

COVER PAGE

Document downloaded by @DAEL

Fri Nov 6 04:30:58 2026
For personal use

When automatic English translation is provided, only the original document is authentic.
The EAA cannot be held responsible of any translation error

Bibliographical reference

The Acoustical Effect of Musicians' Movements During Musical Performances, David Ackermann, Christoph Böhm, Fabian Brinkmann and Stefan Weinzierl, Acta Acustica **vol. 105** (Number 2), 2019, pp. 356-367

DOI
https://doi.org/10.3813/AAA.919319

Acta Acustica united with Acustica
Vol. 105 (2019) 356 – 367

DOI 10.3813/AAA.919319

The Acoustical Effect of Musicians' Movements During Musical Performances

David Ackermann, Christoph Böhm, Fabian Brinkmann, Stefan Weinzierl
TU Berlin, Audio Communication Group, Einsteinufer 17c, 10587 Berlin, Germany,
david.ackermann@tu-berlin.de

Summary
Acoustic musical instruments act as dynamic sound sources communicating the expressive intentions of a performer to the audience in a dedicated spatial environment. From an acoustical point of view, the directivity of musical instruments is relevant both for the loudness and timbre of an instrument at a certain position in the loudness and timbre of an instrument at a certain position in the audience, as well as for the spatial characteristics of the generated sound field. Musical instruments, however, are dynamic sound sources always moved by musicians as an element of their performance on stage. This work aims at assessing the acoustical effect and the perceptual relevance of these movements. For this purpose, we have recorded solo musical performances with all standard orchestral instruments with an optical motion tracking system, as well as the corresponding audio signals. The effect of the movements was evaluated by analysing the spectral fluctuation and the time-dependence of room acoustical parameters in a virtual acoustic environment in anechoic and reverberant conditions. In a subsequent listening test, an auralization of the static and dynamic musical performance was presented to listeners by binaural synthesis, showing that the signal-related fluctuations are clearly audible both in anechoic and reverberant situations. We discuss different approaches how to consider these effects for the simulation of natural acoustic sources in virtual acoustic reality.

© 2018 The Author(s). Published by S. Hirzel Verlag · EAA. This is an open access article under the terms of the Creative Commons Attribution (CC BY 4.0) license (https://creativecommons.org/licenses/by/4.0/).

PACS no. 43.55.HJy, 43.55.Ka, 43.75.Yy

1. Introduction

Various attempts have been made to measure the directivity of musical instruments, dating back to the 1970s, when Meyer measured the radiation of different musical instruments in averaged octave bands [1]. Other approaches include those by Cook and Trueman using an icosahedral array of 12 microphones to measure directional impulse responses of stringed musical instruments with force hammer excitation [2], by Otondo and Rindel using 13 microphones in two circular arrays (vertical and horizontal) to measure the directivity of three wind instruments with natural excitation [3], by Hohl using a spherical array of 64 microphones to measure the directivity of a violoncello, three brass and five woodwind instruments [4]; and by Patynen and Lokki, who used a spherical array of 22 microphones to measure the directivity of 14 orchestral instruments and a singer [5]. More recently, a German consortium has published an acoustic radiation pattern database of 41 musical instruments, including all standard orchestral instruments and their historical precursors, based on measurements with a spherical array of 32 microphones [6, 7].

The relevance of correctly implemented directivities for room acoustical measurements or room acoustical simulations was investigated by Otondo and Rindel [3], and Wang and Vigeant [8], who demonstrated presumably audible changes in room acoustical parameters when compared to an omnidirectional source. Even different dodecahedron loudspeakers, all of them compatible with the requirements for omnidirectional sound sources according to ISO 3382, were shown to produce significantly different room acoustical parameters [9].

Musical instruments, however, are *dynamic* sound sources. Their directivity varies not only with the played note (pitch), but also with the movements that musicians always make - to a greater or lesser extent - during the performance on stage. First attempts to investigate the effect of such time-variant characteristics were made for the human voice, by evaluating the perception of phoneme-dependent directivities in a virtual environment [10]. For a representation of time-dependent directivities in auralizations, Otondo and Rindel proposed multi-channel anechoic recordings of the audio content with a subsequent segmentation of a virtual, spherical sound source in room acoustical simulations. Compared to the static directivity,

Received 13 April 2018,
accepted 22 January 2019.

© 2018 The Author(s). Published by S. Hirzel Verlag · EAA.
This is an open access article under the terms of the CC BY 4.0 license.

Document downloaded by @DAEL, 2026-11-06 04:31:00 - personal use only

fed with a single-channel anechoic recording they found an improvement in the perceived timbral qualities of the musical instrument, which could result, however, both from the time-dependent timbral fluctuations of the auralization as well as from the increased spectral resolution of the directivity representation itself [11]. A similar experiment was done by Postma *et al.*, using a multi-channel auralization approach with 12 overlapping beams instead of segments with rectangular weighting, which are prone to discrete changes in sound level and timbre when the source orientation is altered. They used a pre-measured voice directivity controlled in orientation by measured motion tracking data for an actor on stage in room acoustical simulation, and found differences in several perceptual attributes such as the 'plausibility' of the auralization between static and dynamic source representation [12].

Movement-induced fluctuations make *one* contribution to the dynamic behaviour of musical instruments and speech, as emulated in [11] and [12]. Both authors, however, did not try to isolate this effect from other contributions such as (in [11]) the note-dependent differences in directivity or the spectral and spatial resolution of the directivity itself, nor did they try to quantify the resulting sound field modulations, as a first indication of the perceptual significance of these effects.

This, however, is the aim of the current study. Therefore, we have investigated the extent of time-variant modulations of the sound field generated by musical instruments, as induced by the movements of musicians and their instruments. At a particular listening position, these modulations will manifest acoustically in two ways: The pivoting of the frequency-dependent directivity of musical instruments will lead to a spectral modulation of the sound of the instrument. At the same time, also the room will be excited in a different way, if, for example, the main lobe of the radiation pattern is directed to surfaces with different absorption and scattering properties, or reflected by different surface geometries. Thus, we analysed the spectral modulations at a frontal listening position as well as the fluctuations of room acoustical parameters in a typical performance space. These modulations were examined by capturing the movements of 20 professional musicians playing in seating and standing condition, corresponding to a solo performance and an orchestral performance situation.

We can suspect that the distance between the sound source and the listener will have a notable impact on the two effects described above. With increasing source-receiver distance, the increasing contribution of the diffuse sound field can be expected to smooth out spectral modulations at the receiver, where different parts of the radiation pattern, deflected by sound reflections from the room, overlap. At the same time, modulations of the room acoustic "embedding" of the musical instrument will only be noticeable if sound reflections from the hall are sufficiently strong compared to the direct sound. To describe the range of possible room acoustic conditions, we performed the analysis both for the anechoic case and for a listener position at twice the critical distance in a typical, mid-size performance venue.

In order to isolate the impact of instrumental movements of other influences such as note-dependent changes of the directivity or spectral changes of the audio content itself, the acoustical analysis was performed in a room acoustical simulation rather than the performance venue itself. The acquired motion tracking data was first analyzed with respect to the strength and speed of the musicians movements. To estimate the perceived spectral differences, the tracking data and directivities measured in high spatial resolution were used to virtually rotate the instruments in synchrony with the natural movements of the musicians. The spectral differences between the rotated and a static instrument orientation were calculated, and room acoustic simulations were performed to analyse the fluctuation of room acoustic parameters. The physical analysis was followed by a perceptual evaluation to test the audibility of the movements. For this test, auralizations of three of the recorded instruments in static and dynamic condition and in anechoic and reverberant conditions were used as stimuli for a 2-Alternative Forced Choice (2AFC) test with 11 subjects.

