Coloration in Wave Field Synthesis

Hagen Wierstorf[1], Christoph Hohnerlein[1], Sascha Spors[2], Alexander Raake[1]

[1]*Assessment of IP-based Applications, T-Labs, TU Berlin, Ernst-Reuter-Platz 7, 10587 Berlin, Germany*

[2]*Institut für Nachrichtentechnik, Universität Rostock, R.-Wagner-Str. 31, 18119 Rostock, Germany*

Correspondence should be addressed to Hagen Wierstorf (`hagen.wierstorf@tu-berlin.de`)

ABSTRACT

Wave Field Synthesis systems are known to suffer from coloration. Depending on the applied number of loudspeakers comb-filter effects are present in the amplitude spectrum. In a listening experiment participants rated the perceived coloration of a synthesized point source compared to the reference case of a single loudspeaker. The point source was presented by two-channel stereophony and Wave Field Synthesis applying a circular loudspeaker array ranging from 14 up to 3 584 loudspeakers. Test participants were placed at different positions in the audience area for Wave Field Synthesis. The whole experiment was done using binaural synthesis.

1. INTRODUCTION

Speech communication and music plays an important role in our everyday lives. In both cases sound is often created via electroacoustical transducers. In order to preserve the original spatial arrangements of the presented sound more than one transducer is necessary. The importance of the spatial aspect is highlighted by the short time it took until a two-channel presentation technique was invented after the first telephones were available: T. du Monel [1] arranged two parallel telephone channels to transmit recorded music for binaural listening in 1881—only five years after the invention of the telephone.

After more channels became a technical possibility different spatial sound presentation techniques were proposed that use more channels to enhance the spatial perception compared to the case of a two-channel presentation. This started with quadrophonie and developed further to sound field synthesis methods that apply a large number of loudspeakers such as Wave Field Synthesis [2]. Nonetheless, stereophony is still the most prominent presentation technique today. This includes not only the classical two-channel stereophony, but also techniques such as 5.1 surround. The advantage of enhanced spatial impressions coming with sound field synthesis methods does not seem large enough compared to its disadvantages. The high number of loudspeakers is an obvious disadvantage, but also the impact of the sound field synthesis techniques on the timbre of the presented sound has to be considered. Rumsey et al. [3] showed that for quality ratings in the context of surround sound, timbre could explain about 70% of the given quality ratings.

This paper investigates timbre in Wave Field Synthesis. We start with a general discussion of timbre and coloration. Then the influence of the number of loudspeakers on timbre will be discussed and investigated with binaural synthesis in a listening experiment including also two-channel stereophony.

2. TIMBRE AND COLORATION

All definitions of timbre are implicit ones, they state what timbre is not. This leads to the circumstance that the definition of timbre already has a big influence on the resulting research questions. Timbre is most often defined as "that attribute of auditory sensation which enables a listener to judge that two nonidentical sounds, similarly presented and having the same loudness and pitch, are dissimilar" [4]. Plomp added "same duration" to the properties of the nonidentical sounds [5, p. 285]. To highlight the wide range of features that are included in such a definition Patel [6] provides an analogy. Timbre is similar as if describing "looks" of human faces, "looks" is that attribute which enables an observer to judge that two nonidentical faces with the same height, width and complexion, are dissimilar. It is obvious that timbre is a multidimensional percept and the number of dimensions that can be detected in an experiment depends highly on the stimuli.

If the difference of two points in the timbral space is assessed it is described as *coloration*, whereby one of the points is the reference and the other point is colored. The reference point can be explicitly presented to a listener or it is implicitly known to the listener through experience. The latter is known as internal reference. One of the complicating aspects of coloration is that the metric of the timbral space is not known and could be non-trivial. In the literature an euclidean metric [7] or a weighted euclidean metric [8] is commonly assumed, but cannot be assured. Another questionable assumption that is often made is a negative connotation of coloration. For example Brüggen [9] defined the reference as the desirable point and coloration as the move in timbral space to an adverse point. This statement makes the implicit assumption that there is only one point in timbral space that corresponds with a high perceived sound quality and that the reference should be placed always at this point. Brüggen relativized his opinion by stating that for performances such as music played in a concert hall the coloration due to the room is a desired one and the perceived quality of the sound is better for the colored case.

