Influence of listener task on ratings of pleasantness for everyday sounds

HELEN G. CONNOR[1] AND TORBEN POULSEN[2]

[1] *Audiological Research, Widex A/S, DK-3500 Værløse, Denmark and Centre for Applied Hearing Research, Ørsted•DTU, Technical University of Denmark, DK-2800 Lyngby, Denmark*

[2] *Acoustic Technology, Ørsted•DTU, Technical University of Denmark, DK-2800 Lyngby, Denmark*

Objective: To develop a method to investigate the influence of subject task on the evaluation of sound stimuli. The method is for use in future hearing aid experiments.
Method: Twenty listeners with normal hearing rated real-life sound stimuli under different conditions. The sound stimuli were binaurally-recorded soundscapes with low-level target sounds mixed in. The conditions were:

1. Listening only to sound stimuli without any other tasks. This condition is similar to the method used in typical hearing aid studies.
2. An 'auditory detection' paradigm, where listeners detect low-level target sounds (e.g. a microwave beep) within the sound stimuli.
3. The 'irrelevant sound' paradigm, where listeners perform cognitive tasks (e.g. simple addition of numbers), while the sound stimuli are presented.

After listening to each sound stimulus under these three conditions, listeners rated the pleasantness of the sound stimulus.
Results and Conclusions**:** Ratings of auditory pleasantness were lower under the irrelevant sound condition and under the auditory detection condition than in the listening only condition. However, there was a large degree of variability associated with the ratings, which reduces the sensitivity of the method for use of evaluating hearing aid settings.

INTRODUCTION

This experiment is part of a Ph.D. project that aims to investigate hearing aid wearers' preference for the audibility of soft sounds. In order to assess preference for hearing aid settings, a suitable method must first be found that can provide the results of interest. Typically in hearing aid laboratory studies, listeners evaluate hearing aid settings by passively listening to sound stimuli and then assessing the hearing aid settings based on their perception of the stimuli. However in real life, there may be many signals competing for the listeners' attention and "listeners must locate, identify, attend to and switch attention between signals" (Noble and Gatehouse, 2004, p. 86). Also in real life, listeners may find auditory information from the environment to be distracting because it directs attention away from another task (e.g. reading). This attentional

Auditory signal processing in hearing-impaired listeners. 1st International Symposium on Auditory and Audiological Research (ISAAR 2007). T. Dau, J. M. Buchholz, J. M. Harte, T. U. Christiansen (Eds.).
ISBN: 87-990013-1-4. Print: Centertryk A/S.

aspect of hearing is not present in current typical hearing aid methodologies. It is a general aim of this experiment to create a research paradigm that combines the attentional complexity of the real world with the experimental control of the laboratory. In order to do this, the subjects are given tasks, where sound stimuli are relevant or irrelevant, in order to direct the subjects' attention either to or from the sound stimuli. In the experiment, normal-hearing listeners heard binaurally-recorded real-life sound stimuli under three conditions:

1. Listening only to the sound stimuli in a manner similar to typical hearing aid studies. This condition acts as the reference condition.
2. An auditory detection paradigm, where listeners detect low-level target sounds (e.g. a microwave beep) within the sound stimuli. This is similar to situations in real-life, in which listeners must listen to the auditory environment in anticipation of an auditory event.
3. The 'irrelevant sound effect' paradigm, in which listeners perform visual cognitive tasks (e.g. simple addition of numbers), while sound stimuli are presented. This is similar to a situation in real-life where listeners are engaged in a task, and sound is not relevant to the task. Extraneous sound has been consistently shown to impact performance on cognitive tasks, particularly short term memory tasks (see Beaman, 2005, for a recent review).

After listening to each sound stimulus under these three conditions, listeners rated the pleasantness based on their perception of the sound stimuli. The ratings in each of the three conditions are then compared in the analysis. Pleasantness was chosen as the listening criterion because it is a sound attribute that is easy for test subjects to understand and has been used as a listening criterion in a number of previous hearing aid studies. Additionally, it has been demonstrated that 'auditory unpleasantness' can be judged consistently over a wide range of stimuli (Ellermeier *et al.*, 2004).

