Audio Engineering Society
Convention Paper
Presented at the 121st Convention
2006 October 5–8 San Francisco, CA, USA

Ensemble hand-clapping experiments under the influence of delay and various acoustic environments

Snorre Farner[1,2], Audun Solvang[1], Asbjørn Sæbø[1], and Peter Svensson[1]

[1] *Centre for Quantifiable Quality of Service in Communication Systems*, NTNU, O.S. Bragstads plass 2E, NO-7491 Trondheim, Norway*

[2] *Presently at IRCAM, Analysis/Synthesis Group, 1 place Igor Stravinsky, FR-75004 Paris, France*

Correspondence should be addressed to Peter Svensson (svensson@q2s.ntnu.no)

ABSTRACT

Hand-clapping experiments were performed by pairs of subjects under the influence of a delay up to 68 ms in various acoustic environments. The mean tempo decreased close to linearly as function of the delay. During each sequence the tempo slowed down to a degree that increased with the delay but for delays shorter than about 15–23 ms, the tempo increased during the sequence. For the timing imprecision, and for the subjects' judgements of their own ensemble performance, no effect of the delay could be observed up to 20 ms. Above 32 ms the effects were observed to increase with the delay. Virtual anechoic conditions lead to a higher imprecision than the reverberant conditions, and real-reverberation conditions lead to a slightly lower tempo.

1. INTRODUCTION

With available network capacity increasing, real-time sound and video transfer over the Internet is getting common. Video conferences are now possible between low-end networked computers, and musicians envisage to play together over the Internet.

However, for musical ensemble playing, the audio latency between the players is usually of crucial importance. Even for non-networked performance, the time needed for propagation of sound in air (about three milliseconds per meter) makes a delay between musicians virtually inevitable. In a small ensemble, players are typically separated by a few meters, while for a large orchestra the separation between

* *Centre for Quantifiable Quality of Service in Communication Systems, Centre of Excellence* appointed by The Research Council of Norway, funded by the Research Council, NTNU, and UNINETT. http://www.ntnu.no/Q2S/

the musicians farthest apart may be more than 20 meters, corresponding to an inter-musician delay of more than 60 ms. In the latter case, synchronism between the musicians is typically aided by the use of a conductor. Audio latency due to physical distance is a well-known challenge for organists accompanying other musicians in a church.

For geographically distributed musicians playing together over the Internet, the network connection imposes additional latencies. The speed of light gives a lower limit for the signal transport delay, with a value of 67 milliseconds for transmission around half the globe (20 000 km, theoretically the longest necessary distance on Earth). In practice, the delay will be greater. Network routes typically do not follow straight lines between network endpoints, and time is lost in processing in routers along the network line. Times is also spent in converting the signal from analogue to digital and back, in any applied compression and decompression of the signal, and in buffering for countering network transport time jitter. Experiments with geographically distributed ensemble playing have been conducted [1, 2]. It is interesting to note that software for jam sessions over the Internet is already available [3, 4].

Several studies have tried to answer questions related to how inter-musician delay affects ensemble playing. Chew et al. [5] have conducted an interesting experiment with a professional piano duo with years of experience of playing together. The two players were placed in the same rooms with visual contact but hearing the other on headphones with a latency up to 150 ms, and were asked to play different movements of a music piece from their repertoire. They had visual contact, though. After each repetition, they judged the difficulty of playing together with the applied latency unknown to them. 50 ms was found to be the upper limit for acceptable delay for a rapid movement while it increased to 75 ms for a slow one. When the sound of their own play was delayed as well, the tolerance increased to 65 ms for the rapid movement as they were more comfortable when the audible result was rhythmically correct.

Somewhat earlier, Chafe and Gurevich [6] devised an experiment involving two subjects in separate rooms connected by an audio line with a variable delay from 0 to 77 ms. Their task was to perform rhythmic hand clapping together (with a pair of simple complementary rhythms) while 'keeping the rhythm going evenly'. Randomly, one was chosen to start (supported by a metronome signal) and the other joined in after a few measures. The tempo was found to decrease rapidly for delays above about 11.5 ms. Interestingly, however, at lower delays, the tempo increased. Three possible reasons were suggested: the sound from the other clapper arriving earlier than expected from habitual delays in normal interhuman communication, an intrinsic tendency to speed up to compensate for normal delays, or a destabilization because the natural reverberation was eliminated from the acoustical setting. On their Internet site, the SoundWIRE Group [7] presents similar experiments where the task was also to sustain the tempo. This led to a stable tempo but an asymmetry where the second clapper lagged after the initiating clapper.

In a pilot experiment, Winge [8] placed two experienced musicians (piano and drums) in separated rooms with audio contact (over headphones) with a latency from 4 to 92 ms. The musicians' task was to play Lennon's 'Yesterday' together as well as possible. A reanalysis of his data showed a small tempo decrease, independent of the delay. However, the analysis also showed that the piano to a large extent lagged after the drums. The lag was close to the delay and suggests that the drummer kept a steady tempo independently of the piano while the piano player synchronized to the (delayed) drum rhythm and did not experience a delay.

In the present work, we wanted to expand on the work by Chafe and Gurevich [6] using an experimental setup similar to theirs. One modification compared to ref. [6] was that we wanted to study a situation that was symmetrical in the sense that no participant was giving a leading or lagging role. Furthermore, we wanted to study the influence of reverberation as well.

Our hypotheses for the effect of a delay on the performance are presented in Sec. 2. Section 3 describes the experimental setup, the signal analysis is described in Sec. 4, and Sec. 5 presents the results and a discussion of the results. Conclusions are drawn in Sec. 6.

2. HYPOTHESES

Previous studies have based their experiments on asking the musicians or clappers how difficult it was to play together under the prevailing conditions. Chafe et al. [6] and Winge [8] did also qunatify the way the participants performed by analyzing the recordings made. In this study we hypothesize that when the delay increases, the players tend to react in one or more of three ways:

- one player leads and the other one follows, causing a temporal *asymmetry*,
- both players try to stay ahead to avoid slowing down, which might reduce the *precision* in the players' timing, or
- none of the players try to stay ahead, resulting in a *tempo decrease*.

One reaction does not exclude another, and they may be more or less pronounced. The choice of strategy may be conscient, as it may vary as time passes. Musical experience is expected to help the players perform well even when the delay makes it difficult. All three may be quantified as described later (Sec. 4).

Finally, the introduction of reverberation might affect the influence of the delay.

3. HAND-CLAPPING EXPERIMENTS

In the experiments, pairs of subjects were set to perform simple ensemble playing by clapping rhythmical patterns with their hands. This was done in three different acoustic conditions:

- real reverberant conditions (RR): the subjects were placed symmetrically in a large lecture hall at different distances, thus exploiting the relatively low speed of sound,
- virtual anechoic conditions (VA): the subjects were placed each in either an anechoic or an acoustically dry room, connected to the other by a microphone and headphones through filters that delayed the sound, and
- virtual reverberant conditions (VR): the subjects were placed as in the VA conditions, but in addition to the delay, the computer also added artificial reverberation simulating that of the lecture hall with the clappers in the same positions as in the RR experiments.

The lecture hall had a reverberation time around 1.2 s at mid-frequencies, and a diagonal of about 30 m allowing distances up to 23 m without placing the subjects close to a wall. A sufficiently large anechoic room was unfortunately not available for performing experiments in real anechoic conditions (RA).

In all three conditions, the subjects were placed standing in front of the microphone. They were blindfolded in order to enhance their concentration on the aural input and to avoid that they knew their positions in the RR condition.

The VR condition was based on impulse responses (IRs) measured in the lecture hall used for the RR condition. Ideally, the IRs should have been measured with a sound source that had the same directivity as hand clapping. Here, a two-way active loudspeaker (Dynaudio BM6A) was used. Its directivity was verified in a simplified way by measuring the octave-band levels of both a person clapping hands and the loudspeaker reproducing anechoic recordings of the hand claps of the same person in the reverberation field of the lecture hall. Levels averaged over 26 hand claps were within ± 4 dB of each other for the two sound sources, and this was considered to be acceptably close.

The reverberation filters in the VR conditions was therefore found by measuring the binaural room impulse responses (BRIR) with the mentioned loudspeaker and a dummy head (Neumann KU80) in the position of the head of subject 1 and the lower loudspeaker diaphragm positioned at the position of the hands of subject 2. When measuring subject 1's own BRIR, the loudspeaker was moved to the position of subject 1's hands. This routine was repeated for each pair of positions. Since the lecture hall was symmetric, the BRIR was measured only for one side of the room, and the left and right channel interchanged for the other side. Headphone compensation was performed in Matlab [9] as was the time compensation for the internal delay of the

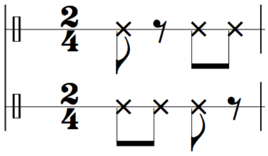

Fig. 1: Complementary rhythms used in the experiments

measurement system and the reproduction system. The direct sound was subtracted from the subject's own impulse response as the subject would hear their own clap almost without attenuation as the headphones were of an open type (Beyerdynamic DT990 Pro).

Since the reverberation in the lecture hall was no longer than 1.2 s for all frequencies, FIR filters of length 34728 at a sampling frequency of 44.1 kHz were used. The same filter length was used consistently to ensure a constant system delay. The filters were used in a Lake Huron real-time convolver [10] connected to AD/DA converters (RME TDIF-1). The VR and VA conditions were randomly interchanged in the same experiment and thus experienced by the same subject pair, while the RR experiments were conducted at a separate occasion and with different pairs of subjects.

To initiate each trial, both clappers were simultaneously given the tempo (called the *given tempo*) by an initiating recorded clapping sequence (recording of six 2/4-measures with a clap at each 1/4-note including a count-down before start), ensuring a symmetric start, i.e., avoiding biasing one of the subject to lead and the other to lag. They were asked to listen to the other and perform one of two complementary hand-clapping rhythms together in ensemble, while keeping the tempo as steady as possible. The rhythms are the same as those of Chafe and Gurevich [6], as shown in Fig. 1.

After each trial, they were asked to grade how well the ensemble playing was accomplished: good, ok, or bad. In the lecture-hall experiments (RR), they showed a thumb up or down, or an open hand if 'ok', in order to avoid influence from the other subject. In the virtual experiments, the subjects reported their judgement orally when the test assistant asked for it while muting the connection between the subjects.

Three different tempi of 86, 90, and 94 beats per minute (same as Chafe et al. [6]), six different delays of 5.9, 20.6, 32.4, 44.1, 55.9, and 67.6 ms (corresponding to distances 2, 7, 11, 15, 19, and 23 m), and three different room conditions (RR, VA, and VR) were used. Altogether, this led to 18 trials for each of the 12 pairs of subjects that participated in the lecture-hall experiment (of which one pair was discarded before the analysis as explained below) and 36 trials for the 11 pairs in the headphone-experiment. The order of the trials was randomized, while randomization of pairs was assumed accomplished by letting the persons choose time without knowing who the partner would be.

Tempo was changed by playing different recordings, which were constructed by recording a clapping/count-down sequence in an anechoic room and editing the recording by repositioning the claps corresponding to the wanted tempi. For the VR condition, the claps were filtered with the impulse response of the lecture hall at 7 meters so that the sound corresponded to the room where they were supposed to be. The filter and start-up recording was changed between each trial. In the lecture hall, the distance was simply changed by moving the subjects to their new position. They were always kept blindfolded, and some care was taken to confuse the subjects by taking detours and turning the subjects, although we could not completely inhibit them from guessing their position from aural input.

Assuming that musical experience might influence the results, we searched to avoid pairing persons without music experience with persons having some experience with an instrument. As music is a mandatory subject in the Norwegian school up to about 12 years of age, we used the term 'non-musicians' for those not having played an instrument or participated in a music ensemble such as a choir outside the mandatory school practice. All other subjects were denoted 'musicians', whether they had only played in a janissary band 15 years ago or they were actively practicing music on a high level. No professional musicians were used as subjects. This classification was done when recruting subjects.

Before the experiment, the subject pair was allowed some training together with eye contact and at close distance. This was partly to ensure that they were able to do the task and to let them experience clapping together under easy conditions. In 10 of the 22 accepted experiments, the subjects showed difficulties in clapping different rhythms together and were therefore asked to clap the same rhythmic pattern (second line in Fig. 1). (The 23rd pair, whose results were discarded, seemed to cope with clapping different rhythms during the training, but failed repeatedly during the experiment.) This unexpected problem for some of the participants to carry out the task, and subsequent modification of the task, implies that there were two slightly different tasks used in the experiment. We still decided to include the results from all subjects. However, unfortunately, the two tasks coincided almost completely with the two groups 'musicians' and 'non-musicians' (except for one pair in each category), which means that when these two groups are compared, there is also a confound variation in task.

In total, $6 \times 3 \times (1 + 2) \times 11 = 594$ trials were performed, of which one had to be discarded because there was an error with the sound file. The clappers were manually stopped after about 15 s, and the number of 1/8-notes finally recorded varied from 30 to 79 with a mean of 50.

The microphone signals for both subjects were recorded as a stereo signal with channel 1 for the left clapper or the one in the acoustically dry room (the 'Aura lab'), and channel 2 for the other. In the VA/VR experiments, the clappers exchanged room half-way through the experiment, and the channels as well as the statistics and judgements were exchanged accordingly before the analysis.

4. SIGNAL ANALYSIS

In the signal analysis, the time instant for each hand clap was extracted by a simple peak search. In the lecture-hall experiments, it was verified that the instants corresponded to three even 1/8-notes and an 1/8-note pause to avoid picking up claps from the other subject, especially at small distances and if the claps of one of the subjects were louder than those of the other, which was often the case. The tempo varied to some degree in most trials, and many clappers tended to deviate from a strict 2/4 rhythm, making this verification difficult. It was rather successfully achieved by watching three subsequent notes from note i, calculating the ratio of the standard deviation s_i of their inter-onset interval (IOI) times to their average m_i (denoted $\rho_i = s_i/m_i$), and comparing this to the same ratio for a correct measure consisting of two 1/8-notes and one 1/4-note ($\rho_{112} = s_{112}/m_{112}$):

$$\delta_i = \rho_i - \rho_{112} = \frac{s_i}{m_i} - \frac{s_{112}}{m_{112}}. \qquad (1)$$

The closest ratios ρ_{ijk} to the expected ρ_{112} are those of the note configurations $ijk = 113$ and $ijk = 122$. Note that the order of the notes does not matter for the calculations, nor will a moderately changing tempo be important. A criterion for acceptable IOI times is therefore:

$$\delta_{122} < \delta_i < \delta_{113} \qquad (2)$$

where $\delta_{ijk} = \rho_{ijk} - \rho_{112}$ and $\delta_{122} < 0$. This step was only meant to remove false claps, and no clap was removed if it would lead to a δ_i further from 0. It was not necessary in the virtual experiments.

For the clapping sequences to be exploitable, all the pauses (also claps that failed to make a loud enough sound) had to be filled to obtain two even sequences of 1/8-notes. Thus the sequences of the clapping times of the two subjects were aligned in pairs of almost simultaneous claps while new clapping instants were interpolated between the existing ones to pair with unpaired claps of the other subject or to fill in if the distance between two claps was too high.

Once the clapping instants and interpolated missing 1/8-notes were acquired, the various global measures for each trial, i.e., the tempo change, the rhythmic precision, and the asymmetry as well as the initial tempo, could be calculated. Let us denote the N timings for each subject k by $t_{k,i}$, where $i = 0, \ldots, N-1$, the $N-1$ inter-onset intervals (IOI) $\Delta t_{k,i} = t_{k,i+1} - t_{k,i}$, the N inter-*subject* differences (ISD) $\delta t_i = t_{2,i} - t_{1,i}$, the N mean timings $\bar{t}_i = (t_{2,i} + t_{1,i})/2$, and finally the $N-1$ mean IOIs $\Delta\bar{t}_i = \bar{t}_{i+1} - \bar{t}_i$. We may then express the wanted measures as follows:

Initial tempo τ_0: the average of the tempo of the first measure of four mean IOI times $\Delta\bar{t}_i$:

$$\tau_0 = \frac{1}{4}\sum_{i=0}^{3} \frac{60}{2\Delta\bar{t}_i} \qquad (3)$$

Tempo change $\Delta\tau$: the tempo increase from the first to the sixth measure divided by the time passed:

$$\Delta\tau = \frac{\frac{1}{4}\sum_{i=24}^{27}\frac{60}{2\Delta\bar{t}_i} - \tau_0}{\bar{t}_{24} - \bar{t}_0}. \qquad (4)$$

The choice of six measures was made because the shortest recorded sequence measured 30 1/8-notes, i.e., 6.5 measures.

Mean asymmetry α: the average of the ISDs δt_i, i.e., the mean time that subject 2 lags after subject 1:

$$\alpha = \frac{1}{N}\sum_{i=0}^{N-1}\delta t_i. \qquad (5)$$

We implicitely use the absolute asymmetry $|\alpha|$ in the following as it makes more sense when averaging over subject or other factors.

Ensemble imprecision σ: the standard deviation of the ISDs δt_i:

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{i=0}^{N-1}\delta t_i^2}. \qquad (6)$$

Precise clapping would give constant ISD, i.e., constant asymmetry and zero ensemble imprecision.

5. RESULTS AND DISCUSSION

Figure 2 shows four examples of clapping sequences performed by an arbitrary musician pair in the RR condition at different distances with a given tempo of 90 beats per minute (bpm). The circles show the instant tempo $60/2\Delta\bar{t}_i$ (cf. Sec. 4). The squares indicate the mean tempo of measures 1 and 6, and the solid sloping lines show the estimated tempo change.

Before going further, we should note that the initial tempo deviated from the given tempo and showed a decrease with respect to delay, as seen in Fig. 3. The error bars show the two-sided confidence intervals at 5 % confidence limit for each average, assuming a normal distribution. The comparisons were performed by a four-way Anova with distance, given

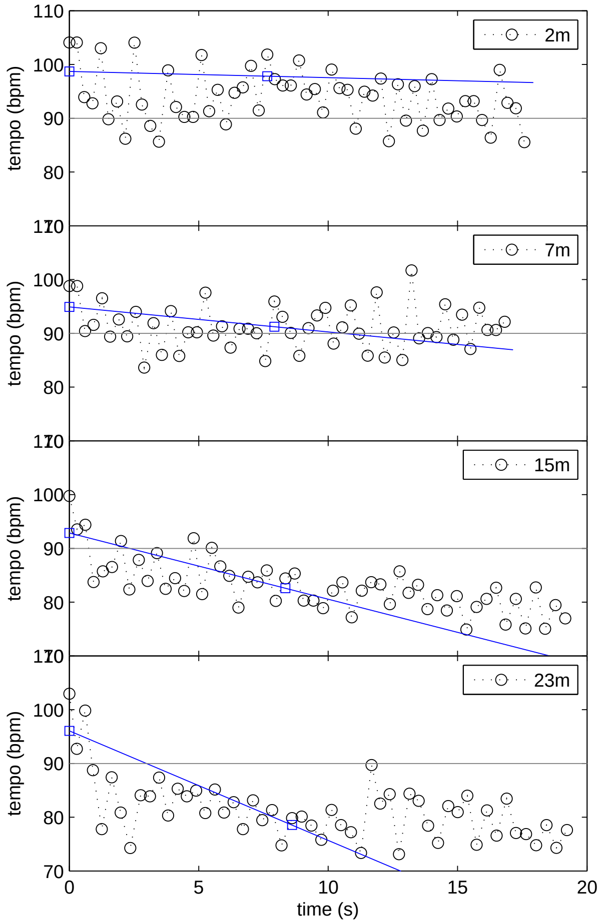

Fig. 2: Tempo change for four RR trials at 90 bpm at different distances performed by an arbitrary musician pair. Circles show instant tempo ($60/2\Delta\bar{t}_i$), squares indicate mean tempo per measure one and six, and solid lines show the given tempo (horizontal line) and the tempo change from the two squares (sloping line).

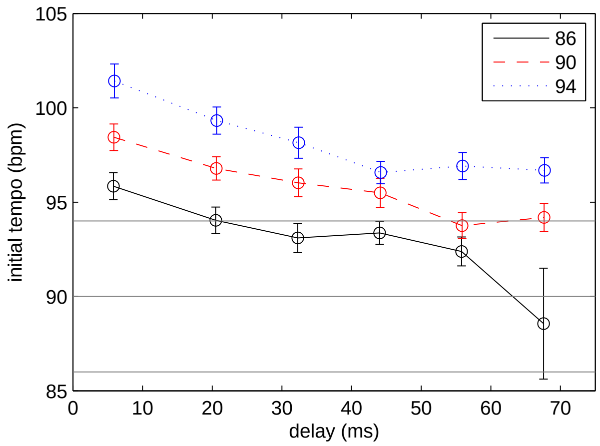

Fig. 3: The average *initial tempo* vs. *delay* (or distance/340 m/s) and *given tempo*. Error bars show confidence intervals (5 %).

tempo, musical experience, and acoustical conditions as the four factors. Differently from Chafe and Gurevich [6], we observed a statistically significant dependence between the initial tempo and the delay, which was a little surprising since the subjects were not supposed to be influenced by the delay before starting to clap. But from Fig. 2 it may be seen the tempo changed already during the first measure. We suggest two explanations for this dependence: an initial hesitation at long delays when the other does not seem to start, and the tempo change due to a delay that was manifested already in the first measure.

Although the initial tempo showed a significant dependence on the given tempo, as is evident from Fig. 3, we do not have a good explanation for the fact that the initial tempo on average was significantly higher than the given tempo. We may suggest a lack of anticipation of the first clap, but also the mean tempo for each clapping sequence was significantly higher than the given tempo. This was true even for the musicians, whose initial tempo and mean tempo was significantly lower (about 2.5 bpm) than for the non-musicians. The tempo of the recordings that gave the tempo was double-checked, and the question still resides.

The other characterizing measures: tempo change, ensemble imprecision, and asymmetry, were extracted from the 593 stereo clapping sequences (trials) as explained in Sec. 4. The averages of these measures for each delay are shown in Fig. 4 together with the averaged judgements of both clappers. The first graph in Fig. 4 shows the mean tempo change, i.e., the mean slope of curves like the ones in Fig. 2. The results show that the tempo decreased faster for longer delays. It may be mentioned that there was a small but significant tendency for the tempo to decrease more when the given tempo was higher. More interestingly, we observe an increase in tempo at delays less than about 18 ms on average. This confirms the observation by Chafe and Gurevich [6] that short delays tend to speed up ensemble clapping although their zero intersection was at 11.5 ms. As discussed further below, the zero-intersection from our experiments was around 15 ms when only the 'musicians' group was analyzed, and only this group did the complementary-rhythm clapping, as explained

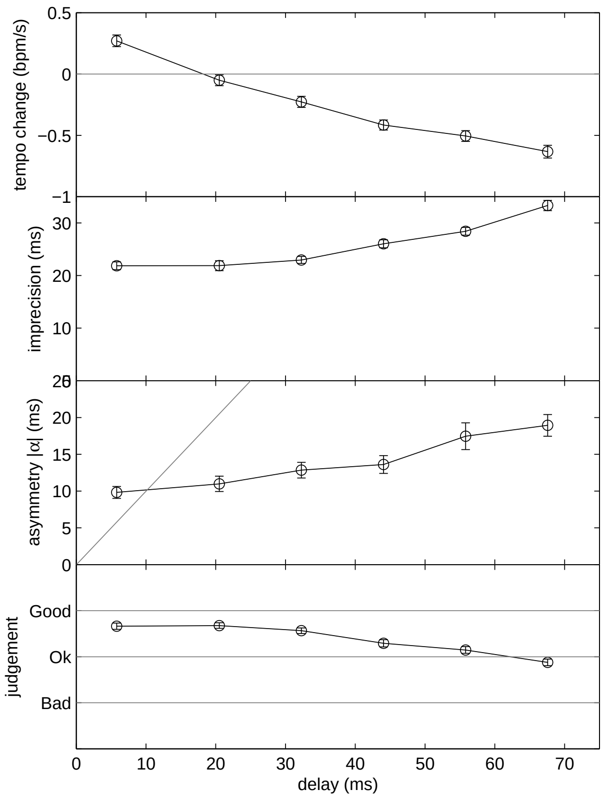

Fig. 4: Mean *tempo change*, *ensemble imprecision*, *asymmetry*, and *judgement* vs. delay. The sloping line indicates asymmetry equal to delay. Error bars show confidence intervals (5 %).

in Sec. 3.

In the two middle graphs in Fig. 4, we observe that the mean ensemble imprecision increased with increasing delay as did the mean asymmetry. There was no significant dependence on the given tempo. When it comes to the timing imprecision, there was a low-delay imprecision of about 22 ms up to the 32.4 ms delay, which may be considered as an average minimum imprecision for all our subjects. Above this delay, the imprecision started to increase, suggesting a threshold around 25 ms. The asymmetry did not manifest a threshold, merely a slow increase with the delay. At the lowest delay (5.9 ms), the mean asymmetry was about 10 ms, which is higher than the delay (illustrated by the gray sloping line). This implies that there is also a mean minimum asymmetry that is accepted as a normal deviation from symmetry.

Finally, the average subject judged the ensemble performance to be close to 'good' for delays up to about 20 ms, and gradually worse above this, as shown in the last graph in Fig. 4. Similar to the imprecision results, the judgement seems to exhibit a threshold around 25 ms. Reviewing the hypotheses stated in Sec. 2, we conclude that on average, all three reactions took place. In addition, there is evidence of a threshold above which the ensemble playing becomes gradually more difficult.