From a musical point of view, knowledge about the acoustical effect of musicians' movements can be valuable for understanding the role of performative *gestures* during a musical performance. In the recent decade, a variety of studies has addressed the role of gestures for the communication of emotions [13], the relation of gestures to the notated musical text [14], the role of 'embodied cognition' of music via internal models of motor action [15], also as an example for the general link between action and perception [16], and the design of musical interfaces translating gestures into electronic sound synthesis [17]. These studies, however, always considered the *visual* communication of gestural meaning, putting aside the question to what extent these gestures are also communicated in the acoustical domain.

2. Method

2.1. Motion capturing and audio recording

The movements of 20 musicians playing 11 different musical instruments (Table I), including all standard orchestral instruments, were captured during solo performances by means of a motion capturing system under concert-like conditions. The recordings were made in the Curt-Sachs-Saal (Tiergartenstraße 1, 10785 Berlin), a chamber music hall for an audience of 250. Each instrument was played by two different professional musicians, with the exception of bass and trombone that were played by only one. Each musician played three different music pieces of his or her own choice, one time standing and one time sitting, except for the cello, that is usually played in a sitting position, and the double bass, that is usually played standing.

The movements of the musical instruments during the performance were captured with an optical motion tracking system (OptiTrack) and the associated Motive soft-

Document downloaded by @DAEL, 2026-11-06 04:31:00 - personal use only

Table I. Orchestra instruments invited for the motion capturing of different musical performances.

Instrument	Quantity	Group
Violin	2	Strings
Viola	2	
Cello	2	
Double bass	1	
Flute	2	Woodwinds
Oboe	2	
Clarinet	2	
Bassoon	2	
Trumpet	2	Brass
French horn	2	
Trombone	1	

ware. The system consisted of eight cameras and reflective markers. Each instrument was equipped with several markers, so that they were always visible by at least two of the cameras for reliable tracking. The position of the instruments are provided by the software in Cartesian coordinates and the orientation in quaternions in 120 Hz temporal resolution. For later analysis the quaternions were converted to the Euler angles yaw, pitch and roll (Figure 2) and Cartesian coordinates.

For the audio recording a professional lavalier microphone with omnidirectional polar pattern (AKG C417) was attached to the instrument, in order to obtain an almost anechoic signal which is unaffected by the movements of the musicians and their instruments during the performance.

The movements of the musical instruments were analysed in range and speed. For this purpose, we indicate the ranges of the orientations in the three degrees of freedom yaw, pitch and roll by means of 5%/95% quantile ranges. Furthermore, we calculated the angular velocity in degree per second indicating the mean speed of the instrument rotations in the three degrees of freedom yaw, pitch and roll. The median of the orientation of each instrument was used as an indication of the 'natural' orientation of the instrument, and used as a reference position defined as 0°/0°/0° yaw, pitch and roll.

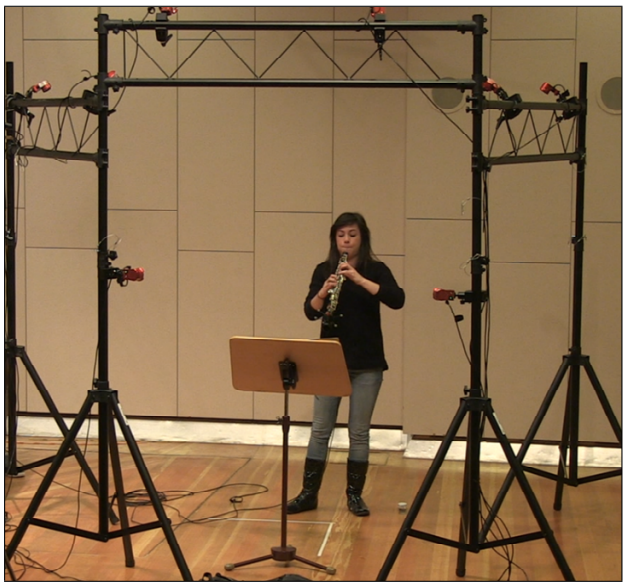


Figure 1. (Colour online) Setup for audio recording and motion tracking. The microphone is clipped to the instrument, which is surrounded by eight motion tracking cameras mounted to the racks.

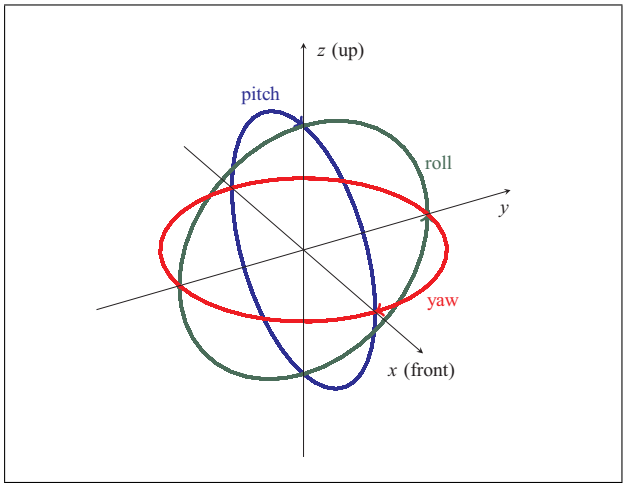


Figure 2. (Colour online) The Euler angels yaw, pitch and roll used to describe the orientation of the musical instruments. The 0°/0°/0° orientation was set to the median of the time-variant values, indicating the 'natural' orientation of the instrument.

2.2. Radiation patterns

The radiation patterns for the 11 orchestral instruments were taken from measurements of sound power and directivity of 41 different musical instruments with a spherical array of 32 microphones, which are available as an open access data publication [6, 7]. In this database, for each instrument the acoustic radiation pattern of every playable tone is presented in the spherical harmonics (SH) domain by coefficients up to order 4 for each fundamental and its first nine overtones. A compact representation of the instrument directivities in third octave bands is additionally provided and was used for this investigation.

To analyze the spectral behavior and to process the directivity in a room acoustical simulation software, the radiation pattern was evaluated along a spatial grid with a resolution of 2° in azimuth and elevation based on the SH representation. The magnitude spectrum of the extracted radiation pattern was stored in the Open Directional Audio File Format (openDAFF) [18], with the phase information ignored.

The directional characteristic of a sound source can be described through its directivity index (DI)

$$DI(f) = 20\log\frac{p_0(f)}{p_{\text{diff}}(f)}, \tag{1}$$

where p_0 is the on-axis sound pressure and p_{diff} the mean sound pressure over the entire spatial grid at a specific frequency f. Thus, a source with an omni-directional sound

Document downloaded by @DAEL, 2026-11-06 04:31:00 - personal use only

propagation has a DI of 0 dB, a source focused in frontal direction has a $DI > 0$ dB, whereas sources with strong lateral radiation can have a $DI < 0$ dB.

2.3. Room acoustic simulation

In order to analyse the modulations induced by instrument movements not only in anechoic conditions, but in a real-world musical setting, we carried out room acoustical simulations, using the RAVEN software [19]. All simulations were done in third octave resolution and a hybrid simulation algorithm using image sources up to third order and ray tracing using 86000 rays. These settings, and the stochastic part of RAVEN's simulation algorithm were kept constant to ensure that the results only depend on the different instrument directivities, and their orientation in space. From every simulation, an impulse response was derived and stored for later analysis.

