



HAL
open science

Influence of context recognition on the representation of acoustic horizon: investigations on auditory distance perception

Pierre Fleurence, Mitsuko Aramaki, Richard Kronland-Martinet

► To cite this version:

Pierre Fleurence, Mitsuko Aramaki, Richard Kronland-Martinet. Influence of context recognition on the representation of acoustic horizon: investigations on auditory distance perception. 1st AUDITIVE Conference, Jun 2023, AACHEN, Germany. pp.77. hal-04138813

HAL Id: hal-04138813

<https://hal.science/hal-04138813v1>

Submitted on 23 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Influence of context recognition on the representation of acoustic horizon: investigations on auditory distance perception

Pierre Fleurence^{1,2}, Mitsuko Aramaki¹, Richard Kronland-Martinet¹

¹ Aix Marseille Univ, CNRS, PRISM, 31 Chemin Joseph Aiguier, 13402 Marseille, France, Email: lastname@prism.cnrs.fr

² GMEM, 41 rue Jobin, 13003 Marseille, France, Email: firstname.lastname@gmem.org

Introduction

How do we describe auditory space around us, this question has been deeply investigated for a number of years and from different viewpoints. In fact, an auditory space is generally studied either: (1) by its acoustic footprint characterized by room acoustics attributes such as reverberation [1] or (2) by the perceptual characteristics of the contained sound sources in terms of their localization [2], width [3], motion [4] or distance to the listener [5]. Yet, one could argue that in one hand, we do not perceive auditory space simply as a sum of sound sources and in the other hand, that auditory space does not rely solely on room acoustics. Inline with this, in [6], the authors showed that there is a strong influence of the sound source on the perception of reverberation. If we change our point of view from room acoustics to auditory space in general, would it be possible that the global context has also an influence on the perception of some auditory space perceptual parameters ?

In the 1970s, a group named the World Soundscape Project (WSP) directed by Canadian composer Raymond Murray Schafer theorized an approach of the sound environment with an emphasis on how it is perceived by an individual or a society of individuals. From this approach emerged a lexicon used to describe perceived soundscapes [7]. In this lexicon, we can find the term *acoustic horizon* defined as the farthest distance in every direction from which sounds may be heard. Interestingly, this percept seems to better characterize the auditory space previously evoked since it may refer to the spatial counterpart of the perceptual source distance.

Therefore in this work, we investigate the acoustic horizon of different soundscapes thanks to a perceptual experiment. Do people discriminate the acoustic horizon between different soundscapes and does their ability to recognize them has an influence on this discrimination ? Our hypothesis is that subjects are able to do so and that the context do have an influence on the reported acoustic horizon.

Perceptual evaluation

Forty-nine volunteers participated to the experiment with a meaning age of 35 years old (std = 10.35), 82.9% of them were right-handed. During this perceptual experiment, subjects were asked to listen to eight different soundscapes and to draw the perceived acoustic horizon for each soundscape.

Stimuli

Firstly, we designed four different soundscapes corresponding to four shapes characterizing the auditory space: an omnidirectional shape, a cardioid one, a bidirectional one and an asymmetric bidirectional one. Among these four soundscapes, we decided to create two outdoor and two indoor environments. The choice of the environments was led by their recognition capacity. We wanted to create soundscapes easily recognisable by most of people. The soundscapes were composed in collaboration with a sound artist based on these criteria. The designed soundscapes are listed in Table 1, each of them lasted 60 seconds.

Table 1: List of composed soundscapes

Shape	Soundscape
Omnidirectional	Urban parc
Cardioid	Port
Bidirectional	Hospital corridor
Asymmetric bidirectional	Small kitchen

Then, to test our hypothesis on the influence of context on the acoustic horizon, we altered the previous soundscapes so that they can no longer be recognized. For that purpose, for each sound object composing the soundscape, we measured the temporal envelop then replaced the spectrum by white noise thanks to a 1-band vocoder. We obtained what has been called EMN (Event Modulated Noises) by Gygi et al. in [8]. Thanks to that method, we kept the temporal aspect (the temporal envelop) which facilitate the sound source localization by listeners [9]. We also kept the spatial aspect, or the spatial position, of each sound source in the soundscape while we could remove the ability to recognize what is the emitter of the sound. In total, we obtained 8 different stimuli: the four initial soundscapes and the four EMN soundscapes.

