PSYCHOACOUSTICS
A Perceptual Modeling of Acoustic Events Focused on Spatial Impression
Manabu Fukushima\textsuperscript{a}, and Hirofumi Yanagawa\textsuperscript{b}

\textsuperscript{a} Fukuoka Institute of Technology, 3-30-1 Wajiro-higashi Higashi-ku Fukuoka 811-0295 Japan
\textsuperscript{b} Chiba Institute of Technology, 2-17-1 Tsudanuma Narashino Chiba 275-0016 Japan

This paper describes the relation between physical/acoustic parameters and psychological scale for the sound fields in order to create an artificial impulse response of the room based on the perception. First, 19 specific words were chosen expressing subjective impressions of the sound field from a Japanese language dictionary with 42,000 vocabularies. To classify the 19 words, speech sounds are compared in the way of dichotic listening. The speech sounds are convolution of an anechoic speech and impulse responses of rooms measured by using a dummy head microphone. The words are clustered into 4 categories, 1) high tone timbre, 2) low tone timbre, 3) spaciousness and 4) naturalness or clearness. Then, the 'spatial impression' was selected among 19 words and a scale of it was obtained by way of Thurston's case V since it is one of the important factors in the sound field design. Second, to create an impulse response corresponding to the 'spatial impression', we investigate the relation between the 'spatial impression' and physical/acoustic parameters. As a result, we found that the initial part of impulse response is an important part for controlling 'spatial impression'. The result is confirmed by listening test using artificial impulse responses.

INTRODUCTION

The purpose of our study is to realize a system that provides a virtual sound space sharing in a network environment. We call the system ISFN (Interactive Sound Field Network). An interactive use of the sound space sharing such as a virtual conference room is needed real time reproduction of room impulse responses.

When we share the sound space by using a network with limited bandwidth, the amount of the data to reproduce the sound space should be reduced. As the data reduction technique, MPEG-4 proposes structured sound description\textsuperscript{7}. However it deals less with the sound field than with sound signal. Acoustic Events Modeling Language (AEML)\textsuperscript{3}, which is the language for describing acoustic events including description of the sound field and reproduces impulse responses of the sound space. The characteristic of a human perception for the sound field can be applied to reduce the amount of the data by describing the sound field using AEML. A concept of ISFN and its modules are figured out as figure 1.

![Fig.1 Concept of Interactive Sound Field Network](image)

It is necessary to scale the characteristic of a human perception (psychological scale) for the AEML. In this study, the relation between physical parameters and psychological scale for the sound field is examined in order to create the room impulse response based on the perception of the sound field.

WORDS EXPRESSING THE SOUND FIELD

We chose words expressing subjective impressions of the sound field from a Japanese language dictionary\textsuperscript{4} which contains about 42,000 vocabularies. For the first step, native Japanese speakers chose candidate words from the dictionary. The words were judged from their usability when they were used as adjective for a sound space. Furthermore, the words were selected on the condition that a scene of the sound field could be pictured with the word. Finally, 19 words were picked out as shown in table 1.

\begin{table}[h]
\centering
\begin{tabular}{|l|l|l|l|}
\hline
\textbf{English words} & \textbf{English words} & \textbf{English words} \\
\hline
spacious & clear & heavy \\
reverberate & rough & dry \\
echoic & hard & flutter echoic \\
deep & heavy & natural \\
hollow & brilliant & beautiful \\
clean & bright & pleasant \\
booming & & \\
\hline
\end{tabular}
\caption{Selected words that expressing subjective impression of the sound field}
\end{table}

![Fig.2 The result of clustering the specific words](image)
headphones used are Pioneer SE-900D. Test signals were generated as a convolution of impulse responses of 7 rooms which volume ranges from 0.53\([\text{m}^3]\) to 6041\([\text{m}^3]\) and an echoic female speech signal\(^5\). The impulse responses were measured by using a dummy head microphone (OSS HATS). The words are clustered into 4 categories by using a statistical analysis software SPSS as shown in figure 2, which are (1) high tone timbre, (2) low tone timbre, (3) spaciousness and (4) naturalness or clearness.

**PSYCHOLOGICAL SCALE FOR 'SPACIOUS'**

We focus on a word 'spacious'\(^7\) that belongs to (3) spaciousness, since it is one of the important factors in the sound field design. It is necessary to scale the psychological result for creating impulse responses. For scaling the 'spacious', we add other 8 rooms (totally 15 rooms) in the hearing test. Thurstone's case V\(^5\) was applied to scale.

In order to find the relation between the psychological scale and the physical parameters, we apply 9 physical parameters, 1) room volume, 2) D: deutlichkeit, 3) C: clarity, 4) R: hallmass, 5) STI: speech transmission index, 6) EDT: early decay time, 7) Ts: center time, 8) RT: reverberation time and 9) IACC: inter aural cross correlation. Figure 3 shows the relation between the result of the scaling 'spacious' and physical parameters.

From figure 3, the 'spacious' correlates with most of the physical parameters and seems to be controlled by them. To confirm it, the hearing test was done by using simulated impulse responses instead of room impulse responses. The impulse response was calculated by an image method for a rectangular room. Then, it was changed in two ways. (i) The ratio between the direct sound and the reverberation was altered. This causes change of Ts, D, R and so on. (ii) The decay rate was altered between an initial part of the impulse response and the other part of it. This means change of EDT. The result of the hearing test by diotic listening is shown at an upper half of figure 4 where Ts is a value changed in the way of (i) and also EDT in the way of (ii). Lower two figures in figure 4 are extracts from figure 3. Figure 4 shows that the 'spacious' is controlled by changing Ts and EDT.

**SUMMARY**

This paper described the relationship between physical parameters and psychological scale for the sound field to create the room impulse response that causes a specific subjective impression of the sound field. Firstly, we clustered 19 words expressing subjective impressions of the sound field into 4 categories, which are (1) high tone timbre, (2) low tone timbre, (3) spaciousness and (4) naturalness or clearness. Secondly, we focused on a word ‘spacious’ that belongs to (3) spaciousness, since it is one of the important factors in the sound field design. We scaled the 'spacious' by Thurstone's case V. As a result, 'spacious' correlates with most of the physical parameters of the sound field and can be controlled by them.

**REFERENCES**

5. DENON Professional Test CDs COCO-75084-~86,DISK2-Track37
Detection Thresholds for Pure Tones Presented against a Broadband Noise - A Comparison of Young and Elderly Listeners -

Kurakata K\(^a\), Matsushita K\(^b\), Shibasaki-K. A\(^c\) and Kuchinomachi Y.\(^a\)

\(^a\) Natl. Inst. of Advanced Industrial Science and Technology, 305-8566 Tsukuba, Japan; kurakata-k@aist.go.jp
\(^b\) Natl. Inst. of Technology and Evaluation, 305-0044 Tsukuba, Japan
\(^c\) University of Tsukuba/JSPS, 305-8573 Tsukuba, Japan

Thresholds for detecting pure tones presented against a broadband noise were investigated with young and elderly listeners. Target tones were manipulated in terms of (1) frequency, (2) duration, and (3) the number of repetitions. The results showed the following: (1) The thresholds for the elderly listeners were 5-10 dB higher for target tones with frequencies of 250-4000 Hz. (2) As the duration of target tone increased, the detection threshold for 1000-Hz tone decreased at the same rate in both age groups. However, the threshold decrease for the 4000-Hz tone was smaller in elderly listeners with severe hearing loss, indicating a deterioration of temporal-summation ability in the high-frequency region. (3) When the target tone was repeatedly presented, detection thresholds decreased in both groups. The threshold for the elderly with severe hearing loss dropped with longer repeated sequences than the young listeners and the elderly listeners with moderate hearing loss. This suggests that the deterioration in the temporal process occurs not only at an early stage of auditory processing, but also at a higher, cognitive stage that utilizes information from input sounds ranging over a few seconds. These results provide a basis for determining the appropriate properties for auditory signals, such as those used in domestic appliances.

INTRODUCTION

Recent consumer products such as domestic appliances and personal electric equipment often employ auditory signals in their user-interface. The signals notify the user when something is wrong with the machine or when some process is complete. Such interfaces with auditory signals are thought to improve the usability of these products, especially for the elderly who are not accustomed to such electric appliances.

In designing suitable auditory signals for the elderly, we need to take into account their hearing ability and the masking effect of background noise. The elderly usually have hearing loss in the high frequency region and their hearing ability can be negatively affected by external noises. Therefore, auditory signals need to have acoustic properties to ensure that the elderly can easily perceive them even in noisy situations.

Although there have been many studies on the effect of aging on auditory functions, there is still insufficient data for the practical designing of auditory signals for the elderly. In this study, we measured the detection thresholds of the elderly for pure tones against a broadband noise. The target tone was modified in terms of (1) frequency, (2) duration, and (3) the number of repetitions. Comparison of the thresholds in elderly listeners with those in young listeners provides criteria for choosing auditory signals suitable for the elderly as well as for young people.

EXPERIMENT

The target tones were pure tones with frequencies of 250, 500, 1000, 2000, or 4000 Hz. The durations were 50, 100, 200, 500, or 1000 ms, with a rise/fall time of 20 ms. To measure the effect of repetition, a 1000-Hz tone of 100-ms duration was repeated 2, 5, or 10 times, with 100-ms intervals. These target tones were presented against a continuous broadband noise of 60 dBA.

A two-alternative forced choice method with the PEST procedure[1] was used to determine detection thresholds for the target tones. The target tones and background noise were presented via a loudspeaker placed at the distance of 2.5 m from the listener. At each trial, target tone was presented at one of two intervals indicated by a red LED. Listeners were asked to judge after each presentation which interval had contained the tone. A stimulus level of 75% correct response was adopted as the threshold level. Every listener participated in all stimulus conditions and one measurement series was run for each condition.

Fourteen university students aged from 19 to 24 and 45 elderly persons aged from 60 to 83 participated in the experiment. Young listeners were screened by a hearing test prior to the experiment and confirmed to have normal hearing. The elderly listeners were divided further into two groups according to their hearing threshold level (HTL) of better ear averaged over 250-4000 Hz: (1) 30 listeners with an average HTL of 25 dB or less (group HTL\(_L\)), (2) 15 listeners with an average HTL over 25 dB (group HTL\(_H\)).

RESULTS AND DISCUSSION

Figure 1 shows the detection thresholds of the three groups of listeners as a function of the frequency of the target tone. The duration of the tone was 1000 ms. The
elderly groups show higher thresholds at all frequencies than the young group, indicating that they have more difficulty detecting a pure tone against noise. This threshold elevation is expected due to the fact that the width of auditory filters is larger for the elderly[2], so that the "signal-to-noise ratio" at the target frequency becomes worse. The difference in thresholds amounts to 5-10 dB for the HTL_{40} group. Therefore, in order to make auditory signals audible for elderly users, their sound levels need to be 5-10 dB higher than those designed for young users.

FIGURE 1. Detection thresholds of the three groups of listeners as a function of the frequency of the target tone. The median values for thresholds relative to that of the young group and the interquartile range within each group are shown.

Figure 2 (a) and (b) show the detection thresholds of the three groups as a function of the duration of the target tone. The frequency of the tone was (a) 1000 or (b) 4000 Hz. The curves for both elderly groups are almost parallel to that of the young group for the 1000-Hz tone, although the curve for the HTL_{40} group for the 4000-Hz tone significantly deviates from those of young and HTL_{40} groups. This suggests that the temporal-summation ability deteriorates in the elderly with severe hearing loss. Thus, although using a longer signal may help for detection in the elderly with moderate hearing loss, it is not for those with severe losses. Rather, increasing the level of signal may be more beneficial for the latter group.

FIGURE 2. Detection thresholds of the three groups of listeners as a function of the duration of the target tone, (a) 1000 Hz, (b) 4000 Hz. The median values for thresholds relative to that for the 50-ms tone and the interquartile range within each group are shown.

Figure 3 shows the detection thresholds of the three groups as a function of the number of repetitions. In general, as the target tone was repeated more times, the threshold decreased in every group. The threshold of the HTL_{40} group remained roughly constant up to repetitions of 5 times and dropped at 10 times, whereas the thresholds for the other two groups began to decrease at 5 times. As each stimulus tone was the same for every repetition sequence, the detection threshold seems to depend on the listener's ability at a higher, cognitive stage to process the information from several input sounds. The elderly with severe hearing loss may have difficulty to retain information from repeated tones and effectively utilize this to enhance detectability. Longer sequences of signals would be desirable for elderly listeners to attract their attention.

FIGURE 3. Detection thresholds of the three groups as a function of the number of repetitions. The median values for thresholds relative to that for a single tone (number of repetitions = 1) and the interquartile range within each group are shown.

REFERENCES

Pitch strength of ordered click-train regular interval stimuli

D. Mapes-Riordan\textsuperscript{a,b} and W. A. Yost\textsuperscript{a}

\textsuperscript{a}Parmly Hearing Institute, Loyola University Chicago, Chicago, IL, 60626 USA, wyost@luc.edu
\textsuperscript{b}DMR Consulting, Inc., Evanston, IL, 60201 USA, danmr@mediaone.net

Regular interval stimuli (RIS) contain temporal regularities that are known to produce a pitch percept. Previous experiments\cite{1} have shown that shuffled, wideband first-order click-train RIS have a stronger pitch strength than the original, unshuffled stimuli; where \( k \) is a fixed duration interval and \( x \) is a random interval. This result provides evidence for the importance of short-term periodicity within RIS in determining its pitch strength. Further pitch strength comparison experiments were run using ordered click-train RIS. The purpose of these experiments was to compare the height of the first peak in the autocorrelation of the stimuli (AC1) and the number of consecutive regular intervals as measures of pitch strength in click-train RIS. This was accomplished by creating stimuli with lower AC1 peak height but also have more consecutive regular intervals than the other compared stimuli. The results show that the pitch strength of these stimuli is more a function of the number of consecutive regular intervals than a function of the AC1 peak height.

\section*{INTRODUCTION}

RIS are useful for studying pitch since the amount of temporal regularity (i.e., pitch strength) can be controlled. One measure that has been used to predict pitch strength is the peak height of the first non-zero delay peak in the normalized, long-term autocorrelation of the stimulus (AC1). However, previous experiments\cite{1,2} have shown that AC1 may not always accurately reflect changes in pitch strength. For example, shuffling first-order click-train RIS has been shown to increase its pitch strength even though the long-term autocorrelation remains the same\cite{1}. These and other results suggest that short-term periodicity may be a better predictor of pitch strength for these stimuli.

The purpose of the current experiments was to compare the AC1 peak height and the number of consecutive regular intervals as measures of pitch strength in click-train RIS. This was accomplished by creating stimuli with lower AC1 peak height yet more consecutive regular intervals than the other compared stimuli. The AC1 peak height is equal to the number of regular \( k \) intervals divided by the total number of regular \( k \) and random \( x \) intervals. In the first set of trials, the pitch strength of \( kx \) stimuli with an AC1 peak height of 0.5 was compared to stimuli with a lower AC1 peak height (0.33) but consisted of multiple consecutive regular \( k \) intervals. In a second set of trials, \( 4k8x \) stimuli that has an AC1 peak height (0.33) was compared to stimuli with a higher AC1 peak height (0.5) but fewer multiple consecutive regular \( k \) intervals.

\section*{EXPERIMENT}

Pitch strength comparisons were made using a 2AFC task; one stimulus was the standard and the other an ordered stimulus with a different AC1 peak height made up from a different proportion of regular and random intervals. Listeners were asked to pick the stimuli that had the stronger pitch strength. A 300-ms silent gap was inserted between stimuli. Listeners were provided visual reinforcement that indicated when each stimulus was being presented. A 1-s delay was inserted after each response before the next trial began.

Regular ICIs with fundamental frequencies 125, 250, 500, and 1000 Hz were used. In the first set of trials, the standard stimulus was ordered \( kx \) click-train RIS (AC1 = 0.5) and was paired with ordered \( 2k4x, 3k6x, 4k8x, \) or \( 5k10x \) click-train RIS (AC1 = 0.33). In the second set of trials, the standard stimulus was ordered \( 4k8x \) click-train RIS (AC1 = 0.33) and was paired with ordered \( kx, 2k2x, 3k3x, \) or \( 4k4x \) click-train RIS (AC1 = 0.5). A block of trials consisted of the four different pairings at a single fundamental frequency. Ten trials were run at each condition resulting in forty trials per block. Listeners ran four blocks of trials...
without a break. Trials were randomly selected without replacement in each block. A block of forty trials was repeated five times throughout the experiment resulting in fifty trials per condition for each listener.

Results and Discussion

The overall results of the first set of trials are shown in Figure 1. The results are presented as the percentage of trials the target stimuli (AC1 = 0.33) were judged to have a stronger pitch strength than the standard stimuli (AC1 = 0.5). The 50% line in Figure 1 represents equal pitch strength. Figure 1 shows that, when f0 is equal to 500 and 1000 Hz, only two consecutive regular intervals were required to make the pitch strength greater than the standard stimuli which had a higher AC1 peak height. At f0 equal to 125 and 250 Hz, four consecutive regular intervals were required for the mean pitch strength to be stronger than the standard stimuli.

