Temporal weighting in loudness perception: effect of bandwidth

Jesko L. Verhey[1], Jan Rennies[1,2]

[1] *Institute of Physics, University of Oldenburg, Germany, Email: {jan.rennies, jesko.verhey}@uni-oldenburg.de*

[2] *Fraunhofer IDMT, Hearing, Speech and Audio Technology, Oldenburg, Germany, Email: rns@idmt.fraunhofer.de*

Introduction

Recent studies investigated perceptual weights in the context of loudness experiments. One of the main results was that temporal portions of a broadband noise stimulus do not contribute equally to the perception of loudness. Rather, the onset of the stimulus has a significantly larger weight when the overall loudness is judged, i.e. a primacy effect is observed [1, 2]. The present study examines if the higher perceptual weight at stimulus onset depends on bandwidth by repeating a previous experiment with both broadband and narrowband noise. If the primacy effect is reduced in the narrowband condition, then the effect may be related to spectral loudness summation (SLS). SLS has been shown to depend on duration and is usually larger for short stimuli (e.g. [3]).

Neither the primacy effect nor the dependence of SLS on duration is accounted for in commonly used current loudness models. Rennies et al. [4] presented a modified loudness model, which was able to quantitatively predict SLS for short and long signals. The consequences of such a modified model developed for temporal effects in SLS are discussed in the light of the observed primacy effect.

Experiment

Stimuli

The temporal properties of the stimuli were similar to the ones used by [1]. They used white noise of 1 s duration and varied the level every 100 ms. The levels for each of the ten segments were randomly chosen from a normal distribution with a mean value of either 61 (“signal”) or 60 dB SPL (“noise”). The standard deviation was 2 dB in both cases. The same parameters of level distributions, segments and duration were used in the present study. In order to have the spectral properties of the stimuli comparable to those in the study on duration-dependence of spectral loudness summation by [3], the bandwidth and geometric center frequency for the broadband stimulus were set to 6400 Hz and and 2 kHz, respectively. This corresponded to the broadest spectral condition used in their study. In addition to the broadband stimulus, a narrowband stimulus with the same center frequency and a bandwidth of 400 Hz was used. The remaining stimulus parameters were the same as in the broadband condition. For the smaller bandwidth of the noise, slow intrinsic envelope fluctuations could have disturbed the random levels assigned to the individual temporal segments. In order to minimize the influence of the intrinsic fluctuations, double-iterated low-noise noise was used. It was generated dividing a portion of Gaussian noise by its Hilbert envelope on a sample-by-sample basis and a subsequent restriction to the original bandwidth. This procedure was repeated once. Each level step in the stimulus was gated with a 2.5-ms $\cos^2$-ramp, as were stimulus onset and offset. Figure 1 shows an example of time course (top row) and random level (bottom row) of a “noise” for both the broadband (left) and the narrowband condition (right).

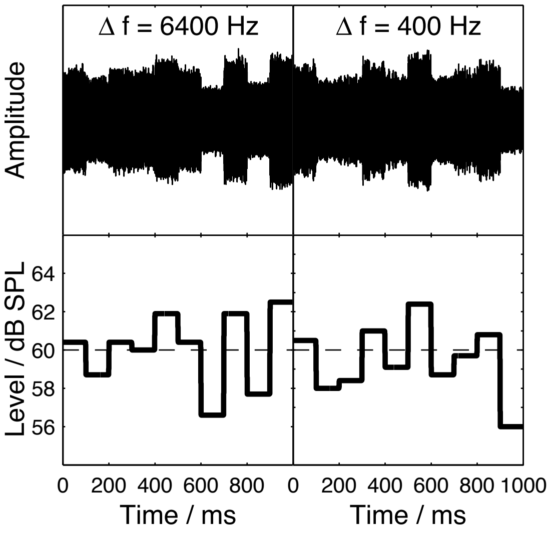

Figure 1: Example time course of a ”noise” stimulus (top panels) and the corresponding random levels (bottom panels) for a bandwidth of 6400 Hz (left) and 400 Hz (right). Dashed lines indicate the mean value of the distribution from which the random levels were drawn.

Method

A two-alternative forced choice procedure was used. In each trial, the subjects heard two sounds, “signal” and “noise”, separated by 500 ms of silence. Their task was to indicate which of the two sounded louder by pressing the corresponding button on the keyboard. The intervals were highlighted on the screen during stimulus presentation. However, the subjects were unaware which interval contained the “signal” and which the “noise”. This was set randomly with equal probability for the two permutations. New noise samples and envelopes were generated for each trial; “signal” and “noise” were generated independently from each other. No feedback was provided as to which interval contained “signal” or “noise”. In the first experiment, the subjects made 3000 comparisons between “signal” and “noise” for stimuli with a bandwidth of 6400 Hz. The same procedure was used in the narrowband condition with 400-Hz wide stimuli. The subjects were given the same task as in

