CROSSCULTURAL EFFECTS IN AUDIO-VISUAL INTERACTIONS

T. Rader[1,2], M. Morinaga[2], T. Matsui[2], H. Fastl[1], S. Kuwano[2] and S. Namba[2]

[1]AG Technische Akustik, MMK, Technische Universität München,
Arcisstr. 21, D-80333 München, Germany
[2]Department of Environmental Psychology, Osaka University,
1-2 Yamadaoka, Suita, Osaka, 565-0871, Japan

In this study, the evaluation of audio-visual interaction regarding to the sounds and colours of trains was investigated. Different pictures of trains were modified in their colors and presented on a liquid crystal screen in a sound proof room with the appropriate sound of the train. The subjects judged the loudness of the sound-picture stimuli with the method "free magnitude estimation". At first, a previous study with German subjects and German trains was repeated under the same conditions with Japanese subjects. Furthermore different experiments with more Japanese trains were performed. To conclude differences between Japanese and German subjects in the estimation of loudness, a semantic differential experiment was arranged. The results show that an influence of the color to the perception of loudness for most subjects does exist while the averaged effect over all subjects is small. In addition an experiment with the semantic differential can explain the different favour of German and Japanese subjects for the most quiet sound.

INTRODUCTION

Pervious studies ([1], [2], [3], [4], [6]) show that factors other than sound can influence the perceived loudness of acoustical stimuli. In this study the dependence of this influence from different cultures of subjects (Japanese and German) was analyzed. Three different experiments with Japanese subjects were performed to find out quantitatively if different colors of trains as optical stimuli are able to influence the perceived loudness of an acoustical stimulus presented at the same time. The results were compared with results of a previous study with German subjects [6].

EXPERIMENTS

Experiments 1, 2 and 3 were performed in a sound proof room with a distance of 1.8m between the subject and the 32" LCD-Screen (Panasonic TH-32LX20). A personal computer (Toshiba S3000-514) realized the sound reproduction over an external digital-to-analog converter (Roland UA-30) and the control of the LCD-screen over an integrated S-Video output connected to the LCD-screen. A numerical keypad positioned in front of the subject was used to submit the output from the subject to the personal computer. The sequence control was taken over by a program written for "MatLab 6.5" from "The MathWorks".

Experiment 1: German Train (ICE) presented with one sound level.

Material:
As acoustic stimulus for *Experiment 1* the same sound as in [6] was used: an "ICE" (Inter City Express) train passing and recorded monophonically at a distance of 25 m. All audio stimuli were presented diotically monophone via a freefield equalized [7] and electrodynamic headphones (Beyer DT 48) to obtain the same conditions as in Germany. The sound with a duration of 11 seconds was presented with a maximal sound level of $L_{F,A,max}$=82dB. As seen in Fig. 1, pictures with the original color of the "ICE" (white), electronically modified colors of the train (red, blue and green) and a complete black screen (without picture) were presented as visual stimuli. 15 Japanese subjects (8 male and 7 female) from 21 to 35 years old (median: 23) carried out this experiment.

Fig. 1: Visual stimuli: ICE, JR Chuo Line and Nozomi Express

Methods:
The method "absolute free magnitude estimation" was chosen, to get a quantitative result how much

From: ASJ 2004, Tokyo, Japan

visual stimuli can change the perceived loudness of the acoustical stimuli. The subjects were asked to estimate the loudness of the presented audio-visual stimuli. They were instructed that "each positive real number is possible".

In this experiment, 30 audio-visual stimuli (5 colors * 4 repetitions + 5 colors * 4dB_up + 5 colors * 4dB_down) were used. The stimuli increased and decreased by 4dB provide a "confusion" of the subject, so that the aim of this study is not obvious for them. The audio-visual stimulus pairs were presented in three randomized orders changed every five subjects.

Results:
Figure 2 shows the free estimated loudness of the stimuli "ICE" with the modified colors of the train for both population groups. The values of each subject were normalized to the mean of the acoustical stimuli without visual stimuli. The data for German subjects could be found in [6].

For each subject, normalization was done to the arithmetic mean (Normalization Value) of the 4 estimated values without optical stimuli. The relative loudness as seen in Figure 2 is the deviation of normalized value from the estimated loudness without optical stimulus.

