PERCEPTUAL RELEVANCE OF ASYNCHRONY BETWEEN ORCHESTRAL INSTRUMENT GROUPS IN TWO CONCERT HALLS

Henna Tahvanainen, Aki Haapaniemi, Jukka Pätynen, Tapio Lokki

Department of Computer Science
Aalto University School of Science, Finland
`henna.tahvanainen@aalto.fi`

ABSTRACT

Timing in a music ensemble performance is asynchronous by nature. Asynchrony is generated by the players themselves, and further delays to listeners are introduced by the location and orientation of the instruments on stage. While the musicians aim to an accurate mutual synchronization, deviating from the perfect synchrony may even produce desirable effects. For one, the timbre can appear broader as with orchestra string sections. The perceived asynchrony within an ensemble varies between 20 to 50 ms.

This paper studies the perceptual relevance of asynchrony between three orchestral instrument groups in two concert halls. Perfect synchrony was compared to 1) the bass-register instruments (double basses and timpani) played first with delays of 20 ms for middle-register instruments (cellos, bassoon), and 40 ms for treble-register instruments (winds, brass, violas, violins), and 2) the treble-register instruments played first with delays of 20 ms for middle-register, and 40 ms for bass-register instruments.

Listener preference was investigated with a paired comparison online listening test using binaural renderings of the concert halls over headphones. The results were analysed with a probabilistic choice model with latent preference groups. The analysis shows that listener preference generally depends on the asynchrony: the bass-register instrument starting first is the most preferred option in both halls while the treble-register starting first is the least preferred. The results also imply that preference on timing depends on the concert hall, and this requires future listening tests with a spatial audio system in order to reproduce the spatial characteristics of the concert halls more accurately.

1. INTRODUCTION

Perfect synchrony is practically never obtained across all players of a music ensemble. Musicians aim to synchronise their playing by their internal clock, as well as acoustic and visual cues, such as body movements of other players [1, 2]. In many cases, these cues are obstructed. For example, the maximum distance between two members of a symphony orchestra on stage translates to a propagation delay of about 30 to 50 ms which hampers the synchronisation by hearing [1]. In addition, different string sections are typically on different sides of the stage, which makes visual cues from other players difficult to follow.

Consequently, asynchrony in ensemble performance is mostly unintended. A typical value of asynchrony between players is 20 to 50 ms [3, 4, 5, 6, 7, 8, 9]. For example, the asynchrony of the bowing onsets within a string ensemble members settles somewhere between ±30 ms respective to nominal synchrony [9]. Asynchrony of a few milliseconds can be heard, but determining the temporal order of two sounds requires a time difference of about 15 to 20 ms [10]. Furthermore, asynchrony is more difficult to detect when the lead sound is louder than the lagging sound, and this is thought to be caused by forward masking [11, 12].

Asynchrony can also be intentional, and it is often employed in ensembles for expressiveness. For instance, in string, wind, and piano duets, the melody line is typically played about 20 ms earlier in order to emphasize the melody [4, 5, 7, 13]. In jazz ensembles, the drummer often advances the melody instruments at every other beat, and this makes the music swing [6]. Informal discussions with several conductors have revealed that they occasionally utilize intentional asynchrony within the orchestra in order to achieve a desired musical expressiveness, such as the enhancement of bass.[1]

Playing in asynchrony may have benefial effects. For example, asynchrony is applied in orchestra auralisation to create instrument sections that sound natural, for example to obtain a wider timbre for string sections [14, 15]. Furthermore, when the orchestra plays in synchrony the frequencies of higher harmonics of the bass-register instruments and the fundamentals of the treble-register instruments coincide [16, 17, 18], and this may lead to undesirable simultaneous masking effects [12]. In this case, the perceived loudness of the bass-register instruments lies solely on their low frequency components which may not radiate sufficiently as the wavelengths at the low frequencies are much greater than the dimensions of the instruments [16, 19]. Thus, playing in asynchrony may help in making bass-register instruments more audible.

This study explores the perceptual relevance of intentional asynchrony within a symphony orchestra via binaural renderings in online listening test. An anechoic loudspeaker orchestra is auralised in two concert halls. The musical instruments are divided into three groups according to their register, and artificial asynchrony is introduced between the groups. The results indicate that listeners prefer asynchrony to synchrony and the preferred asynchrony is such that the bass-register instruments start to play before mid- and treble-register instruments. It seems this preference is hall-dependent, but further testing with a larger selection of concert halls is required using a spatial sound system.

