In: Elliott, M. A., Antonievié, S., Berthaud, S., Mulhacy, P., Bargary, B., Martyn, C., & Schmidt, H. (Eds.) (2009). *Fechner Day 2009. Proceedings of the 25th Annual Meeting of the International Society for Psychophysics*. Galway, Ireland: International Society for Psychophysics.

THE TEMPORAL WEIGHTING OF THE LOUDNESS OF TIME-VARYING SOUNDS REFLECTS BOTH SENSORY AND COGNITIVE PROCESSES

Daniel Oberfeld
Department of Psychology, Johannes Gutenberg-Universität Mainz, 55099 Mainz, Germany
oberfeld@uni-mainz.de

Abstract

In two experiments, listeners judged the loudness of sounds consisting of 10 contiguous 100-ms wide-band noise segments. On each trial, the sound pressure levels of the segments were drawn independently from a normal distribution. Analyses of the trial-by-trial data were used to estimate temporal perceptual weights ("molecular psychophysics"). In Experiment 1, three level profiles were presented (flat, and increasing or decreasing over the first three segments). The temporal weights showed a primacy effect and an increase of the weights with mean level, compatible with previous results [Oberfeld, Canad. J. Exp. Psych. 62, 24-32 (2008)]. Trial-by-trial feedback had no significant effect on these patterns of weights, indicating that the potential for top-down control is limited. In Experiment 2, weights for a flat level profile were compared to weights for profiles with a 3, 6, or 9 dB increase in level during the first three segments. The weights in the latter two conditions differed significantly from the weights for the flat profile. It is demonstrated that these results indicate two independent mechanisms: a primacy/recency weighting pattern compatible with processing of the segment levels as serially sorted information, and an increase of the weights with mean level that could be due to either the specific sensory continuum underlying the decision process, or to selective attention to the louder elements.

Building on the pioneering work of G. T. Fechner and S. S. Stevens, an excellent understanding of the loudness of simple "laboratory type" sounds has been achieved (cf. Glasberg & Moore, 2006; Scharf, 1978). The spectrum and/or the sound pressure of such simple sounds (e.g., a sinusoid or a burst of wideband noise) typically remain relatively constant across the presentation duration. For the loudness of "dynamic" sounds changing temporally, spectrally, or on both dimensions during presentation, just as many environmental sounds do, a much smaller amount of data is available. Technical measures proposed as estimates of the loudness of fluctuating sounds, as for example the energy-equivalent level of a steady sound (L_{eq}) or the 95th percentile of the loudness distribution N_5 (cf. Zwicker & Fastl, 1999) assume that all temporal portions of a sound contribute *equally* to overall loudness. Recent studies using level-fluctuating noise stimuli which remained constant in spectrum but changed in level every 100 ms or so showed, however, that this conjecture is not correct. Listeners' judgments of the global loudness of a level-fluctuating noise are more strongly influenced by the first 100-300 ms of a sound than by its middle portion (Dittrich & Oberfeld, in press; Ellermeier & Schrödl, 2000; Pedersen & Ellermeier, 2008). In other words, the temporal weighting of loudness shows a primacy-effect like pattern, and to a weaker extent also a recency effect (Pedersen & Ellermeier, 2008). This weighting pattern differs from the behavior of an ideal observer, who would apply identical weights to all temporal portions of a wideband sound (Berg, 1989). Dittrich and Oberfeld (in press) demonstrated that the prediction of loudness can be improved significantly by taking into account this non-uniform temporal weighting pattern.

A potential explanation for the primacy effect would be that due to the abrupt onset of the noise, attention is directed to the beginning of the sound, in the sense of an orienting response. In this line of thinking, Oberfeld and Plank (2005) introduced a gradual increase in level (fade in) over the first few hundred milliseconds (see Fig. 1, left panel). However, instead of an approximately uniform pattern of weights, a *delayed primacy effect* was observed. The weights assigned to the attenuated fade-in part were close to zero, and the maximum weight was assigned to the first segment presented at the full level (Oberfeld, 2008a, 2008b; Oberfeld & Plank, 2005). In contrast, for noise stimuli beginning with an "inverse fade in", that is, with a gradual decease in level over the first segments (Fig. 1, right panel), Oberfeld (2008b) reported the maximal weight to be assigned to the first segment.

