Proceedings of the 9th International Conference of Students of Systematic Musicology (SysMus16), Jyväskylä, Finland, 8th - 10th June 2016. Birgitta Burger, Joshua Bamford, & Emily Carlson (Eds.).

How Important is the Reproduction Technique for the Perception of Spaciousness in Music?

Claudia Stirnat[1]

[1]Institute of Systematic Musicology, University of Hamburg, Germany
Claudia.stirnat@uni-hamburg.de

Abstract

Sound fields are preferred when spaciousness is perceived. "Spaciousness means that auditory events, in a characteristic way, are themselves perceived as being spread out in an extended region of space". The author's former study (2012) investigated the perception of spaciousness of five musical genres. Participants evaluated spaciousness while listening to music with loudspeakers. Resulting, these genres allowed classifications of spatial features such as "big", "wide" and "open". A new idea of conducting listening tests has arisen using wave field synthesis instead of headphones. Firstly, this study aims to reveal the perceptional characteristics of spaciousness in music itself with headphones, loudspeakers and wave field synthesis. Secondly, it will find the differences of perception when participants listen with headphones, loudspeakers and wave field synthesis. Thirdly, this study's goal is to investigate the possibility of replacing headphones with wave field synthesis for listening tests. A hearing test has been conducted asking 28 participants for their spacious impression. The Participants rated 30 music excerpts on a 7 Point-Likert-Scale from "little spacious" to "much spacious". Dummy head measurements for an objective comparison were made. 3x3 ANOVA repeated measures revealed a significant Within-Subject effect for the technical devices ($F = 4.541$, $p < 0.05$), different instrument groups ($F = 71.281$, $p < 0.01$) and also for the interaction technical devices-instrument groups ($F = 7.700$, $p < 0.01$). The kind and number of music instruments on the one hand and the reproduction technique on the other hand influences the perception of spaciousness.

Keywords: Music Perception, Psychoacoustics, Spaciousness, Acoustics, Wave Field Synthesis

Introduction

Listeners prefer sound fields when spaciousness is perceived (Blauert and Lindemann, 1986). "Spaciousness means that auditory events, in a characteristic way, are themselves perceived as being spread out in an extended region of space" (Blauert 1997, p. 348). According to Winckel (1970), spacious sensation occurs when timbre features a high ratio of low frequently energy. A dependence between sound level and spaciousness has been found but has not been "investigated quantitatively enough to be included in the index of spaciousness" (Blauert 1997, p. 355).

The spaces experienced while listening to music are distinguished by Wellek (1982). He separates the physical space from psychological space. Blauert (1997) introduces a model about the auditory information processing of spatial hearing. It consists of three different aspects: the physical, psychophysical, and the psychological aspects.

The author's former study (2012) investigated the perception of spaciousness of the musical genres Classic, Ethno, Electro, Jazz and Rock. Participants evaluated spaciousness represented by twelve adjectives while listening to 30 music excerpts for each genre with loudspeakers. Resulting, these genres allowed classifications of spatial features such as "big", "wide" and "open" (Stirnat, 2012).

Loudness as a related subjective parameter of sound level is a very subjective size differing throughout participants. Von Ruschkowski (2013) found a significant difference in the loudness perception among others between male and female participants. As size and shape of a head and ears vary between participants, every participant has a different head related transfer function (HRTF), which "characterizes how a particular ear (left or right) receives a sound from a point in space" (Potisk 2013).

When conducting studies with technical devices, the technical features have to be considered. They give sounds a certain characteristic and quality. If the quality is low,

distortions appear quickly for example. Thus, features as the driver, the frequency response and the radiation are important cues (Kleiner, 2012). A complex frequency response describes the response characteristics of the amplitude along the frequency spectrum and phase (Friesecke, 2007). Headphones measurements have shown that the way of wearing headphones influences the frequency response up to approximately 20dB especially for low frequencies (Kleiner, 2012).
In order to avoid this issue, a new idea of conducting listening tests has arisen using wave field synthesis instead of headphones (e.g. Laumann, Theile and Fastl (2008)).

Firstly, this study aims to reveal the perceptional characteristics of spaciousness in music itself with headphones, loudspeakers and wave field synthesis. Secondly, it will find the differences of perception when participants listen with headphones, loudspeakers and wave field synthesis. Thirdly, this study's goal is to investigate the possibility of replacing headphones with wave field synthesis for listening tests.
Therefore, the following hypotheses have been formulated:

- Music specific characteristics will occur referring to the perceived spaciousness.
- The headphones, loudspeakers and wave field synthesis will be perceived differently and reveal specific characteristics for the used technical devices.

