AN INVESTIGATION INTO THE PERCEPTION OF REPRODUCED URBAN SOUNDFIELDS

T Ackroyd University of Salford, Salford, UK
Y W Lam University of Salford, Salford, UK

1 INTRODUCTION

1.1 Soundscape and Soundfield

A soundscape is a sonic vista, inherently contextualised. A soundfield recording may be taken of a soundscape, and when the soundfield is reproduced, a new soundscape is created (from the modified soundscape of the new context.) Westerkamp aligns the realms of acoustic ecology and soundscape composition, both curating place within sound[1]. The researcher would expand the link to state that one who reproduces a soundfield with the objective of simulation is simultaneously taking on the role of composer, by creating new context. From this paradigm, the research is motivated by neither imitation of existing soundscape perceptions nor creation of new perceptions, but understanding of the very nature of the differing perception.

1.2 Semantic Differential Analysis

Semantic differentials are a popular method for assessing the perception of urban soundscapes, owing to the quantitative nature of the data and hence the ease of comparison with acoustic parameters. Kang conducted a large survey using 18 semantic scales at two soundscapes with 491 participants[2]. A principal component analysis (PCA) produced dimensions that were offered as a framework on which to rate and categorise soundscapes. This four-dimensional space had the factors relaxation/calmness, dynamics/vibrancy, communication, and spatiality. Similarly, Axelsson et al - using 116 semantic scales - produced the dimensions pleasantness, eventfulness, and familiarity[3]. Davies et al tested how such soundscape dimensions manifest in the perception of reproduced soundfields by repeating Kang's questionnaire with 15 participants listening to 8-channel ambisonic reproductions of four local urban soundscapes[4]. There is yet to be any experimentation that uses the same soundfields to compare soundscape dimensions derived from semantic differential analyses for in situ listening and amidst reproduction. Furthermore, the phenomenon is yet to be explained or linked to other forms of analysing the differing perception of reproduced soundfields.

1.3 The Cognitive Approach

Guastavino et al offered the categorisation of open responses into three semantic categories as a means of empirically assessing the ecological validity of soundscape reproduction[5]. This approach is 'a linguistic analysis of complex phrasing, rather than a lexical analysis of words [as per the semantic differential analysis],… to infer properties of mental representations of acoustic phenomena'[6]. It is, however, yet to be conducted with the same participant group in situ and in laboratory reproduction nor is it yet to be conducted alongside another approach for cross-validation.

1.4 Aims

This research compares two methods – semantic differential analysis and semantic categorisation of open response - in measuring the difference in perception of in situ and reproduced soundfields. A complimentary method of semantic categorisation analysis is suggested. Furthermore, the researcher explores what combined inferences can be made by the approaches on the factors responsible for the differing perception of urban soundscapes and reproduced soundfields.

2 METHOD

2.1 In Situ

The first phase of testing was in situ. Five sites were chosen such that there was sufficient variation in the soundscape, activity or purpose, and demographic of users: the entrance to Newton Building, Peel Park campus, University of Salford; a central foyer of Salford Precinct; the public green space at MediaCity, Salford; the west corner of Manchester's 'Triangle'; and the garden of The King's Arms pub, Salford. The focus group consisted of five city-dwelling postgraduate students aged 24-42, each with knowledge, or awareness, of soundscapes and acoustics. The tests took place on an autumn weekday afternoon. The participants were taken, as a group, to the five sites where they individually answered the same questionnaire of six open questions. Finally, at The King's Arms they also completed a semantic scales enquiry. The final site was chosen on the interest of its temporal nature: pubs are less frequented and quieter in the afternoon than in the evening thus the afternoon soundscape may appear incongruous with expectations of place. At each site, B-format recordings were made with a multi-capsule Soundfield microphone.