The overall results of the second set of trials are shown in Figure 2. The results are presented as the percentage of trials the target stimuli (AC1 = 0.5) were judged to have a stronger pitch strength than the standard stimuli (AC1 = 0.33). Note that the kx condition in the second set of trials (Fig. 2) is the same as the 4k8x condition in the first set of trials (Fig. 1). Figure 2 shows that three consecutive regular intervals can have a larger pitch strength than four consecutive regular intervals if the proportion of regular intervals is 0.5 as compared to 0.33. The change in pitch strength is most striking for the 500 and 1000 Hz fundamental frequency conditions which show an abrupt change in preference from the 2k2x to the 3k3x conditions.

The results of these ordered click-train RIS pitch strength comparisons confirm the importance of consecutive regular intervals in determining the pitch strength of these stimuli, which is more a function of the number of consecutive regular intervals than a function of AC1 peak height. These results are consistent with a statistical analysis of shuffled click-train RIS pitch strength comparisons which indicate that short runs of consecutive regular intervals can have a large effect on pitch strength[1].

FIGURE 1. Ordered click-train RIS pitch strength comparisons: kx vs. 2k4x, 3k6x, 4k8x, 5k10x. Wide band stimuli, total duration = 500 ms, bars are mean percentage of four listeners, error bars are standard errors.

FIGURE 2. Ordered click-train RIS pitch strength comparisons: 4k8x vs. kx, 2k2x, 3k3x, 4k4x. Wide band stimuli, total duration = 500 ms, bars are mean percentage of four listeners, error bars are standard errors.

ACKNOWLEDGMENTS

This work was supported by a National Institute of Health (NIDCD) Program Project Grant.

REFERENCES


Thresholds of Hearing for FM Tones Coming from Various Directions

T. Letowski, K. Abouchacra, T. Tran, and J. Kalb

Human Research and Engineering Directorate, U.S. Army Research Laboratory
Aberdeen Proving Ground, MD 21005-5425 USA

Sound-field hearing thresholds were measured using frequency-modulated tones (FM tones) and five horizontal incidence angles (0°, 45°, 90°, 135°, and 180°). The tests were conducted in a large audiometric room meeting ANSI criteria for hearing testing in a sound field. Monaural and binaural thresholds of hearing were measured at 125, 250, 1000, 4000, and 8000 Hz for ten young listeners (age 18 through 25) during two test sessions. In Session I, the test signals were generated by five loudspeakers located on a semi-circle and separated by 45°. In Session II, only one loudspeaker was used for signal generation while the listener was facing the five directions corresponding to the loudspeaker positions in Session I. The results of the study provided reference equivalent threshold sound pressure levels (RETSPLs) for 125-8000 Hz FM tones arriving at the listeners in the 0-180° azimuth range. Intersubject variability of the data obtained in Session I was larger than that obtained in Session II.

INTRODUCTION

Sound-field audiology is an area of audiology where there is little consensus regarding test methodology as is shown by the wide variation of clinical practice [1,2]. Especially troubling is the large uncertainty and variability of reference equivalent threshold sound pressure level (RETSPL) values used in sound-field audiology. Existing standards [3,4], recommended clinical practices [5], and numerous scientific reports differ greatly regarding the RETSPLs for specific signals, frequencies, and angles of signal incidence. The signals of choice are frequency-modulated tones (FM tones). Yet, their RETSPLs are based on pure-tone studies and a few piecemeal warble-tone data. No single FM-tone study has addressed the whole audiometric range of frequencies and the whole range of incidence angles used in clinical practice (0°, 45°, and 90°). There are also no hearing threshold data for FM tones arriving outside the 0-90° incidence range. Yet, the knowledge of the whole 360° range of hearing thresholds is important for both civilian (hearing aids) and military (situation awareness) applications.

The purpose of this study was to determine the effect of an angle of incidence on the monaural and binaural thresholds of hearing for FM tones. A single speech signal (“northwest”) was also included for comparison. The data will be used as a baseline in our future studies aimed to determine the effects of protective headgear on auditory perception.

METHOD

Ten listeners (five males and five females) between the ages of 18 and 25 participated in the study. All listeners had pure-tone hearing thresholds better than or equal to 20 dB HL at audiometric frequencies from 125 Hz through 8000 Hz (ANSI S3.6-1996). The difference in hearing thresholds in both ears did not exceed 5 dB at any of the test frequencies.

The test signals were 125Hz, 250Hz, 1000Hz, 4000Hz, and 8000Hz FM tones (f_m=5Hz, Δf=±10%) and a speech signal The speech signal was a spondee word “northwest” that is used as a standard speech signal in our studies.

Test signals were generated by five matched loudspeakers (BOSE 108515K) mounted 45° apart on a semicircle with the radius of 1.1 m. The main axis of each loudspeaker was perpendicular to the tangent line at the loudspeaker location. The loudspeakers were located 1.2m above the floor level, which is ear level of an average person in a sitting position.

A Bekesy tracking procedure was used for data collection. The test signal was played for 30 sec. The signal level changed at the rate of 5 dB/sec and the direction of change was controlled by the listener using a hand-held switch. The listener’s task was to track the hearing threshold by pressing and releasing the switch. The hearing threshold was determined as the mean level of the longest uniform part of the tracking period.

Each listener participated in two 30-minute listening sessions. In Session I, the signals were generated by five loudspeakers oriented at 0°, 45°, 90°, 135°, and 180° re. the listener position. In Session II, the signals were generated by a single loudspeaker while the listener was oriented at 0°, 45°, 90°, 135°, and 180° re. the loudspeaker position. In both cases, the angular position of 0° indicated the configuration in which the listener was facing the loudspeaker. Monaural (left ear) and binaural hearing thresholds were measured during each test session. The order of signals, orientations, and sessions was counterbalanced.
RESULTS

Results of the study are presented in Tables 1-4.

**Table 1.** Monaural hearing thresholds (dB SPL) for FM tones emitted from five loudspeakers placed around the listener.

<table>
<thead>
<tr>
<th>Signal frequency</th>
<th>Signal incidence angle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0°</td>
</tr>
<tr>
<td>125 Hz</td>
<td>25.9</td>
</tr>
<tr>
<td>250 Hz</td>
<td>14.1</td>
</tr>
<tr>
<td>1000 Hz</td>
<td>10.8</td>
</tr>
<tr>
<td>4000 Hz</td>
<td>4.5</td>
</tr>
<tr>
<td>8000 Hz</td>
<td>13.4</td>
</tr>
<tr>
<td>Speech</td>
<td>15.1</td>
</tr>
</tbody>
</table>

**Table 2.** Binaural hearing thresholds (dB SPL) for FM tones emitted from five loudspeakers placed around the listener.

<table>
<thead>
<tr>
<th>Signal frequency</th>
<th>Signal incidence angle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0°</td>
</tr>
<tr>
<td>125 Hz</td>
<td>25.1</td>
</tr>
<tr>
<td>250 Hz</td>
<td>13.1</td>
</tr>
<tr>
<td>1000 Hz</td>
<td>10.9</td>
</tr>
<tr>
<td>4000 Hz</td>
<td>3.5</td>
</tr>
<tr>
<td>8000 Hz</td>
<td>10.5</td>
</tr>
<tr>
<td>Speech</td>
<td>12.9</td>
</tr>
</tbody>
</table>

**Table 3.** Monaural hearing thresholds (dB SPL) for FM tones emitted from a single loudspeaker at five listener orientations.

<table>
<thead>
<tr>
<th>Signal frequency</th>
<th>Signal incidence angle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0°</td>
</tr>
<tr>
<td>125 Hz</td>
<td>26.6</td>
</tr>
<tr>
<td>250 Hz</td>
<td>15.0</td>
</tr>
<tr>
<td>1000 Hz</td>
<td>11.3</td>
</tr>
<tr>
<td>4000 Hz</td>
<td>4.9</td>
</tr>
<tr>
<td>8000 Hz</td>
<td>13.0</td>
</tr>
<tr>
<td>Speech</td>
<td>15.0</td>
</tr>
</tbody>
</table>

**Table 4.** Binaural hearing thresholds (dB SPL) for FM tones emitted from a single loudspeaker at five listener orientations.

<table>
<thead>
<tr>
<th>Signal frequency</th>
<th>Signal incidence angle</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0°</td>
</tr>
<tr>
<td>125 Hz</td>
<td>23.4</td>
</tr>
<tr>
<td>250 Hz</td>
<td>13.1</td>
</tr>
<tr>
<td>1000 Hz</td>
<td>9.1</td>
</tr>
<tr>
<td>4000 Hz</td>
<td>2.8</td>
</tr>
<tr>
<td>8000 Hz</td>
<td>10.2</td>
</tr>
<tr>
<td>Speech</td>
<td>12.6</td>
</tr>
</tbody>
</table>

DISCUSSION AND CONCLUSIONS

The data presented in Tables 1-4 do not clearly support any single set of previously published directional hearing thresholds although there are some similarities. In general, the results of this study show less dependence of the hearing threshold on the angle of incidence than the data reported previously in the literature. This result may be explained by the nature of the test signals and acoustic properties of the test room used in this study. Since the test conditions and signals investigated are commonly used in clinical sound-field audiometry, it can be argued that the reported data are representative of the settings used in audiology clinics.

The average directional thresholds obtained in Session I (five loudspeakers) and Session II (single loudspeaker) were very similar. However, the variability of the data obtained with five loudspeakers was typically much greater than that of the data obtained with a single loudspeaker under the same test conditions. Assuming that this effect was not caused by the differences between loudspeakers (all loudspeakers were carefully matched), this finding may indicate that the orientation of the loudspeaker in the test room had a greater effect on the hearing test data than the orientation of the listener.

Binaural hearing thresholds were 1-3 dB better than monaural thresholds when measured at 0° incidence angle. There were two cases (out of 12) where the difference was less than 1 dB and both cases corresponded to threshold conditions measured in Session I (five loudspeakers). The directional pattern of binaural thresholds was shallower than that of monaural thresholds and this difference affected directional relationship between the thresholds.

ACKNOWLEDGMENT

We wish to thank Amanda Rowe and Beth Schafer for their help with data collection. The study was conducted as a part of ANSI S3/WG83 “Sound Field Audiometry” data collection activities.

REFERENCES

Effects of Preceding and Following Tone on Laterality Threshold

M. Ebata\textsuperscript{a}, K. Takahashi\textsuperscript{b}, T. Ishiodori\textsuperscript{b}, H. Miyazono\textsuperscript{b}, O. Komoda\textsuperscript{b}
and T. Usagawa\textsuperscript{b}

\textsuperscript{a}Kumamoto National College of Technology, 2659-2 Suya, nishigoshi, Kumamoto 861-1102 Japan
\textsuperscript{b}Graduate School of Science and Technology, Kumamoto University, 2-39-1 Kurokami, Kumamoto 860-8555 Japan

In this paper, the thresholds of detection and laterality are explored using a tone signal with interaural time difference and a diotic tone masker in the three stimulus configurations: in simultaneous, forward and backward masking paradigms. The results show that the characteristics of the laterality thresholds are similar to the well-known pattern of temporal masking. The laterality thresholds with an interferer, however, are higher than those without an interferer at a wider frequency range around the frequency of an interfering tone.

INTRODUCTION

Many studies have been reported on the precedence effect in the past few decades. Tolin and Henning(1998) \cite{1} recently measured the time course of the precedence effect in terms of the lateralizability of the sound produced when dichotic clicks were either preceded or followed by another diotic click. Under the condition that a dichotic click with a given interaural time delay (ITD) followed a diotic click (in a forward masking paradigm), observers’ sensitivity to ITD was degraded for a period from approximately 1 to 10 ms after the onset. They also measured threshold ITDs when a dichotic click preceded a diotic click (in a backward masking paradigm), and obtained the results that the threshold ITDs in this condition are almost the same as those in quiet conditions. Almost all of the results obtained show the deterioration of sound localization by noise presented simultaneously or temporally. From the view point of information processing in an auditory system, however, the precedence effect can be regarded as the forward masking of binaural information. It has been suggested that the precedence effect is related to the properties of temporal masking, but the relationship has not been explored extensively. In this paper, the interference in the binaural information is explored from the view point of temporal masking. Specifically not only the forward masking but also the backward masking is investigated.

METHODS

A signal detection experiment was done in three stimulus configurations: in simultaneous, forward and backward masking paradigms. For each stimulus configuration, a masker was presented diotically. The test tone was a 5 ms tone burst with an interaural time difference of 0.5 ms and with 1 ms-ramps in rising and falling. The frequencies of the tone signal were 0.5, 1, 2, and 4 kHz. The masker was a 100 ms tone burst of 1 or 4 kHz, which also has 1 ms-ramps and was presented through headphones. The steady-state level of the masker was 70 dB SPL. A two-interval two-alternative forced-choice procedure was used and subjects were asked to detect the tone signal with the interaural time difference by responding to one of the two intervals. The interval between tone pairs was 500 ms.

The laterality threshold was also explored in three stimulus configurations. A single-interval two-alternative forced-choice procedure was used and subjects were asked to indicate to which side of midline the test tone was lateralized. The same stimuli used in the detection experiment were used. A 3-down 1-up adaptive procedure was used to measure the threshold of the detection and laterality of the tone burst.

RESULTS AND DISCUSSION

Figure 1 and 2 are the results showing the characteristics of the threshold of detection and laterality as a function of the frequency of a test tone in the forward masking paradigm. In the figures, laterality thresholds are shown by open circles and detection thresholds are shown by cross points. The parameter in the figures is ISI. The detection and laterality thresholds without a masker or an interferer are also shown by solid lines. Figure 1(a) presents the results obtained in condition where the frequency of a masker or an interferer is 1 kHz and Fig.1(b) presents the data obtained when the frequency is 4 kHz. The thresholds of laterality without an interferer become higher than those of detection without a masker as the frequency of a test tone increases. In Fig. 1 (a) and (b), the thresholds of detection with a masker are...
higher than those without a masker only at the same frequency as the masker. On the other hand, the thresholds of laterality are higher than those without an interferer in a wider frequency range around the frequency of an interfering tone. The same thing can be said of the results obtained in backward masking paradigm, as shown in Fig. 2 (a) and (b), but the net value of masking or interfering is smaller.

From the view point of information processing in an auditory system, the precedence effect can be regarded as forward masking of binaural information. This effect is well explained as follows. In the auditory system, the information on interaural differences is extracted and sent to the more central nervous system. On the way to the central system, the succeeding signals interfere with the information of interaural difference.

CONCLUSION

In this paper, we measured the threshold of laterality of a tone as the minimum level of the tone at which the subject could reliably discriminate the laterality, right or left, and explored how the lateralization of the tone burst with ITD was affected by the preceding or following tone burst using forward and backward masking paradigms. Both the threshold of laterality and the detection threshold were determined as a function of ISI. The threshold of laterality is almost the same as the detection threshold in the low-frequency region (500 Hz), but in high-frequency region (higher than 1 kHz), the threshold of laterality is much higher than the detection threshold. The deterioration of lateralization as a function of ISI, almost parallels that of detection. On the other hand, the thresholds of laterality without an interferer also becomes higher than those of detection without a masker as the frequency of a test tone increases. It can be said, however, that the net interfering effect of a preceding or following sound on lateralizability is spread over a wider frequency range than detection.

Further studies are needed to explore the effect of a preceding tone and the following one on the lateralization of the sound with interaural intensity difference in order to compare the effect of interaural time and intensity difference on laterality threshold.

REFERENCES

Echoes in the Dark

Ramos, O. A.; Arias, C. a,b; Ortiz Skarp, A. H a; Frassoni, C. A.a


bCentro de Investigación de la Facultad de Filosofía y Humanidades. CIFFyH. Facultad de Filosofía y Humanidades. Universidad Nacional de Córdoba. Córdoba. Argentina

Abstract: The main aim of our interdisciplinary long-term project is to study the human obstacle perception ability -i.e., echolocation- in order to develop a special program, using virtual and real obstacles, to safely train the visually handicapped person for the achievement of his/her independent mobility. This paper presents some relevant theoretical issues related to this topic and a brief report of our experimental research work with blind participants.

INTRODUCTION

Most of the blind persons have a severe limitation for travelling independently. This is one of the most serious consequences imposes by blindness.

We are involved in an long-term project on the human auditory obstacle perception ability -i.e. echolocation- and its underlying mechanisms, in order to lay on the theoretical and practical basis of a training program for the assessment and development of the echolocation ability of the visually handicapped person. [1]

Human echolocation

Echolocation is the ability to detect, discriminate and localize obstacles by processing the acoustic information contained in echoes produced by the reflection of self-generated sounds on the surrounding obstacles. [1, 2]

The self-generated sound is called the direct or emitted signal and the echo is called the reflected signal in an echolocation paradigm. Visually handicapped people spontaneously and intuitively generate sounds such as tongue clicks, snaps, hissings or vocalizations when they move around in order to gather spatial information.

