ANNOYANCE PERCEPTION AND MODELING FOR HEARING-IMPAIRED LISTENERS

Srikanth Vishnubhotla, Jinjun Xiao, Buye Xu, Martin McKinney, and *Tao Zhang*

Starkey Laboratories, Inc.
6600 Washington Ave. S., Eden Prairie, MN 55344, USA

ABSTRACT

Perceptual annoyance of environmental sounds is measured for normal-hearing and hearing-impaired listeners under iso-level and iso-loudness conditions. Data from the hearing-impaired listeners shows similar trends to that from normal-hearing listeners, but with greater variability across individuals. A regression model based on the statistics of specific loudness and other perceptual features is fit to the data from the normal-hearing listeners, and is used to predict annoyance for the hearing-impaired listeners. Differences across the subject populations are discussed.

***Index Terms*—** Psycho-acoustic annoyance, modeling, hearing impaired, hearing aids.

1. INTRODUCTION

The annoyance of sounds is an important topic in many fields, including urban design and development, transportation industries, environmental studies and hearing aid design [1, 2, 3, 4, 5, 6]. There exist established methods for subjective measurement of annoyance (e.g., [7]) and a wide variety of data on annoyance has been collected in these various fields.

The study of annoyance has been extended to include computational models that predict the annoyance of sounds based on their acoustic characteristics [8, 9] or through intermediate *psychoacoustic* models [10, 11]. While current models have limitations and are imperfect, they offer a cost-effective approach to estimating annoyance under a wide variety of conditions. This is especially helpful for those applications wherein iterative measures of annoyance are required to evaluate successive stages of system development.

A significant limitation in our current understanding of annoyance and in our ability to model it is in the treatment of hearing-impaired (HI) listeners. Most previous research has dealt with normal-hearing (NH) listeners. However, an important application of annoyance assessment is in the development of hearing aid algorithms. It is well known that HI listeners have a low tolerance for high ambient noise. This becomes especially challenging with open fittings where ambient noise can propagate directly to the ear drum without going through hearing aids. Instead of minimizing the noise level it is more effective to minimize the annoyance to improve HI subjects' comfort in such an environment. In order to do this effectively, there is a need to develop a better understanding of annoyance in HI listeners, and build computational models that reflect this understanding.

In this study, we collect and compare data on the perceived annoyance of realistic environmental noise from both NH and HI listeners to examine if hearing impairment affects annoyance perception. We focus on low-frequency noises because they can be especially troublesome for HI listeners who wear open-fit hearing aids. We also develop a simple model for annoyance based on a loudness model that takes hearing impairment into account.

2. METHOD

The test setup for the assessment of noise annoyance is described in this section.

2.1. Subjects and Stimuli

18 subjects (12 NH and 6 HI) participated in this study. They were all employees of Starkey Labs, Inc. and were not paid for their participation in this study. Fig. 1 shows the hearing loss profiles of the 6 HI subjects.

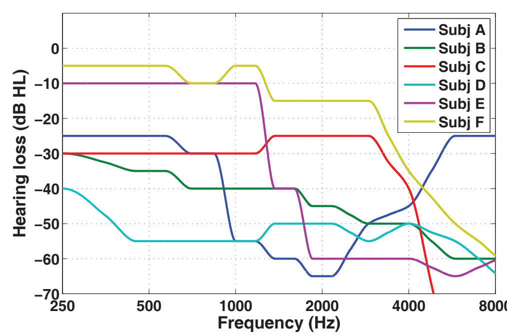

Fig. 1. Hearing loss profile of 6 HI subjects.

The stimuli set consisting of eight everyday environmental noises is given in Table 1. Each stimulus had a duration of 5 seconds and was taken from a longer recording. The stimuli were processed to produce 4 different conditions for each subject: 2 iso-loudness conditions (10 and 20 sones) and 2 iso-level conditions (NH subjects: 60 and 75 dB SPL; HI subjects: levels were chosen to match the average loudness of iso-level stimuli for NH subjects). Thus, a total of 32 stimuli were used for each subject.

