inter·noise 2021
1–5 August | Washington, DC

Temporal integration of partial loudness of helicopter-like sounds

Josef Schlittenlacher[1]
Division for Human Communication, Development and Hearing
University of Manchester
M13 9PL Manchester, UK

Brian C. J. Moore[2]
Cambridge Hearing Group
Department of Psychology
University of Cambridge
CB2 3EB Cambridge, UK

ABSTRACT

When developing new vehicles that are to be operated in existing background noise, such as electric vertical take-off and landing aircraft in cities, a sound design goal should be to minimize the loudness in the given background noise. Rotorcraft sounds are characterized by short bursts of noise and the choice of rotor size and number allows variation of their temporal characteristics. We asked participants to compare the loudness of a reference sequence of bursts of noise with a burst duration of 20 ms with that of a test sequence for which the burst duration was 1, 2, 5 or 10 ms. For both sequences there were 20 bursts/s. A two-interval, two-alternative forced-choice task and a 1-up/1-down procedure was used. Simulated street noise was presented simultaneously with the noise bursts, and had the same root-mean-square (RMS) level as the fixed reference train of bursts of about 65 dB SPL. Initial results indicate that for equal loudness of the test and reference sequences, the level is markedly lower for a short burst duration than for a longer burst duration. This means that at a fixed equivalent sound pressure level partial loudness increases with decreasing burst duration.

1. INTRODUCTION

In many circumstances machinery and vehicles that produce sounds are operated in a noisy environment, for example in a city. In these cases, it is desirable to design a new source such that the impact of the additional noise it produces is as small as possible. Current examples of new sound sources are electric vertical takeoff and landing aircraft (eVTOL), also known as “air taxis”, that several companies plan to operate in urban areas within the next few years. The sound of eVTOLs depends on the composition of their rotors which, when combined and operated in a smart way, affords the opportunity for a sound design that is considerably quieter than that of a helicopter. Noise regulations should be created to reward designs that are perceived as quiet, and market forces may lead to the choice of eVTOLs that blend well into an existing sound environment.

---

[1] josef.schlittenlacher@manchester.ac.uk

[2] bcjm@cam.ac.uk

Many regulations rely on metrics that are based on equivalent sound pressure level, for example the sound exposure level (SEL), the day-night-average sound level (DNL), or the community noise equivalent level (CNEL; ANSI/ASA S1.1-2013, sections 3.22, 3.19, 3.20). These metrics ignore the presence of background noise, as if the vehicle was operated in quiet, and they make a very simplified assumption about temporal integration for loudness by simply averaging energy over time. Rotorcraft sounds are characterized by a regular sequence of noise-like bursts. Because eVTOLs have several rotors or similar electric propulsors, their sounds can be made more or less impulsive. However, the effects of this on loudness are unlikely to be reflected in measures like the SEL, DNL, or CNEL.

The time-varying partial loudness model of Glasberg and Moore (2005) offers the potential to overcome these limitations; partial loudness refers to the loudness of a given sound in the presence of a background sound. The 2005 model is based on the model of Glasberg and Moore (2002; see also Moore et al. 2016, 2018) and uses time constants that are based on the perception of loudness. Furthermore, it explicitly models the perception of loudness in the presence of background sounds. It has been used for various sound sources, such as the ringtones of mobile phones (Glasberg and Moore, 2005), wind turbine noise (Bolin & Nilsson, 2010), water fountains (Radsten-Ekman, 2015), and music (Ma et al., 2014). Without background sounds, it has also been used for aircraft noise (e.g., Marshall & Davies, 2011; Swift & Gee, 2011).

The present study investigates the partial loudness of bursts of noise of various durations in background noise, i.e. artificial sounds that are similar to those of rotorcraft in an urban environment. The results are compared with the predictions of the model of Glasberg and Moore (2005).

2. METHOD

Because the study focused entirely on temporal aspects of loudness perception and relied mainly on relative levels that were well above the detection threshold, it could be done via the internet, since the exact characteristics of the electroacoustic equipment used (e.g. the frequency response of the headphones) should have had, at most, a minor effect. Further measures, as described below, were implemented to control the absolute level and to rule out undesired spectral effects resulting from the use of unknown headphones.