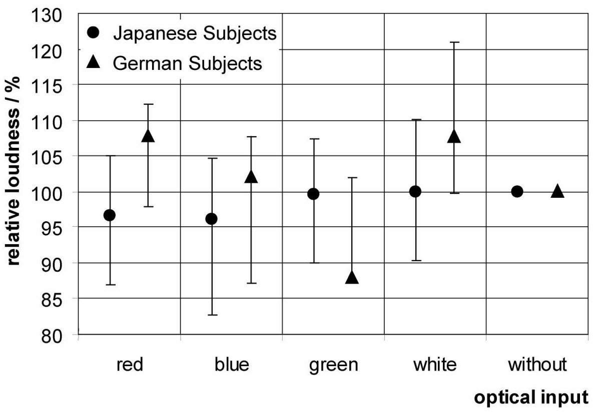

Fig. 2: Medians and interquartile ranges (over all estimations of all subjects) of the loudness for the stimuli "ICE" for the different optical inputs for Japanese (dots) and German (triangles) subjects.

For Japanese subjects, the blue "ICE" train was estimated most quiet with 96%, followed from the red stimulus with 96.6%. The green (99.5%) and white (100.0%) optical input have nearly the same value as the acoustical stimulus without visual component (per normalization 100%) and are estimated the loudest. German subjects judge the color green 12% lower and the colors red and white 7.5% higher than without stimuli.

Independent from culture group, the blue stimulus shows the lowest lower interquartile with a value of 82.8%; but it also shows the highest interquartile span of 21.8%.

Analysis of variance was used to test for statistically significant differences among colors, and when appropriate the Scheffe's post hoc test was applied. No significance between the colors could be found.

To investigate further influences with more colors and Japanese trains, the following experiment was designed.

Experiment 2: Japanese Trains (JR Chuo Line and Nozomi Express).

Material:
In this experiment the stereo sounds of two passing trains, "JR Chuo Line" (duration 15.5s) and "Nozomi Express", (duration 10.5s) recorded stereophonically at a distance of 25m were used and presented diotically with headphone amplifier unit (STAX SRM-313) with the appropriate headphones (STAX SR-303). The sounds were presented again with a maximum sound level of $L_{F,A,max}$=82dB. As shown in Fig. 1 pictures with the original color of the "JR Chuo Line" (orange) and with electronically modified colors of the train (red, yellow, dark blue, blue, green and white) as well as the original color of the "Nozomi Express" (white) and again electronically modified colors (red, orange, yellow, dark blue, blue, green), and a complete black screen (without picture) were presented as visual stimuli. 14 Japanese subjects (10 male and 4 female) from 22 to 31 years old (median: 24) carried out this experiment.

Method:
The same method as described in Experiment 1 (absolute free magnitude estimation) was used in this experiment.

For each kind of train, 48 audio-visual stimuli (8 colors * 4 repetitions + 8 colors * 4dB_up + 8 colors * 4dB_down) were used. The stimuli increased and decreased by 4dB again provide a "confusion" of the subjects, so that the aim of this study was not obvious for them.

Results:
Figure 3 shows the estimated loudness of the stimuli “JR Chuo Line” (dots) and “Nozomi Express” (triangles) with the modified colors of the train. The ordinate represents the medians and interquartile ranges of all subject data (n=60). The relative loudness as seen in Figure 3 is again the deviation of normalized value from the estimated loudness without optical stimulus.

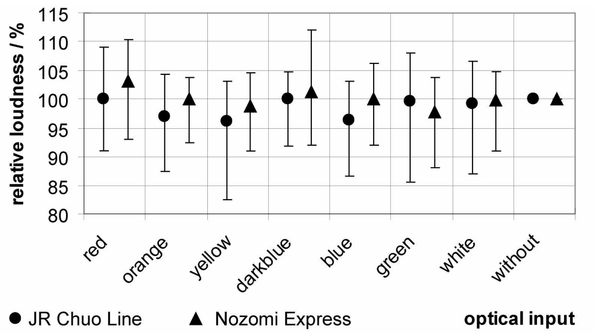

Fig. 3: Medians and interquartile ranges (over all estimations of all subjects) of the loudness for the stimuli “JR Chuo Line” (dots) and “Nozomi Express” (triangles) for the different optical inputs.

