CHALMERS
UNIVERSITY OF TECHNOLOGY

Perceptual Significance of Tone-Dependent Directivity Patterns of Musical Instruments

Citation for the original published paper (version of record):
Corcuera Marruffo, A., Chatziioannou, V., Ahrens, J. (2023). Perceptual Significance of Tone-Dependent Directivity Patterns of Musical Instruments. AES: Journal of the Audio Engineering Society, 71: 293-302. http://dx.doi.org/10.17743/jaes.2022.0076

A. Corcuera, V. Chatziioannou, and J. Ahrens, "Perceptual Significance of
Tone-Dependent Directivity Patterns of Musical Instruments"
J. Audio Eng. Soc., vol. 71, no. 5, pp. 293–302, (2023 May).
DOI: https://doi.org/10.17743/jaes.2022.0076

Perceptual Significance of Tone-Dependent Directivity Patterns of Musical Instruments

ANDREA CORCUERA,[1] *AES Student Member*,
(corcuera-marruffo@mdw.ac.at)
VASILEIOS CHATZIIOANNOU,[1] **AND JENS AHRENS,**[2] *AES Member*
(chatziioannou@mdw.ac.at) (jens.ahrens@chalmers.se)

[1] *University of Music and Performing Arts, Vienna, Austria*
[2] *Chalmers University of Technology, Gothenburg, Sweden*

Musical instruments are complex sound sources that exhibit directivity patterns that not only vary depending on the frequency, but can also change as a function of the played tone. It is yet unclear whether the directivity variation as a function of the played tone leads to a perceptible difference compared to an auralization that uses an averaged directivity pattern. This paper examines the directivity of 38 musical instruments from a publicly available database and then selects three representative instruments among those with similar radiation characteristics (oboe, violin, and trumpet). To evaluate the listeners' ability to perceive a difference between auralizations of virtual environments using tone-dependent and averaged directivities, a listening test was conducted using the directivity patterns of the three selected instruments in both anechoic and reverberant conditions. The results show that, in anechoic conditions, listeners can reliably detect differences between the tone-dependent and averaged directivities for the oboe but not for the violin or the trumpet. Nevertheless, in reverberant conditions, listeners can distinguish tone-dependent directivity from averaged directivity for all instruments under study.

0 INTRODUCTION

Accurate representation of the directivity characteristics of sound sources is fundamental to achieve authentic simulations of virtual acoustic environments [1]. Variations in the directivity characteristics of a source can potentially influence the perceived localization [2] and auditory distance [3]. Sound sources, such as the human voice, loudspeakers, or musical instruments, have distinctive directivity patterns that vary significantly across the frequency range and can also change depending on other aspects.

Several researchers have measured the directivity of musical instruments, dating back to the 1970s with the pioneering work of Meyer [4], who conducted an extensive investigation on the radiation of numerous musical instruments. Further studies have measured the directivity of musical instruments, generally using repeated capturing methods with artificial excitation that allow the directivity of the instrument to be obtained at high resolution [5, 6] or spherical microphone arrays that allow measurements to be made in a natural situation with a musician [7, 8].

Studies on the influence of source directivity on the perception of acoustic simulations have demonstrated that listeners are able to perceive differences caused by different directivity representations. Wang and Vigeant [9] demonstrated that subjects can distinguish between omnidirectional and extremely directional sources. Otondo and Rindel [10] showed that varying the directional characteristics of sound sources affects the room acoustic parameters and can lead to audible differences in terms of loudness, reverberance, and clarity.

Musical instruments are dynamic sources, such as the voice, whose directivity varies according to the movement, the played tone in the case of musical instruments [11], or the phonemes in the case of the voice [12]. One of the first investigations on the perceptual implications of such dynamic characteristics of musical instruments on auralizations was carried out by Otondo and Rindel [10]. They showed that listeners can perceive changes caused by different directivity representations (averaged or tone-dependent directivity patterns) in static auralizations. However, to evaluate the audibility of different directivity representations, the authors only tested one tone-dependent directivity pattern with melodies that did not always include the tone that corresponded to the directivity pattern studied.

More recently, Ackermann et al. [13] demonstrated that the fluctuations created by the movement of the musicians during solo musical performances are audible both under anechoic and reverberant conditions. Similarly, Ehret et al. [14] performed a perceptual evaluation involving static and dynamic phoneme-dependent voice directivities. They showed that participants were not able to distinguish phoneme-dependent directivities from averaged directivities and that their subjective preference might not be dependent on the realism of the directional rendering.

To better understand the perceptual requirements of musical instrument directivities in virtual acoustic environments, this paper analyzes the differences between tone-dependent directivities and the directivity averaged over all tones. To validate the classification of the instruments into three categories proposed in [15], the patterns of several musical instruments were derived and analyzed based on their maximum directivity index averaged over all tones, the variation of maximum directivity per tone and the similarity of the main radiation region among tones. After selecting a single instrument representative of each category, the spectral differences between the averaged and tone-dependent directivities were estimated and discussed. Additionally, the spectral differences with respect to an omnidirectional source were calculated. Subsequently, to test the audibility of source directivity variations, auralizations using omnidirectional, averaged, and tone-specific directivities of the three selected instruments were evaluated in a listening test in a virtual environment with anechoic and reverberant conditions.

This paper is structured as follows: SEC. 1 analyzes the directivities of 38 musical instruments based on three categories. SEC. 2 describes the stimuli utilized in the listening test and SEC. 3 presents the results. SEC. 4 further discusses the results and draws conclusions.

1 SORTING OF MUSICAL INSTRUMENTS BASED ON TECHNICAL UNIVERSITY OF BERLIN DATABASE

1.1 Instrument Database

The analysis of directivity patterns in this study is based on the measurements from the open-access Technical University of Berlin (TU Berlin) database [15]. This database contains scales and single-tone recordings of 41 symphonic orchestral instruments at two dynamic levels (*pianissimo* and *fortissimo*), along with their calculated directivities and audio features.

