Are absolute thresholds and loudness judgements influenced by different colours?

Daniel Menzel, Elias Faccinelli and Hugo Fastl

AG Technische Akustik, MMK, TU München, Arcisstr. 21, 80333 München, Germany
menzel@tum.de

Previous experiments showed that images of differently colored trains and sports cars can influence the loudness ratings given by subjects via free magnitude estimation. Red vehicles caused subjects to rate the loudness of simultaneously presented train or car sounds higher relative to green vehicles. To investigate whether these loudness differences correspond to shifts in absolute threshold, subjects' threshold in quiet was measured via Békésy-tracking while viewing red, green, and neutral color patches. Also, the influence of color on the loudness of broadband noise was measured using a method of adjustment, in which subjects had to adjust the level of a test sound until it was perceived as loud as a reference sound. In both cases, no influence of color on either absolute threshold or loudness perception measured via adjustment could be found. These results support the hypothesis that, compared to previously used methods like magnitude estimation, methods that require subjects to concentrate on the auditory stimulus do not seem to be as applicable for measurements of audio-visual interactions.

1 Introduction

The measurement of loudness is an essential tool used routinely in areas like noise abatement, product sound design, and sound-quality engineering. The main contributing factors for loudness perception are of course the physical properties of the sound field. Numerous studies however show that non-acoustical influences like visual stimulation can change the subjective loudness judgment and thus have to be taken into account in the context of sound-quality evaluations (Blauert and Jekosch, [1]).

For example, Namba et al. ([9]) studied the loudness of interior car sounds using a continuous rating scale. In one condition, no visual stimulus was visible, while in another condition a video of the scenery surrounding the driving car was presented. Subjects tended to avoid extreme categorical ratings in case of additional visual stimulation, i.e. loud sounds were rated as softer while soft sounds were rated as louder. These results are in line with data reported by Böhm et al. ([2]). Here, magnitude estimates of the loudness of passing trains were evaluated with and without simultaneous presentation of still and moving pictures of trains. Moving pictures caused subjects to underestimate the loudness by about 5% in comparison to the audio-only case.

The effects of audio-visual interactions on semantic parameters of environmental sounds were investigated by Suzuki et al. ([13]). Using semantic differential, it could be shown that ratings along the "volume"-factor generally decreased when subjects were presented a matching video of the sound source.

The content of images shown to subjects during loudness rating tasks can also be of importance. Employing an eleven-point categorical scale, Höger and Greifenstein ([5]) asked subjects to rate which of two trucks was louder in an experimental task involving images of small and large trucks combined with truck-sounds of different loudness. It was found that a sound associated with pictures of big trucks frequently was judged as being louder than the same sound associated with small trucks.

Patsouras et al. ([10]) performed experiments with images of differently colored passenger trains. Subjects gave higher magnitude estimates of the loudness of passing trains when an image of a red train was presented simultaneously, in comparison to light-green trains, which caused an underestimation of loudness with a relative difference of about 10%. These results could be repeated in a cross-cultural study by Rader et al. ([11]). Japanese subjects also tended to overestimate the loudness of red trains in contrast to light-green trains. An overview of these results is given in [3].

Loudness expectations could also be influenced by altering the visual stimulus. The color of sports cars for example affected loudness ratings as reported by Menzel ([7]). Here, red and dark-green cars produced higher magnitude estimates of loudness compared to light-green or blue cars. As light green or blue are rather untypical colors for a sports car, vehicles presented in these colors might not provide an immediate impression of a strong and therefore loud sports car. (cf. Menzel et al. [8])

To further investigate loudness differences associated with the presentation of colors, two different experiments were performed:

1. To see whether additional optical stimulation with selected colors can influence the absolute threshold of hearing - which in turn could cause or contribute to changes in loudness - absolute threshold measurements were conducted under different optical conditions.

2. To obtain equivalent level differences in dB, i.e. to answer the question "What level adjustment in dB is necessary so that red and green stimuli elicit the same loudness?", loudness matching via level adjustment of broadband stimuli combined with different colors was carried out.

2 Absolute threshold measurements

2.1 Setup and procedure

A Békésy audiometer was used to measure absolute thresholds of hearing from 0.2 Bark (20 Hz) to 24 Bark (15.4 kHz) in steps of 0.05 Bark, leading to a resolution of 5 Hz at the lower end of the frequency range and 278 Hz at the upper end. A gated test tone was presented diotically through calibrated free field equalized electrodynamic headphones Beyer DT 48 in a sound proof booth (cf. [12], [4]). When the tone was audible, a button had to be pressed which caused the test tone level to decrease. When the tone became inaudible, the button had to be released and the test tone level started to increase again. The average between the points of audibility and inaudibility over the abovementioned frequency range then yielded the absolute threshold of hearing.

