Comparison between four methods of loudness estimation of stationary and non-stationary sounds

Isabelle Boullet

GENESIS, Bâtiment Beltram BP 69, Domaine du Petit Arbois, F-13545 Aix en Provence Cedex 04, France
CNRS-LMA, 31 Chemin Joseph Aiguier, F-13402 Marseille Cedex 20, France
Email: isabelle.boullet@genesis.fr

Introduction

It is well known that loudness is an important parameter for annoyance and sound quality. Currently, loudness can only be measured with psychoacoustic tests. Various models exist but they present limits, in particular, for impulsive sounds. This research is part of a study designed to propose a new model of loudness estimation for impulsive sounds.

The aim of this study has been firstly to determine with which precision could psychoacoustic tests permit to estimate the loudness. Another objective has been to know the inter-subject variability of each of four methods of loudness estimation, in order to determine if this variability depends on temporal patterns (stationary or non-stationary sounds). Thus, four methods often used were compared: magnitude estimation [1], the method of adjustment [2], an adaptive procedure in a two-alternative forced choice [3], and an adaptive procedure with multitracking [4].

Methods

Stimuli

Three type of sounds were studied:

Stationary sounds were one real sound (gravels being poured) and three synthesized sounds (pure tones at 200 Hz and 4 kHz, and a narrow-band noise centred on 1 kHz, 40-Hz wide). The duration for each sound was 1 second and their sound pressure levels were between 47.8 and 77.9 dB SPL.

Nine **impulsive sounds** were also compared which had a short rise time (less than 5 ms) and an exponential decay. These included five real sounds (champagne cork, two stones struck together...) and four synthesized sounds (a pure tone of 1 kHz and three white noises). Their durations were between 203 ms and 501 ms. Their peak levels were between 62.7 and 89.9 dB SPL.

The final type sounds were four **repeated impulsive sounds**, which were a real sound (a hydraulic rock breaker with 8 impulses per second) at two different sound levels and two synthesized sounds (4 and 20 impulses per second). Their durations were between 1 and 3.8 s. Their peak levels were between 62.7 and 77.7 dB SPL.

Procedure

Seven experiments were made on the measurement of the loudness level for stationary, impulsive and repeated impulsive sounds. The various methods are summarised in Table 1.

The first method was **magnitude estimation** without reference. This method permits the loudness to be measured in sones. The loudness level was calculated with the relationship: $\text{Phons} = 40+10*\log_2(\text{sones})$ (1)

The **method of adjustment** was divided into two sessions where the test sound (T) was presented before (Adjustment1) or after (Adjustment2) the sound of comparison (C). The loudness level was the average of the two estimations. In this way the bias caused by the louder judgement of the second-played sound was reduced.

The third method was an adaptive method in a two-interval, two-alternative forced-choice (**2I-2AFC**) paradigm. This experiment was also divided into two sessions : 2down-1up procedure, which permitted to determine the loudness level just louder [5]; and 1down-2up procedure, which permitted to determine the loudness level just softer. The loudness level was the mean of these two values. A single measurement of the loudness level just louder (or just softer) was taken as the average across three tracks, each one ended after six reversals.

The principle of the last tested method, a **multitracking method**, was similar to the 2I-2AFC method, but instead of following a single sequence, several sounds (four or five) were mixed and then several sequences are simultaneous.

Apparatus

The apparatus was the same for all tested methods. Sounds files were converted into analog signals with an OROS sound card. The comparison sound level was varied using a Wilsonics (model PATT) programmable attenuator before being fed to a Genelec loudspeaker in a frontal position, 1.5 meters from the listener, in an anechoic chamber. The sound system was calibrated with a 1-kHz pure tone with an SPL of 90 dB, using of a Bruel & Kjaer (type 2669) microphone at the center of what would be the place of the listener's head during the experiment.

Results

Figure 1 illustrates the loudness level obtained with the four methods tested and the seventeen sounds. This study showed that the methods give equivalent loudness levels when the loudness level is directly measured in phons. The magnitude estimation consistently underestimates the loudness level compared to the three other methods. This bias could be introduced by the formulae (1), which permits phons to be calculated from sones. Indeed, formulae (1) is based on a slope of 0.6 of the loudness function for a 1-kHz pure tone.