For “JR Chuo Line” yellow (96.1%) was judged as the most quiet sound, followed of the colors blue (96.3%) and orange (97.1%). The visual input of red, dark blue, green, white and without optical stimulus, was estimated as the loudest level with nearly the same value (100%). The smallest interquartile range could be found for the dark blue (92.0%-104.7%) and the biggest interquartile range for the green (85.5%-108.0%) input. Yellow (82.5%) shows the lowest and red (109.0%) the highest interquartile.

For the “Nozomi Express” stimuli, the subjects judged green (97.6%) as the most quiet and red (103.1%) as the loudest sound. Orange, yellow, dark blue, blue, white and without stimuli have nearly the same level of 100%. The smallest interquartile range could be found for the orange train (92.3%-103.7%), the biggest for the dark blue (92.0%-112.0%). The lowest interquartile is located at the green (88.1%), the highest at the dark blue (112.0%) stimulus.

For each train, analysis of variance was used to test for statistically significant differences among colors, and when appropriate the Scheffe’s post hoc test was applied. Neither between the colors of “JR Chuo Line” nor those of the “Nozomi Express” significance between the colors could be found.

After the experiment, some subjects reported that they had difficulties to estimate the sound, because the level was always nearly the same. From this reason the next experiment was designed with a bigger dynamic range.

Experiment 3: Japanese Train (JR Chuo Line) and German Train (ICE) presented with four different sound levels.

Material:
This experiment used the same hardware setup as in *Experiment 2*. The acoustical stimuli (JR Chuo Line and ICE) were now presented with 4 different maximum sound level of $L_{F,A,max}$=60dB, 70dB, 80dB and 90dB. For visual stimuli, electronically modified color of the trains (red, blue, green and white) of “JR Chuo Line” and the “ICE” were presented. 13 Subjects (5 male and 8 female) from 21 to 27 years old (median 23) carried out this experiment.

Method:
The same method as described in Experiment 1 (absolute free magnitude estimation) was used in this experiment.

For each kind of train, 48 audio-visual stimuli (4 colors * 3 repetitions * 4 levels) were used.

Results:
Figure 4 shows the geometrical mean (over all estimations of all subjects) of the different sound levels for the stimuli “JR Chuo Line” (upper part) and “ICE” (lower part) for the four different kinds of optical inputs.

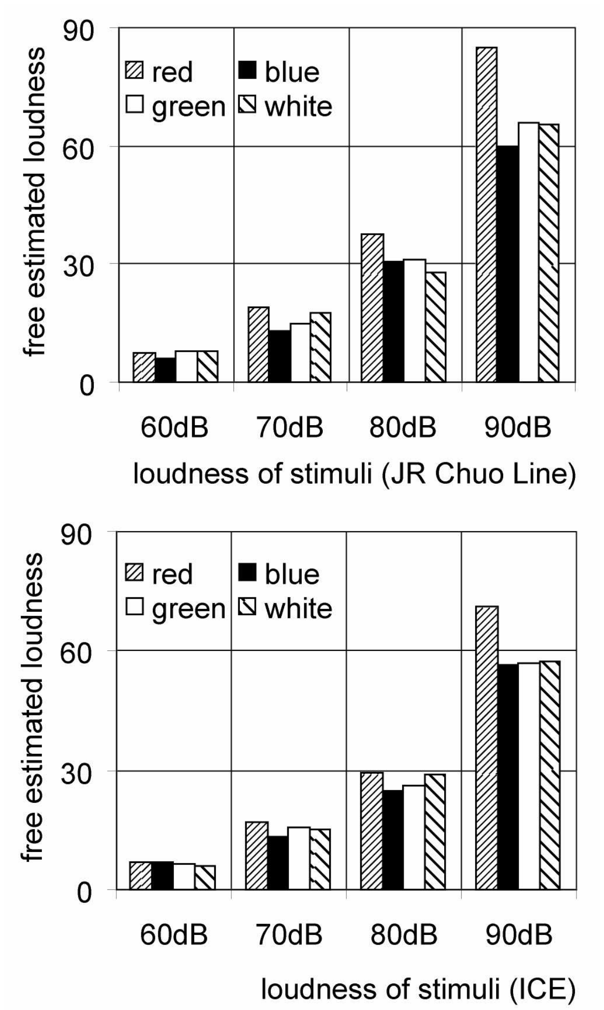

Fig. 4: Geometrical mean (over all estimations of all subjects) of the different loudness levels for the stimuli "JR Chuo Line" (upper part) and "ICE" (lower part) for the four different kinds of optical red, blue, green and white inputs.

