Science of the Total Environment 523 (2015) 191–200

Contents lists available at ScienceDirect

Science of the Total Environment

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

# Audio-visual interactions in environment assessment

Anna Preis [a], Jędrzej Kociński [a,*], Honorata Hafke-Dys [a], Małgorzata Wrzosek [b]

[a] Institute of Acoustics, Faculty of Physics, Adam Mickiewicz University in Poznań, Umultowska 85, 61-614 Poznań, Poland
[b] Faculty of Social Sciences, Adam Mickiewicz University in Poznań, Szamarzewskiego 89, 60-568 Poznań, Poland

## HIGHLIGHTS

- People differentiate samples based rather of the sound sources than on sound level.
- Object identification is responsible for landscape grouping.
- Audio input in a A/V assessment has a stronger influence than video input.

ARTICLE INFO

*Article history:*
Received 7 January 2015
Received in revised form 17 March 2015
Accepted 29 March 2015
Available online 10 April 2015

Editor: P. Kassomenos

*Keywords:*
Soundscape
Audio-visual assessment
Subjective assessment of environment
Environmental noise

ABSTRACT

The aim of the study was to examine how visual and audio information influences audio-visual environment assessment. Original audio-visual recordings were made at seven different places in the city of Poznań. Participants of the psychophysical experiments were asked to rate, on a numerical standardized scale, the degree of comfort they would feel if they were in such an environment. The assessments of audio-visual comfort were carried out in a laboratory in four different conditions: (a) audio samples only, (b) original audio-visual samples, (c) video samples only, and (d) mixed audio-visual samples. The general results of this experiment showed a significant difference between the investigated conditions, but not for all the investigated samples. There was a significant improvement in comfort assessment when visual information was added (in only three out of 7 cases), when conditions (a) and (b) were compared. On the other hand, the results show that the comfort assessment of audio-visual samples could be changed by manipulating the audio rather than the video part of the audio-visual sample. Finally, it seems, that people could differentiate audio-visual representations of a given place in the environment based rather of on the sound sources' compositions than on the sound level. Object identification is responsible for both landscape and soundscape grouping.



## 1. Introduction

The environment we live in and create is increasingly polluted, by many factors. Nowadays we can measure the pollution connected to chemicals, biohazards, radiation, etc. There are many procedures, standardizations and limits prescribed by law, which must be obeyed. Taking this into account, two approaches to environmental acoustics were introduced. The first one is based on the assumption that sound in the environment is waste which must be managed and reduced. This is often called environmental noise management (Traux, 1998), or environmental noise control, and is based on measurements of objective characteristics of the sound which, in general, must be reduced. A commonly used index in this approach is based on the averaged A-weighted equivalent sound pressure level ($L_{AeqT}$). It is widely known (Berglund and Berglund, 1976; Brambilla and Maffei, 2006; Carles et al., 1999; Dittrich and Oberfeld, 2009; Fastl and Zwicker, 2007; Marquis-Favre et al., 2005; Viollon et al., 2002) that $L_{AeqT}$ is well correlated with the perceived loudness of a sound, and that loudness is generally correlated with the annoyance caused by a sound. Thus, parameters based on $L_{AeqT}$ are used in most of the standardizations and limitations. However, one must bear in mind that for two different sounds with the same loudness, other factors like sharpness, roughness or fluctuation strength play a role in the subjective assessment of annoyance.

However, it appears that our preference is not as related to sound level as the noise management approach suggests. Although considerable effort has been put into noise control, recent research has shown that reducing the noise level does not necessary lead to a better acoustic comfort assessment (Ballas, 1993; de Ruiter, 2000, 2004; Dubois, 2000; Gaver, 1993; Kang, 2006). It turns out that the relationship between the percentage of people highly annoyed by noise (%HA) vs. ($L_{DEN}$), (which is based on the yearly averaged A-weighted equivalent sound pressure level during day-evening-night) %HA may be significantly different for the same $L_{DEN}$ value and different sound sources (European Environment Agency, 2010; European Parliment, 2002; Janssen et al., 2011;

* Corresponding author.
*E-mail address:* jedrzej.kocinski@amu.edu.pl (J. Kociński).

http://dx.doi.org/10.1016/j.scitotenv.2015.03.128
0048-9697/

Miedema and Oudshoorn, 2001; Miedema and Vos, 2004; Office for Official Publications of the European Communities, 2002; WHO Regional Office for Europe, 2009). For example, railway noise at the level of $L_{DEN} = 65$ dB(A) is assessed as highly annoying by 7% of people exposed to that noise, while aircraft noise at the same $L_{DEN}$ value makes 28% of people highly annoyed. From the acoustical point of view, the spectra-temporal characteristics of different sound sources are responsible for the obtained dose–response curves (the relationship between $L_{DEN}$ and %HA) for each sound source. Therefore, it seems reasonable to state that both the aforementioned complex sound structures, as well as the high-level cognitive processes responsible for the recognition of particular sources, influence the overall annoyance caused by a sound (Preis et al., 2008; Van Renterghem et al., 2013). Therefore, in environmental acoustics, a second approach is also used, based on the assumption that a sound in the environment is not necessarily waste, but a resource that can be modified and used again. This approach refers directly to the subjective perception of environmental sounds and takes its origins from the concept of the soundscape, introduced originally by Schafer (1977). Furthermore, in contrast to environmental noise control, in which the main goal is to reduce the sound level of noise (related perceptually only to discomfort or annoyance), this approach differentiates between wanted and unwanted sound sources (Brown, 2010, 2011). In general it can be stated that a soundscape is the aural equivalent of a landscape (Anderson et al., 1983; Dubois et al., 2006). As suggested by Brown (2011), *a soundscape is the acoustic environment of a place (or area), as perceived by people, whose character is the result of the action and interaction of natural and/or human factors*.

