24th INTERNATIONAL CONGRESS ON SOUND AND VIBRATION
23–27 July 2017, London

AN EVIDENCE-BASED SOUNDSCAPE TAXONOMY

Oliver Bones, Trevor J. Cox and William J. Davies

University of Salford, Acoustics Research Centre, School of Computing, Science and Engineering, Salford, M5 4WT, UK
email: o.c.bones@salford.ac.uk

In an attempt to cultivate standardization in soundscape reporting Brown, Kang and Gjestland [1] offered an influential schema by which the acoustic environment is divided initially into indoor and outdoor environments, and within each into further categories; urban, rural, wilderness, and underwater. Within each of these, sounds are categorised with increasing levels of detail. However, this schema is offered as an organizational framework for further elaboration, rather than as an evidence-based account of semantic categories. In other soundscape categorisation research semantic differential data are used to examine how perceptually similar sounds are. However, this approach typically involves prescribing descriptive terms, taken from previous research or simply prescribed by the researcher, rather than eliciting descriptive terms and semantic categories *per se*. Another common method for identifying semantic categories is to perform multidimensional scaling and cluster analysis of similarity data generated by a pairwise comparison task. This approach avoids prescribing attributes with which to rate sounds; however the absence of semantic labelling in the task means that interpretation of the data is necessarily subjective, and the amount of time required to perform pairwise comparisons on a large number of sounds is potentially prohibitive. Here we present an evidence-based account of semantic categories of sounds commonly described in the soundscape literature, using a robust method based upon perceptual data generated by an online sorting and category-labelling task. This method is relatively quick to perform and elicits rather than prescribes categories and descriptive terms.

Keywords: taxonomy, soundscape, perception, methodology

1. Introduction

There has been much recent interest in developing an organising account of the multitude of quotidian sounds experienced in everyday life. In part this is driven by the need to standardize reporting in the relatively new and multidisciplinary area of soundscape research. To this end Brown, Kang & Gjestland [1] provide an influential initial schema for categorising sounds by spatial context and sound source. However, from the perspective of sound perception research it is of interest to understand whether these categories have ecological validity based upon experimental data. The current paper presents preliminary data from a study of semantic categories of sounds commonly reported in the soundscape literature. Whilst a comprehensive review of existing research does not fall within the scope of the current paper (a more comprehensive review can be found elsewhere [2]) , the motivation for the approach taken is briefly discussed with respect to the methodology commonly used in this area.

A common approach to identifying semantic sound categories is to test how similar different sounds are considered to be to one another. The pairwise comparison method [3] produces a similarity matrix which can be further analysed using multidimensional scaling (MDS). MDS identifies the dimensions of the similarity data which account for the most variance, and the proximity of sounds within the resulting space can be made explicit via cluster-analysis. One drawback to the pairwise comparison method, in which each sound is compared and rated for similarity to every other sound, is that it is time-intensive. Moreover, whilst cluster-analysis of the dimensions resulting from MDS provides insight into which sounds are perceived as being similar to one another, the interpretation of what the clusters represent is necessarily subjective, in that it is left to the researcher to label each cluster.

The semantic differential technique [4] involves rating stimuli using attribute scales, the semantic descriptors for which are typically determined *a priori*, based upon previous research [5-8] or the researcher's assumptions [9]. Semantic differential ratings are typically analysed using principle component analysis (PCA) in order to identify the underlying structure. Whilst this approach goes some way towards addressing the issue of subjective researcher interpretation of the resulting categories since the semantic meaning of clusters of sounds can be inferred from the ratings of sounds within the cluster on each of the attribute scales, a drawback is the prescriptive nature of predefining the scales with which to rate each sound.

Qualitative approaches such as conducting interviews [10-12], verbal descriptions [13], and linguistic analysis [14-17] avoid this pitfall by allowing participants to generate their own descriptive terms. Similarly, methods involving free-sorting of sounds by similarity into categories which the participant then labels [16, 18, 19] is advantageous with respect to ecological validity. An extension of the method described by Kawaii et al. [19] is employed here, whereby correspondence analysis of a contingency table of data resulting from a free-sorting task is used to extract dimensions for cluster-analysis. The ecological validity of the approach lies in the use of the descriptive words generated by the sorting task in the objective interpretation of the resulting categories. This approach was employed previously to investigate categories of frequently used search terms on the Freesound.org audio database [20]. Here we present preliminary data using the same method to investigate sounds frequently reported in the soundscape literature.