The instruments were recorded at the anechoic chamber of the TU Berlin using a spherical array of radius 2.1 m, consisting of 32 microphones placed on the faces of a truncated icosahedron [8]. By using a microphone array, the instruments could be measured in a performance situation, which allowed the acoustic effect of the musician and the natural excitation of the source to be included in the measurements. The resolution of the measurements is limited by spatial aliasing, which is apparent over large parts of the frequency range of the instruments. Nevertheless, the measured data can be interpolated to a higher spatial grid using, for instance, spherical harmonic (SH) decomposition.

1.2 Overall Analysis of Musical Instruments

In order to investigate the differences between time-varying (tone-specific) and static (averaged) directivities, and to reach general conclusions about symphonic instruments, a set of instruments was selected from groups with similar radiation characteristics. The conventional classification divides the symphonic musical instruments into four groups or families: strings, woodwind, brass, and percussion instruments. However, this and other traditional classifications are not based on the radiation of the instruments but on other criteria, such as the morphology of the instruments or the way the sound is generated [16, 17]. Shabtai et al. [15] made a preliminary sorting into three groups depending on how the instruments radiate sound (see Table 1). To validate this classification, this section presents a general analysis of the musical instruments according to their maximum directivity index, variation of maximum directivity per tone, and similarity of the principal radiation regions.

From the TU Berlin dataset, 38 musical instruments were selected for analysis, all but the timpani, which did not contain single-note recordings, and the singer, which was excluded to focus on musical instruments. The recordings of single tones in *ff* were used for the analysis to guarantee a good signal-to-noise ratio. Different methods can be used to calculate the directivity patterns from the multichannel recordings [18, 19].

For the overall analysis of the instruments, this study follows a directivity derivation procedure based on the extraction of the partials of each tone and the definition of the directivity as the spectral envelope of the partials, similar to the one described in [15, 19]. To identify the fundamental and overtone frequencies, the stationary parts of each single-tone recording, provided by the authors of the database in sample indices, were windowed using a Hamming window and transformed to the frequency domain. Then, the magnitudes of the partials below 10 kHz and higher than –50 dB were extracted and averaged over one-third octave bands.

1.2.1 Maximum Directivity Index (DI_{max})

As a measure of the overall degree of directionality of the instruments studied, the maximum directivity index (DI_{max}) of the averaged directivity was calculated. To calculate this metric, the directivity averaged over all tones was used and smoothed into one-third octave bands. In this study, the directivity index (DI) is defined as the ratio between the sound power at a certain direction and the average power over all measured directions [20]. The DI of a source indicates the extent to which the source's radiation is biased toward a

Table 1. List of musical instruments belonging to a specific category (suggested by [15]).

Category	ID	Instruments
I	1	Alto trombone historical
	2	Bass trombone historical
	3	Bass trombone modern
	4	Basset horn
	5	English horn
	6	French horn
	7	Natural horn
	8	Trumpet historical
	9	Tenor trombone historical
	10	Tenor trombone modern
	11	Trumpet modern
	12	Tuba
II	13	Alto saxophone modern
	14	Baroque bassoon
	15	Baroque transverse flute
	16	Bass clarinet
	17	Modern bassoon
	18	Clarinet historical
	19	Modern clarinet
	20	Classic bassoon
	21	Classic oboe
	22	Contrabassoon
	23	Dulcian
	24	Historical transverse flute
	25	Modern oboe
	26	Romantic oboe
	27	Tenor saxophone
	28	Modern transverse flute
III	29	Acoustic guitar
	30	Historical cello
	31	Modern cello
	32	Harp
	33	Historical double bass
	34	Modern double bass
	35	Historical viola
	36	Modern viola
	37	Historical violin
	38	Modern violin

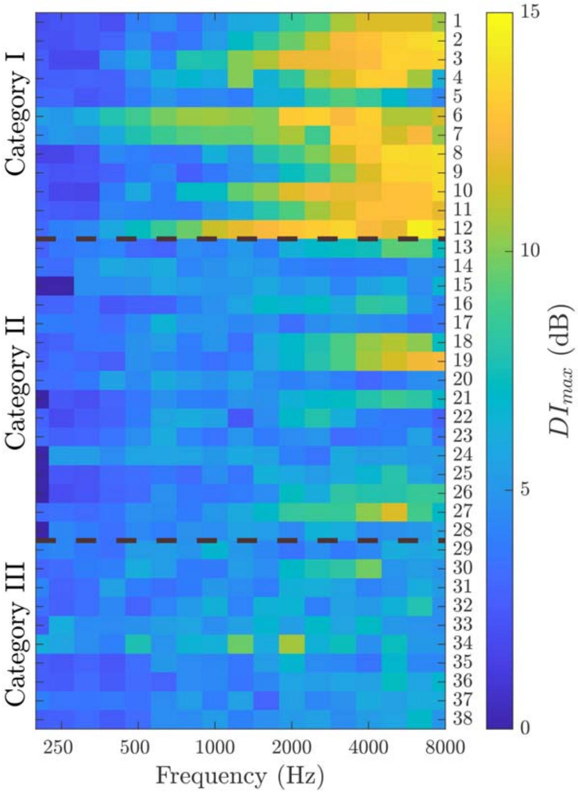


Fig. 1. Maximum directivity index (DI_{max}) of the TU Berlin database, grouped in three categories suggested by Shabtai et al. in [15]. The names of the instruments associated to the ID number can be found in Table 1.

certain direction, as a function of angle and frequency. It is defined in decibels as

$$DI_{\theta,\phi}(f) = 10\log_{10}\left(\frac{\left|p_{\theta,\phi}(f)\right|^2}{\tilde{p}(f)}\right), \quad (1)$$

with $|p_{\theta,\ \phi}(f)|^2$ being the power at azimuth and elevation angles θ and ϕ, and $\tilde{p}(f)$ the average power over all L directions $\tilde{p}(f) = \frac{1}{L}\sum_{\theta,\phi}\left|p_{\theta,\phi}(f)\right|^2$. The *DI* was calculated for each instrument, frequency band, and measurement position. Then the DI_{max} values were obtained for each frequency band by selecting the highest *DI* value from all directions ($DI_{max} = max(DI_{\theta,\ \phi})$).