One problem of the above definition of timbre is that only three perceptual aspects are directly named that should be constant between different stimuli. It is not specified which other dimensions the phrase "similarly presented" should include. For example, is it still similar if one stimulus is presented in an anechoic chamber and the other in an office? In order to clarify this situation some authors have included more aspects in the indirect definition of timbre. Letowski [10] gives a definition of timbre that explicitly adds spatial perception to the similar presented dimensions. Emiroglu [11, p. 89] has a similar approach stating: "The label timbre combines all auditory object attributes other than pitch, loudness, duration, spatial location and reverberation environment."

In Wave Field Synthesis the signal of the presented source is weighted, delayed and played back over several loudspeakers which adds unwanted high-frequency wave fronts to the sound field. Hence, the influence of the additional wave fronts on its perceived timbre is of special interest. This situation is comparable to the influence of a room on the timbre of a sound source. In the literature there are mainly two phenomena investigated in this context. One is the influence of different rooms on the coloration of a sound source. The other deals with the fact that the coloration of a sound source placed in a room is different depending on the listener's usage of only one or both of her ears when listening to the sound. The second one is summarized under the term *binaural decoloration* [12].

A straightforward explanation of both phenomena is presented by Brüggen [9]. He defined timbre after the ANSI definition [4] and subsumed any kind of spatial impression due to the room under coloration. He followed Berkley [13] who found the dimensions *echo* (related to reverberation) and *color* (related to spectral deviation) with a multidimensional scaling method for sound events with reflections. The results showed that the echo dimension is mainly influenced by late reflections and the color dimension by early reflections. The binaural decoloration phenomenon is then explained via a blind system identification that tries to identify the spectrum of the room and removes it from the spectrum of the sound source placed in that room. With this mechanism Brüggen was able to explain his results regarding the coloration of a sound source placed in different rooms [9, Fig. 5.11]. To explain the binaural decoloration phenomenon for stereophony, Theile [14] proposed his association model. The model states that a listener is able to identify the location of two sound events even if they have the same signal. In the case of stereophony the listener is then able to do a binaural decoloration of the single corresponding perceived auditory event. The association model has problems to predict the same amount of binaural decoloration for a sound source placed in a room, where the number of different locations of the sound events, including reflections, is higher than two.

