Evaluation of the acoustic comfort of air-treatment systems with a multi-task procedure

Antoine Minard, Benjamin Robin, Francette Fournier
CIAT, Avenue Jean Falconnier, 01350 Culoz, France.

Jean-Jacques Embrechts
University of Liege, Campus du Sart-Tilman, B28, 4000 Liege 1, Belgium.

Alexis Billon, Bertrand Goujard, Anas Sakout
LaSIE, Avenue Michel Crépeau, 17042 La Rochelle Cedex 1, France.

Summary
The perceptual evaluation of sound quality is a necessary step to define acoustic specifications for industrial products. However, usual procedures for this include experimental tasks where the participants are asked to perform their perceptual judgment while paying exclusive attention to the produced sound. As far as air-treatment systems (ATS) are concerned, this kind of task is not representative of the normal listening conditions. Indeed, ATS users are hearing the sound while performing other tasks that require partially or entirely their attention. In this study, a dedicated procedure was designed in order to address the influence of the listener's attention towards the judged sound. This procedure includes a subsidiary task on the purpose of shifting the listener's attention away from the sound. Participants were then asked to evaluate a posteriori the acoustic comfort in terms of how much the sound interfered with the subsidiary task. The results of this experiment as regards sound preferences were finally compared to those obtained without the subsidiary task. This study was conducted through the Vaicteur Air^2 project supported by OSEO.

PACS no. 43.66.Lj

1. Introduction

Sound quality evaluation has been an issue of great interest in many studies due to important applications in the industrial world. Indeed, acoustics is nowadays of significant importance for companies because consumers' choice criteria include, among other things, the acoustic comfort associated with the product, in the present purpose the air-treatment systems (ATS). In order to differentiate themselves from competitors, manufacturers are now inclined to follow a process of "sound design" [1] in which sound quality evaluation is a critical step in defining the acoustic specifications of the product. Thus, they try to define to which extent perceptual parameters affect the annoying, intrusive or unpleasant character of the sound produced by an object designed for a particular function and not specifically created to have a "nice" sound, unlike musical instruments for example.

Many studies addressing sound quality evaluation are found in the literature. However usual experimental methodologies put listeners in a particular condition where they have to perform their judgement while listening attentively to the sound. For the specific case of ATS, this condition is rather unrealistic. Indeed, in a real environment, ATS users judge the sound emitted by the ATS as a source of intrusion in their everyday life. As a consequence ATS users do not listen carefully to the sound but are nonetheless annoyed when it interferes with their current activity.

The aim of the present work is to address the question of how different annoyance judgements are between an attentive listening condition and a condition where listener's attention is shifted away from the sound. This implies the inclusion in the experimental procedure of a subsidiary task aiming to take over the attention of the listener so that he/she is unable to focus on the sound. As far as psychological acoustics is concerned, though the use of such a subsidiary task is sporadic in the literature, a few examples of

ISSN 2226-5147
ISBN 978-80-01-05013-2

experimental procedures in a multi-task context exist. Although their subject of interest is very different from that of the present work, Suied et al. used a subsidiary task to measure precisely reaction times to warning sound [2]. The subsidiary task consisted in following with the mouse a moving dot on the screen. In an older study [3] Susini and McAdams used a subsidiary task of memorization of digit sequences when trying to perceptually measure the overall loudness of time-varying sounds.

This paper describes the experimental work conducted in order to compare the annoyance judgements obtained in usual attentive listening conditions to those obtained with a multi-task procedure.

2. Experiment description

The experimental methodology of the work described here is strongly inspired by the study of Susini and McAdams [3]. A dedicated procedure was specifically designed in order to shift the listeners' attention away from the sound that was being played and to ask them to perform their annoyance judgement once both the memorization task and the sound playback had stopped.

2.1. Stimuli

The stimuli set corresponds to a fraction of the larger dataset of ATS sound recordings established in a previous study [4]. Five sounds out of the 8 used in this previous work were selected. Each sound corresponded to a different ATS. The reason of this dataset reduction lies in a necessary trade-off between experiment duration and the number of sound evaluations required by the procedure described below. Each sound had been previously recorded by AKG C451B cardioid microphones in a semi-anechoic room. Moreover the listeners' judgement should not be based upon loudness variations. Consequently the sounds had been experimentally equalized in loudness with 5 participants. Finally the duration of the sounds corresponded to that of the memorization task, which is between 50 and 70 sec. according to the listener's promptness at entering the digit sequences.