Fig. 1 shows the DI_{max} of all selected instruments, grouped into the three categories proposed by Shabtai et al. Each row in the figure corresponds to an instrument, shown in the order and with the ID specified in Table 1. In general, brass and many woodwind instruments present low DI_{max} at low frequencies and higher DI_{max} as the frequency increases. Musical instruments in the Category I (all brass instruments, English horn, and basset horn) show the highest DI_{max}, which increases considerably with frequency. However, the English horn, with a constant DI_{max} value over the entire frequency range, does not show the same behavior as the rest of the instruments in this category. It is therefore surprising that this instrument falls into the same category as all brass instruments.

Some woodwinds (tenor saxophone, the modern and classical clarinets) also exhibit high DI_{max} at frequency bands above 3,000 Hz, but to a lesser extent than brass instruments. In contrast, strings and some woodwinds, such as the flute, tend to exhibit low DI_{max} values over the entire frequency range, suggesting that they are less unidirectional (they radiate less in a single direction, like brass instruments). It should also be noted that although measurements were carefully done with the instruments pointing in a specific direction (for example, brass pointing at one specific microphone), the DI_{max} of some instruments may vary slightly, as the measurement point may not coincide to the maximum point of radiation of the instrument.

1.2.2 Variation of Maximum Directivity

To describe the variability of the directivity patterns per tone, the variance of maximum directivity among tones is calculated. That is, the variance of the maximum directivity index of each tone-specific directivity was determined. Low variance indicates a source with a similar DI_{max} per note.

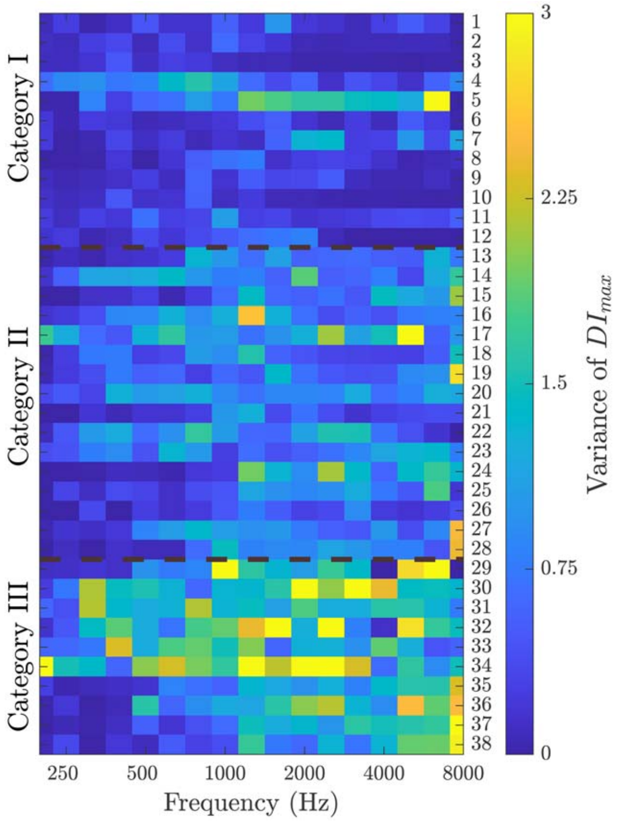


Fig. 2. Variance of DI_{max} among notes of the instruments in the TU Berlin database, grouped in three categories suggested by Shabtai et al. in [15].

Because not all tones have partials in every frequency band, the number of tones included in the calculation varies from one frequency band to another. Hence, the analysis of each frequency band was limited to tones with partials in those bands.

As shown in Fig. 2, it is clear that the directional characteristics of the different tones of most brass instruments are almost the same (maximum variance <1 dB) regardless of the tone being played. The English horn (maximum variance 2.6 dB), and to a lesser degree the Basset horn (maximum variance 1.6 dB), show more variation among tones. Category III instruments show the greatest variation of all instruments in the dataset (maximum variance between 2.2 and 8.1 dB), suggesting that the directional characteristics of stringed instruments change the most as a function of tone played. Among them, the acoustic guitar (maximum variance 4.7 dB), the harp (maximum variance 8.1 dB), and the modern double bass (maximum variance 4.2 dB) show the most variance. Category II instruments, on the other hand, show slightly less variation (maximum values between 1.2 and 3.2 dB) than category III instruments, with the highest variation found in the classic bassoon (maximum 3.2 dB).

1.2.3 Similarity of Principal Radiation Region

Following Meyer's methodology [4], the authors define the principal radiation region as the region for which the normalized directivity pattern does not drop more than 3 dB relative to its maximum value, as a function of frequency. This approach using the radiation region is also used by Pezzoli [21] to compare the sound radiation of historical violins. The principal radiation region is calculated for each tone and frequency band as

$$R_t = \begin{cases} 1, d(\theta_l, \phi_l) > \kappa \\ 0, \text{otherwise} \end{cases}, \tag{2}$$

where θ is the azimuth and ϕ the elevation angle of the l microphone, t is the tone, $\kappa = -3$ is the threshold in decibels, and $d(\theta_l, \phi_l)$ is the normalized directivity in decibels.

For each frequency band, the similarity index (*SI*) is used to compare the patterns of different tones. The *SI* is calculated by dividing the principal radiation region by the total number of tones with partials on that direction as

$$SI = \frac{\sum_t R_t}{B_f}, \tag{3}$$

where B_f is the total number of tones per frequency band f_C (referred as f for notational simplicity), obtained by adding the number of tones with partials laying on that band.

Finally, the averaged similarity index $\overline{SI}$ is calculated by averaging over all L directions.

$$\overline{SI} = \frac{\sum_{l=1}^{L} SI}{L}. \tag{4}$$

Instruments with the exact same directivity pattern for all the tones would have $\overline{SI} = 1$, whereas for an instrument with changing patterns depending on the tone, the $\overline{SI}$ would be close to zero. The values of the $\overline{SI}$ for each instrument and frequency band are shown in Fig. 3.

Overall, all instruments in category I, except the English horn, show the highest $\overline{SI}$ in most bands, suggesting that for the given bands the tone-dependent directivity patterns match for all or most of the tones. Some instruments in category II and III also show high $\overline{SI}$ values in frequencies up to about 500 Hz. These results suggest that the given instruments exhibit a similar radiation behavior at low frequencies.