A shortcoming of the proposed binaural decoloration mechanism is its independence from the task or context of the listener. For example, Olive et al. [15] published a study where they investigated the influence of room acoustics on absolute quality ratings of loudspeakers via dynamic binaural synthesis. The listeners were asked to rate the sound quality using a 10 point scale. In a first test the listeners could make direct comparisons in one trial only between the different loudspeakers. In a second test the direct comparison was only possible between the rooms having a fixed loudspeaker per trial. The dominating factor for the quality ratings were the loudspeakers in the first experiment and the rooms in the second one. These experiments show that listeners often answer differently towards the same signals and relative coloration has a larger influence than absolute coloration.
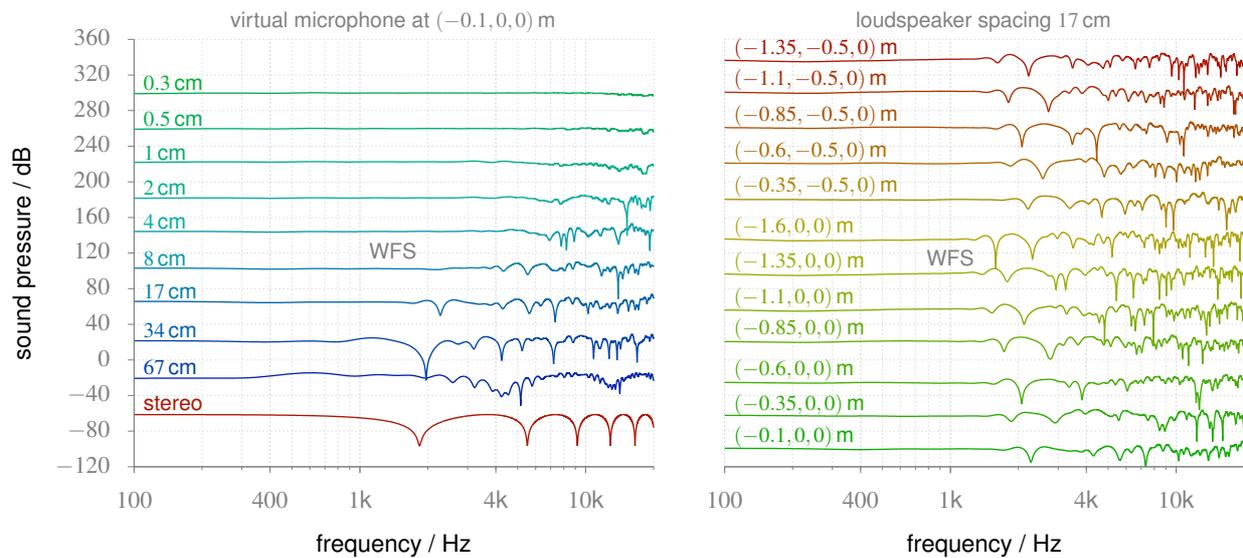
Fig. 1: Amplitude spectra for WFS and stereophony. In the left graph the loudspeaker arrays vary, in the right graph the position of the listener varies. The position of the listener and the distance between single loudspeakers is indicated in the figure. The spectra are shifted in absolute magnitude in order to display them.

3. WAVE FIELD SYNTHESIS

Wave Field Synthesis (WFS) allows for the control of a sound field in an extended area given that the area is at least partly surrounded by loudspeakers. If the distance between the loudspeaker is small enough, under 1 cm, there will be no deviations in the sound field compared to the desired one. Practical loudspeaker setups have larger spacings between the single loudspeakers. The result will be additional high-frequency wave fronts arriving with different delays at the listener position. This process leads to a comb-filter-like amplitude spectrum of the synthesized sound above the so-called aliasing frequency. The aliasing frequency is depending on the distance between the loudspeaker and the position of the listener [16]. The influence of the loudspeaker setup on a synthesized source can be analyzed by simulating the impulse response of the complete system synthesizing a source with Wave Field Synthesis. This is presented in Fig. 1 for the loudspeaker setups and listener positions that are included in the listening experiment. Different circular loudspeaker arrays with a diameter of 3 m were driven by Wave Field Synthesis to synthesize a point source at a position of $(0, 2.5, 0)$ m. In all cases the pre-equalization filter of Wave Field Synthesis was manually

optimized for the central listening position. The impulse responses were calculated for positions of the left ear of the test participants, excluding any head-related transfer function (HRTF). The calculation is identical to placing microphones at these positions and measure the impulse responses. This has been done for a two-channel stereophonic setup as well.

In the left graph the distance between the loudspeakers was varied from 67 cm to 0.3 cm, ranging from 14 to 3 584 loudspeakers. The amplitude spectra highlight that the loudspeaker setup with an inter-loudspeaker distance of 0.3 cm has a more or less flat frequency spectrum, whereas for lower numbers of loudspeakers comb-filter-like deviations in the spectrum occur. The lower the number of loudspeakers the earlier these deviations occur, starting around 400 Hz for 67 cm. The stereophonic amplitude spectrum has a more regular comb-filter structure due to the involvement of only two loudspeakers. Deviations of the spectrum in the form of large dips occur above 1.8 kHz.

The right graph of Fig. 1 summarizes the amplitude spectra for Wave Field Synthesis applying a loudspeaker array with 56 sources and a fixed distance of 17 cm between them. The spectra are plotted for twelve different

listening positions as indicated in the figure. The further the listener will move to the left of the audience area, the slightly lower the spatial aliasing frequency, which is visible in the form of an earlier start of the spectral deviations. By comparing the first dips of the spectra it can also be observed that the dips are shifted to higher frequencies for listener positions further to the back of the audience area.

