Methods for Measuring Loudness and their Application to Sounds Fluctuating in Space and Time

W. Ellermeier et J. Schlittenlacher
Technische Universität Darmstadt, Alexanderstr. 10, Institut für Psychologie, 64283 Darmstadt, Allemagne
ellermeier@psychologie.tu-darmstadt.de

The various methods used to measure the subjective intensity of sound, i.e., loudness, are reviewed, among them axiomatic approaches, intramodal and cross-modality matching, direct and indirect estimation, and simple reaction time. Subsequently, the merit of using molecular psychophysics to the study of loudness is discussed. As an example, the binaural loudness of non-stationary dichotic sounds varying both from moment to moment and in their interaural level difference is investigated. To that effect, two streams of independently varying 1-s broadband noise samples randomly changing in level every 100 ms are delivered to the left and right ear, respectively. Subjects have to categorize the overall loudness of these time-varying sounds into a 'loud' and 'soft' group. Subsequently, psychometric functions relating the overall judgment (a) to the random level fluctuations in the ten 100-ms segments, and (b) to the lateral position changes are derived. That way, both temporal and lateral weights operating in the determination of overall loudness may be obtained. The results are related to those obtained with other methods studying binaural loudness or the loudness of non-stationary sounds.

1 Introduction

Binaural loudness has been studied employing nearly the entire methodological repertoire available to study suprathreshold hearing sensations: Axiomatic measurement, intramodal and cross-modality matching, direct and indirect estimation, and simple reaction time (for reviews, see [1, 2]). None of the methods enumerated, however, provides a sufficiently detailed view of how the auditory system integrates level information from the two ears on a moment-to-moment basis.

That, precisely, is the domain of 'molecular psychophysics,' a group of techniques by which the observer's (global) decisions in a perceptual task involving several channels of information are analyzed with respect to the contribution of each channel. In studying binaural loudness, these information channels might be operationalized as the input received (a) from the two ears, and (b) at several, successive points in time. Specifically, noises superimposed with independent, random level fluctuations (x_i) are dichotically delivered to the ears, and the resulting "noisy" stimulus is submitted to a perceptual judgment. By analyzing the observer's decisions as a function of these random fluctuations x_i, and by doing so separately for each of the information channels involved (lateral position, moment in time), conditional psychometric functions (*COSS functions*, s. [3] and method section below) are obtained, the shapes of which reflect the contribution of each of the information channels to the overall decision. Based on these COSS functions, or - alternatively - on statistical regression analysis, weights w_i may be estimated which determine, how strongly, and in which direction, a given stimulus component influences the global judgment.

Molecular psychophysics techniques have been used to study intensity discrimination of level-fluctuating sounds [4], auditory spectral shape discrimination ('profile analysis'; [5]), and the loudness of sounds simultaneously varying in level across time and the frequency spectrum [6]. The present investigation focuses on the joint operation of temporal integration and loudness summation across the two ears. To that effect, noise samples fluctuating in level over time and doing so independently for each ear are subjected to COSS analysis, thus yielding both temporal and 'lateral' weights for the loudness of short sounds.

2 Method

2.1 Participants

Five volunteers (age 25–57 years, median age 32, all male) participated in the experiments. Each of them passed a hearing test, confirming that thresholds did not exceed 20 dB HL for all audiometric frequencies from 125 Hz to 8 kHz, the exception being listener WE who had a 25-dB hearing loss in his left ear at 8 kHz. All participants were members of the laboratory who - with one exception - had no prior experience with the kind of sample discrimination task employed.

2.2 Apparatus and procedure

Stimuli were generated digitally at a sampling rate of 48 kHz and with 16 bits resolution. They were D/A converted by an external sound card (RME multiface II), passed through a headphone amplifier (Behringer Powerplay Pro 8000) and were dichotically delivered to Beyerdynamic DT 990 headphones. Sound levels were calibrated using a 94-dB sinusoid the level of which was measured at the headphones using an artificial ear (Brüel & Kjær type 4153) connected to a sound level meter (Brüel & Kjær type 2250).