2. Methods and analysis

2.1 Stimuli and participants

Stimuli were five second clips taken from Freesound.org, equalised in RMS, and selected so as to be representative of sounds described in a number of studies from the soundscape literature [1, 3, 19, 21, 22]. In previous taxonomies an emphasis is placed on sounds occurring in multiple contexts [1]. Therefore an effort was made to include examples of sounds recorded indoors and outdoors where this was possible. In some cases these were recordings of sounds occurring outside, recorded from indoors e.g. 'Fireworks_2'. In other cases these were recordings which were audibly recorded in different sized spaces e.g. 'Laughter_1' sounds like it was recorded in a large room due to the audible reverberations, whereas 'Laughter_2' does not contain audible reverberations.

The study was run online, and data from 20 participants is presented here. Future work using this method will include a larger number of participants; however, the emergent categories with 20 participants are consistent with categories resulting from a previous study of frequently used audio search terms [20].

2.2 Card-sorting

At the onset of the experiment all sounds, each of which was represented by a tile containing a single word descriptor (e.g. Road_1), were arranged in a random order in a 'Sound Bank' panel on the left hand side of the screen. Instructions at the top of the screen directed participants to:

- group similar sounds together by dragging them from the Sound Bank into the five groups;
- use all five groups;
- give each group a label to describe the group;
- click the words once to hear the sound.

No time limit was imposed. The average time taken was approximately 20 minutes.

2.3 Contingency table

Following the methodology described by Kawai et al. [19] the data from each participant were initially collected as a 60 (sounds) x 5 (category names) contingency table of 1s and 0s, where a 1 indicated that a sound corresponded to a given category. At the end of the study data from all 20 participants were collated into a 60 x 100 contingency table of 1s and 0s. In an extension to this approach, data from the contingency table was pooled column-wise in order to combine categories that were the same and to increase the power of the analysis. The 100 category names were initially processed by:

- removing white space;
- removing special characters;
- removing the words ‘sound’ and ‘sounds’;
- removing numbers;
- converting to lower-case;
- correcting spelling.

Category names were then stemmed (e.g. ‘natural’ and ‘nature’ were reduced to ‘natur-’) before restoring each stem to the most common pre-stemming version of that word (e.g. ‘natural’). Categories which had either the same name following this process, or which were identified as synonyms by Microsoft’s synonym checker were then pooled. This process resulted in a contingency table with 70 categories, and which contained numbers other than 1s and 0s. Hereafter category names are referred to as ‘descriptive words’.

2.4 Correspondence analysis and hierarchical cluster analysis

Correspondence analysis (CA), a method similar to PCA but suitable for categorical rather than continuous data, was used to identify the principle dimensions of the data and to visualise row points (sounds) and column points (categories) in a low-dimensional space. A number of dimensions sufficient to retain 50% of the variance was selected. Agglomerative hierarchical cluster analysis was performed on the dimensions resulting from CA using Ward’s criterion. Clusters were interpreted by the descriptive words that contributed to each cluster. The contribution of each descriptive word to each cluster was assessed by comparing global frequency (the sum of a descriptive word’s column in the contingency table i.e. the total number of times a sound was assigned to the descriptive word) to the internal frequency for a given cluster (the number of times sounds within a cluster were assigned to that descriptive word). Significance of over- or under-representation of each descriptive word within each cluster was assessed using the hypergeometric distribution.

3. Results

The variance retained by the first five dimensions of the CA is displayed in Table 1. Although the variance retained is relatively low (38.62% cumulatively for the first three dimensions) compared to that reported from PCA of semantic differential data [5-7], it is comparable to reported results using CA [19]. The correlation coefficient between rows and columns, calculated as the square root of the sum of eigenvalues, was high (1.82), indicating a strong association between sounds and categories. Similarly, the chi-square statistic (4660.04) was found to be significant (p=0.005). Figure 1 displays sounds plotted on the first two dimensions.

Table 1: Eigenvalues and variance retained by the first five dimensions

	Dim. 1	Dim. 2	Dim. 3	Dim. 4	Dim. 5
Eigenvalue	0.48	0.46	0.35	0.27	0.19
% of Variance	14.42	13.79	10.42	8.02	5.77
Cumulative %	14.42	28.2	38.62	46.64	52.42

The results from the hierarchical cluster analysis are displayed in Figure 1, and also as a dendrogram in Figure 2. Five descriptive words were significantly over-represented in Cluster 1, ten in Cluster 2, five in Cluster 3, eight in Cluster 4, and 13 in Cluster 5. In the interests of space, only the four most significantly associated descriptive words for each cluster are presented here in Table 2, in descending order of internal frequency relative to global frequency, and therefore in ascending order of *p*-value. All *p*-values were <0.001 and are therefore not presented.

