inter•noise 96
25th Anniversary Congress – Liverpool

INCE: 61

LOUDNESS AND SHARPNESS AS DETERMINANTS OF NOISE SIMILARITY AND PREFERENCE

B Berglund (1), P Hassmén (2) & A Preis (1 & 2)

(1) Institute of Environmental Medicine, Karolinska Institute, Stockholm & Dept of Psychology, Stockholm University, Sweden, (2) Institute of Acoustics, A Mickiewicz University, Poznan, Poland

1. INTRODUCTION

Although unclear to its definition, noise annoyance has become an accepted psychological concept. Much research has been devoted to the search for adequate physical properties of complex sounds that may explain why some sounds evoke a feeling of annoyance. Accumulated research results seem to agree to that (perceived) loudness, and thus also its correlate sound pressure, is the main feature of environmental sounds that determines annoyance [2]. However, as advocated by Zwicker [10] and Preis [6], other properties of sound as well as psychological factors [2,3] have to be taken into consideration. We contend that properties of sound related to character and quality must play a primary role in evoking feelings of annoyance in real-life situations simply because they determine what sound is perceived. Thus, a pertinent issue in noise measurement must be to prove that it is sound to collapse the qualitative properties into a unidimensional scale of sound pressure, as is done, for example, in various frequency weighting networks, Leq, or Zwicker's loudness.

For the benefit of a good noise abatement policy, especially with regard to noise annoyance predictions, it may be wise to link noise measurement more firmly to the perceptual aspects of *complex sounds* in environmental settings [6]. After all, an improvement of current noise measurement and annoyance reduction techniques may require that other properties of sound be abated than loudness correlates, that is, properties related to time or frequency domains.

The present research approaches the annoyance problem indirectly and from a different angle than before, in the hope to reveal key principles involved in noise perception. The question posed is to what extent psychological attributes (loudness, annoyance [2]) and/or physical properties (Zwicker's loudness [9], Aures's sharpness [1]) may explain how similar sounds are perceived or what sound is preferred in a listening situation. With the aid of psychoacoustical experiments it was thus asked to what extent the structure of similarity data matched the structure of preference data, and to what extent the dimensions extracted to describe these structures are associated with perceptual and/or acoustic scales.

A major difference between the concepts of similarity and preference is their primarily cognitive and affective nature, respectively. The two underlying psychological processes are believed to be independent [8]. Typically only a minor part of the judgmental variance is common in similarity and preference data for the same set of stimuli. Therefore, it seems plausible to expect that perceived loudness and Zwicker's loudness should be associated with the similarities of sounds (cognitive aspects), and annoyance with sound preferences (emotional aspects). If Aures's sharpness is related to the intrusiveness of sounds [5], it might influence preferences more than similarities. In order to obtain a judgmental focus on perceived quality of environmental sounds, all stimuli were matched with regard to sound pressure level (SPL) such that the variation in perceived loudness would be restricted allowing other potential variables to express themselves.

2. METHOD

Stimuli and Apparatus. Fifteen complex sounds (11 community noises & 4 model spectra), all of 4-s duration were combined in 210 unique pairs (identity pairs excluded, reversed pairs included). The sounds were: A–passing car, B–office printer, C–ventilation fan, D–alarm clock, E–passing subway, F–departing subway, G–leaf blower, H–lunch restaurant, I–road traffic, J–food mixer, K–coffee maker; L, M, N, & O are different model spectra created from white noise.

The 210 pairs of sounds were recorded in 3 random orders on 3 corresponding experimental tapes in which a sound pair covered a duration cycle of 15 s (sound 4 s, pause 1 s, sound 4 s, pause 6 s). The tapes were played back by a Fostex DAT tape recorder using a Pioneer (A-878) amplifier which fed two Cerwin-Vega loudspeakers (PD-3). All sounds were presented at 65 dBA SPL to a group of subjects, seated in a large conference room (8 x 8 x 2.75 m; reverberation time 0.9 s).

Subjects. Twenty-five university students (10 men, 15 women) participated as volunteers in both the similarity and the preference experiments. Their mean age was 27.2 (*sd*=7.2) years and all were screened for normal hearing with an audiometric test before the experiment.

Similarity. The perceived similarity of sound pairs was reported in percent, on a scale from 0 to 100, i.e., from complete dissimilarity (0% similar) to complete similarity (100% similar). The instruction was to judge the degree of "overall similarity". In all, each of the 25 subjects made 630 similarity judgments (210 x 3).

Preference judgments. Preference, or rather its opposite, nonpreference, was obtained by asking the participants to mark which of the two sounds presented in a pair they would have *preferred to switch off*, if given the possibility right now. The more frequently the particular sound was wished to be switched off, the higher the nonpreference percent. In all, each of the 25 subjects made 630 nonpreference choices (210 x 3).

