Soundfield reproduction: the limits of the physical approach

Jean-Dominique Polack[1], Michèle Castellengo[1],
Valérie Maffiolo[2], Catherine Guastavino[1], Brian FG Katz[3]

[1] *Laboratoire d'Acoustique Musicale, F-75015 Paris, France, Email: polack@ccr.jussieu.fr*
[2] *FranceTelecom*
[3] *LIMSI-CNRS, BP 133 – F 91403 Orsay Cedex, France, Email: brian.katz@limsi.fr*

Introduction

Sound field reproduction techniques are usually based on the physical reproduction of the sound heard at the ears of the listeners. This is often accomplished through binaural techniques, by means of headphones and implementation of the Head Related Transfer Function (HRTF). However, many listeners have experienced that this reproduction does not equate to real life listening. One of the major shortcommings is front back ambiguities, due to the use of transfer functions that do not correspond to the individual listener's head.

A very different approach to sound field reproduction is offered by the recording engineers' tradition. This tradition does not aim at reproducing the sound pressure field at the ear of the listeners, but to create an illusion: the illusion of natural sound. Building on that approach, our laboratory has developed two recording and reproduction techniques that aim at reproducing different aspects of natural soundscapes. They are both based on Gibson's concept of ecological validity [1], revisited by mean of a semantic approach to cognition [2]. Thus, our techniques aim at obtaining the same reactions from the listeners as in real life, attested by their description of soundscapes in language: only when listeners use the same linguistic categories as in real life are they not perturbated by the simulation context.

Stereophony

We use stereophonic recordings when identification of sources present in a sound field is a key issue. A series of listening tests have confirmed the importance of the microphone directivities, as well as the distance and the angle between the two microphones. We preferably use cardioid microphones with an angle of about 100 degrees and separated by 60 to 70 cm – wider apart than classical microphone configurations such as ORTF or NOS [3]. This is based on the subjective results of a listening test conducted with 12 subjects and involving 12 sequences of traffic noise [4]. The listening set-up was the classical stereophonic triangle, with the listener at 1.5m from the loudspeaker; however, the listening room was highly damped. Subjects were instructed to select the most realistic sequence for each microphone configuration: short spacing between the microphones resulted in comments like "narrow image"; large spacing in comments like "hole in the middle". Further tests confirmed that omnidirectional microphones or narrow spacing between the microphones, are not pertinent configurations for sounscapes that present moving sources, such as traffic soundscapes [4].

Figure 1 presents an application of our stereophonic recording and reproduction techniques to urban soundscapes. Subjects were asked to group together sequences according to their level. They were free to listen to the sequences as many times as desired. Even when asked to focus on loudness, subject could not help blending signification with level: sequences with the number in a square, corresponding to soundscapes where specific events take place, are grouped together; and sequences with the number in a circle, corresponding to amorphous soundscapes where no specific events take place, are grouped together. Notice that sequences with similar levels, such as sequences 6 and 7, are not grouped together. A second test, presenting the same sound sequences at equalized levels, confirmed the results.

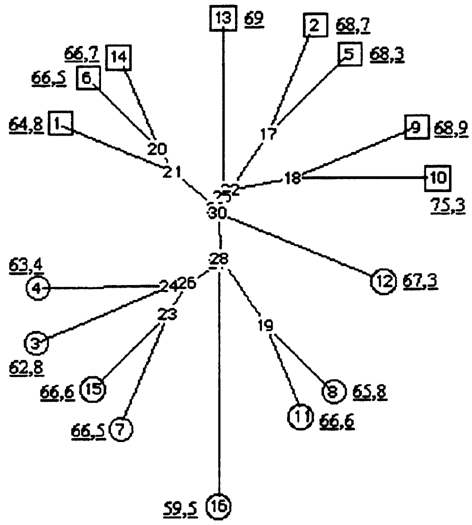

Figure 1: Loudness similarity between 16 soundscapes – 8 event sequences indicated in squares, and 8 amorphous sequences indicated in circles (see text). Mean levels in dBA are indicated for each soundscape (from ref. [5]).

In another application, a variety of 15 warning signals were played to subjects in different background contexts: alone, mixed with a traffic soundscape, or mixed in a park soundscape. The levels of the warning signals were gradually increased and subjects were instructed to halt the name the signal. Figure 2 presents the results for one category of signals, named "ambulance" by the subjects, and corresponding to different warning signals: signal 3 is most frequently named as such in each case, signal 4 is only named as such in the stand-alone case, and signal 2 is only named in context. The fact is that signals 3 and 4 effectively are "ambulance" signals, while signal 2 is a police signal, signal 5 a fire brigade signal, and signals 8 and 9 are car horns. Thus, the results of Figure 2 prove that the identification and classification of a sound depends on the soundscape it belongs to.