2. METHODS

2.1. Participants

Thirty nine subjects (5 female) performed the online listening test (mean age=32, SD=9). 61 % reported to have taken part

---

[1] See for example the video "'Die Kunst des Dirigierens"' with Herbert von Karajan and Vienna Philharmonic Orchestra between 1:00-1:37 – `www.youtube.com/watch?v=Shc-4AZVaNk`

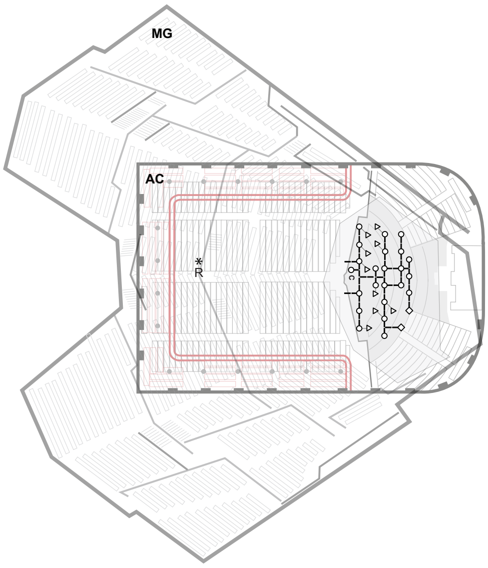

Figure 1: Plans of the concert halls superimposed with the loudspeaker orchestra position. The measurement position is marked with R and it is located at 19 m from the stage edge.

Table 1: The properties of the measured unoccupied concert halls. Reverberation time RT, and strength G are averaged over 63-8000 Hz octave bands. Clarity C_{80} is averaged over 500-1000 Hz octave bands. SDE stands for the frequency of maximum attenuation of the seat-dip effect. Note that the level of the frequency response in AC has been reduced by 2 dB.

Abbr.	**Name**	**N**	**RT** [s]	**G** [dB]	**SDE** [Hz]	**C$_{80}$** [dB]
AC	Amsterdam Concertgebouw	2040	2.1	-1.4	125	-2.7
MG	Munich Gasteig	2400	2.0	-1.5	99	-0.1

in listening tests before, and 82 % reported they play a musical instrument. 28 % reported they prefer listening to classical music 5 or more points on a scale of 0 to 10. Because of the online listening test, the participants could not be screened with audiometry and were thus assumed to have normal hearing. The participants used their own headphones that ranged from cheap in-ear models to professional open-ear models.

2.2. Stimuli

The stimuli for the experiment consisted of an anechoic symphony orchestra auralised in two concert halls: Amsterdam Concertgebouw (shoebox), and Munich Gasteig (fan). Their hall plans are shown in Fig. 1 and some objective parameters are listed in Tab. 1.

These concert halls were chosen because they have very similar spectral responses, when the level of the overall response in AC is reduced by 2 dB. The left-hand side of Fig. 2 shows the 20-ms and the full frequency response in both halls. The middle and the right-hand side show the time-frequency development of the frequency response at 10-ms time-window increments in AC and MG, respectively.

The symphony orchestra comprised a 24-channel loudspeaker orchestra (LSO) on stage [20]. The LSO positions on stage are shown in Fig. 3 with the corresponding channel numbers. The content of the channels is listed in Tab. 2. The orchestra played the opening chord of Beethoven's 7th symphony (see Fig. 4 for score).

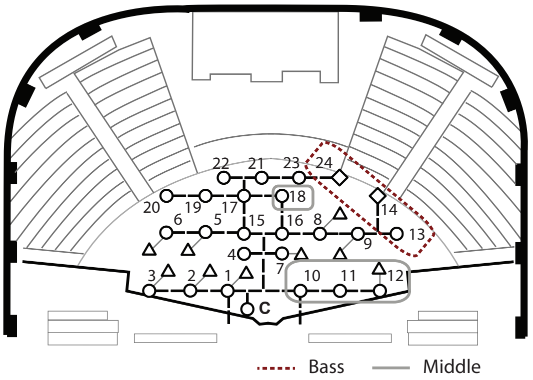

Figure 3: The position of the loudspeaker orchestra on stage with the channel numbers. The bass-register instrument (double bass and timpani) and the middle-register instruments (cellos, bassoon) are shown for this study. All the other loudspeakers contain the treble-register instruments (violas, violins, clarinets, horns, oboe, flute). The content of each channel is shown in Table 2.

Table 2: The instrument registers and their corresponding loudspeaker orchestra channel numbers.