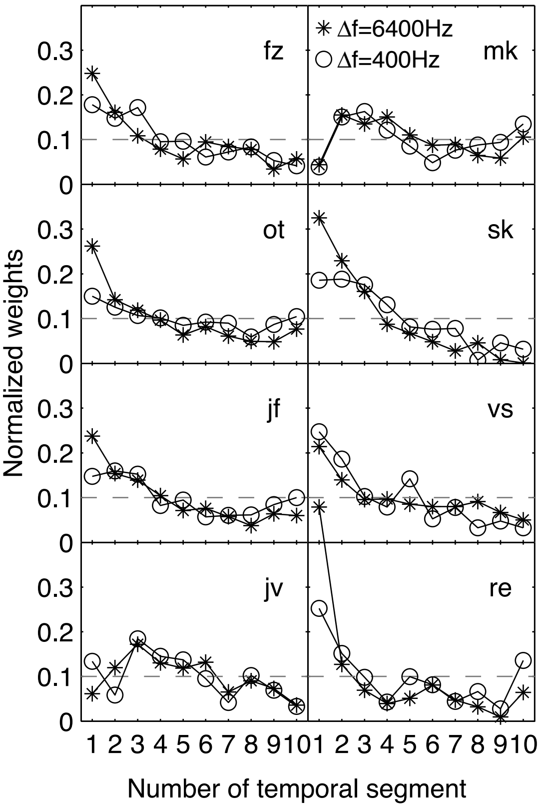

Figure 2: Individual data of eight test subjects. The normalized weights are shown for the ten temporal segments. Asterisks and circles represent data for 6400 and 400 Hz bandwidth, respectively. Dashed lines indicate equal weights for all segments.

the broadband condition. During the experiment, they were not informed about the outcome. All together, approximately ten hours of measurement time were needed for each subject.

Eight normal-hearing, naive test subjects between 22 and 26 years of age participated in the experiment. All had hearing thresholds $\leq$ 15 dB HL at standard audiometric frequencies between 125 and 8000 Hz.

The results were analyzed using the COSS-analysis(conditional on single stimulus), which allows to estimate how the individual temporal segments contribute to the overall loudness perception (see [1] for details).

Results

Figure 2 shows individual data for a stimulus bandwidth of 6400 Hz (asterisks) and 400 Hz (circles). Normalized weights ($\sum_{i=1}^{10} w_i = 1$) are shown as a function of temporal segment for each subject. For comparison, dashed lines indicate equal weights for all segments. In the broadband condition, all but two subjects (mk, jv) assigned the highest weight to the first temporal segment. For later temporal segments weights decrease. The inter-subject difference between the highest and lowest weight varies substantially, ranging from 0.11 (mk) to 0.47 (re). Some subjects tend to have higher weights at the last segments than at segments in the temporal center.

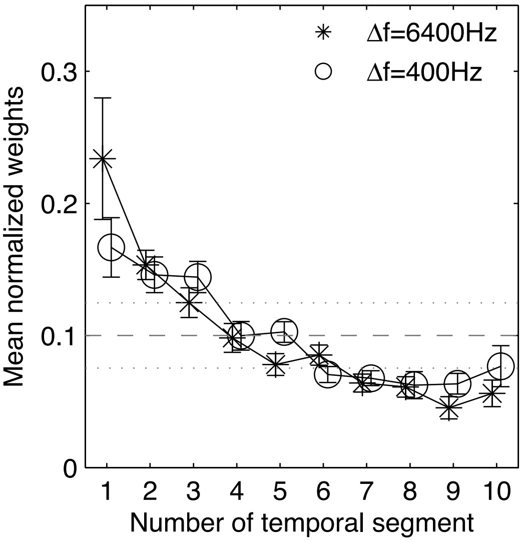

Figure 3: Mean normalized weights across all subjects for the ten temporal segments. Asterisks and circles represent data for 6400 and 400 Hz bandwidth, respectively. Error bars represent plus and minus one inter-individual standard error. The dashed line indicates equal weights for all segments, dotted lines represent 99% confidence limits for equal weights obtained from Monte-Carlo simulations.

For the smaller bandwidth, the two subjects with a lower weight at the first segment in the broadband condition show weights similar to the broadband condition. For all other subjects, the higher weighting of the onset can still be observed. However, the magnitude of the first weight is considerably reduced in all but one subject (vs) compared to the broadband condition. The differences between the highest and lowest weight are smaller than for the larger bandwidth and range from 0.09 (ot) to 0.23 (re). In general, curves for the two bandwidths are very similar except at the first temporal segment. The main effects are also reflected in the mean data over all subjects as shown in Fig. 3. The highest weights are assigned to the first temporal segment for both bandwidths, and this effect is more pronounced for a bandwidth of 6400 Hz. To investigate the significance of the obtained results, Monte-Carlo simulations were made by using a simple model which assumes equal weights of all ten temporal segments. The mean weights and 99% confidence limits of the Monte-Carlo-simulations are indicated by dashed and dotted lines in Fig. 3, respectively, and show that for both bandwidths, the mean weights and standard error ranges of the first segment are clearly outside the confidence limits such that the hypothesis of uniform weights is comfortably rejected for both bandwidths.

Discussion

The experimental data show that temporal segments of acoustic stimuli are weighted differently when forming a global loudness judgment. The stimulus onset contributes to a significantly higher degree to the overall perception. This is in line with results in [1] and [2].