2.1. Participants

The study was pre-registered at osf.io/4cfn9 for twenty participants. At the time of writing, nine participants had completed the experiment, four of whom were female. Their ages ranged from 20 to 52 years, with a mean of 28 years and a median of 24 years.

2.2. Apparatus

The experiment was conducted online. Participants were asked to wear headphones and to use a computer or laptop. The wav files were presented via the web browser, with underlying code in JavaScript and the Web Audio API.

2.3. Stimuli

The background noise was a 2-s long recording of a highway, which was a rather stationary sound. Because we could not control the headphones that the participants used in the online experiment, the noise was bandpass-filtered with cut-off frequencies of 250 and 4000 Hz to limit the frequency range to that reproduced well by consumer headphones. Raised-cosine rise and fall times of 20 ms were applied to the background. The background was the same for all trials.

The helicopter-like signals had the same average spectrum as the background noise. First, a stationary noise with the same spectrum as the background noise and a low crest factor was generated using the procedure described by Moore et al. (2004). Second, noise pulses with durations of 1, 2, 5, 10 and 20 ms were cut out of that stationary noise using rectangular windows. These pulses were repeated at a rate of 20 Hz to result in a helicopter-like sound. Every second pulse was the same as the preceding pulse except for a phase shift of 180°; this avoided any overall DC offset.

The pulse trains had durations of 1.6 s, and started 200 ms after the background noise commenced. All levels are expressed as root-mean-square (RMS) levels over these 1.6 s, unless stated otherwise. All stimuli were presented diotically.

2.4. Procedure

The pulse trains with pulse durations of 1, 2, 5 and 10 ms (target sounds) were each compared in loudness to a pulse train with a pulse duration of 20 ms (reference sound) in a two-interval, two-alternative forced-choice task. The target and reference sounds were presented in a random order. Participants were told that both intervals contained a background street noise and a helicopter-like sound. In each trial, the participant was asked to indicate the interval in which the helicopter-like sound was louder.

Each target sound was compared to the reference sound in four runs of a 1-up/1-down procedure. In two runs the reference was fixed and the target was varied, while in the other two runs the reference was varied and the target was fixed. The fixed signal had a level of 0 dB relative to the background noise, and the starting relative level of the variable sound was either +10 dB or −10 dB relative to the background noise. For each run, the 1-up/1-down procedure terminated after eight reversals. The step size was 5 dB until the second reversal occurred, then 3 dB until the fourth reversal occurred, and 1 dB thereafter. The mean level at the last four reversals was used to calculate the level difference needed for equal loudness (LDEL) of the target and reference. The sixteen runs were performed in pseudo-random order.

To set the overall level, participants were asked to adjust their system settings so that running speech from a female speaker had a natural loudness. We assume that most participants adjusted the level to be 60 to 65 dB SPL. The speech sound had an RMS level of −36 dB relative to a full-scale sinusoid. The background noise was presented at the same level (and thus also the level of the fixed signal).

3. RESULTS

To facilitate comparison of runs where the target was varied and runs where the reference was varied, the sign of the LDEL was inverted when the 20-ms bursts (reference) were varied in level and the 1, 2, 5 or 10-ms target bursts were fixed in level. Thus, a negative sign means that the sequence with the shorter pulses was louder at an equal level, and needed a lower level than the sequence with the 20-ms bursts for equal loudness.

