# TOWARDS A COMPREHENSIVE COGNITIVE ANALYSIS OF DELAY-INFLUENCED RHYTHMICAL INTERACTION

*Alexander Carôt and Christian Werner*
University of Lübeck, Germany
Institute of Telematics

*Timo Fischinger*
University of Kassel, Germany
Institute of Music

## ABSTRACT

Within the last decade, interdisciplinary researchers have been exploring the area of remote musical interaction. The current network infrastructure already provides high quality audio transmission but it also implies significant transmission latencies depending on the physical distance, network capacities and actual usage. If the latency exceeds a certain value, a realistic musical interplay becomes impossible. Despite numerous valuable investigations it has so far not been possible to figure the precise dimension of this value and in how far it might underly further cognitive or purely musical aspects. This paper refers to the most significant findings in this area, extracts problems and unsolved questions in order to establish an own comprehensive cognitive analysis of delay influenced rhythmical interaction.

## 1. INTRODUCTION AND PROBLEM

As a side effect of the constantly evolving Internet bandwidth capacities and the respective transmission quality, the domain of distributed music is becoming more and more attractive to musicians, engineers and computer scientists. Various projects exist, which aim at achieving live music conditions for displaced musicians as if they were in the same room. In that context and apart from technical solutions, a number of researchers have investigated cognitive aspects related to latency, which represents the most significant limiting factor in a distributed music performance. Each tried to identify a latency threshold up to which a real musical interplay is possible by having two persons performing in two places, separated by a delay processor, which was adjusted from zero ms up to a maximal delay threshold. The experimental setup and the outcome outcome of such experiments is diverse: The first fundamental work carried out by Chris Chafe [4] worked with a random choice of clapping persons and states acceptable latencies up to 35 ms in order to prevent a slow down of the performance speed. This value was confirmed by Alvaro Barbosa in [1], who worked with four semiprofessional musicians. Elaine Chew, however, works with two professional classical pianists and states a maximal commonly acceptable threshold of 50 ms while already taking the speed of a song into account [5]. Moreover, our own experiments suggested latencies not to exceed 25 ms [2] but we concluded that – in terms of defining a setting, which is able to deliver a general valid statement – the involved musicians should have a professional attitude in order to precisely judge about an actual playing condition. Secondly, the music should be as beat-driven as possible to represent the most timing-critical situation. In that context the speed of a tune and the actual note resolution might have an impact as well. Furthermore, with respect to the sound studio setup the direct and dry signal could be perceived as unnatural, which is why especially singers often ask for an artificial reverb in order to compensate this effect. Hence, an experiment should take this perceptual phenomenon into consideration. Finally, additional alternative interaction approaches according to [3] should be evaluated in case the total latency resides in unacceptable dimensions. Based on these requirements, this paper will present an approach, the goal of which is to overcome the general confusion about valid latency thresholds in delay-influenced musical interaction.

## 2. CONCEPT

We decided to ask 5 professional drummers to perform with a single professional bass player to be the rhythmical counterpart. This way a direct comparison of each rhythm section constellation is possible. The choice of the drummers assumed professionalism as a collective term for playing regular commercial concerts, playing tracks for professional sound studio productions, teaching students and an overall acceptance in the north german music scene. Furthermore, each musician held a university degree in music.

In terms of the trial procedure we choose a test scenario, which consists of subsequent $\frac{1}{4}$, $\frac{1}{8}$ and $\frac{1}{16}$ bass sequences. This basic test pattern is made of $3 \cdot 12 = 36$ bars. Furthermore, we composed a reference one bar drum pattern for each of the 36 bars shown in figure 1, however, drummers have the choice to play in such a way they consider as most suitable and convenient. The drum kit itself consists of bass drum, snare drum and hit hat.