Register	**LSO Channel**	**Instruments**
Bass	13	double bass
	14	double bass
	24	timpani
Middle	10	cello
	11	cello
	12	cello
	18	bassoon
Treble	1-6	violin
	7-8	viola
	9	viola, cello
	15	violin, flute
	16	viola
	17	clarinet
	19-20	trumpet
	21-22	trombone
	23	empty

In brief, the stimuli were created in the following manner: the room impulse responses (RIRs) from each loudspeaker (24 channels) on the concert hall stage were measured using the logarithmic sine sweep technique with a G.R.A.S. probe at the measurement position R located 19 m from the stage edge. The RIRs were then analysed with the Spatial Decomposition Method that estimates the direction of arrival of the sound field at discrete time samples [21]. The directional estimates were then allocated to a 3D spatial sound system with 24 channels, with directions matching the nearest loudspeaker in the set-up. The binaural RIRs were obtained by filtering the loudspeaker set-up with head-related transfer functions (HRTFs) of subject no. 40 in the CIPIC database [22]. Then, the binaural RIRs were convolved with anechoic musical instrument recordings.

Because the CIPIC HRTFs are obtained in anechoic con-

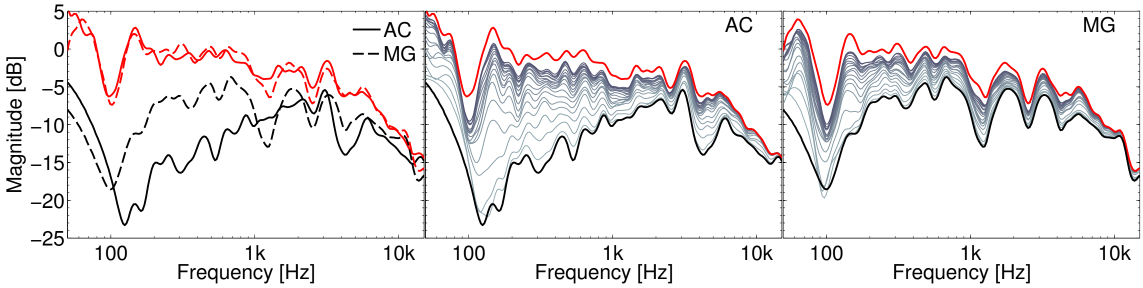

Figure 2: The time-frequency development of the frequency response of the concert halls. The top curve shows the frequency responses of the whole impulse response. The thin curves show the frequency responses at 10-ms increments with the lowest cure in bold starting at 20 ms after the direct sound. The second highest curve in bold shows the frequency response at 200 ms after the direct sound. Note that the level of the frequency response in AC has been reduced by 2 dB.

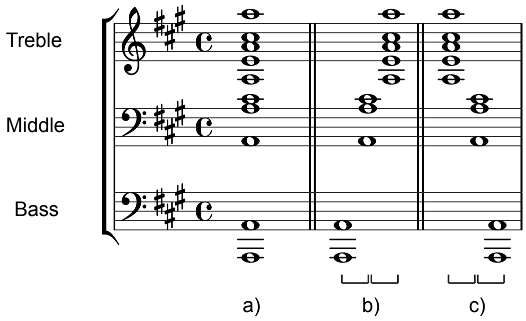

Figure 4: Score for a) perfect synchrony SYNC, b) bass-register first BMT, and c) treble-register first TMB.

ditions, the HRTFs were preprocessed with an equalisation response to better resemble listening in a normal room. The equalisation response was obtained by 1/3-octave smoothing the average of eight HRTFs on the lateral plane at 45 degree intervals, and raising the smoothed average to the power of $\alpha = 0.6$ (for more details, see [23]). Several listeners have informally evaluated the equalisation to produce a plausible binaural reproduction using CIPIC responses.

Furthermore, due to the positioning of the loudspeakers on the stage and the corresponding path length differences to the measurement position, the initial delays of the RIRs vary between the LSO channels. While this corresponds to the situation with an actual orchestra, it was undesired for the present study where exact control over the timing between channels was required. Therefore, the LSO channel IR’s were synchronized by the position of their initial peaks. Similarly, the relative timing of the anechoic orchestra channels were adjusted to have synchronous physical onsets of notes. A perfect synchrony between the instruments was obtained by finding an energy threshold for both the instrument and room impulse onsets.

The convolved audio was divided into three instrument groups based on note heights/instrument registers: bass, middle, and treble, see Table 2. Various delays were applied for the three instrument groups by shifting the convolved tracks in time and summing the result. Three timings were used with delays of 0 ms, 20 ms, and 40 ms across groups. The 20-ms time difference between consecutive instrument groups was chosen based on earlier research on existing asynchrony within ensembles [4, 5], as well as on the threshold for the time difference of two events whose temporal order can be judged [10].

