Applied Acoustics 74 (2013) 248–254

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

An exploratory evaluation of perceptual, psychoacoustic and acoustical properties of urban soundscapes

Deborah A. Hall [a,b,*,1], Amy Irwin [c,1], Mark Edmondson-Jones [c], Scott Phillips [a], John E.W. Poxon [d]

[a] Division of Psychology, School of Social Sciences, Nottingham Trent University, Burton Street, Nottingham NG1 4BU, UK
[b] NIHR National Biomedical Research Unit in Hearing, Ropewalk House, 113 The Ropewalk, Nottingham NG1 4DU, UK
[c] MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, UK
[d] WMG, International Digital Laboratory, University of Warwick, Coventry, West Midlands CV4 7AL, UK

ARTICLE INFO

Article history:
Available online 7 May 2011

Keywords:
Well-being
Sound quality
Affective responses

ABSTRACT

The present study reports an exploration of the multi-dimensional space involved in listening to soundscape recordings made in different city-based settings. A range of perceptual, psychoacoustic and acoustical properties were examined using a range of statistical methods including principal components analysis and multiple regression. Just as the affective responses to individual sound sources have often been described in terms of pleasantness and arousal (or vibrancy), so could the acoustic scenarios of urban life. However, different from previous research, the acoustic and psychoacoustic variables contributed very little to judgements about pleasantness and vibrancy. We surmise that the perceived quality of the soundscape is very much an individual subjective experience that is likely to be embedded within contextual domain and that it probably relates to personal preference, past history and other social and cultural factors.

1. Introduction

A soundscape can be construed as the acoustic equivalent of a visual landscape and the concept of the soundscape is generally credited to Schafer in his study of auditory ecology [25]. His approach to characterising environmental sound attempts to capture the rich complexity of those variables contributing to the experience. It should give weight to the positive as well as the negative attributes of, what has been termed by others, 'environmental noise'. This broad philosophy has spread beyond the field of auditory ecology to acoustics (e.g. [17]), health [11] and architectural development (e.g. [23,26]). At the intersection of health and architecture is a shared concern about the social well-being of those who live and work in towns and cities. However, professionals lack reliable tools for measuring sound quality and so lack the necessary guidance about how to best improve environmental quality [23]. While urban planners might make a concerted effort to include human evaluations of environmental quality, subjective evaluations are not mandatory by law, nor are they systematic. A simple measurement of sound level (A-weighted sound pressure level, SPL) remains the most common measure of sound quality or annoyance. For example, noise programs in the European Union are setting up strategic noise maps through estimation of A-weighted SPL measures, even though it is difficult to interpret the associated impact without further information about the nature of the sound source and its context.

It is widely appreciated that simply reducing the sound level of an urban space does not necessarily increase a listener's degree of 'acoustic comfort' [29] and over the past few years, researchers have begun systematic studies to define the key factors that influence the individual experience of a soundscape. This approach acknowledges the multi-factorial nature of the experience including the characteristics of the dominant sound source, the meaning of the sound as interpreted by the listener and the context in which the sound is heard [23,1,6]. The main objectives have been to identify the acoustic, perceptual and affective properties that contribute to the soundscape experience and to develop measurement scales that adequately describe the soundscape experience. Acoustic science has so far done a good job of quantifying the physical properties of the signal in terms of attributes such as its frequency spectrum, temporal envelope, directivity and amplitude (e.g. [12,13]. A number of psychoacoustic properties, such as loudness and sharpness, can also be quantified using objective measurement criteria [30,19,24]. Affective reactions to sound stimuli have mostly been quantified using visual analogue scales on which people rate their responses along a number of semantic dimensions

* Corresponding author at: Division of Psychology, School of Social Sciences, Nottingham Trent University, Burton Street, Nottingham NG1 4BU, UK. Tel.: +44 115 8482377, fax: +44 115 8486829.
E-mail address: deborah.hall@ntu.ac.uk (D.A. Hall).
[1] These authors contributed equally to this research.

0003-682X/$ - see front matter Crown Copyright © 2011 Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.apacoust.2011.03.006

[3,8,22,29]. Results consistently indicate that pleasantness and arousal best describe the range of emotional experiences [21,3,27] see also [4] for a review). However, although these factors explain most of the variance in responses, some authors suggest that the complexity of emotion cannot be entirely explained by pleasantness and arousal alone [27].

Perhaps unsurprisingly, results confirm that natural sound sources such as human speech and animal calls are generally preferred over traffic and machinery noise [29,9,20]. For example, [29] noted that when the dominant sound source was judged to be pleasant (e.g. music or running water), the level of listening comfort was rated far better than when the dominant sound source was judged to be unpleasant. Moreover, this relationship was true even when the overall sound level was high. Such a modulatory effect highlights the important contribution to overall quality of the individual's perception and understanding of the sound, rather than simply relying on a quantitative analysis of the sound's acoustical properties.