The objectives of the present study were (*a*) to find out to which extent the temporal weights are under top-down control and (*b*) to determine the minimum deviation from a flat level profile resulting in a change in the weighting pattern. Equally important, the results of both experiments were analyzed to see whether the patterns of weights could in part be explained by listeners' use of a sensory continuum/decision variable giving higher weight to louder elements.

Experiment 1: Effects of trial-by-trial feedback?

Experiment 1 was a within-subjects comparison of temporal weights in the presence versus the absence of trial-by-trial feedback, in order to examine for which level profiles listeners are able to adjust their weights towards the optimal set of weights (Berg, 1989). For a flat level profile, (Pedersen & Ellermeier, 2008) reported evidence for such an effect. It is conceivable, however, that the primacy effect observed with a flat level profile is caused by a different mechanism than the increase of the weights with mean segment level observed in the fade-in and inverse fade-in condition (see General Discussion). The latter mechanism might not be under top-down control, so that feedback would not alter the pattern of weights for the respective level profiles.

Method

Eight normally hearing listeners (6 women, 2 men, age 20-30 years) participated voluntarily for course credit. The stimuli were presented to the right ear via Sennheiser HDA 200 headphones. The experiment was conducted in a single-walled sound-insulated chamber.

In a two-interval forced-choice task, two noises consisting of ten contiguous 100-ms segments were presented, separated by a silent interval of 500 ms (Fig. 1, left panel). On each trial and for each interval, the sound pressure levels of the ten temporal segments were drawn independently from a normal distribution. The task was to select the louder noise. In such a setting, the *perceptual weight* is defined as the relative influence that the level of a given temporal segment had on the decision of the listener. These weights can be estimated from the trial-by-trial data by means of "molecular" analyses (Ahumada & Lovell, 1971; Berg, 1989).

For the *flat level profile* (Fig. 1, right panel), in the interval containing the less intense noise the mean of the distribution was $\mu_L = 49.5$ dB SPL and the standard deviation (*SD*) was 2.0 dB. In the interval containing the more intense noise, the mean was $\mu_H = 50.5$ dB SPL, also with *SD* = 2.0 dB. The noise sampled from the "louder" distribution was presented in interval 1 or interval 2 with identical a priori probability. For the *inverse fade-in condition*, the level of the first three segments was increased by 15, 10, and 5 dB, respectively, after the levels had been drawn from the same distributions as in the flat condition. For the *fade-in condition*, the first three segments were attenuated by 15, 10, and 5 dB, respectively, and the means of the level distributions were increased ($\mu_L = 59.5$ dB SPL, $\mu_H = 60.5$ dB SPL) so that the mean level of the softest segment was at least 30 dB above threshold. For each listener, sessions with and without trial-by-trial feedback alternated. The feedback was given on the basis of the interval containing the higher average level in the current trial (Oberfeld & Plank, 2005; see Pedersen & Ellermeier, 2008 for a discussion of alternative types of feedback in this task). In each session, two 105-trials blocks of each of the three level profiles were presented, in randomized order. 840 trials were collected for each Level Profile × Feedback combination.