2.2. Apparatus

The experiment was performed through a Labview 2010 Graphical User Interface (GUI) handling sound playback and recording of the listeners' response data. The sounds were played through an RME Fireface 800 audio interface and Sennheiser HD650 Pro headphones in an IAC double-walled audio booth.

2.3. Participants

Twenty-nine participants (20 men, 9 women, aged between 20 and 25) volunteered as listeners for this experiment, all of which reported having normal hearing.

2.4. Procedure

At the beginning of the experiment, participants were given written instructions presenting the context of the study and explaining the task to accomplish. They had to perform 2 successive tasks.

In the first one, digits were displayed one by one on the screen thus forming a sequence to be entered with the keyboard at the end of the digits display. After validating, another sequence was presented and had to be memorized, and so forth until the time limit was reached. The time limit was 20 sec. and did not include the display time (only the time for entering the digits and validating was counted). The duration of the display time could vary since it was related to the number of sequences displayed which depended on how fast the listener's responses were. The memorization task lasted 50 to 70 seconds overall (diplay time and response time). The first sequence was made out of 4 digits. It incremented each time the participant answered correctly to 2 sequences of the same length. An ATS sound was played through the headphones over the whole duration of this task, except for a silent condition where no sound was played. This silent condition was used in order to observe a possible influence of the presence of sound on the performance in the realization of the memorization task

At the end of the allotted time, the second task consisted in evaluating how much the sound interfered with the realization of the first task on a 5-item categorical scale. The 5 items were labeled with the French equivalent of 'not at all', 'a little', 'fairly', 'much', and 'very much'. It is important to note that the sound was not played during this task anymore, and that the participant had no means to listen to it again. Moreover, the participants were only asked to perform this task when a sound had indeed been played during the memorization task, which is not the case of the silent condition mentioned above.

These 2 tasks were repeated 3 times for each of the 5 sounds in a random order, but each successive memorization task for the same sound was necessarily separated by a memorization task with another sound (or with no sound). The silent condition was repeated 5 times. Thus 20 successive alternations of the two tasks (except for the silent condition for which the second task was not presented) were to be fulfilled by the participants.

2.5. Results

For each participant and each sound condition, there are two points of interest. First, his/her performance in the memorization task can be estimated as the number of sequences correctly memorized in the allotted time, averaged over the 3 or 5 presentations of the same sound condition. Second, his/her evaluation of the degree at which the sound interfered with the realization of the memorization task (for ATS sound conditions only) is averaged over the 3 presentations of the same sound condition.

As for the first point of interest, none of the participants exhibited a significant influence of the presence or absence of sound on the performance in the memorization task. Their performance did not show strong variations between the different ATS sounds either.

As for the second point of interest, there are strong variations among participants and no significant difference can be observed between the evaluation of each sound. As a consequence, it is necessary to look at how scattered the individual evaluations are around its mean values. For this purpose, the inter-participant correlation matrix of the sound evaluations was calculated. In order to display these correlations in a meaningful manner, a cluster analysis (see [5]) was used. First the correlation matrix was linearly transformed into a distance matrix, with the minimum value of '0' for a correlation coefficient of 1, and the maximum value of '1' for a correlation coefficient of -1. This distance matrix serves as input for an unweighted arithmetic average clustering (UPGMA) analysis algorithm whose result is displayed through a hierarchical tree representation shown in Figure 1. On this figure, each leaf (at the bottom) represents a participant, and the height of the node (horizontal line) that links 2 leaves corresponds to the cophenetic distance between the 2 associated participants. The cophenetic distance is the distance modeled by the clustering algorithm so as to approximate the input distance.

The tree representation shows a rather strong separation (displayed with the vertical dotted line on the figure) of two groups of participants. This separation suggests two different trends in the participants answers that probably explain in part the high variations mentioned earlier. As a consequence, the evaluations will subsequently be considered separately. Figures 2 and 3 show the mean evaluations of the 5 sounds and their standard deviations for respectively group 1 with 18 listeners and group 2 with 9 listeners (the results of 2 of the initial 29 participants were preemptively removed from the data because of inconsistencies).