There are two success criteria to continue using this method in future hearing aid experiments:

1. It should be demonstrated that ratings of auditory pleasantness depend on the listening condition.
2. The variability between conditions should be sufficiently low that the method can be used to detect perceptual differences between hearing aid processed sound stimuli.

METHOD

Subjects

Twenty listeners with normal or near-normal hearing were used as subjects. Nineteen of the twenty subjects were either students or employees of the Technical University of Denmark (DTU). Five of the twenty subjects were female. The age of the subjects ranged from 21 to 44 years (mean=30 years).

Materials

Equipment

The experiment was performed in a soundproof booth at the Department of Acoustic Technology, Ørsted•DTU. The presentation of the tasks and sound stimuli was controlled via MATLAB on a stationary computer. The computer sat outside the listening booth to minimise extraneous noise while the screen, keyboard and mouse were inside the booth. The screen was a 17 inch LCD screen. The sounds were presented via a good quality soundcard and HD580 precision circumaural headphones.

Sound stimuli

Each sound stimulus consisted of a background soundscape and a target sound (table 1). The soundscapes were all from the ICRA natural sound recordings (Bjerg and Larsen, 2006) and were recorded using a Head and Torso simulator. The target sounds are taken from the Digiffects CD sound effects library (Digiffects, 2007) and were mixed in at levels determined in a previous pilot experiment to give an average 70% detection rate.

Soundscape	Loudness (sone)	SPL (dB)	Fluctuation (vacil)	Target Sound	SNR
Dishwasher	12.9	72	0.55	Glass breaking	-27
Supermarket	13.5	61	1.19	Baby cry	-21
Kitchen	28.7	71	0.62	Microwave beep	-25
Pneumatic drill	43.5	78	1.57	Whistle	-22
Traffic, high	47.4	78	0.88	Car horn	-25

Table 1: Overview of the sound stimuli: soundscapes and corresponding target sounds.

The cognitive tasks

The tasks used in the ’irrelevant sound effect paradigm’ were taken from the Walter Reed Performance Battery described in Thorne *et al.* (1985) and coded into MATLAB using the Psychophysics Toolbox extension version 2.54 (Brainard, 1997). The Walter Reed battery was selected because it is designed to compare intra-subject differences across test conditions and the tests are short and do not require any prior knowledge or training.

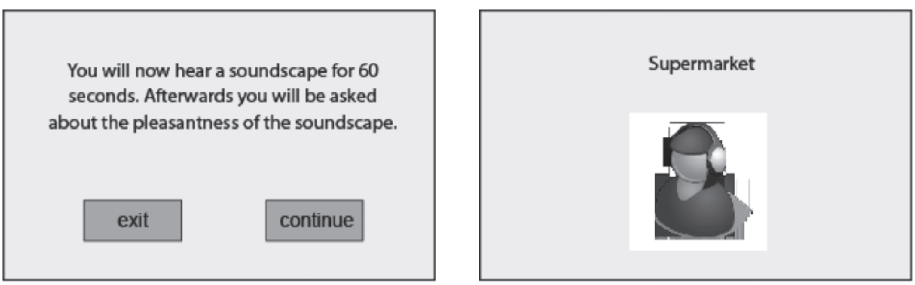
Procedure

Prior to testing, each subject listened to and rated each sound stimulus for 45 seconds to become familiar with the stimuli and the pleasantness scale. For the actual testing, subjects listened to each sound stimulus in randomised order for one minute each under the following listening conditions, ordered in a counter-balanced latin square design.

1. Listening only to the stimuli in a manner similar to typical hearing aid studies (Fig. 1.)
2. An auditory detection paradigm, where listeners detect a target sound (e.g. a microwave beep) within the sound stimulus. See Fig. 2. The target sound appears five times at randomised intervals within the one minute of listening. MATLAB registered the hit and miss rates. Prior to testing in this condition, subjects were given one training round with an easy example of a dog barking in a forest.
3. The 'Irrelevant Sound' paradigm, where listeners perform cognitive tasks in the presence of the sound stimuli. Subjects had one training round with the cognitive tasks prior to testing. Three cognitive tasks were performed in a counter balanced order: two column addition, missing picture and missing letter tasks. For brevity, only the missing letter task will be described here.