Long and short distance echolocation are the two basic modalities of human echolocation based on different underlying psychoacoustic mechanisms. [3, 4]

The short distance modality (≤ 2 to 3 m between subject and object) is the most relevant for every day purposes: it allows not only orientation but also self protection. Two fusion auditory phenomena seem to be involved in this modality: repetition pitch and the precedence effect.

Repetition pitch and short distance echolocation

Repetition pitch occurs when a sound (direct sound) and the repetition of this sound (reflection) are added after a brief delay and presented to a listener [5]. It seems to be involved for obstacles located just in front of the subject face. He/she hears only one sound which is perceived as a pitch shift of the direct signal. If the emitted signal is noise, the fused sound acquires a pitch-like quality. [6]

Short distance echolocation and the precedence effect

The precedence effect occurs when two similar sounds are presented from different locations with a brief delay. The sound that arrives first, called the leader (direct signal), dominates the perceived location of the fused image. However, recent findings suggest that the auditory system maintains information about the lag stimulus (reflection) even when fusion and leader dominance occur. Certain changes in the acoustic environment break down the precedence effect and the subject can extract directional cues of the lagging component. [7]

Lag discrimination suppression -i.e., another aspect of the precedence effect- and the break down phenomena seem to be involved in the short distance echolocation mode for obstacles located outside the median sagittal plane.

EXPERIMENTAL WORK

Four main experiments with blind participants were carried out:

1) Echolocation using real obstacles: a classical study of detection, localization and discrimination of real obstacles was carried out in the anechoic chamber of our Lab. Six good obstacle detector blind persons participated as subjects. The results obtained were in good agreement with that of obtained in previous researches: it was easier to detect the presence or absence of an obstacle than to discriminate its...
characteristics. Besides, the obstacle size was the easiest characteristic and its shape was the most difficult.

2) Auditory functioning in blind and sighted subjects: Peripheral and central auditory functioning of 8 good obstacle blind detectors and 8 sighted subjects were assessed. An echolocation paradigm in short latency evoked response was included. It was found that the blind subjects were better (faster) than sighted subjects in auditory processing and that the echolocation signals may be processed at a lower level in the auditory pathway, possibly in the superior olivary complex of the pons. [8]

3) Psychoacoustic tests: several psychoacoustic tests - constructed and administered with the ECOTEST (one of the Rousettus' module) [9] were especially designed to simulate acoustic conditions assumed to be involved in the short distance echolocation modality. Echolocation stimuli were used in all the tests. They were administered to blind and sighted subjects.

a) Echolocation and repetition pitch
In a first serie, four tests that measure the subject performance in repetition pitch detection and discrimination tasks using an echolocation paradigm as sound stimuli, were administered to 30 sighted subjects with and without musical training and to one blind person. The results suggested that: a) echolocation is a genuine ability that does not require a "privileged ear"; b) the subjects do indeed perceive a repetition pitch when they are stimulated with echolocation stimuli; c) noise signals seemed to give better information than click signals and d) the performance of the blind subject in the matching repetition pitch test was, as we anticipated, the most relevant result. [10]

b) Echolocation and spatial audition
Sound lateralization under precedence effect condition: The performance in a lateralization task under precedence condition of 20 sighted subjects with and without musical training and of 2 blind subject with musical training, was studied. The results displayed good agreement with previous ones and point towards the perceptual training hypothesis: the blind subjects performance was better than sighted ones, especially in the more difficult task, i.e. the precedence condition.

Evolutive aspects of the precedence effect: A replica of Litovsky study [11] is been carrying out at present. The performance of 36 sighted subjects (6 children and 30 adults) are been estimating using the minimum audible angle task (MAA) in a localization test for a single-source condition and for two precedence conditions: lag and leader discriminations and for 3 stimulus conditions (4 ms, 25 ms and a real echolocation signal). The preliminary results showed good agreement with the original study. One blind participant obtained lower MAAs in almost all the more difficult experimental conditions. Both of them obtained lower MAAs in the single-source condition and the leader discrimination one, with real echolocation signal.

CONCLUSIONS AND DISCUSSION
Our results as a whole can be summarize as follow:
1. Echolocation is a genuine human ability that does not require a "privileged ear". It is one of the most important factors for the blind person independent mobility achievement.
2. Two auditory fusion phenomena seemed to be involved in the short distance echolocation modality: repetition pitch and the precedence effect.
3. The blind subjects performed better than the sighted ones in several auditory tasks, especially in the most difficult ones.
4. A good obstacle blind detector might have learnt, not even bewared of it, to perceive obstacles through the refined and sustained auditory training which is daily exposed to.

ACKNOLEDGMENTS
Grateful to: CONICET and CONICOR of Argentine and to all the blind and sighted participants.

REFERENCES

---

1 Rousettus is a PC based system specially developed by us in order to study the human echolocation processes.
Simulating the Franssen Illusion
William M. Hartmann and Scott R. Lawton

Department of Physics and Astronomy, Michigan State University, East Lansing, MI, 48824, USA.

The Franssen illusion is created with sine tones, presented by loudspeakers in a reverberant room. In this illusion, a listener localizes a tone at the position of its onset, even though the onset may have faded away and a continuous steady-state tone is arriving from a different location. The illusion has been attributed to a reweighting of localization cues, caused by the plausibility of direct onset cues in contrast to steady-state cues that are made implausible by standing waves in the room. This explanation was tested by comparing the Franssen illusion in a room to a headphone simulation with reduced complexity. In both configurations, slow-onset tones were used to determine the uncertainty associated with standing waves (room) and conflicting interaural time and intensity cues (headphones). The effect of adding the abrupt onset was compared for conditions of comparable uncertainty for the room and headphone configurations. The comparison showed that headphone listening can lead to a clear analog to the Franssen illusion, but with somewhat reduced strength. The difference might be attributed to the temporal fluctuations that occur in the room but are absent in the headphone simulation.

INTRODUCTION

The Franssen illusion is one of the most dramatic effects in psychoacoustics [1, 2]. It is created with two loudspeakers in a standard stereophonic configuration in a room that must not be too dry. The signal is a sine tone.

To create the illusion, the tone is turned on with an abrupt onset at one of the two speakers (leading). Immediately, the tone begins to fade away while a tone in the other speaker (lagging) grows in such a way as to keep the total power approximately constant. The duration of the transition is not important; transitions from 30 to 2000 ms can work. After the transition, the leading speaker is off and the lagging speaker is on. The illusion is that the listener continues to localize the tone at the position of the leading speaker, even though it is completely silent.

The illusion fails if the room is too dry. It also fails if the stimulus is noise instead of a tone. An explanation for the illusion that is consistent with the facts [3] is that signals are localized on the basis of weighted binaural cues. The weighting is a central operation so that cues can be evaluated for self-consistency. In the case of a sine tone in a room, the standing waves lead to steady-state differences in interaural level (ILD) and interaural time (ITD) that are invariably inconsistent. By contrast, a transient event, such as an abrupt onset, can be evaluated by the binaural system from its direct transmission path without interference from room reflections. The result, with respect to the Franssen illusion, is that the steady-state tone from the lagging speaker is discounted in the weighting process because of its implausible binaural cues. The onset receives high weight because its ILD and ITD cues are mutually consistent, and they are also consistent with visual cues. If this explanation is correct then it ought to be possible to create an analog to the Franssen effect using headphones by putting an implausible steady-state (or lagging) sound in competition with an onset (or leading) transient having plausible binaural cues.

EXPERIMENTS

The goal of the experiments was to compare a headphone analogy to the Franssen effect with a real Franssen effect as it occurs in a room. Experiments were done with sine tones, with frequencies of 250, 500, and 4000 Hz. There were two types of trial: abrupt onset (instantaneous) and slow onset (1-s ramp). The abrupt-onset trials were intended to create the Franssen effect. The slow-onset trials served to establish a baseline lateralization/localization for the steady-state tone, the natural competition for the onset. Immediately after the onset, the transition to the steady-state occurred, a 30-ms linear sweep. Steady-state tones were 4 s in duration to give listeners a good opportunity to evaluate the position. Tones were turned off together with a noise burst from the forward direction to mask the offset. After the noise burst the listener was required to indicate the location of the steady-state tone, left or right. The above conditions apply to both the room experiment and the headphone experiment.

Room Experiment

The room experiment was done in a large room 8.5 by 7.3 by 4.6 meters high, with concrete walls and ceiling and a vinyl floor. In this room left (L) and right (R) loudspeaker positions and the listener's square formed an equilateral triangle 5 meters on a side. The listener
changed locations within the 1-meter square after each trial to randomize the standing wave pattern. Leading-lagging tones were L-L, L-R, R-L, and R-R, equally in a balanced design. Trials of type L-R and R-L tested the Franssen effect.

**Headphone Experiment**

The headphone experiment was the simplest possible analog to the Franssen effect. The onset was lateralized by a finite ITD but zero ILD. The transition (30 ms) retained the ITD but introduced an ILD for the steady state, either agreeing (L-L or R-R) or disagreeing (L-R or R-L) with the onset.

**RESULTS AND CONCLUSIONS**

Because the headphone conditions were manufactured whereas the room conditions were uncontrolled, it is necessary to compare the two experiments on an equivalent basis. Headphone and room conditions were called equivalent for comparable responses to slow-onset tones.

**Ambiguity-conditional analysis**

Many headphone conditions failed to produce ambiguous lateralization - listeners responded consistently. In the ambiguity conditional analysis, data from abrupt-onset trials of the form L-R or R-L were accepted only if the corresponding conditions for slow-onset trials produced between 20 and 80 percent correct responses. The percentage of responses following the steady-state ILD for headphones or the percentage of correct responses in the room is shown in Fig. 1. The smaller the percentage, the larger the Franssen effect. The comparison in Fig. 1 indicates similar effects in the room and with headphones. Both tend to disappear at higher frequency.

**Unambiguity-conditional analysis**

The unambiguity-conditional analysis in Fig. 2 is for headphones only. It shows that as the steady-state ILD cue becomes more effective (70 to 100 percent following ILD for slow onsets) the abrupt onset results vary from 20 to 65%. This variation, only 45%, is rather small compared to the strength of the Franssen effect. We expected that for 100% correct given slow onsets listeners would score much better than 65% given abrupt. In the end, the two analyses suggest that an analog to the Franssen effect can be produced with headphones but that it is weak compared to the Franssen effect in a room.

**REFERENCES**


**FIGURE 1.** Ambiguity-conditional analysis, percent vs sine tone frequency. Filled symbols are for headphones open symbols are for the room. Triangles for listener S; squares for T; circles for W; star for N, and diamond for X.

**FIGURE 2.** Unambiguity-conditional analysis. For four listeners in the headphone experiment the plot shows the percentage of lateralization judgements that followed the ILD for abrupt onsets given that the percentage for slow onsets (with the same ILD and ITD) was 70-100 percent, as given on the horizontal axis. The heavy curve called MEAN is the average of all data points.
Observation of neural activities in pure-tone search and band-noise search

N. Asemi\textsuperscript{a}, Y. Sugita\textsuperscript{b} and Y. Suzuki\textsuperscript{a}

\textsuperscript{a}Research Institute of Electrical Communication/ Graduate School of Information Sciences
Tohoku University, Katahira 2–1–1, Aoba-ku, Sendai, Japan
\textsuperscript{b}The National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan/ PREST, JST

In our previous study, auditory search asymmetry between pure tones and narrow band noises was clearly seen, indicating that there might be some pre-attentive processes for quick detection of a narrow band noise. For investigating the mechanisms for quick detection, in this paper, neural activities in auditory search tasks were observed.

INTRODUCTION

Some sounds with specific features may be more easily detected than sounds with other features. Clarification of such sound features and their underlying mechanisms for detection would help clarify the properties of the auditory system. We have found clear auditory search asymmetry\cite{1,2}. For example, a narrow-band noise among pure tones was easily detected, while it took significantly longer to detect a pure tone among narrow-band noises (Fig. 1). This result suggests that time fluctuation may be a feature which facilitates detection.

In this study, neural activities for pure-tone search and band-noise search were measured by using functional MRI. Differences of activities in various areas were examined, and the mechanisms for the quick detection of a sound with time fluctuation are herein discussed.

METHODS

Three male subjects with normal hearing and ranging in age from 20 to 45 years participated in this study.

Pure tones and narrow-band noises (1/4 oct.) were used as stimuli. The subjects were presented with two kinds of tasks: (1) searching a pure tone among narrow band noises (PT search) and (2) searching a narrow band noise among pure tones (BN search). The number of stimuli was 2 or 3. The stimuli were mixed down to a single channel and monophonically presented to the subject’s left ear through an air-conductive earphone. The duration of stimuli was 2.5 s. The search tasks were conducted under the four conditions shown in Table 1 (A–D). For each condition, a subject made 30 judgments as to whether the target stimulus was present or absent.

Images were acquired using a 3T GE Signa scanner. Echoplanar images (64 × 64, 5 mm × 5 mm pixels) comprised of 15 axial slices were taken every 7 mm. All volumes were realigned to the first volume and smoothed. Data were analyzed using Statistical Parametric Mapping (SPM99)\cite{3}. According to a cross-correlation method, areas showing significant differences between two conditions were extracted.

<table>
<thead>
<tr>
<th>Table 1. Experimental conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>condition</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
</tbody>
</table>
RESULTS AND DISCUSSION

Four kinds of comparisons were possible between the experimental conditions. They are summarized in Fig. 1 and are labeled as Comparison #1 to #4, hereafter. Fig. 2 and 3 show voxels with significant differences in Comparison #1 and #2 and Comparison #3 and #4, respectively.

Obviously, there were many voxels with significant differences in Comparison #3 (top of Fig. 3). Referring to Fig. 1, this comparison is between two and three stimuli in the PT search. An increase in the number of stimuli in the PT search resulted in the activations in the primary auditory cortex and the prefrontal association cortex. This result seems to be consistent with the increment of search time with the number of stimuli in the PT search. In the visual search study, the prefrontal association cortex was activated by active attention [4]. The activations of the prefrontal association cortex were consistent with those in visual attention [4] and may be due to the active attention required in the PT search.

In Comparison #1, only a few voxels show significant differences. This corresponds to the slight difference in the search times between the PT and BN searches when the number of stimuli was two (Fig. 1). This would mean that the level of difficulty of the tasks in these conditions were similar. Only a few voxels show significant differences in Comparison #2 as well. The search times, however, are quite different between the PT and BN searches when the number of stimuli was three (Fig. 1). The authors have no clear explanation for this “discrepancy” at this time.

In Comparison #4 (bottom of Fig. 3), there were a few voxels showing significant differences in the primary auditory cortex. This result may reflect the neural responses of the primary auditory cortex; the increase in the number of the stimuli may have caused some change in the auditory characteristics of the input signal.

CONCLUSIONS

Neural activities for auditory search tasks were measured by fMRI. A clear difference of activations was observed in the PT search when the number of stimuli was increased, which is consistent with the increment of search time. Comparing the present results with results of neurophysiological studies in the visual search, it is suggested that the prefrontal association cortex is activated in the PT search process which requires active attentions.

REFERENCES

3. http://www.fil.ion.ucl.ac.uk/spm/
Psychoacoustic product sound quality evaluation
L. Kynčl, O. Jiříček,

Department of Physics, CTU in Prague, 166 27 Technická 2, Czech Republic

In our research, six vacuum cleaners with similar parameters were recorded in the setting of a typical office. The data was acquired by means of an artificial head with microphones located at the beginning of the outer ear canal of its artificial ears. The first test consisted of description of sounds in the subjects' own words. These words were then used as basic words in the next test. In the second experiment, the subject obtained a five-points response scale to evaluate each of 33 pairs of opposite descriptive adjectives with respect to appropriateness of the pair for description of a vacuum cleaner. Grade 1 at the scale corresponds to an absolutely suitable and grade 5 corresponds to an absolutely unsuitable pair. In the third test, the average intensity of perception of auditory attributes was tested. Five psychoacoustic attributes from the previous experiment were presented to subject: rippleness, untypicality, inefficiency, loudness, pleasantness. The subject evaluated the intensity of perception of these attributes on a 10-point scale for each vacuum cleaner sound. Some of the attributes were excluded from tests using reliability tests (distribution of values, t-test for mean values).

From the mid-Eighties, designers of new products emitting sound energy have been interested not only in better design, lower price and higher power but also in the aesthetic pleasurable quality of the sound produced. This interest has created a new term: product-sound quality.

Design work in this field makes use of, besides the usual sound pressure and sound power analysis, sound quality studies based on subjective listening tests resulting in a set of auditory attributes (e.g. loudness, sharpness, pitch) separating several characteristics of auditory event. In the next step, a quantitative description of psychoacoustic quantities is performed. The sound of the product is then assessed through large-scale testing of listeners, originating from psychological methods, as every subject has different perceptual conceptions and experiences with auditory attributes. The results of the tests are then statistically analyzed.