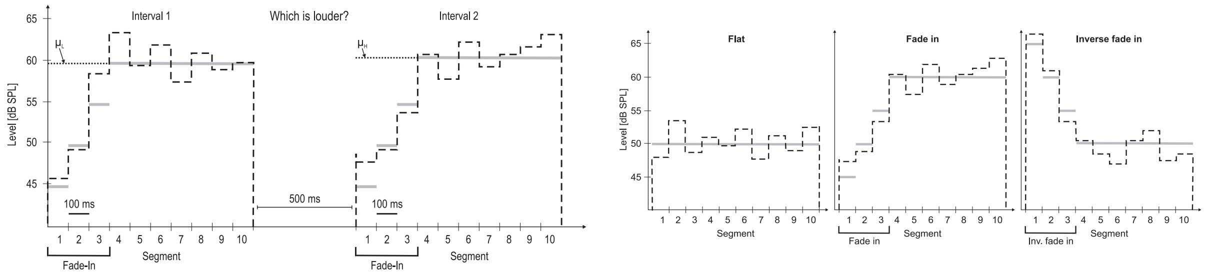

Fig. 1. Left panel: Example trial from the fade-in condition in Experiment 1. Gray lines show mean segment levels. Dashed lines represent the actual levels. The noise sampled from the "louder" distribution (mean μ_H) was randomly presented in interval 1 or interval 2. Right panel: Level profiles presented in Experiment 1.

Multiple binary logistic regression (PROC LOGISTIC, SAS 9.2) was used to estimate the weights from the trial-by-trial data (Dittrich & Oberfeld, in press; Pedersen & Ellermeier, 2008). For each trial and each segment ($i = 1 \ldots 10$), the difference between the level of segment i in interval 2 and the level of segment i in interval 1 was computed ($\Delta L_i = L_{i2} - L_{i1}$). The binary responses served as the dependent variable, and the ten within-trial segment level differences served as predictors. This analysis assumes that listeners use the decision variable

$$D(\Delta \boldsymbol{L}) = \left(\sum_{i=1}^{10} w_i \, \Delta L_i \right) - c \qquad (1)$$

where w_i is the weight assigned to the i^{th} segment, c is a constant, and the listener responds that the loud noise had been presented in interval 2 if $D(\Delta \boldsymbol{L}) > 0$. Due to the difference in mean level between the two intervals, the within-trial segment level differences were correlated. To avoid problems with multicollinearity, separate logistic regression analyses were conducted for the trials in which the noise with the higher mean level (μ_H) occurred in interval 1, and for the trials in which the position of the noise with mean level μ_H was interval 2. Thus, a logistic regression was conducted for each factorial combination of subject, level profile, feedback, and position μ_H. The regression coefficients for the ten segment level differences were taken as weight estimates. The weights were normalized such that the sum of the ten absolute values was unity (Oberfeld, 2008a).

Results and discussion

The average relative temporal weights for the three level profiles are displayed in Fig. 2. A repeated-measures ANOVA using a univariate approach and the Huynh-Feldt correction for the *df*s showed a significant effect of segment, $F(9, 63) = 48.5$, $p < .001$, $\tilde{\varepsilon} = .39$, compatible with the expected non-uniform temporal weighting patterns. The Segment × Level Profile interaction was also significant, $F(18, 126) = 48.5$, $p < .001$, $\tilde{\varepsilon} = .39$, reflecting the influence of the level profile on the temporal weights. As in previous experiments (Oberfeld, 2008a, 2008b; Oberfeld & Plank, 2005) a primacy effect was observed for the flat profile, a delayed primacy effect in the fade in condition, and a very strong weight on the first segment in the inverse fade in condition.