By inspecting the amplitude spectra the hypothesis arises that the perceived coloration will directly depend on the number of applied loudspeakers and that it will be most prominent for the case of a distance of 67 cm between the loudspeakers. Also for the case of stereophony, a perceived coloration is expected. Looking at the literature only few investigations of the coloration properties of spatial sound techniques are available. Wittek [17] has investigated the differences in intra-system coloration between Wave Field Synthesis and stereophony, using linear loudspeaker arrays with different spacings. He asked the listeners if they perceive a timbral difference between a reference source coming from 5° and the given test stimuli coming from other directions. The reference source and the test stimuli were always presented by the same system, leading to an assessment of the coloration differences. These differences were rated on a scale ranging from *no difference* towards *extremely different*. The listeners were centrally seated at a distance of 1.5 m to the array, and pink noise bursts were presented. The test stimuli were generated via dynamic binaural synthesis. Figure 2 summarizes the results, the intra-system coloration is given by the average coloration of the sources coming from different directions than the reference source. For a loudspeaker spacing of 3 cm the intra-system coloration of WFS was comparable to the case of stereophony and single loudspeakers. For larger loudspeaker spacings ranging from 12 cm to 48 cm, the intra-system coloration was perceived as being stronger but independent of the different loudspeaker spacings.

De Bruijn [18] investigated the variation of timbre for WFS within the audience area for linear loudspeaker arrays with different spacings. He found large differences in terms of coloration for loudspeaker spacings of 0.5 m and negligible differences for a spacing of 0.125 m. As source stimuli, speech shaped noise was applied.

4. METHOD

The experiment was performed with the binaural simulation method as described in Wierstorf et al. [19] using
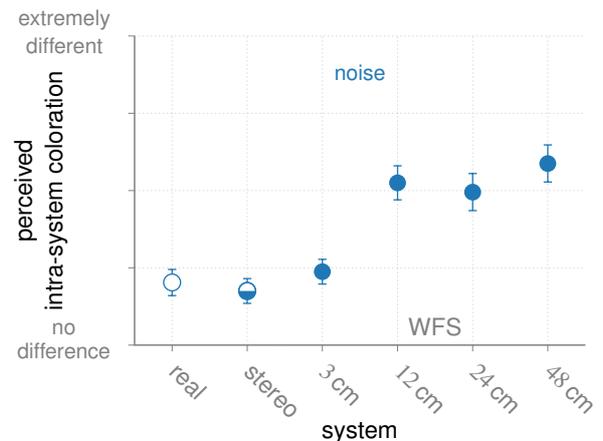

Fig. 2: Average results with confidence intervals for the following question: Is there a timbral difference between the reference and the stimulus? Whereby the reference and the other stimuli were presented by the same system each time, leading to the measurement of intra-system coloration. All loudspeakers, including real, stereo, and WFS, were simulated via binaural synthesis. The results are replotted from Wittek [17, Fig. 8.6].

HRTFs measured with the KEMAR manikin and a resolution of 1°. The only difference is that the dynamic head-tracking part was disabled to exclude changing coloration due to head movements, which occured during informal listening.

The investigation of changes in timbre with binaural synthesis has inherent limitations. The biggest problem is that the synthesis itself introduces changes in timbre, which can only be compensated for to some degree by using individual HRTFs and individual headphone compensation [20]. The investigation of absolute coloration judgement without explicit reference will not be possible with binaural synthesis, because the measured coloration could be due to the synthesis process itself or due to the system under investigation, and there is no way to distinguish between both causes. If the absolute coloration due to the binaural synthesis could be limited, the differences in coloration between simulated systems could be investigated under the assumption that the binaural synthesis has the same influence on coloration for all systems.

One promising result from the literature is the study by Olive et al. [21] who found no difference in the preference ratings for four different loudspeakers between

the measurement with real loudspeakers and their bin-aural simulations. They applied non-individual HRTFs and non-individual headphone compensation filters. This was further supported by the study presented in Wittek [17, Fig. 8.4] who found the same amount of intra-system coloration for a stereophonic setup realized by real or simulated loudspeakers.