For all measures, there was a significant difference between musicians and non-musicians, as is shown in Fig. 5. It should be recalled that almost all the 'non-musicians' performed a different task, using the same clap rhythm rather than complementary rhythms, as explained in section 3. Persons with musical experience tended to slow down more than non-musicians. The upper limit for tempo increase (the 'Chafe effect') was lowered to about 15 ms. It may be somewhat surprising that musical experience did not stabilize the tempo. Indeed, the musically experienced slowed down more than the non-musicians. An explanation may be that they were less willing to disregard their partner in an urge to succeed with the ensemble play. Another possible cause for the smaller effect of the delay on the tempo of the musically less experienced is that when both subjects clapped the same rhythm, it was slightly more difficult to hear the claps of the other one, and thereby they were less affected by the other subject's timing.

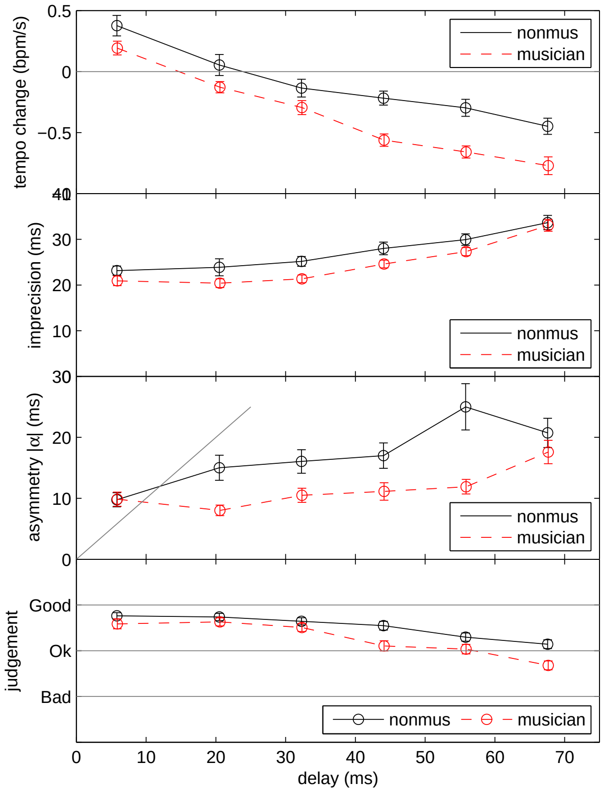

Fig. 5: Mean *tempo change*, *ensemble precision*, *asymmetry*, and *judgement* vs. delay and musical experience. The error bars show the confidence interval (5 %).

At the same time, the musicians were less satisfied with the ensemble play than the others, especially above the threshold established above: the difference between the two groups was higher above the 32.4 ms delay, as is evident in the judgement graph in Fig. 5. Furthermore, the two middle graphs in Fig. 5 indicate that musical experience/complementary rhythms may decrease the ensemble imprecision and the asymmetry. In fact, there was no signficant difference between the asymmetry for the musicians up to 56 ms.

Our final objective with this study was to examine the influence of reverberation on coping with a delay in ensemble playing. The results are summarized in Fig. 6, with one curve for each acoustical condition RR and VR (real and virtual reverberation) and VA (virtual anechoic conditions). Note that the upper graph is the initial tempo, which was significantly

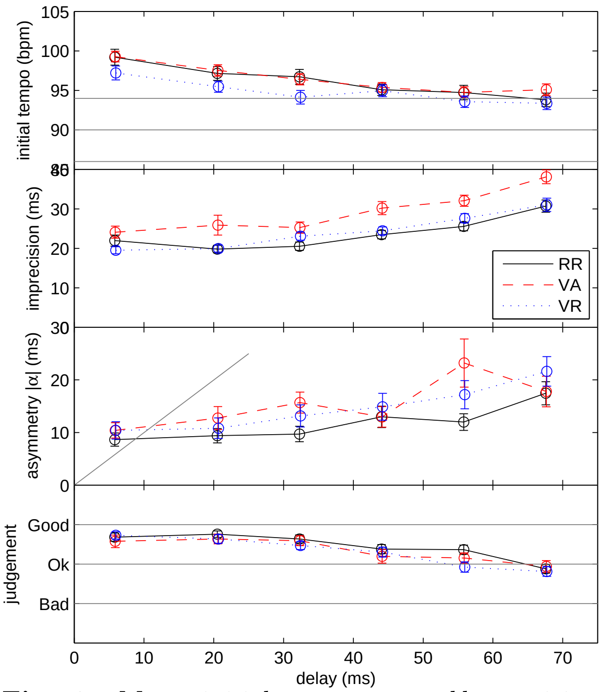

Fig. 6: Mean *initial tempo*, *ensemble precision*, *asymmetry*, and *judgement* vs. delay and acoustical conditions. The error bars show the confidence interval (5 %).

lower in the VR conditions than in the others. We do not have an explanation for this, especially because there were no significant differences in the tempo change, whose graph is not shown.