The first timing represents perfect onset synchrony and it is referred to as SYNC. The second timing is a case where the bass-register (B) instruments start earlier and are followed by middle-register (M) instruments after 20 ms and treble-register (T) instruments after 40 ms, and it is referred to as BMT. The third case represents the opposite case where treble-register instruments start earlier with the same relative delays (TMB). The timings are also shown in Fig. 4.

2.3. Design

The experiment was a paired comparison (two-alternative choice) test with ties allowed. The listeners were presented with 20 pairs of stimuli on a web page. They were asked to listen to each pair and to choose which of the two samples they prefer. It was also possible to choose a tie with ”‘Cannot say”’. The samples could be listened to as many times as needed, and all pairs were available simultaneously. Furthermore, the listeners were asked to listen with headphones in a quiet place.

The test consisted of pair combinations of six stimuli (2 halls × 3 timings = 15 pairs) in a random order. In addition, three pairs with identical samples were randomly added to the sequence in order to monitor the use of ties, and whether the listeners were listening accurately. Two additional pairs were included in the beginning of the test as the training pairs and they were excluded from the analysis.

The online listening test was implemented with HTML5 and JavaScript. Because 16-bit wav-files (sampling rate 48 kHz) and HTML5 <audio>-tags were used for playback, some browsers (e.g., Internet Explorer, Safari) could not be used in the listening test.

The listener was also asked at the end of the listening to comment freely on the basis for their preference. At the end of the web page, the listener was also asked to provide some background information including headphone type, experience in classical music, listening tests, and playing musical instruments.

2.4. Analysis methods

The results of the paired comparison experiment were analysed with a probabilistic choice model developed by Bradley, Terry [24] and Luce [25], that is based on a logistic distribution of response differences. The BTL-model describes the probability of an item to be preferred over the others with a specified attribute. The previous auditory perception studies include various attributes, such as unpleasantness [26, 27]; eventfulness [27]; spaciousness, brightness, naturalness, distance [28]; and level of bass and articulation [29]. In this case, the dimension is *preference*.

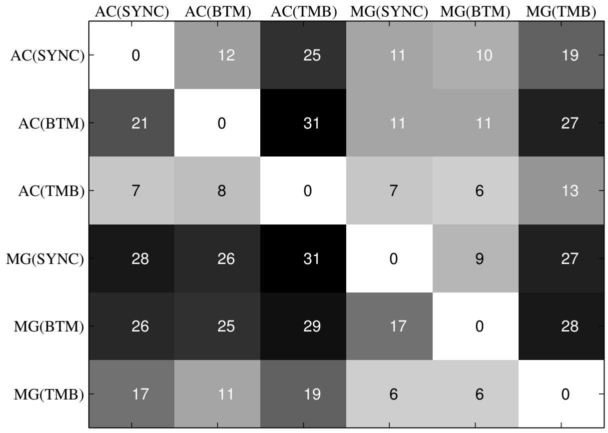

Figure 5: The preference matrix. The number indicates the absolute frequency with which the sample in the row was preferred over the sample in the column. A darker colour indicates a higher frequency. Since each participant was presented the pair once, the maximum frequency is 39.

Allowing ties in the paired comparison experiment makes the analysis of the results more complicated [30, 31, 32]. While a number of packages (for an example for R and Matlab see [33]) are available for the analysis of paired comparison experiments, most of them consider the two-alternative forced choice procedure only, and prohibit ties in the data. In the simplest treatment, the ties can be neglected, split, or apportioned either randomly, or according to some criterion [34]. Rao and Kupper [30] proposed that when the difference between two responses is below a certain threshold, a tie is declared. Davidson [31] proposed that the probability of a tie is proportional to the geometric mean of the preference probabilities. For the data in this study the probabilities do not differ significantly between the Davidson model for ties (implemented in the psychotools-package in R) and discarding the ties. Thus, the ties have been discarded in the following BTL-models.

Moreover, untrained listening test subjects may not necessary form a homogeneous group, and thus the data may include latent preference groups [35, 34]. Consequently, the data was segmented using a latent class model developed by Courcoux et al [35]. This approach includes a Monte Carlo significance test procedure in order to determine the adequate number of classes or groups. Here, 100 Monte Carlo random samples with size of the number of subjects (39) were drawn for models with an increasing number of latent groups. The segmentation that best fits the data can be found with a likelihood ratio statistic between the groups. The BTL-model can then be applied to each latent group separately.

3. RESULTS

Combining the individual data of all the participants yields a preference matrix in Fig. 5 without the ties. It shows the absolute frequencies of the preferred samples. The bottom-left quadrant has the highest frequencies, showing that MG is generally preferred over AC.

The portion of ties was about 10% of all the answers and they are plotted in Fig. 6. The most ties occur in MG between SYNC and BMT (about 33 %), while between other pairs the percentage of ties is about 18% or less.

The top plot of Figure 7 shows the results of the BTL-model of all the answers. The BTL-model shows that MG appears to be preferred with all timings. By implementating the segmentation, a model with three latent groups was found to best fit the data. The bottom plot of Figure 7 shows the BTL-model for each of the three latent groups. The first two groups with 6 and 19 subjects, respectively, prefer MG over AC in general. The third group consists of 14 subjects that somewhat prefer AC to MG. As for the preference of timing, the first group prefers BMT timing in MG. In the second group the only significant result is the SYNC and BMT are clearly most preferred in MG. In the third group, none of the timing adjustments is clearly preferred.

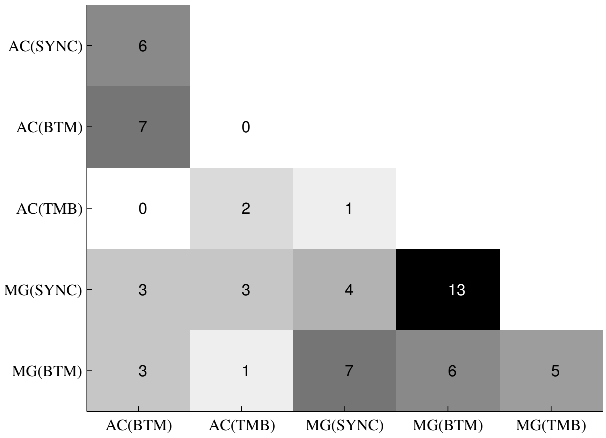

Figure 6: Number of ties for each comparison pair. A darker colour indicates a higher frequency. Note that matrix is triangular, because the presentation order of the pairs has been ignored.

Based on the top plot of Fig. 7, it appears that the differences between hall acoustics were a more dominating feature than synchrony for the preference. Hence, the preference of timing adjustment in each hall is not clearly observed in the BTL-model. Therefore, the between-concert-hall pairs were excluded and the effect of timing was analysed in each hall separately. In a similar manner to all the data, this subset of data can be segmented. In AC, the 39 subjects are best-fitted to a two-group model, and in MG to a single-group model. The BTL-model for these groups is shown in Fig. 8. In AC, eight listeners prefer the TMB timing, while the other 31 listeners prefer the BMT. In MG, the BMT timing is most likely preferred, but is not significantly different from the preference of SYNC.

Neither age, sex, preference of classical music, nor experience in listening tests or playing a musical instrument were significantly different between the obtained groups.

4. DISCUSSION

The results indicate that synchrony between symphony orchestra instrument groups is relevant for preference. In general, the case when the treble-register instruments started to play (TMB) first, was the least preferred. One of the explanations could be that the percussive timpani is played later than most instruments which gives a sensation of being ”off-beat”. However, there was a group of eight people in AC that actually preferred the treble-register instruments starting first. It is possible that these listeners prefer a soft attack of the timpani, as some of them reported. The soft attack is likely caused by forward-masking of the other instruments.

The results further indicate the preferred timing, or asynchrony, may differ across concert halls. In the two halls studied, the bass-register instruments playing first (BMT) was gen-

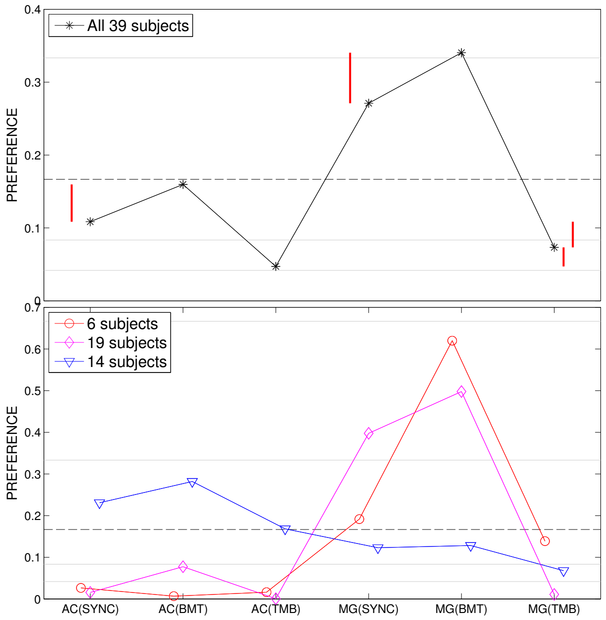

Figure 7: The BTL-model for the probability of preference over all stimuli before and after grouping. The results are displayed on a ratio scale, where the sum of ratios is normalised to unity. The black dashed line indicates the probability of chance (1/6). The 95% confidence intervals are not shown in order to retain clarity. The non-significant differences between the values are indicated with vertical lines on top plot. For the bottom plot, the thin lines indicate the doubling of the probability (on a logarithmic scale). If two values fall within the same area formed by two lines, generally no significant difference exists between the values.

erally preferred but in MG the difference to perfect synchrony (SYNC) is insignificant. The high amount of ties in MG between BMT and SYNC shows that adjusting the timing may not be as important in MG as in AC, as long as it is natural (not ”off-beat”).

The preference differences between the concert halls may be at least partially explained by the seat-dip effect. Namely, the time-frequency responses of the two halls differ in the temporal increase of the level at both low and mid-high frequencies from about 100 Hz until 1 kHz (see Fig. 2), and this frequency range is mostly affected by the seat-dip effect [36]. The seat-dip effect refers to the attenuation that the direct sound arriving at grazing angles to the seats undergoes at low frequencies. This leads to a more audible lack of bass in some concert halls than in others [37, 38, 29]. The effect is observed as a dip in the frequency response by 20 ms after the direct sound, and it is caused by a destructive interference between the direct sound and reflections from the tops of the seat backs. Depending on the further reflections in the concert hall that do not arrive at the listener at grazing angles, the dip can diminish. Previous research shows that the perception of bass can be enhanced if direct sound lack the low frequencies, but early frequencies retain it [39].

In MG, the seat-dip effect at 20 ms is a narrow band dip centred around 100 Hz. While the magnitude of frequency response increases steadily with time, the response retains its shape, including the dip. On the contrary, in AC the dip is at a slightly higher frequency and it covers a wide frequency range. What is more, the frequency response in AC becomes flatter with time and the dip becomes smaller. Consequently, in AC the main boost of the low and mid-high frequencies is delayed. This means that when the orchestra plays simultaneously, the low frequencies are initially hampered by the seat-dip effect. If the bass-register instruments anticipate a little, their low and mid-high frequencies will arrive at the listener closer in time with the treble instruments.

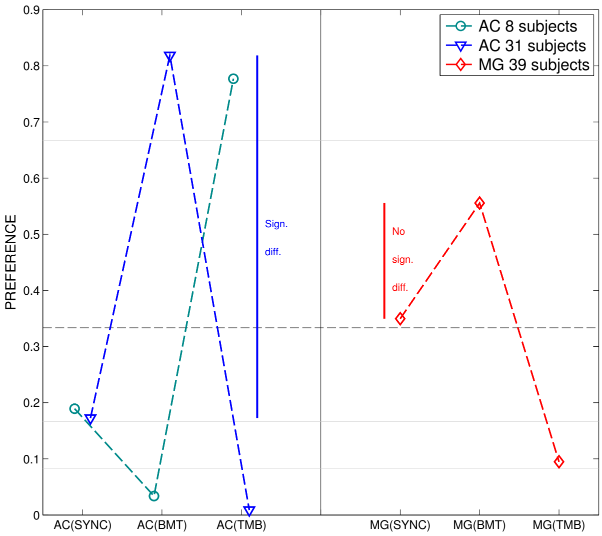

Figure 8: The BTL-model for the probability of preference over timing for each hall separately. The significances of the value differences are indicated with vertical lines.

In other words, it is more difficult to play in AC than in MG, since timing of the instruments is crucial. The bass-register instruments should play first in AC because of the strong time-frequency development of the frequency response. And when asynchrony is applied to playing chords in AC, the effect is much more salient than in MG where the time-frequency development of the frequency response mostly consists of increasing level.

An alternative explanation is that AC might benefit from the bass-register instruments starting first due to some masking effects. Because the mid-high frequency response is delayed, the higher partials of the bass-register instruments are less masked by the treble-register instruments, when the bass-register instruments start to play earlier.

The overall preference of MG over AC is not in line with the other published results on concert hall preference [40]. One of the reasons could be the chosen stimuli. The stimuli consisted of onlyone chord which makes it susceptible to small variations. For example, in MG the timpani can be heard sharper than in AC, which is likely to increase overall preference towards MG.

Another reason for the inconsistent preference results could be the reproduction method. The correct spatiotemporal development in the concert hall auralisations is essential and its quality may deteriorate from the lack of both individualised HRTFs and head tracking, as well as the lack of headphone equalisation. Finally, the listeners participated in the test in an uncontrolled environment. Thus, there may have been many kinds of distractions. Consequently, a listening test in a controlled environment with a more elaborate spatial sound reproduction is required to confirm the results of the internet experiment. Further listening tests could also focus on the possible enhancement of bass with asynchronous playing.

5. CONCLUSIONS

Intentional asynchrony within three groups of symphony orchestra instruments were studied via binaural renderings of concert halls. The results show that in a shoebox-shaped hall AC, it is preferable to play chords asynchronously so that the double basses and timpani start first, and are then followed by the cello and bassoon, and finally by violins, violas, woodwinds and brass. In a fan-shaped hall MG, there is no clear preference towards sychrony or asynchrony as long as the low frequencies are not too late. However, in neither of the halls the treble-register instruments should start first. The preference differences can at least partially be account for by the seat-dip effect and the subsequent time-frequency development of the frequency response.

6. ACKNOWLEDGMENTS

The research leading to these results has received funding from the Academy of Finland, project no. [257099].

7. REFERENCES

[1] G. M. Naylor, "The achievement of ensemble," *Applied Acoustics*, vol. 23, no. 2, pp. 109–120, 1988.

[2] P. E. Keller and M. Appel, "Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles," *Music Perception*, vol. 28, no. 1, pp. 27–46, 2010.

[3] D. Deutsch, "Hearing music in ensembles," *Physics Today*, vol. 63, no. 2, pp. 40–45, 2010.

[4] R. A. Rasch, "Synchronization in performed ensemble music," *Acustica*, vol. 43, no. 4, pp. 122–131, 1979.

[5] L. H. Shaffer, "Timing in solo and duet piano performances," *The Quarterly Journal of Experimental Psychology*, vol. 36, no. 4, pp. 577–595, 1984.

[6] A. Friberg and A. Sundström, "Swing ratios and ensemble timing in jazz performance: Evidence for a common rhythmic pattern," *Music Perception*, vol. 19, no. 3, pp. 333–349, 2002.

[7] W. Goebl and C. Palmer, "Synchronization of timing and motion among performing musicians," *Music Perception*, vol. 26, no. 5, pp. 427–438, 2009.

[8] G. P. Moore and J. Chen, "Timings and interactions of skilled musicians," *Biological Cybernetics*, vol. 103, no. 5, pp. 401–414, 2010.

[9] J. Pätynen, S. Tervo, and T. Lokki, "Temporal differences in string bowing of symphony orchestra players," *Journal of New Music Research*, vol. 41, no. 3, pp. 223–237, 2012.

[10] I. J. Hirsh, "Auditory perception of temporal order," *The Journal of the Acoustical Society of America*, vol. 31, no. 6, pp. 759–767, 1959.

[11] W. Goebl and R. Parncutt, "The influence of relative intensity on the perception of onset asynchronies," in *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney, Australia*, 2002.

[12] B. Moore, *An introduction to the psychology of hearing*, Brill, 2012.

[13] W. Goebl and R. Parncutt, "Asynchrony versus intensity as cues for melody perception in chords and real music," in *Proceedings of the 5th Triennial ESCOM Conference, Hanover, Germany*, 2003.

[14] M. C. Vigeant, L. M. Wang, and J. Holger Rindel, "Investigations of orchestra auralizations using the multichannel multi-source auralization technique," *Acta Acustica united with Acustica*, vol. 94, no. 6, pp. 866–882, 2008.

[15] J. Pätynen, S. Tervo, and T. Lokki, "Simulation of the violin section sound based on the analysis of orchestra performance," in *Proceedings of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2011)*, New Paltz, NY, USA, 2011, pp. 173–176.

[16] J. Meyer and U. Hansen, *Acoustics and the Performance of Music*, Springer, New York, NY, USA, 2009.

[17] N. Nishihara and T. Hidaka, "Loudness perception of low tones undergoing partial masking by higher tones in orchestral music in concert halls," *Journal of the Acoustical Society of America*, vol. 132, no. 2, pp. 799–803, 2012.

[18] L. J. Trainor, C. Marie, I. C. Bruce, and G. M. Bidelman, "Explaining the high voice superiority effect in polyphonic music: Evidence from cortical evoked potentials and peripheral auditory models," *Hearing Research*, vol. 308, pp. 60–70, 2014.