2.2 Reproduction

Four months later, further soundfield recordings (of 10-15 minutes) were made at the same five sites on a weekday afternoon. Two-minute clips were extracted from these later recordings. Selection was based on sound quality (since wind noise was an issue) and the typicality of representation. Two minutes has been deemed sufficient for one to perceive and judge a soundscape[7]. The laboratory listening tests took place in a well diffused listening room. The recordings were played back through an 8 channel ambisonic loudspeaker configuration. Since the four of the five soundscapes were outdoors, a 2D reproduction was chosen a better perceived 'naturalness'[8]. Decoding was conducted in MaxMSP with an RME digital to analogue converter. Sound levels (L_{Aeq}) were measured and balanced to that of the in situ soundscape using peaks and 10 second samples. Each participant sat at a desk in the centre of the room. A single ceiling lamp lit the desk such that the loudspeakers, and the rest of the room, were less apparent. The questionnaire was completed with pen and paper, as per the in situ responses. The researcher, illuminated only by a desktop monitor, remained in view in the corner of the room. (Note that the presence of the researcher is consistent with the first phase, in situ tests.) The participants were invited one at a time. The clips were played twice with the option of more repeats if desired (although no one chose to do so.) There was casual conversation before, during and after the tests. The participants were not informed of the origin of the recording - namely, whether they had heard them or not. This apparently gave a slight sense of mystery and intrigue. The playback was indeed of unheard soundfield recordings so as not to trigger memory of particular (sonic or non-sonic) events but only of the holistic experience. Furthermore, this allows the influence of day to day variation of the soundscapes to be compared to the influence of using the same participants.

2.3 Soundscape Selections

The precinct offered an indoor soundscape; the Kings Arms beer garden offered a mid-afternoon soundscape that appears incongruous with expectations; the Triangle and Quays offered rather typical examples used by those in soundscape research; the University campus offered a very familiar soundscape. Essentially, the selection offered a range in terms of dynamics; familiarity; contextual congruousness; indoor-outdoor; demographics; and proximity to listening room of the ambisonic reproductions.

2.4 Participant Group

The participant group consisted of five people with whom the researcher is well acquainted. All five were in post-graduate study of scientific background: one in soundscape research; one in theoretical physics; and three in acoustics. The incentives were informal: helping a friend; general interest in the research; tea-making services; and beer. The group and researcher rendezvoused outside their place of work at the first soundscape location, the front of the Newton Building, University of Salford. The researcher provided transport between sites by means of a hired minibus.

The use of a focus group meant that the researcher could look into the individual agencies without variation between in situ and reproduction listening environments. The effects of memory are not ignored despite the four month period between the in situ and laboratory listening tests. One participant claimed a particularly good memory of his in situ responses whilst responding in the laboratory environment. Using this focus group allows memory to be explored – alongside place-relationship – when comparing phase one and two.

2.5 Semantic Differentials

The semantic differential enquiry took place after completion of the open questionnaire at the final soundscape, the beer garden at the Kings Arms pub, Salford. This closed form of enquiry was only used at the final soundscape so as to not influence responses to the open questions. Six scales were used: unpleasant – pleasant; uneventful – eventful; social – unsocial; boring – chaotic; happy – sad; noisy – quiet; exciting – tranquil; and finally unpleasant – pleasant with regards to the environment overall. The polarity of the scales was varied to encourage responses that focus on each individual scale rather than a develop left-right polarity for any overarching and simplified expression, of say good-bad.

2.6 Cognitive Categories

The open responses were initially analysed by categorisation into three semantic classes: object-centred; subject-centred; and source identification. This enables a comparison with Guastavino's results[5] and hence an assessment of the influence of using different participants.

Another method of categorisation was applied to the participants' responses. Firstly, this was in order to increase the rigour of the categorisation process and, secondly, to produce a broader and more varied array of categories that – derived from the experience of the experiment itself – may offer a richer insight into the process of soundscape interpretation. The categories themselves were for the researcher to choose. The decision was reached from an epistemological foundation in ontology. The researcher takes the (post-structural) stance that there is an external reality and each person within it has their own individual internal reality. Existence is the interplay between the two realities - comprising of experiences. Perception is the organisation, identification and interpretation of sensory information[9]. Perception and memory exist only in the same instance. When a participant is questioned on the soundscape, their interpretations manifest on the paper. The survey itself is asking for their interpretation. So it then becomes very difficult to separate opinions and shared

understandings. Where does one draw the line? Therefore in the categorisation technique 'opinions' are not entered as subject-orientated. The phrase "I like the sound of the choir" would only have one subject entry, 'I'. However, an effect of the environment on the earwitness is entered as a subject entry. "I find the choir very relaxing" would have two subject entries, 'I' and 'relaxing'.