We mixed the soundscapes inside PRISM's spherical array of loudspeakers [10, 11] in an oriented-object format to be able to encode and decode it in high-order ambisonics (HOA) formalism inside the spherical array. The whole apparatus is described in the next paragraph.

Apparatus

The mixing of the soundscapes and the perceptual experiment took place in a 4m-radius geodesic spherical array of

speakers, the “Sphere”, composed by 42 Genelec 8020C two-channel active loudspeakers. The soundscapes were decoded following HOA formalism at 4th order. An energy-preserving decoder [12] has been used with a basic normalization below 700Hz and a max-Re [13] normalization above. Spherical harmonics are sorted within the ACN ordering convention. The “Sphere” is located inside an semi-anechoic chamber ($RT_{60} < 0.1s$ and cutoff frequency $f_c = 80Hz$) [11].

The experiment was developed inside Max¹ where we used the Spat for Max library [14] for the spatialization aspect. A graphical tablet was used whose interface was developed in Processing².

Protocol

The experiment lasted between 55 minutes and 2 hours depending on subjects, mean duration of 1h30. It was split in two parts: a training one and the experiment itself. The protocol was the same for training and experiment, only stimuli differed between them. A session was decomposed in three different tasks: first, subject had to describe each sound object perceived in the soundscape. For this, he had to name the sound object, write its name then categorize it. He had to categorize within two attributes, a temporal one and a provenance one, inspired by [15]. For the temporal attribute, subjects had three different choice: “continuous” meaning a drone-like sound, “rhythmic” meaning with a perceptible structural rhythmic and “single event” as an event which appears only once in the soundscape. For the provenance attribute, they could choose between “natural” or a sound coming from something biologic which is not human, “industrial” is a sound coming from a tool made by human and “human” which are the sounds made by human body. The subject could listen the stimulus as many times as needed. Stimulus remains the same during a whole session and in total, subjects had to solve exercises for 12 sessions (4 in the training part and 8 in the experiment).

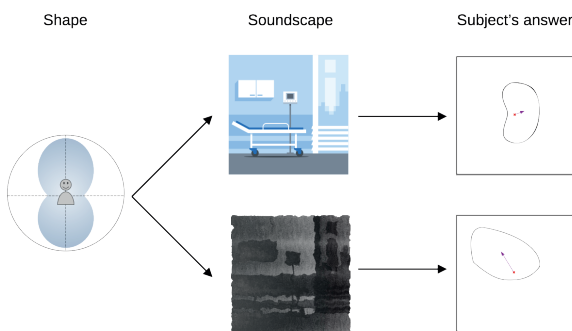


Figure 1: Summary of the protocole: hospital case

Once this categorical exercise was done, subjects had to position each sound previously listed around him thanks to a drawing interface. On this interface, the listener’s

¹<https://cycling74.com>

²<https://processing.org>

representation was at center and subjects had to draw where they perceived the origin of the incoming sound object based on azimuth and distance by writing the sound object’s name. If they could perceive a width for the sound object, they were instructed to draw an arc around the write name. If a movement was perceived, they could draw an arrow following the sound object’s movement while pointing toward the movement’s direction.

