Listener's preferences with respect to frequency shaping and loudness adjustments for music and speech reproduction

A. Kubiak[1,2], M. Saft [1,2,3], J. Rennies[1,2], B. Kollmeier[1,2,4]
[1] Fraunhofer IDMT Project Group Hearing, Speech and Audio Technology, Oldenburg, Germany
[2] Cluster of Excellence "Hearing4all", Oldenburg, Germany
[3] Jade Hochschule, Institute for Hearing Technology and Audiology, Oldenburg, Germany
[4] Medical Physics, Department of Medical Physics and Acoustics, University of Oldenburg, Germany

Introduction

It has been observed that there is a large variation of listening preferences, e.g., with respect to equalizer and volume settings, even among groups of subjects with clinically normal hearing [1]. The factors underlying these differences are not at all understood yet. However, only a deeper understanding of the relevant factors can lead to the development of individualized perceptual models [2] and consequently to model-based, adaptive algorithms, which may have a great potential for individual sound adjustment in high-end audio equipment and communication devices.

The goal of this study was to perform a follow up on the research described in [1], which investigated preferred loudness and equalization settings for normal-hearing subjects taking into account their hearing abilities, namely performing an audiogram and categorical loudness scaling procedure. Here the goal was to investigate the impact of three similar low-frequency environmental noises onto subjects' choices in order to check if such interferers would affect their listening preferences, as well as if the SNR ratio of the preferred playback settings would be similar between subjects and depend on the noise characteristics. The general motivation for the selection of the employed background noises was the usage trend of the audio consumers listening to the music via headphones, which occurs mainly during outdoor activities. That is why the chosen noises consist of those that may be frequently met by the subject in the urban space, namely traffic, train and crowd noise. Also music pieces were chosen so that could cover the wide variety of stimuli by originating from different genres.

Material

Fifteen normal-hearing subjects were invited to the experiment (three female, twelve male). The subjects listened to different pieces of music in silence or in background noise. The stimuli were presented as diotic .wav files at a sampling rate of 44100 Hz over headphones (Sennheiser HD650). The starting level of the music signals was equalized to 65 dB SPL in the 1/3 octave band centered at 1 kHz. The music probes contained three music pieces: Ludwig van Beethoven "Moonlight Sonata", Elvis Presley "Jail House Rock" and Bob Marley "Buffalo Soldier". Additionally, four types of background were used: silence plus three types of low-frequency noises at levels presented in brackets: crowd (72.5 dB SPL), traffic (75 dB SPL) and train (67.5 dB SPL). The passive damping characteristics of the headphones were accounted for to simulate the scenario of listening to music via headphones in the presence of environmental noise.

Method

This first study investigated the inter-individual variability of normal-hearing subjects´ preferences with respect to frequency shaping and loudness adjustments for different types of music listened via headphones. The subjects were asked to adjust the linear gain and spectral shaping to their individually preferred setting. A perceptually continuous 2D matrix was used consisting of 19x19 equalization presets assigned so that the user can adjust the preferred sound reproduction settings in real-time using the touch screen along two perceptually continuous dimensions. The participant chooses the gain of the low or high frequencies on the abscissa and the boost of the mid frequencies on the ordinate, while overall gain was controlled via a rotary knob. Additionally, compression was introduced to the X axis, settings at high input levels exceeding 80 dB SPL

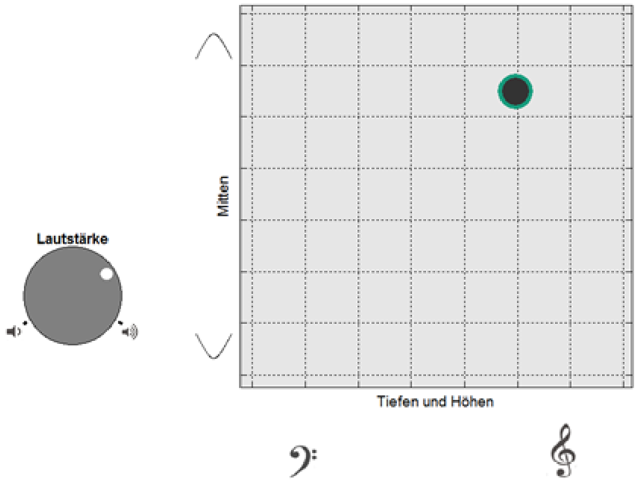

Figure 1: Graphical layout of the user interface for self-adjusting the mid-frequency gain (ordinate) and the gain of low- and high-frequency components. Overall gain was adjusted by a rotary knob.