The descriptive words from Table 2 with the largest internal frequency relative to global frequency for each cluster were taken as the cluster name (see Fig. 2), except in the case of Cluster 5 where ‘Mechanical’ was deemed more descriptive than ‘Constant’.

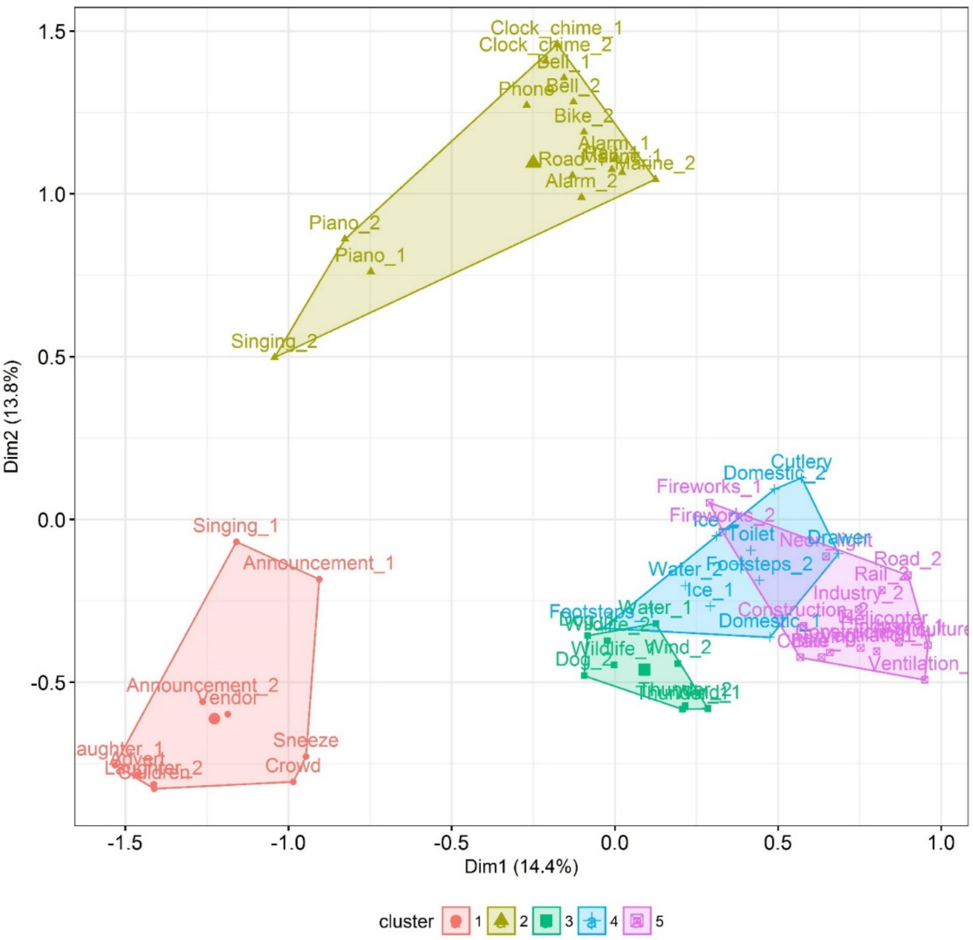

Figure 1: Sounds arranged on the first two dimensions

Table 2: Descriptive words associated with each cluster. Category names are indicated by an asterisk.

	Descriptive Word	Internal Freq. (%)	Global Freq. (%)
Cluster 1	Human*	23	4.67
	People	10.5	1.83
	Vocal	7.5	1.42
	Living	4	0.833
Cluster 2	Music*	19.33	5.67
	Tones	5	1.25
	Alerts	4	1
	Rushing	3.67	1.08
Cluster 3	Natural*	30	7.08
	Scary	3.33	0.67
	Sudden	3.33	0.75
	Weather	1.67	0.25
Cluster 4	Objects*	5	1
	Short	8	2.67
	Random	5.5	1.42
	Inside	8	2.83
Cluster 5	Constant	7.5	2.75
	Mechanical*	6.56	2.25
	Industrial	3.75	1.25
	Machinery	2.5	0.67

4. Discussion

Previous work with a sample size of 101 investigating the semantic categories of frequently used audio search terms using the method described here produced the categories: Nature; Music; Urban; Human; and Effects. In the present study with a relatively small sample size of 20, similar categories were generated: Mechanical; Objects; Natural; Human; and Music. The sound-source categories proposed by Brown, Kang & Gjestland [1] are: Motorized Transport; Human Movement; Electro-Mechanical; Voices and Instruments (amplified / non-amplified); Social/Communal; Other Human; Nature; and Domesticated Animals.

