universität wien

# MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

## "Match me if you can: Continuous dynamics of learning to coordinate auditory-motor rhythms with a partner"

verfasst von / submitted by

Constanze Leeb, BSc BA MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2021 / Vienna 2021

| | |
|---|---|
| Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet: | UA 066 013 |
| Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet: | Masterstudium Joint Degree Programme MEi:CogSci Cognitive Science |
| Betreut von / Supervisor: | Prof. Dr. Natalie Sebanz (Central European University (CEU), Vienna) |

# Abstract

People in every culture not only listen to music, but also produce music with one another (Greenberg et al., 2015). Group music-making requires that partners coordinate the timing of actions with one another. The underlying mechanisms of interpersonal auditory-motor synchrony have been studied across a range of contexts, typically involving coordination of discrete movements and sounds, e.g. piano duets (Goebl & Palmer, 2009; Loehr et al., 2013; Loehr & Palmer, 2011; Zamm et al., 2015) and finger tapping (Konvalinka et al., 2010; Zelic et al., 2019). One mechanism that has been shown to influence coordination between partners is auditory feedback coupling (Demos et al., 2019; Goebl & Palmer, 2009; Zamm et al., 2015). However, the influence of auditory feedback coupling has been investigated primarily in the context of discrete actions. One open question is how auditory feedback coupling influences coordination of continuous motion.

The current experiment aims to fill this gap by looking at continuous auditory-motor rhythms from a dynamical systems perspective. Partners rhythmically moved their finger on a ring-shaped touch sensor under different Auditory Feedback Conditions: (i) Uncoupled (partners heard themselves), (ii) Unidirectional 1 (both heard Partner A), (iii) Unidirectional 2 (both heard Partner B) and (iv) Bidirectional (partners heard each other). The stability of synchrony was influenced by auditory feedback coupling, resulting in an enhancement in stability of synchrony for Bidirectional relative to other feedback conditions. In addition, synchrony was influenced by auditory feedback coupling: the Bidirectional Condition resulted in the smallest phase offset (i.e. highest synchrony). The Uncoupled Condition showed large phase offsets, reflecting chance-level phase alignment between movements occurring at the same frequency.

The phenomenon studied in this thesis tackles a central question within cognitive science, namely how people manage to successfully coordinate actions with one another. The thesis is rooted within the realm of social interaction, but insights from various disciplines were needed to answer the posed question. Specific disciplines drawn upon were: psychology (for theoretical motivation), computer science (for programming of the experiment), physics and engineering (for construction of auditory stimuli). Findings yield interesting insights into mechanisms of interpersonal entrainment, which are relevant for cognitive research, where a current focus is on understanding the relationship between entrainment at the behavioral and neural levels. Further, findings about sensorimotor synchrony can provide important insights for rehabilitation and treatments of motor coordination diseases, such as Parkinson (Fujii & Wan, 2014).

# Zusammenfassung

Menschen aller Kulturen hören und produzieren Musik miteinander (Greenberg et al., 2015), indem sie ihre Bewegungen zeitlich koordinieren. Die Mechanismen der zwischenmenschlichen auditiv-motorischen Synchronisation wurden in verschiedenen Kontexten untersucht, typischerweise mit Fokus auf diskrete Bewegungen und Klänge, z.B. Klavierduette (Goebl & Palmer, 2009; Loehr et al., 2013; Loehr & Palmer, 2011; Zamm et al., 2015) oder Fingertappen (Konvalinka et al., 2010; Zelic et al., 2019). Ein Mechanismus, der die Koordinierung zwischen Menschen beeinflusst, ist die auditive Kopplung (Demos et al., 2019; Goebl & Palmer, 2009; Zamm et al., 2015), welche jedoch hauptsächlich für diskrete Aktionen untersucht wurde. Offen bleibt, wie sich auditive Kopplung auf die Koordinierung von kontinuierlichen Bewegungen auswirkt.

Das beschriebene Experiment versucht diese Lücke zu schließen, indem es kontinuierliche auditiv-motorische Rhythmen aus der Perspektive dynamischer Systeme betrachtet. PartnerInnen bewegten je einen Finger rhythmisch auf einem ringförmigen Drucksensor unter verschiedenen akustischen Feedback-Bedingungen: (i) Ungekoppelt (PartnerInnen hören sich selbst), (ii) Unidirektional 1 (beide hören Partner A), (iii) Unidirektional 2 (beide hören Partner B) und (iv) Bidirektional (PartnerInnen hören einander). Die Stabilität der Synchronität wurde durch die auditive Kopplung beeinflusst, was zu einer besseren Stabilität unter der Bidirektionalen Bedingung im Vergleich zu anderen Feedback-Bedingungen führte. Zusätzlich wurde die Synchronität durch die akustische Kopplung beeinflusst: die Bidirektionale Bedingung führte zum kleinsten Phasenunterschied (d.h. beste Synchronität). Die Ungekoppelte Bedingung zeigte den größten Phasenunterschied, der mit einem Zufallsniveau vergleichbar ist.