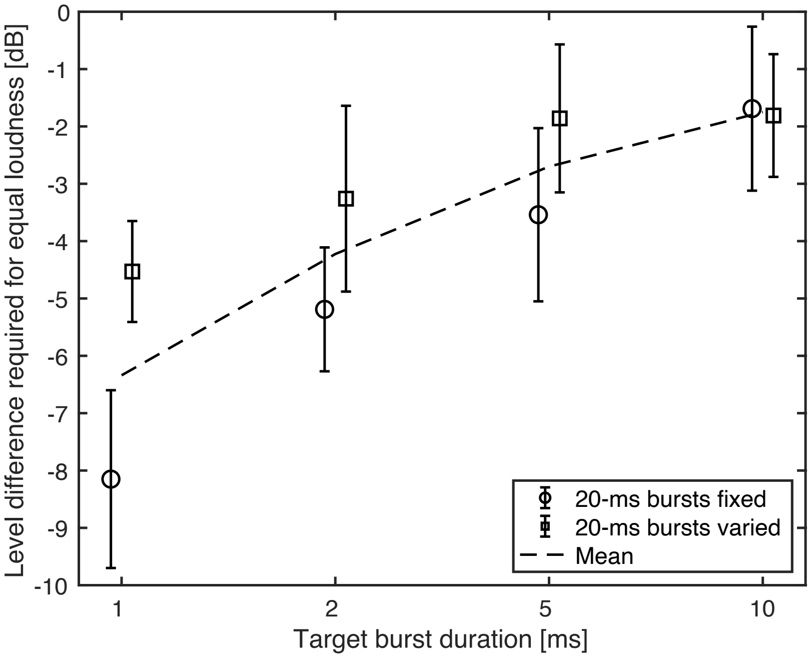

Figure 1: Level difference required for equal loudness (LDEL) as a function of target burst duration.

Figure 1 shows the LDELs averaged across participants and the two runs that differed only in the relative starting level of the adjustable sound. Squares show results for the conditions in which the level of the 1, 2, 5 or 10-ms target was varied and circles show results for the conditions in which the level of the 20-ms reference was varied. Error bars show standard errors of the mean, based on standard deviations across runs and participants of about 3 to 5 dB. The dashed line shows the mean across the runs with the different variable sounds. The LDEL was most negative for the smallest burst duration of 1 ms, with a mean LDEL of –6.3 dB, and became closer to 0 with increasing burst duration, reaching –1.8 dB for the target duration of 10 ms.

With all LDELs being negative, the runs with the different variable sounds (circles versus squares) differed not only in which of the two sound was fixed but also in absolute level and level relative to the background noise: at the point of equal loudness with the 20-ms bursts, the 1-ms bursts had a relative level of –8.2 dB when they were varied compared to –4.5 dB when they were fixed. Thus, the level relative to the noise and the partial loudness were lower when the shorter bursts were varied.

A two-way within-subjects analysis of variance with factors target duration and fixed/variable sound showed a significant effect of target duration, $F(3,24) = 13.5, p < 0.001$, and a significant effect of fixed/varied sound, $F(1,8) = 69.1, p < 0.001$, but no significant interaction, $F(3,24) = 1.81, p = 0.17$.

It is clear from Figure 1 that the equivalent sound pressure level of the bursts, which involves integration over the whole stimulus, is not a good predictor of partial loudness. We next assessed whether the model of Glasberg and Moore (2005) could predict the results. The model involves the initial calculation of instantaneous partial loudness, which is assumed to be an intervening variable that is not accessible to conscious perception. The instantaneous partial loudness is calculated from short-term spectra using window durations from 2 to 64 ms, depending on the frequency range; the shortest window is used for the highest frequencies and the longest window is used for the lowest frequencies. The short-term partial loudness, which corresponds to the momentary impression of loudness, is calculated from the instantaneous loudness using an averaging mechanism similar to an automatic gain control system, with an attack and release time. It was found that the short-term partial loudness calculated using the model did not predict the data; the predicted LDELs were all close to zero. The discrepancy between short-term partial loudness and the current results may indicate that the durations of the windows used to calculate the short-term spectra need to be reduced or that the time constants used to calculate short-term partial loudness from instantaneous partial loudness need to be reduced.

The LDELs were in fact closer to the values predicted from the peaks of the instantaneous partial loudness values as illustrated in Figure 2. This shows instantaneous partial loudness as functions of time at the levels of equal loudness for the target sounds (assuming an SPL of 63 dB for the background noise).

The maxima of the calculated instantaneous partial loudness functions vary over only a small range, from about 16 sone (10-ms bursts) to about 22 sone (1-ms bursts). This corresponds to a change in relative level of 3 dB, which is comparable to the confidence intervals of the measurements. However, the trend for the instantaneous partial loudness to decrease with increasing burst duration suggests that slightly longer temporal integration is needed than the one that is imposed by the window size used to calculate instantaneous loudness.