As a virtual acoustic environment, a model was used based on the Theater an der Wien (Figure 3), a popular venue used for music theater and concerts in Vienna, with a reverberation time of $T_{\mathrm{m}} = 1.0$ s. For the current study, a perfect match between the simulation and the real room was not necessary, because only relative parameter changes are considered. Room acoustic parameters according to ISO 3382-1 [20] were calculated at the receiver position used for the later analyzes, based on a simulated impulse response generated with an omni-directional source and receiver, or a figure-of-eight receiver for J_{LF} and L_{J} (Table II).

To analyze the movement-induced modulations of the sound field, simulations were caried out with the directivity patterns of the respective instrument, as outlined in Section 2.2. The source was positioned in the center of the stage facing the audience. The receiver was placed 10 m away from the source in the audience area facing towards the source. This equals twice the critical distance, and was chosen to slightly emphasize the diffuse field for the analysis.

The simulation of impulse responses corresponding to different source rotations could efficiently be done with RAVEN's animation module, by passing text files with tracking information for the sources and receivers. To limit the simulation time, the tracking data was downsampled to 3 Hz, which was still a sufficient representation of the instrument movements as can be seen in Figure 4. This resulted in 143 to 701 simulations for each performance depending on the musical piece and instrument. For the simulations, only the angular movements were considered, because the translational movements, resulting in purely distance-related changes, were assumed to be inaudible.

2.4. Spectral analysis

For the analysis of the movement-induced spectral fluctuation of the musical instruments in anechoic conditions, the available third octave radiation patterns were rotated sequentially in the SH domain according to the measured Euler angles yaw, pitch and roll. After each rotation, the on-axis magnitude spectrum in front of the virtual instrument was obtained by applying an inverse SH transform using AKtools [23]. The spectral fluctuation was thereby calculated using the relative deviation between spectra of the source in motion, and the spectrum of the static source in its median orientation. The overall amount of spectral fluctuation that occurred for each instrument, playing

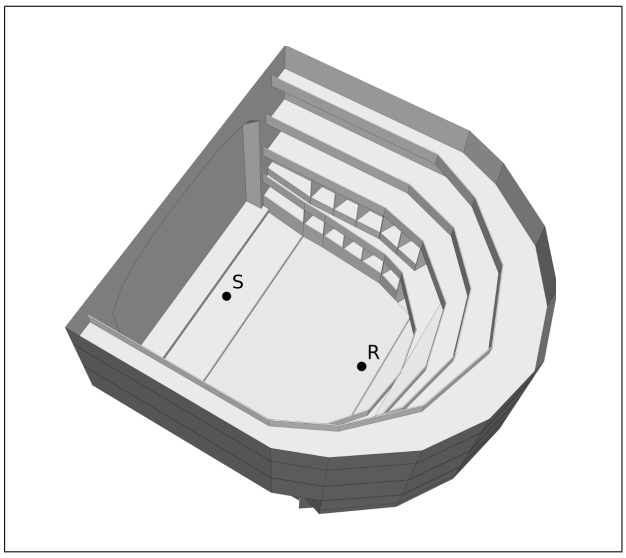


Figure 3. The model used for room acoustic simulation based on the geometric structure of the Theater an der Wien, with the source (S) and receiver (R) positions indicated.

Table II. Room acoustic properties of the virtual concert venue and JNDs according to ISO 3382-1, with recommended frequency averaging.

Parameter	Value	JND
T_{20}	1.0 s	Rel 5%
C_{80}	1.9 dB	1 dB
D_{50}	48%	5%
EDT	1.4 s	Rel 5%
G	5.3 dB	1 dB
L_{J}	−2.2 dB	not known
J_{LF}	0.07	0.05

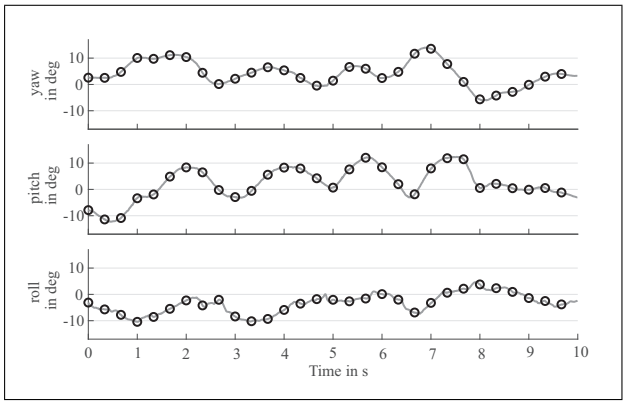


Figure 4. Example of the captured tracking data (yaw, pitch and roll) and the resampled data for room acoustic calculations.

Document downloaded by @DAEL, 2026-11-06 04:31:00 - personal use only

style (standing and sitting), and third octave band was then specified by means of the 5%/95% quantile range.

For the reverberant environment, every impulse response derived from the room acoustics simulations was filtered with a third octave band filter, using the ITA Toolbox [24], and the RMS level in dB was calculated for each band. These energy levels were subtracted from the level corresponding to the median orientation. The overall amount of fluctuation is again indicated by the 5%/95% quantile range for each third octave band, instrument, and playing position.

2.5. Room acoustic analysis

To characterize the influence of the musicians' movements on the spatial sound field in the room, the modulation of acoustic parameters during the musical performance was calculated. For this purpose, we chose standard parameters according to ISO 3382-1 [20], including reverberation time T_{20}, early decay time EDT, clarity C_{80}, definition D_{50}, sound strength G, early lateral energy fraction J_{LF}, and late lateral sound level L_{J}.

The impulse responses were simulated using the source directivity patterns of the respective instrument (section 2.2) and an omni-directional receiver, or a figure-of-eight receiver for J_{LF} and L_{J} (Table II). The parameters were determined from the simulated impulse responses using the ITA toolbox [24], and averaged across third octaves according to [20]. Afterwards, the 5%/95% quantile range of every parameter, instrument, and playing position was calculated.

2.6. Perceptual analysis

The measures of dispersion describing the movement-induced spectral and room acoustical variation over the entire duration of the musical performance (Table V, Table VI) are a first indication of the perceptual relevance of the effect. They are, however, no direct evidence of the audibility, since it would be conceivable that these changes take place so gradually that they are never audible during a musical performance.

For this reason, an additional perceptual evaluation was carried out to test if differences between static and dynamic representations of the musical instruments are audible when using the original temporal development of these modulations. To this end, binaural impulse responses for auralizations were simulated in RAVEN for static and dynamic directivities according to the procedure described in section 2.3, using the FABIAN HRTF database [21, 22] as receiver.

In order to avoid selecting music excerpts as stimulus where spectral modulations would be inaudible at this very moment, a pre-analysis was conducted by calculating movement-induced spectral modulations in third octave bands in windows of two seconds length, with the initial orientation of the instrument used as a reference. The spectral changes were weighted with an inverse A-curve [26] and the third octave RMS level of the audio content at the corresponding time instance, in order to consider both the sensitivity of the auditory system and the energy of the current audio content, and then added across frequency bands. Based on this rudimentary auditory model, we selected one instrument from each of the three instrument groups (strings, woodwinds, brass), with the violin, the flute, and the trumpet in standing position. Moreover, we selected a five second excerpt from the recording showing large predicted spectral modulations. As can be seen in Figure 5, however, the spectral modulation in the selected five second excerpt for the violin is by no means exceptional.

