ELICITING INDIVIDUAL LANGUAGE DESCRIBING DIFFERENCES IN AUDITORY IMAGERY ASSOCIATED WITH FOUR MULTICHANNEL MICROPHONE TECHNIQUES

William L. Martens, Sungyoung Kim, Kent Walker, David Benson, and Wieslaw Woszczyk
Schulich School of Music, McGill University, 555 Sherbrooke Street W., Montreal, QC, Canada H3A 1E3
wlm@music.mcgill.ca

1. INTRODUCTION

When different listeners hear several versions of a single piano performance, each of which is different from the others in a number of ways, those individual listeners might use any number of words to describe the differences that they hear. It is worth questioning whether they are hearing the same distinctions on identifiable auditory attributes, even though they name those distinctions differently. This question generally needs to be addressed in a systematic perceptual evaluation of modern techniques used for multichannel sound recording. The work described in this paper attempted to address this question through a controlled verbal elicitation that was followed by an examination of ratings on individualized attribute scales constructed from the results of the verbal elicitation for each of five listeners.

In contrast to a recent related investigation that focused upon the development of a consensus language for a group of listeners with similar training (Martens & Kim, 2007), the current study relied upon the modern descriptive analysis technique termed the Repertory Grid (RepGrid) Technique. This approach, perhaps first introduced to the sound quality evaluation community by Kjeldsen (1998), was employed with good results for the sensory evaluation of multichannel sound reproduction by Berg & Rumsey (2000; 2006). In their verbal elicitations they collected both descriptive and attitudinal features of multichannel reproduced sound, and used cluster analysis to organize the obtained terms. In the current study, five trained listeners were asked to focus their attention primarily upon spatial attributes that could be used to discriminate between stimuli presented in groups of three. Always, a given triadic comparison was made between three different versions of a single solo piano performance. By focusing only upon multiple versions of a solo piano performance, the current investigation was not too ambitious in its exploration of reproductions that were very similar to one another in many respects, and were indeed all acceptable reproductions of piano performances that were captured simultaneously using four different multichannel microphone techniques. A full explanation of the RepGrid approach is beyond the scope of this paper, and so the reader is referred to the manual for commercial software enabling the analysis of such data (Fransella et al, 2003).

2. METHOD

Two short excerpts of each of four solo piano pieces composed in the European concert musical tradition were chosen for this study: works by Bach, Schubert, Brahms, and a contemporary improvisation by Plaunt. It was hypothesized that some microphone techniques might be preferred for certain musical selections within the performance space, which was the 600-seat Pollack Concert Hall located at McGill University, and expert advisors agreed that the following four surround microphone arrays were appropriate selections for comparison: Fukada Tree, Polyhymnia Pentagon, Optimized Cardioid Triangle combined with a Hamasaki Square, and a SoundField microphone. All musical excerpts were performed in the same concert hall by a single musician, and played on a single piano. The details of the recording procedure are well documented in Kim, et al. (2006).

The 32 five-channel stimuli were presented via five active full bandwidth loudspeakers (Dynaudio model BM15A) positioned at a height of 1.2 m from the floor and at a radius of 1.5 m from the central listening position (no LFE signal was prepared, and no subwoofer was employed). This surround sound reproduction system was located in MARLAB (Multichannel Audio Research LABoratory) at McGill University. The five listeners had all participated in previous listening tests conducted in MARLab, and were all either professors in the Sound Recording Area within the Schulich School of Music of McGill University, or were students enrolled in the School's doctoral program in Sound Recording.

The details of the experimental method are similar to those that were well described in the recent report by Martens & Kim (2007), although that study was focused upon the development of a consensus language for a group of listeners with similar training. None of the five listeners had participated in that previous study, though they were all well aware of that study's results. For the current study, no examples were presented prior to listening sessions, such as stimuli exhibiting extreme positions on any of the attributes identified in the previous study. Also, no group discussions were held, in order to avoid any bias other than that which comes from training in sound recording.

3. RESULTS

The following table gives a summary of some of the bipolar adjectives that were elicited from the five listeners, along with the relative frequency with which they were generated.