Two reference stimuli, namely pink noise at 60 and 75 dB SPL, were used for the NH subjects to compare the annoyance of the stimuli set with respect to each reference. For the HI subjects, the levels were again chosen to match the loudness of that of a NH subject. We chose 2 reference stimuli under the assumption that subjects would rate annoyance more consistently if the perceived annoyance of a test stimulus is similar to that of the reference. Thus, multiple references provide greater coverage and higher probability that a test stimulus would be similar in annoyance to at least one reference stimulus. The choice of iso-loudness and iso-sound pressure levels was motivated by the desire to understand the effect of level and loudness on the annoyance experienced by both NH and HI subjects.

#1	Airplane	#5	Hair Dryer
#2	Bathroom Fan	#6	Motorcycle
#3	Car	#7	Vacuum Cleaner
#4	Engine (Diesel)	#8	Washer (Laundry)

Table 1. Stimulus set used in the test.

2.2. Procedure

The stimuli were played through a headset unilaterally in a sound treated room. In front of a computer screen, the subjects *rated the annoyance of the test stimuli relative to one of the 2 reference stimuli*. Each subject was asked to listen to one reference and a test stimulus during each trial. The annoyance of each test stimulus was rated relative to that of the reference. If the test stimulus was twice as annoying as the reference, a rating of 2 was given; If the test stimulus was half as annoying as the reference, a rating of -2 was given.

The study had a duration of about 60 minutes. A *Training block* was used to acclimatize the subjects with all the 34 stimuli (32 test stimuli and 2 reference stimuli) in the experiment. A *Testing block* then consisted of 102 ratings, wherein the subject rated each stimulus according to its annoyance level relative to that of the reference stimulus. Part of the *Testing block* was used for the subject to get acquainted with the rating task, and part of the *Testing block* was used to check the consistency of the subject on the task. Eventually 64 ratings out of the *Testing block* (32 ratings for each of the 2 references) were used in the final analysis and modeling.

3. ANNOYANCE DATA

To obtain a unique annoyance rating for each stimulus, the ratings of each stimulus based on both references were combined as a weighted average, where the weights were derived from the relative distance from each reference. This method is in line with our assumption that a rating for a test stimulus similar in perceptual annoyance to the reference is more consistent than one where the test stimulus annoyance rating is dissimilar to that of the reference stimulus. The resultant rating is the (perceptual) average relative annoyance of the stimulus. This average rating was then mapped into the logarithmic domain.

The last 18 ratings in the testing block were repetitions of the first 18 trials and were used to check the rating consistency of each subject. The correlation coefficient r between the two sets of ratings was calculated for each subject. Among the 18 subjects, 14 subjects (9 NH and 5 HI) produced high r values > 0.7. The average correlation among these 14 subjects is 0.86. Four subjects had correlations $r < 0.7$ and were deemed unreliable. Data from these four subjects were excluded from further analysis.

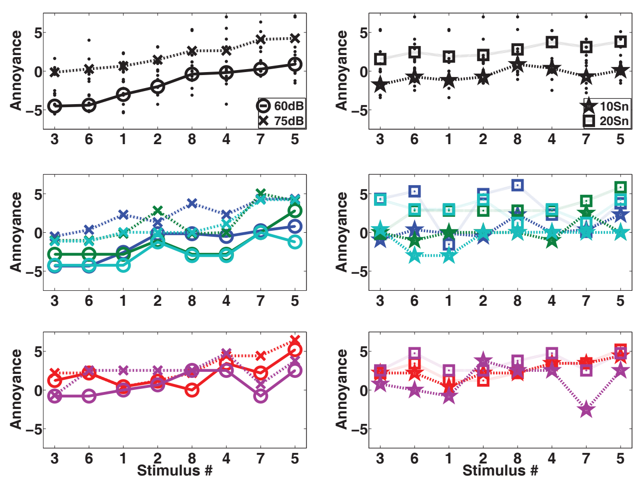

Fig. 2. Iso-level (Left) and Iso-loudness (Right) annoyance distribution for NH (black) and HI (color-coded as in Fig. 1) subjects. The legend in the first plot of each column applies to the plots in that entire column.