4.1. STIMULI

In order to ask the listeners to judge changes in timbre a point source placed at $(0, 2.5, 0)$ m was chosen as a reference stimulus, which was realized by using a single HRTF. The same point source was synthesized with Wave Field Synthesis for several circular loudspeaker arrays. Each array had the same geometry with a radius of 3 m with its center at $(0, 0, 0)$ m, but different numbers of loudspeakers, namely 14, 28, 56, 112, 224, 448, 896, 1 792, and 3 584. For the array with 14 loudspeakers this corresponds to a distance of 67 cm between the individual loudspeakers going down to 0.3 cm for the array with 3 584 loudspeakers. In addition, a stereophonic setup with two loudspeakers placed at $(1.4, 2.5, 0)$ m and $(-1.4, 2.5, 0)$ m was included leading to a total number of 10 different conditions, not counting the reference. All impulse responses were normalized to the same maximum absolute amplitude before convolving them with the audio material during the experiment.

Three different audio source materials were used. A pulsed pink noise train composed of 800 ms noise bursts with 50 ms windowing at the beginning and end and a pause of 500 ms between the bursts. This stimulus was also used by Wittek [17, Sec. 8.2]. As a second stimulus a twelve second clip from the electronic song "Luv deluxe" by "Cinnamon Chasers" was chosen. It is an instrumental song including cymbals and subtle white noise which may help revealing coloration to a similar degree as the pink noise stimulus. The third stimulus was an eight second long female speech sample.

4.2. PROCEDURE

The listeners were asked to rate the difference in timbre between the reference stimulus and the other conditions on a continuous scale with the attribute pair *no difference* and *very different* at its endpoints. This was accomplished with a MUSHRA test design, including a hidden reference and a lower anchor. The low anchor was created by high-pass filtering the reference condition with a second order Butterworth filter with a cutoff frequency of 5 kHz. The listeners were instructed to rate the coloration
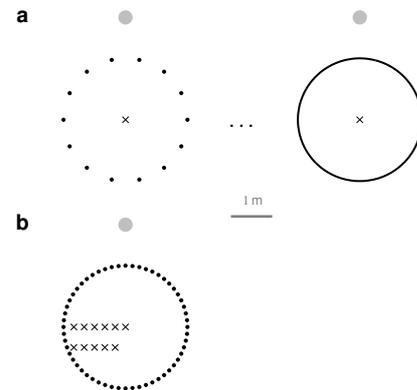

Fig. 3: Experimental setup for coloration experiment. In the first part of the experiment the listener position was fixed (a) and the number of loudspeakers varied, in the second part this was ivice versa (b). The black dots are the loudspeakers, the black crosses the listener positions and the grey dots the position of the synthesized point source.

and not the differences in loudness or perceived externalization of the stimuli. They started with one training run before the actual experiment began. The training consisted of a run at the central listening position, with varying number of loudspeakers and a different music track.

During a single run in the experiment the participants had to rate all 10 different conditions, the hidden reference and the lower anchor for one given audio material. The stimuli were looped during the experiment and the listener could switch instantaneously between the conditions as often as she liked. For the first part of the experiment the listeners were placed centrally in the audience area, at $(0, 0, 0)$ m, see Fig. 3. This part was repeated two times, resulting in a total of six runs.