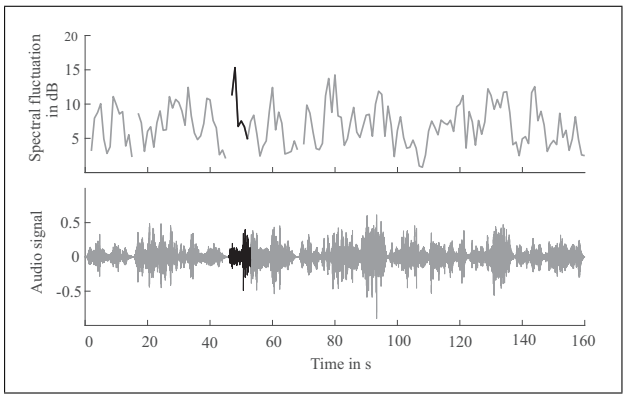


Figure 5. Spectral fluctuation (top) and audio signal (bottom) for the recording of violin 1, used to select a 5 s segment (black) for the listening test.

The stimuli for the listening test were obtained by block-wise and time-variant convolution of the quasi-anechoic audio recording (cf. section 2.1) of these musical segments with the binaural impulse responses under both anechoic and reverberant conditions using RAVEN. Auralizations of the static and dynamic sources always started with the same source orientation to make sure that no differences were audible at the beginning of a stimulus.

The binaural stimuli were played through an RME HDSPe MADI interface without head-tracking (static binaural synthesis). Open, circum-aural headphones (Sennheiser HD800) were used with a frequency compensation filter, which is part of the FABIAN HRTF database [25, 21]. The playback level was equal for all participants, and was adjusted by an expert group to avoid differences between the six test conditions.

Although the generated auralizations sound plausible with respect to the instrumental timbre and the room acoustical environment, we must assume that the spectral shape of the musical signal is, to a certain degree, affected both by the position of the microphone and by its frequency response. Since the microphone was located in the near field of each instrument and the frequency-dependent directivity of the instrument is valid only in the far-field, there is no obvious solution how to compensate for this influence with an inverse filter. Since the listening test, however, aimed at movement-related spectral *fluctuations* of the audio signal rather than the absolute timbre, there is no reason to expect that slight deviations in tone color would influence the results of the test.

Document downloaded by @DAEL, 2026-11-06 04:31:00 - personal use only

The perceptual evaluation was implemented as an ABX test, i.e. a 3-Interval/2-Alternative Forced Choice (3I/2AFC) test, with the intervals A, B and X, and the two possible answers (forced choices) A equals X, and B equals X. Each subject performed 6 ABX tests (3 instruments × 2 acoustical environments) in randomized order with 23 trials each. During each trial the binaural simulation of the static source was placed on button A, the dynamic source on button B, and one of the two stimuli was placed randomly on button X. Before clicking the buttons and making a decision, participants could listen to A, B and X as often as desired, and the playback could be restarted at any time.

ABX is a criterion free and sensitive test design for detecting small differences between stimuli. The test was designed for a critical number of 18 or more correct answers to reject the null hypothesis (differences are not audible, $p_{\text{hit}} = 0.5$), whereby the specific alternative hypothesis (differences are audible, $p_{\text{hit}} = 0.9$) can be rejected for less than 18 correct answers. The type I error level (wrongly concluding that there was an audible difference although there was none) was set to 0.05 and the type II error level (wrongly concluding that there was no audible difference although there was one) was set to 0.20, corresponding to a test power of 80%. Both error levels were corrected for multiple testing of 6 test conditions by means of Bonferroni correction.

The test was conducted with 11 subjects (average age of 32, 9 male and 2 female). All participants had experience with listening tests and binaural synthesis, and no subject had a known hearing impairment. The tests were conducted in an acoustically dry and quiet studio environment at TU Berlin. For each instrument and room 253 trials were conducted in total (11 participants, 23 trials for each of the 6 ABX tests).

3. Results

3.1. Directivity

Movement-induced modulations of the sound field are the result of a frequency-dependent directional pattern of musical instruments in combination with movements of the instrument during the performance. To illustrate the degree of directionality of the 11 orchestral instruments of the current study, Figure 6 shows the directivity index for each instrument in third octave bands. Brass instruments (trumpet and trombone) show a continuous increase of the directivity index over frequency. Although the physics of the radiation for the french horn is the same, the behavior of the index is different, because we chose the frontal viewing direction of the musician as the reference direction, not the axis of the bell. Most other instruments have a mainly frontal sound radiation towards the audience, with large variations across frequency, due to the frequency-dependent modal patterns of the string instruments and the complex radiation by the bell and the tone holes of woodwind instruments.

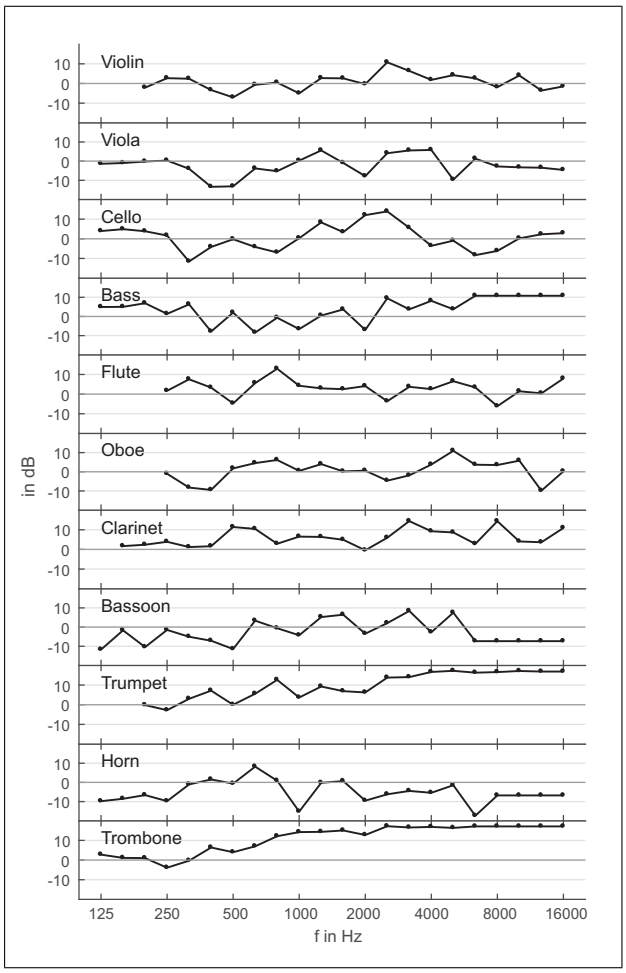


Figure 6. Frequency-dependent directivity index (DI) in dB for each of the 11 orchestral instruments of the current study in third octave bands

3.2. Motion data

The range of motion of the 20 musicians, as indicated by the 5%/95% quantile ranges of yaw, pitch and roll, is between a few degrees and 47° for the yaw movement of violin 1 (Table III). The values, however, are not only different for different musical instruments, but also for different musicians playing the same instrument. As expected, the celli and the double bass, fixed in their position by a spike, exhibit the lowest values. Standing musicians tend to show larger movements than sitting musicians.

The influence of the performer is also visible in the values for the mean angular velocity: Clarinet player 1 exhibits an average yaw/pitch/roll velocity of 11.4/6.5/6.3 in deg/s while clarinet player 2 moves considerably slower with 2.1/3.1/1.7 deg/s (Table IV).