Object entries are simply references to nouns. Descriptive entries are descriptions of anything: objects, feelings, verbs, timbres etcetera. Verb entries include silent and sounding actions. "A choir is singing" would have one object entry and one verb. The sound event category is only for direct references to a sound and its source or words that describe the sound itself. "The sound of the trees" would have 'sound' as an entry and "a rustling nearby" would have 'rustling' as a sound event entry. The latter example identifies another issue addressed in the categorisation: 'rusting' is a verb. Since 'rustling' is never used in a literal reference to anything other than sound creation, it is categorised as a sound event. Onomatopoeic verbs are identified as sound event entries.

3 RESULTS AND DISCUSSION

3.1 Semantic Differential Enquiry

3.1.1 Results

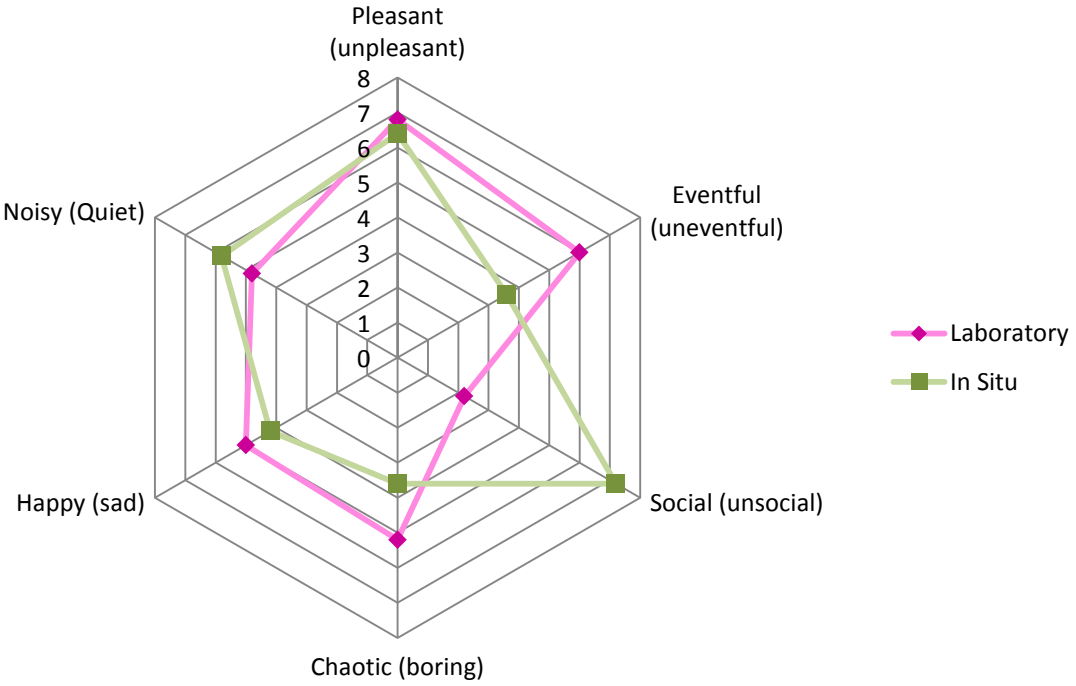

Figure 1 - Semantic enquiry responses for in situ and laboratory soundscapes

The semantic enquiry, of the King's Arms beer garden soundscape, showed the most significant variation between the in situ soundscape and the reproduced soundfield, was accounted for by the social descriptor. This is owing to the recording having been made without the participant group present at the site and that the participants are listening to the laboratory reproductions alone in the listening room, with the researcher being the only other person present opposed to six others with whom a couple of hours had been shared.

It is interesting to note that the lab soundscape was deemed more eventful and chaotic whilst also quieter. The apparent eventfulness and chaos were potentially enhanced by contrast with the

uneventful and dull non-sonic envrionment of the listening room. These may be deemed descriptors that gesture more significantly towards the other dimensions of the environment, beyond the soundscape, and that one is not accustomed to attaching exclusively to sound. Similarly, the moderate isolation of the soundfield, thus 'alternative perspective', facilitated the unveiling of more detail and texture that, in turn, increased the perceived eventfulness and chaos. Furthermore, the in situ soundscape of the Kings Arms was percieved as so uneventful and boring because it did not meet the expectations of the physical environment (owing to the time of day) and therefore contrasted with the participants' expectations.