Experiment

Methods

A listening test was conducted at the wave field synthesis laboratory at Hamburg University of Applied Science (HAW) in July 2015. Participants were asked to rate how spacious they perceived music excerpts on a 7-Likert-Point-Scale.

Participants

28 participants took part in this study. The data of one participant had to be excluded because a technical problem occurred during one session. The other 27 participants were 19 to 62 years old (mean age = 30,59 years, SD = 12.87). According to their answers, 18 participants had had experiences in hearing tests and 26 were normal listeners without any diagnosed hearing impairment. 21 participants had attended musical training at least at amateur level.

Stimuli

30 music excerpts of various anechoic recorded instruments were used. Each music excerpt was 15 seconds and was played once for each stimulus with 5 seconds silence (15s music – 5s silence). The stimuli were either self-recorded in the anechoic chamber of the Institute of Systematic Musicology or elsewhere recorded in an anechoic environment (Bernschütz et. al, 2012; Bang & Olufsen, 1992). When necessary they were faded out at the end of the stimulus. Various kinds of instruments were chosen for a variety of stimuli: guitar, banjo, flute, oboe, e-piano, violin, cello, accordion, vocals, trumpet, and brass. They were played either solo, in a duet or in an ensemble.

Setup

Each participant sat at the same listening position for all three listening conditions (see *Figure 1*). They listened through Beyerdynamic DT100 2x400Ω headphones, two professional A8X loudspeakers and Fouraudio Wave Field Synthesis Model 28-243 with 26 (loudspeaker) modules including 26 loudspeakers for each module (more information in Fohl, 2013; Nogalski, 2012). The Wave Field Synthesis modules surrounded the listener position (top row in Picture 1 and the loudspeakers were located in front of it.

Figure 1. The listening test's setup

A tracking system was used for the wave field synthesis condition so that the music was heard near the grey sensors above the head (see

Figure 2). The participants wore it as shown on the picture. Infrared cameras detected the sensors so that the music excerpts were reproduced by the loudspeaker modules as set in the software "xWonder". Here, a headphone-alike setup was chosen that made the music excerpts being heard near the grey sensors above the head.

Figure 2. The tracking system used in the wave field synthesis condition on the dummy head.

Procedure

At the beginning the participants were introduced to the experiment and were asked to answer some background questions about e.g. sex, age, experiences with listening tests. Before the start of the listening test they could ask any questions in case something was unclear. Their task was to rate music excerpts on how spacious (from 1="little spacious" to 7="very spacious") they perceived them.

The participants listened to all 30 music excerpts in three conditions in a random order. One third of the participants started with the headphones condition, one third of the participants heard through wave field synthesis first and the others began with the loudspeakers condition. The 30 Music excerpts were played in random order for each condition as well. The order within one condition remained the same for all participants.

The first two music excerpts were always pre-testing excerpts so that the participants could get used to the task and the listening conditions. The volume level was set on a quite equal level which was tested by listening as there was no absolute decibel display. The whole experiment took between 30 and 45minutes. As a little reward the participants received a small gift for their participation.

Data Analysis

The music excerpts were categorized into groups by their number and kind of musical instruments in order to decrease the α-Error by reducing the data. Therefore, similar music excerpts were summarized into the groups *wind instruments with one voice, stringed instruments with one voice* and *mixed instruments with several voices*. 3x3 ANOVA-repeated measures was carried out in order to analyze the data in respect to the hypotheses. Two factors with three levels each were chosen: *TechnicalDevices* and *InstrumentGroups* Bonferroni was used as adjustment for multiple comparison. Additionally, the stimuli's frequency spectrum and sound level was checked using the software Audacity.

Results

The overall means by condition are shown in *Figure 3*. The headphones condition reveals a mean value $\bar{x} \approx 4.08$, the wavefield synthesis condition a mean value $\bar{x} \approx 4.49$ and the loudspeakers condition a mean value $\bar{x} = 4.191$.

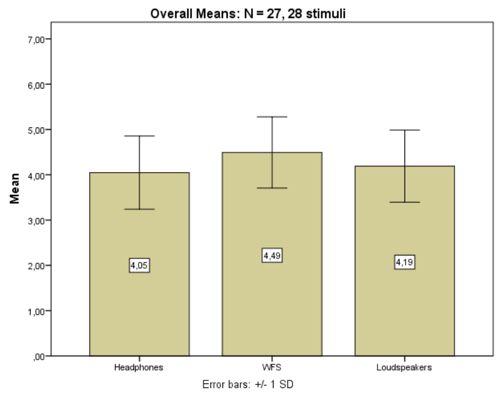

Figure 3. Overall means by condition for spacious perception of 28stimuli for N=27.