Before the trials start both players practice the given sample piece under conventional conditions in the same room until they feel comfortable with it. As the trial starts the drummer first listens to a metronome for 4 bars before he plays 2 bars by himself. After these two bars the bass player starts joining in with the test pattern. The experiments are performed at speeds of 60, 100, 120 and 160 bpm and the delay between the two players is increased from 0 ms in steps of 5 ms until one of the player feels uncomfortable with the playing situation or players tend to slow down. In that context the players have to evaluate the actual delay situation as “excellent”, “tolerable” and “not tolerable”. In the latter case we additionally apply an artificial reverb to the received signals and force the drummers to modify their played rhythms regarding a shuffled or straight phrasing. If this still leads to an unacceptable situation the end of the trial is reached.

Subsequently, in order to evaluate compromised playing conditions, we firstly choose an unacceptable one-way delay (OWD) according to the results of the first evaluation phase and increase this one-way delay in steps of 15 ms. Twice this value corresponds to the roundtrip time delay (RTT). This is added as an artificial delayed feedback to the bass player. In the subsequent trials the bass player‘s self delay is decreased in steps of 5 ms in order to figure a relation between the self-delay and the OWD. After the completion of this experiment we use the same OWD but apply it as well as the delayed feedback to both players. Respectively the self-delay decrease will be applied to both player‘s delay in steps of 5 ms.

The last evaluation concept is based on the same performing musicians and an additional professional saxophone player. A “laid back” impression of a musical solo is the subject of this experiment. Altogether 5 rhythm section constellations play in one room with the delayed remote saxophone player in another room of the studio. Subjects are jazz standards at speeds of 60, 100, 120 and 160 bpm. As in the previous experiments the delay is increased in steps of 5 ms until the the rhythm section complains about the saxophone player‘s solo. Again the player‘s have to evaluate the situation as “excellent”, “tolerable” and “not tolerable”.

## 3. EVALUATION

The applied evaluation scheme for a successful trial assumed a stable timing at the given bpm. Musicians were asked to figure their personal limit of acceptance compared with a conventional scenario on stage. In case of an inconvenient situation musicians were forced to end the trial. However, negative tests were repeated with an artificial reverb added to the direct signal in order to simulate the natural room reverb. The results for each performing constellation are illustrated in figure 2. For each constellation it shows the delay acceptance value in ms (x-axis) against the speed and the note resolution of the performed pattern: Each constellation corresponds to a block of three columns, where the left column shows the performance in $\frac{1}{4}$ bass note resolution, the middle column in $\frac{1}{8}$, and the right column in $\frac{1}{16}$ resolution. The black color indicates a perfect playing condition, while

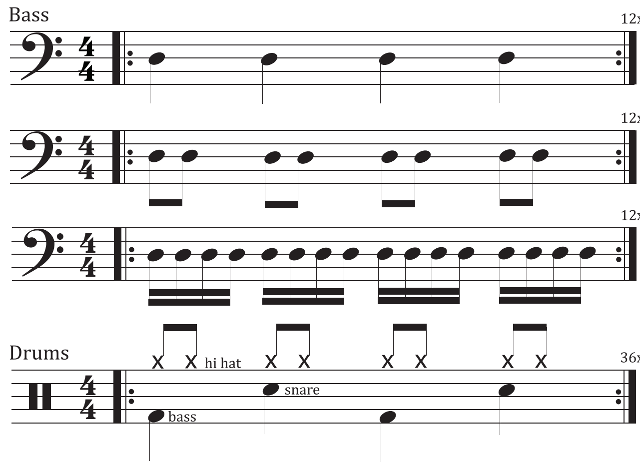


**Figure 1**. Performed evaluation bass pattern with reference one bar drum pattern

the grey color indicates a tolerable situation. Beyond the maximal grey value the musical interaction was not possible anymore. The results show that the overall delay thresholds ranges between a minimal delay of 5 ms and a maximal delay of 65 ms.