The annoyance ratings reported by the subjects are shown in Fig. 2 for the iso-level and iso-loudness cases as a function of stimulus type. For the NH subjects, the data points for the 9 subjects are shown in the top two plots of Fig. 2. Among these two plots, the spread of the annoyance data is shown by dots, while the mean of the scores are connected by lines. Each line represents the data for a loudness or sound pressure level, as shown in the legend. The data for HI subjects is split into two plots each for clarity. The stimuli are arranged in increasing order of their perceived annoyance by subjects. As expected, greater loudness causes subjects to report increased annoyance. Interestingly, though, it is seen that for the iso-loudness case (i.e., when all stimuli are of the same loudness) the annoyance still varies across stimuli. Similar observations can be drawn from the iso-level stimuli. The acoustic features proposed in this study are aimed at capturing the factors beyond loudness which contribute to annoyance.

Ratings for iso-loudness stimuli from HI and NH subjects were analyzed in terms of their range across subject and stimuli. For the 10-sone condition, HI and NH subjects showed similar ranges across stimuli and subjects. For the 20-sone condition, there was a tendency for NH data to show a greater range across subject and smaller mean range across stimulus, i.e., NH listeners showed less agreement overall, although the trend was not significant.

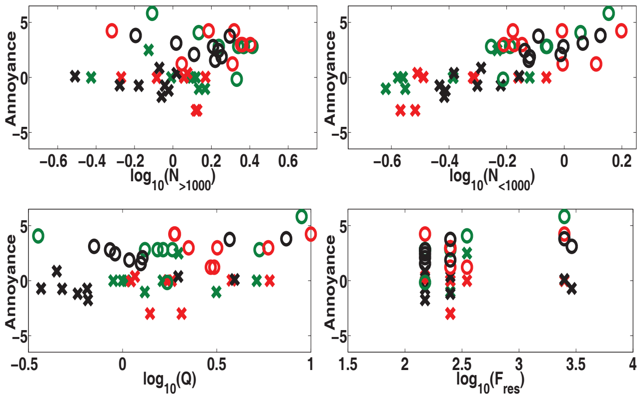

Fig. 3. Distribution of annoyance versus various features, for NH and 2 HI subjects. The data from subjects are color-coded as shown in Fig. 1. "x" and "o" symbols represent data from the 10 Sone and 20 Sone loudness stimuli respectively. The symbols for the features are explained in Sec. 4.1.

Fig. 3 shows the annoyance ratings as a function of some of the proposed features for a NH subject and 2 HI subjects, for the 2 iso-loudness cases combined across all stimuli. It can be seen that for each iso-loudness case, the annoyance is in the similar range for both NH and HI subjects. This is expected since in the iso-loudness case, the stimuli have been scaled to match each other in loudness - thus resulting in similar annoyance. Another observation is that for each of the features, annoyance varies roughly linearly with the feature value. For example, increasing specific loudness causes higher annoyance for both NH and HI subjects.

4. ANNOYANCE MODEL

In this section we propose a preliminary linear regression model for the annoyance perceived by NH subjects. We also use this model as a first attempt to predict the scores of the impaired subjects.

4.1. Feature Selection

The proposed model uses psycho-acoustically motivated features to model annoyance (c.f. [10] for a model of annoyance for NH subjects). Particularly, the features we consider are based on information derived from the Specific Loudness profile [12], including:

- $N_i : 1 \leq i \leq 24$, the *90-Percentile Specific Loudness* of the 24 critical bands, obtained by calculating the 90-Percentile level of the temporal Specific Loudness profile.
- The *Maximum Modulation Peak Value* (V_{mod}) and *Modulation Rate* (F_{mod}), which describe the spectro-temporal modulation characteristics of the stimulus. This pair of features is obtained by computing the modulation spectrum of the Specific Loudness profile, and then identifying the maximum value and location respectively, across all the bands, of the modulation spectrum. This pair of features helps in characterizing the *roughness* of a stimulus [10].
- The *Resonant Frequency* F_{res}, defined as the frequency i with the maximum value for N_i, and the *Q-Factor* (Q), defined as the ratio of the *Resonant Frequency* to the bandwidth of the N_i profile. The above two features capture the *sharpness* [10] of a stimulus.