Methods	Estimation	Adjustment 1	Adjustment 2	2I-2AFC 2down 1up	2I-2AFC 2up 1down	Multitracking 2down 1up	Multitracking 2up 1down
Test sound (T)	4 stationary sounds, 9 impulsive sounds and 4 repeated impulsive sounds						
Choice of T	Randomly, different for each listener						
Sound of Comparison (C)		Spectrum: Narrow-band noise centred on 1 kHz (120-Hz-wide) Duration: 1 s for stationary sounds and repeated impulsive sounds and 500 ms for impulsive sounds					
Presentation order of T & C	T alone	T & C	C & T	T & C or C & T, randomly			
Silence between T & C		500 ms					
Start level of C by report to T		Randomly 10 dB above or under		15 dB above	15 dB under	15 dB above	15 dB under
Task	Give a number proportional to the loudness of T	Adjust the C-level to have the same loudness as T; (step was + or – 5, 2, 1 dB)		Which of these 2 sounds is louder (first or second); Step was 5 dB and 2 dB after the 2nd reversal			
Test duration for 10 sounds	**3 min**	**10 min**	**10 min**	**40 min**	**40 min**	**40 min**	**40 min**
Listeners	14 (5 women and 9 men) who were 25 to 58 years old with a normal hearing						

Table 1: Experimental conditions for the four methods of loudness estimation

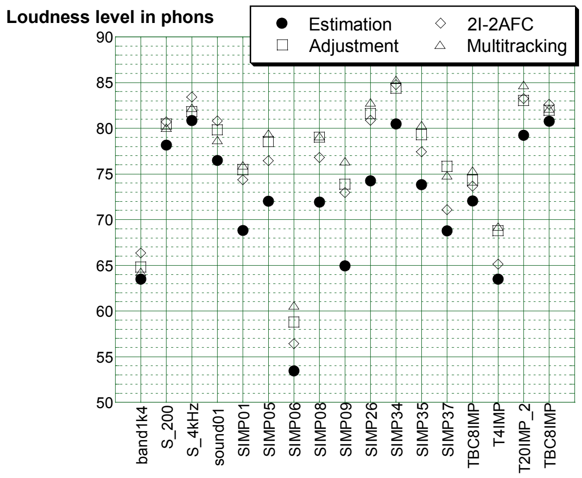

Figure 1: Loudness level estimated for four stationary sounds (the first 4), nine impulsive sounds (the following 9) and four repeated impulsive sounds (the last 4).

Standard deviation	Estimation	Adjustment	2I-2AFC	Multitracking
Stationary	5.4	**3.4**	3.5	3.9
Impulsive	7.1	5.25	**4.9**	5.3
Repeated impulsive	5.5	4.7	4.3	**3.4**
Mean	**6**	**4.4**	**4.2**	**4.2**

Table 2: Standard deviations in phons for each method and each type of sound. The last raw is the mean of the standard deviations for all sounds but for one method.

But it has been showed that the slope of the loudness function could vary between 0.5 and 0.7 [6], (0.5 for our group of listeners). That could explain the underestimations of the loudness level obtained with the magnitude estimation. It would be better to calculate the loudness level with the relation (1) taking into account the individual slope of the loudness function.

Another important point of this research was to study the variability among subjects. The standard deviations for each of the four methods are given in table 2 for each type of sound. The magnitude estimation produces the most inter-subject variability. The three other methods present equivalent standard deviations. The variability of each method changes with the type of sound. In particular, when the test sound and the sound of comparison are very different, like for impulsive sounds, each method presents a higher variability.

Conclusions

In this paper four methods of loudness estimation for stationary, impulsive and repeated impulsive sounds were compared. The experiments showed that the method of adjustment presents the best precision-duration compromise. The prejudged precision that we plan for our new loudness model could be about 4 phons for an absolute judgment. It was also confirmed that the relationship between phons and sones must be used carefully.

Acknowledgments

I would like to acknowledge Sabine Meunier and Georges Canévet from the Laboratory of Mechanics and Acoustics from Marseille for their help in this research.

References

[1] Stevens S.S., "Calculation of the loudness of complex noise", J. Acoust. Soc. Am. **28** (1956), 807-832

[2] Gescheider, G.A., Psychophysics, Method, Theory, and Application, Second Edition, Lawrence Erlbaum Associates, New Jersey, USA, 1985

[3] Jestaed, W., "An adaptive procedure for subjective judgements", Percept. Psychophys. **28** (1980), 85-88

[4] Buus, S, Florentine, M., and Poulsen, T., "Temporal integration of loudness, loudness dicrimination, and the form of the loudness function", J. Acoust. Soc. Am. **101** (1997), 669-680

[5] Schlauch, R. S., and Wier, C.C., "A method for relating loudness matching and intensity-discrimination data", J. Speech Hear. Res. **30** (1987), 13-20

[6] Canévet, G., Hellman, R., and Scharf, B., "Group estimation of loudness in sound fields", Acustica **60** (1986), 277-282