Das in dieser Arbeit untersuchte Phänomen beschäftigt sich mit einer zentralen Frage der Kognitionswissenschaften: wie es Menschen gelingt, erfolgreich Aktionen miteinander zu koordiniern. Die Arbeit beschäftigt sich mit sozialer Interaktion und benötigt zur Beantwortung der Forschungsfrage verschiedenen Disziplinen, wie etwa Psychologie (theoretische Motivation), Informatik (Programmierung des Experiments) oder Physik und Ingenieurwissenschaften (Konstruktion der akustischen Stimuli). Die Ergebnisse der Arbeit liefern interessante Einblicke in Mechanismen des zwischenmenschlichen Entrainments, wobei ein aktuelles Augenmerk in der Forschung hier auf dem Verständnis der Beziehung zwischen Entrainment auf der Verhaltens- und der Neuronalenebene liegt. Zusätzlich können Resultate bezüglich sensomotorischer Synchronisation wichtige Erkenntnisse für Behandlung von Erkrankungen der motorischen Koordination, wie etwa Parkinson (Fujii & Wan, 2014), liefern.

Acknowledgments

I want to thank Natalie Sebanz for the possibility to be part of the SOMBY-Lab and for supervising this thesis. Further, I want to thank my co-supervisor Anna Zamm. She provided me with the most incredible co-supervision one could wish for. Anna, thank you for always having time for me (online and offline, in person and remote), all your support and the amount of knowledge you passed on to me, ranging from academic insights to soft-skill tips essential for a career in academia. In addition, I want to thank you for all the nice times we had together when running the experiments as well as during our extensive zoom-coding sessions.
Fanni Takatsy, thank you for your constant and fast help with all my questions in the lab (or per mail) and for the nice times we shared in the lab! I also want to thank Anna-Katharina Bauer for her suggestions regarding the analysis and Thomas Wolf, both part of this project. Thanks are also due to the Bela team for their technical support on the Bela Forum as well as to all the participants both of the experiment and of the pilots.

As this thesis also marks the end of my Cognitive Science studies, I also want to thank Lisl, Toma and in general the MEi:CogSci team for their help and support throughout my studies. Further, I want to thank the people I was able to work with as part of this program. Many thanks to David Garcia, Max Pellert, Uroš Kovačič and Maja Derlink.
The past two years would not have been the same without my great cohort from all over the world who made logging on/going to university always a pleasure. Thanks to all the Ljubljana-Vienna-people (remote and really there) for an incredible semester abroad in a pandemic - the shared walks, coffees and lunches were the support I needed! I also remember all the shared coffees here in Vienna, every single one of my colleagues has made MEi:CogSci amazing.

Thanks to all my other friends for their support, ranging from having an open ear, being part of the pilots or bringing Sachertorte in stressful times. Special thanks to Marlene, Vici, Daniela, Jolly, Ela and Mara, I count myself very lucky to have such incredible friends in my life.
Manu, thank you for all your (statistical) support (also in the middle of the night), for finding the best pun-titles and for the incredible team we make - I cannot wait to start our new chapter together (see Physics thesis for more).
Lastly, I want to thank my family for all their support. Thanks to my parents for teaching me how to look at everything from more than one angle. Thank you to my sisters and brothers in law for all the life lessons they share without realizing it and thanks to tiny Theo, who was my motivation to finish my thesis quickly so that I could look at more cute pictures of him.

# Contents

# 1. Introduction

Humans are a deeply social species and are immersed in a social context much of the time. In many situations, social interaction of some form takes place. This often requires individuals to coordinate their behaviors (Loehr & Palmer, 2011), for example when rowing a boat, playing an instrument in an orchestra, playing a card game or lifting a heavy box together. If two or more people start to coordinate their behavior and their actions to achieve a common goal, joint action has been established (Loehr et al., 2013). The emerging field of Joint Action research takes an interdisciplinary approach to investigating the mechanisms that underlie interpersonal coordination of actions at the levels of brain and behavior (Sebanz et al., 2006). The field typically distinguishes between two forms of joint action, namely turn-taking (such as during spoken conversation) and simultaneous action (which occurs when people play music together). This thesis focuses on the latter form of coordination, namely on the mechanisms of how social partners coordinate simultaneous joint actions.

Different approaches to coordination of simultaneous joint action can be taken. While representational approaches focus on cognitive representations of joint action plans, dynamical approaches focus on coordination as an emergent property of systems (i.e. humans) that are coupled. This coupling arises from information exchange, more specifically via sensory information (Strogatz & Stewart, 1993). The study within this thesis was inspired by dynamical systems approaches.

Performing actions simultaneously with other people requires to monitor not only one's own but also the behavior of others in parallel to achieve a shared goal (Loehr et al., 2013). Interestingly, the coordination of movements often happens quickly and without much conscious effort (Konvalinka et al., 2010). Motor simulation, in which one simulates the movements of another person, can help to facilitate interactions, such as dancing with a partner or playing in an ensemble (Novembre et al., 2014). This can lead to synchronized behavior and movements, both unintentionally (e.g. rocking of chairs (Richardson et al., 2007)) and intentionally (e.g. playing a piano duet (Loehr & Palmer, 2011)), which need to be coordinated temporally.

## 1.1. Interpersonal synchrony of auditory-motor rhythms

#### Sensorimotor synchronization

There are many situations in which people need to coordinate their movements in time, such as dancing, walking or playing instruments together (Goebl & Palmer, 2009). This often requires people to move parts of their body to a certain spatial point at a specific time event (Torre & Balasubramaniam, 2009) in synchrony. Individuals require sensory information from the signal they are trying to coordinate actions with (Konvalinka et al., 2010). Synchronization of actions with an external signal is therefore referred to as sensorimotor synchronization (Torre & Balasubramaniam, 2009). It is an integral part of our everyday life (Zelic et al., 2019).