Procedure. A balanced design was used with regard to the order of the similarity and preference experiments which also were conducted in separate days. The participants responded individually by giving each similarity (or nonpreference) response in written form (a 21-page booklet was used which allowed only 10 responses per page; 6 booklets were needed per subject). The 3 experimental tapes were presented in irregular order to the groups. Three 17-min sessions (70 pairs) with a 5-min break was needed for one tape. A 10-min break was taken between tapes. During breaks, subjects were allowed to move around freely in the room, but not to communicate with each other.

3. RESULTS AND DISCUSSION

The test-retest reliability of individual scales was good both for similarity (mean Pearson coefficient of correlation r=0.70, range: 0.49-0.85, n=25) and nonpreference (r=0.65, range 0.40-0.91, n=25). The concordance within the similarity and nonpreference sets of individual scales (based on 6 judgments per sound pair) was determined with the aid of a principal components analysis of the intercorrelations between scales (defined by 105 combinations of sounds). All except one subject's preference scale agreed with a classification of the subjects in two groups (eigenvalues >1 was used as criterion for number of extracted factors), discerned in plots of all possible pairs of the two extracted factors for similarity (77% of the variance explained) and nonpreference (70% of the variance explained).The deviant subject was kept in his "best" fitting group. Group 1 and 2 consisted of 19 and 6 subjects, respectively, who in the further data analyses were kept apart because of their seemingly characteristically different scales.

The two similarity as well as two nonpreference matrices were analyzed individually with multidimensional scaling [4,7] using the Euclidian distance model. From Kruskal's stress values it was appropriate to represent each of the four solutions in two dimensions only. Unexpectedly, the solutions for the two groups did not differ very much, a larger difference could be seen in the nonpreference than the similarity data.

In the four MDS solutions, clusters of sounds may be discerned, although distinctly only for the similarity data ("traffic", "motors of equipments", "alarm clock", "model spectra"). These clusters seems to have some reference to the frequency spectra of the sounds, for example, the alarm clock (D) gives a close to typical harmonic complex whereas the road traffic (I) gives relatively few harmonic components (criterion: >3 dB SPL change) and the model spectra (L-O) created from white noise lack harmonic components.

Table 1. Results of principal components analysis of a matrix of pairwise correlations of 12 unidimensional scales characterizing 15 complex sounds (squared loadings showing common variance with ideal factor).

Scale	Factor 1	Factor 2	Factor 3
Similarity, Dim 1, Group 1	*0.89	0.00	0.00
Similarity, Dim 1, Group 2	*0.80	0.05	0.02
Nonpreference, Dim 1, Group 1	*0.65	0.15	0.05
Nonpreference, Dim 1, Group 2	*0.75	0.00	0.00
Perceived annoyance	*0.93	0.00	0.03
Perceived loudness	*0.81	0.06	0.04
Zwicker's loudness (sone)	*0.40	(*)0.39	0.08
Similarity, Dim 2, Group 1	0.00	*0.89	0.00
Similarity, Dim 2, Group 2	0.03	*0.76	0.06
Aures's sharpness (acum)	(*)0.32	*0.48	0.06
Nonpreference, Dim 2, Group 1	0.14	0.08	*0.74
Nonpreference, Dim 2, Group 2	0.00	0.07	*0.88

**Factor has high common variance for scale.*

In order to efficiently compare the four MDS solutions as well as to try to interpret the differences exhibited in the similarity and nonpreference structures, their dimensions were entered as unidimensional scales in a principal components analysis together with four external unidimensional scales of the same 15 sounds. The latter refer to (a) two acoustic scales, Zwicker's loudness [9] and Aures's sharpness [1], and (b) two perceptual scales, loudness and annoyance obtained by free-magnitude estimation in an independent group of 9 subjects. Kaiser's criterion (eigenvalues >1) was used and three factors were extracted after varimax rotation, explaining most of the variance of the intercorrelations of scales (≈90% accumulated from 48, 25 & 17%), see Table 1 for these results.

Obviously, the first factor is strongly associated with perceived annoyance, as scaled unidimensionally, and thus also with perceived loudness (the two scales share 81% in common variance). Consequently, the first dimension of the MDS solutions for similarities as well as nonpreferences for both subject groups could be named "loudness/ annoyance". Also Zwicker's loudness belongs here due to its underlying association with perceived loudness (67%) and perceived annoyance (36%).

The second factor is best characterized by the second dimension of the MDS solutions for similarity in both subject groups; and the third factor by the second dimension of the MDS solutions for nonpreferences in both subject groups. These results clearly show that the perceptual aspects considered in similarity are different from those in nonpreferences, but, as shown for the first factor, both are also strongly influenced by loudness/annoyance. Although Aures's sharpness shows its highest association with the second factor, it is week and only somewhat higher than for the first factor. Interestingly, both perceived loudness and annoyance lack correlation with the second and third factor.