In our research, six vacuum cleaners with similar parameters were recorded in the setting of a typical office. All vacuum cleaners were recorded at full power operation. The data was acquired by means of an artificial head with microphones located at the beginning of the outer ear canal of its artificial ears. Recording was performed using Sony DAT (sampling frequency 44.1 kHz, 16 bit signal quantizing) and for the presentation (play back) to the subject headphones Sennheiser HD250 was used.

The first test consisted in descriptions of the sounds in the subjects' own words. 35 listeners took part in this test. They had to fulfill the following conditions: age from 18 to 50, non-damaged hearing, and personal experience with the regular use of a vacuum cleaner in the household. These conditions were used for all sequential experiments. Every person heard 6 sounds one by one, with duration of one minute each, and described these sounds in their own words. The most frequent words (e.g. pleasant-annoying-unpleasant "příjemný-rušivý-nepříjemný", loud-silent "hlasitý-tichý", efficient-inefficient "výkonný-nevýkonný", sharp "ostrý") were then used as basic words in the following test. A nonnegligible percent age of words compared the evaluated sound to other household appliances, meaning that atypicality "netypnost" is one of the important attributes of vacuum cleaner sound. These list of words were classified into four groups, according to their connotative meaning. Then the opposite descriptive adjective was subjoined to every word's item.

In the second experiment, the subject obtained a five-point response scale to evaluate each of 33 pairs of opposite descriptive adjectives (e.g. pleasant/nonpleasant, loud/quiet) with respect to appropriateness of the pair for description of a vacuum cleaner. Grade “1” in the scale corresponds to an absolutely suitable and grade “5” to an absolutely unsuitable pair. This experiment results in 33 values for each pair; mean value and dispersion were calculated. Pairs with high dispersion were excluded for nondefinite evaluation, and the first five pairs with the lowest mean value were used for the third test.

In the third test, the average intensity of perception of auditory attributes (e.g. pleasant/unpleasant) was tested. Five psychoacoustic attributes (four from each connotative group and pleasantness as a attribute for the global evaluation) from the previous experiment were presented to subject: fuzziness, atypicality, inefficiency, loudness, pleasantness. The subject evaluated the intensity of perception of these attributes on a 10-point scale for each vacuum cleaner sound. Some of the attributes were excluded from tests using reliability tests (distribution of values, t-test for mean values).

Using correlation analysis, we determined the quantity of influence of four attributes on the global attribute – pleasantness - of the sound (see Table 1 and Figure 1), which represents the sound quality of each
vacuum cleaner. Shown in the figure is a high proportion of attribute “inefficiency” in the global evaluation of pleasantness. Most of the correlation coefficients are negative, meaning that the evaluation of these attributes influenced “unpleasantness” instead of “pleasantness”.

Table 1. Descending sequence of attributes by magnitude of coefficient of correlation ρ, (z)-motor placed behind recording head, (l)-motor placed to the left recording head

<table>
<thead>
<tr>
<th>Vac. clean. type</th>
<th>attribute</th>
<th>ρ</th>
<th>Vac. clean. type</th>
<th>attribute</th>
<th>ρ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ETA Ecoline II (z)</td>
<td>Inefficiency</td>
<td>-0,006</td>
<td>ETA Maestro (z)</td>
<td>Inefficiency</td>
<td>-0,298</td>
</tr>
<tr>
<td></td>
<td>Fuzziness</td>
<td>-0,287</td>
<td></td>
<td>Fuzziness</td>
<td>-0,368</td>
</tr>
<tr>
<td></td>
<td>Atypicallity</td>
<td>-0,535</td>
<td></td>
<td>Atypicallity</td>
<td>-0,453</td>
</tr>
<tr>
<td></td>
<td>Loudness</td>
<td>-0,674</td>
<td></td>
<td>Loudness</td>
<td>-0,555</td>
</tr>
<tr>
<td>ETA Ecoline II (l)</td>
<td>Inefficiency</td>
<td>0,002</td>
<td>ETA Maestro (l)</td>
<td>Inefficiency</td>
<td>0,017</td>
</tr>
<tr>
<td></td>
<td>Fuzziness</td>
<td>0,141</td>
<td></td>
<td>Fuzziness</td>
<td>0,17</td>
</tr>
<tr>
<td></td>
<td>Atypicallity</td>
<td>-0,207</td>
<td></td>
<td>Atypicallity</td>
<td>-0,031</td>
</tr>
<tr>
<td></td>
<td>Loudness</td>
<td>-0,488</td>
<td></td>
<td>Loudness</td>
<td>-0,34</td>
</tr>
<tr>
<td>ETA Quinto (z)</td>
<td>Inefficiency</td>
<td>0,21</td>
<td>ZELMER (z)</td>
<td>Inefficiency</td>
<td>0,026</td>
</tr>
<tr>
<td></td>
<td>Fuzziness</td>
<td>-0,378</td>
<td></td>
<td>Fuzziness</td>
<td>-0,17</td>
</tr>
<tr>
<td></td>
<td>Atypicallity</td>
<td>-0,789</td>
<td></td>
<td>Atypicallity</td>
<td>-0,26</td>
</tr>
<tr>
<td></td>
<td>Loudness</td>
<td>0,002</td>
<td></td>
<td>Loudness</td>
<td>0,467</td>
</tr>
<tr>
<td>ETA Quinto (l)</td>
<td>Inefficiency</td>
<td>0,214</td>
<td>ZELMER (l)</td>
<td>Inefficiency</td>
<td>-0,178</td>
</tr>
<tr>
<td></td>
<td>Fuzziness</td>
<td>0,219</td>
<td></td>
<td>Fuzziness</td>
<td>-0,237</td>
</tr>
<tr>
<td></td>
<td>Atypicallity</td>
<td>-0,24</td>
<td></td>
<td>Atypicallity</td>
<td>-0,452</td>
</tr>
<tr>
<td></td>
<td>Loudness</td>
<td>-0,459</td>
<td></td>
<td>Loudness</td>
<td>-0,618</td>
</tr>
<tr>
<td>ETA Astro (z)</td>
<td>Inefficiency</td>
<td>-0,065</td>
<td>ETA Ecoline [I] (z)</td>
<td>Atypicallity</td>
<td>-0,13</td>
</tr>
<tr>
<td></td>
<td>Fuzziness</td>
<td>-0,238</td>
<td></td>
<td>Fuzziness</td>
<td>-0,146</td>
</tr>
<tr>
<td></td>
<td>Atypicallity</td>
<td>-0,335</td>
<td></td>
<td>Atypicallity</td>
<td>-0,34</td>
</tr>
<tr>
<td></td>
<td>Loudness</td>
<td>-0,347</td>
<td></td>
<td>Loudness</td>
<td>-0,502</td>
</tr>
<tr>
<td>ETA Astro (l)</td>
<td>Inefficiency</td>
<td>0,465</td>
<td>ETA Ecoline [I] (l)</td>
<td>Inefficiency</td>
<td>0,253</td>
</tr>
<tr>
<td></td>
<td>Fuzziness</td>
<td>-0,078</td>
<td></td>
<td>Fuzziness</td>
<td>-0,268</td>
</tr>
<tr>
<td></td>
<td>Atypicallity</td>
<td>-0,404</td>
<td></td>
<td>Atypicallity</td>
<td>-0,304</td>
</tr>
<tr>
<td></td>
<td>Loudness</td>
<td>-0,583</td>
<td></td>
<td>Loudness</td>
<td>-0,598</td>
</tr>
</tbody>
</table>

REFERENCES


ACKNOWLEDGMENTS

This research work has been partially supported for the research program No: I04/98/212300016 and Research Center of Combustion Engines and Automobiles No LN00B073.
Parameters Influencing Perception of Highway Traffic Noise

C. Nathanaila and F. Guyotb

aIMPEDANCE S.A, 20 rue Jean Rostand, 91893 Orsay Cedex, France
bLAPS, 54 rue René Boulanger, 75010, Paris, France

Recent research in traffic noise perception points out the complex character of the evaluation process. Quantitative and qualitative parameters of sound, but also visual and cognitive factors seem to play an important and often combined role. The present research aims to investigate these questions by a series of perceptive audiovisual tests. Sound stimuli are binaural recordings of highway noise; variations are introduced by varying the nature of the field between the recording position and the highway (free field, tree or noise barrier) and the percentage of heavy vehicles in the total flow. Visual influence is studied by coupling video sequences representing the three corresponding visual environments, to the same sound stimuli. Cognitive factors (expectations, etc.) are investigated by comparing test results of an exposed and a non-exposed population. The method of semantic differential scales is used; the proposed descriptors are based on preliminary testing using open interviews. A factorial analysis shows that most of the variance of the responses is related to one single dimension mainly due to the set of descriptors used and to non-negligible differences in level. Analysis of variance reveals a complex visual influence depending on the sound level and character. No significant differences are observed between the two participating populations.

INTRODUCTION

Highway traffic noise is an important environmental factor closely related to the quality of life of increasingly larger samples of the population. Studies on subjective traffic noise assessment involve either field investigations either laboratory experiments. Field studies respect the environmental context but fail to control the experimental material. Laboratory experiments are experimentally valid, since enabling the control and analysis of the presented sounds, but do not take into account number of environmental factors susceptible to influence sound perception [1,2]. Likewise, while some experimental results show a rather good correlation with standardised metrics, other results, and mainly those issued from field studies, seem to point out its limits [2,3]. Several questions seem to arise, concerning the qualitative aspect of sound on one hand (number and type of sound sources) and contextual or cognitive factors on the other (visual influence, expectations or interest of the subjects).

The present work aims to investigate the influence of some of these parameters on highway noise perception and study their interactions. For this reason, controlled subjective experiments involving visual simulation were organised. In order to study parameters related to experience they were partly realised in Reims (subjects leaving close to the highway A4) and partly in Paris, that is with the participation of an exposed and of a non exposed population respectively.

EXPERIMENTS

Three experiments were realised. The first uses an ‘open interview’ method and aims primarily to select a set of descriptive adjectives of highway noise. The other two use the semantic differential scaling method in order to investigate the perceptive dimensions of highway noise and test the effect of various experimental factors susceptible to influence them (these two experiments concern actually the same experiment realised once in Paris and once in Reims, and are considered as one in the following).

Sound and visual stimuli are common to all experiments. Sound stimuli are six binaural recordings of highway noise realised in 3 different sites at a distance of approximately 100m from it. The sites are an open field close to the highway, a site behind a thick vegetation and a site behind a noise barrier. All stimuli were selected to represent the same number of total flow but with two different percentages of heavy vehicles for each site (less than 10% and 30% or more). Stimuli duration is 20 s. Levels vary from 53 to 62.5 dB(A). Five visual sequences were used, 2 represent open views of the highway (corresponding to the first two sound stimuli), two represent views of thick vegetation and one shows a view of the noise barrier. These sequences were coupled to create fourteen audiovisual stimuli presented over headphones for the sound and LCD screen glasses for the video.
After a short training session, the six sound stimuli are presented to the subjects without visual stimulation (unimodal session). Follows the presentation of the fourteen audiovisual stimuli (bimodal session). Subjects were asked, depending on the experiment, to comment or judge the sound stimuli presented to them, they were told that the video images were simply present to help them imagine that they are in the actual place, but that they were not meant to be judged.

‘open interview’ experiment

Eighteen subjects from Paris participated in this experiment. An ‘open interview’ method was used: subjects were asked to comment freely on the sound environment presented on the audio or audiovisual stimuli. The main outcome of this experiment was a set of descriptive adjectives of highway noise selected on the basis of the total number of occurrences. These words appeared at least 18 times in the whole test and are ‘fort’, ‘lointain’, ‘continu’, ‘gênant’, ‘supportable’, ‘vivable’ (loud, distant, continuous, annoying, pleasant, tolerable and liveable-with). Obviously, certain meanings are similar, but our choice was to retain the words used by the subjects because they made sense to them. A simple verbal analysis realised on these results is useful in the validation and/or the interpretation of the results of the semantic differential experiments.

semantic differential experiments

Eighteen subjects from Paris and eighteen subjects from Reims participated in these experiments. They were asked to judge the sound environment of the 20 stimuli presented to them on the basis of 8 differential scales. Seven, representing the descriptive adjectives selected in the previous experiment and one, the adjective ‘naturel’ (natural) added in order to denote the presence of sources of the nature in the stimuli (mainly birds).

PCA analysis in both the results of the unimodal and the bimodal parts of the two experiments shows that in all four cases approximately 90% of the total variance is explained by one factorial dimension strongly correlated to ‘fort’ (loud), ‘gênant’ (annoying) and inversely correlated to ‘éloigné’ (distant), ‘agréable’ (pleasant), etc. Differences in level account certainly for that; also the set of the adjectives used by the subjects in the previous experiment.

An ANOVA realised with ‘site’ (Paris or Reims) as between subjects factor and ‘sound stimulus’ and ‘visual sequence’ as within subjects factors. ‘Site’ was in none of the cases significant, indicating that there are no main differences between the two populations. ‘Sound stimulus’ was in most of the cases significant indicating that the adjectives used had a meaning for the subjects, who claimed indeed that they had no difficulty to judge them. No significant differences were observed due to the different percentages of heavy vehicles in the total flow. Finally either ‘Visual sequence’ either its interaction with sound stimulus had in most of the cases a significant effect indicating clear but complex visual influence. Sounds considered really loud or annoying do not seem to benefit from pleasant visual settings – they can even show an opposite tendency – while less loud sounds seem to benefit. Figure 1 shows this influence for ‘fort’ (loud). Similar patterns of influence are observed for ‘gênant’ (annoying), ‘éloigné’ (distant), etc.

FIGURE 1. Visual influence on the subjective ‘fort’ (loud). Planned comparisons for the two sound groups (A1, A2 and A3-A6) show significant effect of vision (F(2,68)=7.3, p<.01 & F(1,34)=10, p<.01 respectively).

In conclusion, the two experiments seem to reveal that the highway – at least at the distance at which it was recorded – is represented by one single dimension and considered rather as one single source. Small, complex visual influence is observed on most descriptors and even on the robust perceptive dimension of loudness.

REFERENCES

Sounds transmitted through a steel plate: perceptual judgments analysis

J. Faure and C. Marquis-Favre


For reducing disagreement caused by noise, acoustic engineers have decreased the overall sound pressure level of sources. But disagreement can also be diminished by modifying sounds characteristics; this is the aim of sound quality studies. If sound quality is now a current practice in several industries, qualitative studies of building structures are scarce. This paper presents the study of the acoustical behavior of a simple component of an exterior wall, a steel plate, from both physical and perceptual points of view. Simulated sounds have been submitted to a panel of subjects asking them to rate the dissimilarity and the preference between sounds using the paired-comparison method. The influence of the plate structural parameters variations on perception for sounds transmitted through the structure is investigated. The results of the perceptual test are confronted to the ones of the physical analysis of the plate acoustical behavior.

INTRODUCTION

This work aims to link a physical analysis and a perceptual test in the case of sound transmitted through a steel plate. The purpose of the present paper is to illustrate these combined analyses when a design parameter, the thickness, is varied. Presentation of perceived distance between sounds and preference trends will be done.

SOUND PERCEPTION TEST

Four seconds length sounds were synthesized from the acoustical response spectra of the plate in two points corresponding to ears position of a human. The sound pressure level spectra were analytically calculated for a simply supported plate embedded in a rigid baffle, radiating in a semi-infinite field and excited by a normal incidence plane wave with a white noise spectrum. Except for thickness, all the structural parameters are kept constant. The considered values of thickness are: $0.5 \times 10^{-3} \text{m}, 1.1 \times 10^{-3} \text{m}, 1.5 \times 10^{-3} \text{m}, 2.1 \times 10^{-3} \text{m}, 2.5 \times 10^{-3} \text{m}, 3.1 \times 10^{-3} \text{m}, 5 \times 10^{-3} \text{m}, 7.1 \times 10^{-3} \text{m}$ and $10.1 \times 10^{-3} \text{m}$.

The perceptual test, based on the method of paired comparisons, was performed in a semi-anechoic room. After 3 training representative pairs, a judgment of dissimilarity between sounds was given by 20 subjects on a 7-point scale with verbal labels at the end points. Their preference was gathered by a dichotomous procedure.

ANALYSES AND RESULTS

To characterize perceptual dimensions, multidimensional scaling analysis has been performed using the ordinal information of the dissimilarity scale. Judgments according preference have been treated to obtain preference on an interval scale [1]. To explain perceptual dimensions, physical and psychoacoustical metrics have been computed using Sound Quality ©2000 MTS Systems Corporation.

In a previous test [2], with a smaller range of thickness from $1.1 \times 10^{-3} \text{m}$ to $3.1 \times 10^{-3} \text{m}$ and inducing loudness variation from 80 to 123 sones, the perceptual dimension was correlated with a linear combination of loudness and Natural Frequencies Deviation metrics, denoted by N.F.D. [3]. This test aims at studying a larger variation of thickness inducing loudness variation from 46 to 145 sones. Two series of stimuli have been submitted to subjects:

1. one with sounds without any modification, in order to investigate if results are similar to the ones obtained for a smaller variation of thickness [2].