Finally for the third exercise, subjects had to draw their perception of the acoustic horizon on a graphical tablet. They were instructed precisely, in french, “*For this part, you’ll have to draw your perception of acoustic horizon. Acoustic horizon is the farthest audible distance around you, you can make a parallel with visual horizon but in the audible domain. In other terms, acoustic horizon is the limit after what you can’t hear what’s happening.*”. We gave them an example, knowing “*If you are in a room with two peoples speaking nearby you and you can’t hear what’s happening behind them because they are masking it, your acoustic horizon will be just after them.*”

At the end of the experiment, a short debriefing session was conducted with each subject to collect their global impression about the experiment and their potential free comments. A summary of the protocol is shown in Figure 1: from one auditory space shape, we had two different soundscapes with and without the context which lead to the subject’s drawing.

Results

Of the three exercises described before, only the third one was used to test our hypothesis. We retrieved the drawings from the data collected on the graphic tablet and measured some descriptors on the drawn shapes. In particular, we considered the descriptor “drawing’s area” expressed in *pixel*². We hypothesize that this descriptor is linked to the perceived size of the auditory space, this would allow us to definer later some label going from “small” to “large” for example.

A first take on the data showed that some subjects used the whole interface to draw while some of them used only part of it. Because we were interested on the difference between the natural and the noisy soundscapes, we therefore computed a ratio A_r between the natural case A_{nat} and noisy case A_{noi} for each shape and for each subject.

$$A_r = \frac{A_{nat}}{A_{noi}}$$

As shown in Figure 2, we found no significant influence of the soundscape nor the context on the ratio, except for the kitchen for which a marginal tendency was revealed by a Wilcoxon t-test to zero ($p \approx 0.052$).

Actually, we deduced that subjects adopted different strategies to report their acoustic horizon. During the debriefing session, some subjects explained that it was easier to draw the acoustic horizon for a noisy soundscape because they just had to focus on the distance auditory

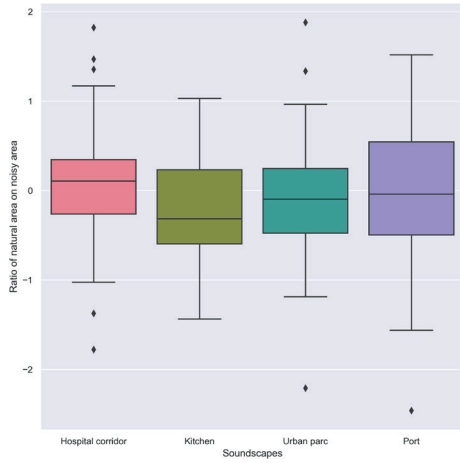


Figure 2: Boxplot of area ratio for each soundscape (log)

cues while a natural soundscape bring some uncertainty (e.g. someone whispering but localized far from the subject). At the opposite, some subjects reported that it was easier for the natural soundscape because they could not recreate a mental image of the noisy soundscape. Based on those comments, we thought that we would be able to find these tendencies through the data by conducting a clustering analysis on the subjects.

For that purpose, we conducted a hierarchical clustering analysis (HCA) to find groups and then test the influence of context (two levels: natural and noisy) and of the soundscape (four levels) our factors on these groups. For this, we did a measure of correlation between each subject's variable and did our clustering based on the correlation matrix obtained. Three groups composed by 19, 14 and 14 subjects, noted group A1, A2 and A3, respectively was obtained from the HCA (Figure 3).

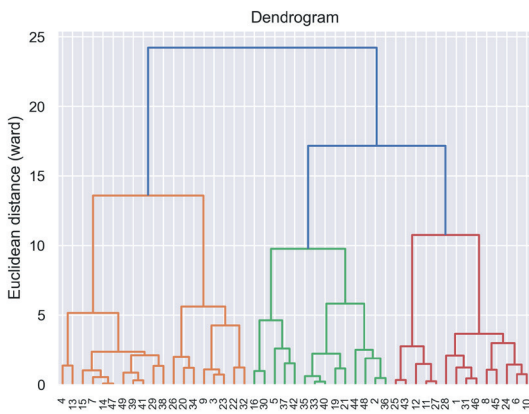


Figure 3: Dendrogram from the HCA

We did some Friedman tests to evaluate the influence of soundscape in each group and paired t-tests or Wilcoxon tests to evaluate the factors' influence on the area ratio.

