Discriminating temporal loudness patterns in the absence of overall level cues

Tina Plank[1], Wolfgang Ellermeier[2]

[1] *Inst. für Exp. Psychologie, Univ. Regensburg, D-93040 Regensburg, Germany, Email: tina.plank@stud.uni-regensburg.de*

[2] *Sound Quality Research Unit, Department of Acoustics, DK-9220 Aalborg, Denmark, Email: we@acoustics.dk*

Abstract

To investigate the capability of the auditory system to discriminate the temporal "shape" of short non-stationary sounds, in a One-Interval-Forced-Choice procedure, samples of white noise having a level increment either near the beginning, or near the end of the noise bursts were presented. The subjects' task was to discriminate between these two kinds of "loudness profiles". Applying COSS-analysis [1] to the data, weights were assigned to each stimulus component, yielding a measure for the strength of the influence that each segment has on the listener's decision. The weighting functions obtained assigned high positive weights to the incremented components and low, or negative weights to the non-signal components. The results confirm the assumption that the auditory system is capable of analyzing individual stimulus components selectively with high temporal resolution.

Introduction

How can listeners use information from well-defined temporal components in short sounds fluctuating in level, when they have to perform different tasks?

While former investigations focused on loudness integration of non-stationary sounds [2], we have recently started to investigate how well listeners can distinguish different "temporal loudness profiles" in a roving-level paradigm [3].

The current experiment extends this work to investigate whether listeners can discriminate temporal loudness patterns without overall level cues, and whether they can still discriminate loudness profiles, when the temporal segments they consist of have a duration of only 10 ms each.

Method

White noise signals with appropriate amplitude changes were computed in software, and D/A-converted via a Tucker Davis Technologies (TDT RP 2.1) signal processor with 24bit resolution and 50 kHz sampling rate. Subsequently they were diotically presented to subjects via (Beyerdynamic DT 990) headphones. The subjects were seated in a double-walled, sound-proof cabin.

The total duration of the stimuli varied in two conditions. In condition 1 stimuli with a total duration of 200 ms were presented, consisting of ten temporal segments, each 20 ms long. In condition 2 the stimuli were 100 ms long, consisting of ten 10 ms long temporal segments.

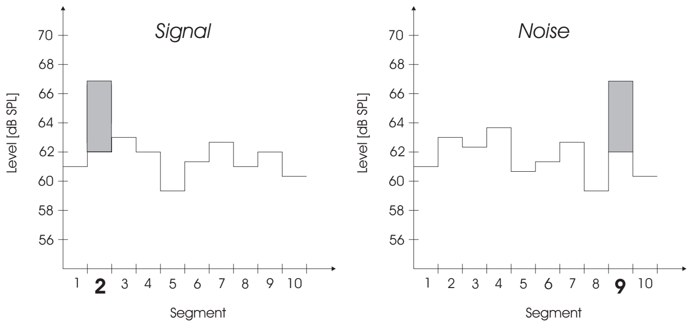

Figure 1: Schematic representation of the two kinds of stimuli presented in random order in the One-interval-forced-choice-task.

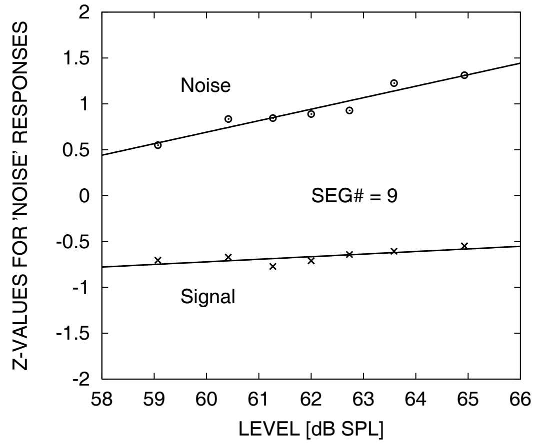

Figure 2: Example for COSS functions (subject 1's data for the 9th segment in condition 1; 200 ms total duration). The proportion of the z-values of "Noise"-responses is plotted as a function of the random level perturbations.

The sound pressure levels of the ten segments of the stimuli were independently drawn from a normal distribution with $\mu = 62$ dB SPL and $\sigma = 2$ dB. The different level patterns were defined as "Signal" and "Noise", with the "Signal" characterized by an increment on segment 2, the "Noise" by an increment on segment 9. The increments were 5 dB SPL for the 200 ms condition and 8 dB for the 100 ms condition. Figure 1 shows a schematic depiction of "Signal" and "Noise" stimuli. The sounds were presented in a One-Interval-Forced-Choice Procedure. The subjects' task was to decide whether the interval they have just heard was a "Signal" or a "Noise".