3.1.2 Principal Component Analysis

	In situ responses				Reproduction responses			
	Component				Component			
	1	2	3	4	1	2	3	4
% of Variance	38.5	30.3	22.5	8.7	42.5	20.3	18.6	18.6
Pleasant (unpleasant)	**-0.931**	-0.267	-0.193	-0.156	0.390	-0.161	**0.889**	-0.178
Eventful (uneventful)	0.115	0.682	-0.568	0.446	**0.924**	-0.239	-0.196	-0.226
Unsocial (social)	0.166		**0.982**		-0.800	-0.406	-0.362	-0.255
Chaotic (boring)	-0.097	**0.983**	0.141		**0.909**	-0.222	0.342	
Sad (happy)	**0.919**	-0.194		-0.344	-0.236	**0.943**	-0.146	0.184
Quiet (noisy)	0.742	-0.519		0.419	-0.148	0.187	-0.135	**0.962**

Figure 2 - Comparison of in situ and laboratory responses. Kings Arms soundfield; N=5; PCA with quartimax rotation and Kaiser normalisation.

The principal component analyis was performed in SPSS. A Quartimax rotation was empolyed to spread the commponent loadings and encourage clearer contributions from the semantics scales. The number of scales used was greater than the number of participants. However, the subject-observation ratio has no impact on the 'goodness' of the principal component anlaysis[10]. Furthermore, the low number of observations has no impact on factor stablity[11]. **Error! Reference source not found.** shows the first four factors and their variances, with quartimax rotation and Kaiser noralisation. In situ the first factor positions itself from the contributions of unpleasantness, sadness, and quietness; the second from chaos (and eventfulness); and the third from unsocialness. For the laboratory reproduction, the first factor positions itself from the contributions of eventfulness, chaos, and socialness; the second from sadness; the third from pleasantness; and the fourth from quietness. The in situ first two components, pleasantness/happiness (39%) and chaotic/eventful (30%), correspond with Axelsson's first two components, pleasantness (50%) and eventfulness (18%).

It can be seen that in situ, the process of response is more evenly accounted for by the first three factors. In the laboratory reproduction, increased variance is explained by the first factor suggesting a more simplistic process of evaluation. This skewed component loading is the same as can be seen in the comparison of Kang's and Davies' results. Kang et al's in situ results produced four components accounting for 26%, 12%, 8% and 7% of variance. Davies et al's reproduction results produced 41%, 10%, 7% and 7%. This experiment, with a small number of participants and semantics scales, has managed to produce the same soundscape dimensions as larger-scale existing experiments as well as the trend for skewed component loading for reproduction soundscape responses.

The first component in situ represents the subects' pathetical repsonses. In contrast, the reproductions first component represents dynamics/vibrancy whilst personal feelings and opinions are relegated to components two and three. This could suggest that the laboratory soundscape interpretations are principally objective and cold. However, this effect may be owed to the nature of the location – a pleasant and familiar pub garden – that influenced the interpretation of an otherwise unpleasant and unfamiliar soundscape in situ. Eitherway, the reproduction has shifted 'objective'

ahead of 'subjective' in the ranking of components. (The semantic categoristion analysis supports, and elorates upon, this trend.)

3.1.3 Conclusions

This semantic differential enquiry, with a small number of participants and semantics scales, can produce the same soundscape dimensions as larger-scale existing research as well as the trend for skewed component loading for reproduced soundfield responses. Reproduced soundfields were evaulated increasingly in terms of dynamics/vibrancy and decreasingly in terms of appreciation.

3.2 Linguistic Discourse Analysis

3.2.1 Memory versus Soundscape Sample: a Comparison with Existing Research

Here a comparison is made with existing research in which different participants were surveyed in situ and at laboratory reproductions, opposed to a focus group. Furthermore, the soundfield was recorded at the time of the in situ survey.