Discussion

In the following, we discuss the results from the HCA.

On the influence of soundscape

A strong significant difference ($p < 0.001$, cf. Table 2) is highlighted between each group of subjects which validates our hypothesis on our ability to assess the acoustic horizon. Within a group, subjects were able to differentiate the acoustic horizons between different soundscapes. In particular, subjects in group A3 discriminated the hospital corridor from the three other soundscapes while it was less the case for groups A1 and A2 (Table 3).

Table 2: Influence of soundscape on ratio area per group - Friedman tests

Group	W	dof1	Q	p-unc
A1	0.307	3	17.526	0.0006
A2	0.514	3	21.6	0.0001
A3	0.544	3	22.886	0.0001

Table 3: Post-hoc tests for soundscape influence, only significant differences (Bonferroni correction)

Group	A	B	Test	p-cor
A1	Hospital	Kitchen	t-test	0.009
	Kitchen	Parc	Wilcoxon	0.0001
A2	Hospital	Parc	Wilcoxon	0.005
	Kitchen	Parc	t-test	0.0001
A3	Hospital	Kitchen	Wilcoxon	0.001
		Parc	Wilcoxon	0.001
		Port	Wilcoxon	0.004

Table 4: Influence of context on ratio area per group

Stim	Group	Test	p-val
Kitchen	A1	t-test	0.001
	A2	t-test	0.0001
	A3	Wilcoxon	0.0001
Hospital	A1	t-test	0.262
	A2	t-test	0.106
	A3	t-test	0.0001
Parc	A1	t-test	0.019
	A2	t-test	0.006
	A3	Wilcoxon	0.0001
Port	A1	t-test	0.064
	A2	Wilcoxon	0.412
	A3	t-test	0.434

On the influence of context

Each group seems to have a particular strategy in regard to the drawing's area (Figure 4). While the differences between the natural case and noisy (EMN) case are close in group A1, it seems that group A3 discriminates the two cases more easily. Interestingly, it seems that for the Port soundscape, the context had no effect for all groups (Figure 4). The Port was linked to the cardioid shape but most importantly, a boat was driving toward the listener from the far front during approximately 30

seconds. This sound object movement, no matter the context (natural vs noisy), could have attracted the listeners' attention which prevents them to equally evaluate the auditory space in all directions.

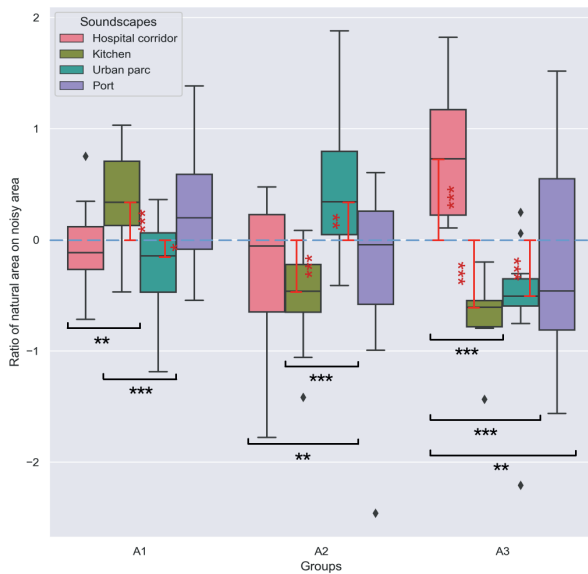


Figure 4: Boxplot of area ratio per group (log scale)

Conclusion

In this work, we presented an experiment designed to question the influence of context on the perception of the acoustic horizon, described as the farthest distance that we are able to hear in all directions. We composed four different soundscapes, removed their contextual attributes by noisy counterparts and asked subjects to report their perception of the acoustic horizon by drawings for both types of soundscapes (natural and noisy). Based on significant differences, the results confirmed that the auditory space is relevantly characterized by the acoustic horizon and that the semiotic content of the sound sources affects the perception of this acoustic horizon. Using more detailed analyses, different groups of subjects who adopted different strategies were also highlighted. Going further, these groups need to be explained and the influence of each sound object need more investigations.