Rhythm is used as a communication tool in many social interactions and can facilitate sensorimotor synchronization. It is not only a common feature when people are dancing or performing music together, but can further be found in other everyday social interactions, such as speech (Fujii & Wan, 2014; Haegens & Golumbic, 2018) or the unintentional alignment of rocking chair sway between people sitting side-by-side (Richardson et al., 2007). Rhythm helps to mutually adapt and anticipate behavior even on a small time-scale (Demos et al., 2019), resulting in temporal coordination.

The process of synchronization or entrainment of behavior with another person unfolds over time (Bauer et al., 2018). Movements can be characterized by frequency and phase. One speaks of synchronization of motions, when phase and frequency of these movements become entrained, e.g. rhythmic behavior (Konvalinka & Roepstorff, 2012). An underlying mechanism of interpersonal synchrony is the alignment in frequency and phase between neuronal signals and a sensory input (Haegens & Golumbic, 2018). Phase-locking between the stimulus and the brain oscillations can occur, as well as an increase in power of the neuronal signal at the frequency of the sensory stimulus, which in turn affects sensory processing and behavior (Bauer et al., 2018; Haegens & Golumbic, 2018). This shows that social interaction can shape through synchronization with a stimulus, such as another person, basic cognitive processes such as perception (Loehr & Vesper, 2015).
The study presented in this thesis focuses on rhythmic auditory-motor synchronization of partners.

### Dynamical Systems Theory

A common theoretical framework to understand how people adapt to one another and anticipate each others' actions when synchronizing their behavior in time is dynamical system theory (Pikovsky et al., 2001; Schuster & Wagner, 1989; Strogatz & Stewart, 1993). Following van Gelder, 1998, a system consists of a set of interdependent variables that can be in different states at different time points. If any variable of a system changes, any other might change their state as well. The behavior of a dynamical system is the transition between states. In addition, systems can couple to their environment which can in turn consist of another system. This coupling results in state changes of the system whenever its environment changes as well (van Gelder, 1998). In the example of two people interacting, one individual can be viewed as a system that is coupling to their environment, which is another person/system. The partners' synchronized behaviors are the result of the transitioning between states of each system. Using the idea of coupling from dynamical systems theory enables us to connect two separate systems with one another and explain their common behavior.

The described dynamical systems framework allows for an explanation of synchronization of rhythmic behavior between individuals. In this approach, individuals are treated as harmonic oscillators. One speaks of entrainment if two oscillators, which can generate a rhythm themselves even without a rhythmical input, synchronize their oscillations (Haegens & Golumbic, 2018). An oscillator introduces a coupling strength, as well as an inherent frequency. The strength of the coupling determines how well oscillators (i.e. people) can synchronize. The inherent frequency of an oscillator/person can be understood as individual differences in preferred movement (Richardson et al., 2007). Interestingly, the more similar the spontaneous frequencies of people/oscillators are, the better they are at synchronizing their movements to one another (Zamm et al., 2015). In addition, the Dynamical Systems Theory makes it possible to introduce a metric for quantifying synchronized behavior (van Gelder, 1998).

### Role of sensory feedback in synchronization

Sensory feedback between oscillatory systems is essential for sensorimotor synchronization. This is of interest as the Dynamical Systems Theory can be applied to humans. The nature of feedback from a partner, for example in piano duet performance, has a direct impact on the pattern of coupling (Demos et al., 2019). The feedback humans receive from their partner can take on different physical forms. Haptic information usually leads to intentional synchronization, for example in ballroom dancing or when pushing weight together (van der Wel et al., 2011), whereas visual or auditory input can result in both intentional and unintentional entrainment. This thesis focuses on intentional auditory-motor synchronization.

It can help to synchronize movements if one is able to see what the partner is doing (Goebl & Palmer, 2009; Richardson et al., 2007; Richardson et al., 2005). This visual information can be directly linked to the movement of interest, such as seeing the actual rocking frequency of the partner's chair or the other person when performing a duet (Palmer et al., 2019; Richardson et al., 2007), or indirectly linked, such as head movements of performing musicians (Goebl & Palmer, 2009). Musicians often use body movements to express their intent (Goebl & Palmer, 2009). The amount of visual information, for example focal or peripheral vision, can influence the coupling pattern between individuals (Richardson et al., 2007).

In the case of auditory-motor synchronization, auditory feedback is critical for synchronization. Musicians use their own and their partners' auditory outcomes to entrain their movements with one another (Loehr & Palmer, 2011). People monitor their own and their partners' actions and are sensitive to errors that affect the common goal in musical performance (Loehr et al., 2013), which arises from joint auditory outcomes of one's own and a partner's actions.