Missing letter

Nine randomised letters appear in a row for 3.3 seconds. After a 1.7 second retention interval, eight of the nine letters are re-displayed in a different random order and the subject enters the missing letter. See Fig. 3. For each sound stimulus, each subject did five missing letter tasks.



Instruction Screen — Screen while listening

Fig. 1: Screens shown to subjects during the 'listening only' test condition.

Instruction Screen. The target is presented both pictorially and as a sound over the headphones. — Screen while listening.

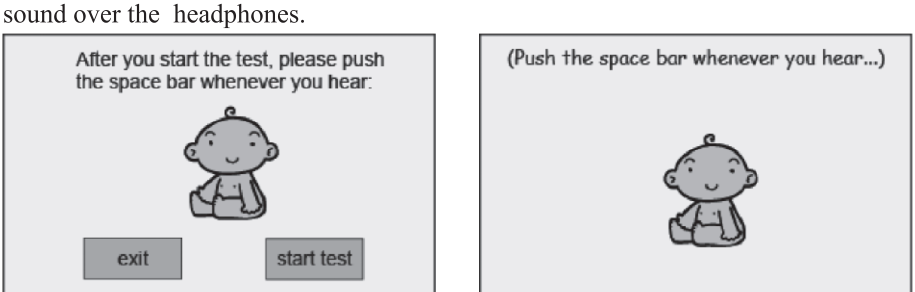


Fig. 2: Screens shown to subjects during the 'auditory detection' test condition.

Nine letters are presented in random order for 3.3 seconds.

Eight of the nine letters reappear and subject should type which letter is missing.

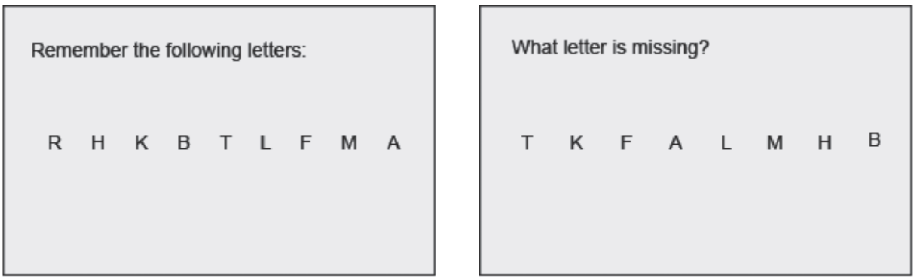


Fig. 3: Screens for the missing letter task during the 'irrelevant sound' condition.

After listening to each sound stimulus under the conditions listed above, listeners rated the pleasantness of the sound using the scale shown in Fig. 4. To avoid ceiling and floor effects, the ends of the scale are not fixed.

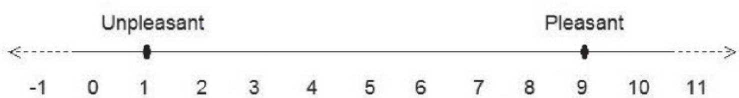


Fig. 4: Pleasantness scale used for rating the sound stimuli.

RESULTS

Effect of listening condition on ratings

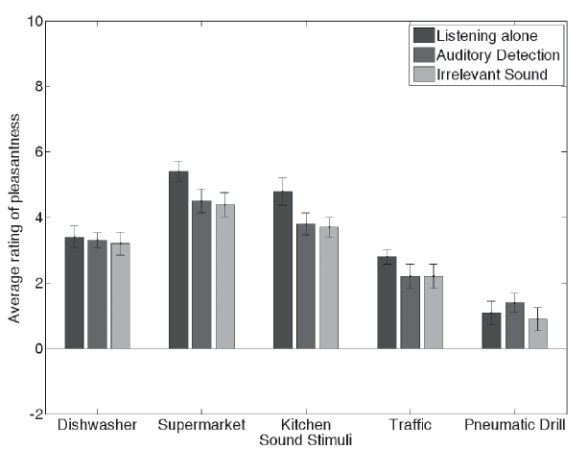


Fig. 5: Average ratings of auditory pleasantness for the five sound stimuli under the three listening conditions. The error bars represent the 95% confidence intervals.