2. one with equal-loudness sounds, in order to be free of the predominant effect of loudness on other aspects of sounds [4].

For the first series of nine sounds, the analysis shows one perceptual dimension. This one is highly correlated with the Zwicker loudness (R=0.991 and p<0.001), see fig.1. Preference trend is also explained with the Zwicker loudness (R=0.941 and p<0.001), see fig.2.

For the second series, with the nine equal-loudness sounds, it also appears one perceptual dimension which is correlated with the logarithm of the N.F.D. metrics (R=-0.984 and p<0.001), see fig.3. Considering the variation of thickness between $1.1 \times 10^{-3} \text{m}$ and $3.1 \times 10^{-3} \text{m}$, perceived distances and N.F.D. metrics are linearly correlated as in the previous test [2]. Fig.4 displays preference versus thickness values, preference is also correlated with the logarithm of N.F.D. (R=0.92 and p<0.001).
CONCLUSION

For loudness variation from 80 to 123 sones corresponding to an increase of thickness from $1 \times 10^{-3} \text{m}$ to $3 \times 10^{-3} \text{m}$, both shift in frequency, represented by the N.F.D. metrics, and sound reduction, represented by loudness, act on preference and dissimilarity judgments [2]. For a larger variation of thickness, from $0.5 \times 10^{-3} \text{m}$ to $10 \times 10^{-3} \text{m}$, loudness variation from 45 to 145 sones is too important and hides other aspects of sounds. For such considered variation, both preference and perceived distances are correlated with loudness. For equal-loudness sounds, perceived distances and preference are correlated with the logarithm of the N.F.D. metrics.

The physical analysis advises higher thickness according to the transmitted sound level reduction. Due to the corresponding loudness reduction, the preference trend, obtained after the perception test, gives the same advice. In addition, the qualitative analysis goes further by explaining that subjects prefer also higher thickness because of the variation of sounds frequency content. Increase of plate thickness has thus a dual benefit effect on preference.

REFERENCES

Loudness of Impulsive Environmental Sounds

S. Meunier\textsuperscript{a}, I. Boullet\textsuperscript{b}, G. Rabau\textsuperscript{a}

\textsuperscript{a} LMA-CNRS 13420 Marseille cedex 20, France
\textsuperscript{b} GENESIS, CEEI Provence, Domaine du Petit Arbois BP 88 13545 Aix-en-Provence cedex 04, France

An experiment on the loudness of impulsive environmental sounds has been run. Loudness was measured using magnitude estimation. A forward regression showed that energy, and fall time are correlated with loudness. The study can help to design a model of loudness of impulsive sounds.

1- INTRODUCTION

Many studies on loudness have been done for steady sounds. But most environmental sounds are dynamic. Impulsive sounds are a special case of dynamic sounds which are characterized by a very short duration. Impulsive sounds are usually very annoying. It is well known that annoyance is proportional to loudness \cite{1} et \cite{2}. Thus, the determination of the loudness of impulsive noises will help to develop a model of annoyance of these sounds.

Studies have been done on loudness of impulsive noises, but most of them used synthesized sound as stimulus. These studies showed that as the duration of a sound increases up to a critical duration, loudness increases while the amplitude of the sound is kept constant (see [3] for a review). Scharf [3] summarized different studies (including a round robin test [4]) and showed how loudness varies as a function of duration in each study. Three different relationships between loudness and duration were found in these studies : a) loudness stays constant when sound energy (which is the product of time and intensity) is maintained constant, b) loudness stays constant when energy decreases as duration increases, c) loudness stays constant when energy increases with duration. The different results are highly discordant, but it is usually admitted that loudness depends on sound energy up to approximately a critical duration of 150 to 300 ms (the dependance of the loudness of a brief sound on duration can be described by an exponential function with a time constant of 80 ms).

The loudness of an impulsive noise decreases when the rise time of the signal increases for durations greater than (or equal to) 0.3 ms \cite{5} or 1.5 ms \cite{6}, depending on the study, while the amplitude of the sound is kept constant.

The aim of our work was to extend to environmental sounds the results obtained with synthesized sounds. The first part was to define the physics of impulsive environmental sounds, especially the temporal pattern. Secondly, loudness was evaluated by subjects and we tried to find a correlation with several physical parameters. A model of loudness including both energy and fall time of the impulsive signal is proposed.

2- EXPERIMENT

Twenty-four impulsive environmental sounds (blow of a hammer on different surfaces, noise of the uncorking of a bottle, hands clapping, broken glass, musical percussions) have been recorded in an anechoic chamber using a DAT recorder (Tascam), with a B\&K microphone (omnidirectional, 4190) and a preamplifier B\&K (Nexus). Twelve synthesized white noises and four pure tones, with different temporal pattern (similar to the temporal patterns of the environmental sounds) have been added to the environmental sounds.

![Model of the temporal pattern of the impulsive environmental sounds.](image_url)

The temporal pattern of each environmental sound have been analysed and can be modelling as shown in fig. 1. We have observed that for our sample of sounds, the pattern is always the same : a short attack (modelized by an exponential function), following by a longer decrease (modelized by a decreasing exponential function), without steady part at the maximum amplitude. It is interesting to note that in most of the studies concerning loudness of impulsive
noise, the temporal pattern of the synthesized sounds included a steady part at the maximum. This observation points out the importance to determine whether the results found for synthetic sounds hold for environmental noises.

The loudness has been evaluated using magnitude estimation without reference. Fifteen listeners took part in the experiment.

3- RESULTS

Different parameters have been calculated to be compared with loudness: peak level (in dB), energy (in dB), fall time (logarithm), rise time (logarithm). Fall and rise times are defined as the time between 10% and 90% of the maximum of the signal envelope. Table 1 shows the correlation (linear regression) between all parameters plus the loudness. Peak level, energy and fall time are correlated with loudness. But some of them are also correlated with each other (energy and peak level; energy and fall time). To eliminate the influence of these correlations on the correlations between loudness and physical parameters, a forward regression was run (table 2). Then, only the energy and the fall time was found to be correlated with loudness. Seventy percent of the variance of the loudness can be explained by a linear relationship to energy (table 2, $R^2$ mod=0.698). When fall time is added to the model, the prediction of the variance increases by 5% (table 2, $R^2$ mod=0.753).

A multiple linear regression between the logarithm of loudness, the energy (in dB) and the logarithm of the fall time of the signal gives the following model for the loudness of impulsive environmental sounds:

$$S = K.E^{0.25}.T_d^{0.16}$$

No influence of rise time on loudness was observed. It should be noted that the rise time of the sounds used in our study was rarely longer than 1.5 ms, the critical value found by Gustaffson [6] below which the rise time does not influence loudness.

4- CONCLUSION

We have proposed a model for loudness of impulsive environmental sounds based on the energy and on the fall time of the signal. Our model suggests that, for the kind of sounds used in this study, loudness stays constant when energy decreases as duration increases. So, the second relationship, given by Scharf [3] (see introduction), between loudness and duration is observed. The rise time in environmental noises is too short to have any influence on the loudness. Eventually, it is important to note that the sounds recorded had no steady maximum amplitude, opposed to synthesized sounds used in most studies. Thus, the results of these studies cannot always be extended to environmental sounds.

REFERENCES

Comparison of Hearing Thresholds by Individual Equalization of the Audiometric Equipment

B. L. Karlsen, M. Lydolf and A. O. Santillán

Department of Acoustics, Aalborg University, Frederik Bajers Vej 7 B4, DK-9220 Aalborg Ø, Denmark, Email: blyk@acoustics.au.dk

Currently, hearing thresholds are measured with different earphones in the hearing clinics. Earlier experiments have shown that this can result in different thresholds. In this paper, five different audiometric earphones are compared by using threshold measurements in the frequency range 100 Hz - 8 kHz. The chosen psychophysical method was the ascending method nearly implemented as described in ISO 8253-1 [4]. Twenty seven young otologically normal test subjects participated in the experiment. The resulting data are meant for validation of the reference threshold for regular pure-tone audiometry given in ISO 389-1 [3]. All of the five earphones used in the experiment were originally calibrated on a coupler conforming to IEC 60318-1 [1]. The coupler response for each earphone was also recorded. Furthermore, the phone-to-closed-ear-canal-transfer-function (PTF) of all earphones on all subjects was measured. In order to equalize for the difference between the earphones, the mean PTF across subjects is added and the coupler response is subtracted from the mean absolute threshold data. After this equalization has taken place, the resulting absolute threshold data obtained with the different earphones are much more similar than they were before the equalization.

INTRODUCTION

This document describes a set of data which are a contribution to the ongoing work on ISO 389 parts 1 and 8 in ISO/TC 43/WG 1. A selection of five different audiometric earphones is compared by using threshold measurements in the frequency range 100 Hz - 8 kHz. The earphones were Sennheiser HDA200, Telephonics TDH39, Telephonics TDH39 with noise capsules and cushions (TDH39C), Holmberg 95-01 and Beyerdynamic DT48. All of the earphones were supra-aural in nature except for Sennheiser HDA200 which was circum-aural. The data presented here are meant for validation of the reference threshold for supra-aural pure-tone audiometry given in ISO 389-1 and for setting up a new reference threshold for circum-aural pure-tone audiometry in ISO 389-8. Apart from measuring absolute thresholds the coupling between the earphones and the ears of the subjects was also determined by measuring phone-to-closed-ear-canal-transfer-functions (PTFs) of all earphones on all subjects. This initiative was undertaken in order to equalize the audiometric equipment to a uniform ear exposure of all the subjects and not to a sound pressure in a coupler. The experiment reported here was almost performed according to the "Preferred Test Conditions for Determining Hearing Thresholds for Standardization" as issued by ISO/TC 43/WG 1 [5].

METHOD

The absolute thresholds measured in this experiment were obtained using a direct detection paradigm and the ascending method of presentation nearly as described in ISO 8253-1 [4].

There were 27 subjects who participated in the experiment. The minimum and maximum age of the subjects were 18 and 25 years respectively. The mean age was 22.2 years and the median age was 23 years. There were 13 males and 14 females who participated. All the subjects were found to be otologically normal.

All the stimuli were presented only to the left ear of each of the subjects. The stimuli consisted of pure tones of 1 s duration at the nominal frequencies corresponding to the 1/3 octave band center frequencies from 100 Hz to 8 kHz as well as the audiometric frequencies 0.75, 1.5, 3.0 and 6.0 kHz in accordance with IEC 60645-1 [2].

The data in these experiments were collected on three separate days for each subject. The days were in no case consecutive. On the first two days, the subjects listened to the tones presented from each of the five earphones in succession with increasing frequency range (lower-mid-higher). On the last day a 1 kHz anchorpoint was established by four repeated measurements of threshold for each earphone and for each subject. The PTFs of each earphone on each subject were also measured on this last day.

The transfer functions between each earphone and the IEC 60318-1 coupler (Brüel & Kjær type 4153) were measured using a Brüel & Kjær 4134 microphone and a PC-based MLSSA measuring system. The same system was used for measuring PTFs, but in this case the measuring microphone was a Sennheiser KE4-211-2 miniature microphone with a custom made battery powered microphone amplifier. The miniature microphone was mounted in a standard earmold and the earmold was inserted into...
the subjects left ear canal so that it was flush with the end of the canal.

**RESULTS**

The means across subjects of the measured pure-tone thresholds for each of the different earphones can be seen in Figure 1 (left). It is clear from this figure that there is a considerable difference between the different earphones especially at the lower frequencies and that this difference is most pronounced for the circum-aural earphone, Sennheiser HDA200.

The reason for the difference between the threshold data obtained with the five earphones is expected to be found in the calibration technique. If the IEC 60318-1 coupler provided the same acoustical load of the earphones as human ears do on average, the thresholds in Figure 1 (left) would principally be identical.

In order to verify this hypothesis, the thresholds are re-calibrated to a sound pressure level existing at the entrance of the ear-canal of the listener, with the ear-canal blocked. The re-calibration includes measurements of two transfer functions, 1) the transfer function of the earphone placed on the IEC 60318-1 coupler $H_{coupler}$ and 2) the transfer function of the earphone on a human ear $H_{PTF}$, measured with a miniature microphone. If the parameters are considered in the frequency domain and averaged across subjects, the re-calibration is as follows:

$$T_{cor} = (|H_{PTF}| - |H_{coupler}|) + T_{ETSPL} \quad [\text{dB re. } 20\mu\text{Pa}]$$

where $T_{cor}$ is the corrected threshold values and $T_{ETSPL}$ is the standardized threshold measured with each earphone.

The corrected mean thresholds can be seen in Figure 1 (right).

**CONCLUSION**

In conclusion, the hypothesis that the IEC 60318-1 coupler does not provide the right load of the earphone is true. Furthermore, until a coupler is designed which provides a load corresponding to human ears, it should be recognized that different absolute thresholds will be measured when using different earphones.

**REFERENCES**

In order to investigate sensory dimensions of low frequency sound, a kind of three-way principal component analysis, PARAFAC (Parallel Factor Analysis), was applied to the subjective rating data of infrasound at 3, 5, 10 Hz and low frequency pure tones at 20, 40 Hz on 5-point rating scales of 22 scale items. The experiment was done in a low frequency pressure-field chamber (4.8 m³), and 30 data-sets of subjective ratings of 20 stimuli on 22 scales were obtained from 17 subjects. The obtained 20 (stimuli) x 22 (scales) x 30 (subjects) three-mode data were analyzed by PARAFAC model, and the best-fitted three factors solution was obtained. It is characteristic of this method that the directions of factors (axes) are uniquely determined and different weights of the factors are permitted for each subject. The obtained factors were interpreted as auditory and pressure-feeling, vibration-feeling and pressure-feeling. All the factors were also related with unpleasantness. It was also observed that the factors corresponded to the particular frequencies of the stimuli and sex differences exist in the weights of factors.

INTRODUCTION

Sensory impression of low frequency sound is a compound of various feelings, and it is supposed that individual differences of sensitivities to the various aspects of the feelings will affect the evaluation of low frequency sound [1,2,3]. The purposes of this paper are to extract the common factors to all the subjects on subjective feelings of low frequency sound, and to clarify the individual differences of the weights to the common factors. For this purpose, the subjective rating data [1] were reanalyzed by using PARAFAC [4], which provides a set of factors with unique orientation and individual subjects’ weights or, perceptual importance, of the factors.

METHOD

Apparatus: The experiment was carried out in a 4.8 cubic meter pressure-field chamber (1.95 x 1.4 m base with a height of 1.75 m). The stimulus sound was generated by four 46 cm loudspeakers, which were mounted in the ceiling, and driven by DC amplifiers. The waveform distortion at infrasonic frequencies was improved by motion feedback.

Subjects: Eight male and nine female subjects participated in the experiment. Thirteen subjects of them repeated two rounds of experiment and thirty sets of data in total including repetitions were analyzed.

Stimuli: Table 1 shows 22 low frequency pure tones used as stimuli in the experiment.

<table>
<thead>
<tr>
<th>Frequency Hz</th>
<th>Sound Pressure Level (dB(L))</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>110 115 120 125</td>
</tr>
<tr>
<td>5</td>
<td>110 115 120 125</td>
</tr>
<tr>
<td>10</td>
<td>100 105 110 115</td>
</tr>
<tr>
<td>20</td>
<td>80  90  100 110</td>
</tr>
<tr>
<td>40</td>
<td>70  80  90 100</td>
</tr>
</tbody>
</table>

Rating scale-items: Table 2 shows 22 rating-scale items used in the experiment. Subjective impressions of the stimulus sound were rated on a continuous scale with respect to 22 scale items listed in Table 2. The scale was used as continuous one, but it was assigned numbers with labels at equal intervals, such as 1: insensible, 2: slight, 3: moderate, 4: considerable, 5: extreme.