The most important observation is, that the players do not exhibit a common latency acceptance value as the direct comparison for each bpm trial shows. E.g. the largest difference of 35 ms occurs between constellation 3 and 5 at a speed of 60 bpm with $\frac{1}{4}$ note resolution and between constellation 3 and 4 at a speed of 100 bpm with $\frac{1}{4}$ note resolution. As each player‘s test results vary in such an extreme way it is not possible to define a general valid delay threshold. It must instead be considered as an individual acceptance value, which depends on the player‘s rhythmical attitude. Apart from that outcome, each player‘s graph shows a similar progression but the actual dimension differs significantly. Furthermore, for each player it is obvious that – apart from a few exceptions – the delay acceptance threshold is directly related to the speed and the note resolution of the performed pattern. Generally a faster bpm and a higher note resolution lead to a lower acceptance threshold for each of the players, however, the transitions are more or less obvious: In some cases the threshold remains equal between subsequent trials, where either the bpm or the note resolution was increased (e.g. see constellation 4 at 100 bpm and constellation 1 at $\frac{1}{16}$ note resolution for 100 and 120 bpm. A surprising effect occurs with 4 constellations at 120 bpm: Here the $\frac{1}{8}$ note pattern could be better performed than the remaining patterns. Later the players‘s stated that speed of 120 bpm with a slightly laid back $\frac{1}{8}$ bass pattern represents the typical rock music style, which is why they felt more comfortable in this situation. Equally constellation 5 had problems with $\frac{1}{4}$ notes at 120 bpm: Although the bass player had no problems with delays at 5 and 10 ms, the drummer had stylistic issues. Finally, a repetition of the same trial

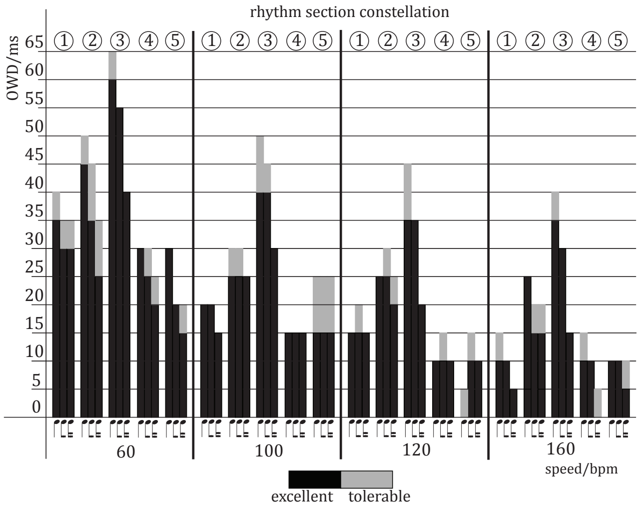


**Figure 2**. Test results of the performing drum/bass constellations

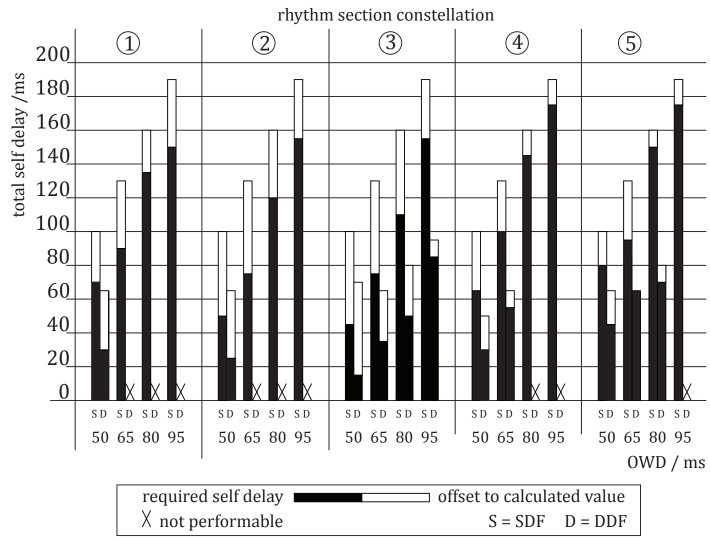


**Figure 3**. Test results of the performing drum/bass constellations with single (SDF) and dual delayed feedback (DDF)

with a delay of 0 ms didn't lead to any complains.