In situations in which all auditory information is available, pairs tend to mutually adapt to one another (Konvalinka et al., 2010; Zamm et al., 2015) and to form bidirectional coupled systems (Demos et al., 2019), which can be explained through models with harmonic oscillators and are only delayed by an inherent frequency difference (Demos et al., 2019). Each partner uses the information about their own movements, as well as information about their partner's behavior. Individuals anticipate and adapt to their partner's movements (Konvalinka et al., 2010; Zamm et al., 2015). However, this situation changes if the auditory feedback of one partner is removed. Demos et al., 2019 and Konvalinka et al., 2010 showed that the removal of auditory feedback results in a unidirectional coordination pattern. The partner whose feedback has been removed for both participants starts to anticipate the other's movement and becomes the follower, while the person whose feedback is still provided becomes the leader (Demos et al., 2019; Konvalinka et al., 2010). When auditory feedback is restored, this unidirectional system becomes again a bidirectional system. If auditory information from both participants is removed, no coupling can occur, resulting in an uncoupled system which can also revert to a bidirectional one after restoration of the feedbacks (Demos et al., 2019).

## 1.2. Discrete vs. continuous auditory-motor rhythms

Much of the research on interpersonal synchrony of auditory-motor rhythms has focused on how partners coordinate discrete movements and sounds (Demos et al., 2019; Goebl & Palmer, 2009; Konvalinka et al., 2010; Loehr et al., 2013; Loehr & Palmer, 2011; Loehr & Vesper, 2015; Zamm et al., 2015; Zelic et al., 2019). Specifically, much of the literature investigates synchronization of piano duets, where pianists produce discrete piano key strokes (Demos et al., 2019; Goebl & Palmer, 2009; Loehr et al., 2013; Loehr & Palmer, 2011; Loehr & Vesper, 2015; Zamm et al., 2015). In addition, a discrete metronome signal was often used as a rhythmical pacing signal to study interpersonal synchrony of auditory-motor rhythms with a partner (Demos et al., 2019; Goebl & Palmer, 2009; Loehr et al., 2013; Loehr & Palmer, 2011; Zamm et al., 2015). Dyadic finger-tapping is another widely used paradigm for studying how partners synchronize auditory-motor rhythm (Heggli et al., 2019; Konvalinka et al., 2014; Konvalinka et al., 2010; Novembre et al., 2017; Repp, 2010; Repp & Doggett, 2007; Zelic et al., 2019), either with a discrete metronome or with a partner under different auditory feedback conditions.

Research on auditory-motor synchronization in the discrete realm showed that the amount of auditory information (i.e. auditory feedback) is crucial for movement synchronization and influences the resulting coordination patterns, as described above. However, in many natural auditory-motor synchronization contexts (such as music), movement and sound are not discrete but rather continuous (e.g. violinist playing legato tones, flutist playing legato tones, etc.).

Not only are discrete and continuous sounds perceptually distinct, but the motions required to produce them are different. A discrete motion consists of events that occur singularly and are preceded and followed by a period of time without this motion (Huys et al., 2008). Examples would be finger tapping (Konvalinka et al., 2010; Zelic et al., 2019) or the pressing of piano keys (Demos et al., 2019; Goebl & Palmer, 2009; Loehr et al., 2013; Loehr & Palmer, 2011; Loehr & Vesper, 2015; Zamm et al., 2015). In contrast, continuous motions do not have such endpoints and noticeable periods without motion. Furthermore, they consist of repetitions of specific events and motions which often result in sinusoidal patterns of the spatial location (Huys et al., 2008). Huys et al., 2008 showed that dynamical systems theory offers the possibility to theoretically distinguish these two motions into distinct classes of discrete and continuous motions. A behavioral study investigated synchronization of rhythmically discrete (finger tapping) and rhythmically continuous (forearm oscillation) movements with a discrete metronome (Torre & Balasubramaniam, 2009). The authors showed that the trajectories of the movements contrasted from one another. Namely, the velocities of movement around the time point of the metronome differed, showing a more equal distribution before and after the metronome time point for the continuous motions and a speeding up in the discrete motion before the time point of the metronome (Torre & Balasubramaniam, 2009). This suggests that the way in which individuals synchronize discrete and continuous motions with external auditory signals differ from one another.

One explanation for differences in how individuals synchronize discrete versus continuous movements is that distinct timing mechanisms may be associated with these two movement types (Torre & Balasubramaniam, 2009). Specifically, for discrete movement an event-based timing is useful, whereas an emergent form of timing constitutes continuous movement. Event-based timing is often linked to the image of an internal timekeeper which keeps track of the time intervals. In contrast, emergent timing is lacking this hierarchical image and is directly associated with continuous motion (Maes et al., 2015; Teki et al., 2011; Torre & Balasubramaniam, 2009; Zelaznik et al., 2005).
Zelic et al., 2019 matched discrete and continuous motions with discrete and continuous stimuli. Discrete pacers yielded better results than continuous ones, suggesting that event-based timing might be easier for synchronization. However, continuous movements (in this case forearm oscillation) resulted in higher accuracy and less variability than discrete motion (Zelic et al., 2019). These results suggest that distinct timing mechanisms underlie discrete versus continuous movements. However, auditory-motor synchronization has primarily been studied in the context of discrete actions, leaving the question open how partners synchronize continuous auditory-motor rhythms. More specifically, to the best knowledge of the author it has not been studied how auditory feedback coupling influences continuous auditory-motor rhythm coordination.

Therefore, this thesis focuses on how humans synchronize continuous movements that produce continuous sounds.

## 1.3. Research Problem

This thesis addresses how partners coordinate the timing of continuous auditory-motor rhythms. Specifically, the thesis addresses how auditory feedback coupling between partners influences patterns of continuous movement synchronization. The thesis conceptually replicates the paradigm used by Konvalinka et al., 2010, but in the context of a continuous rather than a discrete auditory-motor rhythm production task. The research problem at hand, namely how auditory feedback coupling between partners influences patterns of coordination of continuous auditory-motor rhythms, leads to the following research questions that were investigated in this thesis:

**Research Questions**

- ❖Can previously observed patterns of temporal adaptation for partners with discrete auditory-motor rhythms (Konvalinka et al. 2010) be replicated for coordinating continuous auditory-motor rhythms?
- ❖How do auditory feedback conditions influence outcomes?
- ❖How do the patterns of temporal adaptation differ depending on musical education of the partners?

To address these questions, partners synchronously produced continuous auditory-motor rhythms in the following feedback conditions: (i) Uncoupled (partners only hear themselves), (ii) Unidirectional 1 (both partners hear Partner A), (iii) Unidirectional 2 (they hear Partner B) or (iv) Bidirectional (they hear each other). If partners' auditory feedback coupling influences continuous synchronization in the same way as discrete synchronization, then manipulations of feedback coupling should yield similar influences on continuous coordination as previous observed ones in discrete coordination tasks (Konvalinka et al., 2010), meaning that relative phase offsets computed on continuous data should reflect patterns previously observed in discrete data (Konvalinka et al., 2010). In particular, when partners have full auditory feedback (Bidirectional Condition), coordination should be optimal and partners should show the greatest stability of synchrony and optimal synchronization (smallest relative phase offsets). When only one partner can hear the other (Unidirectional conditions), the sign of the relative phase is expected to reflect leader-follower roles, i.e. the partner who can hear the other is expected to anticipate the other (Repp & Su, 2013). No synchronization is expected for the Uncoupled Condition, in which partners can only hear themselves but not their partner, serving as a chance estimate of synchronization between two individuals moving at the same frequency who have no sensory information about one another's actions. Methods for the present study are described below.

# 2. Methods

## 2.1. Experiment Overview

In order to investigate the research questions stated above (see Section 1.3), a sensorimotor synchronization experiment was conducted, in which pairs of individuals synchronously produced circular rhythmic motions on a ring-shaped touch sensor in two separate rooms. Partners' touch locations on the ring-shaped touch sensor were converted into sound, enabling partners to synchronize their movements based on the auditory feedback they received, similarly to music-making in groups. The data reported in the current thesis represent preliminary findings from a larger project of A.Z. and N.S.. The experimental apparatus and procedure are described in the following sections.

## 2.2. Participants

Ten participants (5 pairs) were recruited through the Research Participation System SONA (CEU, 2021) of the Central European University Private University (CEU PU). Participants were included if they passed a prescreening questionnaire in the online SONA system, which confirmed that they were between 18-45 years old, to ensure that participants were adults within the general range of healthy motor-perceptual coordination. In addition, it ensured that they were right-handed, proficient in English and/or German, had normal or corrected-normal vision, normal hearing and did not report on any history of neurological or psychiatric disorders. Demographic information is summarized in Table 2.1 on a participant-level and in Table 2.2 on a pair-level. The tables report on gender and age distribution as well as years of musical training. On a pair-level the difference between partners in years of musical training is reported.

| | **Gender** | | **Age [years]** | **Musical training [years]** |
|---|---|---|---|---|
| **Female** | 7 | **Range** | 18-29 | 0-11 |
| **Male** | 3 | **Mean ± *sd*** | 24.80 ± 3.22 | 3.6 ± 3.89 |

**Table 2.1.:** Demographic information of participants (N = 10), such as gender (female, male), age in years (range, mean and standard deviation (*sd*)) and years of musical training (range, mean and standard deviation (*sd*)).

| | **Pair 1** | **Pair 2** | **Pair 3** | **Pair 4** | **Pair 5** |
|---|---|---|---|---|---|
| **Gender** | Mixed | Female | Mixed | Mixed | Female |
| **Age [years]** | 26 & 26 | 26 & 18 | 29 & 21 | 28 & 25 | 25 & 24 |
| **Musical training [years]** | 8 & 6 | 0 & 2 | 6 & 0 | 2 & 1 | 11 & 0 |
| **Diff. in musical training [years]** | 2 | -2 | 6 | 1 | 11 |

**Table 2.2.:** Demographic information of pairs (n = 5), such as gender distribution within a pair (only female, only male, mixed), age in years and number of years of musical training of each participant within a pair. Further, the difference (diff.) (here Partner A - Partner B) of musical training in years within each pair.

Recruited participants were randomly assigned into pairs (5 pairs in total) and received 15 Euros compensation upon completion of the study, in accordance with standard CEU PU subject recruitment procedures. The study was conducted in the Social Mind and Body Lab (SOMBY-Lab) in Vienna at the CEU PU. Five pilot pairs, which fulfilled the specified criteria, were recruited to fine-tune the experimental set-up and clarify task instructions. Their data was neither analyzed nor included in the present study.

## 2.3. Experimental set-up

The used hardware can be seen in Figure 2.1 (the touch sensor) and Figure 2.2 (the embedded computing platform). Figure 2.3 reports on the timing of the hardware and software system and Figure 2.4 displays the experimental set-up.

### 2.3.1. Hardware & Software

Ring-shaped touch sensors (Augmented Instruments Ltd., 2021e) produced by Cypress Semiconductor Corporation (Cypress Semiconductor Corporation, 2015) were used. They are nominally referred to as “Trill” sensors and were used to capture partners’ finger movements in a sensorimotor synchronization task. Figure 2.1 shows a drawing of the ring-shaped Trill touch sensor: the outer diameter is 52.0 mm and the inner one 28.2 mm (Augmented Instruments Ltd., 2021e). Trill sensors are sensitive to human body capacitance values and detect touch by comparing these capacitance signals to a threshold parameter. The capacitance signal increases when a sensor is touched on a sensory pad and gets registered as a touch when it is above this threshold. A centroid algorithm interpolates the exact touch location by taking not only the largest registered value into account but also the signals from the adjacent sensor pads (Cypress Semiconductor

Corporation, 2015). The Trill sensors transmit touch data to an embedded computing platform via the I2C protocol, which is a common digital communication protocol for transmitting and receiving data between devices, such as sensors, within a shared network (Augmented Instruments Ltd., 2021c). For this, a custom Trill library was used (Augmented Instruments Ltd., 2021a).

**Figure 2.1.:** Drawing of the used ring-shaped Trill touch sensor by Augmented Instruments Ltd. (taken from (Augmented Instruments Ltd., 2021e)).

The embedded computing platform used was the Bela audio computing platform. This hardware and software environment enables one to acquire low-latency audio data. It consists of a BeagleBone Black board, which runs a Debian Linux system with real-time capabilities via a Xenomai real-time kernel extension (Moro et al., 2016). This board is connected to an expansion "cape", responsible to ensure stereo audio in- and output. Further, it provides 8 channels, each consisting of 16-bit ADC and DAC for actuators and sensors (McPherson & Zappi, 2015). The hardware and software of Bela are produced by the Augmented Instruments Laboratory (Augmented Instruments Ltd., 2021a).

A single Bela (version 0.3.8.a (BelaPlatform, 2021a)) was used for data acquisition during the main experimental tasks. Partners had to complete practice sessions and tasks prior to the main experiment individually, which were completed and recorded in parallel to minimize experiment duration. Therefore, two separate Belas were used for the Individual Tasks for data acquisition, one partner's data was recorded on the same Bela used for the main experimental tasks and the other partner's data on a second Bela (version 0.3.1.b (BelaPlatform, 2021a)). For technical reasons, one of the partners used two different but identical Trill sensors in the experiment, one for the Individual and one for the Joint Task of the experiment, which were labeled accordingly. The embedded computing platform

Bela can be seen in Figure 2.2.

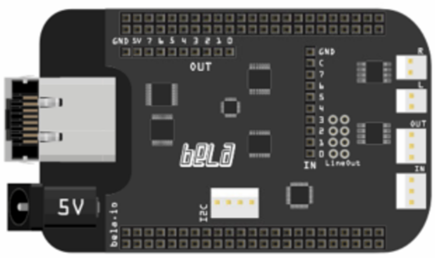


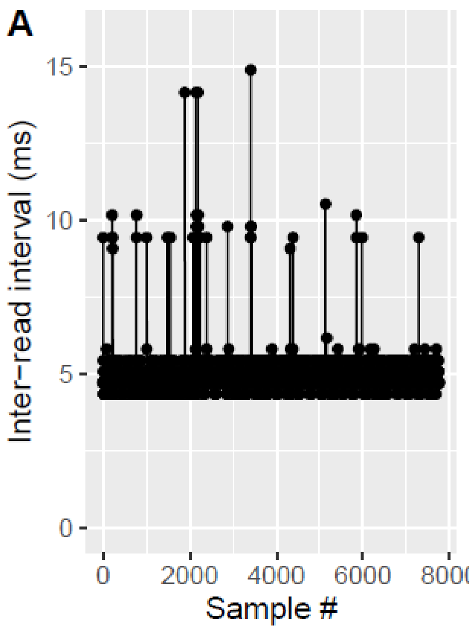

**Figure 2.2.:** The embedded computing platform Bela used for the experiment (taken from (BelaPlatform, 2021b)).

The Bela software provides open-source code for interfacing with the Trill sensors (see https://github.com/BelaPlatform/Bela). The configurations of the Bela were set to a block size of 16 audio frames, a headphone level of -6 dB, a DAC and ADC level of 0 dB and a PGA gain (left and right) of 10 dB (Augmented Instruments Ltd., 2021b). Fixed IP addresses within a shared wireless network were used for the Belas.

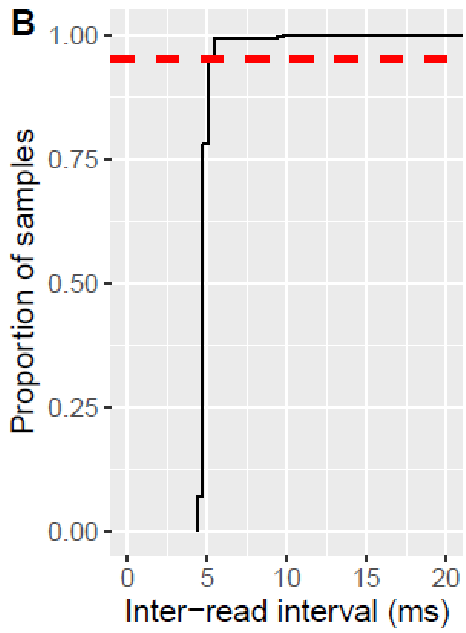
#### 2.3.1.1. Timing of the system

The time between touch location updates required by the system is illustrated in Figure 2.3.