Based on the values of the DI_{max}, variation of DI_{max} per tone, and $\overline{SI}$, the preliminary sorting proposed by Shabtai et al. seems to be adequate in almost all cases, except in the case of the English horn. This instrument, with a lower and steady DI_{max} at all frequencies and a generally low $\overline{SI}$, is more similar to instruments in the Category II.

2 TONE-DEPENDENT DIRECTIVITY ANALYSIS

This section presents a study of the directivity influence on auralizations. The observed differences between omnidirectional, averaged, and tone-specific directivities are discussed and evaluated in a listening test to determine whether listeners are able to reliably distinguish these auralizations. Additionally, the listening test setup, the simulated room acoustic conditions and test procedures and results are presented.

The listening test and tone-dependent objective analysis were conducted using the open-access database of spherical harmonic (SH) representations of sound sources provided

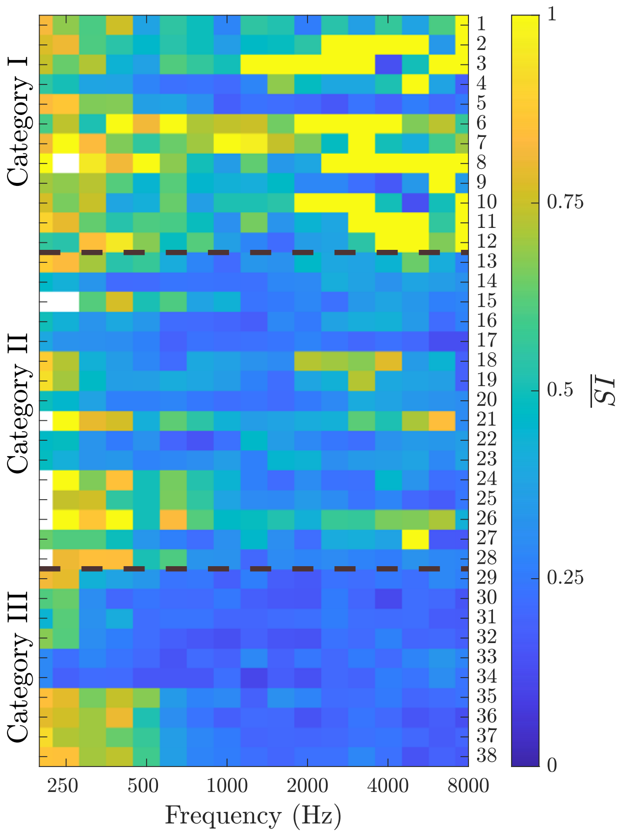


Fig. 3. $\overline{SI}$ for each frequency band and instrument of the TU Berlin database.

in [22]. This database contains impulse responses of various notes of several instruments, based on the TU Berlin measurements [15] and representing the directivity of the sound source in various given discrete directions [23].

To obtain the directivity patterns used for the auralizations, the spherical harmonic representation of fourth-order of the directivities was first obtained from the data using the toolbox provided in [22]. Then, the magnitude directivity was computed from the SH representation in various directions and included in the auralization software. One musical instrument representative of each of the three aforementioned categories was used for the tone-dependent analysis and listening test: a trumpet, oboe, and violin.

2.1 Generating Stimuli for the Listening Test

Auralizations were obtained by convolving anechoic recordings from the three chosen instruments with the simulated Binaural Room Impulse Responses (BRIRs) using their different directivities. Simulations were carried out using the RAVEN [24] software (50,000 rays, second-order image sources), a hybrid algorithm that uses image sources for direct sound and early reflections and ray tracing for late reverberation. The room model was taken from the BRAS database [25, 26]. This model corresponds to the small hall of the *Konzerthaus Berlin*, used for chamber music, with a reverberation time of about $T_m = 1.3$ s.

The source was located on a side of the stage facing the audience. If the listener and the musician were facing each other, the directivity variation in the direct sound would have little effect on the signal as a result of the normalization of the directivity in the direction of the recording microphone. Therefore, the listener was positioned in the audience on one side of the stage, facing the source, at a distance of about 2.5 times the critical distance to accentuate the effect of the room. This configuration was kept constant for all instruments and for both scenes.

The anechoic recordings used for the listening test were obtained from the denoised versions of anechoic orchestral recordings [27] provided in [28]. To avoid colorizing the spectrum, the directivities were normalized by the directivity in the direction of the microphone used in the dry recordings (azimuth $= 0^\circ$ and elevation $= 11^\circ$) [27], obtained from SH interpolation. For the listening test, the authors selected a melody from the recordings of each instrument of about 5 s each.

Static directivity auralizations were obtained by convolving excerpts of anechoic recordings with the BRIRs derived using their corresponding averaged directivity. The averaged directivities were calculated for each instrument by averaging the magnitude across all available tones before performing the normalization by the recording microphone. Using a fourth-order spherical harmonic decomposition, directivity was determined along a spatial grid with a resolution of 5° in azimuth and elevation. In order to include the directivity patterns in the RAVEN software, the magnitude data was smoothed into third-octave bands and encoded in the Open Directional Audio File Format (OpenDAFF) [29].

Time-varying auralizations using tone-dependent directivities require knowing the tone being played at every moment in order to use their corresponding directivity patterns. Therefore, in this study the monophonic pitch tracker CREPE [30] was used to estimate the pitch of the chosen sound excerpts, with a time step of 10 ms.

The output of the pitch tracker contains the timestamps, the predicted fundamental frequency in Hertz, and the confidence (value from 0 to 1). Before using this information to generate the stimuli, the predicted fundamental frequencies with a confidence lower than 0.5 were set to the previous predicted frequency with a higher confidence level. Predicted pitch with frequencies higher than the expected highest frequency per instrument were considered outliers and were replaced by a lower neighboring value. In order to avoid misleading results derived from the use of vibrato in the recordings, the estimated pitch of the anechoic excerpts was smoothed by applying a median filter. The predicted pitch of the excerpts was then manually revised and fixed if needed and linked to their corresponding tones and BRIRs, obtained in the same way as for the averaged version. Finally, tone-specific stimuli were generated by block-wise and time-variant convolution of the anechoic recordings with the BRIRs of each corresponding tone.