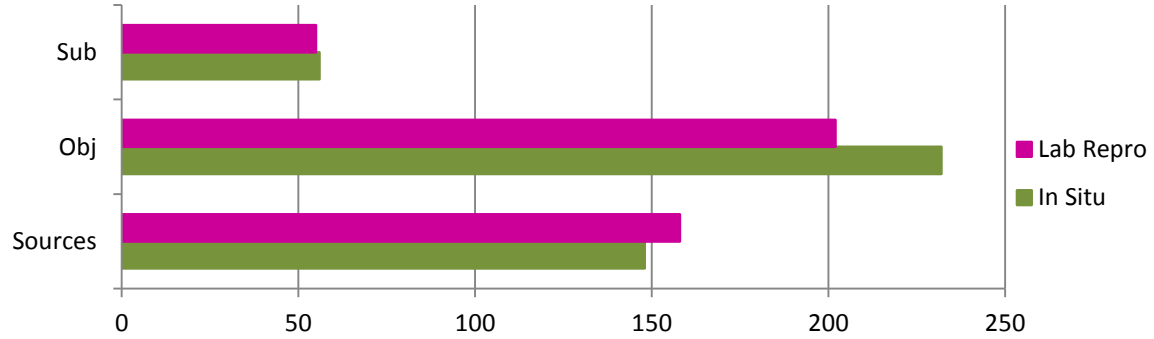

Figure 3 – Open responses into Guastavino's three semantic categories

Overall, there is a good correlation between the responses to the perception of the in situ soundscapes and of the laboratory soundscapes (see Figure 3). There was less reference to objects and singular sources since they cannot be identified as easily without as full a sense of the environment, namely without sight; smell; touch; habitual expectation; and more acute spatial auditory perception. There was also more reference to aesthetics in response to the laboratory soundscape since more attention was given to the sound itself and less to the sound in relation to the non-sonic environment.

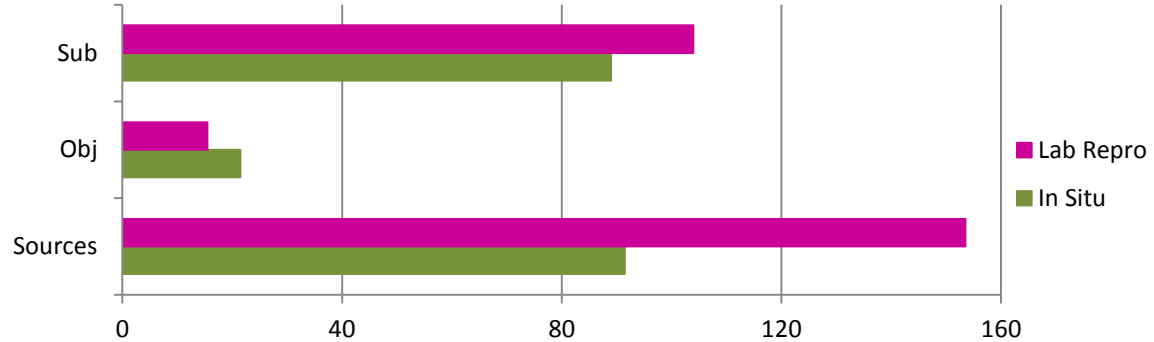

Figure 4 – Guastavino's open responses (ref) into three semantic categories

Figure 4 shows the results of previous research by Guastavino's similar experiment with various soundscapes, 42 participants in situ and 29 for the reproduction. The differences between the two experiments are firstly, that Guastavino used different participants for the reproduced soundfields to

those used in situ at the original soundscape and secondly, that the same soundfield – recorded in situ - was reproduced (with a similar 8-channel ambisonic setup).

The distribution of subject-, object-, and source-centered responses will vary due to the construction of the questionnaire. It is the in situ to laboratory variations that are of interest. The principal differences in test conditions are that this investigation using the same focus group of participants, rather than users of the space and volunteers in the listening room, and that the soundfield recordings are not from the time of in situ questioning. It appears that reproducing soundfield recordings of exactly the same soundscape experienced by the participants in situ, has less influence in aligning the in situ and laboratory responses than using the same participants. This infers that memory of the in situ experience is more important than physical soundfield accuracy (in terms of typical daily and seasonal variation). However, the margin of confidence in the former statement is dependent upon the absence of significant variation between the two groups participating in Guastavino's study of ecological validity. Furthermore, the vastly greater variation for Guastavino's source-orientated responses suggests she employed a somewhat different technique of categorisation. Potential differences in categorisation techniques is limited for 'object' references, and to similar extent 'subject' references, but for sound source identification there is larger scope. (For instance, she analyses foreground and background sounds separately. Sources could be objects, timbres, physical acoustic phenomena, verbs etc.)