4. CONCLUSIONS

The results show that equivalent sound pressure level is not suitable for predicting the temporal integration of loudness for sequences of noise bursts in background noise. The discrepancies were up to 8 dB in the present experiment. Also, the LDELs differed by about 4.7 dB for the target sounds with 1 and 10 ms burst durations (6.5 dB when both were fixed and the 20-ms bursts were varied), which cannot be explained by a procedural bias. The discrepancy would probably be even higher when comparing a sequence of short pulses like helicopter sounds to a more continuous sound like that of some eVTOLs.

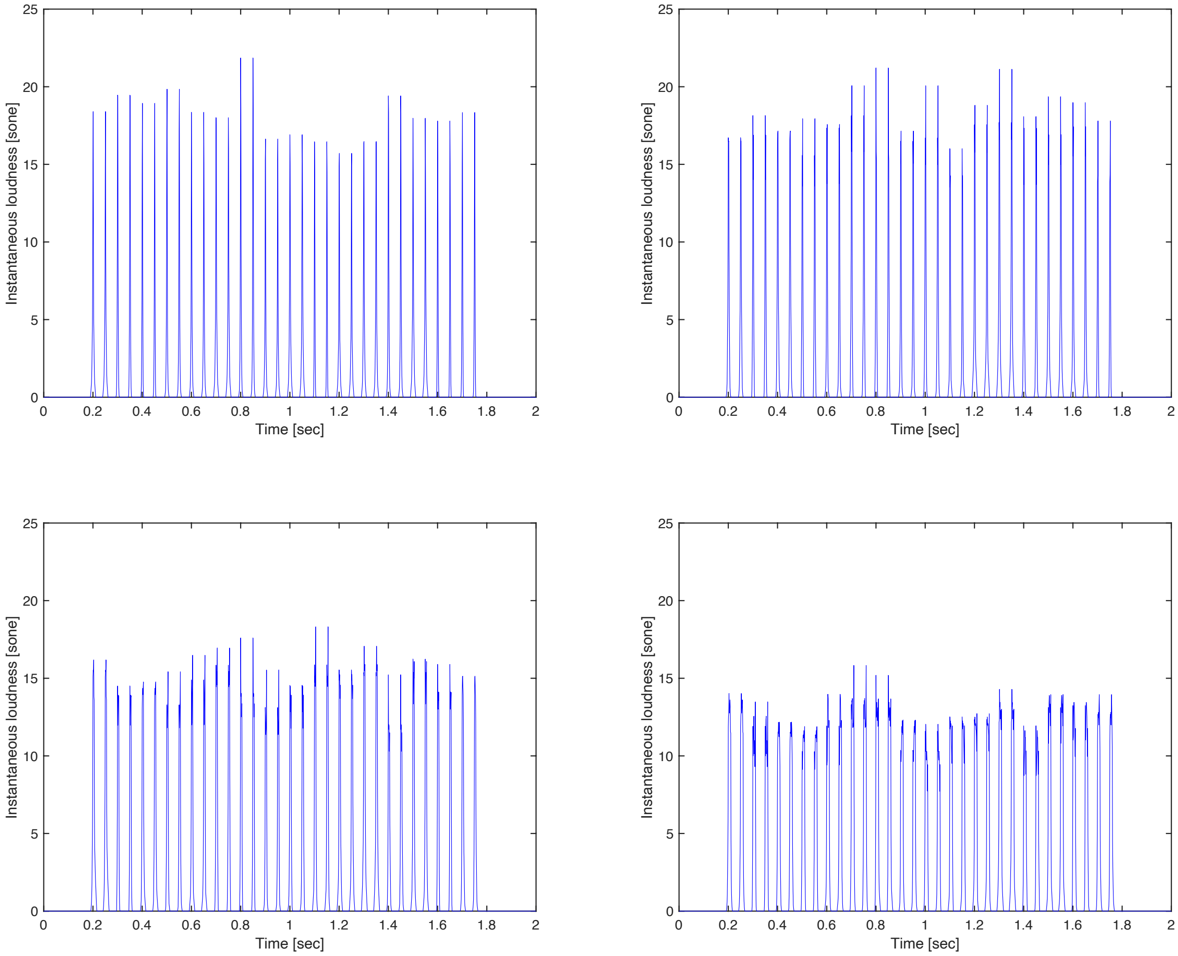

Figure 2: Instantaneous partial loudness as a function of time for the sequences with burst durations of 1 ms (top left), 2 ms (top right), 5 ms (bottom left) and 10 ms duration (bottom right) when they were adjusted to have the same loudness as the reference with 20-ms bursts.