This result shows that the wave field synthesis condition was rated the highest, the headphones condition the lowest and the loudspeaker condition in between. The mean values of each stimulus point to the same tendency. The mean value of the wave field synthesis condition is higher than the mean value of the headphones condition in 23 cases.

The tendency was checked for significance analysing the data with 3x3 ANOVA-repeated measures. Therefore, three instrument groups were used:

Wind instruments – one voice:
In *Table 1*, the headphones condition reveals the lowest mean value $\bar{x} \approx 2.96$ with a standard deviation $SD \approx 0.92$, followed by the loudspeakers condition with a mean value $\bar{x} \approx 3.42$ and a standard deviation $SD \approx 1.03$. The highest mean value has the wave field synthesis condition with $\bar{x} = 3.75$ and a standard deviation $SD \approx 1.03$.

Descriptive Statistics

	Mean	Std. Deviation	N
Headphones_WOV	2,9636	,91669	27
Wave field synthesis_WOV	3,7500	1,07473	27
Loudspeakers_WOV	3,4213	1,03145	27

Table 1. Mean values and standard deviation of spacious perception for wind instruments one voice (WOV) by condition.

Stringed instruments – one voice:
Here, the mean values are higher than those of the wind instruments (see *Table 2*). The loudspeakers condition shows the lowest mean value $\bar{x} \approx 4.06$ with a standard deviation $SD \approx 0.93$. The headphones condition has a mean value $\bar{x} \approx 4.11$ in the middle of the other conditions with a standard deviation $SD \approx 1.05$. Again, the wave field synthesis condition reveals the highest mean value $\bar{x} \approx 4.57$ with a standard deviation $SD \approx 1.01$.

Descriptive Statistics

	Mean	Std. Deviation	N
Headphones_SOV	4,1173	1,04984	27
Wave field synthesis_SOV	4,5679	1,01196	27
Loudspeakers_SOV	4,0593	,92516	27

Table 2. Mean values and standard deviation of spacious perception for stringed instruments one voice (SOV) by condition.

Mixed instruments – several voices/polyphony:

These mean values are higher than those of the wind instruments and stringed instruments (see *Table 3*). The headphones condition reveals the lowest mean value $\bar{x} \approx 4.63$ with a standard deviation $SD \approx 0.94$, followed by the loudspeakers condition with a mean value $\bar{x} \approx 4.75$ and a standard deviation SD ≈ 0.79. The wave field synthesis condition has the highest mean value $\bar{x} \approx 4.82$ with a standard deviation $SD \approx 0.84$.

Descriptive Statistics

	Mean	Std. Deviation	N
Headphones_MISV.P	4,6337	,94019	27
Wave field synthesis_MISV.P	4,8189	,83669	27
Loudspeakers_MISV.P	4,7490	,79444	27

Table 3. Mean values and standard deviation of spacious perception for mixed instruments several voices/polyphony (MISV.P) by condition.

As the Mauchly test is not significant for *TechnicalDevices* and *InstrumentGroups* sphericity is assumed. The interaction *TechnicalDevices*InstrumentGroups* is significant with $p < .05$ so that sphericity is not assumed but Greenhouse-Geisser value ($\varepsilon < .75$). A significant Within-Subject effect was found for *TechnicalDevices* ($F = 4.541$, $p < .05$), *InstrumentGroups* ($F = 71.281$, $p < .01$) and also for the interaction *TechnicalDevices*InstrumentGroups* ($F = 7.700$, $p < .01$). Estimated Marginal means confirm this result and reveal the factor levels showing the effect. *Headphones* and *wave field synthesis* reveal a significant ($p < .05$) mean difference of d = .47. *InstrumentGroups* shows a significant effect for all three levels. *Wind Instruments-One Voice* and *Stringed Instruments-One Voice* have a significant ($p < .01$) mean difference of $d = 0.87$. *Wind Instruments-One Voice* and *Mixed Instruments-Several Voices* show a significant ($p < .01$) mean difference of d = 1.36. *Stringed Instruments-One Voice* and *Mixed Instruments-Several Voices* have significant ($p < .01$) mean difference of $d = 0.49$. *Figure 4* displays the Estimated Marginal means and makes the interaction between the variables obvious. The distance between the *InstrumentGroups* variables varies for all three conditions which indicates an interaction effect for all instrument groups and listening conditions (Janssen & Laatz, 2013). It means that the instrument groups have an impact on the perception in the listening condition and vice versa.