2.2 Spectral Analysis

For the analysis of the spectral differences in the direct sound caused by the variation of the source directivity in the auralizations, the impulse responses obtained from

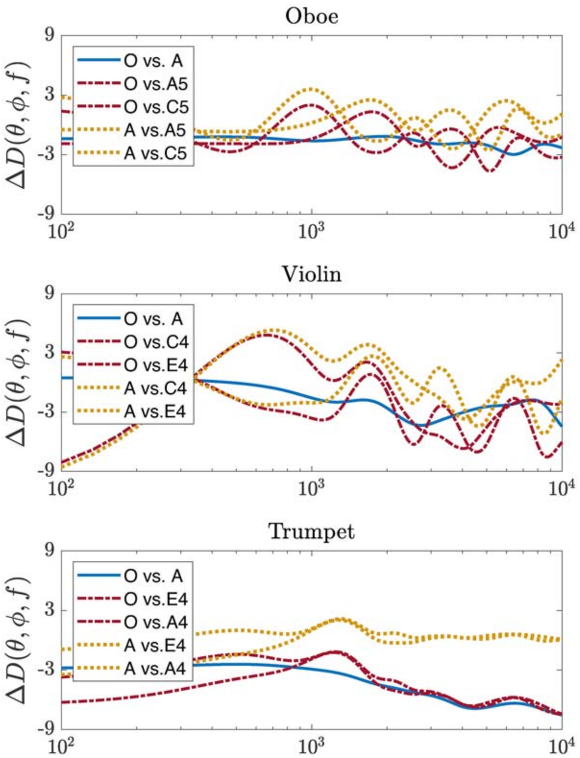


Fig. 4. Spectral differences $\Delta D(\theta, \phi, f)$ between omnidirectional (O), the averaged directivities (A), and the directivity patterns of the first tones of selected melodies for an oboe (C5 and A5), violin (C4 and E4), and trumpet (E4 and A4).

the acoustic simulations were studied and compared. The spectral differences between the tone and the averaged directivities were calculated in decibels as

$$\Delta D(\theta, \phi, f) = 20 \log_{10} \left(\frac{|D_{ton}(\theta, \phi, f)|}{\left|D_{avg}(\theta, \phi, f)\right|} \right), \qquad (5)$$

where $D_{ton}(\theta, \phi, f)$ is the spectrum of the RIR using the directivity pattern per tone and $D_{avg}(\theta, \phi, f)$ is the spectrum of the RIR using the directivity pattern averaged over all tones. The authors calculated the spectral differences $\Delta D(\theta, \phi, f)$ of each tone with respect to the averaged directivity as well as the spectral differences of the averaged directivity with respect to the omnidirectional case.

Fig. 4 shows the spectral differences $\Delta D(\theta, \phi, f)$ of different directivities that were included in the listening test. The plots show that spectral differences between auralizations with omnidirectional and the averaged directivity for the trumpet are higher than 6 dB for high frequencies above 2 kHz. The spectral differences for the oboe and the violin are relatively small, with maximum differences of 3 and 4 dB, respectively. Although similar results are obtained for the comparison between the omnidirectional and tone-specific directivities for the trumpet, the violin and oboe show higher spectral differences, higher than the omnidirectional compared to the averaged. Spectral differences between the averaged and tone-specific directivities of the trumpet are smaller than 1 dB, whereas the oboe and the violin present spectral differences up to about 4 and 5 dB, respectively.

2.3 Listening Test

An ABX listening test was conducted to determine whether the differences between tone-specific and averaged directivities of musical instruments were audible. Three auralization versions were studied: omnidirectional, i.e., with a source emitting sound equally in all directions; averaged, using the averaged directivities of the instruments; and tone-dependent, using the tone-dependent directivity patterns.

Listeners were presented with stimulus A, B and X and two forced answers: X = A or X = B. For each trial, the simulation with the omnidirectional, tone-specific and averaged directivities were randomly assigned to the A and B buttons and one of them was randomly repeated on button X. The participants could listen to the sound samples as often as desired before giving an answer.

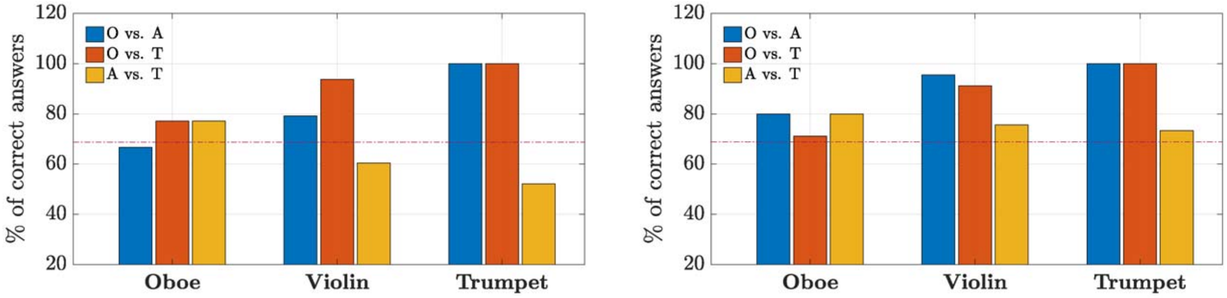
The experiment was conducted in an acoustically damped audio laboratory at the University of Music and Performing Arts, Vienna. A graphical user interface was developed in MATLAB for the playback of the test signals and the collection of participants' responses. Three buttons labeled *A*, *B*, and *X* were used to start the auralizations. Subjects could control playback both with the keyboard and mouse. Stimuli were presented through headphones (Beyerdynamic DT990) without head-tracking, equalized with a compensation filter [31, 32], and with the same playback level for all listeners.

To familiarize themselves with the test procedure and stimuli, participants underwent a training session with three conditions (one per instrument) prior to the listening test. Each participant was presented with a total of 54 test trials, 27 trials (3 instruments × 3 directivities × 3 repetitions) per scene (anechoic and reverberant). In the test, the order of the conditions was randomized for each listener, but during training the conditions were presented in the same order for all participants. After completing the listening test, participants were asked to fill in a questionnaire indicating in their own words which auditory cues they had used for discriminating between sounds. The experiment took on average 40 min per participant, including instructions, training, and an optional short break.