[19] A. Askenfelt, "Eigenmodes and tone quality of the double bass," in *Dept for Speech, Music, and Hearing – Quarterly Progress and Status Report*, Roayl Institute of Technology, Stockholm, Sweden, 1982, vol. 23, pp. 149–174.

[20] J. Pätynen, *A Virtual Loudspeaker Orchestra for Studies on Concert Hall Acoustics*, PhD dissertation, Aalto University School of Science, Espoo, Finland, 2011.

[21] S. Tervo, J. Pätynen, and T. Lokki, "Spatial decomposition method for room impulse responses," *Journal of Audio Engineering Society*, vol. 61, no. 1, pp. 1–13, 2013.

[22] V. R. Algazi, R. O. Duda, D. M. Thompson, and C. Avendano, "The CIPIC HRTF database," in *IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics (WASPAA2001)* , New York, NY, USA, 2001, pp. 99–102.

[23] J. Pätynen, S. Tervo, and T. Lokki, "Amplitude panning decreases spectral brightness with concert hall auralizations," in *Proceedings of the 55th International Conference of the Audio Engineering Society: Spatial Audio*, Helsinki, Finland, 2014.

[24] R. A. Bradley and M. E. Terry, "Rank analysis of incomplete block designs: I. The method of paired comparisons," *Biometrika*, pp. 324–345, 1952.

[25] R.D. Luce, *Individual Choice Behaviour. A Theoretical Analysis.*, Wiley, New York, 1959.

[26] W. Ellermeier, M. Mader, and P. Daniel, "Scaling the unpleasantness of sounds according to the btl model: Ratio-scale representation and psychoacoustical analysis," *Acta Acustica united with Acustica*, vol. 90, no. 1, pp. 101–107, 2004.

[27] B. De Coensel, S. Vanwetswinkel, and D. Botteldooren, "Effects of natural sounds on the perception of road traffic noise," *The Journal of the Acoustical Society of America*, vol. 129, no. 4, pp. EL148–EL153, 2011.

[28] S. Choisel and F. Wickelmaier, "Evaluation of multichannel reproduced sound: Scaling auditory attributes underlying listener preference," *The Journal of the Acoustical Society of America*, vol. 121, no. 1, pp. 388–400, 2007.

[29] H. Tahvanainen, J. Pätynen, and T. Lokki, "Studies on the perception of bass in four concert halls," *Psychomusicology: Music, Mind and Brain*, 2015.

[30] P. V. Rao and L .L. Kupper, "Ties in paired-comparison experiments: A generalization of the Bradley-Terry model," *Journal of the American Statistical Association*, vol. 62, no. 317, pp. 194–204, 1967.

[31] R. R. Davidson, "On extending the Bradley-Terry model to accommodate ties in paired comparison experiments," *Journal of the American Statistical Association*, vol. 65, no. 329, pp. 317–328, 1970.

[32] D. E. Critchlow and M. A. Fligner, "Paired comparison, triple comparison, and ranking experiments as generalized linear models, and their implementation on GLIM," *Psychometrika*, vol. 56, no. 3, pp. 517–533, 1991.

[33] F. Wickelmaier and C. Schmid, "A Matlab function to estimate choice model parameters from paired-comparison data," *Behaviour Research Methods, Instruments, and Computers*, vol. 36, no. 1, pp. 29–40, 2004.

[34] H. T. Lawless and H. Heymann, *Sensory Evaluation of Food: Principles and Practices*, Springer Science & Business Media, 2010.

[35] Ph. Courcoux and M. Semenou, "Preference data analysis using a paired comparison model," *Food Quality and Preference*, vol. 8, no. 5, pp. 353–358, 1997.

[36] H. Tahvanainen, J. Pätynen, and T. Lokki, "Analysis of the seat-dip effect in twelve European concert halls," *Acta Acustica united with Acustica*, vol. 101, no. 4, pp. 731–742, 2015.

[37] T. Schultz and B. Watters, "Propagation of sound across audience seating," *Journal of the Acoustical Society of America*, vol. 36, no. 5, pp. 885–896, 1964.

[38] G. Sessler and J. West, "Sound transmission over theatre seats," *Journal of the Acoustical Society of America*, vol. 36, no. 9, pp. 1725–1732, 1964.

[39] A. Walther, P. Robinson, and O. Santala, "Effect of spectral overlap on the echo threshold for single reflection conditions," *Journal of the Acoustical Society of America*, vol. 134, no. 2, pp. EL158–EL164, 2013.

[40] L. L. Beranek, *Concert Halls and Opera Houses: Music, Acoustics, and Architecture*, Springer- Verlag, New York, USA, 2nd edition, 2004.