3.3. Spectral effects

Spectral differences between static and moving sources according to Section 2.4 are given in Figure 7 for anechoic and reverberant condition, and sitting and standing performances. The mean spectral fluctuation over all third octave bands as well as the maximum spectral difference for each instrument and playing position are shown in Table V.

Document downloaded by @DAEL, 2026-11-06 04:31:00 - personal use only

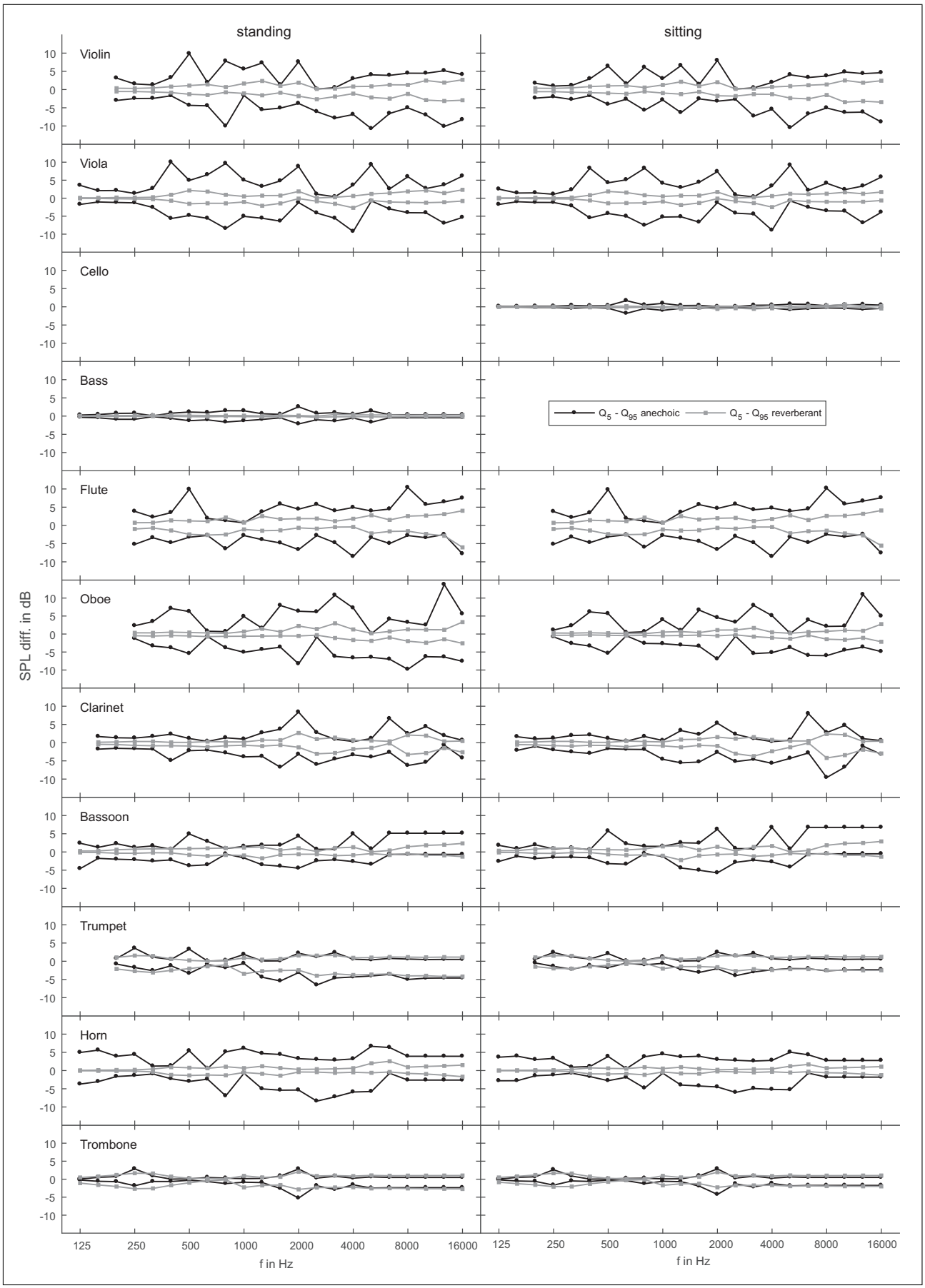


Figure 7. Movement-induced spectral differences of musical instruments, relative to their mean orientation for anechoic and reverberant conditions.

Document downloaded by @DAEL, 2026-11-06 04:31:00 - personal use only

In the anechoic condition, the mean spectral differences for the string instruments range from 0.9 dB for the cello and 1.6 dB for the double bass (both fixed with a spike on the floor) to 7.8 dB for the viola and 8.1 dB for the violin in sitting position, and 8.8 dB resp. 9.7 dB in standing position. For the woodwind section, the spectral modulations for flute and oboe are similar to those of violin and viola, with mean spectral differences of 9.1 dB for flute and 7.6 dB for oboe during sitting performances, and 9.2 dB resp. 10.1 dB in standing position. Clarinet and bassoon showed smaller ranges for their movements (Table III), corresponding to smaller spectral modulations of 5.7 resp. 5.5 dB in sitting and 5.7 resp. 5.1 dB in standing position. For the brass instruments, showing the smallest movements (Table III), also the spectral fluctuations are relatively small, with mean spectral differences for the trumpet of 2.9 / 4.4 dB (sitting / standing), 1.9 dB / 2.3 dB for the trombone, and 5.9 / 7.7 dB for the french horn. With the exception of the clarinet and the bassoon, the mean spectral differences were always higher for standing than for the sitting performance.

If we consider the reverberant condition, the added room reflections can theoretically influence the spectral modulation in two ways. In a room with homogenous absorption, i.e. with the same frequency-dependent absorption coefficients for all boundaries, the spectral modulations can only be reduced by the diffuse field, which effectuates an averaging over different radiation angles (as the origins of different room reflections) of the source directivity. In a room with heterogenous absorption, however, certain parts of the directivity of the musical instrument can be effectively amplified or damped, if they are directed at a highly reflective or highly absorptive wall, thus increasing the spectral fluctuation in case of a pivoted beam pattern.

For the performance venue we chose, which is neither completely homogenous nor extremely heterogeneous but can be regarded as a typical, real-world situation, with a highly absorptive audience area and geometrically complex but similarly absorptive side walls and ceiling, both effects can be observed, but the smoothing effect clearly outweighs the potential amplification. For most instruments, the mean spectral differences in reverberant condition are between 0.5 dB (for the cello) and 8.0 dB (for the oboe) smaller than for the anechoic case. For the trumpet (sitting), no difference could be observed, and for the trombone, the spectral modulation in the reverberant condition was 0.5 dB / 0.6 dB larger (sitting / standing) than in the free field case. In this case, a strongly directional instrument (Figure 6), with a rotational movement mainly in the direction of pitch (Table III), was directing the main lobe sometimes *on* and sometimes *above* the audience area. In this case, the room acoustical environment slightly emphasized the spectral modulation.

3.4. Room acoustic effects

Table VI shows the 5%/95% quantile ranges of the room acoustic parameters according to ISO 3382-1. These values can be compared to just notable differences (JNDs, cf. Table II and [20]), and values exceeding these JNDs are marked in bold in Table VI. It can be seen, that parameters primarily evaluating the early part of the room impulse response such as C_{80}, D_{50}, J_{LF}, and EDT are more sensitive towards movement-induced fluctuations than parameters evaluating the later part of the impulse response such as T_{20} or L_{J}. This corresponds to the effect that the diffuse

Table III. 5%/95% quantile range for yaw/pitch/roll movements, all values in degrees.