A total of 16 listeners, six men and ten women, aged 21–49 years (average age of 29.5 years) participated in the listening experiment. All of them reported normal hearing and had at least 10 years of musical experience (19.4 years on average) or extensive experience with listening tests, and were therefore considered as trained listeners. Written informed consent was received from all participants at the beginning of the session.

3 RESULTS

The results of the listening test, given in percentage of correct answers, are summarized in Fig. 5 for all participants and test conditions. Results of one participant were discarded from the reverberant scene analysis due to a tech-



(a) Anechoic

(b) Reverberant

Fig. 5. Percentage of correct answers for all participants for comparisons between auralizations using an omnidirectional (O) source, the averaged directivity (A), or the tone-specific directivity (T) of the instruments. The horizontal dash-dotted line shows the threshold for statistical significance.

nical issue. The results of all listeners were pooled, yielding a total of 48 answers per test condition for the anechoic scene and 45 for the reverberant scene. To determine the statistical significance of the test results, the authors conducted a binomial test [33]. For a 1% significance level, the critical number of correct answers in order to reject the null hypothesis is 33 [detection accuracy 68.7%, see horizontal line in Fig. 5(a)] for the anechoic condition with 16 participants and 31 [detection accuracy 68.9%, see horizontal line in Fig. 5(b)] for the reverberant condition with 15 participants. Therefore, if the number of correct answers is above the critical value, the differences between the conditions are considered to be significant.

As shown in Fig. 5(a), the results for the anechoic environment indicate a clear distinction between omnidirectional (O) and averaged (A) directivity (leftmost bars), and between omnidirectional and tone-specific (T) directivity (middle bars) of the trumpet. This instrument shows the highest detection accuracy for the omnidirectional directivity, as all participants were able to correctly detect differences. Differences between A and T are most noticeable on the oboe, whereas the listeners were not able to detect differences in the trumpet, which were not significantly difference from chance. These results follow the same trend found in a previous pilot study in anechoic conditions [34]. Contrary to expectations, listeners could not distinguish between auralizations made with the tone-specific directivity and the averaged directivity for the violin, although the detection accuracy is higher than for the trumpet.

The detection accuracy for both O vs. A and O vs. T follows a trend opposite to that observed in the results between A and T, with the trumpet showing the highest detection accuracy, followed by the violin and the oboe. The violin exhibits very high detection accuracy for T vs. O and a lower but significant accuracy for O vs. A. The oboe, however, shows a lower detection accuracy for O vs. A and O vs. T than the other instruments. These results are in accordance with the study of the spectral differences in SEC. 2.2, which showed that the smallest differences between O and A are found in the oboe.

In general, a similar trend is observed for the reverberant environment as for the anechoic environment, with the trumpet being the instrument in which the largest differences can be distinguished between O vs. A and O vs. T, followed by the violin and oboe. Similarly, the results of T vs. A show the same trend, with the oboe being the instrument with more correct answers, followed by the violin and the trumpet. A chi-square test did not reveal a significant effect of the environment for the oboe ($\chi^2 = 0.44$, $p = 0.51$, $df = 1$), violin ($\chi^2 = 4.83, p = 0.03$, $df = 1$), or trumpet ($\chi^2 = 4.15, p = 0.04, df = 1$). As seen in Fig. 5(b), the pooled detection accuracy is above the critical value (indicated in a red dotted horizontal line) not only for the oboe, but also for the violin and trumpet, indicating that differences between tone-specific and averaged directivity representations are audible.

After completing the listening test, participants wrote in their own words which auditory cues influenced their decisions. All listeners identified timbre or color as the main cue for distinguishing sounds for the anechoic scene (also reported as brilliance, overtones, and how muffled the sounds were). Two participants mentioned that the differences were specifically in the timbre of certain tones, one noted changes in the nuances of the tones, and another participant commented on the difference in richness of the sounds. In the case of the reverberant scene, seven participants indicated that the reverberation (also referred to as the resonances of the room, harmonic reflections, or echo) was a major difference. One participant perceived differences in the sound width, whereas another noted that some sounds *felt emptier* and their clarity was different. Three participants also noted differences in source direction in the case of the trumpet. Most of the participants also mentioned the timbre and loudness as a main cue for detecting differences in the reverberant scene.

4 DISCUSSION

This paper assessed the audibility of differences between auralizations using omnidirectional, tone-dependent directivity patterns and averaged directivity patterns under anechoic and reverberant conditions. To this end, three instruments (an oboe, violin, and trumpet) representative of groups of instruments with similar directivity characteristics were investigated. The spectral deviations between auralizations with different directivity representations were determined. Inspired by previous studies that demonstrated the differences between the individual directivities of the tones [11, 19] and their potential perceptual significance [10], a subsequent listening test was conducted to compare simulations using omnidirectional, tone-specific, and averaged directivities.

Results of the listening test with 16 participants proved the audibility of the tone-specific directivity patterns for the oboe under anechoic conditions, but not for the violin or the trumpet [see Fig. 5(a)]. Results of the trumpet are in line with the literature, demonstrating the strong similarity of the tone-specific directivity patterns of brass instruments [11, 19]. Listeners reported that, under anechoic conditions, they can perceive differences in timbre influenced by directivity variations, and differences in timbre, loudness, and reverberance under reverberant conditions.

Results show that listeners are not able to reliably distinguish tone-dependent directivities from the averaged directivity of a violin and a trumpet under anechoic conditions. These results suggest that auralizations using tone-dependent directivities are not perceptually required for some instruments. However, whether the use of a dynamic directivity representation with tone-dependent directivities would have a perceptual benefit remains to be studied. Therefore, future work could examine not only the detectability of directivity variations, but also whether different directivity representations can influence, for instance, the authenticity of a virtual scene. In future studies, it would be interesting to use head tracking and moving sources along with tone-dependent directivity patterns.