3.2.2 Responses into Five Categories

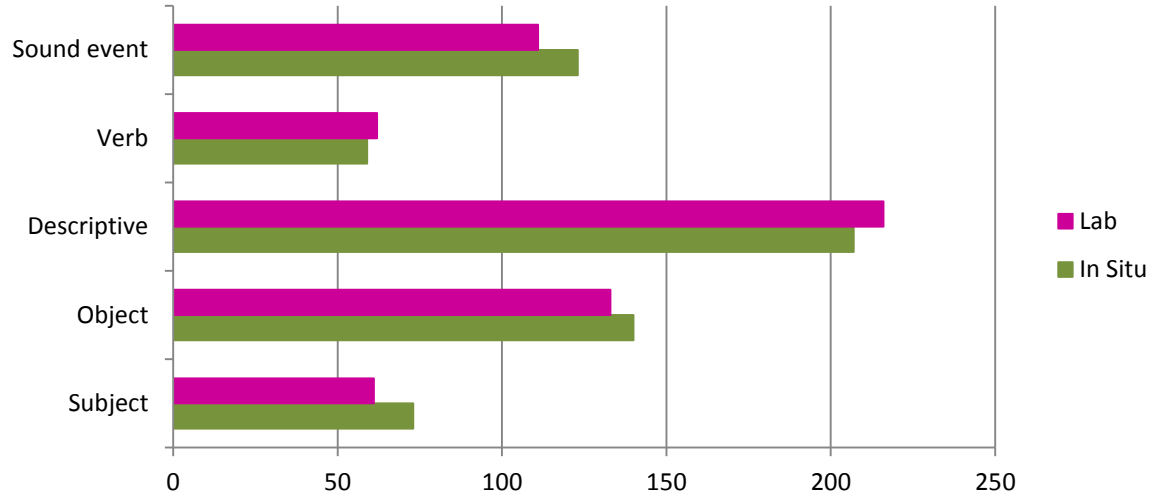

Figure 5 – Overall response categorisations for in situ and lab

The open responses were re-analysed, this time with the newly proposed categorisations: sound event; verb; descriptive; object; and subject. On first glance there appears to be a good correlation between the in situ and reproduction soundscape responses shown in Figure 5. The most prominent variation is the decreased 'object' and 'sound event' reference during the laboratory reproductions. The reproduction environment doesn't offer visual accompaniment for the direct identification of objects. Instead, objects can only be inferred from the sounds and the subjects memory and understanding of the original environment. At this stage, one can only attribute the decrease in sound event references to the reduced clarity and spatiality of the reproduction soundscape.

Summating the total responses, and thus overlooking the individual cases, can lose trends within the data. It next transpires that looking at the inter-relationships between the category references, by means of a principal component analysis, unravels more variation between in situ and reproduction listening experiences.

3.2.3 Principal Component Analyses

The principal component analyses (PCAs) expose the inter-relationships between the five category references. The PCAs treat all soundscape-participant combinations as 50 discrete observations, 25 for in situ and 25 for laboratory reproduction as shown in Figure 6. The category contributions are visualised in Figure 7 for the in situ responses and in Figure 8 for the reproduction responses.

	In situ open responses			Reproduction open responses		
	Component			Component		
	1	2	3	1	2	3
% of variance	30.0	25.4	22.0	29.0	27.9	22.2
Subject		-0.149	**0.937**		**0.835**	-0.189
Object	**0.606**	**0.642**	0.162	**0.877**	0.117	-0.289
Descriptive	**0.672**		0.197			**0.922**
Verb	-0.134	**0.901**	-0.229	**0.819**	-0.341	0.257
Sound event	**0.815**	-0.152	-0.318		**0.753**	0.273

Figure 6 - Principal component analyses of in situ and lab response categorisations with Varimax rotations and Kaiser normalisation