Colored sheets of cardboard (width: 68 cm, height: 50 cm) were used as visual stimuli. The colors red and light green were chosen for this experiment, as these colors showed the strongest influence on loudness in previous studies. The sheets could be fixed to a wall of the sound proof booth and were positioned to fill a large part of the field of vision of the subjects. Standard incandescent lighting was used to illuminate the sheets and care was taken to ensure constant lighting conditions throughout the experiment. The wall of the booth, which had a neutral beige color, provided a baseline visual condition to test for any irregularities in the subjects' hearing.

Twenty-one subjects participated in this experiment. Each subject completed three threshold measurements under the visual conditions red (R), green (G), and neutral in random order. The data of two subjects could not be used in the following analysis due to failure to correctly follow the experimental instructions. The remaining 19 subjects were aged 24 to 28 years (median 24 years).

2.2 Results

For each subject, the absolute threshold obtained under the green visual condition ($L_{TH,G}$) was subtracted from the absolute threshold measured under the red visual condition ($L_{TH,R}$). The results are given in Fig. 1.

Regarding the median shown in Fig. 1 as bold curve, there seems to be no difference between $L_{TH,R}$ and $L_{TH,G}$ throughout the examined frequency range. Variations of about ± 2 dB (indicated by the interquartiles) suggest that the subjects' absolute thresholds were rather similar in the range from 2 to 15 Bark, while greater variability of up to ± 5 dB can be seen above 15 Bark.

Statistical treatment via repeated measures analysis of variance (ANOVA) with critical-band rate and color as within-subject factors shows no significant effect of color. A closer look at the individual thresholds shows that no subject was uniformly influenced by any of the two colors over the complete frequency range.

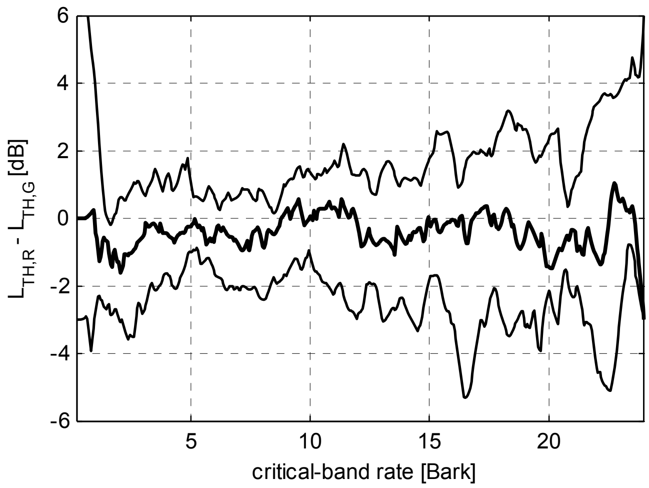

Fig. 1 Difference between absolute thresholds measured while viewing red ($L_{TH,R}$) or green ($L_{TH,G}$) color patches, plotted as median (center line) and interquartiles (top and bottom line)

3 Loudness matching

3.1 Setup and procedure

To avoid possible influences of image content or identifiable sound sources, synthetic optical and acoustical stimuli were used in this experiment in contrast to pictures and sounds of trains or cars as in previous studies. To isolate the effects of different colors, uniform red or light-green color-patches were presented on a 19 in. CRT monitor with a viewing distance of 70 cm. The monitor was positioned outside a sound attenuating booth and was visible through a window. The subjects were seated inside the booth and were instructed to keep their eyes focused on the image throughout the experiment.

Uniform exciting noise (UEN) as described in Fastl and Zwicker ([4], p. 171) was used as auditory stimulus. It was presented diotically through calibrated electrodynamic headphones (Beyer DT 48), using a free field equalizer (cf. [4], p. 7).

The subjects' task was to iteratively adjust the level of a test sound in comparison to a previously heard reference sound until both were perceived as having the same loudness. Simultaneously, red or green color patches were presented on the monitor. The reference and the test sound were always associated with different colors, so that a red reference stimulus would be paired with a green test stimulus and vice-versa. Under the hypothesis that red causes an increase in loudness, the following result should show up: a red reference color combined with a green test color should lead to an increased adjusted test sound level, as the test sound would require a higher level to produce the same loudness perception as the reference sound. Accordingly, a green reference color, combined with a red test color, would imply a decrease in adjusted test sound level.