<table>
<thead>
<tr>
<th>Table 2. Rating items for subjective impressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Pressure feeling</td>
</tr>
<tr>
<td>2) Pressure feeling on eardrum</td>
</tr>
<tr>
<td>3) Pressure feeling on head</td>
</tr>
<tr>
<td>4) Pressure feeling on breast</td>
</tr>
<tr>
<td>5) Vibration feeling</td>
</tr>
<tr>
<td>6) Vib. feeling on eardrum</td>
</tr>
<tr>
<td>7) Vib. feeling of head</td>
</tr>
<tr>
<td>8) Vib. feeling of breast</td>
</tr>
<tr>
<td>9) Vib. feeling of abdomen</td>
</tr>
<tr>
<td>10) Vib. feeling of buttock</td>
</tr>
<tr>
<td>11) Vib. feeling of legs</td>
</tr>
<tr>
<td>12) Loudness</td>
</tr>
<tr>
<td>13) Muddiness of sound</td>
</tr>
<tr>
<td>14) Unpleasantness</td>
</tr>
<tr>
<td>15) Unpleasantness by vib.</td>
</tr>
<tr>
<td>16) Unpleasantness by sound</td>
</tr>
<tr>
<td>17) Noisiness</td>
</tr>
<tr>
<td>18) Oppression</td>
</tr>
<tr>
<td>19) Suffocation</td>
</tr>
<tr>
<td>20) Sway of body</td>
</tr>
<tr>
<td>21) Heart-beat</td>
</tr>
<tr>
<td>22) Total intensity</td>
</tr>
</tbody>
</table>
Analytical procedure: Obtained three mode data, i.e. I (20-stimuli) x J (22-scale-items) x K (30-subjects) matrix, X = (x_{ijk}) were analyzed by PARAFAC. In order to exclude the variations of scale origins and units between subjects, the data-preprocessings were done for each subject by equations, (1) and (2).

$$\sum_{i}^{I} x_{ijk} = 0 \text{ for all } j, k$$  \hspace{1cm} (1)

$$\sum_{i}^{I} x_{ijk}^2 = I \cdot J \text{ for all } k$$  \hspace{1cm} (2)

Under the following conditions,

$$\sum_{i}^{I} a_{it} = 0 \text{ , } \sum_{i}^{I} a_{it}^2 = I \text{ , } \sum_{j}^{J} b_{jt}^2 = J \text{ for all } t$$  \hspace{1cm} (3)

the least squares criterion (4) was minimized by alternative least square method assuming T factors.

$$E^2 = \frac{1}{I \cdot J \cdot K} \sum_{i}^{I} \sum_{j}^{J} \sum_{k}^{K} (x_{ijk} - \sum_{t}^{T} a_{it} b_{jt} c_{kt})^2$$  \hspace{1cm} (4)

RESULTS AND DISCUSSION

Figures 1, 2 and 3 shows obtained three-factor solutions of the three modes, A, B and C, respectively. The goodness of fit of the model, $1 - E^2 / \sigma^2$, was 0.710.

Figure 1 shows that the 1st, 2nd and 3rd factors of A-mode were most related with the stimulus SPL’s at 40 Hz, with those at 10 Hz and 20 Hz, and with those at 3 Hz and 5 Hz, respectively. Figure 2 shows that the 1st, 2nd and 3rd factors of B-mode were most related with the auditory and pressure-feeling items, with the vibration-feeling items, and with the pressure-feeling items, respectively. All of the factors were also related with unpleasantness. Figure 3 shows the individual differences in the loadings to the three factors of C-mode, and sex differences were clear in the loadings to the third factor. By examining X of male and female subjects, it was shown that the female subjects felt pressure and unpleasantness more by the stimuli at 40 Hz. On the other hand, the male subjects were more sensitive than the female subjects in the evaluations of pressure-feelings and unpleasantness by the stimuli at 3, 5 and 10 Hz. It was also shown that unpleasantness was highly correlated with the pressure-feelings of the stimuli in this experiment.

**REFERENCES**

Threshold, Unpleasantness and Acceptable Limits of Low Frequency Pure Tones for Elderly People

Y. Inukai, N. Nakamura and H. Taya

Institute of Human Science and Biomedical Engineering
National Institute of Advanced Industrial Science and Technology (AIST)
Tsukuba Central 6, Higashi 1-1, Tsukuba, Ibaraki 305-8566, Japan

Threshold and equal unpleasantness sound pressure levels of pure tones from 10 Hz to 100 Hz were obtained for 12 elderly subjects by the method of adjustment (production) in which subjects adjusted sound pressure levels equating to subjective degrees of unpleasantness on a 5-point rating scale. And the equal unpleasantness contours will be estimated for the obtained data. In addition, the maximum acceptable sound pressure levels for the elderly subjects were measured on each frequency, assuming two kinds of situations, that is, a living room and a bedroom. The obtained results were compared with those of younger subjects, which were obtained from our previous experiment. It was shown that elderly people’s thresholds of sound pressure levels at low frequencies were higher but the dynamic ranges of unpleasantness at low frequencies were narrower than those of younger subjects. The differences were examined in the relations with the difference of the hearing loss of subjects.

INTRODUCTION

In our previous study, it was found that not only audible loudness, but also oppression and vibration feelings are dominant subjective impressions to low frequency pure tones[1]. And unpleasantness was considered as one of the most representative evaluation scale of these negative feelings[2]. The purpose of this study is to obtain equal unpleasantness contours and acceptable limits of low frequency sound pressure level for elderly subjects, and to compare the results with previously obtained results for younger subjects[3].

METHOD

Stimuli: Sixteen pure tones were used as stimuli at frequencies with 1/3-octave intervals in the range from 10 Hz to 250 Hz and at 500 Hz.

Subjects: Twelve paid volunteers, six females and six males aged from 60 to 75, participated as subjects in the experiment.

Apparatus: The experiment was carried out in the low frequency pressure chamber. The internal dimensions were 3.5 x 2.5 m base with a height of 2.6 m. The chamber was constructed of concrete wall covered with grass wool. The absorption rates were 0.6/125Hz and 0.95/500Hz, and the background noise was less than 16dB(A). The stimuli were generated with a sine/noise generator (B&K1049) and presented to subjects through the sixteen 46 cm diameter loudspeakers mounted in the wall of the chamber.

Subjects produced their sound pressure levels by manipulating a remote volume controller of the power amplifiers. The controller can set levels of stimuli at steps below 1dB. Sound pressure levels were measured by the microphone in the chamber and calibrated at the position of subject’s head.

Levels of unpleasantness: The three levels of unpleasantness, U2, U3 and U4, were used. They were selected from five categories defined on 5-point rating scale such as U1: not at all unpleasant, U2: somewhat unpleasant, U3: unpleasant, U4: quite unpleasant, U5: very unpleasant.

Assumed situations for acceptable limits of sound pressure levels: Two different situations were assumed as follows. AL: A living room, assumed reading a newspaper quietly sitting on a sofa. AB: A bedroom, assumed lying down on the bed to get to sleep.

Procedure for measurement of unpleasantness: Subjects were required to produce a sound pressure level of a given stimulus relevant to a given category on the 5-point rating scale of subjective unpleasantness by manipulating the volume controller. The produced sound pressure levels were measured by the computer controlled recording system. Considering the subjects’ ages, U5 were not used in this experiment and the experimenter assigned one of the other three categories to subjects in each experiment.

Procedure for measurement of acceptable limits: Subjects were required to produce a sound pressure level of a given stimulus, which reflects their acceptable limit of sound pressure level for the assumed situation by manipulating the volume controller. In each experiment, one of the above
mentioned two situations was indicated by the experimenter to subjects. Every situation was assigned once to all the subjects.

RESULTS AND DISCUSSION

As the results of preliminary analysis of the observed data, two subjects with high thresholds were separated as Group B from Group A of the other 10 subjects.

Figure 1 shows the obtained results from this experiment for Group in comparison with the equal unpleasantness contours calculated by equation (1) which was obtained from younger subjects in the previous report [3].

\[ L = \frac{y+43.5-37.4x+10.7x^2-0.988x^3}{0.312 - 0.176x + 0.0360x^2}, \]

(1)

Where L indicates the estimated sound pressure level in dB (L) at the frequency h for a given rating (y) of unpleasantness and x indicates the logarithm of the frequency h (10 \( \leq h \leq 500 \)).

The equal unpleasantness contours obtained by (1) for ratings, 2, 3, 4, 5 were shown as four curves denoted by \( U_2[3] \), \( U_3[3] \), \( U_4[3] \) and \( U_5[3] \) in Figure 1. It was shown that elderly people’s thresholds (T) of sound pressure levels at low frequencies were higher than those (\( T[3] \)) of younger subjects. But the dynamic ranges of sound pressure levels of equal unpleasantness at low frequencies for elderly subjects were narrower than those for younger subjects. The sound pressure levels of acceptable limits for the two situations were observed at very low level unpleasantness, which were similar to the results of younger subjects observed in the previous report [3].

Figure 2 shows the mean values of the observed threshold, equal unpleasantness and acceptable limits from the two subjects with greater hearing loss. It was shown that the thresholds (T) were high, and the dynamic range of unpleasantness was very narrow. Unpleasantness rapidly increased from \( U_2 \) to \( U_3 \), \( U_4 \) as the sound pressure levels increased over thresholds.

Figure 3 shows the group means of hearing loss obtained from the two groups of subjects by the audiometer. The results show that there is a large difference in hearing loss between the two groups of subjects. Although the sample size was small, it is conjectured that individual sensibilities to low frequencies below 100 Hz might be closely related to the hearing loss at higher frequencies than 100 Hz.

REFERENCES

Acoustic Features and Impressions of Musical Timbres by Children with Hearing Impairment

Y. Ota and Y. Kato

Institute of Disability Sciences, University of Tsukuba
1-1-1 Tennodai, Tsukuba, 305-8572, Japan

This study aimed to clarify the relation between the acoustic features of musical timbres and impression on the timbres by children with hearing impairment. Ten different musical instrumental timbres (Piano, Flute, Violin, Bell, Harmonica, Trumpet, Harpsichord, Recorder, Xylophone, and Organ) were employed with MIDI system as stimuli. Twenty-nine subjects who were hearing impaired children were asked to rate the impression on the timbre of stimuli on 7-point adjective scales based on Semantic Differential. The stimuli were analyzed by DSP Sona-Graph (KAY 5500) to investigate acoustic features of them. The results obtained were follows. The Harpsichord tone had a short attack and a long decay and contained strong components higher than 8000Hz. The subjects evaluated the Harpsichord timbre as “harmonious” “good” and “beautiful”. The result showed that “decaying sounds” was impressed as “harmonious” “clear” “beautiful” “easy” and “good” compared to “steady-state sounds”. Therefore it was apparent that acoustic features of envelope pattern (decaying or being steady) become effective to impression on the timbres for children with hearing impairment.

INTRODUCTION

There are many issues how to evaluate musical sounds by hearing-impaired children. To investigate their subjective aspects for musical sounds is necessary as well as their musical abilities (such as discrimination or recognition of pitch, loudness and timbre). In this study, we investigated the relation between the acoustic features of musical timbres and impression of the timbres by children with hearing impairment. And we used the MIDI system. Its system may be preferable in educational situation because it can easily be used without the real musical instruments and reproduced.

METHOD

Subjects

Subjects were 29 children with hearing impairment at the School for the Deaf. Criterion for inclusion in this study was no other known impairment. Subjects were 12 females and 17 males ranged in age from 13 to 15 years. Hearing loss for all children was classified as severe or profound (range in better ear pure tone average from 87 to 117 dB Hearing Level). All subjects daily use personal hearing aids.

Stimulus

Ten different musical instrumental timbres were employed with MIDI system as stimuli. The ten tones (Piano, Flute, Violin, Bell, Harmonica, Trumpet, Harpsichord, Recorder, Xylophone, and Organ) were generated by a music keyboard (Roland SK-50) and edited with software (Roland Ballade). The tones were controlled the pitch and duration (middle C=260Hz, a quarter note=60). The stimuli were recorded with a DAT (Sony DTC-1000ES) and presented trough an amplifier (Audio-Technica AT-SA50) and a loudspeaker (Audio-Technica AT-SP50). Because the stimuli were the same pitch and duration, timbre is the only factor that changes in each stimulus. The stimuli were analyzed by DSP Sona-Graph (KAY 5500) to investigate acoustic features of them. Parameters in the analysis were waveform, wideband spectrogram (bandwidth was 300Hz) and narrowband spectrogram (bandwidth was 29Hz).

Procedure

The stimuli were presented to a subject through a loudspeaker in a soundproof room and the timbre of the stimuli was evaluated with the Semantic Differential method. Subjects were instructed to rate their impression of timbre on 7-point adjective scales. The ten pairs of adjective were showed in Table 1. Each stimulus was presented ten times until the subjects had rated the impression for all adjective pairs.
Table 1. Pairs of adjective to rate the impression of the timbre.

<table>
<thead>
<tr>
<th>Pairs of adjective</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>weak</td>
<td>strong</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>clear</td>
<td>thick</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dull</td>
<td>distinct</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>harmonious</td>
<td>discordant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ugly</td>
<td>beautiful</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>annoying</td>
<td>quiet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>easy</td>
<td>difficult</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hard</td>
<td>soft</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>calm</td>
<td>shrill</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>good</td>
<td>bad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

RESULTS and DISCUSSION

The scores by all subjects were averaged for each pair of adjectives. The mean scores plotted the profile curves. Based on the acoustic analysis, all timbres were divided into two types of envelope pattern (“decaying sounds” and “steady-state sounds”). Figure 1 and 2 show the impression of the timbres categorized as “decaying sounds” and “steady-state sounds”. The timbre of “decaying sounds” was evaluated variously compared to “steady-state sounds”. These timbres were characteristic for the hearing-impaired. These timbres were particularly impressed as “easy” and “good”. In the opposite, subjects didn’t show strong impression for “steady-state sounds”. These timbres were impressed as featureless or relatively showed negative impression. And most positive impression was showed on Harpsichord timbre. Children with hearing impairment evaluated the Harpsichord timbre as “harmonious” “good” and “beautiful”. Its timbre has a short attack and a long decay and contained strong components higher than 8000Hz. Therefore it is probable that other parameters such as spectrum also have effects on impression of the musical timbre for hearing-impaired children in addition to the envelope pattern (decaying or being steady).

CONCLUSION

We reported on the relation of the acoustic features and impression of the timbre by hearing-impaired children. It was apparent that acoustic features of envelope pattern (decaying or being steady) become effective to impression of the timbres for children with hearing impairment.

REFERENCE

Calculation Model of Time-varying Loudness by Using Critical-band Filters

H. Jeonga, and J. G. Ihb

a Virtual Reality R&D Center, Electronics and Telecommunications Research Institute, Korea
b Department of Mechanical Engineering, Korea Advanced Institute of Science and Technology, Korea

In this study, a new loudness model is suggested for dealing with the transient sound for a unified analysis of various practical sounds. A signal processing technique is introduced for this purpose, which is required for the band subdivision and the prediction of band-level change of transient sounds. In addition, models for the post-masking and the temporal integration are adopted in the analysis of the loudness of transient sounds.

LOUDNESS MODEL

The schematic procedure of loudness model for unsteady sounds based on Zwicker's is suggested in FIGURE 1. The frequency range for the loudness analysis of arbitrary noise sources has to be covered about from 20 Hz to 16 kHz. Therefore the acoustic signal acquired by a microphone is converted to the digital signal with the sampling frequency above 40 kHz. The loudness calculation procedure for unsteady sounds contains digital 47 critical bands filtering, envelope extraction, calculating main loudness, post-masking, spectral masking, loudness summation and temporal integration. Each step will be described in detail.

47 Critical Bands Analysis

In this study, 47 band-pass filters overlapping each boundary are used for the improvement of calculation accuracy. FIGURE 2 shows the procedure of critical bands analysis by using 47 critical band filters. For the stability of filtering process, the half-rate resampling technique is required [2]. Cut-off frequencies of each filter are center frequencies of neighboring bands.

Next, the envelope extraction of band-pass filtered signals is needed for calculating the sound pressure level change of each signal. In this study, as the square-law rectification process, the low-pass filtering is conducted by using the one pole low pass filter with time constant of 2 ms after calculating the power of band-pass filtered signal. One pole low pass filtering acts as an exponential averaging. The filter outputs are smoothed by using this filter and given by:

\[ P(n) = (1-A)P(n-1) + Au(n) \]  

where \( u(n) \) is the instantaneous power input and \( P(n) \) is the average power output [2]. The factor A can be expressed as \( A = t/\tau \), where \( t \) is the original sampling interval and \( \tau \) is the desired time constant for the exponential averaging.

Then sound pressure level at each critical band is given by:

\[ L = 10 \log_{10}(P/P_{ref}) \]  

where \( P \) is the mean power of instantaneous band-pass filtered signals and \( P_{ref} \) is the reference value, \( 4 \times 10^{-5} \text{ Pa}^2 \). Since loudness values for unsteady sounds are required on every 2 ms according to Zwicker [1], the

![FIGURE 1. Schematic diagram of time-varying loudness model.](image)

![FIGURE 2. Procedure of critical band analysis by using 47 critical band filters for the present time-varying loudness model. Each number in the figure represents the center frequency in Hz of each critical band filter.](image)
sound pressure level at each band is acquired on every 2 ms.

Calculation of Main Loudness

The critical band levels in dB are converted to main loudness in sone, which have linear relation with perceptual loudness by the following:

\[ N' = N_0 \cdot 10^{0.025 L_{10}} \cdot \left[ 1 - s + s \cdot 10^{0.1 (L - L_{10})} \right]^{1.25} - 1 \]  \hspace{1cm} (3)

where \( L_{10} \) is the excitation level at threshold in quiet, \( L_z \) is the excitation level of the signal obtained by the critical band analysis, \( s \) is the threshold factor \( [1] \). The constant \( N_0 \) is chosen so that the loudness of 1 kHz pure tone at 40 dB will be exactly 1 sone.