2.6. Discussion

The results displayed in Figures 2 and 3 have to be looked at in regards of a comparable evaluation of the preferences of listeners among the sounds. As a reference condition, Figure 4 shows the results of an experiment conducted with the same sounds in a comparative study of experimental procedures to be published. In the corresponding experiment, the listeners' preferences were obtained through a paired-comparison procedure without any attention-focusing subsidiary task. This means that this experiment was conducted in attentive listening conditions. The preference scale is obtained by a linear methodology that basically estimates the preference degree of each sound as the

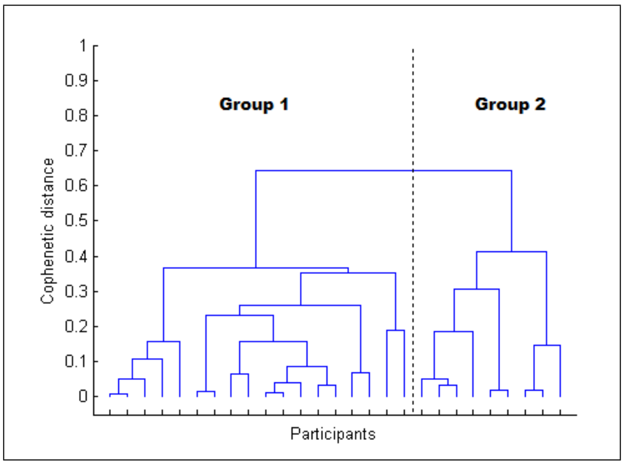

Figure 1. Tree representation of the inter-participant distance matrix.

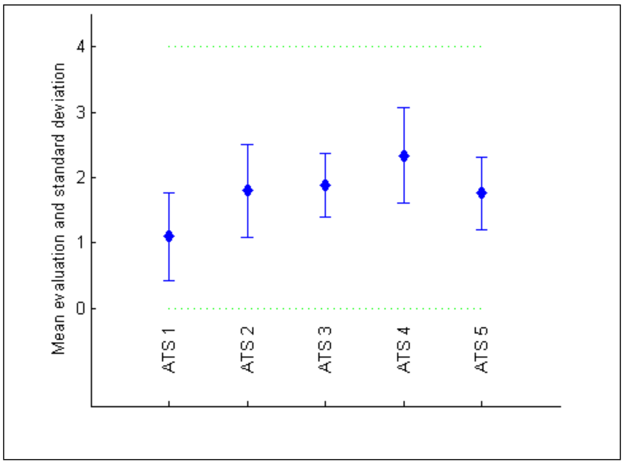

Figure 2. Mean evaluations and standard deviations for group 1. The values on the Y axis represent the items of the categorical scale from '0' for 'not at all' to '4' for 'very much'.

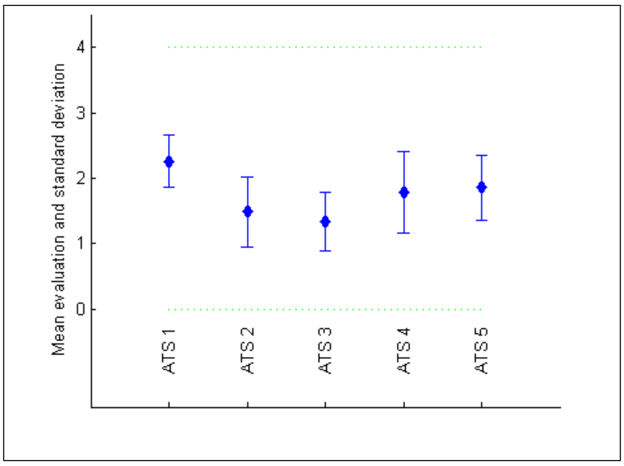

Figure 3. Mean evaluations and standard deviations for group 2. The values on the Y axis represent the items of the categorical scale from '0' for 'not at all' to '4' for 'very much'.