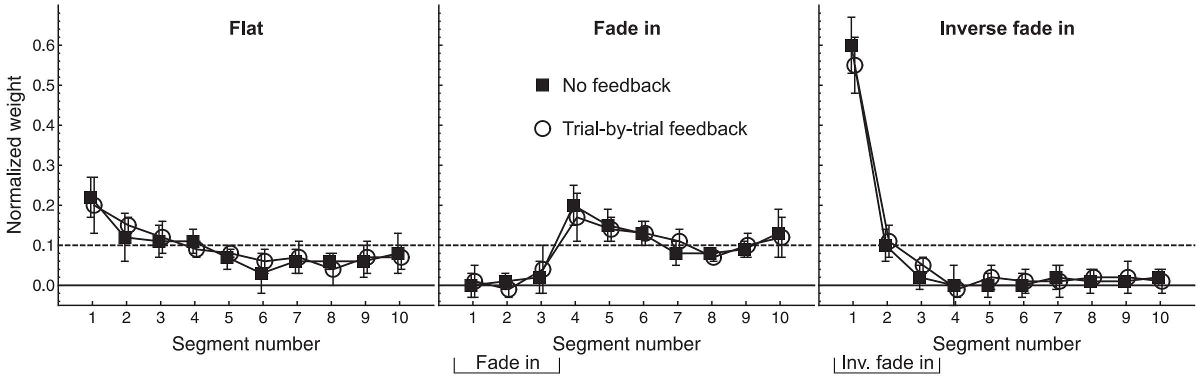

Fig. 2. Mean normalized weights in Experiment 1 as a function of segment number. Panels represent level profiles. Squares: No feedback. Circles: Trial-by-trial feedback. Error bars show 95% confidence intervals.

However, the effect of feedback was not significant, $F(1, 7) = 3.1$, $p = .12$, and all interactions with feedback also failed to reach significance ($p > .05$). Post-hoc separate ANOVAs conducted for each level profile showed that feedback did not have an effect at any level profile. Thus, listeners did not use trial-by-trial feedback to adjust their weights closer to the optimal uniform weighting pattern (Berg, 1989). This result differs from findings by Pedersen and Ellermeier (2008) who reported trial-by-trial feedback to result on average in more uniform weights, for a noise with a flat level profile. However, only two of the five listeners in their group of listeners receiving feedback showed a clear absence of a primacy effect.

Experiment 2: The minimum deviation from a flat profile resulting in altered weights

Experiment 1 again showed that deviations from a flat level profile result in an altered set of temporal weights. The purpose of Experiment 2 was to determine the minimum change in mean segment level eliciting this effect. A flat level profile was contrasted with three fade in conditions, which differed in the attenuation of the first three segments.

Method

Seven normally hearing listeners (5 women, 2 men, age 19-31 years) participated voluntarily for course credit. The same apparatus, stimuli, and procedure as for the flat and the fade in condition in Experiment 1 were used, with three exceptions. First, no trial-by-trial feedback was provided. Second, three different types of fade in were presented, with the mean segment level increasing across the first three segments in steps of 1 dB (3-dB fade in), 2 dB (6-dB fade in), or 3 dB (9-dB fade in). Third, the means of the two level distributions were $\mu_L = 59.5$ and $\mu_H = 60.5$ dB SPL for all level profiles. 700 trials were collected per listener and level profile.

Results and discussion

Fig. 3 shows the mean normalized weights for the four level profiles. Three separate repeated-measures ANOVAs with the within-subjects factors level profile and position μ_H were conducted, comparing each fade-in condition to the flat level profile. For the 6-dB and the 9-dB fade in, a significant Segment × Level Profile interaction indicated that the weights differed from the weights for the flat level profile [$F(9, 54) = 5.8$, $p < .001$, $\tilde{\varepsilon} = .78$, and $F(9, 54) = 10.4$, $p < .001$, $\tilde{\varepsilon} = .93$, respectively]. For the 3-dB fade in, the interaction was not significant, however, $F(9, 54) = 1.3$, $p = .27$. Thus, level steps between 1 dB and 2 dB are necessary to cause a significant change in the weighting pattern. As the confidence intervals in Fig. 3 show, weights on the fade in segments not significantly different from zero were observed only for the 9-dB fade in.

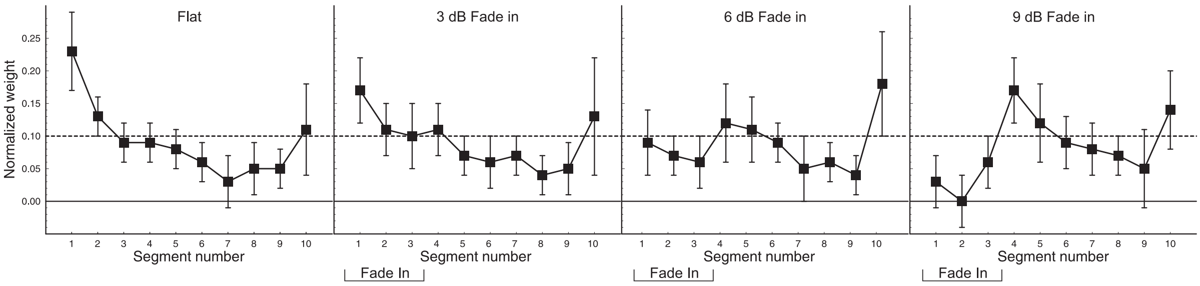

Fig. 3. Mean normalized weights in Experiment 2 as a function of segment number. Panels represent level profiles. Error bars show 95% confidence intervals.