Listener 1:
wide vs. narrow 10/32
focused vs. diffuse 9/32
tight-bass vs. muddy-bass 8/32
distant vs. close 3/32
bright vs. dark 2/32
not reverberant vs. reverberant 2/32

Listener 2:
sparkling, bright vs. dull, muddy 10/32
wide vs. narrow 16/32
close, near vs. far, distant 4/32
more spacious vs. less spacious 2/32

Listener 3:
narrow vs. wide 14/32
defined vs. diffused 6/32
open vs. closed 4/32
far vs. close 4/32
bright vs. dark 4/32

Listener 4:
strong-centre vs. well-spread 15/32
clear vs. blurry 5/32
distant vs. close 4/32
dull vs. bright 3/32
surrounding vs. less surrounding 3/32

Listener 5:
narrow vs. wide 12/32
covered, muffled vs. clear 8/32
thin, narrow bass vs. fat, wide bass 5/32
spiky, pointy vs. smooth,warm,even 4/32

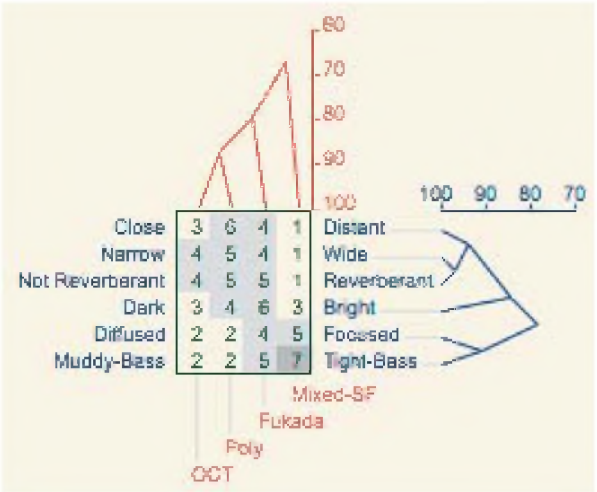

Fig. 1. Results of RepGrid cluster analysis relating Mic. Techniques to ratings on the 6 attribute scales constructed for Listener 1 from his elicited bipolar constructs.

Figure 1 shows the results of Repertory Grid (RepGrid) cluster analysis performed on ratings obtained from just one of the five listeners. Note that instead of running the RepGrid analysis on ratings for all 32 stimuli, the centroid response dataset was calculated from the combined ratings for all 8 of the musical programs for each of the four microphone techniques (and rounded to integer values as required). Thus, the average responses given for each of the four microphone techniques, regardless of which musical selection was being rated, were submitted to RepGrid analysis, thereby giving a general overview for each. This analysis for Listener 1 gave a tight cluster between obtained ratings on three attribute scales: on the <wide-narrow> scale, the <distant-close>, and the <not reverberant-reverberant> scale. A second tight cluster was observed between ratings on the <focused-diffuse> and the <tight(bass)-muddy(bass)> scale. Ratings on the <bright-dark> scale separated out from the other attributes in that this attribute described timbral rather than spatial character of the reproduced musical performances.

REFERENCES

Martens, W. L., and Kim, S. (2007) Verbal Elicitation and Scale Construction for Evaluating Perceptual Differences between Four Multichannel Microphone Techniques. In Proc. Audio Engineering Society 122nd Int. Conv., Vienna, Austria.

Kjeldsen, A. D. (1998) The measurement of personal preferences by repertory grid technique. In Proceedings of 104th Conv., Amsterdam, The Netherlands.

Berg, J., and Rumsey, F. (2000) In search of the spatial dimensions of reproduced sound: Verbal protocol analysis and cluster analysis of scaled verbal descriptors. In Proc. Audio Engineering Society 108th Conv., Paris, France.

Berg, J., and Rumsey, F. (2006) Identification of Quality Attributes of Spatial Audio by Repertory Grid Technique. J. Audio Eng. Soc., 54(5):365–379, 2006.

Fransella, F. , Bell, R. and Bannister, D (2003) A Manual for Repertory Grid Technique, 2nd Edition. John Wiley & Sons, Inc., New York, NY.

Kim, S., Martens, W. L., Marui, A. and Walker, K. (2006) Predicting Listener Preferences for Surround Microphone Technique through Binaural Signal Analysis of Loudspeaker-Reproduced Piano Performances. In Proc. Audio Engineering Society 121st Int. Conv., San Francisco, CA.

ACKNOWLEDGEMENTS

This work was supported in part by the Social Sciences and Humanities Research Council of Canada (SSHRC), within the project entitled "*Recording and auralization of the acoustics of concert halls and sonic spaces*." Additional support was provided by the Canada Foundation for Innovation (CFI) New Opportunities Program, and also by McGill University's Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT).