Investigations of the Effect of Surround Microphone Setup on Room Perception

Marie-Josefin Meindl[1], Ulrich Vette[1], Thomas Görne[2]

[1] *University of Music and Performing Arts, Vienna, Austria*

[2] *University of Music, Erich Thienhaus Institute, Detmold, Germany*

Correspondence should be addressed to Marie-Josefin Meindl (`Marie-Josefin@web.de`)

ABSTRACT

A popular approach for multichannel recordings of classical music is to reproduce the stage via the front channels, while the rear channels are limited to the reproduction of early reflections and diffuse sound. This paper investigates the influence of systematic variations of a spaced pair of surround microphones on the spatial impression of the reproduced sound. The investigations include variations of the distance between front and rear microphones and of the spacing of the capsules, as well as variations of the respective polar patterns.

1. INTRODUCTION

Investigations on multichannel recording have been performed for more than 25 years, and many proposals for microphone placement have been published during this time. Popular microphone setups for 5.1 surround consist of five microphones dedicated to L, C, R, LS and RS channels like INA-5, OCT, or the arrays proposed by Williams. Others treat the rear channels separately like Hamasaki Square or IRT Cross (see e.g. [1, 2, 3]).

Disregarding optimized five–channel arrays and disregarding special rear channel setups, many recording engineers treat surround recordings rather pragmatically. Confronted with needs like recording stereo and 5.1 simultaneously, they utilize a stereo setup, supplemented with surround microphones to feed the LS and RS channels.

Aim of this research was to focus on this rather practical approach, and to give an overview of the effects of surround microphone polar patterns and placement on the perceived sound. The results may help the balance engineer with surround recordings when adjusting the spatial microphones in addition to a running stereo setup.

2. SETTING UP SURROUND MICROPHONES

One might expect that the choice of surround microphone polar patterns and placement strongly influences the quality of reproduced sound. But different considerations on setting up surround microphones may lead to contradictory conclusions, as a short view in the literature shows:

According to B. Slotte, direct sound in the surround channels produces blurring and inhomogeneity in surround imaging. He concludes:

> *“The best way to avoid direct sound from the front is to use rear-facing cardioids, but also laterally oriented figure-8’s [...] can work quite well in this respect.”* [4]

On the other hand, the sense of spaciousness is dependent on frequency, as J.M. Potter et al. showed in experiments with different narrow-band and octave-band noise stimuli:

> *“The results of all three experiments show that spaciousness decreases for increasing frequency”* [5]

Thus omnidirectional microphones may be useful for picking up low frequency content and therewith increase spaciousness, but at the same time may cause problems in imaging.

D. Griesinger states that:

> *“[...] to provide a realistic impression of the hall over a large area, the reverberation in the*

> *front left and right speakers should be uncorrelated [...]. Furthermore the reverberation in the rear speakers should be uncorrelated, and at the same time it must be uncorrelated with the front speakers.”* [6]

To avoid correlation of the diffuse sound, the microphone distance must be large in comparison to the critical distance (Hallradius) in the recording venue.

J. Wuttke however fears:

> *“Large distances betweeen the microphones, however, cause special problems [...]. In the worst case one hears an undifferentiated cloud of sound sources looming around each loudspeaker without forming a coherent sterephonic image; there is no ‘envelopment’.” [7]*

G. Martin even suggests:

> *“The seperation between front and surround microphone arrays should not exceed approximately 70cm to maintain front / back cohesion.” [8]*

He suggests however creating decorrelation with other methods like mirophone angle.

In this work, an experiment was carried out to investigate surround setups of two microphones feeding LS and RS channels with variations of microphone positioning and polar patterns.

3. RECORDING

Recordings for this experiment were made in Vienna in the Mozartsaal of the Konzerthaus, which is said to be one of the most beautiful halls for staging and recording chamber music. This medium sized music hall (704 seats) has a mean reverberation time of $RT60 \approx 1{,}9\,\mathrm{s}$ without an audience (frequency dependent reverberation time, measured as $RT20$, is depicted in Fig. 1 [11]) and a volume of $\approx 4410\,\mathrm{m}^3$. Critical distance can be estimated as $0.057\sqrt{V/RT60} \approx 2{,}8\,\mathrm{m}$. The hall has large lateral and rear balconies which influence the early lateral reflection patterns significantly in delay and intensity and thereby intensify the spatial impression [9, 10].

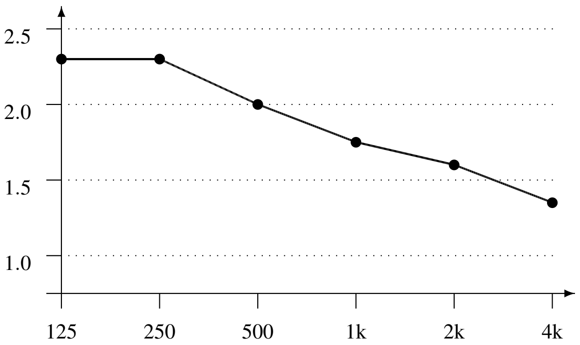

Fig. 1: Reverberation Time RT20

As sound sources were chosen:

- male speech
- solo violin
- small orchestra

In order to allow the subjects during the listening test to concentrate on the sound quality rather than on the text, the speaker recited Bulgarian text. Violin and orchestra were chosen to cover both small single sound sources and a bigger ensemble, whereas orchestra as a bigger ensemble is said to show the influence of lateral reflections more explicitly [10]. A violin music excerpt with rather long breaks between single notes was chosen to render the impression of envelopment in the “spaces between foreground events”, according to Griesinger [6].

The front L/R channels were recorded with an ORTF stereo array, using Sennheiser MKH 800. Distance and height were adjusted subjectively for the various sound sources. The center channel was ignored in order to prevent any questions concerning this wide (and interesting) area of surround recording.

The arrays under investigation consisted of 12 spaced pairs, located at 4 different positions, i.e. two positions on a front plane 1 m behind the ORTF array (spacing 0.6 m and 5.2 m), and another two with the same spacing 7.8 m behind the ORTF array on a back plane (Fig. 2).

Each surround microphone location was equipped with omnidirectional, cardioid and bidirectional microphones, consisting of Sennheiser MKH 20, 40 and 30. The cardioids and bidirectionals were arranged so that the angle of maximum rejection pointed towards the source to

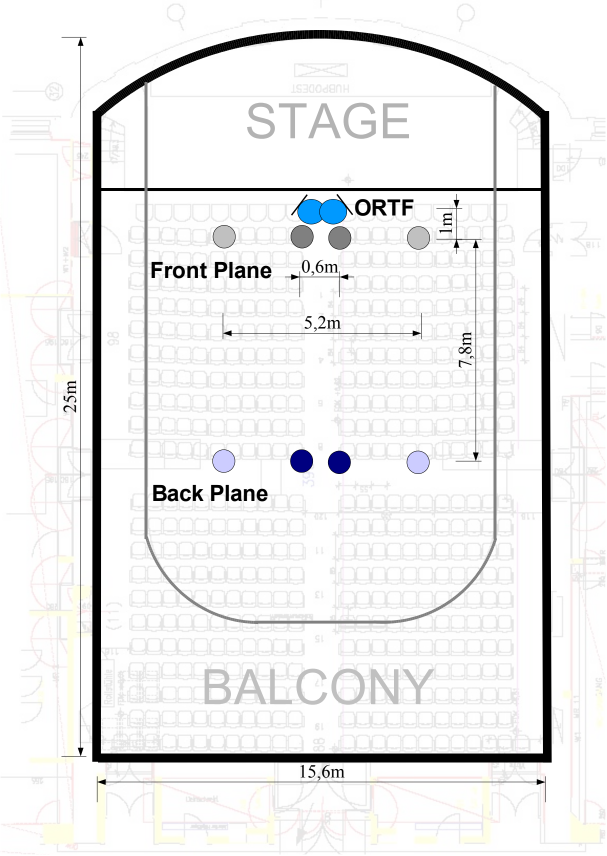

Fig. 2: Microphone Setups, Mozartsaal of the Konzerthaus, Vienna

mask the direct sound from the stage as recommended in [4]. All surround-microphones were at a height of 3.3 m.

Recording levels (same level for 0° sound incidence on the respective capsule on all recording tracks) were adjusted utilizing an acoustic noise generator [12].

During the editing process the ORTF signal was routed to the front L/R channels. The spaced pairs were each routed to LS and RS channels which led to a total of 12 investigated setups, each combined with the same front array. Center and LFE channels were not used.

Three Tonmeister students were asked to adjust the surround levels to ensure that the surround signals had the same subjective level, despite differing directivity indices of polar patterns and of differing direct sound levels due to microphone positioning. As the differences between the three values didn't exceed 1.5 dB, the mean was taken as the resulting value.

4. EXPERIMENT

For subjective evaluation of the recorded samples, a randomized blind test was performed according to recommendations on psychological research [13]. The test took place with students of Tonmeister and Music at Detmold University of Music, guided by a test administrator.

4.1. Sound Reproduction Conditions

The test took place in the standard ITU 775-1 listening room [14] of Erich Thienhaus Institute at Detmold University of Music, equipped with Geithain RL 901K monitor loudspeakers. Subjects were seated exclusively in the sweet spot. All stimuli were played at the same reproduction level for all tests, except for one, where the subject requested a 2 dB reduction of the orchestra stimuli to feel comfortable.

4.2. Test Design

To check the significance and practicability of the design, a pretest with 9 Tonmeister students was carried out. Following the questionnaire the test administrator had a short talk with each subject to hear their problems, suggestions and experiences during the test. Furthermore, duration and annoyance of the experiment were discussed.

As result of the pretest, the main experiment was modified somewhat in the wording of one parameter, and the training phase was augmented from 3 stimuli for each sound sample at the beginning of the test, to each 5 stimuli just before the corresponding sound sample.

In the questionnaire a direct rating method with continuous scale was employed, utilizing adjective pairs (two extreme terms for a single parameter) [13].

The parameters and adjective pairs were chosen inspired by the descriptive language for spatial sound reproduction developed by J. Berg and F. Rumsey using the "Repertory Grid Technique" [15], as well as by recommendations in ITU-R BS.1284-1 [16]. The questionnaire language was German.

The ITU Recommendation was followed as closely as possible in terms of test realisation, e.g. like choice and duration of stimuli, or number of subjects.

As stimuli were chosen short, coherent excerpts of approx. 10 sec from each of the three sources (speech, violin, orchestra). Test duration was one hour including a

Fig. 3: Microphone Test Setup in the Mozartsaal

break. For each subject dice were thrown to select one of six specific test versions, varying in stimuli and sample order. In every test, one microphone position was repeated to allow assessment of the subject's reliability.

4.3. Rating Process

The participants had to rate the following 7 parameters:

1. Room Size (*Raumgöße*)
   How large does the room seem to be?
   rather small – rather large (*eher klein – eher groß*)
2. Spatial Distribution of Reverberation[1] (*Räumliche Verteilung des Halls*)
   How is the distribution of reverberation, considering the different room directions? uneven - even (*ungleichmäßig – gleichmäßig*)
3. Room Envelopment (*Räumliche Umhüllung*)
   How strong is the envelopment of sound?
   weak – strong (*schwach – stark*)
4. Front Image Sharpness (*Frontale Abbildung*)
   How sharp is the image of the individual instruments?
   muddy – clear (*verwaschen – klar*)
5. Apparent Source Width (*Breite der Quelle*)
   How wide is the whole sound source on stage?
   narrow – wide (*schmal – breit*)
6. Timbre (*Klangfarbe*)
   How is your overall impression of timbre?
   rather dark – rather light (*eher dunkel – eher hell*)
7. Preference (*Präferenz*)
   Do you like the sound as a whole?
   don't like – like very much (*gefällt nicht – gefällt sehr*)

The subjects were allowed to listen to the looped stimulus for one minute. In most cases completing the questionnaire for 7 parameters did not exceed 45 sec.

4.4. Subjects

A total of 29 subjects participated in the experiment. 24 of them were Tonmeister students, 5 Music students. 9 of them only participated in the pretest (with its different wording for one parameter). 2 of the completed questionnaires were excluded because of poor performance in recognizing the repeated stimuli.
This gives results for 27 participants for 6 parameters, and for 18 participants for the parameter that was changed.

4.5. Data Quantization

As continuous scales were employed, the rating had to be quantized to allow statistical interpretation. For data quantization the graphical rating was converted into an integer number from 0 to 10, where 0 corresponds to one of the extreme adjectives (the left) and 10 to the other

---

[1] wording changed after the pretest

(the right). The intermediate values were equally distributed in between.

4.6. Selection of Data

All parameters except Spatial Distribution (2) and Apparent Source Width (5) gave highly significant results in ANOVA [17]. Consequently, (2) and (5) were disregarded, and further analysis concentrated on the five parameters Room Size, Envelopment, Front Image, Timbre and Preference.

5. RESULTS

Following abbreviations will be used in the diagrams:

omnidirectional microphone	O
cardioid microphone	C
bidirectional microphone	8
narrow spaced pair	n
wide spaced pair	w
front plane	f
rear plane	r
example:	
omnis, rear plane, narrow spaced	*O rn*

5.1. Presented Diagrams

For the statistical analyses cumulative frequencies were looked at first. For further interpretations means with 95% confidence intevals were calculated, whereof the results of the Orchestra sample can be found at the end of this document.

However, in the following diagrams direct comparisons between the 12 microphone setups can be made more understandable than with error bars. Displayed are the means of all parameters, where the values of the diagrams correspond to the quantization of subject ratings: the low value inside the center corresponds to rather small room size, weak envelopment, muddy front image, rather dark timbre and less preference. The high value at the outside border corresponds to rather large room size, strong envelopment, clear front image, rather light timbre and great preference.

5.2. Differences between Sound Samples

Some changes in microphone setup influence the sound samples in completely different ways and some effect exactely the same changes for all sound samples.

Variations of polar patterns for example can't be generalised at all for different instruments.

But for example moving the widely spaced bidirectional

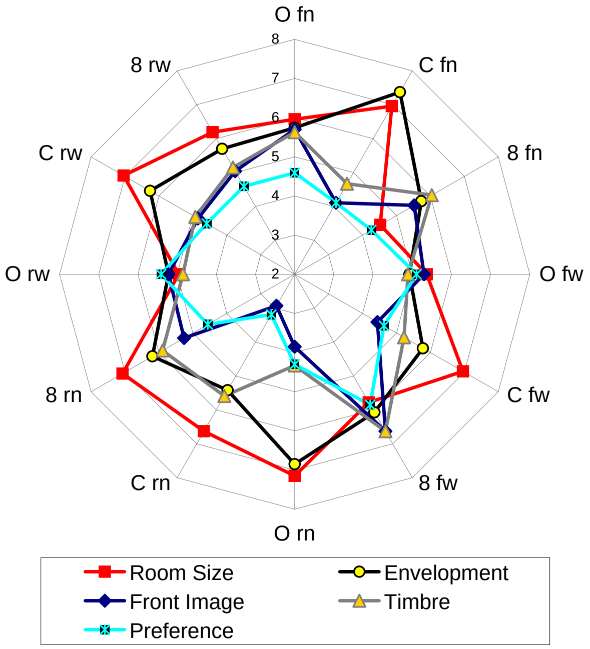

Fig. 4: Overview: Comparison of Means - Orchestra

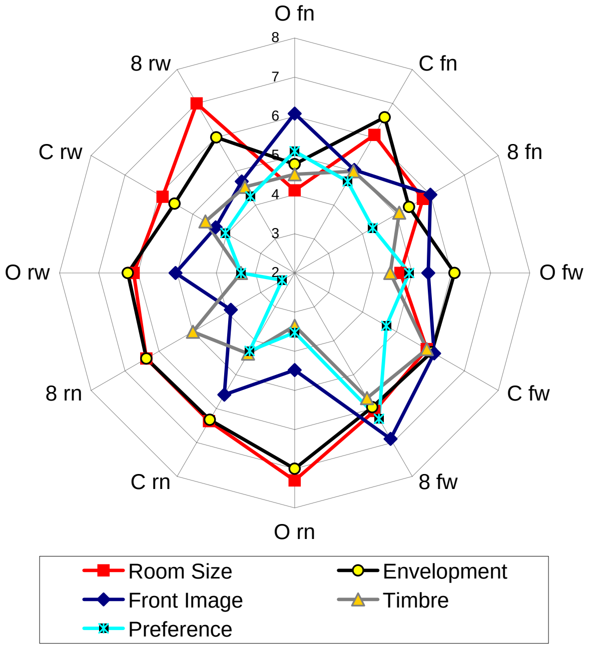

Fig. 5: Overview: Comparison of Means - Speech

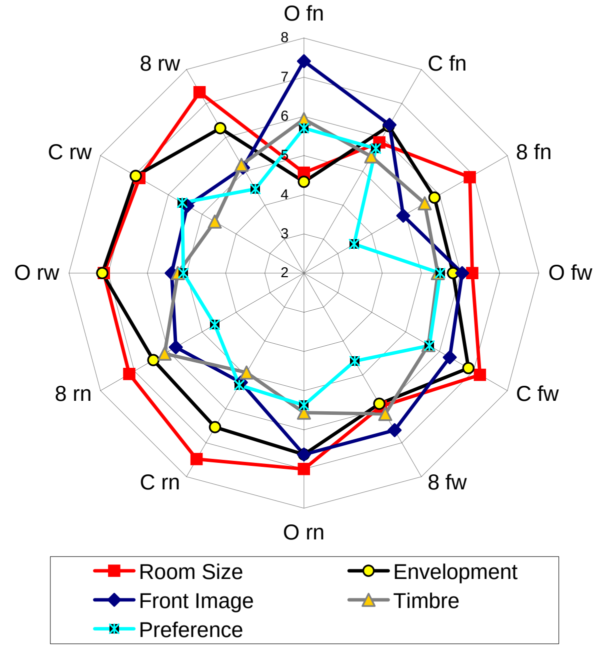

Fig. 6: Overview: Comparision of Means - Violin

microphones from the front plane to the rear plane gives exactly the same types of perceived changes for all sound sources.
Sometimes the results of the orchestra samples differ from the results of speech and violin samples, this may be caused by its differing width.

5.3. Differences between Polar Patterns

The biggest differences between omnidirectionals, cardioids and bidirectionals for the parameters room size and envelopment were found at the front positions. Cardioids and bidirectionals seem to create both a bigger room and stronger envelopment for the single sound sources (see Fig.7). With narrow spacing for the orchestra sample however, omnidirectionals seem to make the room bigger than bidirectionals.