	Standing			Sitting		
	yaw	pitch	roll	yaw	pitch	roll
Violin 1	47	33	32	30	30	32
Violin 2	24	31	26	13	35	20
Viola 1	-	-	-	16	14	21
Viola 2	31	24	27	33	21	26
Cello 1	-	-	-	4	2	6
Cello 2	-	-	-	2	1	5
Bass	3	6	10	-	-	-
Flute 1	33	41	29	33	36	28
Flute 2	15	31	17	16	32	19
Oboe 1	31	33	26	17	23	15
Oboe 2	25	25	28	16	21	15
Clarinet 1	35	24	23	22	36	13
Clarinet 2	12	18	8	11	27	6
Bassoon 1	21	17	19	12	20	9
Bassoon 2	13	10	11	10	15	9
Trumpet 1	25	23	9	12	16	13
Trumpet 2	24	20	18	19	18	12
Horn 1	11	9	11	9	9	6
Horn 2	21	16	14	17	12	7
Trombone	7	21	5	3	20	7

Table IV. Mean angular velocity for yaw/pitch/roll movements, all values in degree per second.

	Standing			Sitting		
	yaw	pitch	roll	yaw	pitch	roll
Violin 1	10	8.7	8.5	4.7	7.4	8.1
Violin 2	6.3	8.4	7.3	3.0	7.6	4.8
Viola 1	-	-	-	4.2	4.2	5.2
Viola 2	7.2	7	6.3	8.6	6.0	6.6
Cello 1	-	-	-	0.3	0.4	0.5
Cello 2	-	-	-	0.5	0.3	0.9
Bass	0.7	1.4	1.5	-	-	-
Flute 1	6.1	9	6.1	5.1	7.5	5.6
Flute 2	3.0	6.3	3.9	2.5	5.9	2.9
Oboe 1	6.5	6.9	4.8	3.5	5.2	3.2
Oboe 2	5.5	6.1	4.9	3.5	5.8	2.9
Clarinet 1	11.4	6.5	6.3	7.3	11.1	4.8
Clarinet 2	2.1	3.1	1.7	2.4	3.7	1.6
Bassoon 1	4.3	3.8	3.9	2.5	3.8	2.0
Bassoon 2	2.3	1.8	1.7	1.9	2.7	1.7
Trumpet 1	4.4	4.2	1.8	1.8	2.5	2.3
Trumpet 2	2.4	2.3	1.9	2.2	1.9	1.2
Horn 1	2.2	2.3	1.8	1.4	2.0	0.9
Horn 2	5.2	3.1	3.0	3.6	2.5	1.5
Trombone	0.9	4.9	1.0	0.5	4.0	1.1

Document downloaded by @DAEL, 2026-11-06 04:31:00 - personal use only

Table V. Mean spectral fluctuation over all playable third octave bands and maximum spectral difference for each instrument and playing position.

	Standing						Sitting					
	anechoic			reverberant			anechoic			reverberant		
	mean in dB	max in dB	f_{max} in Hz	mean in dB	max in dB	f_{max} in Hz	mean in dB	max in dB	f_{max} in Hz	mean in dB	max in dB	f_{max} in Hz
Violin	9.7	17.8	500	2.8	5.6	10000	8.1	14.5	3150	2.7	6.0	6300
Viola	8.8	18.1	800	2.0	3.7	500	7.8	15.8	800	1.7	3.3	500
Cello	-	-	-	-	-	-	0.9	3.5	630	0.4	0.8	10000
Bass	1.6	4.7	2000	0.4	0.6	3150	-	-	-	-	-	-
Flute	9.2	15.3	8000	3.7	10.0	8000	9.1	15.2	8000	3.6	9.7	8000
Oboe	10.1	20.1	6300	2.1	5.9	8000	7.6	14.6	6300	1.4	4.9	8000
Clarinet	5.7	11.6	1600	2.1	5.3	6300	5.7	12.3	6300	2.3	6.6	6300
Bassoon	5.1	8.9	2000	1.7	3.6	16000	5.5	11.9	2000	1.8	4.1	16000
Trumpet	4.4	7.7	1600	4.0	5.5	1600	2.9	5.4	1600	2.9	4.2	1600
Horn	7.7	12.4	5000	1.6	3.2	16000	5.9	10.3	5000	1.2	2.3	16000
Trombone	2.3	8.1	2000	2.9	4.9	2000	1.9	7.1	2000	2.4	4.2	2000

Table VI. 5%/95% quantile ranges of room acoustic parameters according to ISO 3382-1. Values exceeding the JND (not available for L_J) are bold-faced. stn: standing, sit: sitting.

	T_{20} in ms		C_{80} in dB		D_{50} in %		EDT in ms		G in dB		L_J in dB		J_{LF} in %	
	stn	sit	stn	sit	stn	sit	stn	sit	stn	sit	stn	sit	stn	sit
Violin	20	17	**2.0**	**1.6**	**12.2**	**9.8**	35	42	**1.8**	**1.4**	4.3	4.4	**8.4**	**7.4**
Viola	20	20	**3.4**	**3.4**	**24.2**	**23.6**	**77**	54	**4.9**	**4.5**	4.3	3.7	**7.7**	**6.4**
Cello	-	6	-	0.3	-	1.9	-	34	-	0.4	-	0.4	-	2.4
Bass	2	-	0.3	-	1.8	-	20	-	0.4	-	0.6	-	0.5	-
Flute	29	30	**2.9**	**2.9**	**12.6**	**13.0**	**461**	**461**	**1.9**	**1.9**	3.7	3.8	**16.4**	**16.4**
Oboe	3	3	**1.0**	0.6	**6.8**	3.8	**57**	29	**1.2**	0.7	2.4	1.7	3.0	2.0
Clarinet	6	7	**1.1**	**1.3**	**7.9**	**9.6**	50	50	0.6	0.5	1.4	1.8	1.8	1.4
Bassoon	23	31	**1.4**	**1.3**	**9.7**	**8.9**	35	31	**1.4**	**1.4**	3.9	4.3	**6.3**	**7.5**
Trumpet	18	16	**2.3**	**1.6**	**9**	4.8	**360**	**242**	**1.1**	0.7	3.8	2.4	4.7	2.8
Horn	6	3	**1.6**	**1.2**	**8.1**	**5.8**	**147**	**125**	**1.2**	0.8	4.3	3.1	**8.9**	**7.6**
Trombone	18	13	**1.5**	**1.3**	3.6	2.8	**152**	**155**	0.6	0.5	3.3	2.7	1.5	1.1

field tends to smooth out movement-induced sound field modulations. Nevertheless, almost all modulations of C_{80}, D_{50}, J_{LF}, EDT, and also modulations of the strength factor G are above the related JNDs, except for the cello and the double bass, due to their limited range of movement.

Since the respective JNDs are usually determined based on paired comparisons [27] rather than the gradual modulations provided by the movements of musical instruments, we can not consider these values, describing the range of conditions over the total duration of the musical segment, as a direct proof of the audibility of these changes. However, as illustrated by Figure 4, these movements are not *slow*. They often exhibit a rather periodic behavior, in which the entire range of motion is exhausted in periods of 2–3 s.