Anyhow, the graph does not illustrate aspects related to an artificial signal reverb: In none of the trials we could observe a positive impact of this effect. The players rather complained about a lack of clear note onset and in turn preferred playing with the pure unmodified signal. Hence, we did not consider this effect as relevant and useful. Neither rhythm scheme changes from binary to ternary beats lead to an improvement of the playing conditions, which is why we did not consider them as a useful information in the graph either.

After the completion of the experiment‘s first part we approached the compromised interaction evaluation: As a first step we choose a speed of 120 bpm at $\frac{1}{8}$ note resolution and applied a one-way delay of 50 ms as due to the first experiment‘s outcome this parameter combination resulted in a not performable situation for each of the constellations. Then we applied the self delays as described in the concept and plotted the results in figure 3:s The term SDF describes the case, in which only the bass player is delayed, while DDF symmetrically provides the self delay to both players. According to our expectations the players could cope with a delayed feedback up to a certain limit. However, only constellation 3 was able to perform in each of the proposed trials but claimed that self-delays beyond 30 ms could not be considered as a natural situation anymore. Other drummers canceled the respective trials, where the delayed feedback exceed their personal acceptance limit. Such trials are marked with an "X" and declared as not performable. Furthermore, the main outcome of this second experiment is that the delayed feedback can be decreased up to a point, where it significantly undershoots the expected theoretical value. Asymmetric self delay trials are not considered since an additional practical analysis exhibited situations equal to the symmetric self delays.

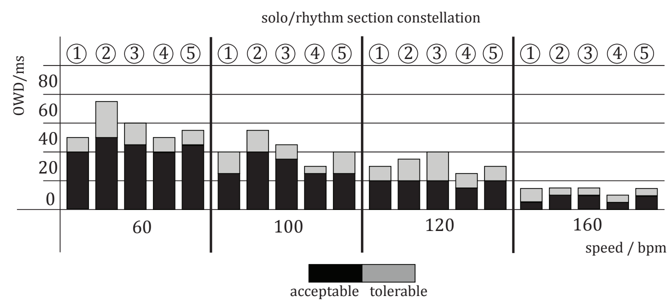


**Figure 4**. Test results of the performing drum/bass constellation with a remote saxophone solo player

Finally, in the last experiment we evaluated the compromised "laid-back" principle. Subject of this trial was the jazz standard "The days of wine and roses" of again 60, 100, 120 and 160 bpm, which consisted of one chorus of melody, one chorus of solo and another final melody chorus. The delay between the rhythm section and saxophone player was increased until the rhythm section complained about a confusing melody or solo. The results were plotted as illustrated in figure 4. Again the graph indicates a perfect situation in black color and a tolerable solo in grey color. Beyond this maximal value the solo was perceived as "out of time". Again it is clear that each constellation‘s delay acceptance significantly decreases with the speed of the tune. Apart from the fact that the players again exhibit slight threshold deviations, the general delay acceptance resides at higher dimensions in comparison to an uncompromised interaction. Nevertheless, this statement looses significance beyond a speed of 100 bpm.

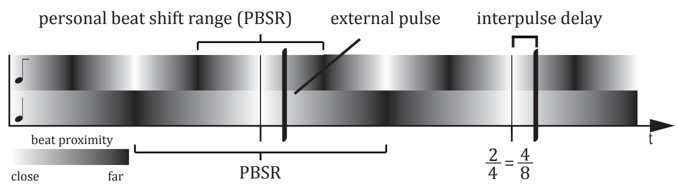


**Figure 5**. Visualization of human musical interaction

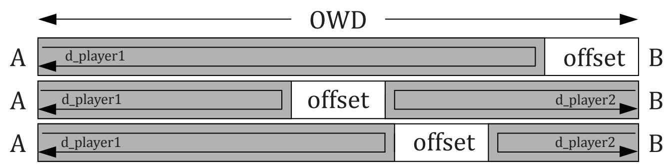


**Figure 6**. Visualization of the SDF (top) and DDF (middle and bottom) principle including the concept of a delay offset.