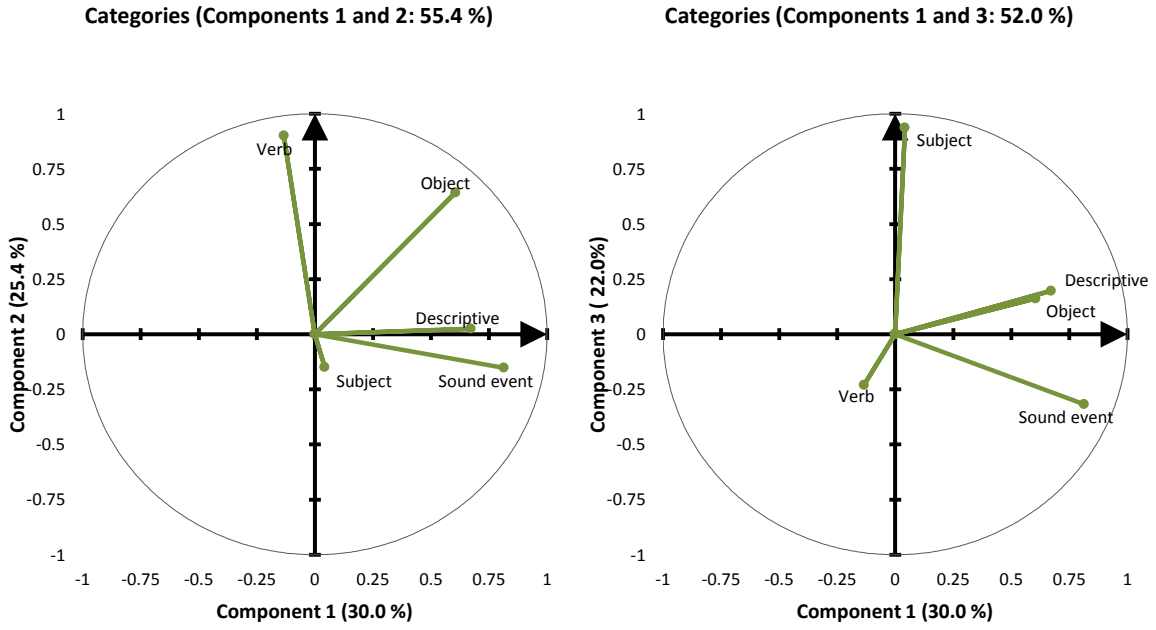

Figure 7 – Plots of the category contributions on the first three principal components from the in situ responses

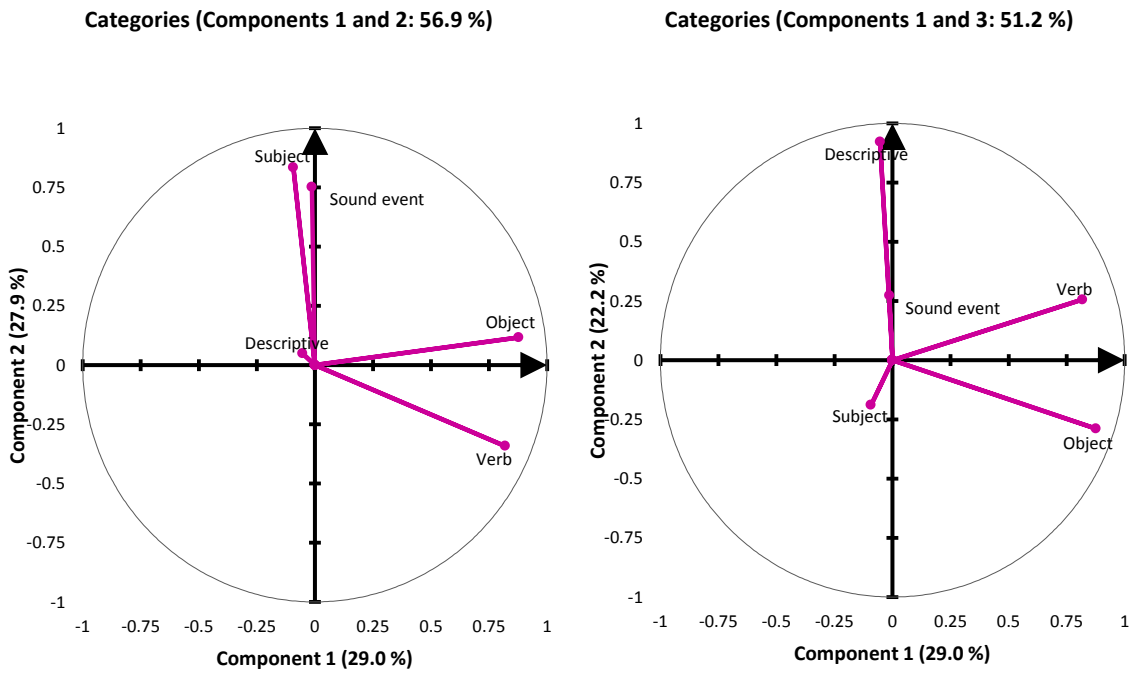

Figure 8 – Plots of the category contributions on the first three principal components from the reproduction responses

It can be seen in Figure 6 that the component loadings are more heavily skewed to the first component (C1) for the in situ responses than for reproduction. However, the in situ first component, C1(in situ), does account for the variance of three, rather than two, of the semantic categories. The trend is opposite to that of the semantic differential analysis but not pronounced enough to be deemed significant.