The succession of auditory and visual stimuli is shown in Fig. 2. At the beginning of each iteration of the procedure, the reference color (either red or green) was shown on the monitor. After a time Δt, the reference sound with a duration T = 2.5 s was presented with a reference level L_{ref} of 60 dB. An additional waiting time t_n = 0.5 s was inserted after the sound, during which the reference color remained visible on screen. This was done to take into account possible postmasking effects. This sequence was then repeated for the test stimuli.

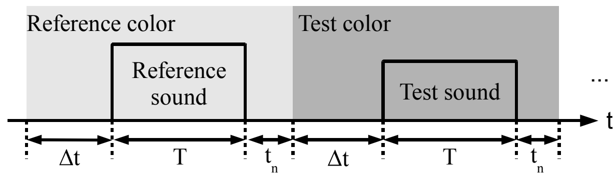

Fig. 2 Stimulus pattern used during one iteration of the loudness matching procedure.

Following the test color, the display was switched to a graphical user interface on a neutral gray background. The subjects now had the possibility of adjusting the level of the test sound using a fader displayed on the screen. The test sound was not audible at that time. The minimum and maximum values of the fader were initialized randomly, to avoid visual cues of the test sound level through the position of the fader. A representation of the user interface is shown in Fig. 3:

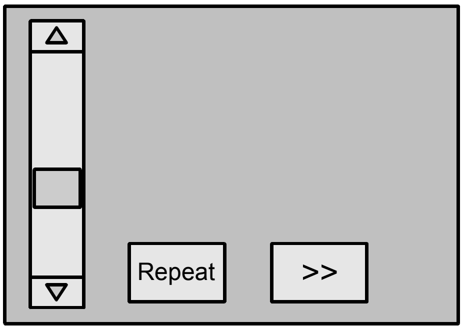

Fig. 3 Graphical user interface: The fader was used to adjust the test sound level; the button labelled "Repeat" triggered a repetition of the stimulus sequence with the new test sound level; the button labelled ">>" was used to indicate a satisfactory loudness match.

This procedure could be repeated until a satisfying loudness match has been achieved, yielding the adjusted test sound level L_{test}. After that, a new trial with new stimulus combinations followed. The initial test sound level at the beginning of each trial was set pseudo-randomly to either $L_{ref} + 6$ dB or $L_{ref} – 6$ dB. Both initial levels were chosen equally often to avoid bias effects.

The time between the onset of the visual and auditory stimuli Δt was varied between 0 s, 1 s, and 2 s to see if additional visual habituation time is beneficial for audio-visual interactions. Together with the waiting time t_n of 0.5 s this resulted in pauses between the reference sound and the test sound of 0.5 s, 1.5 s, and 2.5 s.

Sixteen normal hearing subjects participated in this experiment. Their age ranged from 22 to 42 years, with a median of 25 years.

3.2 Results

As shown in Fig. 4 and supported by statistical analysis (repeated measures ANOVA, Δt and color as within-subject factors), no influence of color on the adjusted test sound level can be found. For values of Δt of 1 and 2 s, subjects were able to match test and reference level with high accuracy, regardless of color. For Δt = 0 s however, L_{test} is about 0.5 dB higher than L_{ref}, suggesting that in this case the test sound was perceived as being somewhat softer than the reference sound. Statistical analysis via t-tests confirms this observation: a significant deviation from 0 can only be found for Δt = 0 s.

The pause between the end of the reference sound and the beginning of the test sound in this case was 0.5 s, so that effects due to postmasking and temporally partial masked loudness can be ruled out. Another possible explanation could involve the observation that, in case of Δt = 0 s, the auditory and visual stimuli started simultaneously. This means that an abrupt change of displayed color occurred at the same moment as the onset of the test sound, which could have diverted attention away from the auditory modality, thus reducing the remembered loudness at the time of level adjustment.