Overall, bidirectionals tend to be preferred and are perceived as giving a clearer image for the orchestra samples and parts of the speech samples (e.g. Fig. 8).
However, for violin samples at the front narrow position (see Fig. 7), omnidirectionals seem to create a sharper image and, along with cardioids, are preferred when compared with a bidirectional microphone.
Statements from Slotte [4] as noted in section 2 can therefore be confirmed only for parts of this experiment.

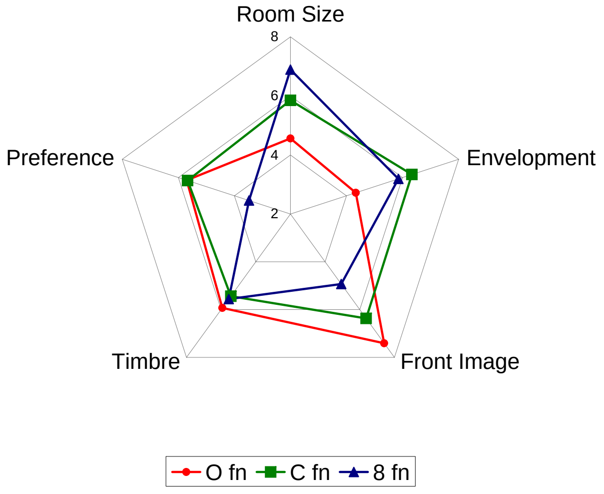

Fig. 7: Violin - Patterns. Omnis, cardioids, fig8s; narrow spacing, front plane (O fn, C fn, 8 fn)

As expected, omnidirectionals create a darker timbre impression overall. The influence of low frequency content on spaciousness however, weren't answered in this experiment.

Interesting is the fact that for the narrow spaced rear setup, the effects of bidirectional and cardioid mirophones for the speech sample appear to be inverted when compared to the orchestra sample. (see Fig. 8 and 9)

5.4. Differences in Position

It can be stated that aside from a few exceptions, the positions with more distance to the stage create a comperatively muddier front image (especially with narrow spacing), althought direct sound is less of a problem then at the front positions (see Fig. 10, 11, 12).

Also at the rear plane, contrary to expectations of Wuttke and Martin [7, 8] (see section 2), omnidirectionals and bidirectionals seem to create more envelopment and room size than at the front plane.
Althought cardioids show fewer differences between the planes at these parameters, the orchestra sample with the front narrow setup gives the best envelopment of all. This, on the other hand, agrees with statements by Martin [8] about front / back cohesion as mentioned in section 2.

The wide spaced setups were preferred, even though

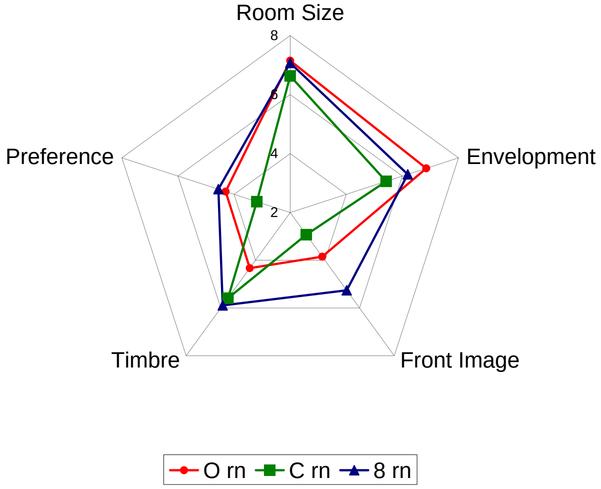

Fig. 8: Orchestra - Patterns. Omnis, cardioids, fig8s; narrow spacing, rear plane (O rn, C rn, 8 rn)

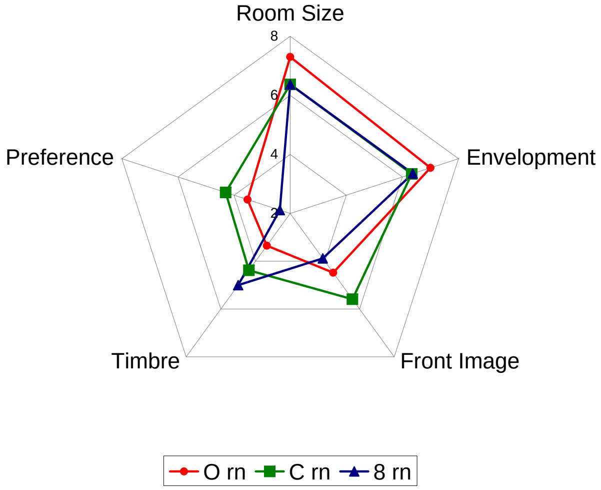

Fig. 9: Speech - Patterns. Omnis, cardioids, fig8s; narrow spacing, rear plane (O rn, C rn, 8 rn)