Sound events account for most variation along C1(in situ). In situ, the sound events are identified alongside descriptive words and reference to objects: the participants elaborate on their description of the sound events; associate sounds to physical sources in the environment; and describe the visual environment. However, in the reproduction listening experience, sound events are perceived as detached from the physical environment: 'sound event' shifts to C2(repro) and 'object' and 'verb' align strongly together on C1(repro).

For the reproductions, the physical environment – represented by C1(repro) – has the highest percentage of variance, 29%. This could be attributed to the participants remembering or imagining the physical environment in order to interpret the soundscape: they are doing so by their own choice, rather than being presented with visual information. One might have hypothesised that without visual cues in the case of reproduction, the focus would be on the sound itself, and hence 'sound event' would represent the first component. In fact 'sound event' has aligned itself with 'subject' on C2(repro). However, C2(repro) accounts for only 1% less variance than C1(repro) therefore their variances can be approximately equated. In the reproductions 'descriptive' - accounting for most variation along C3(repro) - becomes more a measure of how much the participants write, rather than elaborations of particular interpretations.

3.2.4 Conclusions

The discourse analyses' initial categorisation counts appear to show little variation between the cognitive process of perception of in situ soundscapes and reproduction soundscapes. However, the new categorisation method along with the principal component analysis unraveled interesting differences. The most significant phenomenon of soundfield reproduction was that sound events

were perceived independently of objects: a clear separation of the sonic and visual environments in the participants' cognitive processing of soundscape interpretations.

4 CONCLUSIONS

The semantic differential analyses do show some variation between the evaluation of in situ and reproduction soundscapes. Firstly, a skewed loading on the first component for participant responses in the laboratory reproduction listening experience and, secondly, reproduced soundfields were evaulated increasingly in terms of dynamics/vibrancy and decreasingly in terms of appreciation. The discourse analyses' initial categorisation counts appear to expose very little of the variation between the cognitive process of perception of in situ urban soundscapes and reproduced urban soundfields. However, the new categorisation method along with the principal component analysis unraveled more significant differences. The most interesting difference was the clear separation of the sonic and visual environments in the participants' cognitive processing of the soundscape of reproduced urban soundfields.

5 REFERENCES

1. H. Westerkamp, "Linking soundscape composition and acoustic ecology," *Oragised Sound,* pp. 51-56, 2002.
2. J. Kang and M. Zhang, "Semantic differential analysis of the soundscape in urban open public spaces," *Building and Environment,* pp. 150-157, 2010.
3. O. Axelsson, M. E. Nilsson and B. Berglund, "A principal components model of soundscape perception," *Journal of the Acoustical Society of America,* pp. 2836-2846, 2010.
4. W. J. Davies and J. E. Murphy, "Reproducibility of soundscape dimensions," *Inter-Noise,* 2012.
5. C. Guastavino, B. F. G. Katz, J.-D. Polack and D. J. Levitin, "Ecological Validity of Soundscape Reproduction," *ACTA Acustica United with Acustica,* pp. 333-341, 2005.
6. D. Dubois, C. Guastavino and M. Raimbault, "A cognitive approach to urban soundscapes: using verbal data to access everyday life auditory categories," *ACTA Acustica United with Acustica,* pp. 865-874, 2006.
7. N. Bruce, "The Effect of Expectation on the Perception of Soundscapes," Thesis, Salford, 2011.
8. C. Guastavino and B. Katz, "Perceptual evaluation of multi-dimensional spatial audio reproduction," *Acoustical Society of America,* pp. 1105-1115, 2004.
9. D. L. Schacter, D. T. Gilbert and D. M. Wegner, Psychology, Worth Publishers, 2011.
10. J. W. Osborne and A. B. Costell, "Sample size and subject to item ratio in principal components analysis," *Practical Assessment, Research and Evaluation,* 2004.
11. W. A. Arrindell and J. van der Ende, "An Empirical Test of the Utility of the Observations-To-Variables Ratio in Factor and Components Analysis," *Applied Psychological Measurement,* pp. 165-178, 1985.