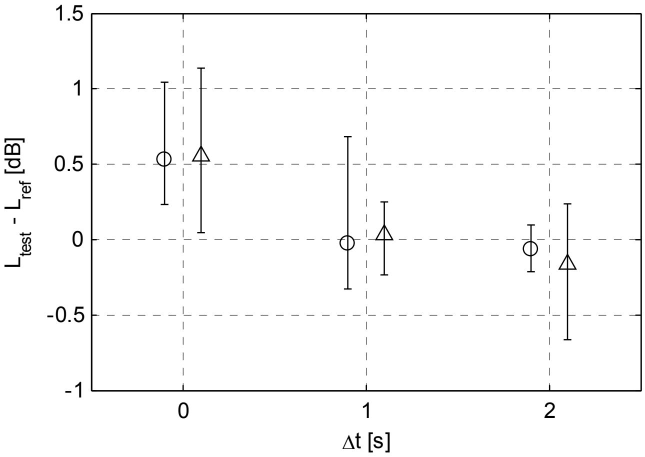

Fig. 4 Level differences between adjusted test sound and reference sound after loudness matching for different values of the time between onsets of the visual and auditory stimuli (Δt), given as medians and interquartiles for reference colors red (circles) and green (triangles).

In a second experiment without color changes (using a gray color patch as reference and test stimulus) however, similar results could be observed: For Δt = 1 and 2 s there were no differences between reference and adjusted test sound level, while for Δt = 0 s a significant increase of L_{test} above L_{ref} of about 0.5 dB was noted.

Additionally, a control experiment using the method of free magnitude estimation was performed under the same acoustical and visual conditions as described above. UEN with levels of 66, 63, 60, 57, and 54 dB was presented to 13 subjects (aged 23 to 33 years, median 27 years) while viewing red or green color patches on a CRT monitor. The results can be seen in Fig. 5.

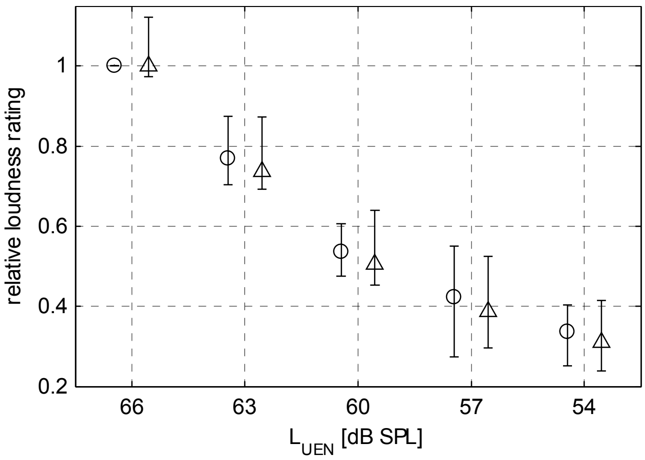

Fig. 5 Median values and interquartiles of magnitude estimates of the loudness of UEN under visual stimulation with red (circles) and green (triangles) color patches, normalized to the stimulus combination 66 dB/red.

The magnitude judgments were individually normalized to the value given for the stimulus combination 66 dB/red, as this was the condition expected to yield the highest ratings. Regarding the median values, in four out of five cases sounds associated with the color red caused an overestimation of loudness compared to sounds heard while seeing a green color. However, these differences can only be interpreted as tendencies, as they failed to reach statistical significance in an analysis of variance. The acoustical stimulus with a level of L_{UEN} = 66 dB could probably be recognized and remembered as the loudest sound in the experiment, so that in this case the additional presentation of colors seems to have only a marginal impact on loudness ratings.

It should be noted that the differences in magnitude estimates between the two visual conditions found in this part of the experiment are smaller than those reported in earlier studies with products of different color (cf. [3], [8], [11]).

4 Discussion

In contrast to previous experiments involving color influences on loudness, the studies presented here showed no evidence of audio-visual interactions between visual stimulation by different colors and absolute thresholds or loudness matching.

A possible reason for this discrepancy can be found by comparing the psychoacoustic methods employed in the experiments. Both threshold measurement via Békésy-tracking and loudness matching via iterative level adjustment require the subjects to strongly concentrate on the auditory modality in order to comply with the given experimental task. Any input from the visual system such as color information might thus be reduced in its importance in contrast to the acoustic stimulus.

In past experiments however, free magnitude estimation was chosen to rate the loudness of the test sounds (cf. [3], [8], [11]), and subjects were instructed to give spontaneous loudness ratings. Stimuli were presented only one time and could not be repeated. It is possible that in these cases subjects were concentrating more on loudness rating than loudness perception.

This could mean that additional visual input mainly affects loudness judgments and only to a lesser extend loudness perception itself. To test this hypothesis, further studies of audio-visual interactions using methods which demand spontaneous loudness judgments, such as ratings on categorical scales, have to be performed. It is expected that these methods, in accordance with magnitude estimates and in contrast to e.g. level adjustments, will more clearly show evidence of color influences on loudness.