Three normal-hearing listeners at the age of 28 to 46 years (including the two authors) participated in the experiment. All listeners had experience in similar kinds of psychoacoustic tasks. After some training 6400 trials

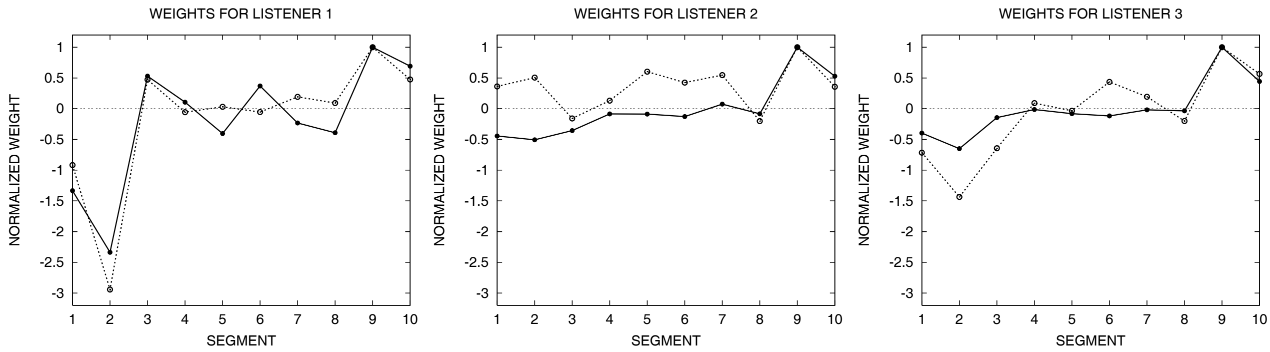

Figure 3: Temporal weights for all three listeners.

were completed by each subject, 3200 trials for the long duration condition (200 ms) and 3200 trials for the short duration condition (100 ms). The order in which sessions with long and short duration trials followed each other was counterbalanced within subjects.

In order to determine the perceptual weight with which each of the temporal segments contributes to the decision of the listener COSS-analysis [1] was chosen. To obtain conditional psychometric functions (COSS functions, COnditional on a Single Stimulus) the random level perturbations serve as independent variable. Plotting the COSS functions for each temporal segment the proportions of 'Noise'-responses show the change of a subject's judgment depending on a single stimulus component, in this case the level of the respective temporal segment. The COSS functions are fitted by a least-squares regression. Figure 2 shows an example for such COSS-functions. The data for trials with "Signal" and trials with "Noise" are analyzed separately (see upper and lower curves in figure 2). Steep COSS functions indicate a strong weight for a given segment, flat COSS functions imply that the segment has no influence on the decision of the subject.

Results

For each subject 40 COSS functions (10 segments × 2 kinds of stimuli "Signal" and "Noise" × 2 durations) were estimated. From the variance estimates of the COSS functions, the weights a_i can be calculated, which give a measure for the strength of the influence that each segment i has on the listeners decision [1].

Figure 3 shows the weights for all three subjects as a function of the ten temporal segments for long duration (200 ms) (solid line) and short duration (100 ms) (dotted line) respectively. The weight for segment 9, the incremented component in the 'Noise'-trials, is normalized to 1, the weights for the other components are computed relatively to this unit weight.

The weighting functions confirm a similar pattern for all listeners with high positive weights for segment 9 and negative weights for segment 2 in almost all conditions. But the results also show that the listeners differ in their strategies used. For listener 1, the 2nd segment has, in absolute values, a much stronger influence than segment 9. Listener 1 also builds contrasts between incremented segments and neighboring segments (see #3 and #8). For listener 2, segment 9 influences the decision the most, both in the long duration condition (200 ms) and the short duration condition (100 ms). Listener 3 relies more on segment 9 in the long duration condition, while he switches to the 2nd segment in the short duration condition, where this segment has a stronger influence than segment 9.

Conclusions

The experiment shows that listeners are able to discriminate different temporal loudness profiles, even in the absence of overall level cues. Further the weighting curves show that the listeners are able to use single temporal components selectively, even within a time frame of only 100 ms. So the results confirm the assumption that the auditory system is capable of analyzing individual stimulus components with high temporal resolution, when it is required by the task, changing temporal weighting patterns adaptively, even though with notable individual differences.

Acknowledgement

We would like to thank Benjamin Pedersen, Aalborg University, for making available and modifying his experimental and analysis programs for the purpose of this study.

References

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

[2] Ellermeier, W. & Schrödl, S. (2000). Zeitliche Gewichtung bei der Lautheitsintegration. *Fortschritte der Akustik, DAGA 2000*, 302-303. Oldenburg: Deutsche Gesellschaft für Akustik

[3] Plank, T. & Ellermeier, W. (2003). Auditive Unterscheidung von Lautheitsprofilen. *Paper presented at the 45th Annual Meeting of Experimental Psychologists (TeaP 2003)*, Kiel, Germany