The results here are consistent with the previous suggestion [16] that acoustic parameters do not necessarily discriminate categorisation at this level e.g. dog sounds and water sounds are categorised as Natural, despite their obvious acoustic differences. On the other hand, the categorisation of tonal sounds such as the car horn of 'Road_1' and the sonar of 'Marine_1' as Music—rather than Motorized transport [1]—suggests the physical parameters of the sound did influence categorisation in these cases. It has been suggested previously that this is likely to occur when source identification is ambiguous [16], although in the present study this is unlikely to have been the case since text descriptors of each sound were provided. The majority of sounds from the Electro-mechanical [1] category are categorised here as Mechanical, with the exception of some domestic sounds, such as 'Toilet' and 'Cutlery', which are categorised as Objects. It is interesting to note that both 'Footsteps' sounds were also categorised as Objects, rather than as Human. This is presumably because they were categorised

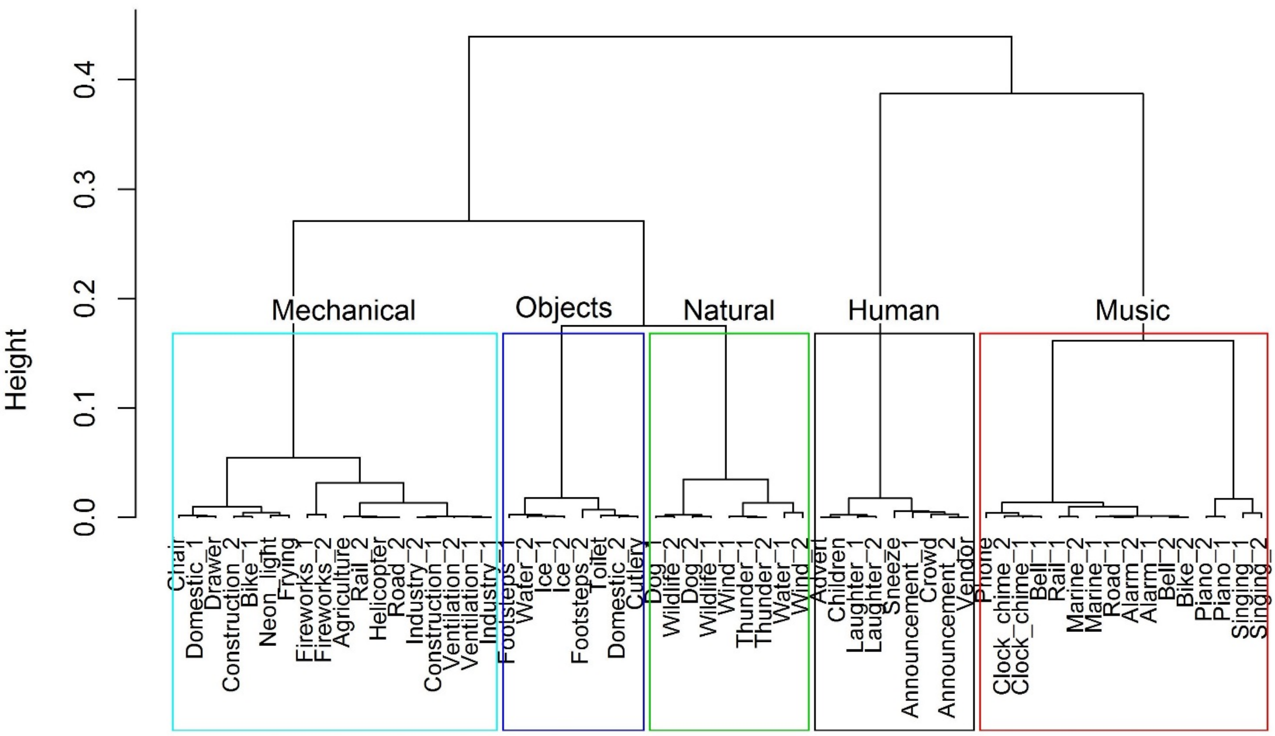

Figure 2: Sounds arranged in a cluster dendrogram

based on the sound of the foot's impact against a surface, rather than as being human sounds *per se*. In general, the results are consistent with multiple categorisation strategies being deployed; primarily based on semantic information, but with inferred source attributes (e.g. 'Footsteps') and acoustic attributes (e.g. 'Road_1') also being used.

Sounds from the Voices and Instruments categories [1] are distributed here between Human and Music, with 'Singing' judged to be a Music sound rather than Human, whilst the Social/Communal category [1] is distributed between Music ('Bell', 'Clock-chime', 'Alarm') and Mechanical ('Fireworks').