Regarding differences between the omnidirectional source and the directivity of the instrument, listeners can clearly detect differences between omnidirectional and averaged auralizations of the trumpet both in anechoic and reverberant environments. The very directive nature of the trumpet leads to a more distinctively different directivity pattern than the other studied instruments, making these differences less apparent for the oboe and violin in comparison to the trumpet. This is especially noticeable in the case of the oboe, which shows no significant difference between the omnidirectional and averaged directivity pattern in anechoic conditions [see Fig. 5(a)]. This instrument-dependent differentiation may be due to the averaging process over all tones, which may smooth the final averaged pattern. Perhaps the lack of differences observed in auralizations with musical instruments in previous studies [9] may be due in part to overlooking the particularities of the directivity of each of an instrument's tones.

The auralizations used for the present study have demonstrated the audibility of directivity variations in one scene with a listener positioned in a given direction relative to the nominal source orientation. In that particular direction, differences owing to directivity may result from both direct sound and room reflections. If the listeners were positioned in the same direction as the microphone with which the melodies were recorded, then the only directivity that could be heard would be due to early reflections and late reverberation. It was shown in [35] that early reflections are significant for detecting the orientation of a sound source. The results of the present study suggest that subjects were more successful in differentiating the variants of the directivity tested in this study. The analysis revealed that this trend was not highly significant, so determining the role of reverberation in the perception of source directivity requires further research.

5 ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 812719. The authors would like to thank Lukas Aspöck for his support with the RAVEN software and Montserrat Pàmies Vilà for constructive criticism on the manuscript. The authors thank the reviewers for their feedback on an earlier draft of this manuscript.

6 REFERENCES

[1] M. Vorländer, *Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality* (Springer, Berlin, Germany, 2007).

[2] F. Wendt and M. Frank, "On the Localization of Auditory Objects Created by Directional Sound Sources in a Virtual Room," in *Proceedings of the 30th Tonmeistertagung VDT International Convention* (Hannover, Germany). pp 127–132 (2018 Nov.).

[3] F. Wendt, M. Frank, F. Zotter, and R. Höldrich, "Directivity Patterns Controlling the Auditory Source Distance," in *Proceedings of the 19th International Conference on Digital Audio Effects (DAFx)*, pp. 5–9 (Brno, Czech Republic) (2016 Sep.).

[4] J. Meyer, *Acoustics and the Performance of Music: Manual for Acousticians, Audio Engineers, Musicians, Architects and Musical Instrument Makers* (Springer, New York, NY, 2009).

[5] H. J. Vos, O. Warusfel, N. Misdariis, and D. de Vries, "Analysis and Reproduction of the Frequency Spectrum and Directivity of a Violin," *J. Acoust. Soc. Neth.*, vol. 167, pp. 1–11 (2003 Sep.).

[6] T. Grothe and M. Kob, "High Resolution 3D Radiation Measurements on the Bassoon," in *Proceedings of the International Symposium on Musical Acoustics*, pp. 139–145 (Detmold, Germany) (2019 Nov.).

[7] J. Pätynen and T. Lokki, "Directivities of Symphony Orchestra Instruments," *Acta Acust. united Acust.*, vol. 96, no. 1, pp. 138–167 (2010 Jan.).

[8] M. Pollow, G. Behler, and B. Masiero, "Measuring Directivities of Natural Sound Sources With a Spherical Microphone Array," in *Proceedings of the Ambisonics Symposium*, pp. 166–169 (Graz, Austria) (2009 Jun.).

[9] L. M. Wang and M. C. Vigeant, "Evaluations of Output From Room Acoustic Computer Modeling and Auralization due to Different Sound Source Directionalities," *Appl. Acoust.*, vol. 69, no. 12, pp. 1281–1293 (2008 Dec.).

[10] F. Otondo and J. H. Rindel, “The Influence of the Directivity of Musical Instruments in a Room,” *Acta Acust. united Acust.*, vol. 90, no. 6, pp. 1178–1184 (2004 Nov.).

[11] F. Hohl and F. Zotter, “Similarity of Musical Instrument Radiation-Patterns in Pitch and Partial,” in *Proceedings of the 36th German Annual Conference on Acoustics (DAGA)*, pp. 701–702 (Berlin, Germany) (2010 Jan.).

[12] C. Pörschmann and J. M. Arend, “Investigating Phoneme-Dependencies of Spherical Voice Directivity Patterns,” *J. Acoust. Soc. Am.*, vol. 149, no. 6, pp. 4553–4564 (2021 Jun.). https://doi.org/10.1121/10.0005401.

[13] D. Ackermann, C. Böhm, F. Brinkmann, and S. Weinzierl, “The Acoustical Effect of Musicians’ Movements During Musical Performances,” *Acta Acust. united Acust.*, vol. 105, no. 2, pp. 356–367 (2019 Mar.). https://doi.org/10.3813/AAA.919319

[14] J. Ehret, J. Stienen, C. Brozdowski, et al., “Evaluating the Influence of Phoneme-Dependent Dynamic Speaker Directivity of Embodied Conversational Agents’ Speech,” in *Proceedings of the 20th ACM International Conference on Intelligent Virtual Agents*, paper 17 (Online) (2020 Oct.). https://doi.org/10.1145/3383652.3423863.

[15] N. R. Shabtai, G. Behler, M. Vorländer, and S. Weinzierl, “Generation and Analysis of an Acoustic Radiation Pattern Database for Forty-One Musical Instruments,” *J. Acoust. Soc. Am.*, vol. 141, no. 2, pp. 1246–1256 (2017 Feb.). https://doi.org/10.1121/1.4976071.

[16] S. Weisser and M. Quanten, “Rethinking Musical Instrument Classification: Towards a Modular Approach to the Hornbostel-Sachs System,” *Yearb. Tradit. Music*, vol. 43, pp. 122–146 (2011).

[17] T. Magnusson, “Musical Organics: A Heterarchical Approach to Digital Organology,” *J. New Music Res.*, vol. 46, no. 3, pp. 286–303 (2017 Jul.).

[18] C. Anemüller and J. Herre, “Calculation of Directivity Patterns From Spherical Microphone Array Recordings,” presented at the *147th Convention of the Audio Engineering Society* (2019 Oct.), paper 10285.

[19] M. Pollow, *Directivity Patterns for Room Acoustical Measurements and Simulations*, vol. 22 (Logos Verlag Berlin GmbH, Berlin, Germany, 2015).

[20] J. G. Tylka, R. Sridhar, and E. Choueiri, “A Database of Loudspeaker Polar Radiation Measurements,” presented at the *139th Convention of the Audio Engineering Society* (2015 Oct.), e-Brief 230.