The time-varying loudness model of Glasberg and Moore (2002) has been reported to provide accurate predictions for aircraft noise in quiet (e.g., Marshall & Davies, 2011; Swift & Gee, 2011). The model for time-varying partial loudness (Glasberg and Moore, 2005) was based on the 2002 model and was successful in predicting the partial loudness of various sounds in background noise. The present study motivates the use of a time-varying partial loudness model although at present more data are needed to pinpoint the time constants for the temporal integration of partial loudness. The present experiment used only one burst rate (20 bursts/s) and a limited range of target levels. Its results suggest that the temporal integration needed in the model is longer than that used for the calculation of instantaneous partial loudness but shorter than that predicted by short-term partial loudness. Data are needed for a wider range of burst rates and levels.

Frequently used metrics for the regulation of environmental noises, like the SEL and the DNL, are based on equivalent sound pressure levels. In order to promote the use of vehicles for urban air mobility that are relatively quiet in the presence of typical environmental background noise, better metrics are needed. Such metrics should be able to give good predictions of partial loudness for vehicles that differ considerably in number, type and setup of propulsors and thus the type of noise that they produce. A model based on partial time-varying loudness is likely to be suitable for this purpose.

5. ACKNOWLEDGEMENTS

J. S. was supported by the Wellcome Trust University of Manchester TPA [209741/Z/17/Z] and by the NIHR Manchester Biomedical Research Centre.

6. REFERENCES

1. American National Standards Institute/Acoustical Society of America (ANSI/ASA), *Acoustical Terminology*, S1.1 2013 Edition.
2. Bolin, K., Nilsson, M. E., & Khan, S. The potential of natural sounds to mask wind turbine noise. *Acta Acustica united with Acustica*, ***96*(1)**, 131 – 137 (2010).
3. Glasberg, B. R. & Moore, B. C. J. A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society,* **50(5)**, 331 – 342 (2002).
4. Glasberg, B. R. & Moore, B. C. J. Development and evaluation of a model for predicting the audibility of time-varying sounds in the presence of background sounds. *Journal of the Audio Engineering Society,* **53(10)**, 906 – 618 (2005).
5. Ma, Z., Reiss, J. D., & Black, D. A. Partial loudness in multitrack mixing. In *Audio Engineering Society 53rd International Conference*, London, UK, January 2014.
6. Marshall, A. & Davies, P. Metrics including time-varying loudness models to assess the impact of sonic booms and other transient sounds. *Noise Control Engineering Journal,* **59(6)**, 681 – 697 (2011).
7. Moore, B. C. J., Glasberg, B. R., Stone, M. A. New version of the TEN test with calibrations in dB HL, *Ear and Hearing*, **25(5)**, 478 – 487 (2004).
8. Moore, B. C. J., Glasberg, B. R., Varathanathan, A., Schlittenlacher, J. A loudness model for time-varying sounds incorporating binaural inhibition, *Trends in Hearing*, 20, 1 – 16 (2016).
9. Moore, B.C.J., Jervis, M., Harries, L. & Schlittenlacher, J. Testing and refining a loudness model for time-varying sounds incorporating binaural inhibition, *Journal of the Acoustical Society of America*, **143(3)**, 1504 – 1513 (2018).
10. Radsten-Ekman, M. *Wanted Unwanted Sounds: Perception of sounds from water structures in urban soundscapes,* Holmberg, 2015.
11. Swift, S. H. & Gee, K. L. Examining the use of a time-varying loudness algorithm for quantifying characteristics of nonlinearly propagated noise (L). *Journal of the Acoustical Society of America*, **129(5)**, 2753 – 2756 (2011).