General discussion

The data from Experiment 1 and Experiment 2 are compatible with the suggestion by Oberfeld (2008b) that the temporal weights for loudness are the result of two different processes. The primacy/recency weighting pattern could be explained by processing of the perceived segment intensities as serially sorted information (Dittrich & Oberfeld, in press), analogously to item lists in working memory (e.g., Postman & Phillips, 1965), thus representing a "cognitive" effect. The effects of attenuating or amplifying certain segments, on the other hand, could be due to segments with higher mean level having a stronger impact on the decision variable. For instance, the decision variable might not be based on within-trial differences in sound pressure level, as Eq. (1) assumes, but instead on differences in intensity ($I_{i2} - I_{i1}$), loudness ($N_{i2} - N_{i1}$), or a power function of loudness ($N^k_{i2} - N^k_{i1}$). The weights estimated above could be due to listeners' use of a sensory continuum which has a higher slope (*re*. sound pressure level) at higher levels. In such a case, the same difference in dB between interval 2 and interval 1 would have a stronger impact on the decision variable at a higher mean segment level, representing a "sensory" mechanism.

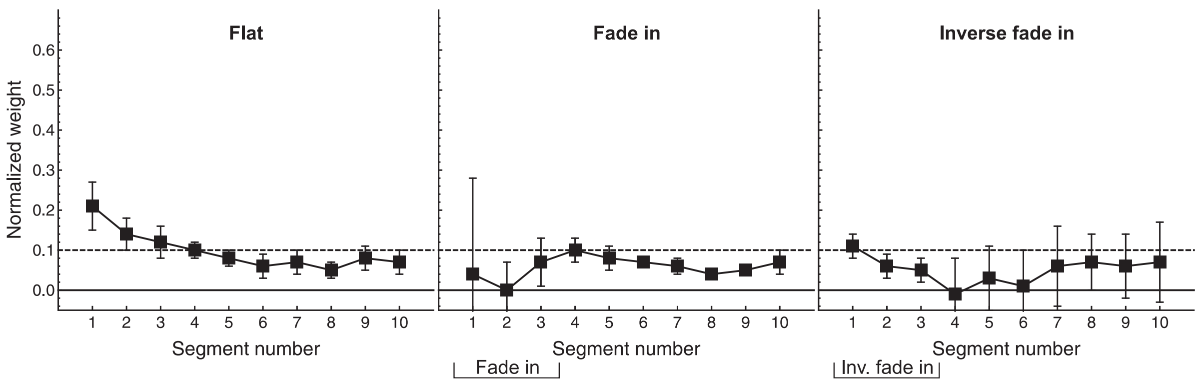

Fig. 4. Re-analysis of the data from Experiment 1: Mean normalized weights with the decision variable based on $N^k_{i2} - N^k_{i1}$ (see text). Panels represent level profiles. Error bars show 95% confidence intervals.

To answer this question, the data from Experiment 1 were re-analyzed using a decision variable based on $N^k_{i2} - N^k_{i1}$ rather than on ΔL_i. First, the individual relative temporal weights for the flat level profile estimated above were taken, aggregated across feedback and position μ_H. Subsequently, a logistic regression model for the relation between the binary response and $N^k_{i2} - N^k_{i1}$ (with N estimated via the Glasberg & Moore, 2006 model) was fitted to the data from the fade in and the inverse fade in condition simultaneously, using SAS PROC

NLMIXED. The intercept was allowed to vary between level profiles and position μ_H. Across listeners, the best fitting exponents k ranged between 2.7 and 3.7 (M = 3.2, SD = 0.36). On a side note, the third power of loudness is approximately proportional to intensity in the range of levels presented here.