The imprecision was higher in the VA conditions than in the reverberant conditions. This was in spite of the fact that the VA signals were not attenuated due to distance in the way that the reverberation signals were. This may be an indication that reverberation simplifies the ensemble playing. For the asymmetry, the VA condition did not deviate from the two other conditions in the same clear way as for the imprecision. Also for the judgement, there were less clear patterns than for imprecision. In total, this could be seen as an interesting indication that reverberation might have an effect in the direction of decreasing imprecision.

6. CONCLUSIONS AND FURTHER WORK

The results of hand-clapping experiments under the influence of a delay up to 68 ms show that the mean tempo decreased close to linearly as function of the delay. During each sequence of around 15 s, the tempo slowed down to a degree that increased with the delay but the 'Chafe effect' was also observed: for delays shorter than about 15–23 ms, subjects actually increased the tempo during the sequence. This was true for pairs of musicians performing complementary clap rhythms (15 ms speed-up threshold) as well as for pairs of non-musicians performing the same rhythmic patterns (23 ms speed-up threshold).

For the timing imprecision, and for the subjects' judgements of their own ensemble performance, a threshold around 25 ms was found. For delays below this value, no effect of the delay could be observed for this group of subjects but above 25 ms an imprecision that increased with the delay was observed. The judgement was observed to decrease when the delay was increased above 25 ms. The performance was, however, judged as 'OK' up to around 50 ms. Virtual anechoic conditions lead to a higher imprecision than the reverberant conditions. Real-reverberation conditions lead to a slightly lower tempo than the two virtual conditions, but only a very small effect of the different conditions on the own judgement resulted. Since significant differences were found between the two groups 'musicians performing complementary rhythms' and 'non-musicians performing the same rhythm patterns', both of these factors should be investigated further.

More studies on the effects of reverberation might be needed to draw more clear conclusions. Virtual anechoic experiments are much easier to set up and perform than reverberant experiments, whether real or virtual, but significant differences were found, as described above. The simple ensemble experiment, as devised by Chafe and Gurevich and as used here, was efficient for quantifying effects on the ensemble of a delay. It could be expanded to three or more performers, and to other types of music performances.

Finally, this study has been based on averages of a great number of trials. We foresee, however, a study of the within-trial balance of subjects' reactions due to a delay: Do possibly asymmetry, imprecision, and tempo decrease sum up to a constant at a given delay?

7. REFERENCES

[1] W. Woszczyk, J. Cooperstock, J. Roston, W. Martens, “Shake, rattle and roll: Getting immersed in multisensory, interactive music via broadband networks,” J. Aud. Eng. Soc. 53, pp. 336-344, 2005.

[2] A. X. Xu, W. Woszczyk, Z. Settel, B. Pennycook, R. Rowe, P. Galanter, J. Bary, G. Martin, J. Corey, J. R. Cooperstock, “Real-time streaming of multichannel audio data over Internet,” J. Aud. Eng. Soc. 48, pp. 627-641, 2000.

[3] Ninjam, Realtime Music Collaboration Software, URL: www.ninjam.com

[4] eJamming Studio, URL: www.ejamming.com

[5] E. Chew, A. Sawchuk, C. Tanoue, and R. Zimmermann, “Segmental tempo analysis of performances in user-centered experiments in the distributed immersive performance project,” in *Proceedings of the Sound and Music Computing Conference (SMC)*, Salerno, Italy, Nov. 2005.

[6] C. Chafe and M. Gurevich, “Network time delay and ensemble accuracy: effects of latency, asymmetry,” in *Preprint no. 6208, 117th AES convention*, San Francisco, CA, USA, Oct. 2004, pp. 2–7.

[7] SoundWIRE Group at CCRMA: ccrma.stanford.edu/groups/soundwire/

[8] H. L. Winge, “Musikksamspill over IP,” Project report, Department of Telecommunication, NTNU, Trondheim, Norway, 2003, (in Norwegian).

[9] Mathworks, URL: www.mathworks.com

[10] Lake Technology, URL: www.lake.com.au