The stimuli presented on each trial were samples of white noise having 1 s duration. Their overall level was randomly varied every 100 ms, thus producing a stepwise level-fluctuating sound consisting of 10 segments (see Figure 1). The overall level of each segment was drawn randomly from one of two normal distributions denoted "signal" and "noise", with the "signal" distribution having a higher mean value. The "noise" distribution had mean value $\mu_n = 60$ dB SPL and a standard deviation of $\sigma_n = 2$ dB. The "signal" distribution had a mean value $\mu_s = 61$ dB SPL and a standard deviation of $\sigma_s = 2$ dB. Consequently, approximately 95% of the segment levels for each distribution fall in the range $\mu \pm 4$ dB. The distributions were truncated at $\mu \pm 15$ dB. In contrast to the schematic shown in Figure 1, the transitions between the segments were smoothed by 2-ms gaussian-shaped ramps. On each trial, all 20 noise segments (10 in the left, and 10 in the right ear) were independently drawn from either distribution, thus allowing the levels in the ears to vary independently.

Samples - derived either from the noise or the signal distribution with equal probability - were presented one at a time in a single-interval 2AFC task. The listener had to decide whether a given noise sample came from the 'signal' or 'noise' distribution, or, in the more colloquial terms employed in the instructions, belonged to a 'loud' or 'soft'

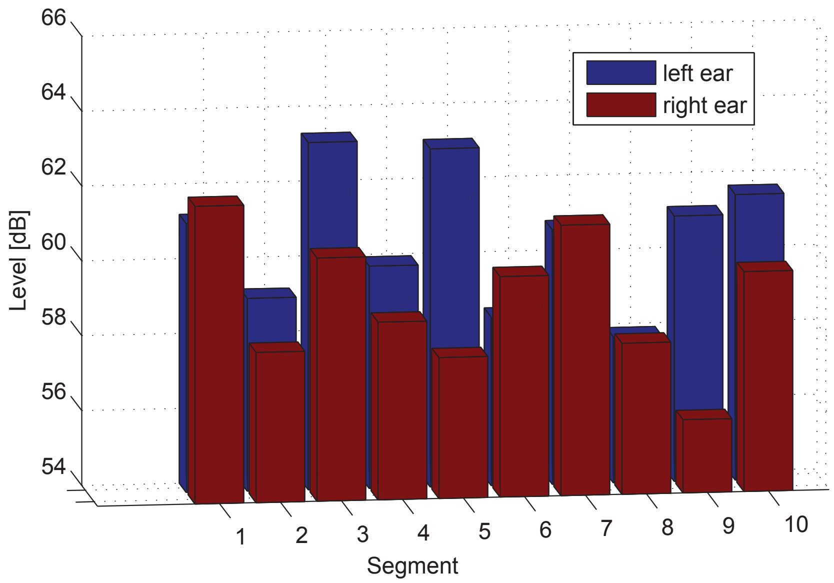

Figure 1: Schematic depiction of the stimuli employed: Ten adjacent noise segments of 100 ms each, randomly fluctuating in level, and independently varying in the two ears were presented on a given trial.

group of level-fluctuating sounds.

For comparison, in addition to the 'standard' condition containing both lateral and temporal variation, three further conditions were run with the same instructions to determine: (1) *Lateral weights* by just picking random sound-pressure levels in the left and right ear, in the same manner as described above, but keeping them constant over time, i.e. across the ten 100-ms segments; (2) *Temporal weights* by randomly varying level from segment to segment, but presenting the sounds diotically, i.e. without lateral variation; and (3) *Monaural weights* by randomly varying the levels at one ear at a time (right or left). These conditions were not mixed, but rather run in blocks of 100 trials each. Five blocks of the same kind constituted a session. Sessions in which data on the additional conditions were collected were interspersed with the standard condition containing both lateral and temporal variation by counterbalancing orders across subjects. In total, each subject completed 3000 trials to determine 'lateral-temporal' weights (standard condition), 400 trials to determine lateral weights, 1000 trials to determine temporal weights, and 300 trials each in the left and right ear to determine monaural weights. That was accomplished in ten sessions per listener, typically run on different days, and requiring approximately 30 min each.

3 Results

3.1 COSS functions

An initial, descriptive analysis of the results may be performed by constructing COSS functions for individual listeners in selected experimental conditions. COSS (conditional on a single stimulus) functions are psychometric functions that, in the present case, depict the proportion of 'loud' jugments as a function of the random level fluctuation in a given experimental condition, say a particular temporal segment and lateral position. Figure 2 shows COSS functions for listener SD, as a function of the level in the second 100-ms segment, when the levels in the two ears covary (purely temporal condition).