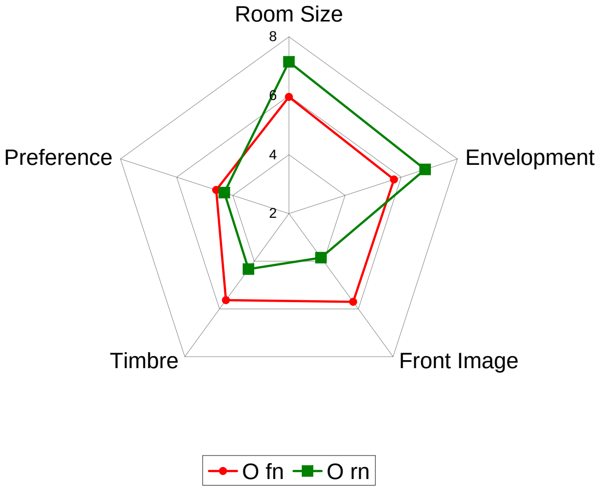

Fig. 10: Orchestra - Distance. Omnis, narrow spacing, front and rear plane (O fn, O rn)

differences in preference aren't big, with the exception that the bidirectional microphones yielded a much higher mean value for the front wide spaced setup (Fig. 13).

6. CONCLUSIONS

In this paper 12 variations of microphone positioning and polar patterns were investigated regarding their influence on room and image perception.

Remarkable are the significant differences in Front Image Sharpeness (4). Bidirectional microphones seem to produce a sharper image than other polar patterns - especially in the front wide spaced position. The other results partly agree with theoretical statements in respect to Room Size (1), Timbre (6) and Envelopment (3). Less significant results in Preference (7) may reflect the subjectivity in perception of surround sound. It is nevertheless worth mentioning that the results for preference correlate to a certain degree with those for front image sharpeness.

Perhaps the results of this experiment will be useful to audio engineers as a guideline when creating their own surround sound recording setup. However, again it must be said clearly that these results are in detail only valid for the Mozartsaal of the Konzerthaus, Vienna.

To find more exact results it is first necessary to design

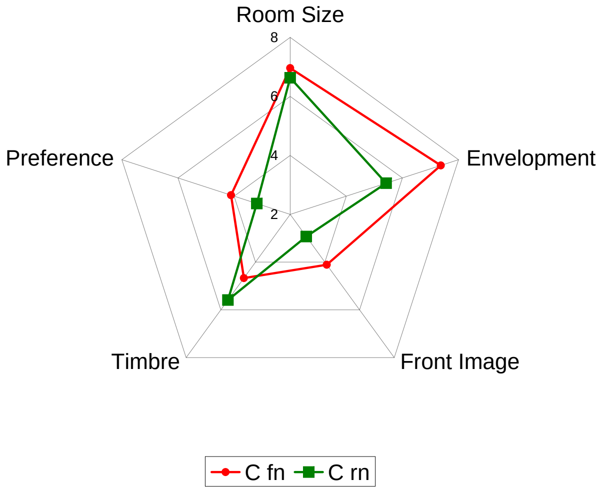

Fig. 11: Orchestra - Distance. Cardioids, narrow spacing, front and rear plane (C fn, C rn)

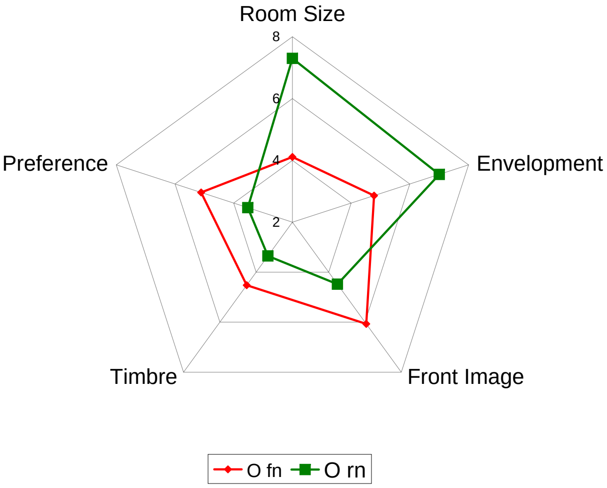

Fig. 12: Speech - Distance. Omnis, narrow spacing, front and rear plane (O fn, O rn)