[21] M. Pezzoli, A. Canclini, F. Antonacci, and A. Sarti, “A Comparative Analysis of the Directional Sound Radiation of Historical Violins,” *J. Acoust. Soc. Am.*, vol. 152, no. 1, pp. 354–367 (2022 Jul.). https://doi.org/10.1121/10.0012577.

[22] J. Ahrens, “Database of Spherical Harmonic Representations of Sound Source Directivities,” *Zenodo* (2020). https://doi.org/10.5281/zenodo.3707708.

[23] J. Ahrens and S. Bilbao, “Computation of Spherical Harmonic Representations of Source Directivity Based on the Finite-Distance Signature,” *IEEE/ACM Trans. Audio Speech Lang. Process.*, vol. 29, pp. 83–92 (2020 Nov.). https://doi.org/10.1109/TASLP.2020.3037471.

[24] D. Schröder and M. Vorländer, “RAVEN: A Real-Time Framework for the Auralization of Interactive Virtual Environments,” in *Proceedings of the Forum Acusticum*, pp. 1541–1546 (Aalborg, Denmark) (2011 Jan.).

[25] L. Aspöck, F. Brinkmann, D. Ackermann, S. Weinzierl, and M. Vorländer, “Benchmark for Room Acoustical Simulation (BRAS),” *FG Audiokommunikation* (2019 May). https://doi.org/10.14279/depositonce-6726.3.

[26] F. Brinkmann, L. Aspöck, D. Ackermann, et al., “A Round Robin on Room Acoustical Simulation and Auralization,” *J. Acoust. Soc. Am.*, vol. 145, no. 4, pp. 2746–2760 (2019 Apr.). https://doi.org/10.1121/1.5096178.

[27] J. Pätynen, V. Pulkki, and T. Lokki, “Anechoic Recording System for Symphony Orchestra,” *Acta Acust. united Acust.*, vol. 94, no. 6, pp. 856–865 (2008 Nov.).

[28] M. Miron, J. J. Carabias Orti, J. Bosch, et al., “PHENICX-Anechoic: Note Annotations for Aalto Anechoic Orchestral Database,” *Zenodo* (2016 Nov.). https://doi.org/10.5281/zenodo.1289821.

[29] F. Wefers, “A Free, Open-Source Software Package for Directional Audio Data,” in *Proceedings of the 36th German Annual Conference on Acoustics (DAGA)*, pp. 1059–1060 (Berlin, Germany) (2010 Jan.).

[30] J. W. Kim, J. Salamon, P. Li, and J. P. Bello, “Crepe: A Convolutional Representation for Pitch Estimation,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 161–165 (Calgary, Canada) (2018 Apr.). https://doi.org/10.1109/ICASSP.2018.8461329.

[31] F. Brinkmann, A. Lindau, S. Weinzierl, et al., “A High Resolution and Full-Spherical Head-Related Transfer Function Database for Different Head-Above-Torso Orientations,” *J. Audio Eng. Soc.*, vol. 65, no. 10, pp. 841–848 (2017 Oct.). https://doi.org/10.17743/jaes.2017.0033.

[32] F. Brinkmann, A. Lindau, S. Weinzierl, et al., “The FABIAN Head-Related Transfer Function Database,” TU Berlin (2017 Feb.). https://doi.org/10.14279/depositonce-5718.5.

[33] L. Leventhal, “Type 1 and Type 2 Errors in the Statistical Analysis of Listening Tests,” *J. Audio Eng. Soc.*, vol. 34, no. 6, pp. 437–453 (1986 Jun.).

[34] A. Corcuera Marruffo and V. Chatziioannou, “A Pilot Study on Tone-Dependent Directivity Patterns of Musical Instruments,” in *Proceedings of the AES International Audio for Virtual and Augmented Reality Conference* (2022 Aug.), paper 22.

[35] H. Steffens, S. van de Par, and S. D. Ewert, “The Role of Early and Late Reflections on Perception of Source Orientation,” *J. Acoust. Soc. Am.*, vol. 149, no. 4, pp. 2255–2269 (2021 Apr.). https://doi.org/10.1121/10.0003823.

THE AUTHORS


Andrea Corcuera


Vasileios Chatziioannou


Jens Ahrens

Andrea Corcuera is a doctoral candidate at the Department of Music Acoustics of the University of Music and Performing Arts in Vienna, Austria. She received her B.Sc. degree in Audiovisual Systems Engineering from Pompeu Fabra University in Barcelona, Spain, and her M.Sc. degree in Sound and Music Computing from Aalborg University in Copenhagen, Denmark. She started her Ph.D. after working in the hearing aid industry at Oticon A/S in Denmark. Her main interests include spatial audio and psychoacoustics.

●

Vasileios Chatziioannou received his Ph.D. from Queen's University Belfast in the field of electronics, electrical engineering, and computer science. Subsequently, he moved to the Department of Music Acoustics in Vienna, where he is teaching and conducting research in the acoustics of musical instruments. He has led two projects on physical modelling of musical instruments, funded by the Austrian Science Fund (FWF), and the ITN project Virtual Reality Audio for Cyber Environments (VRACE), funded by the Horizon 2020 Framework Programme of the European Union. He is an associate editor of the *Journal of the Acoustical Society of America* and the chair of the EAA Technical Committee on Musical Acoustics

●

Jens Ahrens has been an Associate Professor within the Division of Applied Acoustics at Chalmers since 2016. He received a Diplom (equivalent to an M.Sc.) in Electrical Engineering/Sound Engineering jointly from Graz University of Technology and the University of Music and Dramatic Arts, Graz, Austria, in 2005. He completed his Doctoral Degree (Dr.-Ing.) at the Technische Universität Berlin, Germany, in 2010. From 2011 to 2013, he was a Postdoctoral Researcher at Microsoft Research in Redmond, Washington, and in the fall and winter terms of 2015/2016, he was a Visiting Scholar at the Center for Computer Research in Music and Acoustics (CCRMA) at Stanford University, California. He is an Associate Editor of the *IEEE Signal Processing Letters* and EURASIP *Journal on Audio, Speech, and Music Processing*.