3.5. Perceptual evaluation

The detection rates of the 2AFC test for each participant and every condition are given in Figure 8. The anechoic auralization of a static sound source and a sound source in motion could be distinguished by almost every subject. Only two participants missed the significance threshold of 18 correct answers by one. As could be expected based on the related sound field differences, the detection rates for the auralisation in reverberant environment were slightly lower than for anechoic conditions. In 6 out of 33 tests (3 musical instruments × 11 subjects), the participants could not reliably identify the movement-induced differences. In 3 cases, they missed the significance threshold by only 1 or 2 corrects answers, in 3 cases they were close to the guessing rate of 50%.

The pooled detecting rates across all participants and test conditions are given in Figure 9. Detection rates are highly significant ($p < 0.01/6 = 0.0017$) if at least 151 correct answers were observed, which is true for all conditions. The trumpet in the anechoic acoustical environment shows the highest detection rate of 99.6%, whereas the auralization of the violin in the reverberant acoustical environment lead to the lowest detection rate of 88.9%.

To analyze effects between test conditions statistically, the pooled data was analyzed by means of a generalized linear mixed model (GLMM) [28]. A highly significant difference in correct answers was found between the

Document downloaded by @DAEL, 2026-11-06 04:31:00 - personal use only

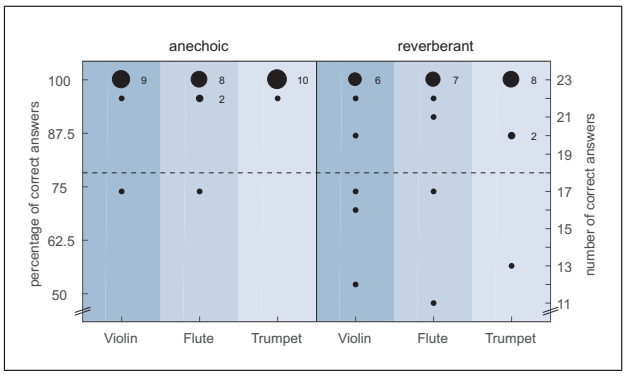


Figure 8. (Colour online) Detection rates of the 2AFC test for all participants and test conditions. The size of the dots and the numbers next to them indicate how many participants scored identical results. Results on or above the dashed line are significantly above chance, indicating that differences between static and moving sources were audible.

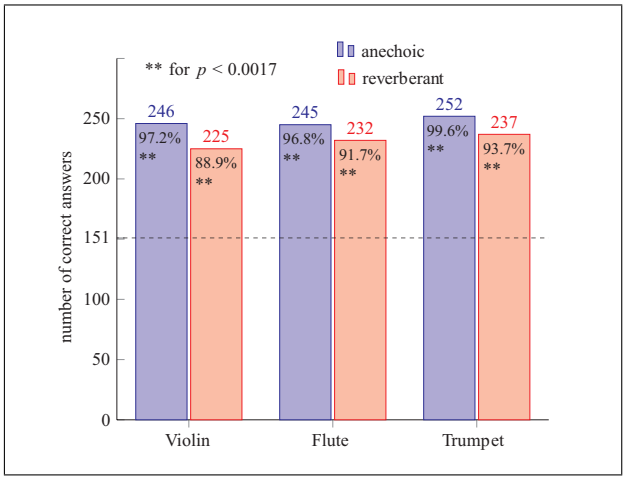


Figure 9. (Colour online) Pooled detection rates of the 2AFC test across all participants for each test condition. Results on or above the dashed line are highly significantly above chance, indicating that differences between static and moving sources were audible.

two acoustical conditions, with detection rates 98% vs. 91% for anechoic and reverberant condition ($F(1.1514) = 30.29$, $p < 0.01$). A significant effect in detection performance was also found between the instruments across the rooms, however only between trumpet (97%) and violin (93%), with ($F(2.1514) = 3.74$, $p < 0.05$). No significant difference in detection performance was found between trumpet and flute and violin and flute with the given sample size and test power ($p = 0.40$). No significant effect was found for an interaction between instruments and rooms.

3.6. Primary research data

All primary research data of the current study are available as a digital publication [29]. It includes the motion tracking data (position in x/y/z coordinates, rotational values in quaternions) and the audio recordings in .flac format, as well as the 3D model of the performance venue used (Theater an der Wien) in Sketchup format with source and receiver position, and the acoustic properties of the surfaces (absorption, scattering). It also contains the stimuli used for the listening test, i.e. the simulation of the performance with and without the musicians' movements.

4. Discussion

Inspired by previous studies indicating the relevance of time-dependent directivities in auralisation [10, 11, 12], we quantified the sound field modulations induced by the movements of musicians during a musical performance for all standard orchestral instruments. As a measure for these modulations, we used the spectral fluctuations in third-octave bands and the fluctuations of room acoustical parameters in a typical performance space. The evaluation was based on previously measured directivities in high resolution and on the motion tracking of 20 different professional musicians playing 11 different musical instruments.

In comparison with the free-field condition, in which movements can induce spectral modulations in single third-octave bands of up to 20 dB, the diffuse field tends to smooth out the time-variant spectral changes, reducing the modulations to 0.6–10 dB in single third-octave bands, and 0.4–4.0 dB as a mean of all third-octave bands. Only in special situations, when strongly directional sources coincide with a heterogeneous absorption of the room, the spectral modulations can be enhanced by sound reflections. This was the case for the trombone, with its main lobe directed sometimes on and sometimes above the highly absorptive audience area, thus leading to a slightly increased spectral modulation in the reverberant condition.

In a medium-size performance venue with $T_{20} = 1.0$ s, the modulations of room acoustical parameters evaluated at twice the critical distance are well above the generally assumed just noticeable differences for all parameters evaluating the early part of the room impulse response, such as the Early Decay Time (EDT), the clarity measures (C_{80}, D_{50}), and the early lateral energy fraction (J_{LF}), as well as for the strength (G).

The audibility of the combined effect of spectral modulations and changes in room acoustic parameters was confirmed by a listening test, where listeners could reliably distinguish between the static and the dynamic auralization of three musical performances. The test was done with representatives of three orchestral groups (a violin, a flute, a trumpet), and auralizations generated with a virtual sound source controlled by the motion tracking data of the corresponding musical performermances. Since both effects (spectral modulation and time-variant excitation of the room) increase with the directionality of the source and the extent of movements of the performer, it can, of course, not be expected that these are audible for any musical instrument, any performer, and any time instance of a musical performance. The listening test, however, proves the audibility for three different instruments, of which two (violin and flute) are not particularly directional (see directivity indices in Figure 6). We would thus assume, that

Document downloaded by @DAEL, 2026-11-06 04:31:00 - personal use only

the effects are audible for most, if not all, orchestral instruments, on the condition that the musicians are not particularly static in their performance and the audio content has sufficient spectral bandwidth to make these modulations audible.

The results demonstrate that movements as constitutive elements of the gestural performance of instrumental music do not only have a visual impact on the perceived musical event, but are also communicated in the acoustical domain. Thus, it seems indispensable to consider the movement- induced, dynamic behavior of musical instruments and musical performances also in the auralization of natural acoustic sound sources, if these are supposed to represent *all* perceptually relevant aspects of the acoustic event.

From a music psychological perspective, visual cues have repeatedly been shown to have an influence on judgements about music performances that we may think are purely auditory [30, 31]. Our investigation shows that even if you can *not* see the source you can detect the difference between a static and a dynamic musical performance. This effect will certainly be enhanced by additional visual cues, if these are consistent in character with the auditory stimulus.