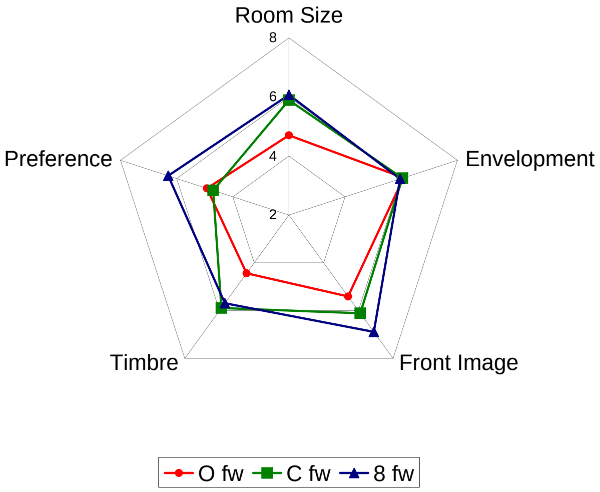

Fig. 13: Violin - Patterns. Omnis, cardioids, fig8s; wide spacing, front plane (O fw, C fw, 8 fw)

several detailed listening tests for parts of the 12 setups and second to repeat the recordings with the same setup in other rooms.
Furthermore it would be very interesting to continue this experiment with different combinations of 4 microphones. The recorded material from this reserch can be provided for further investigations.

7. ACKNOWLEDGEMENT

The authors would like to thank Heinz Repper and the management of the Konzerthaus Vienna, Josef Schütz from ORF Vienna, as well as Sennheiser electronic KG and Jürgen Peissig for their kind support.

8. REFERENCES

[1] F. Rumsey: *Spatial Audio*, Focal Press, Oxford, 2001.

[2] G. Theile: “Natural 5.1 Music Recording Based on Psychoacoustic Principles,” 19th Int. Conf. Audio Eng. Soc. (Schloss Elmau), 2001.

[3] M. Williams: *Microphone Arrays for Stereo and Multichannel Sound Recordings* Vol. 1, Editrice Il Rostro, Segrate, 2004.

[4] B. Slotte: “Sharpening the image in 5.1 surround recording,” Prepr. 6509 118th Conv. Audio Eng. Soc., Barcelona, 2005.

[5] J.M. Potter, F.A. Bilsen, J. Raatgever: “Frequency Dependence of Spaciousness,” *Acta Acustica / Acustica* vol. 3, **5**, pp. 417–427, 1995.

[6] D. Griesinger: “The Psychoacoustics of Listening Area, Depth, and Envelopment in Surround Recordings, and their Relationship to Microphone Technique,” 19th Int. Conf. Audio Eng. Soc. (Schloss Elmau), 2001.

[7] J. Wuttke: “General Considerations on Audio Multi-Channel Recording” 19th Int. Conf. Audio Eng. Soc. (Schloss Elmau), 2001.

[8] G. Martin: “A New Microphone Technique for Five-Channel Recording”, Prepr. 6427 118th Conv. Audio Eng. Soc., Barcelona, 2005.

[9] L. Beranek: *Concert Halls and Opera Hauses: Music, Acoustics, and Architecture*, Springer New York Berlin Heidelberg, 2nd Ed. 2004.

[10] W. Kuhl: “Räumlichkeit als Komponente des Raumeindrucks,” *Acustica* vol. 40, pp. 167–181, 1978.

[11] K. Müller: Konzerthaus Wien Mozart-Saal Raumakustische Messungen, Müller-BBM, Bericht Nr. 22 965/19, 2002.

[12] J. Wallbrecher, O. Gelhard, J. Peissig: “Entwicklung einer handgehaltenen Signalquelle für Prüf- und Messzwecke,” Deutsche Jahrestagung für Akustik DAGA’06 (Braunschweig), 2006.

[13] J. Bortz, N. Döring: *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*, 3rd run, Springer-Verlag Berlin Heidelberg New York, 2003.

[14] ITU Recommendation ITU-R BS.775-1: “Multichannel stereophonic sound system with and without accompanying picture,” Intern. Telecom. Union Radiocom. Genf 1992-1994

[15] J. Berg: *Systematic Evaluation of Perceived Spatial Quality in Surround Sound Systems*, Doctoral Thesis, Luleå University of Technology, 2002.

[16] ITU Recommendation ITU-R BS.1284-1: “General methods for the subjective assessment of sound quality,” Intern. Telecom. Union Radiocom. Sector 1997–2003.

[17] A. Bühl, P. Zöfel: *SPSS 12 - Einführung in die moderne Datenanalyse unter Windows*, 9th run, Pearson Education Germany, 2005.