## 4. CONCLUSIONS AND FUTURE WORK

Rather than playing in precise synchrony, two musician‘s pulses typically exhibit a certain rhythmical inaccuracy. This inaccuracy is illustrated in figure 5 as the “inter pulse delay” (IPD) between the local pulse and an external pulse. Obviously some musicians accept a higher IPD than others but this acceptance is bound to the note resolution: It ranges around a rhythmical entity, which we define as the “personal beat shift range” (PBSR). Hence, it is not possible to define a common valid latency threshold, under which musical interaction is feasible or not. Furthermore – since a musical interaction implies the consideration of two musician‘s PBSR and the corresponding playing styles – it is rather a collective latency-acceptance-limit, which determines this latency threshold. Hence, we introduce the term “ensemble delay acceptance limit” (EDAL) as a typically not a known number, which must be figured in a dedicated test setup. Nevertheless, the EDAL falls with an increasing note resolution and the bpm.

For a compromized self-delayed musical interaction it was clearly obvious that the delayed feedback can be reduced up to a certain amount, which automatically leads to a delay offset between the players. According to our experiment‘s outcome the minimal required total self delay $dF_{min}$ can be calculated by equation 1: Firstly the EDAL is subtracted from the actual OWD and the result is multiplied by the factor 2. Secondly the equation adds a certain amount of delay due an undetermined human error $\varepsilon$, which occurs because of the inconvenience of the delayed feedback principle. In our experiments we could observe that $\varepsilon$ increases with the amount of self-delay and in some cases even leads to a not performable situation. Finally, the actual offset can be calculated, by a subtraction of $dF_{min}$ from the roundtrip time delay (RTT). Nevertheless, although DFA improves the delay situation between two musicians, it is no doubt that a delay of one‘s own signal typically must considered as inconvenient and not natural. The larger the delay gets and the louder the instrument‘s direct noise, the worse the realistic instrument feel and playing conditions.

$$dF_{min} = (OWD - EDAL) \cdot 2 + \varepsilon \tag{1}$$

$$\text{offset}_{max} = RTT - ((OWD - EDAL) \cdot 2 + \varepsilon) \tag{2}$$

The second alternative interaction approach lies in the separation of the rhythm and the solo section. Due to this constellation the rhythm sections could keep a stable beat, while the saxophone player was received with an artificial “laid back” delay effect equal to the RTT. In how far this effect was considered as disturbing or not obviously depends on the speed of the performed tune. As expected, again the results exhibit slight deviations from player to player but as an overall conclusion it is clear, that in this distributed constellation – especially for slow pieces – most of the players accept a higher latency. Beyond 100 bpm, however, this modified setup does apparently not provide better performance conditions.

This comprehensive analysis answers significant questions in context with delay in musical performances and furthermore suggests new forms of musical interaction. Nevertheless, in the future we consider to investigate the domain of conducted classical music. Due to the role of a conductor here a different form of musical interaction is present: Rather than the interplay of a rhythm section, the visual cues form the rhythmical reference of an orchestra. This principle has so far hardly been examined in terms of a displaced scenario and hence requires additional scientific attention.

## 5. REFERENCES

[1] Alvaro Barbosa. Displaced soundscapes, 2006. PhD Thesis.

[2] Alexander Carôt. Livemusic on the internet, 2004. Diplomarbeit.

[3] Alexander Carôt and Christian Werner. Network music performance – problems, approaches and perspectives. In *Proceedings of the “Music in the Global Village“ - Conference*, Budapest,Hungary, September 2007.

[4] C. Chafe, M. Gurevich, G. Leslie, and S. Tyan. Effect of time delay on ensemble accuracy. In *Proceedings of the International Symposium on Musical Acoustics*, Nara,Japan, March 2004.

[5] Elaine Chew and Alexaner Sawchuk. Distributed immersive performance. In *Proceedings of the North American Summer Meeting*, Brown University, Providence, Rhode Island, USA, June 2004.