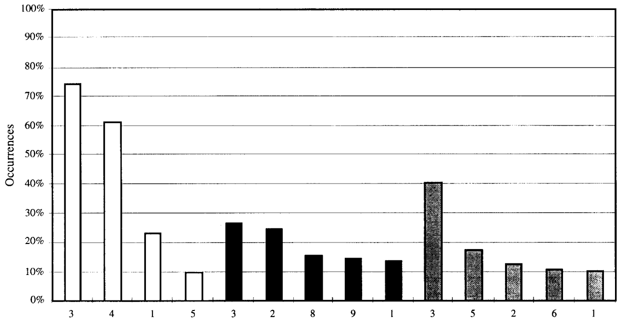

Figure 2: Warning signals most often named as "ambulance" in stand -alone context (left), in traffic (centre) and park (right) soundscapes - see text for the origin of the different signals (from ref. [4]).

Ambisonics

We use Ambisonics recordings when immersion of the listener in a sound field is a key issue. We use a standard SoundField microphone [6] which consists of coincident directional elements, with an optional omni-directional microphone for the low frequency channel. Listening tales place in our very damped listening room over 6 loudspeakers located on a horizontal plane at the ear level; or 12 loudspeakers, 6 on the horizontal planes with 3 above at ceiling level and 3 below at floor level (Figure 3). A subwoofer is located in one corner of the room, using a crossover frequency of 100 Hz.

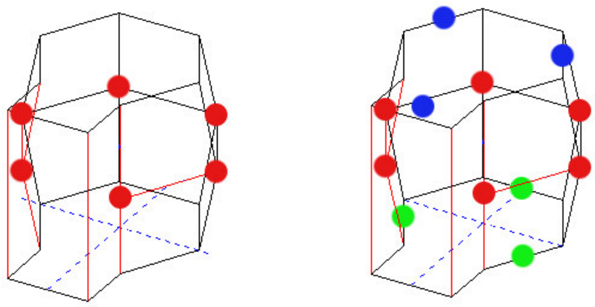

Figure 3: Loudspeaker disposition in the listening room.

Listening tests with 27 subjects, using open questionnaires, confirmed that Ambisonics is better adapted to the reproduction of the spatial properties of soundscapes [7, 8]. Indeed, with Ambisonics, we obtained the same proportion of descriptions of the soundscape in terms of "noise" as on - site (Figure 4). The listening test also confirmed that stereo was valid for descriptions of soundscapes in terms of "sounds".

Further listening tests showed that the 6 loudspeaker configuration was generally more preferred than the 12 loudspeaker configuration for urban sondscapes. However, subjective variations were found depending on the soundscape, as stereophony was often selected for music and 12 loudspeakers chosen for interior soundscapes such as the interior of a train. This is further proof of the influence of context and meaning on perception.

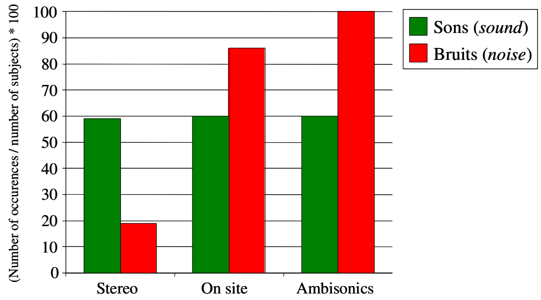

Figure 4: Percentage of descriptions of urban background noise in terms of "sounds" (green) and "noises" (red) on - site as compared with stereo and Ambisonics reproduction (from ref. [7, 8]).

Conclusion

The present paper stresses that the physical approach to sound field reproduction is not always sufficient. The recording engineers' tradition offers an alternative, on which meaningful recording and reproduction techniques can be derived. Of these techniques, stereophony is well adapted to sound source identification and Ambisonics is better adapted for simulating the feeling of immersion in the soundscape. Furthermore, this paper stresses the need for special listening rooms that are adapted to the recording systems, according the recording engineer's motto: "Which reco rding system for which reproduction environment?"

References

[1] J.J. Gibson, The ecological approach to visual perception Lawrence Erlbaum Associates, 1986

[2] D. Dubois, Cognitive approach to soundscape. CFA/DAGA ' 04 , Strasbourg, 2004

[3] Ch. Hugonnet, P. Walder, Théorie et pratique de la prise de son stéréophonique. Eyrolles, 2000

[4] C. Vogel, Etude sémiotique et acoustique de l' identificatin des signaux sonores d' avertissement en contexte urbain. Ph.D. Thesis, Université Pierre et Marie Curie, Paris, 1998

[5] V. Maffiolo, De la caractérisation sémantique et acoustique de la qualité sonnre de l' environnement sonore urbain. Ph.D. Thesis, Université du Maine, 1999.

[6] Reference to Soundfield Microphones.
URL: `http://www.soundfield.com/`

[7] C. Guastavino, Etude sémantique et perceptive de la perception des basses fréquences dans l' environnement sonore urbain. Ph.D. Thesis, Université Pierre et Marie Curie, Paris, 2003

[8] C. Guastavino, B. Katz, J.D. Polack, D. Levitin, D. Dubois, Using verbal data to measure the ecological validity of sound reproduction systems. Submitted to Acustica, 25-11-2003