Also, the role of image content, e.g. the depiction of recognizable everyday objects versus abstract color patches, has to be investigated in greater detail, as thoughts, memories, or expectations associated with presented objects might contribute to the process of forming a loudness judgment. This could also help to explain the differences in effect size found in the magnitude estimates described here (using abstract color patches and synthetic sounds) and those presented in earlier studies (using images of real vehicles and matching technical sounds).

The concept that interactions between different modalities rely on spatial, temporal, or other similarities of the involved stimuli was already noted by Marks ([6], p. 7). The present studies also support Marks's notion that when investigating intermodal interactions, a distinction has to be made between the actual perception of sensory inputs and judgments regarding these perceptions.

5 Summary

The effect of visual stimulation by red and green color patches on absolute threshold and loudness perception was measured. No difference between absolute thresholds obtained under the influence of the two colors was found. Also, loudness matching of broadband acoustic stimuli via level adjustment was not affected by simultaneous presentation of red and green colors.

Taking into account the results of previous studies which did show color influences on loudness, it seems that psychoacoustic methods that require a spontaneous loudness judgment are better suited to investigate audio-visual interactions than methods which necessitate a strong concentration on the auditory modality.

References

[1] Blauert, J., and Jekosch, U.: "Sound-Quality Evaluation - A Multi-Layered Problem." *Acta Acustica united with Acustica* 83, 747-753, 1997

[2] Böhm, M., Patsouras, C., and Fastl, H.: "Beeinflussung des Lautheitsurteils durch Stand- oder Bewegtbilder. (Influence of still and moving pictures on loudness judgments)" In: *Fortschritte der Akustik - DAGA '03*, 612-613, Deutsche Gesellschaft für Akustik e.V., Oldenburg, 2003

[3] Fastl, H.,: "Audio-visual interactions in loudness evaluation," In: *Proceedings of the 18th International Congress on Acoustics*, Kyoto, Japan, 1161–1166, 2004

[4] Fastl, H., and Zwicker, E.: "Psychoacoustics. Facts and Models." 3rd ed., Springer, Berlin, New York, 2007

[5] Höger, R., and Greifenstein, P.: "Zum Einfluß der Größe von Lastkraftwagen auf deren wahrgenommene Lautheit. (Influence of the size of trucks on their perceived loudness)" *Zeitschrift für Lärmbekämpfung* 35, 128-131, 1988

[6] Marks, L. E.: "The Unity of the Senses." Academic Press, New York, San Francisco, London, 1978

[7] Menzel, D.: "Psychoakustische Untersuchungen zum Einfluss der Farbe auf die Lautheit von Sportwagen. (Psychoacoustic experiments on the influence of color on the loudness of sports cars)" In: *Fortschritte der Akustik - DAGA '07*, Deutsche Gesellschaft für Akustik e.V., Berlin, 855-856, 2007

[8] Menzel, D., Fastl, H., Graf, R., and Hellbrück, J.: "Influence of vehicle color on loudness judgments." *J. Acoust. Soc. Am.,* in press, 2008

[9] Namba, S., Kuwano, S., Kinoshita, A., and Hayakawa, Y.: "Psychological Evaluation Of Noise In Passenger Cars - The Effect Of Visual Monitoring And The Measurement Of Habituation." *Journal of Sound and Vibration* 205, 427-433, 1997

[10] Patsouras, C., Filipou, T., and Fastl, H.: "Influences of color on the loudness judgement," In: *Proceedings of the Forum Acusticum*, Sevilla, PSY-05-002-IP CD-ROM, 2002

[11] Rader, T., Morinaga, M., Matsui, T., Fastl, H., Kuwano, S., and Namba, S.: "Cross-cultural effects in audio-visual interactions," In: *Transactions of the Technical Committee of Noise and Vibration of the Acoustical Society of Japan*, N-2004-31, 2004

[12] Seeber, B., Fastl, H., and Koci, V.: "Ein PC-basiertes Békésy-Audiometer mit Bark-Skalierung. (A PC-based Békésy audiometer with Bark scaling)" In: *Fortschritte der Akustik - DAGA '03*, 614-615, Deutsche Gesellschaft für Akustik e.V., Oldenburg, 2003

[13] Suzuki, Y., Abe, K., Ozawa, K., and Sone, T.: "Factors for perceiving sound environments and the effects of visual and verbal information on these factors." In: Schick, A., Meis, M., and Reckhardt, C. (ed.), *Contributions to Psychological Acoustics: Results of the Eighth Oldenburg Symposium on Psychological Acoustics*, 209-232, 2000