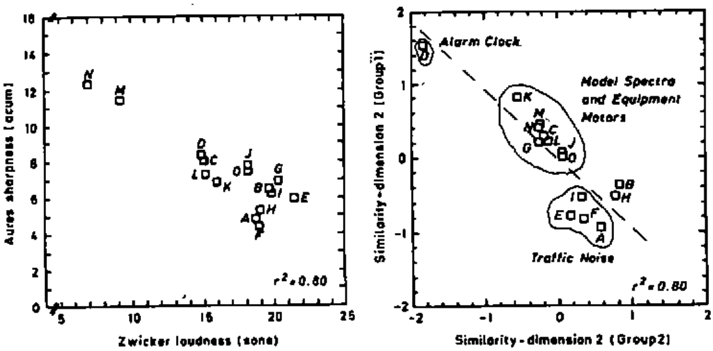

Fig. 1. Illustration of possible influences on similarity of sounds of acoustic properties reflected in Zwicker's loudness and Aures's sharpness. [For symbols see the methods section.]

For the investigated set of 15 complex sounds, Aures's sharpness is inversely related to Zwicker's loudness and the two scales share as much as 80% in common variance, Fig. 1 (left). This is not surprising because Aures's sharpness is partly based on Zwicker's loudness but puts, i.a., a larger weight on the critical bandwidths above 16 bark. The inverse scale relationship do not differentiate the complex sounds easily and understandably. In comparison, the second dimension of similarity for the two groups show a similar relationship (Fig. 1, right), but here clusters of sounds could be understood, the alarm clock and the traffic noises clearly represent opposites.

Some specific inherent characteristics of Aures's sharpness may be relevant for similarities of sound (Dimension 1 & 2) but not to the same extent for nonpreferences. For both groups (third factor), dimension 2 for nonpreferences lacks association with Aures's sharpness (and also Zwicker's loudness). Obviously, important aspects of the sounds are included for nonpreferences but left out for similarities. Since preference judgments are afflicted with emotions [8], the MDS solution of nonpreferences may have singled out the affective part of the subjects' choices in Dimension 2 (Factor 3) from the predominating perceptual part in Dimension 1 (Factor 1). For similarity, it seems simply as if the qualitative aspects of the sounds were differentiated cognitively (without affect) in Dimension 2 (Factor 2), and separated from the loudness aspects in Dimension 1 (Factor 1).

According to a unidimensional nonpreference scale constructed as the percent of all comparisons among which a particular sound was nonpreferred, Group 1 and Group 2 agreed reasonably well (67% common variance). At 65 dBA SPL, the most nonpreferred sounds were in rank order: passing car, ventilation fan, office printer, alarm clock, departing subway, and leaf blower. The two groups did not differ substantially in their MDS solutions of similarities and nonpreferences as reflected in the dimensional structures (Table 1), therefore earlier identified differences must be referred to minor shifts in judgment for certain combinations only.

4. EPILOGUE

Any single extracted factor, common for the dimensions defining the similarity and nonpreference structures, was much less associated with Zwicker's loudness and Aures's sharpness (32-48% common variance) than the independently obtained perceived loudness and perceived annoyance scales (81-93% common variance). This would implicate that interindividual differences in perceptual scales do not make them particularly invalid. Rather perceptual scales by definition refer to psychological constructs that, unfortunately, are difficult to characterize acoustically. A further implication would be that Zwicker's loudness and Aures's sharpness incorporate, into a single indicator variable, correlates of incompatible perceptual aspects of complex sounds. By principles of logic such constructions would be uncertain, crude and potentially invalid. While waiting for improvements of these or other acoustic scales, an alternative approach would be to utilize perceptual scales directly. This could be less costly for society in a longterm perspective. A perceptual approach would be valid, and particularly efficient because today's technology makes it practically possible to build a data base of real-life environmental recordings that keep well and that may be researched repeatedly.

Acknowledgements
This research was supported by grants from the Swedish Environmental Protection Agency and the Swedish Council for Building Research. We are grateful for the research assistance of T Beck Kristensen and Mats E Nilsson. Dr. Piotr Miecznik kindly assisted in the production of the tape recordings and the acoustical measurements.

References
[1] W. Aures, W. Acust., 59, 130-141 (1985).
[2] B. Berglund, & T. Lindvall, Arch. Cent. Sens. Res, 2(1), 1-195 (1995).
[3] R.F.S. Job, Environ. Int., 22, 93-104 (1996).
[4] J.B. Kruskal, Psychomet., 29, 115-129 (1964).
[5] A. Preis, Appl. Acoust., 20, 101-127 (1987).
[6] A. Preis, Arch. Cent. Sens. Res., 2(2), 1-57 (1995).
[7] R.N. Shepard, Psychomet., 27, 125-139 & 219-245 (1962).
[8] R.B. Zajonc, Am. Psycholog., 35, 151-175 (1980).
[9] E. Zwicker, Acust., 10, 304-308 (1960).
[10] E. Zwicker, in A. Schick, J. Hellbrück & E. Weber (eds.), Contributins to Psychological Acoustics V (Bibliotheks- und Informationssystem der Universität Oldenburg, Oldenburg, FRG, 1990).