I, Marslen-Wilson WD, Moore BCJ. Tonotopic representation of loudness in the human cortex. Hear Res. 2017;344:244-54.

9. Moore BCJ, Glasberg BR, Varathanathan A, Schlittenlacher J. A loudness model for time-varying sounds incorporating binaural inhibition. Trends Hear. 2016;20:2331216516682698.

10. Moore BCJ, Jervis M, Harries L, Schlittenlacher J. Testing and refining a loudness model for time-varying sounds incorporating binaural inhibition. J Acoust Soc Am. 2018;143:1504-13.

11. Stecker GC, Hafter ER. An effect of temporal asymmetry on loudness. J Acoust Soc Am. 2000;107:3358-68.

12. Levitt H. Transformed up-down methods in psychoacoustics. J Acoust Soc Am. 1971;49:467-77.

13. Ries DT, Schlauch RS, DiGiovanni JJ. The role of temporal-masking patterns in the determination of subjective duration and loudness for ramped and damped sounds. J Acoust Soc Am. 2008;124:3772-83.

14. Ponsot E, Susini P, Meunier S. A robust asymmetry in loudness between rising- and falling-intensity tones. Atten Percept Psychophys. 2015;77:907-20.