Figure 5 shows the influence of the listening condition on average pleasantness ratings. A mixed model analysis of variance was performed using SPSS with stimuli and listening condition as fixed effects and subjects as repeated random effects. There was a significant sound stimuli effect ($p < 0.001$) and a significant condition effect ($p = 0.039$). The condition effect reflects that ratings of auditory pleasantness worsen, while either monitoring for a target sound or while performing a cognitive task, where sound is irrelevant. There was no significant interaction between sound stimuli and condition. Post-hoc pairwise analysis using a Bonferroni adjustment showed a significant difference between the ratings in the listening only condition and the irrelevant sound condition.

Variability between listening conditions

In order that this method can be used in future hearing aid experiments, the variability should be low enough to detect perceptual differences between hearing aid settings. This was assessed using a statistical power analysis to estimate the number of subjects required in a future hearing aid experiment. Firstly, the difference in ratings between conditions was calculated for each subject and each sound stimulus. The overall standard deviation for all intra-subject differences between conditions was 1.45. Secondly, a power analysis was performed in SAS. It is assumed that the differences that we want to detect between hearing aid settings are as low as 0.5 on the pleasantness rating scale. The power analysis based on a paired t-test indicated that 68 test subjects would be required to detect a difference of 0.5 on the rating scale (estimated using $\alpha = 0.05$, $\beta = 0.2$ and $SD = 1.45$). If a less sensitive test can be accepted the number of required test subjects decrease accordingly (e.g. 19 test subjects for a difference of 1.0).

Performance in the ’irrelevant sound’ condition

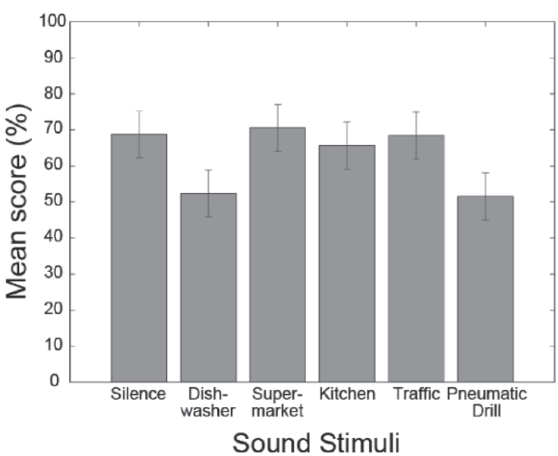


Fig. 6: Average performance for the missing letter task for the five sound stimuli and one silent stimulus. Error bars represents the 95% confidence intervals.

Figure 6 shows the effect of the sound stimuli on the missing letter task scores. The scores are percentage correct out of a total of five questions and therefore is binomially distributed and can not be treated using parametric statistics. Thus, the effect of the sound stimuli on performance was analysed using a Friedman test in SPSS. The Friedman test is a non-parametric test that is similar to the parametric repeating measure ANOVA. The result showed a significant difference between sound stimuli ($\chi 2(13,N = 19)$, $p = 0.24$).

The sound stimuli did not show any effect in the performance of the missing picture task or the two-column addition

DISCUSSION

The objective of this experiment is to develop a method to investigate the influence of the task on the evaluation of hearing aid settings. The method is assessed using the following two criteria.

Criterion 1: Effect of listening condition on ratings of auditory pleasantness

Ratings of pleasantness on average decreased when listeners either had to detect a specific target sound or perform cognitive tasks, where the sound was irrelevant. Thus it appears that the listener task does influence perception of auditory pleasantness. This finding is consistent with audiologists' anecdotal reports that when hearing aid wearers are in attentionally-complex real-world situations their hearing aids "sound worse" than in the clinic.

Criterion 2: Inter-subject variability

As indicated in the results, 68 subjects would be required in order to detect a perceptual difference of 0.5 on the pleasantness rating scale. Perceptual differences between hearing aid settings of this small magnitude have been observed in other hearing aid studies (eg. Neuman et al, 1998). With hearing-impaired subjects, it can reasonably be expected that the standard deviation will be even larger and hence the required number of subjects even larger. Paired comparisons may be a more appropriate method to assess hearing aid settings rather than ratings because paired comparisons are more sensitive to small differences between stimuli (Eisenberg et al, 1997), but it would be too complicated for the test subjects to combine paired comparisons with additional tasks, like cognitive tasks.