The present paper reports one arm of a multi-disciplinary project that has focused on urban soundscapes; those every-day sound experiences perceived by city dwellers. One of the main goals of the project has been concerned with evaluating the relationship between the auditory environment and the responses of the people living within it. Methods have included soundwalks, focus groups and laboratory-based evaluations of soundscapes, including semantic ratings of soundscape clips (see the paper by Cain et al. [7]). The current paper attempts to characterise aspects of the multi-dimensional space involved in soundscape listening and is a companion to a related paper in which we report the physiological responses to the same set of soundscape clips [14]. Here, we report an exploratory evaluation of some of the perceptual and spectro-temporal properties of urban soundscapes measured using visual analogue scales of semantic descriptors and physical analyses of the signal.

2. Methods

2.1. Selecting the urban soundscape materials

A set of recordings made in British urban soundscapes were acquired from a variety of archival sources, including a sound artist with an interest in acoustic ecology (Peter Cusack, personal communication), the British Library (www.sounds.bl.uk/Browse.aspx?collection=Soundscapes), the International Affective Digital Sounds (IADS) (www.csea.phhp.ufl.edu/media/iadsmessage.html), a website managed by Trevor Cox (www.sound101.org/badvibes) and a website containing people's favourite London sounds (www.favouritelondonsounds.org). Selected clips typically contained more than one sound source, often with urban street noise as a background ambient context and a range of different everyday contexts such as the street, market, shopping mall, and park were represented. A small number of clips represented a single sound source and these were digitally mixed with a congruous background noise to provide a suitable ambient context. In each clip, the sound signatures tended to repeat so that an initial evaluation about the scene could be made quite quickly and would most likely represent the same judgement that might be made over a longer listening period. Across the stimulus set, the dominant sound source fell into one of three broad categories; transportation or works (referred to as 'mechanical' sources, $N = 86$), human activity (referred to as 'human' sources, $N = 83$) and nature ('natural' sources, $N = 50$) (c.f. [23]). All recordings were high fidelity (44.1 kHz sampling rate, with 16-bit resolution) two-channel, digital sound files. For comparability with previous research (e.g. [3] a large set of relatively short soundscape clips were required. Hence, the original recordings were edited to create 219 clips, each with an 8 s duration and 50 ms onset and offset ramps. As in the original recordings, the overall (root-mean-square) sound level varied from clip to clip. The range in level across the edited set of materials was 18 dB.

3. Results and discussion

3.1. Validating the cognitive dimensions of the urban soundscape materials

First, we sought to confirm whether pleasantness and arousal were appropriate and sufficient dimensions to capture the principle aspects of the emotional and cognitive response to urban soundscapes. The 219 soundscape clips were each rated by five participants (aged 21–40) according to six different semantic scales; calmness, pleasantness, comfort, vibrancy, intrusiveness and informational content [5]. Each scale was presented as a visual analogue with nine points (c.f. [3]). However, here the anchor-points at the ends of each scale were defined by a set of adjectives instead of a cartoon-like manikin representation of emotion (see Table 1). For example, the 'vibrancy' scale was anchored by "dull, gloomy, bored, dreary, lifeless, tired, artificial" at one end of the scale and "energetic, fun, excited, thrilled, interested, real" at the other. Ratings were carried out using paper booklets, organised such that each of the 219 soundscape clips were rated according to one semantic scale per booklet (i.e. six booklets in total). Participants were asked to circle one of the numbers 1–9 for each soundscape rating. At the first rating, participants were also asked to write down what they judged the soundscape clip to represent. Participants typically described the nature of the environment rather than simply reporting the identity of the dominant sound source.

An extensive time period was required to complete the full rating procedure; about 5 h in total excluding rest breaks, (i.e. 45 min per booklet). Given the practical time constraint and to minimise cognitive fatigue, we allowed participants to perform the sound evaluation in the comfort of their home in a fully self-paced mode, enabling pause and playback of the soundscape clip on any trial. For consistency, sounds were played in quiet listening conditions using a shared laptop equipped with a Conexant high-definition sound card device and set of headphones (Sony Dynamic Stereo Headphones, MDR-V150). The order in which the semantic scales were rated and the sound stimuli were presented was counter-balanced across participants.

Ratings were screened for potential sources of individual bias using correlation methods. We were satisfied that the pattern of ratings given for each sound clip were broadly comparable across participants since sets of ratings for all six semantic scales were significantly correlated across individuals ($P < 0.0001$). To determine the underlying dimensions, ratings were analysed using a principal components analysis. Each semantic variable was transformed to have a mean of zero and a standard deviation of one, and a correlation matrix of these variables was then constructed. Six components were computed from the correlation matrix, each having an eigenvalue that denoted the contribution of that component to the total variance. These outputs are reported in Table 2.

The eigenvalues determined which components were retained for further analysis. Those eigenvalues >1 were retained (Kaiser's criterion, see [15] and hence components 3–6 were discarded. Our analysis then applied a varimax rotation to optimally separate components 1 and 2, whilst retaining orthogonality. Two components explained 71% of the variability in the emotional and cognitive response to the urban soundscapes and these are reported in Table 3. The first component explained 48% of the variance and