However, it must be emphasized that the most important factor in a soundscape is how the acoustic environment is perceived and assessed by people. A soundscape exists through human perception—but always within the context of a particular time, place and activity (Brown, 2010, 2011). From this definition it follows that research into soundscapes relates to many disciplines (Karlsson, 2000), such as acoustics, architecture, anthropology, ecology, communication, design, landscape, law, medicine, political science, urban planning and many more. However, for the purpose of this study the environmental acoustic and psychological aspects are the most important, therefore these aspects are briefly described below.

The perception and understanding of soundscapes have been the topic of research for many years (Axelsson, 2011; Herranz-Pascual et al., 2010; Porteous and Mastin, 1985; Raimbault, 2006; Raimbault and Dubois, 2005; Truax, 1999; Yang and Kang, 2005). Some research was carried out in the area of auditory attention and its influence on the subjective evaluation of environment (Gygi and Shafiro, 2011; Kayser et al., 2005; Oldoni et al., 2013). Obviously, one of the most important topics is the influence of environmental sounds on health (Evans et al., 2001; Kihlman et al., 2001; Lercher et al., 2011, 2013; Lercher and Widmann, 2013). Furthermore, for the last decade this concept has also been employed in urban planning as well as landscape planning (Asdrubali et al., 2014; Easteal et al., 2014; Maffei et al., 2014). For example, Skoda et al. (2014) suggest that road traffic noise annoyance in domestic environments can be reduced by the sound of water.

However, listening is part of a multisensory experience, whether the issue is attention, comprehension or assessment. The listener, embedded in a real environment – in contrast to experimental (usually lab) conditions or noise measurements and standardizations – relies on all the senses to structure a representation of the environment (Driver and Spence, 1998). Carles et al. (1999) suggested two functions of sound in the landscape which provide additional information to visual information: (1) the interpretation of the sound identified, and (2) the function related to the abstract structure of sound information. This means that in certain places with a distinct environmental identity, any acoustic disturbance can lead to a rapid deterioration in quality. Natural sounds, meanwhile, may improve the quality of built up environments, to a certain extent (Kang, 2006).

Therefore, subjective assessment of soundscape should be considered as a part of total environment perception and its influence on human beings. Total environment perception is the combination of all the senses that gives us the final assessment. There are a lot of different interactions found between the senses, especially hearing and vision, which give an insight into the complexity of multisensory stimuli perception. From a subjective point of view, which is the basis of the soundscape approach, assessment of the environment on the basis of all the senses separately, and a simple addition of the results, is not appropriate. Three main audio-visual effects, e.g., the McGurk (McGurk and McDonald, 1976), Ventriloquism (Thurlow and Jack, 1973) and Colavita (1974) effects, are good examples of such interactions. All of them leave no doubt that at the higher levels of the nervous system all the information coming from the different senses is somehow merged together, integrated and analyzed. The senses must interact with each other, thus some information (coming from one sense) can be skipped or ignored in favor of information coming from another sense, leading to completely different reactions or behavior. Interactions can also be found in multisensory attention. Research shows that stimulus from one modality can attract or distract the attention to/from the other one (Santangelo et al., 2010; Santangelo and Spence, 2007; Talsma et al., 2010). For example, Southworth (1969) showed that when aural and visual settings were coupled, attention to the visual stimulus reduced the conscious perception of sound, and vice versa.

Despite these facts, it must be stressed that research currently being conducted on people's assessment of environmental sounds is still mostly limited to separate senses. Nevertheless, some studies of multisensory assessment have been recently carried out. For example, if a place is very hot or very cold, acoustic comfort could become less critical in the overall comfort evaluation (Kang, 2006). Generally, research into audio-visual interaction relating to the assessment of environmental sounds showed that the sounds of nature can improve sensations associated with the landscape (Carles et al., 1999; Maffiolo et al., 1999). Conversely, proximity to green areas significantly improves the sound quality of the evaluated area (Gidlöf-Gunnarsson and Öhrström, 2007). Tsai and Lai (2001) found that for many sounds of nature (e.g., birdsong, water flow, wind, frogs, etc.) good or moderate visual information can enhance the total subjective evaluation of the environment. Furthermore, Viollon et al. (2002) and Carles et al. (1992) suggested that the more urban the visual sights were, the more influenced the auditory assessment was. Moreover, even the meaning of the sound plays an important role in the assessment of environmental sounds (Preis et al., 2008). For example, when Van Renterghem et al. (2013) investigated the assessment of different environmental sounds, it turned out that when the subjects did not know the sound, i.e., could not identify the sound source, they assessed it in a different way than when the picture of the sound source was shown and they could easily recognize it. Furthermore, Fastl (2004) showed that visual input can influence the perceived loudness of environmental sounds. For example, the loudness of a red train can be rated 15% higher than the loudness of a green train (Fastl, 2004). Another important aspect is our expectations concerning the environment. Brambilla and Maffei (2006) showed that the subject's expectation of hearing a specific sound in a specific environment influences their corresponding annoyance. The more the sound is congruent with the expectation, the less the evoked annoyance is. These kinds of interactions were also found in the car industry, where a good image drastically reduced the negative loudness assessment (Hashimoto and Hatano, 2001) or unpleasantness (Hatano et al., 2001). Moreover, when the audio source is not visible it causes lower noise annoyance (Bangjun et al., 2003). Also, Hong and Jeon suggest that audio-visual interaction may influence the overall rating of the environment (Hong and Jeon, 2013a, 2013b) and should be taken into account in urban planning.

The interactions between aural and visual perception, especially when sounds are related to scenes, give people a sense of involvement