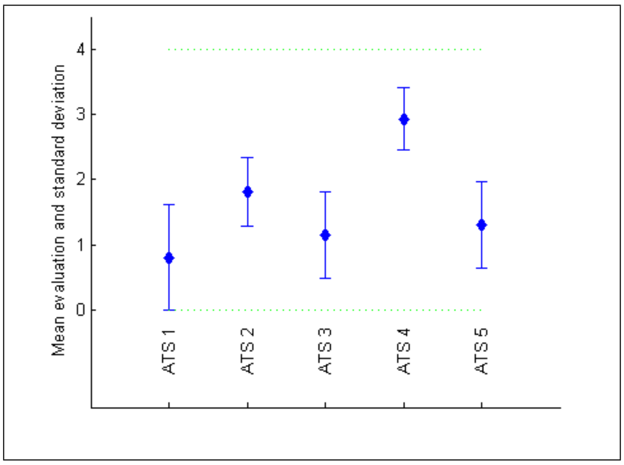

Figure 4. Mean evaluations and standard deviations for the reference condition (paired-comparison without attention-focusing subsidiary task).

number of times it was preferred to the others. However, for comparison purpose, the displayed scale was adapted to that of the previous figures.

In light of these results, one can observe that the evaluation scale for the group 1 (Figure 2) would tend to fit the preference scale of the reference condition in the sense that the extreme points of the scales (ATS 1 and 4) are the same. There are a few disparities among the three other items, but the rather high standard deviation values temper the significance of the differences between the corresponding mean values.

On the contrary, the results of group 2 (Figure 3) exhibit a strong divergence from the reference condition since the rankings of the sounds according to the scales are very different. This scale could even seems reversed when compared with that of group 1 which support the strong disparity of the results of the two groups and consequently of their perception of the sounds in attention-focusing conditions.

In the end, if all participants had shown the same global response as group 1, one could conclude that the attention context in experimental procedures for measuring annoyance or unpleasantness has no influence on the results. However the identification of a minority group that shows a strong discrepancy from the majority and that did not exist with the reference condition prevent us from maintaining such a conclusion. Thus it seems that attention context can have a significant influence on some listeners' judgements, at least as far as annoyance from ATS is concerned.

3. Conclusion

A new experimental procedure was specifically designed in order to address the question of how influent the listening conditions are on annoyance judgements. This procedure included a subsidiary task aiming at preventing the listeners from focusing their attention on the sounds when performing their judgement. This particular condition is closer to that of a real environment of ATS usage, where sound annoyance is expressed through intrusion in everyday activity rather than hedonic judgements.

The results obtained with this procedure were compared with those obtained with a more common paired-comparison procedure. It appeared that attentional conditions may have a significant influence for a minority part of the listeners, whereas the others do not show strong differences in annoyance judgements whatever the listening conditions. It is obvious that the number of participants to this experiment does not allow conclusiveness, particularly for the minority group with only 9 listeners. These results would be reinforced by pursuing this experiment with a larger number of participants.

Another further point of interest is to wonder about what in each sound may induce a different response of participants in different listening conditions. This question leads to the ATS sound description issue and more precisely to the study of the timbre. Indeed, timbre would serve as a starting point for identifying sound parameters (so-called audio features) that are responsible for these disparities.

Finally, it is important to keep in mind that the conclusions drawn here only stand for the particular type of sound under study, that is the sound of ATS. How much the perceptual trends observed here would hold for other types of sounds needs to be investigated. For such purpose, it would be necessary to reproduce such experimental work on a larger and more diversified environment sound dataset.

Acknowledgement

This work has been conducted through the *Vaicteur Air²* project supported by *OSEO*.

References

[1] J. Blauert, U. Jekosch: Sound-quality evaluation – A multi-layered problem. Acta Acustica united with Acustica **85**(5) (1997) 747-753.

[2] C. Suied, P. Susini, S. McAdams: Evaluating warning sound urgency with reaction times. J. Exp. Psychol.-Appl **14**(3) (2008) 201-212.

[3] P. Susini, S. McAdams: Influence of sound-directed attentional focus on overall loudness ratings. Proc. EAA Congr., Sevilla (2002).

[4] A. Minard, A. Billon, J.-J. Embrechts, B. Goujard, A. Sakout: Evaluation of the acoustic comfort of air-treatment systems: influence of factors related to room acoustics. Proc. Forum Acusticum 2011, Aalborg.

[5] P. Legendre and L. Legendre: Numerical Ecology. Development in Environmental Modelling, Elsevier, Amsterdam, 2nd edition, 1998.