Other remarks

It was interesting to observe that performance on the missing letter task significantly decreased for two of the sound stimuli (the industrial dishwasher and the pneumatic drill) but not for the other sound stimuli (supermarket, kitchen and traffic). There is no obvious explanation as to why the dishwasher and drill are the most disturbing because the dishwasher was a quiet stimuli with low fluctuation and the opposite is true about the drill. One possible explanation is that the industrial dishwasher and the pneumatic drill were the least familiar stimuli. However, at least for speech stimuli, the degree of disturbance is only slightly altered by whether the language is familiar or unfamiliar.

Another possible explanation is that the dishwasher and the pneumatic drill recordings each had one dominant sound source, which sometimes turns on and off. Some studies have indicated that the amount of disturbance relates to the degree of change in one or more auditory streams, where one changing stream is more disturbing than three steady streams (Jones and Macken, 1995). A previous study has shown that the degree of disturbance can be reduced using sound processing, such as low pass filtering (Jones et al, 2000). It could be an interesting piece of future research to investigate how hearing aid processing influences performance on cognitive tasks.

CONCLUSION

The proposed method showed that ratings of pleasantness for non-processed real-life sound stimuli decreased when subjects were engaged in additional tasks. However, the ratings showed considerable inter-subject variability, which reduces the usefulness of the method to investigate perceptual differences between hearing aid settings. The performance on the missing letter task was impaired by some of the sound stimuli but not others, which poses an interesting question for future research.

ACKNOWLEDGEMENTS

A special thank you to Andrew Bell and Alice Lhomond who developed the MATLAB code for the cognitive tasks used in the irrelevant sound condition for a special course in environmental acoustics at Acoustic Technology, Ørsted•DTU. Thanks also to the twenty subjects who participated.

REFERENCES

Digiffects. (**2007**).The Digiffects Sound Effects Library [audio CD collection]. Stockholm, Sweden: Ljudproduktion AB. (URL: http://www.ljudproduktion.se. Website last updated in July 2007).

Beaman, C. P. (**2005**). “Auditory distraction from low-intensity noise: A review of the consequences for learning and workplace environments,” Applied Cognitive Psychology, **19**, 1041–1064.

Bjerg, A. P. and Larsen, J. N. (**2006**). “Recording of natural sounds for hearing aid measurements and fitting,” Masters thesis, Acoustic Technology, Ørsted•DTU, Technical University of Denmark. http://server.oersted.dtu.dk/ftp/tp

Brainard, D. H. (**1997**). “The psychophysics toolbox,” Spatial Vision, **10**, 443–446.

Ellermeier, W., Mader, M., and Daniel, P. (**2004**). “Scaling the unpleasantness of sounds ccording to the BTL model: Ratio-scale representation and psychoacoustical analysis,” Acta Acustica United with Acustica, **90**,101–107.

Eisenberg, L. S., Dirks, D. D and Gornbein, J. A. (**1997**). “Subjective judgments of speech clarity measured by paired comparisons and category ratings,” Ear and Hearing, **18**, 294-306.

Jones, D. M.; Alford, D.; Macken, W. J.; Banbury, S. P., and Tremblay, S. (**2000**). “Interference from degraded auditory stimuli: linear effects of changing-state in the irrelevant sequence,” Journal of the Acoustical Society of America, **108**, 1082-1088.

Jones, D. M., and Macken, W. J. (**1995**). “Organizational factors in the effect of irrelevant speech: the role of spatial location and timing,” Memory and Cognition, **23**, 192-200.

Neuman, A. C., Bakke, M. H., Mackersie, C., and Hellman, S. (**1998**). “The effect of compression ratio and release time on the categorical rating of sound quality,” Journal of the Acoustical Society of America, **103**, 2273-2281.

Noble, W., and Gatehouse, S. (**2004**). “Interaural asymmetry of hearing loss, speech, spatial and qualities of hearing scale (SSQ) disabilities, and handicap,” International Journal of Audiology, **43**, 100–114.

Thorne, D. R., Genser, S. G., Sing, H. C., and Hegge, F. W. (**1985**). “The Walter Reed performance assessment battery,” Neurobehaviour, Toxicology and Teratology, **7**, 415–418.