Simulation of Post-masking

Post-masking is described as the masking of a sound by a sound occurring previously. Zwicker [1] proposed the analog non-linear model for simulating post-masking effect. Recently, this analog model was revised for digital sound analysis system by Widdmann et al. [3] and this digital algorithm is used in the present model.

Simulation of Spectral-masking

The procedure making the slope loudness toward higher frequency is similar to the conventional algorithm [1] but the boundary step of each critical band in the previous algorithm is changed to 0.5 Bark instead of 1 Bark. The slope of specific loudness in additional bands is also acquired by using the linear interpolation of previous values. By summation of specific loudness at each time step, the instantaneous loudness value is acquired.

Loudness Integration

For analysis of time-varying loudness, the simulation process for temporal loudness summation is needed. In the previous study [1], 1-pole filter was simulated for loudness temporal integration. In this study, 1-zero, 1-pole filter is suggested for the simulation. The transfer function of the filter is the following:

\[ H(z) = \frac{B(z)}{A(z)} = \frac{b_1 + b_2 z^{-1}}{1 + a_1 z^{-1}} \]  \hspace{1cm} (4)

where \( a_2, b_1, \) and \( b_2 \) are \(-0.952, 0.308 \) and \(-0.260, \) respectively.

For verifying of the present model, the predicted loudness values of shout time burst tones are compared with the experimental data and it is observed that they are in good agreements (FIGURE. 3).

CONCLUSION

In this paper, the digital algorithm of loudness model for unsteady sounds based on Zwicker’s method is proposed for the computer programming. In order to enhance the calculation accuracy, overlapped 47 critical filters are used. The usefulness of present model is confirmed by comparison the calculation results of present model with psychoacoustic facts.

REFERENCES

The Franssen effect in a multisignal environment

K. Jambrošić, B. Ivančević and A. Štimac

aFaculty of EE and Computing, Unska 3, HR-10000 Zagreb, Croatia, kristian.jambrosic@fer.hr
bBrodarski Institute, Av. V. Holjevca, HR-10000 Zagreb, Croatia

A common effect of spatial hearing in enclosed spaces is the so-called Franssen effect [1]. In our research, several parameters of the Franssen effect occurrence were researched. First, the nature of the transient sound was examined. It was found that it is not necessary for the transient signal to be of same wave shape and amplitude level as the steady-state one. Second, measurements of the effect occurrence for more then one loudspeaker playing the transient and/or the steady-state signal simultaneously were made in order to find out how it affects the localization accuracy. Although the Franssen effect signals are sinusoid signals, the existence of the effect by adding two frequency components to the signal was measured.

INTRODUCTION

The Franssen effect is a well-known consequence of the precedence effect (also known as the Haas effect), [2]. It is the mechanism how we localize wide-band sounds in a reverberant enclosed space where strong reflections occur beside the main, direct sound wave. The direction of the direct wave, which comes first to the analyzing listener, gives the localization clue.

Basically, the Franssen effect is produced using two loudspeakers. One is driven with a sinusoidal signal of several seconds duration, with exponential onset and decay curves (the steady-state signal). The other is driven with a signal which, when added to the first one, gives a sinusoid modulated in amplitude by a rectangular pulse of same duration as the first signal. The second signal can be called the transient one, fig. 1. When these signals are played on a loudspeaker pair in stereo setup, only one sinusoid is localized in the direction of the transient signal during the duration of the experiment, thus verifying the precedence rule.

One important point is that this effect can be heard only by loudspeaker reproduction and not also using headphones (due to the lack of later reflections).

MEASUREMENTS

The existence of the Franssen effect for different loudspeaker arrangements was measured. The measurement was done in an acoustically treated room 10 x 7 m (a music listening room) with an average reverberation time of 0.5 s. The loudspeakers were arranged as shown in fig. 2. The test persons were always turned as showed in the figure without moving the head during the test procedure. This is very important because during the measurement of the Franssen effect, a strong standing wave pattern appeared in the room, thus spoiling the occurrence of the effect if the head was not held still. The effect was measured for sinusoids of 125, 160, 250, 500, 1k, 2k
and 4k Hz. The test persons were sitting at different positions between the four loudspeakers with the head turned always in the same direction. Common for all measurements was that both signals, the steady-state and the transient one, were 5 seconds long always shaped in the same manner: the steady-state signal had an exponential onset and offset with a time constant of 50ms, and the transient had an exponential offset and onset complementary to the steady-state one with the same time constant. They were of same peak amplitude. Here are the researched parameters:

**Localization precision**

The localization precision was measured by comparison of the effect while playing the front pair (FL and FR), side pair (FR and RR or FL and RL) and rear pair of loudspeakers (RL and RR). The measured precision of the localization coincides with the localization resolution of the human ear for the same direction. The frontal pair gave an excellent Franssen effect precision. The side pair was not always as good as the frontal (if the transient was fed to the rear side speaker), and the back pair experiment showed the much smaller resolution of men’s hearing comparing to the frontal region.

**Transient signal waveform**

The next measurement was to determine how would the wave shape of the transient signal other than sinusoidal influence the occurrence of the Franssen effect. FL and FR loudspeakers were used. The steady-state signal was always a sinusoid, and the transient was sinusoid, triangle, rectangle and white noise, respectively. The results were not surprising; the best effect was obtained for the sinusoid and the triangle transient. The rectangle was more noticeable because of the richer spectral content, but didn’t spoil the Franssen effect. White noise shaped as the transient made the transient very noticeable, but it made also the localization more confuse, i.e. the direction of the sinusoid was approximately somewhere between the two test loudspeakers.

**Transient signal amplitude**

The aim of this experiment was to determine if the amplitude of the transient signal has to be the same as the amplitude of the steady-state signal to gain the Franssen effect. The amplitude of the transient signal was lowered in 3dB steps until the localization became blurry and the sound image begun to move to the steady-state signal position. The experiment was repeated for various frequencies and the overall result is that at approximately –18 dB less level of the transient signal, the Franssen effect begins to fade.

**Steady-state signal waveform**

The next experiment was to verify the Franssen effect existence when simultaneously playing two sinusoid signals. If the signals were in a whole number ratio (like 500 and 1000 Hz), the effect was still present but not so clear as if there were on sinusoid. When two signals of independent frequencies were used, the localization was very uncertain.

Another experiment was to use a swept sinusoid to check the effect. The sinusoid was swept in the range of 0.5 – 2 kHz. At the beginning, the Franssen effect occurred at the transient signal position, but after approximately 0.8s the sound image begun to move to the other loudspeaker.

**More than channels playing simultaneously**

The last set of measurements was checking the Franssen effect using more then two loudspeakers playing simultaneously. In the first case, the transient signal was played at two loudspeakers, and the steady state at one. A very interesting result was found: the direction of the sound image was either at the first transient signal position or at the second, depending on the position of the test person. The sound image was never perceived somewhere in-between the transient signals, once more verifying the precedence effect.

The second case was playing the steady-state signal at two loudspeakers simultaneously. This experiment didn't give a new sensation of the effect comparing to the 2-loudspeaker arrangement. The reason is simple – the steady-state signal makes a complex standing wave pattern in the reverberant room no matter if playing from one or two loudspeakers.

**CONCLUSION**

Although the Franssen effect is a very well known effect, there were some parameters researched in this paper not previously found in the literature. All of the measured parameters showed the power of the precedence effect having in mind the limitations of the effect depending on the test signals shape and position.

**REFERENCES**

A psycho-acoustical experiment on HVAC noise

K. Ueno\textsuperscript{a}, K. Yasuda\textsuperscript{b} and H. Tachibana\textsuperscript{a}

\textsuperscript{a} Institute of Industrial Science of Tokyo Univ., Komaba 4-6-1, Meguro-ku, Tokyo, 153-0041, Japan
\textsuperscript{b} Taisei Corporation, 1-25-1 Nishi-Shinjuku, Shinjuku-ku, Tokyo, 160-0023, Japan

For the assessment of HVAC noise, such rating methods as NC, NCB and RC curves are proposed. However, their validity has not yet been fully examined from a psycho-acoustical viewpoint. In this paper, subjective experiments on the assessment of loudness and disturbance using a newly developed 6-channel recording/reproduction system are reported.

INTRODUCTION

In order to assess the environmental noises, various kinds of noise indices are proposed and used. As a study to examine the validity of these noise indices from a psycho-acoustical viewpoint, we are making subjective investigation using 6-channel recording/reproduction system developed in our laboratory \cite{1,2}, in which 3-dimentional acoustic properties of the real sound fields can be simulated in an anechoic room. In this paper, the results of the subjective experiments on HVAC noise are introduced.

EXPERIMENTAL SYSTEM

Figure 1 shows the diagram of the experimental system. To realize natural impression actually being in a space, 6-channel recording/reproduction system was employed. The recording system consists of six directional microphones (SONY C48), by which the sounds from spatially orthogonal six directions are received and recorded onto a 6-channel digital data recorder. The signals are reproduced in an anechoic room from six loudspeakers set in the directions corresponding to the six microphones. In this study, one woofer was added in order to reinforce the low frequency components. The subject sat at the center point of the sound field and make judgment.

SUBJECTIVE TEST

Using the recording system mentioned above, HVAC noises with different spectral characteristics were collected in a variety of spaces. The sounds were reproduced at the same levels as in the actual fields and the following two kinds of experiments were performed.

Experiment on Loudness

As the test procedure, the method of adjustment by subjects was adopted. That is, the standard stimulus (Ss) and the comparison stimulus (Sc) were presented alternately, and the subject adjusted the level of Sc with a remote-control attenuator so that it was perceived as being equally loud as Ss. As the test sounds (Ss), 40 kinds of HVAC noises with 28.7 dB to 53.9 dB in $L_A$ were reproduced from the 6 channel loudspeakers. As Ss, a broad band noise with a spectral characteristic of $-5$ dB/octave was adopted since this noise is natural and suitable for representing general noises in living environments \cite{3}. It was reproduced from the loudspeaker set just above the subject. The duration time of Ss and Sc was 6 s each.

Each loudness matching test was performed in both the cases such that Sc was set definitely louder and softer than Ss at the beginning. In each process of loudness matching, the subject was allowed to adjust the attenuator freely, upwards and downwards. The point of subjective equality (PSE) for each Sc for each subject was obtained as the mean value of the results.
of these ascending and descending procedures. Twelve subjects with normal hearing ability participated in this experiment.

The mean values of PSEs judged by all subjects for each test sound were calculated based on several noise indices. Among the results, Fig. 2 shows the correspondence between Ss and Sc in $LL(Z)$ and NC. Due to the difference of the reproductive condition of Ss and Sc as mentioned above, it was confirmed that there is systematic difference between Sc and Ss (3.3 dB in $L_A$) by the preliminary experiment. Therefore, the standard errors for the linear model considering the systematic difference in each index were calculated (Table 1). In the results, it has been confirmed that the precise index, $LL(Z)$, is superior loudness index for HVAC noise. Besides, arithmetic averages of SPLs in octave band, $L_{m(63-4k)}$ and $L_{m(125-4k)}$ [3], are also highly correlated to loudness rather than the tangency methods, such as NC and RC. It has also been confirmed that $L_A$ is a proper estimator of loudness.

![FIGURE 2 Relationship between Ss and Sc on loudness experiment](image)

### Table 1. Standard error in each index on loudness experiment

<table>
<thead>
<tr>
<th>Index</th>
<th>Std. Error</th>
<th>Index</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$LL(Z)$</td>
<td>1.82</td>
<td>RC</td>
<td>3.34</td>
</tr>
<tr>
<td>$L_A$</td>
<td>2.65</td>
<td>$L_{m(63-4k)}$</td>
<td>1.33</td>
</tr>
<tr>
<td>NC</td>
<td>4.91</td>
<td>$L_{m(125-4k)}$</td>
<td>1.84</td>
</tr>
<tr>
<td>NCB (SIL)</td>
<td>3.10</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Experiment on Disturbance**

As a trial to assess the annoyance in actual living condition, an experiment focusing on the disturbance under the condition of writing sentences on a word processor was performed. As the test method, the rating scale method was applied. That is, after the writing work for two minutes, the subject was asked to answer the impression of disturbance of the surrounding noise during the work in five-step categories as follows; “During the work, the surrounding noise is 1) not disturbing at all, 2) little disturbing, 3) moderately disturbing, 4) very disturbing, 5) extremely disturbing.” As the test sounds, 18 kinds of HVAC noises with 28.7 dB to 53.3 dB in $L_A$ were used. 9 subjects with normal hearing ability participated in the experiment.

The mean values of the subjective responses were calculated for each test sound. Figure 3 shows the relationship between the experimental results and noise indices. In the results, it has been indicated that the disturbance impression correlates to loudness estimators on the whole. Concerning the sound quality judgment based on RC criteria ( ○ in Fig. 3), consistent relationship between the rating and disturbance is not observed.

![FIGURE 3 Relationship between the results on disturbance and noise indices](image)

**CONCLUSIONS**

Through the psycho-acoustical experiments on HVAC noise, it has been confirmed that $LL(Z)$ is highly appropriate index of loudness and $L_{m(63-4k)}$ and $L_{m(125-4k)}$ and $L_A$ are also superior indices to such the tangency methods as NC and RC. Regarding the disturbance, it has been indicated that the judgment almost correlates to loudness estimators, whereas the effectiveness of the evaluation of sound quality, “rumble” or “hissy” in RC, has not been found in this experiment.

**REFERENCES**

Auditory Discrimination of Tool Wear: A Case of Applied Psychoacoustics

G. Kidd, Jr., C.R. Mason, B. Fligor and C. Carter

Department of Communication Disorders and Hearing Research Center, Boston University, 635 Commonwealth Avenue, Boston, MA 02215, USA

In industrial settings, the operators of various types of machinery often use auditory monitoring to determine the state of the tools being used, such as the degree of wear or the likelihood of impending failure. In this study, human listeners were asked to discriminate the degree of wear of tools used in a steel milling process. The stimuli were 1-second samples taken from longer recordings of eight tools as they progressed from 'new' to 'worn'. The task was to listen to randomly drawn samples from the set for a particular tool and choose which sounded 'more worn' in a 2AFC procedure. The results revealed a monotonic function relating discrimination to the physical difference in degree of wear. Several manipulations, including low-pass and high-pass filtering, rms-amplitude equalization and random rove in level, helped isolate the acoustic cues underlying discrimination. Moreover, peaks in the function relating discrimination to adjacent milling runs indicated probable local occurrences of 'change of state' and provided markers for regions warranting detailed acoustic analysis. These findings suggest that psychoacoustic experiments can help to identify the acoustic cues that indicate tool wear and provide a means for determining features for automatic pattern recognizers used in sensing and monitoring applications.

INTRODUCTION

This work examined whether trained listeners could accurately discriminate the degree of tool wear that occurs in a machine milling process. The study used forced-choice psychophysical techniques in which the stimuli were accelerometer recordings of passes of the milling process presented as acoustic waveforms through an earphone. The listener judged which of two brief samples taken from different milling passes sounded 'more worn'. Listeners were trained to make this discrimination by response feedback after every trial with a 'correct response' defined as accurately choosing which of the two stimulus presentations was drawn from the later pass.

A second purpose of the study was to learn which acoustic variables formed the basis for the perceptual judgments. Because the physical variable 'degree of tool wear' was defined by the order within the sequence of passes, rather than by specific acoustic parameters, additional experiments were undertaken to learn which spectral or temporal features of the recordings were important for discrimination. In those experiments, the information available to the listeners was limited by low-pass or high-pass filtering and loudness cues were limited by RMS equalization and level randomization. The ability of the listeners to generalize to different and untrained tools (e.g., 1" diameter tools vs. the 1/2" diameter tools used in training) was also tested.

METHODS

The listeners were four young-adult students. The stimuli were 1-sec samples from a set of recordings of milling runs from 1/2" and 1" diameter end mills used in the manufacturing of airplane parts. The samples were equated for RMS and had 10-ms gating ramps. The level was 70 dB SPL excluding the 10-dB rove in level. The stimuli were low-pass filtered at 20 kHz. Listening was monaural through an AKG K-240 earphone. The procedure was two alternative forced choice. Each trial consisted of a pair of passes chosen randomly, without replacement, from all possible passes for a given tool. Each block of trials contained every possible pair of samples one time. A training phase was completed initially using the samples from 1/2" diameter tools S5, S8 and S21. Listeners received response feedback identifying which of each pair of sounds in a trial was drawn from the pass occurring later in the sequence of passes (i.e., greater wear). Following the training phase, in which ten estimates were obtained for every possible pair of passes for each tool, a testing phase was conducted with different tools and three of the four listeners. No response feedback was given during these trials, forcing the listeners to rely on the strategies they had previously learned during training. The test stimuli included one 'new' 1/2" tool (S1) and four 1" tools. Broadband (low pass filtered at 20 kHz), low pass (low pass filtered at 4 kHz), and high pass (high pass filtered at 4 kHz) filtered conditions were tested.