However, due to the high dimensionality (28 dimensions) of the feature vector and limited amount of annoyance data, it is preferable to reduce the number of features before modeling. First we reduce the dimensionality of $N_i : 1 \leq i \leq 24$ to two features: *90-Percentile Specific Loudness for frequencies below 1000 Hz*, $N_{<1000}$, and *90-Percentile Specific Loudness for frequencies above 1000 Hz*, $N_{>1000}$.

Next, sequential feature selection is performed to identify the final set of features. The selection procedure starts with two features for regression, $N_{<1000}$ and $N_{>1000}$. The above four features are sequentially added as explanatory variables. The extra-sum-of-squares F-statistic [13] is calculated for each added feature, and the one with the largest F-statistic value is kept in the model if the F-statistic is greater than 4. This procedure is repeated until no further addition significantly improves the data fitting.

The feature selection process finally yields the following feature set: $\{N_{<1000}, N_{>1000}, Q, F_{res}\}$. Since the majority of stimuli in *this* test contained little modulation, the extracted modulation features F_{mod} and V_{mod} may not have been found to be statistically significant.

4.2. Annoyance Model for NH Subjects

The final linear regression model can be presented as follows:

$$A = 0.37 + 3.20N_{<1000} + 5.19N_{>1000} + 0.97Q + 1.51F_{res}$$

The weights obtained for each feature in the model support the general understanding of annoyance. In particular, an increase in the specific loudness in either frequency region (below and above 1000 Hz) predicts an increase in the annoyance rating. A larger weight for $N_{>1000}$ than that for $N_{<1000}$ implies greater annoyance sensitivity to the specific loudness in the high frequency region. As the Q-factor and the resonant frequency are related to sharpness, which directly influences annoyance [10], the annoyance is expected to increase with their values, which is consistent with the positive weights for the features.

Fig. 4 compares the predictions of the model with NH data. It can be seen that the model prediction fits the average of the NH annoyance ratings very well (e.g. $R^2 = 0.98$ for the iso-level condition), implying that this regression model has likely captured the most significant factors contributing to the average annoyance perception of NH subjects .

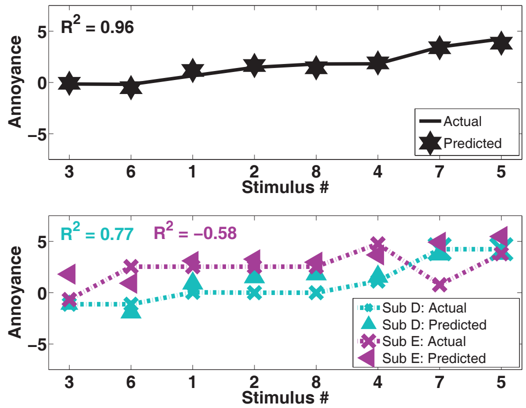

Fig. 4. Prediction of the annoyance ratings for NH and HI subjects using the Proposed Model. Subjects are color-coded as in Fig. 2.

4.3. Annoyance Model for HI Subjects

Since the NH annoyance model is based on features extracted from the specific loudness, the same model can potentially be applied to the HI data. In fact the NH annoyance model does capture the general trend of the HI subjects' annoyance ratings well, but the accuracy varies with subjects. For HI subjects A, B, and D, the NH model predicts their annoyance ratings reasonably well. A comparison between the model prediction and Subject B's annoyance ratings is shown in Fig. 4 as an example - the R^2 statistic for this subject is 0.77. For HI subjects C and E, the accuracy of the model predictions are notably worse, as seen from the example in Fig. 4.