Whereas the colors blue, green and white are largely the same level, the red stimulus is always judged higher. At quiet sound level (60dB) the effect is not very distinct.

A 3-way analysis of variance (ANOVA) was employed. The different sound levels (respectively each color alone) show significant values (Scheffe's Test). Furthermore significance between the colors red and blue could be found.

DISCUSSION OF ALL EXPERIMENTS

Experiment 1:
An obvious discrepancy between the median over Japanese and German Subjects is that Japanese subjects *in an average consideration* are not as sensitive to audio-visual interaction as German subjects.

Further more on German subjects the interquartile ranges are lower (13%-20%) than on Japanese Subjects (17%-23%). This suggests that Japanese subjects might have more *individual* audio-visual interaction than German subjects.

Germans know the original color of the German ICE train (white). That could be the reason that they judge the white train as the same level than the red colored train, as an effect of recognition. Otherwise most Japans don't know the original color of the train so they are not so surprised over red train as German subjects.

The mean influence over all subjects is not as high as for German subjects.

Experiment 2:
It is surprising that most subjects estimate the loudness of the Nozomi Express with optical stimulus averaged over the different colors about the same as without, while the loudness of the JR Chuo Line with optical stimulus is estimated less than without. This may be caused by the optical appearance of the train of Chuo line which gives the impression to be slower than the Nozomi Express.

Experiment 3:
In this Experiment, for both kinds of visual stimuli (German "ICE" and Japanese "JR Chuo Line"), the red trains were judged the loudest.

To understand in which way visual stimuli can interact with the perception of loudness, a further experiment was performed to investigate, if and in which way different colors could influence feelings of a subject.

Semantic Differential Experiment: Ten different colors (presented only by their names) evaluated with the Semantic Differential.

Material:
20 Japanese subjects (11 male and 9 female) from 19 to 56 years old (median 23.5) performed this experiment.

Method:
The SD (Semantic Differential) method was used. Ten different colors (white, yellow, orange, red, green, light blue, blue, brown, purple, grey) written in Japanese language on the questionnaires were used as stimuli. The evaluation occurred with 14 bipolar adjective pairs (Table 1) arranged

in seven category scales. Three orders of the adjective-pairs and randomized orders of the 10 colors for every subject were used to minimize accidental dependencies between the different adjective-pairs and/or the different colors.

Tab. 1: Adjective pairs used in the SD experiment.

deep	-	light
beautiful	-	ugly
loud	-	quiet
graceful	-	coarse
sharp	-	dull
dark	-	bright
vivid	-	dusty
hard	-	soft
pleasing	-	unpleasing
clear	-	vague
heavy	-	light
pleasant	-	unpleasant
hot	-	cool
pure	-	Impure

Results:
The factor analysis with a varimax rotation was applied on the full set of data and, as shown in Table 2, three factors were extracted.

Tab. 2: Rotated Components Matrix of factor analysis and extracted factors.

	Rotated Component Matrix	Component		
		1	2	3
Pleasant	beautiful-ugly	0,86	0,06	-0,15
	pleasing-unpleasing	0,84	-0,11	0,00
	pleasant-unpleasant	0,82	-0,11	-0,22
	pure-impure	0,69	0,08	-0,46
	graceful-coarse	0,60	-0,47	0,09
Powerful	loud-quiet	-0,04	0,85	-0,24
	clear-vague	0,31	0,77	-0,27
	hot-cool	-0,44	0,61	0,21
	vivid-dusty	0,56	0,58	-0,36
	hard-soft	-0,20	0,49	0,04
Metallic	dark-bright	-0,20	-0,17	0,81
	heavy-light	-0,28	0,02	0,76
	deep-light	0,03	0,58	0,65
	sharp-dull	-0,03	0,37	-0,54

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

The first factor, interpreted as "pleasant" factor, contains high loadings on the adjective pairs "beautiful-ugly", "pleasing-unpleasing", "pleasant-unpleasant", "pure-impure" and "graceful-coarse". This factor accounts as 26.6 percent of the variance. The second factor could be described as "powerful" factor. It combines adjective pairs which have relation with energy like "loud-quiet", "hot-cool" and "hard-soft". It accounts as 21.5 percent of the variance. The third factor combines adjective pairs describing "metallic" character like "dark-bright", "heavy-light" or "sharp-dull". This factor accounts as 18.2 percent of the variance.