The logistic regression analyses reported in the results section for Experiment 1 were then repeated with ΔL_i in Eq. (1) replaced by $N^k_{i2} - N^k_{i1}$, and using the individual exponents. Fig. 4 shows mean estimated temporal weights. If the effect of the level profile was accounted for completely by selecting the appropriate decision variable, as assumed by the two-process hypothesis, then the relative temporal weights should be identical for all level profiles. In fact, despite some descriptive differences between the patterns of weights, the Segment × Level Profile interaction was not significant, $F(18, 126) = 1.96$, $p = .10$, $\tilde{\varepsilon} = .28$, unlike in the original analyses. At any rate, the differences between the weighting patterns for the three level profiles were much smaller than in the original analysis based on ΔL_i (Fig. 2). A similar re-analysis of the data from Experiment 2 produced comparable results but is not reported here due to lack of space.

In summary, the results are compatible with the idea that the temporal weights for global loudness judgments of level-fluctuating sounds represent the processing of the segment levels as serially-sorted information, plus a stronger influence of temporal portions with a higher mean level on the decision. For the latter process, it remains for future experiments to show whether the effect is indeed due to the sensory continuum on which the decision variable is based on, or caused by attention to the loudest elements within a sound (Berg, 1990; Lutfi & Jesteadt, 2006).

References

Ahumada, A., & Lovell, J. (1971). Stimulus Features in Signal Detection. *Journal of the Acoustical Society of America, 49*(6), 1751-1756.

Berg, B. G. (1989). Analysis of Weights in Multiple Observation Tasks. *Journal of the Acoustical Society of America, 86*(5), 1743-1746.

Berg, B. G. (1990). Observer efficiency and weights in a multiple observation task. *Journal of the Acoustical Society of America, 88*(1), 149-158.

Dittrich, K., & Oberfeld, D. (in press). A comparison of the temporal weighting of annoyance and loudness. *Journal of the Acoustical Society of America*.

Ellermeier, W., & Schrödl, S. (2000). Temporal weights in loudness summation. In C. Bonnet (Ed.), *Fechner Day 2000. Proceedings of the 16th Annual Meeting of the International Society for Psychophysics* (pp. 169-173). Strasbourg: Université Louis Pasteur.

Glasberg, B. R., & Moore, B. C. (2006). Prediction of absolute thresholds and equal-loudness contours using a modified loudness model. *Journal of the Acoustical Society of America, 120*(2), 585-588.

Lutfi, R. A., & Jesteadt, W. (2006). Molecular analysis of the effect of relative tone level on multitone pattern discrimination. *Journal of the Acoustical Society of America, 120*(6), 3853-3860.

Oberfeld, D. (2008a). Does a rhythmic context have an effect on perceptual weights in auditory intensity processing? *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie experimentale, 62*(1), 24-32.

Oberfeld, D. (2008b). Temporal weighting in loudness judgments of time-varying sounds containing a gradual change in level. [Meeting abstract]. *Journal of the Acoustical Society of America, 123*(5), 3307.

Oberfeld, D., & Plank, T. (2005). Temporal weighting of loudness: Effects of a fade in. In H. Fastl & M. Frühmann (Eds.), *Fortschritte der Akustik: Plenarvorträge und Fachbeiträge der 31. Deutschen Jahrestagung für Akustik DAGA 2005, München* (pp. 227-228). Berlin: Deutsche Gesellschaft für Akustik.

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

Postman, L., & Phillips, L. W. (1965). Short-Term Temporal Changes in Free-Recall. *Quarterly Journal of Experimental Psychology, 17*(2), 132-138.

Scharf, B. (1978). Loudness. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of Perception, Volume IV. Hearing* (Vol. IV, pp. 187–242). New York: Academic Press.

Zwicker, E., & Fastl, H. (1999). *Psychoacoustics: facts and models* (2nd updated ed.). Berlin: Springer.