It should be noted that experimental approaches such as that employed in the present study in which sounds are presented in isolation and without context do not account for the role that an individual's interaction with and activity within a soundscape may have in their assessment of individual sounds [23]. However, by interpreting participant generated sound categories using participant generated descriptors the present approach represents an effort towards objectively eliciting categories of sound, albeit in an artificial context.

REFERENCES

[1] A.L. Brown, J. Kang, T. Gjestland, Towards standardization in soundscape preference assessment, Applied Acoustics, 72 (2011) 387-392.

[2] S.R. Payne, W.J. Davies, M.D. Adams, Research into the practical policy applications of soundscapes concepts and techniques in urban areas, DEFRA report NANR200, in, 2009.

[3] B. Gygi, G.R. Kidd, C.S. Watson, Similarity and categorization of environmental sounds, Perception & Psychophysics, 69 (2007) 839-855.

[4] C.E. Osgood, The nature and measurement of meaning, Psychological bulletin, 49 (1952).

[5] R. Cain, P. Jennings, J. Poxon, The development and application of the emotional dimensions of a soundscape, Applied Acoustics, 74 (2013) 232-239.

[6] B. Yu, J. Kang, H. Ma, Development of indicators for the soundscape in urban shopping streets, Acta Acustica united with Acustica, 102 (2016) 462-473.
[7] O. Axelsson, M.E. Nilsson, B. Berglund, A principal components model of soundscape perception, J Acoust Soc Am, 128 (2010) 2836-2846.
[8] J. Ge, K. Hokao, Applying the methods of image evaluation and spatial analysis to study the sound environment of urban street areas, Journal of Environmental Psychology, 25 (2005) 455-466.
[9] J.Y. Hong, J.Y. Jeon, Influence of urban contexts on soundscape perceptions: A structural equation modeling approach, Landscape and Urban Planning, 141 (2015) 78-87.
[10] B. Schulte-Fortkamp, A. Fiebig, Soundscape Analysis in a Residential Area: An Evaluation of Noise and People's Mind, Acta Acustica united with Acustica, 92 (2006) 875-880.
[11] K. Foale, W. Davies, A listener-centred approach to soundscape evaluation, in: Acoustics, Nantes, 2012.
[12] N.S. Bruce, W.J. Davies, The effects of expectation on the perception of soundscapes, Applied acoustics, 85 (2014) 1-11.
[13] G. Oleksik, D. Frohlich, L.M. Brown, A. Sellen, Sonic interventions: Understanding and extending the domestic landscape, in: Computer Human Interaction 2008, 2008.
[14] C. Guastavino, The ideal urban soundscape: investigating the sound quality of French cities, Acta Acustica united with Acustica, 92 (2006) 945-951.
[15] M. Raimbault, D. Dubois, Urban soundscapes: Experiences and knowledge, Cities, 22 (2005) 339-350.
[16] D. Dubois, C. Guastavino, M. Raimbault, A cognitive approach to urban soundscapes: Using verbal data to access everyday life auditory categories. , Acta acustica united with acustica, 92 (2006) 865-874.
[17] M. Neiessen, C. Cance, D. Dubois, Categories for soundscape: toward a hybrid classification, in: Internoise, Lisbon, 2010.
[18] S.R. Payne, P. Devine-Wright, K.N. Irvine, People's perceptions and classifications of sounds heard in urban parks: semantics, affect and restoration, in: Inter-Noise, Proceedings of Inter-Noise 2007, Istanbul, Turkey, 2007.
[19] K. Kawai, T. Kojima, K. Hirate, M. Yasuoka, Personal evaluation structure of environmental sounds: experiments of subjective evaluation using subjects' own terms, Journal of Sound and Vibration, 277 (2004) 523-533.
[20] O.C. Bones, T.J. Cox, W.J. Davies, Toward an evidence-based taxonomy of everyday sounds, The Journal of the Acoustical Society of America, 140 (2016) 3266-3266.
[21] M. Yang, J. Kang, Psychoacoustical evaluation of natural and urban sounds in soundscapes, J Acoust Soc Am, 134 (2013) 840-851.
[22] J. Salamon, C. Jacoby, J.P. Bello, A Dataset and Taxonomy for Urban Sound Research, (2014) 1041-1044.
[23] R. Cain, P. Jennings, M. Adams, N. Bruce, A. Carlyle, P. Cusack, W. Davies, K. Hume, C. Plack, An activity-centric conceptual framework for assessing and creating positive urban soundscapes, in: Institute of Acoustics Spring Conference 2008, 2008.