The SD experiment shows that for description of colors the same factors as for description of sound can be found (see [5]).

A cluster analysis (using the Ward Method) was conducted and the following groups of colors could be found as shown in Table 3.

Tab. 3: Result of cluster analysis and respective loads of factorscore for the different colors.

group	color	pleasant comp.	powerful comp.	metallic comp.
1	green	-0,46	0,22	-0,43
1	blue	-0,80	0,00	-0,30
2	white	-0,60	0,56	1,14
2	light blue	-0,71	1,08	0,86
3	brown	0,67	0,55	-0,92
3	purple	0,22	0,00	-0,75
3	grey	1,20	1,07	-0,34
4	orange	-0,12	-0,98	0,28
4	red	0,10	-1,46	-0,45
4	yellow	0,51	-0,87	0,93

The result of the cluster analysis (Table 3) is also interpretable with the factor score:

- Group 1 contains the "natural" colors green and blue. Conspicuous is the low loading on the powerful factor and a loading on the pleasant factor.
- Group 2 combines the "bright" colors white and light blue which show a loading on the pleasant factor.
- Group 3 describes the more "dreary" colors like brown, purple and grey.
- Group 4 contains the "warm" colors orange, red and yellow. High loadings on the powerful factor can be found here. Specially red and orange show a small loading on the pleasant factor.

The louder perception of the sound in combination with the color red as optical stimulus (e.g. “Nozomi Express in Exp. 2 and Exp. 3) could be ascribed to the association of red as a color affiliated with the power component including adjectives like “loud-quiet” and “hot-cool”.

Otherwise the color blue with a low loading on the powerful factor and a high loading on the pleasant factor is often judged quieter than other colors (e.g. “JR Chuo Line” in Exp. 2).

This suggests that the perception of loudness has an interacting with the sensation, caused by optical stimulius.

CONCLUSIONS

The effect of audio-visual interaction on the perception of Japanese subjects was examined and compared with German subjects.

In summary, strong individual influences as shown in [6] for German subjects could be found for Japanese subjects, too. However, subjects often judged very distinct from each other: Whereas some subjects show strong impacts (differences in loudness up to 100 %), other subjects show nearly no influence or even an opposite influence.

Japanese subjects, similar to German subjects, show a tendency to consider red as the “loudest” color. In the case of Japanese subjects, this effect could be explained with the SD-experiment in which the color red shows high loadings of the powerful factor and few in the pleasant factor.

Further more basic experiments like e.g. colors (no pictures) and noise/sinus seem necessary to find out more clearly the interconnection between visual and audio perception.

ACKNOWLEDGEMENTS

Many thanks to Dr. Jiro Kaku of the Kobayasi Institute of Physical Research for providing the sounds and the raw data of the photos.

REFERENCES

[1] Abe, K., Ozawa, K., Suzuki, Y. and Sone, T.: The Effects of Visual Information and the Impression of Environmental Sounds. Proc. Inter-Noise Fort Lauderdale 99 (1999).

[2] Hashimoto, T., Hatano, S.: Effects of Factors other than Sound to the Perception of Sound Quality. Proc. International Congress on Acoustics Rom (2001).

[3] Hatano, S., Hashimoto, T., Kimura, Y., Tanaka, T.: Sound Quality Evaluation of Construction Machine. Proc. International Congress on Acoustics Rom (2001).

[4] Morinaga, M., Aono, S., Kato, T. and Kuwano, S.: Psychological Evaluation of Waterside Space Using Audio-Visual Information. Empirical Studies of the Arts, Volume 21, Issue 2, Pages 185-194 (2004).

[5] Namba, S., Kuwano, S. and Hatoh, T.: Sound quality of road traffic noise transmitted through different types of simulated wall. Proc. Inter-Noise Nizza 00 (2000).

[6] Patsouras, C., Filippou, T., and Fastl, H.: Influences of Color on the Loudness Judgement. Proceedings Forum Acusticum Sevilla 2002, CD-ROM (2002).

[7] Zwicker, E., Fastl, H.: Psychoacoustics - Facts and Models. 2nd Updated Edition. Springer, Heidelberg, New York (1999).