In order to investigate the mixed results when applying NH model to HI data, further studies were carried out by fitting the data of each HI subject with the linear regression models that only contain one or two features at a time. Preliminary analysis suggests that loudness is an important contributor to the annoyance perception for some HI subjects (A, B and D) but not as significant for others (C and E), who may be more sensitive to alternative factors. Interestingly, the hearing loss profiles are steeper and more severe above 4000 Hz for the group of subjects C and E than the other group. All these observations may potentially contribute to certain aspects of the cause of the large discrepancy between the NH model and the data from HI subject C and E. However, further studies that would systemically vary stimuli and use a larger number of HI subjects are required to better understand the perceptional annoyance of HI listeners.

5. CONCLUSIONS

The annoyance data of both NH and HI subjects showed a strong dependency on overall loudness. The range of annoyance ratings for HI subjects was smaller than that for NH subjects at the high iso-loudness condition. A linear regression model incorporated with the specific loudness as well as other features was derived based on the annoyance ratings of the NH subjects. An attempt was made to apply this NH model directly to the annoyance ratings of the HI subjects. While the proposed model can account for the data from some HI subjects, it fails to accurately predict annoyance data for all HI subjects. The discrepancy between model predictions and some of the HI data highlights the need for additional investigation and refinement of the current model.

6. REFERENCES

[1] N.J. Versfeld and J. Vos, "Annoyance caused by sounds of wheeled and tracked vehicles.," *Journal Of The Acoustical Society Of America*, vol. 101, no. 5 Pt 1, pp. 2677–2685, May 1997.

[2] M. Alayrac, C. Marquis-Favre, S. Viollon, J. Morel, and G. Le Nost, "Annoyance from industrial noise: indicators for a wide variety of industrial sources.," *Journal Of The Acoustical Society Of America*, vol. 128, no. 3, pp. 1128–1139, Sept. 2010.

[3] E. Pedersen and K.P. Waye, "Perception and annoyance due to wind turbine noise–a dose-response relationship.," *Journal Of The Acoustical Society Of America*, vol. 116, no. 6, pp. 3460–3470, Dec. 2004.

[4] R. Rylander, "Noise, stress and annoyance," *Noise & Vibration Worldwide*, 2006.

[5] F Letourneaux, S Guerrand, and F Poisson, "Assessment of the acoustical comfort in high-speed trains at the SNCF: integration of subjective parameters," *Journal of Sound and Vibration*, vol. 231, no. 3, pp. 839–846, 2000.

[6] C.V. Palmer, R. Bentler, and G.H. Mueller, "Amplification With Digital Noise Reduction and the Perception of Annoying and Aversive Sounds," *Trends in Amplification*, vol. 10, no. 2, pp. 95–104, 2006.

[7] U Sailer and M Hassenzahl, "Assessing noise annoyance: an improvement-oriented approach.," *Ergonomics*, vol. 43, no. 11, pp. 1920–1938, Nov. 2000.

[8] N Versfeld, "A Weighted Equivalent Sound Level As A Predictor Of The Annoyance Caused By Road Traffic Consisting Of Various Proportions Of Light And Heavy Vehicles," *Journal of Sound and Vibration*, vol. 253, no. 2, pp. 389–399, May 2002.

[9] P Davies, "An Overview Of Methods To Quantify Annoyance Due To Noise With Application To Tire-Road Noise," Tech. Rep., 2008.

[10] H. Fastl and E. Zwicker, *Psychoacoustics: Facts and Models (Springer Series in Information Sciences)*, Springer, 3rd edition, Dec. 2006.

[11] W. Ellermeier, M. Mader, and P. Daniel, "Scaling the Unpleasantness of Sounds According to the BTL Model: Ratio-Scale Representation and Psychoacoustical Analysis," *Acta Acustica united with Acustica*, vol. 90, pp. 101–107, 2004.

[12] J. Chalupper and H. Fastl, "Dynamic Loudness Model (DLM) for Normal and Hearing-Impaired Listeners," *Acta Acustica united with Acustica*, vol. 88, no. 3, pp. 378–386, 2002.

[13] David S. Moore and George P. McCabe, *Introduction to the Practice of Statistics*, W. H. Freeman and Company, 3rd edition, 1999.