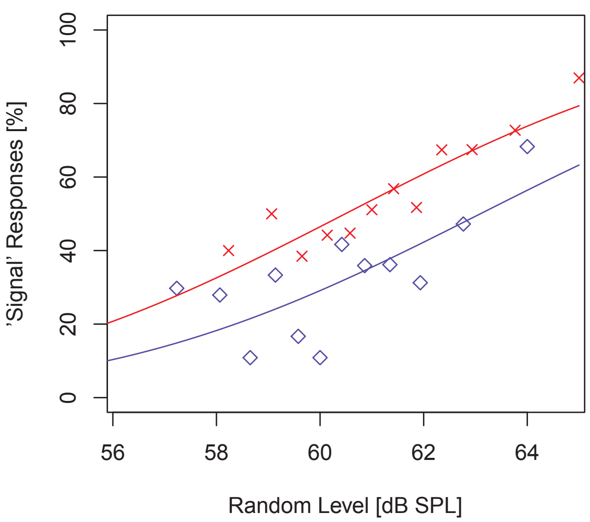

Figure 2: COSS function of listener SD for the second noise segment in the 'purely temporal' condition in which the levels in the two ears were the same. Blue diamonds: 'noise' trials; Red crosses: 'signal' trials. Cumulative normal distributions were fitted to the data.

Figure 3 shows the proportion of 'signal' (or 'loud') responses as a function of the level in the left and right ear, respectively. The figure refers to the 'purely lateral' condition for which temporal variation was absent, so the level remained constant (in each ear) across all 10 temporal segments. It appears that the data depicted show greater sensitivity in that listener's left ear (indicated by the steeper slopes of the psychometric functions) than in his right ear.

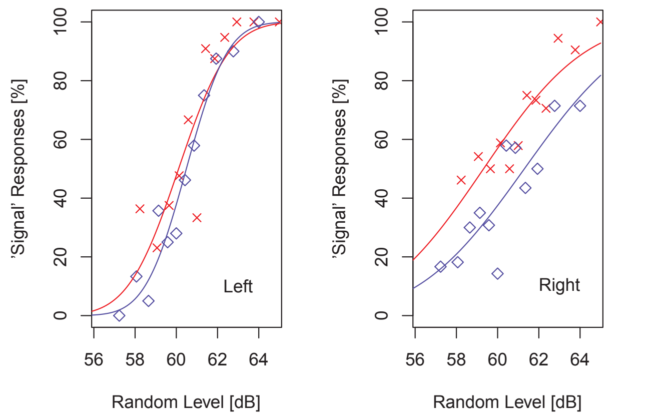

Figure 3: COSS functions of listener WE when the levels in the left and right ear were varied without changing over time ('purely lateral' condition). Left panel: as a function of the sound pressure level in the left ear. Right panel: as a function of the level in the right ear. Blue diamonds: 'noise' trials; Red crosses: 'signal' trials.

3.2 Perceptual weights

While the COSS functions provide an initial indication as to how specific stimulus features influence the overall classification of the sound as 'loud' or 'soft,' the determination of perceptual weights from the entire set of data accumulated in a given experimental condition was accomplished by multiple regression analysis. To that effect, an attempt was made to predict the (global) decision on each trial (0 for 'noise,' and 1 for 'signal') from the 20 sound pressure levels (in 10 temporal segments each at the 2 ears) entering into the construction of each noise sample to be judged. Given the metric quality of the predictors and the binary nature of the criterion, a multiple logistic regression was performed, relating the probability of making a 'loud' judgment to the decibel levels of each of the temporal (or lateral) segments x_i, their perceptual weighting w_i, and a decision criterion c:

$$p(\text{“}loud\text{''} \mid \mathbf{w}, c, \mathbf{x}) = \frac{1}{1 + e^{c - \sum_i w_i x_i}} \qquad (1)$$

Performing the statistical analysis results in 20 regression coefficients to be interpreted as perceptual 'weights'. For ease of comparison, these weights were normalized to sum up to 1.

Figure 4 shows participant WE's weights for the left and right ear as a function of the temporal segment (100-ms portion 1 to 10) of the 1-s noise burst. The most striking feature is that the weights given to each stimulus portion tend to fall from beginning to end. The other interesting feature is that this particular observer (as might have been suspected from the 'lateral' weights depicted in Figure 3) appears to attend more strongly to his left ear in arriving at an overall loudness judgment.