The time points at which touch location values were extracted from the sensors depended on the settings in the software of the Bela (a resolution of 10 bits in the ultra-fast mode with 30 channels). Due to these settings, the scan time was 2.34 ms (30 channels $\cdot$ 78 $\mu$s) (Augmented Instruments Ltd., 2021d). The total delay of the system consisted of this scan time and the amount of time needed for the Bela to communicate with the sensors, which was about 2.5 ms. Therefore, the total delay added up to around 5 ms, which corresponds to the observed time between new touch values registered by the sensor, as can be seen in Figure 2.3.




(a) Inter-read intervals over time.




(b) Proportion of samples below certain inter-read intervals.

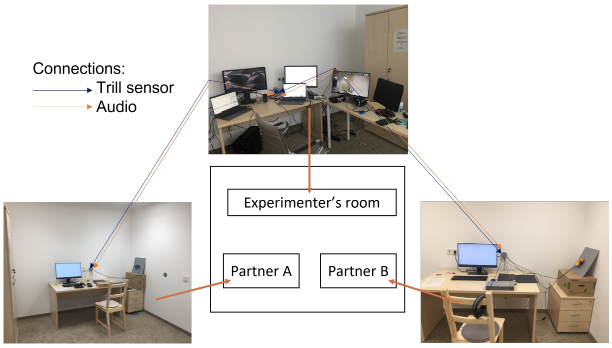
**Figure 2.3.:** The inter-read intervals of new touch locations which are mostly 5 ms. Subfigure (a) shows the inter-read intervals in ms on the y-axis with values at around 5 ms. The x-axis shows the number of samples. (b) reports on the proportion of samples (y-axis) below certain inter-read intervals (x-axis). One can clearly see that most samples are extracted every 5 ms.

This information was used for the pre-processing (see Section 3.2.1).

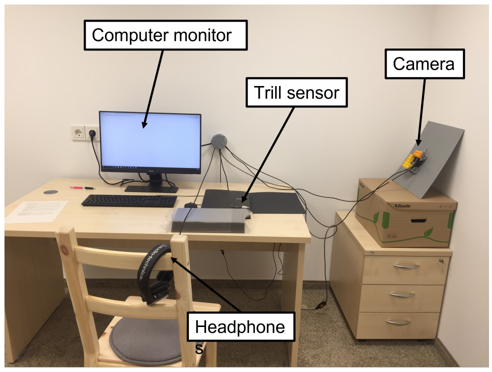
### 2.3.2. Physical configuration

The physical configuration of the experimental set-up is shown in Subfigure 2.4 (a) and (b). Subfigure (c) displays the data flow of the experiment. It was set up with as much distance between partners and experimenters as possible, in order to minimize COVID-19 related risks, while ensuring that partners still experienced social interaction during the experimental tasks. Each partner (namely Partner A and Partner B) completed the tasks in a separate testing room, while two experimenters, one who gave instructions and one who monitored the experimental soft- and hardware, were stationed in another room. The general set-up can be seen in Subfigure 2.4 (a). Each testing room was equipped with a computer monitor, a computer mouse, a computer keyboard, a touch sensor, headphones and a camera, which was used to monitor the partners from the experimenter's room but did not record. The set-up in one of the partners' testing rooms is shown in Subfigure

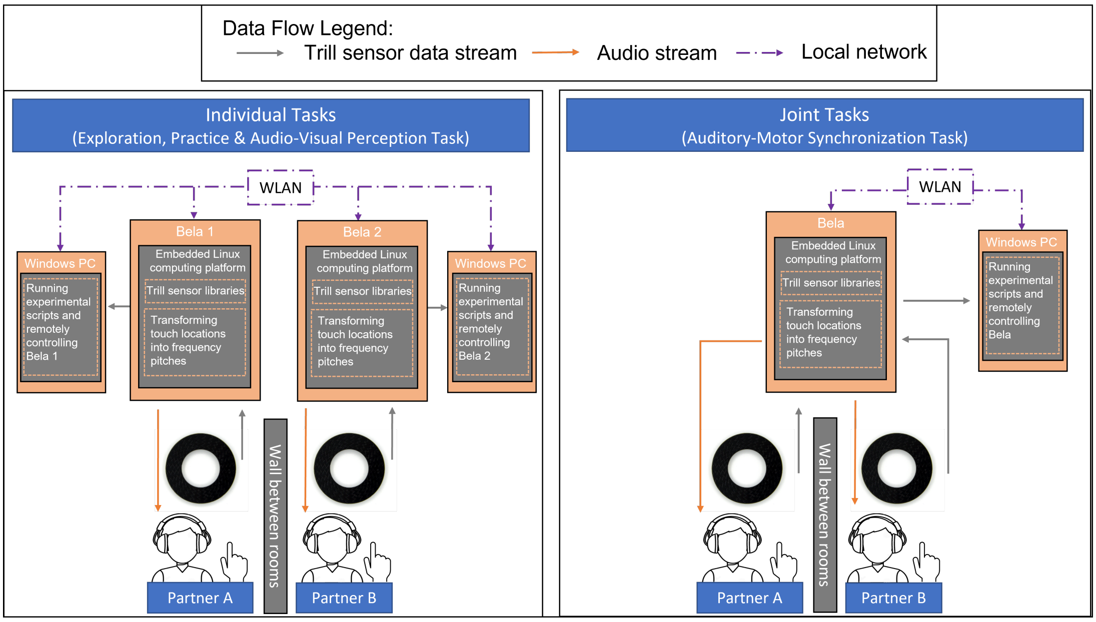
2.4 (b). The partners' touch sensors and audio outputs were connected via cables to a microcontroller located in the experimenter's room, which can be seen in Subfigure 2.4 (a) (see Section 2.3.1 for more details).