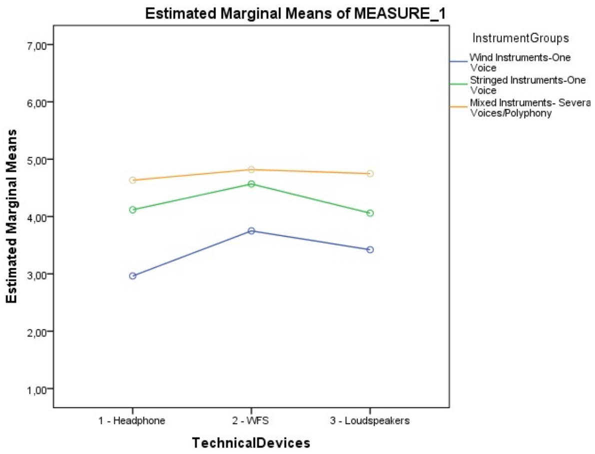

Figure 4. Estimated marginal means for the three instrument groups and the technical devices: blue line shows wind instruments, green line stringed instruments and orange line mixed instruments.

The frequency spectrum and sound level of the stimuli was checked using Audacity. A low amount of low frequency was seen for five of eight wind instrument stimuli. But the amount of low frequency was higher for stringed instrument stimuli and for mixed instruments stimuli. Comparing the overall sound levels of the stimuli with the average ratings revealed that the stimuli of the same instrumentation showed a link between sound level and rating if the level difference was at least 6dB. Thus, the stimuli with an at least 6dB higher sound level was rated more spacious. This link should be investigated more in the future.

Objective Measurements

Dummy head recordings were made in order to better understand and explain the results of the listening test. A dummy head is a normed head including two microphones at the positions of a head's ears. The one used for these measurements was a HEADacoustics dummy head.

Firstly, the reproduction accuracy of wave field synthesis' tracking system was tested with sinus waves in nine different positions (see Figure 5). From 62,5 Hz to 8 kHz sinus waves were played for 5sec. each with 1sec. silence between each frequency. The frequency was raised octave wise.

The overall sound pressure level shows the reproduction accuracy at a listener's position. The differences between both microphones of

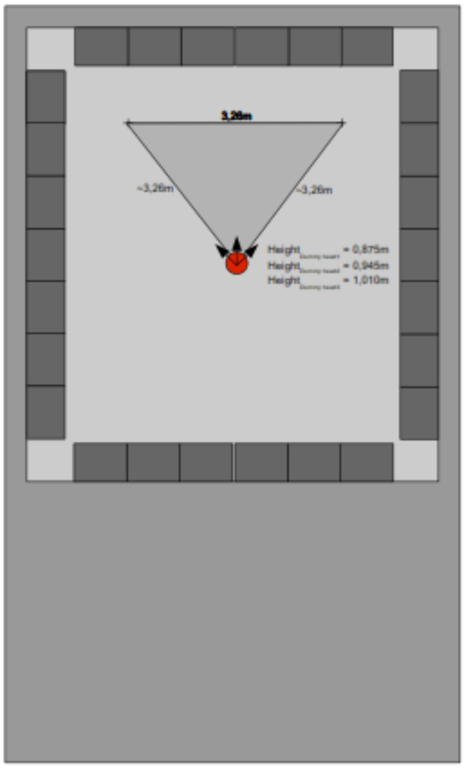

Figure 5. Draft of the measurement setup in the wave field synthesis laboratory using a HEADacoustics dummy head in nine different positions.

the dummy head range from 1,25 dB[V] up to 3,71 dB[V] within one of nine positions. Comparing the positions within on microphone the maximum difference for the left microphone is 3,93 dB[V] and for the right microphone 3,99 dB[V]. This result means that a difference in the sound level is audible for a listener and not exact the same within the listening area.

Another dummy head measurement has been done recording the music excerpt in the listener position for every condition but the data analysis is still in process and will be presented at the conference.