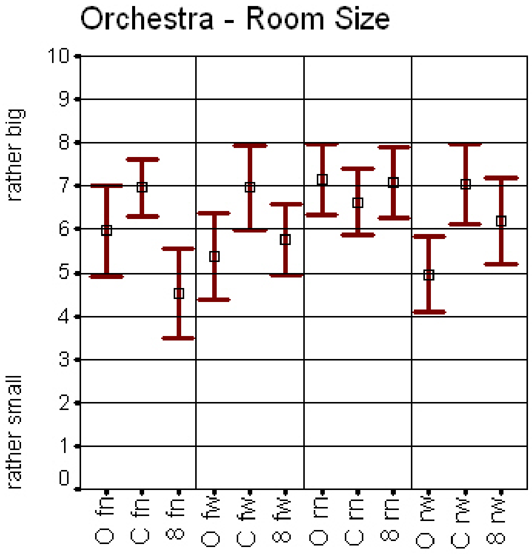

Fig. 14: Mean and 95% Confidence Interval

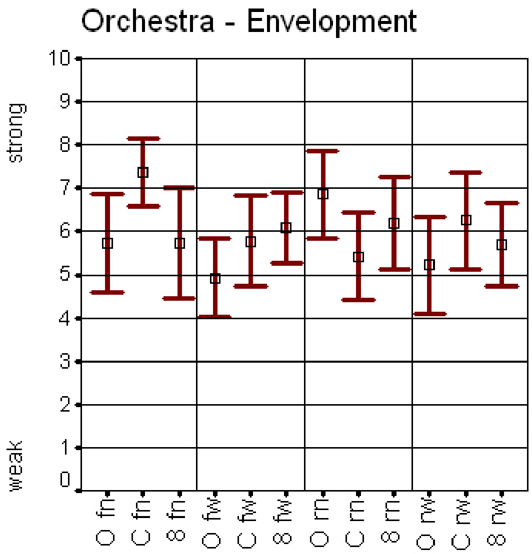

Fig. 15: Mean and 95% Confidence Interval

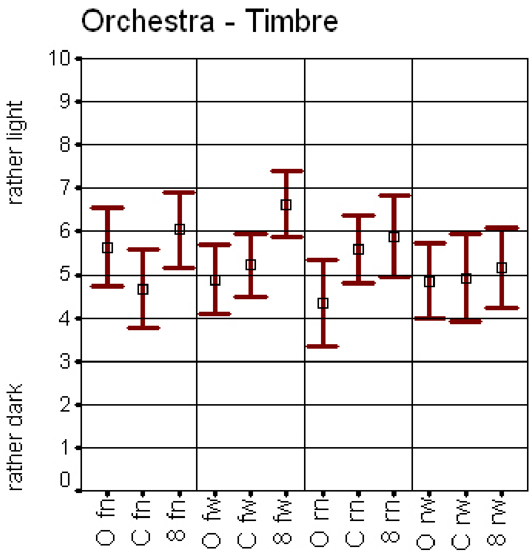

Fig. 17: Mean and 95% Confidence Interval

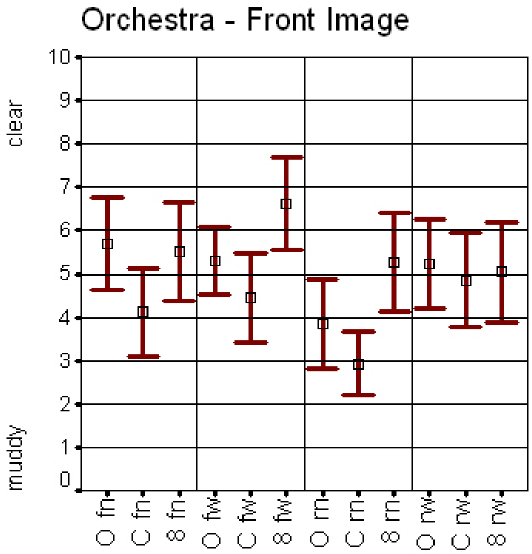

Fig. 16: Mean and 95% Confidence Interval

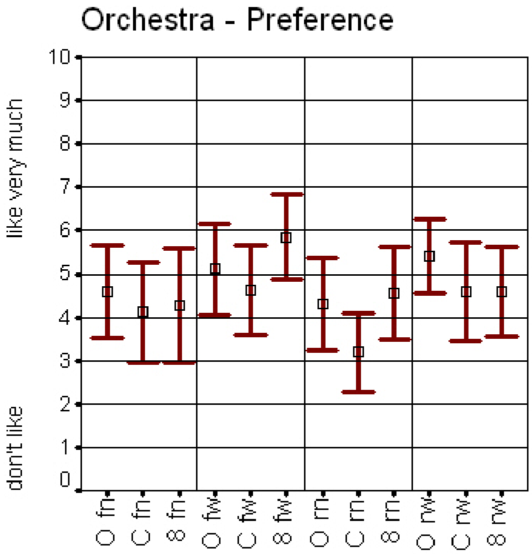

Fig. 18: Mean and 95% Confidence Interval