However, some differences between their data and the present study are worth noting. The accentuation of the first temporal segment is more pronounced in the present study than measured in [1]. A possible source of this deviation is the feedback they provided to the subjects. Pedersen and Ellermeier [2] showed that trial-by-trial feedback indicating whether the subjects had indicated the “signal” interval as the louder one modified the weights toward a more uniform curve, i.e. the first segment was weighted less than in the no-feedback condition. The mean weight at the first temporal segment in the condition without feedback obtained in [2] was similar to the result for the broadband stimuli in the present study. The increase of the weights at the last temporal segments found in [1] and [2] was not observed in the mean data of the present study, although some subjects of the present study show a similar tendency. Considering the inter-subject variability, the different sets of test subjects most likely contributed to the differences between the studies.

Simple Monte-Carlo simulations showed that the hypothesis of uniform weights can be rejected. This is at odds with the commonly applied strategy of modeling loudness on the basis of measures such as peaks or percentiles of the time-varying loudness [2]. Thus, when modeling loudness of time-varying sounds, the onset of the stimuli should be given more weight for stimuli comparable to those in the present study. For the accentuation of the stimulus onset, no underlying mechanism was proposed by [1]. Pedersen and Ellermeier [2] suggested a “multiple look” strategy, in which listeners can weigh looks differently depending on their task. So far, however, the mechanism underlying the primacy effect remained unclear. The results of the narrowband condition in the present study indicate that the bandwidth of the stimuli plays a crucial role for the observed effect. Decreasing the bandwidth to 400 Hz significantly reduced the weight at the first segment, while weights were similar at later segments. Under the assumption that a larger weight corresponds to an increased loudness, this is in line with recent studies showing that spectral loudness summation is larger for short than for long stimuli (see e.g. [3]). These studies measured, among other things, the difference in spectral loudness summation between 1000-ms long stimuli, corresponding to the total duration of the stimuli used in the present study, and 10-ms long stimuli, a duration clearly within the first temporal segment of the stimuli in the present study. In addition, Verhey and Kollmeier [3] found a significantly higher spectral loudness summation for stimuli with a duration of 100 ms (the duration of one segment in the present study) compared to stimuli of 1 s. Thus, an increased spectral loudness summation in the beginning of the stimuli can qualitatively explain the higher weight found in the present study. Further studies are necessary to explore the dependence of larger weights at stimulus onset on overall stimulus duration and number of temporal segments.

Note that the dependence of spectral loudness summation on duration is not observed in every listener [3]. Thus, the fact that two subjects did not show a higher weight at the first segment in the present study might be explained by an absent duration dependence of spectral loudness summation in these subjects.

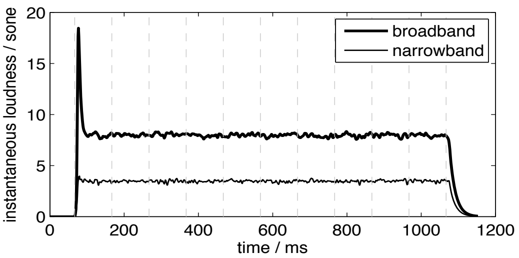

Figure 4: Instantaneous loudness calculated by the modified model proposed by Rennies et al. [4] for broadband (thick line) and narrowband noise (thin line) at a level of 60 dB SPL.

In the light of duration-dependent spectral loudness summation, the ability of current loudness models to predict onset accentuation can be newly discussed. So far, no model was published which correctly predicts the influence of duration on spectral loudness summation. In [4], a mechanism was proposed which quantitatively accounts for this effect. This mechanism is based on the assumption that loudness is larger during the onset for stimuli of larger bandwidths, but not for stimuli whose energy is confined to a single auditory filter (see Fig. 4). The reduction of onset accentuation in the present study as the bandwidth is decreased qualitatively supports the assumption of such a mechanism.

In relation to the influence of spectral loudness summation, the finding of [2] that feedback changed the weights toward a more uniform distribution might be interpreted such that listeners still have a larger spectral loudness summation at stimulus onset but that this is counteracted by a lower weight that subjects learn to give to the first segment due to the feedback.

References

[1] Ellermeier, W. and Schrödl: Temporal weights in loudness summation, Fechner Day 2000, Proceedings of the 16th Annual Meeting of the International Society for Psychophysics, S. Bonnet, C. (ed.), 2000, 169-173

[2] Pedersen, B. and Ellermeier, W.: Temporal weights in level discrimination of time-varying sounds, J. Acoust. Soc. Am., 2008, 123, 963-972

[3] Verhey, J. and Kollmeier, B.: Spectral loudness summation as a function of duration, J. Acoust. Soc. Am., 2002, 111, 1349-1358

[4] Rennies, J.; Verhey, J. and Fastl, H.: Modellierung der Dynamik der spektralen Lautheitssummation (Modeling the dynamics of spectral loudness summation), DAGA 2008 - Fortschritte der Akustik, Proceedings of the 34th Annual Meeting of the Deutsche Gesellschaft für Akustik e.V., 2008, 449-450