To investigate the influence of the listening position on coloration separately, another three runs were added. Here, the loudspeaker array with 56 sources was used and tested for the 11 different listening positions indicated in Fig. 3b: $(0, 0, 0)$ m, $(-0.25, 0, 0)$ m, $(-0.5, 0, 0)$ m, $(-0.75, 0, 0)$ m, $(-1, 0, 0)$ m, $(-1.25, 0, 0)$ m, $(0, -0.5, 0)$ m, $(-0.25, -0.5, 0)$ m, $(-0.5, -0.5, 0)$ m, $(-0.75, -0.5, 0)$ m, $(-1, -0.5, 0)$ m, $(-1.25, -0.5, 0)$ m. The head of the listener was always oriented towards the source at all positions, to exclude a change of the direction the synthesized source was presented from. The synthesized point source for the
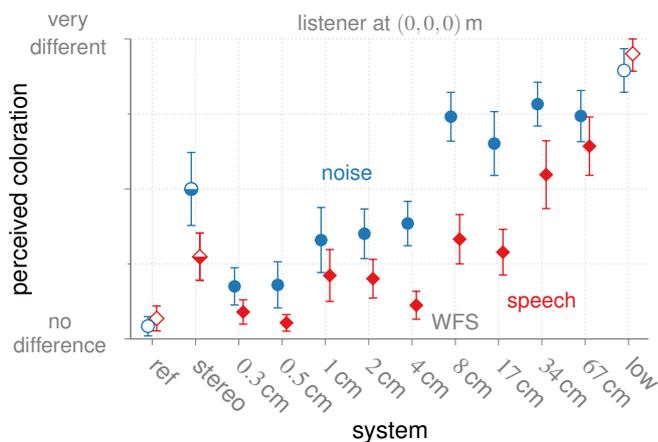
Fig. 4: Average results with confidence intervals for the perceived coloration. Circles show results for noise, diamonds for speech.

Fig. 5: Perceived coloration rated with the attribute pair *very different*, *no difference*. The results are written directly at the listening position where the listener had to rate the coloration and are further highlighted by a corresponding color. The average confidence interval is 1.2 over all positions. The loudspeaker positions are indicated by black circles, where filled circles are the active ones.

listening position at $(0,0,0)$ m was used as the reference which was also included as a hidden reference. In contrast to the other runs, no low anchor was included.

4.3. PARTICIPANTS

15 listeners were recruited for the experiment, aged 23 to 29 years. All test participants had self-reported normal hearing and were financially compensated for their efforts.

5. RESULTS

Figure 4 summarizes the results for the runs of the experiment where the number of loudspeakers was varied and the listener was positioned at the center. Only the results for pink noise and speech as stimuli are presented. The results for music were only significantly ($p < 0.05$) different at two positions from the ones for noise, as indicated by an independent-samples Mann-Whitney U test.

The results are summarized by calculating the average for every listener before calculating the mean and confidence interval over all listeners. An independent-samples Mann-Whitney U test showed that the results of the repeated measurements were not significantly different from each other ($p > 0.05$), highlighting that the listeners were able to answer the task in a reliable way. The test participants rated the hidden reference as not different from the reference and the lower anchor as being very different from the reference. The overall ratings for the Wave Field Synthesis stimuli show a clear dependency
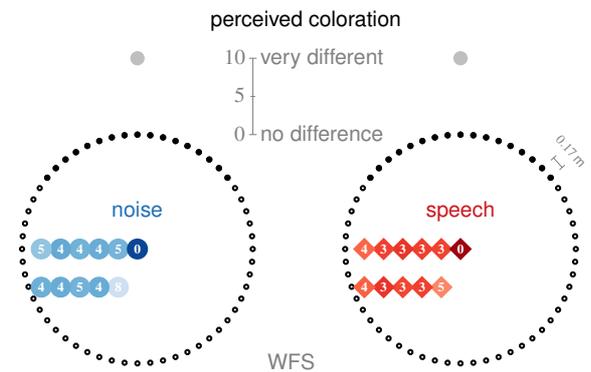
of the perceived coloration on the distance between the loudspeakers. The system with the lowest distance was rated to be only slightly colored, whereas the system with the largest inter-loudspeaker distance was rated the most colored Wave Field Synthesis system. The stereophonic presentation achieved a similar coloration as the Wave Field Synthesis setup with a loudspeaker spacing of 8 cm.

When using the speech stimuli coloration was consistently rated lower in comparison to the case of using noise stimuli. A Wave Field Synthesis system with a loudspeaker distance of 4 cm already achieved a transparent presentation for the speech stimulus in terms of coloration.