Discussion

The significant results of the listening test show that not only the reproduction technique is an important factor but also the kind of music used. On the one hand the wave field synthesis condition is perceived higher than the headphones condition. On the other hand characteristics for the instrumentation of the music were found. Several mixed instruments are perceived more spacious than single stringed instruments and single wind instruments are perceived even less spacious. These findings match with the analysis of the frequency spectrum quite well, as stimuli with a higher amount of low frequency were rated more spacious and confirm Winckel's (1970) result. Also the sound level analysis indicate a link between sound level and perception of spaciousness as Blauert (1997) wrote

Dummy head recordings of the wave field synthesis system revealed fluctuations of max. 3,99dB[V] in the SPL arriving at the listeners ear which are smaller than the differences of approximately 20dB in the frequency response of different types of wearing headphones by Kleiner (2012) as the dummy head recordings did not measure the whole frequency response of the wave field synthesis. But these results indicate a more accurate reproduction technique than headphones for investigating spaciousness and other research areas where the sound level matters.

Conclusions

The hypotheses were confirmed in this study and even an interaction effect between the technical devices and the instrument groups was found. The sound level has turned out to be an important factor for spaciousness and gives a reason for replacing headphones with wave field synthesis. The results just mentioned indicate that the reproduction technique is one cue among several others for the perception of spaciousness. Thus, it is important to use appropriate reproduction technique for the desired sound. But it is also important to consider that reproduction alone influences the perceived spaciousness as well as the kind of music. Both aspects have an impact on the perceived spaciousness, too, leading to a different perception than both aspects separately.

Acknowledgement

I would like to thank SEMPRE for supporting my attendance financially with the SEMPRE Conference Award.

References

Bang & Olufsen (1992). Music from Archimedes. Audio CD 101.

Bernschütz, B., Woirgard, M., Stade, P., Amankwor, J. (2012). Anechoic Recordings,

http://www.audiogroup.web.fh-koeln.de [03.08.2015].

Blauert, J. (1997). Spatial Hearing. The Psychophsysics of Human Sound Localization. Cambridge, London: The MIT Press.

Blauert, J., & Lindemann, W. (1986). Auditory spaciousness: Some further psychoacoustic analyses. In: *JASA 80/2*, 533 – 541.

Fohl, W. (2013). The Wave Field Synthesis Lab at the HAW Hamburg. In: Rolf Bader (Ed.), *Sound, Perception, Performance.* Switzerland: Springer International Publishing, 243 – 255.

Friesecke, A. (2007). Die Audio-Enzyklopädie. Ein Nachschlagewerk für Tontechniker. München: K.G. Saur Verlag.

Janssen, J., & Laatz, W. (2013). Statistische Daten- analyse mit SPSS. Eine anwendungs-orienterte Einführung in das Basissystem und das Modul Exakte Tests. Berlin, Heidelberg: Springer-Verlag[8].

Kleiner, M. (2012). Acoustics and Audio Technology. Fort Lauderdale, FL: Ross[3].

Laumann, K., Theile, G. & Fastl, H. (2008). A virtual headphone based on wave field synthesis. *Proceedings Acoustics 08*. Paris, France, 3593–3597.

Nogalski, M. (2012). Gestengesteuerte Positionierung von Klangquellen einer Wellenfeld-synthese-Anlage mit Hilfe eines kamerabasierten 3D-Tracking-Systems. Bacherlorthesis, Hamburg, https://users.informatik.haw-hamburg.de/~ubicomp/

arbeiten/bachelor/nogalski.pdf.

Potisk, T. (2013). Head-Related Transfer Function. Seminar paper, Ljubljana, 1 – 12,

http://mafija.fmf.uni-lj.si/seminar/files/2014_2015/

Seminar_Ia_Head-Related_Transfer_Function_Tilen_Potisk.pdf [28.08.15].

Stirnat, C. (2012). Percepted spaciousness of different musical genres. Unpublished Bachelor's Thesis, Hamburg.

Von Ruschkowski, A. (2013). Lautheit von Musik. Eine empirische Untersuchung zum Einfluss von Organismusvariablen auf die Lautstärkewahrnehmung von Musik. Dissertation, Hamburg.

Wellek, A. (1982). Musikpsychologie und Musikästethik.Grundriss der Systematischen Musik-wissenschaft. Bonn: Bouvier[3].

Winckel, F. (1970). Akustischer und visueller Raum. Mitgestalter in der experimentellen Musik. In: Fritz Winckel (Eds.), *Experimentelle Musik. Raum Musik. Visuelle Musik. Medien Musik. Wort Musik. Elektronik Musik. Computer Musik*, Berlin: Gebr. Mann Verlag.