References

- [1] T. Lokki, J. Pätynen, A. Kuusinen, and S. Tervo, "Disentangling preference ratings of concert hall acoustics using subjective sensory profiles," *The Journal of the Acoustical Society of America*, vol. 132, pp. 3148–3161, Nov. 2012.
- [2] L. Gandemer, G. Parseihian, C. Bourdin, and R. Kronland-Martinet, "Perception of Surrounding Sound Source Trajectories in the Horizontal Plane: A Comparison of VBAP and Basic-Decoded HOA," *Acta Acustica united with Acustica*, vol. 104, pp. 338–350, Mar. 2018.
- [3] J. Nowak and S. Klockgether, "Perception and prediction of apparent source width and listener envelopment in binaural spherical microphone array au-
- realizations," *The Journal of the Acoustical Society of America*, vol. 142, pp. 1634–1645, Sept. 2017.
- [4] S. Carlile and J. Leung, "The Perception of Auditory Motion," *Trends in Hearing*, vol. 20, p. 233121651664425, Jan. 2016.
- [5] P. Zahorik, D. S. Brungart, and A. W. Bronkhorst, "Auditory Distance Perception in Humans: A Summary of Past and Present Research," *Acta Acustica united with Acustica*, vol. 91, pp. 409–420, May 2005.
- [6] F. Salmon, Hendrickx, N. Épain, Q. George, and M. Paquier, "The influence of the sound source on perceived differences between binaurally rendered sound spaces," in *Conference Paper Presented at the Conference Headphone Technology*, (San Francisco, CA, US), pp. 1–9, 2019.
- [7] B. Truax, "Handbook for Acoustic Ecology," 1999.
- [8] B. Gygi, G. R. Kidd, and C. S. Watson, "Spectral-temporal factors in the identification of environmental sounds," *The Journal of the Acoustical Society of America*, vol. 115, pp. 1252–1265, Mar. 2004. Publisher: Acoustical Society of America.
- [9] O. Santala, M. Takanen, and V. Pulkki, "Effect of Target Signal Envelope on Direction Discrimination in Spatially Complex Sound Scenarios," Audio Engineering Society, May 2013.
- [10] G. Parseihian, L. Gandemer, C. Bourdin, and R. Kronland-Martinet, "Design and perceptual evaluation of a fully immersive three-dimensional sound spatialization system," in *3rd International Conference on Spatial Audio*, (Graz, Austria), pp. 1–8, Sept. 2015.
- [11] S. Fargeot, *Etude de la Perception d'Environnements Acoustiques 3D*. phdthesis, AMU - Aix Marseille Université, Dec. 2022.
- [12] F. Zotter, H. Pomberger, and M. Noisternig, "Energy-Preserving Ambisonic Decoding," *Acta Acustica united with Acustica*, vol. 98, pp. 37–47, Jan. 2012.
- [13] S. Moreau, J. Daniel, and S. Bertet, "3D Sound Field Recording with Higher Order Ambisonics – Objective Measurements and Validation of a 4th Order Spherical Microphone," (Paris), pp. 1–24, Audio Engineering Society, 2006.
- [14] T. Carpentier, "A new implementation of Spat in Max," (Limassol, Cyprus), Zenodo, July 2018.
- [15] J. Salamon, C. Jacoby, and J. P. Bello, "A Dataset and Taxonomy for Urban Sound Research," in *Proceedings of the 22nd ACM international conference on Multimedia*, MM '14, (New York, NY, USA), pp. 1041–1044, Association for Computing Machinery, Nov. 2014.