The other three runs of the experiment investigated the perceived coloration at different positions in the audience area for a WFS system with 56 loudspeakers and a corresponding distance of 17 cm between the loudspeakers. Figure 5 summarizes the results. The condition with the listener at the center was the hidden reference and was not rated as being different. Most of the other positions were rated to be equally colored whereby the noise stimuli are rated to be more colored. Only the position at

$(-0.25, -0.5, 0)$ m is deviating from that pattern by being perceived as more colored than all other positions.

6. DISCUSSION

The results indicate that the number of loudspeakers has a large influence on the perceived coloration for Wave Field Synthesis. This is not a surprising result reconsidering the magnitude spectra of the different systems as shown in Fig. 1. Here, it is obvious that the spectrum deviates from a desired flat frequency response for frequencies above the aliasing frequency, which is directly dependent on the distance between adjacent loudspeakers. In contrast to our localization results, where a distance of 17 cm already resulted in an authentic localization accuracy [19], the perceived coloration never vanishes for Wave Field Synthesis and pink noise as stimulus. Even for an inter-loudspeaker distance of 0.3 cm slight coloration is perceived. Only for the speech stimulus and the inter-loudspeaker spacing of 0.3 cm the perceived coloration was indistinguishable from the hidden reference for both the central and off-center listening positions.

The results for stereophony suggest that sources presented by that method exhibit coloration, meaning that binaural decoloration is not able to suppress it completely. If the amplitude spectrum in Fig. 1 is compared to the ones of Wave Field Synthesis it could be concluded that the binaural decoloration has a larger impact on stereophony, because the perceived amount of coloration seems to be less than what would be predicted by the position of the first dip in the amplitude spectrum. Another possibility might be that the dips in the spectrum for stereophony are more smeared out by the auditory filters as it is the case for Wave Field Synthesis.

The coloration ratings for Wave Field Synthesis with 56 loudspeakers at different listening positions revealed a more or less equal coloration to that obtained at the central listening position. However, this conclusion has to be relativized due to the multi-dimensionality of timbre. Because conditions are rated to have the same coloration compared to a fixed reference condition did not necessarily include that they have no relative coloration between each other.

By averaging the coloration results for noise from Fig. 5, a value of 4.9 is obtained on a scale from 0 to 10. This result is identical to the one Wittek [17] found for loudspeaker arrays with an inter-loudspeaker spacing of

12 cm and 24 cm. The distance for the 56 loudspeakers in the current study is 17 cm.

An inspection of the actual root mean square value of the presented signals revealed that there were fluctuations up to 3 dB between the single conditions. Therefore, it could be that the listeners have included a loudness-related cue in their coloration rating even if they were advised not to do so. To further analyze this, the correlation between the actual root mean square values and the coloration ratings were calculated as an average for the speech and noise stimuli. The result is 0.6 for the first part of the experiment with the fixed central listening position. This indicates that the loudness was not the main cue for the given coloration ratings. The correlation was also calculated for the conditions with different listening positions. When averaging over speech and noise the correlation results to 0.2. This indicates that the loudness did not have a major influence on the coloration ratings for these conditions. A more precise inspection of the listening position that was rated to be most colored, namely $(-0.25, -0.5, 0)$ m, revealed nonetheless that other factors could have influenced the coloration ratings. The position $(-0.25, -0.5, 0)$ m had the loudest signal and was reported as being less externalized compared to all other conditions by two test participants after the experiment.

The music and pink noise stimuli show no significant differences in all but two positions. This indicates that even the usage of music alone might be suitable to investigate the perceived coloration. That is of advantage because most listeners regarded the noise stimulus as unpleasant, as revealed by informal reports after the tests.

7. CONCLUSION

The results show a clear dependency of the perceived coloration of a synthesized point source and the given loudspeaker setup. The lower the inter-loudspeaker spacing, the less coloration will be perceived. This direct relation is due to the connection between the aliasing frequency and the distance between the loudspeakers. The aliasing frequency specifies from which frequency onwards deviations in the amplitude spectrum of the synthesized source will appear, which seems to be a good measure for the perceived coloration of the synthesized source.