Results

The indicated data have a large interindividual spread of preferred adjustments with respect to both loudness and equalization in the group of subjects that according to clinical research would be treated as unified and without further distinction as all the participants were normal-hearing. Meanwhile the interindividual differences observed in the results seem to exceed the influence of the music piece, or noise type. In the subsequent analyses the relation between individual listening preferences and the influence of environmental noises was investigated. The overall results are shown in Figure 2. The figure shows the linear gain adjustments of fifteen participants, each for four different background conditions as an outcome of a task, where subjects were asked to choose their preferred settings using the rotary knob. The results were strongly diversified across subjects, reaching a spread of almost 18 dB between extreme cases. Of course one can say that participants could benefit from both: the boost of the 2D Touch, which introduces narrow band dynamic changes from -15 up to 15 dB, and the knob adjustments, as both tasks were performed simultaneously (this substitution will be investigated in detail in future studies). On the other hand in the presence of noise subjects tend to reschedule not only the knob settings in order to increase SNR, but also change the equalization, as shown in Figure 3.

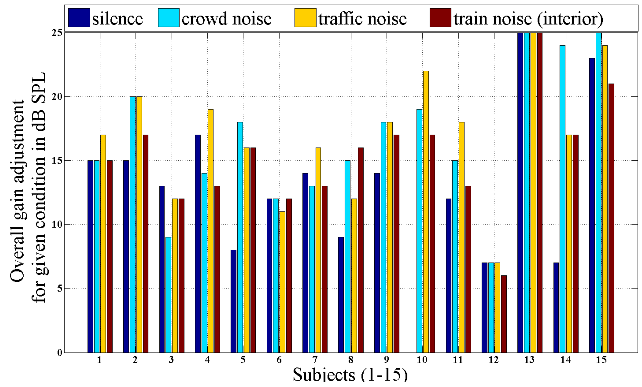

Figure 2: Overall linear gain adjustment scheme for classical music piece across all subjects and conditions. Possible gain adjustment varied from -25 to 25 dB, where zero indicates starting point of subjects' exposition

To further analyze the impact of noise type, the noisy selected presets in the silent condition were compared to the corresponding presets in the noisy conditions. In Figure 3 each symbol indicates an individual subject and the symbol's distance from the center of the plot as well as direction indicate the change introduced by the given subject on the 2D Touch matrix from its preferred adjustments in silent conditions while background changed from silence to noise. Only data for one of the three noises used (crowd noise) are shown. The most appealing trend observed in the data was that subjects tended to change the equalization settings while the background noise was introduced, but that this change seemed not to be related to any systematic trend, but widely spread and perhaps motivated by some personal rationale. What might be found interesting was that at least some of the subjects were quite consistent with their choices regarding the music piece change, although the signals differed significantly representing extremely different genres, instrumentation or recording techniques. Namely, listener indicated by black diamond chose the presents being apart of maximum one preset on the X and one on the Y axis away from each other for all the music pieces in the given noise scenario. Also the choices of the subjects indicated by azure square or pink diamond seem to be very consistent. On the contrary, subjects indicated by green cross and black circle tended to diverse their choices widely as music piece changed.

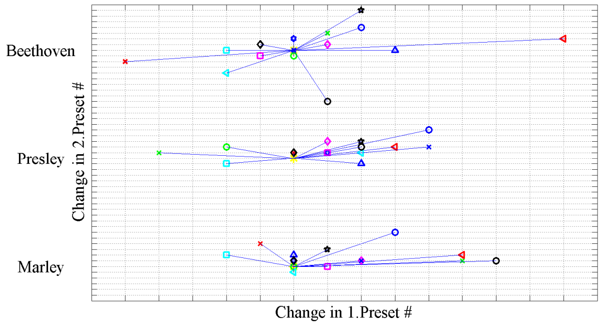

Figure 3: Change of presets for all subjects (symbols) and music pieces for noisy conditions (crowd noise) in comparison to silent condition. The distance from the center of each (yellow star symbol) represents the amount of change due to addition of the noise.

Conclusion

In summary, the main conclusion is that individually preferred frequency shaping and gain adjustment are highly personal. In general, the addition of background noise makes subjects increase the overall gain, but this effect differs considerably between subjects. Noise addition results in a change of EQ settings for the majority of subjects, but the changes are highly individual. The reasons for the large interindividual variability are still unclear and, consequently, building perceptual models on an individual basis is not yet possible. Thus, for practical applications the only way to address the individual preferences is to provide an interface to the user for self-adjusting the playback settings. The present data set will be further analyzed with respect to the interaction of preferred equalizer and gain settings.

Compared to [1], there is a trend that the addition of noise narrowed the interindividual spread, but in order to better understand the factors influencing personal listening preferences, systematic relations to personal hearing abilities (as in [1]), learning process that could occur and possibly other personal factors needs to be investigated.

Literature

[1] Saft, M., Hansen, M. und Rennies, J. (2014). Klang-präferenzen von Normalhörenden bei Musik-wiedergabe. Fortschritte der Akustik - DAGA 2014.

[2] Herrero, P., de Antonio, A., Introducing Human-like hearing Perception in Intelligent Virtual Agents, The Second International Joint Conference on Autonomous Agents &Multiagent Systems (AAMAS '03), July 14-16, 2003, Melbourne, Australia