RESULTS

Figure 1 shows percent correct performance as a function of the separation between passes in the
training phase. The discrimination functions increase monotonically and are similar across listeners and tools. This finding supports the initial hypothesis that there is an orderly perceptual correlate to the degree of tool wear as indicated by the milling pass order within the sequence of passes.

One question of interest is whether the perceptual correlates of tool wear are constant or occur in abrupt steps possibly indicating changes in state of the tool. Inspection of the discriminability of pairs consisting of adjacent passes is one way of examining this issue. Figure 2 shows the data plotted in that format.

This figure clearly shows adjacent passes where discriminability was very high. Close inspection of the acoustic patterns for these passes (e.g., pass 8 vs. 9 for S5) revealed transient acoustic cues that had previously escaped detection by a bandpassed energy measure but that provided sufficient information for the listeners to make accurate judgments of wear. Values significantly below the 50% chance point indicate acoustic cues that were reliably, but inaccurately, associated with greater wear.

Next, the results of filtering are considered for the data collected in the test phase. Although four new tools were tested data are shown only from one representative 1"-tool (see Figure 3). For this tool and the others not shown there was very little difference between the results from the training phase and from the test phase. Also, there was not a substantial difference between the discriminability of the 1/2" tools and the 1" tools. In general, discriminability increases as a function of the separation between passes. Thus, there was sufficient information available in both low- and high frequency regions to perform the task.

FIGURE 3. Percent correct discrimination for one new 1"-tool in the test phase for the three filtering conditions (panels) and three listeners (different symbols). Solid lines are intersubject means.

SUMMARY

Listeners can be trained to reliably discriminate among tools with different amounts of wear based solely on listening to acoustic transductions of accelerometer recordings of a milling process. Thus there is an orderly perceptual correlate to the degree of tool wear as specified by the order within the sequence of passes. Analysis of the discrimination functions plotted as pairs of adjacent passes reveals reliable peaks suggesting that there are abrupt changes of tool state. Listeners trained on one set of tools can then reliably discriminate tool wear on unfamiliar tools. The acoustic cues that underlie auditory discrimination of tool wear appear to be complex, occurring in the frequency regions both above and below 4 kHz. Accurate discrimination performance given RMS equalization and level randomization indicates that the cues are related to features in the spectro-temporal patterns of the stimuli and not solely due to changes in the overall energy of the acoustic output.

ACKNOWLEDGEMENTS

The Office of Naval Research (MURI Z883402) supported this work. The authors thank Dr. Gary Bernard of Boeing Commercial Airplanes for helpful discussions and the recordings used in the study.
In this study, we investigated the effect of time uncertainty in forming the temporal window in signal detection tasks. A band-limited noise was presented for the masker and to identify the observation period. The target tone was presented at the center, before or after the center within observation period. Four tones presented at the different position were used as a probe signal.

In 80% of the trials, the signal was presented as a target. In the remaining 20% of trials, the probe signal was presented at the position differing from that of the target. Five thresholds were determined for four probes and the one target. (test condition). Also, five thresholds were determined using the tone of the same position as that of each probe position (control condition).

The temporal window was defined as the difference between the threshold in control condition and the one in test condition. As a result, the measured temporal window has the peak at the target position. We can confirm the temporal window due to the uncertainty in the time domain. Thus, the shape of the temporal window is affected by the position of target and the duration of the observation period.

**INTRODUCTION**

Many studies have been examined on the dynamics of the auditory system. In the frequency domain, the critical band, critical ratio, is the principal concept of auditory system. In the time domain, while, Penner et al.(1973) defined as Critical Masking Interval that the masking increased with the increasing duration of noise burst, but the duration achieved the critical interval. Moore et al.(1988) examined the temporal resolution of the auditory system using temporal window analogous to the auditory filter in the frequency domain. In frequency domain, when we attend to a particular frequency in signal detection task, the sensitivity to other unattended frequency sounds is diminished.(Ebata 1997) The attention filter in auditory system can be explained this phenomena. In this study, we examine the idea that a time window is introduced in auditory system analogous to the attention filter measured in the frequency domain. Many studies reported the time window ranged from a few ms to several hundred ms depended on the presented stimuli, task and contents of information processing. We note that the attention control the detection and form the temporal window in time domain due to control of uncertainty the signal.

**EXPERIMENTAL METHOD**

A two-interval alternative forced-choice (2IAFC) procedure without feedback was used to determined threshold. Two masker intervals of 2000-ms or 3000 ms duration were separated by pauses of 200-ms. The listener's task was to indicate which of the two intervals contained the signal. The band limited noise, 500-2000 Hz, masker was presented to identify the observation period. The level of masker was a 60dB SPL. The frequency of signal was 1000 Hz which had a duration of 10-ms presented at one of several positions in the masker. In test condition, the signal was temporally placed in the center of the masker(0ms), before(-250ms) the center of the masker or after(+250ms) the center of the masker, target position, at 80% of trials. While, on the remaining trials, the signal was temporally differing from the target position, separated -500,-250,+250 and +500-ms from the target position. The signals were presented at five levels – in random order - spaced 2 dB apart, from -4 dB SL to about +4 dB SL set to each listener. Thus, the classical method of constant was used and measured the psychometric function for each listener in test condition. Also, five thresholds were determined using the signal of the same position as each probe position (control condition).
RESULTS AND DISCUSSION

The psychometric functions for each listener were measured in control and test conditions. These functions were obtained for five positions, then the thresholds corresponding to the 75% correct-response point were estimated. The shift of the thresholds from control to test conditions indicated the change of detectability according to time uncertainty and the attention in time domain. Figure 1 shows the threshold shifts from the control condition in case that target position is placed center in the masker of the duration of both 2000 ms and 3000 ms. The datum are averaged over all listeners. The ordinate shows shift of the threshold in dB SPL and the detectability of the signal at a given position. The abscissa shows the relative time position based the center. It is obvious that the uncertainty and attention in time domain cause the higher detectability at the target position and lowers those at the others and form the temporal window. As the duration of the masker is larger, the shape is broader. Figure 2 shows the shifts of threshold in case of target placed the center of the masker(1000 ms) and before(750 ms) the center of the masker which the duration is 2000ms. As the time from onset of the masker is longer, the shape of temporal window is wider. Figure 3 shows the shifts of thresholds in case that target is placed the center of the masker (1000 ms), before (-250 ms) and after (+250 ms) the center of the masker for one listener. It is sure that as the time from onset of the masker is longer, the shape of temporal window is wider.

CONCLUSION

In this study, we examine the idea that a time window is introduced in auditory system analogous to the attention filter measured in the frequency domain. As a result, the higher detectability was measured at target position according to the higher rate of presentation and the window is observed in the time domain. The window has the peak at the target position. Therefore we can confirm the temporal window form to an attention controlled by uncertainty in the time domain.

REFERENCE

Acoustic Quality Comparison between Natural Sound and Noise with the Frequency-gradient Rate

J. H. Hwang*

* Department of Electronics, Hanbat National University, Dukmyung-dong, Yuseong, Daejeon, Korea

Abstract: In general humans are inclined toward natural sounds, but hold abhorrence in hearing noises. Natural sounds are very pleasant to hear and give happiness to human beings. On the other hand, noises are very detestable. These two sorts of sounds are the same acoustic impacts and physical energies. Nevertheless the psychological responses are far different. Pleasant and unpleasant are the contradictory extremes. This paper is described to clarify acoustic qualities of these two sorts of sounds. The typical natural sounds and noises are selected with psychological inquiries. Natural sounds are those of a cricket, birds and so on. Noises are those of a rock drill, creaking and so on. Each sound is analyzed in a frequency domain. With FFT results, a frequency-gradient rate 'gs' is defined. 'gs' is a frequency variant rate with a typically bounded frequency range. And the results are arranged into statistical domain. It is verified that these method are very powerful in classifying sounds into pleasant or unpleasant. Natural sounds generally have gentle slopes, not flat, and low statistical frequency regular factor. But noises have sharp slopes comparing to natural sounds, or flat, and high statistical one.

INTRODUCTION

Every live being has been living in acoustic environment. So many sounds may be met in daily life. Nevertheless each man has a different acoustic sensibility and response. Both sound and noise are all physical quantities [1,2]. The only difference between them is depended on a human psychological aspect. Sound is the same case. Each man has his favorite music or sound. One likes classic, but dislikes pop or rock music. The other likes pop. But there are few that dislike natural sound, for example, wind, rain, bird or sea waves sound. Natural sound gives happiness to every humans being and very pleasant to hear. On the other hand, noises are not. In this paper, the reason why so different between natural sounds and noises is verified by frequency domain approach. A new discriminative factor, “frequency-gradient rate” is introduced. With this parameter, it is to be classified that two type acoustic signals have their own distinctive signal characteristics.

FREQUENCY – GRADIENT RATE

A new analyzing parameter is introduced, named as “frequency-gradient rate”, gs. Parameter gs is defined as such.

\[ gs \leq \frac{\Delta M}{\Delta f} \quad \text{[ Pa/Hz]} \]  

where, \( \Delta f \) : frequency band (Hz)  
\( \Delta M \) : magnitude (Pa)

This parameter gs describes the frequency gradient. The high value means that frequency variance is high. The low value means a low frequency variance. That is, by this parameter the relation between frequency-gradient and psychological response is to be verified. This is the main purpose of this study. Each signal is analyzed into frequency domain. At first FFT algorithm is used. Frequency domain approach provides many signal information and qualities. By the frequency domain analysis, gs is calculated in every magnitude-variant interval. Frequency interval is to be selected according to gradient value. (+) or (-) value is not important, but only gradient rate is main factor. Rate and range are the analyzing factor.

SELECTION OF SOUND

Natural sounds and noises are selected as table 1.

Table 1. Natural sounds and noises

<table>
<thead>
<tr>
<th>Natural sounds</th>
<th>Noises</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cricket, Birds, Brook</td>
<td>Rock drill, Creaking</td>
</tr>
<tr>
<td>Wind, Cataract, Rain</td>
<td>Motor car, Muffler</td>
</tr>
<tr>
<td>Sea wares</td>
<td>Washing machine</td>
</tr>
<tr>
<td>Electric saw, Traffic</td>
<td>Electric saw, Traffic</td>
</tr>
</tbody>
</table>

Natural Sound

As shows in table 1, natural sounds time signal example is as bellow, fig.1. The FFT analysis is fig.2.

FIGURE 1. Natural Sound Signals
FIGURE 2. FFT Analysis

Noises

Noise time signal example is as fig.3, and fig.4 is FFT result.

FIGURE 3. Noises

FIGURE 4. FFT Results

ANALYSIS RESULTS BY “gs”

With FFT results, that is fig.2 and fig.4, “frequency – gradient rate, gs” are calculated. The analysis results are table 2,3.

Table 2. “gs” calculation: noises

<table>
<thead>
<tr>
<th>Freq. interval</th>
<th>gs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4, -0.26, -0.31, 1.63</td>
</tr>
<tr>
<td>2</td>
<td>-0.41, 0.3</td>
</tr>
<tr>
<td>3</td>
<td>-0.35, -0.65, 0.28</td>
</tr>
<tr>
<td>4</td>
<td>0.38, -0.3, 0.16</td>
</tr>
</tbody>
</table>

Table 3. “gs” calculation: natural sounds

<table>
<thead>
<tr>
<th>Freq. interval</th>
<th>gs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.27, 0.027</td>
</tr>
<tr>
<td>2</td>
<td>0.06, 0.02, 0.03</td>
</tr>
<tr>
<td>3</td>
<td>0.09, 0.4, 0.035</td>
</tr>
<tr>
<td>4</td>
<td>0.1, 0.2, 0.03</td>
</tr>
</tbody>
</table>

Discussion: In case of noises, “gs” is high, but natural sounds have a gentle slope. Frequency variations are slow in natural sounds.

CONCLUSION

Every physical signal has its own signal quality. The qualities are to be calculated with so many analysis algorithm. The “frequency-gradient rate, gs” is the one of these. This parameter shows the difference between noises and natural sounds. Natural sounds have a slow varying frequency variation but noises are very steep. Human response is sensitive. Humans hear the step frequency variant signal as noises, but feel comfortable and pleasant with slow-varying rate.

REFERENCES

Selective attention to sound events
by alternate appearance of subject-object voices:
ERP activity comparison

N. Sano a, T. Oshima b, T. Akita c and K. Hirate a

aDepartment of Architecture, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
bDepartment of Civil Engineering and Architecture, Faculty of Engineering, Niigata University, 8050 Igarashi-Ninocho, Niigata City 950-2181, Japan
cDepartment of Architecture, Faculty of Engineering, Tokyo Denki University, 2-2 Kanda-Nishikicho, Chiyoda-ku, Tokyo 101-8457, Japan

There is little information on how to find out the information causing selective attention to voices, which may form communication with subject in sound environment. Subjective communication sound events occur alternately with subjects speaking. Thus relation between subjective eventual time structure of speaking and event may be a factor of selective attention to find out specific communicative voice from many other voices in environment. We focused on alternate appearance of subject-object voices. Three 'a' voices were provided intermittently with 17 healthy students. Subjects were instructed to speak 'a' intermittently between 1s and 3s tempo hearing sound events. The target sound was one of the three voices and it was produced about 0.3s, 0.6s, 0.9s after subjects speaking vowel of 'a' by conditions. As the non-target sounds, random voices were produced simultaneously. No information was provided with subjects regarding the target voices. Multi-channel ERPs (event related potentials) were carried out. As the result, the target voices were most attracted subjectively among three kinds of voices. ERPs of the target voices indicate two changes: increase of negative CNV-like component at -200-0ms latency and N1 component suppression. These results suggest that alternate appearance of subject-object voices cause selective attention to communicative sound in environment.

1. INTRODUCTION

It has not been clarified that information processing system of selective attention to find out communicative voices and sounds with subject from the other voice-sound events in environment. There seems to be different processes of selective attention from passive listening. We focused on alternate appearance of subject-object voices as communication time structure. In the present paper, subjects' neuronal processes is investigated through an experiment using evoked-potential method, assuming the time structure between subject's utterance and occurrence of vocal events to be a clue to distinguish the communicative voice from other voices.

2. METHOD

Subjects were required to perform 'a' vowel in their self-paces but without counting time mentally. Through the 70 times of subject's vowel, three different 'a' voices were delivered from the loudspeakers in front of the subject. One of the three was produced at constant latency after subject's vowel, which is defined as the target sound. 0.3s, 0.6s, 0.9s interval target conditions were performed among conditions. Non-target voices were given at random intervals (1-5s). As the reference, non-target condition, no comparative random voice condition and no subject's vowel condition were performed. The orders of these conditions were arranged by Latin square method among subjects. Subjects were requested to keep their attention to the voices and they were not informed of any timing of occurrence voices. After each condition, subjects were ordered to evaluate most attracted voice among three 'a' voices. The 'a' voice stimuli duration were 100ms, and its' sound pressure levels were adjusted around 40dBA. EEGs were recorded from Fp1, Fp2, Fz, T3, C3, Cz, C4, T4 and Pz referred to linked earlobes (A1+A2). Triggers for averaging EEGs were set as the onset of voices and averaged for 1s (-200~800ms latency to trigger) at least over 30 epochs. The subjects were 17 right-handed healthy students for subjective evaluation and the 10 of those students were measured EEG.

3. RESULTS AND DISCUSSION

Figure 1 shows the ratio of evaluation of the most attracted 'a' voice per session. As the result of Chi-square test, the difference of the most attracted voice was observed among conditions (p<0.06). Target voices, which appear alternately with subjects' vowel
FIGURE 1. The most attracted voice among three 'a' voices by condition evaluated by subjects. After about 0.6s were most attracted. 0.3s and 0.9s alternate appearance voices are also slightly higher evaluation than the other random non-target voices. On the other hand, non-reaction condition which shows random voice listening has no significant difference among the ratio of three voices. The results show environmental sounds, which appear around 0.6s time interval after subjects speaking bring selective attention.

Figure 2 shows the comparison of grand average ERPs at Fp1 among target voices listening with and without other random voices by conditions. The bottom figure shows random voice ERPs with and without subject vowel. Significant N1 suppressions were observed at target voice ERPs on 0.3s, 0.6s, 0.9s conditions which appear alternately with other random voices. CNV-like negative component at frontal area between -200-0ms latency were also observed on target voice ERPs listening with other random voices on 0.3s and 0.6s conditions. These changes are only observed on listening voices, which appear alternately with random voices. Thus N1 suppression and CNV-like component occurrence may relate with information processing for making selective attention to communicative voice in environment.

Kikuchi and Kita (1994)\(^1\) shows that N1 suppression is observed at vocalization related potential. Selective attention to voices, which appear alternately, may have common information processing system with vocalization.

REFERENCES


FIGURE 2. Grand average ERPs on target voices at Fp1.