The aliasing frequency changes only to a small extent at nearby positions in the audience area, which seems to correspond with the results showing that the perceived coloration is similar at different positions in the audience area for Wave Field Synthesis and 56 loudspeakers.

For stereophony, the amount of coloration seems to be less than for a WFS system with a similar position of the first dip in the amplitude spectrum. This indicates that binaural decoloration may be more pronounced for stereophony than for WFS.

8. ACKNOWLEDGEMENTS

9. REFERENCES

[1] T. du Moncel. The international exhibition and congress of electricity at Paris. *Nature*, (October 20):585–89, 1881.

[2] S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter. Spatial Sound With Loudspeakers and Its Perception: A Review of the Current State. *Proceedings of the IEEE*, 101(9):1920–38, 2013.

[3] F. Rumsey, S. Zielinski, R. Kassier, and S. Bech. On the relative importance of spatial and timbral fidelities in judgments of degraded multichannel audio quality. *The Journal of the Acoustical Society of America*, 118(2):968–76, 2005.

[4] ANSI. American National Standard Acoustical Terminology, ANSI S1.1-1994, 1994.

[5] B. C. J. Moore. *An Introduction to the Psychology of Hearing*. Emerald, Bingley, 2012.

[6] A. D. Patel. *Music, Language And The Brain*. Oxford University Press, New York, 2010.

[7] R. Plomp, L. C. W. Pols, and J. P. van de Geer. Dimensional Analysis of Vowel Spectra. *The Journal of the Acoustical Society of America*, 41(3):707–12, 1967.

[8] S. McAdams, S. Winsberg, S. Donnadieu, G. De Soete, and J. Krimphoff. Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes. *Psychological Research*, 58(3):177–92, 1995.

[9] M. Brüggen. *Klangverfärbungen durch Rückwürfe und ihre auditive und instrumentelle Kompensation*. PhD thesis, Ruhr-Universität Bochum, 2001.

[10] T. R. Letowski. Sound quality assessment: concepts and criteria. In *87th Audio Engineering Society Convention*, page Paper 2825, 1989.

[11] S. S. Emiroglu. *Timbre perception and object separation with normal and impaired hearing*. PhD thesis, Carl von Ossietzky Universität Oldenburg, 2007.

[12] W. Koenig. Subjective Effects in Binaural Hearing. *The Journal of the Acoustical Society of America*, 22(1):61–62, 1950.

[13] D. A. Berkley. Hearing in rooms. In William A Yost and G Gourevitch, editors, *Directional Hearing*, pages 249–60. Springer, New York, 1987.

[14] G. Theile. *Über die Lokalisation im überlagerten Schallfeld*. PhD thesis, Technische Universität Berlin, 1980.

[15] S. E. Olive, P. L. Schuck, S. L. Sally, and M. Bonneville. The Variability of Loudspeaker Sound Quality Among Four Domestic-Sized Rooms. In *99th Audio Engineering Society Convention*, page Paper 4092, 1995.

[16] S. Spors and J. Ahrens. Spatial Sampling Artifacts of Wave Field Synthesis for the Reproduction of Virtual Point Sources. In *126th Audio Engineering Society Convention*, 2009.

[17] H. Wittek. *Perceptual differences between wavefield synthesis and stereophony*. PhD thesis, University of Surrey, 2007.

[18] W. P. J. de Bruijn. *Application of Wave Field Synthesis in Videoconferencing*. PhD thesis, Technische Universiteit Delft, 2004.

[19] H. Wierstorf, A. Raake, and S. Spors. Binaural assessment of multi-channel reproduction. In J Blauert, editor, *The technology of binaural listening*, pages 255–78. Springer, New York, 2013.

[20] B. S. Masiero. *Individualized Binaural Technology*. PhD thesis, RWTH Aachen, 2012.

[21] S. E. Olive, T. Welti, and W. L. Martens. Listener Loudspeaker Preference Ratings Obtained in situ Match Those Obtained via a Binaural Room Scanning Measurement and Playback System. In *122nd Audio Engineering Society Convention*, page Paper 7034, 2007.