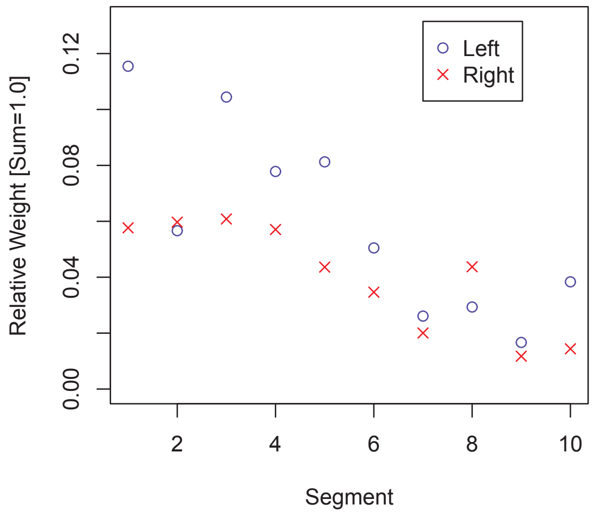

Figure 4: Lateral-temporal weights for listener WE. Red crosses represent regression weights for the right-ear components of the dichotic, time-varying noises, blue circles represent weights for their left-ear components.

Figure 5 shows mean lateral-temporal weights for all 5 listeners participating. They show the same 'primacy' effect observed in the literature [4, 6] for diotic sounds, and - on average - hardly any lateral preference.

4 Conclusions

By applying a 'molecular psychophysics' analysis to the loudness classification of short time-varying, dichotically shifting noise samples, a 'primacy effect' with respect to temporal weighting, as commonly found in the research literature, was observed. Furthermore, idiosyncratic 'lateral weights' favouring the left or right ear were determined for individual listeners. Since these are not related to absolute thresholds, they may reflect a higher attentional weighting of one side when listening to binaural sounds. Further analysis of data collected using the present paradigm might revewal implications for models of binaural summation such as Moore and Glasberg's [7] recent 'inhibition' model.

Acknowledgements

We would like to thank Angelika Kern for writing a number of data-analysis scripts for the present research.

References

[1] V.P. Sivonen, W. Ellermeier, Binaural Loudness, In: M. Florentine, A. Popper, R. Fay (Eds.) *Springer Handbook of Auditory Research: Loudness* **37**, 169–197. Springer, New York (2011).

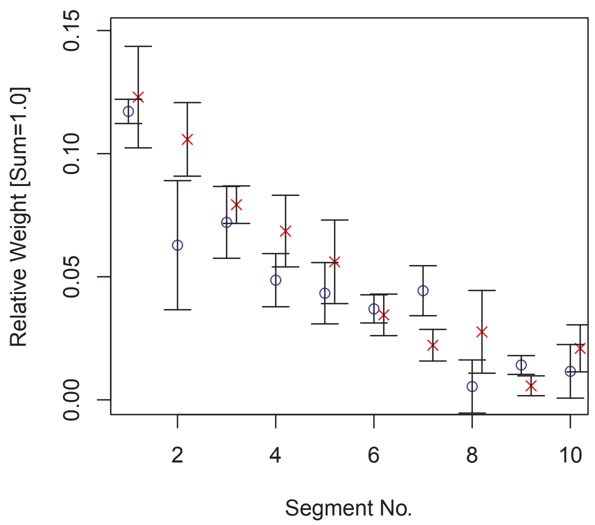

Figure 5: Mean lateral-temporal weights of all 5 listeners along with standard errors, averaged across 'signal' and 'noise' trials.

[2] J. Schlittenlacher, W. Ellermeier, J. Arseneau, Binaural loudness gain measured by simple reaction time, *Attention, Perception & Psychophysics* (in press).

[3] B.G. Berg, Analysis of weights in multiple observation tasks, *Journal of the Acoustical Society of America* **86**, 1743–1746 (1989).

[4] B. Pedersen, W. Ellermeier, Temporal weights in the level discrimination of time-varying sounds, *Journal of the Acoustical Society of America* **123**(2), 963–972 (2008).

[5] B.G. Berg, D.M. Green, Spectral weights in profile listening, *Journal of the Acoustical Society of America* **88**, 758-766 (1990).

[6] D. Oberfeld, W. Heeren, J. Rennies, J. Verhey, Spectro-Temporal Weighting of Loudness, *PLoS One* **7**(11), (2012).

[7] B.C.J. Moore, B.R. Glasberg, Modeling binaural loudness, *Journal of the Acoustical Society of America* **121**, 1604–1612 (2007).