The auralization prepared for the listening test of the current study has demonstrated one approach for the rendering of dynamic sources, by using the motion tracking data of a musical performance to control the orientation of a previously measured directivity. Compared to the multichannel auralisation suggested by [11], our method does not require multi-channel recordings while, at the same time, achieving a potentially much higher spatial resolution than those typically available with multi-channel recordings. The directivities used for the current study were measured with a 32-channel spherical microphone array and are openly available for similar kinds of applications [6].

If the dynamics of the source can not be represented in the room acoustical simulations, calculating a spatial average of the directivity according to the typical range of musicians' movements might be preferable to the static application of directivities in high spatial resolution.

Future work should now be dedicated not only to the quantity but also to the qualities which are 'acoustically embodied' in the gestural performance of music. It might turn out that the sometimes observed lack of liveliness in auralizations is, at least in part, due to the static representation of musical instruments in virtual music performances.

Acknowledgement

This work was funded by the German Research Fundation (grant WE 4057/3-2). The authors would like to thank Dominik Steger for acquiring the tracking data and Steffen Lepa for his advice on statistical issues. We would like to thank the Staatliches Institut für Musikforschung, Berlin, for providing the concert hall for the motion tracking of musical performances, and the 20 musicians for their willingness to participate in the study.

References

[1] J. Meyer: Acoustics and the Performance of Music - Manual for Acousticians, Audio Engineers, Musicians, Architects and Musical Instrument Makers. 5th Edition. Springer, New York, 2009.

[2] P. Cook, D. Trueman: A database of measured musical instrument body radiation impulse responses, and computer applications for exploring and utilizing the measured filter functions. Proc. of Intl. Symp. Musical Acoustics, Leavenworth WA, USA, 1998, 303–308.

[3] F. Otondo, J. H. Rindel: The influence of the directivity of musical instruments in a room. Acta Acust united Ac **90** (2004) 1178–1184.

[4] F. Hohl: Kugelmikrofonarray zur Abstrahlungsvermessung von Musikinstrumenten. Diplomarbeit, IEM Institut für Elektronische Musik und Akustik, Universität für Musik und darstellende Kunst, Graz, Austria, 2009.

[5] J. Pätynen, T. Lokki: Directivities of Symphony Orchestra Instruments. Acta Acust united Ac **96** (2010) 138–167.

[6] S. Weinzierl, M. Vorländer, G. Behler, F. Brinkmann, H. von Coler, E. Detzner, J. Krämer, A. Lindau, M. Pollow, F. Schulz, N. R. Shabtai: A Database of Anechoic Microphone Array Measurements of Musical Instruments. http://dx.doi.org/10.14279/depositonce-5861.2 (2017)

[7] N. R. Shabtai, G. Behler, M. Vorländer, S. Weinzierl: Generation and analysis of an acoustic radiation pattern database for forty-one musical instruments. J. Acoust. Soc. Am. **141** (2017) 1246–1256.

[8] L. M. Wang, M. C. Vigeant: Evaluations of output from room acoustic computer modeling and auralization due to different sound source directionalities. Applied Acoustics **69** (2008) 1281–1293.

[9] R. San Martin, I. B. Witew, M. Arana, M. Vorländer: Influence of the Source Orientation on the Measurement of acoustic parameters. Acta Acust united Ac **93** (2007) 387–397.

[10] M. Noisternig, B. F. Katz, C. D'Alessandro: Spatial rendering of audiovisual synthetic speech use for immersive environments. J. Acoust. Soc. Am. **123** (2008) 3939–3939.

[11] F. Otondo, J. H. Rindel: A New Method for the Radiation Representation of Musical Instruments in Auralizations. Acta Acust united Ac **91** (2005) 902–906.

[12] B. N. Postma, H. Demontis, B. F. Katz: Subjective Evaluation of Dynamic Voice Directivity for Auralizations. Acta Acust united Ac **103** (2017) 181–184.

[13] B. W. Vines, C. L. Krumhansl, M. M. Wanderley, I. M. Dalca, D. J. Levitin: Music to my eyes: Cross-modal interactions in the perception of emotions in musical performance. Cognition, **118** (2011) 157–170.

[14] E. C. Teixeira, M. A. Loureiro, M. M. Wanderley, H. C. Yehia, H. C.: Motion analysis of clarinet performers. Journal of New Music Research, **44** (2015) 97–111.

[15] M. Leman: Embodied music cognition and mediation technology. MIT Press, 2008.

[16] P. J. Maes, M. Leman, C. Palmer, M. Wanderley: Action-based effects on music perception. Frontiers in psychology **4** (2014), article 1008.

[17] M. M. Wanderley, P. Depalle: Gestural control of sound synthesis. Proceedings of the IEEE, **92** (2004) 632–644.

[18] F. Wefers: OpenDAFF. A free, open-source software package for directional audio data. Fortschritte der Akustik: Tagungsband der 36. DAGA, Berlin, Germany, 2010, pp. 1059–1060.

[19] D. Schröder, M. Vorländer: RAVEN: A Real-Time Framework for the Auralization of Interactive Virtual Environ-

Document downloaded by @DAEL, 2026-11-06 04:31:00 - personal use only

ments. Proc. of Forum Acusticum, Aalborg, Denmark, 2011, 1541–1546.

[20] ISO 3382-1:2009 Acoustics - Measurement of room acoustic parameters - Part 1: Performance spaces, 2009.

[21] F. Brinkmann, A. Lindau, S. Weinzierl, G. Geissler, S. van de Par: A High Resolution and Full-Spherical Head-Related Transfer Function Database for Different Head-Above-Torso Orientations. J. Acoust. Soc. Am. **65** (2017) 841–848.

[22] F. Brinkmann, A. Lindau, S. Weinzierl, G. Geissler, S. van de Par, M. Müller-Trapet, R. Opdam, M. Vorländer: The FABIAN head-related transfer function data base. http://dx.doi.org/10.14279/depositonce-5718 (2017)

[23] F. Brinkmann, S. Weinzierl: AKtools - An Open Software Toolbox for Signal Acquisition, Processing, and Inspection in Acoustics. 142nd AES Convention, Berlin, Germany, 2017, e-Brief 309.

[24] M. Berzborn, R. Bomhardt, J. Klein, J. G. Richter, M. Vorländer: The ITA-Toolbox: An Open Source MATLAB Toolbox for Acoustic Measurements and Signal Processing. Fortschritte der Akustik: Tagungsband der 43. DAGA, Kiel, Germany, 2017, 222–225.

[25] A. Lindau, F. Brinkmann: Perceptual evaluation of headphone compensation in binaural synthesis based on non-individual recordings. J. Audio Eng. Soc. **60** (2012) 54–62.

[26] IEC 61672-1:2013, Electroacoustics - Sound level meters - Part 1: Specification, 2013.

[27] I. B. Witew, G. K. Behler, M. Vorländer: About just noticeable differences for aspects of spatial impressions in concert halls. Acoustical science and technology **26** (2005) 185–192.

[28] W. W. Stroup: Generalized linear mixed models: modern concepts, methods and applications. CRC press, Boca Raton, 2013.

[29] D. Ackermann, C. Böhm, S. Weinzierl: A Database on Musicians' Movements During Musical Performances. http://dx.doi.org/10.14279/depositonce-7469 (2018)

[30] C.-J. Tsay: Sight over sound in the judgment of music performance. Proceedings of the National Academy of Sciences **110** (2013) 14580–14585.

[31] C. Ryan, E. Costa-Giomi: Attractiveness bias in the evaluation of young pianists' performances. Journal of Research in Music Education **52** (2004) 141–154.

Document downloaded by @DAEL, 2026-11-06 04:31:00 - personal use only