**(a)** General set-up of the rooms.




**(b)** Set-up in one of the partner's rooms.




**(c)** Data flow chart.

**Figure 2.4.:** The experimental set-up. (a) shows the general set-up of the rooms and the connection cables. Orange arrows indicate audio cables, blue arrows show connection cables for the Trill sensors. (b) displays one partner's testing room with headphones, camera and a touch sensor, (c) shows the data flow in the experiment.

The data flow within the experiment is displayed in Subfigure 2.4 (c) and summarizes the descriptions given in Subsection 2.3.1 and Subsection 2.3.2.

### 2.3.3. Stimuli

The experiment consisted of an Auditory-Motor Synchronization Task and an Audio-Visual Perceptual Task. The audio-visual and audio-motor stimuli used in each respective task were generated using C++ custom-code developed by A.Z. (thesis co-supervisor, see Acknowledgments), which used Bela's native C++ development libraries. Figure 2.5 shows the labelling convention of the sensor, Figure 2.6 the stimuli sonification process and Figure 2.7 the visual stimulus. The stimuli are described in detail in the following sections.

#### 2.3.3.1. Auditory-Motor Synchronization Task

In the Auditory-Motor Synchronization Task, partners moved their fingers around a ring-shaped sensor. The spatial location of their finger on the ring-shaped sensor was sonified to produce a tone. The sonification process is displayed in Figure 2.6. First, the finger locations on the sensor were identified by the Bela Trill library. The Bela system recorded these as values between 0 and 1, where 0 indicated the 0° point and 1 the 360° point (see Subfigure 2.6 (a)). 3584 unique locations between 0 and 1 were distinguished on the ring-shaped sensor (Giuliomoro, 2021). The Trill library coded 0° per default at the point on the ring which corresponds to the uppermost y-axis position (see the point labeled as 90° in Figure 2.5). Custom code was implemented to rotate the recorded spatial locations by 90° counter-clockwise so that the 0° point aligned with the leftmost side of the circle on the x-axis. The reason for placing 0° at the leftmost side of the circle is that moving the finger in a clockwise manner resulted in an increase in pitch frequency at the start of the motion, followed by a decrease in frequency as the finger moved downwards, corresponding to the trajectory of a typical Frequency Modulated (FM) tone (see below). The used positions of the degree points can be seen in Figure 2.5.

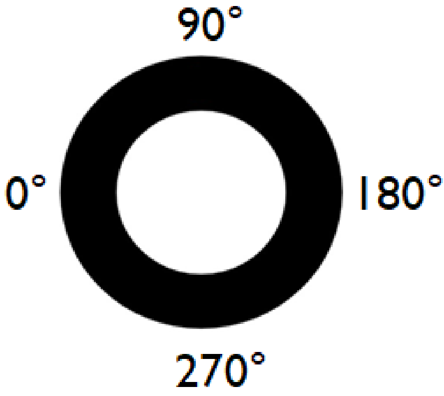
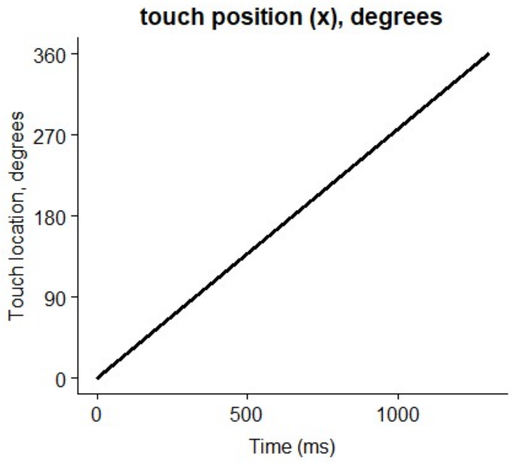
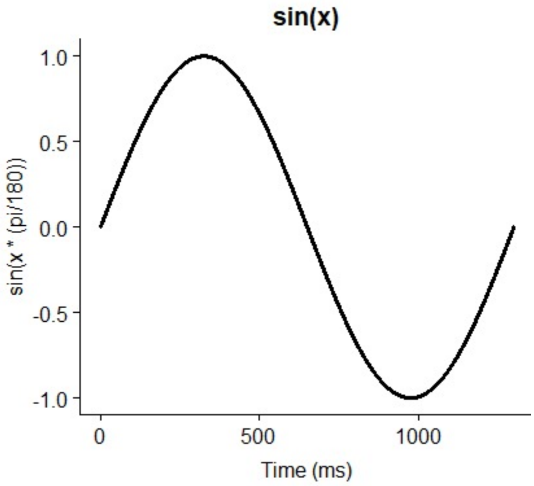



**Figure 2.5.:** The Trill sensor with the labelling convention used. The 0° point was to the left, 90° at the top, 180° to the right and 270° at the bottom of the ring-shaped sensor.

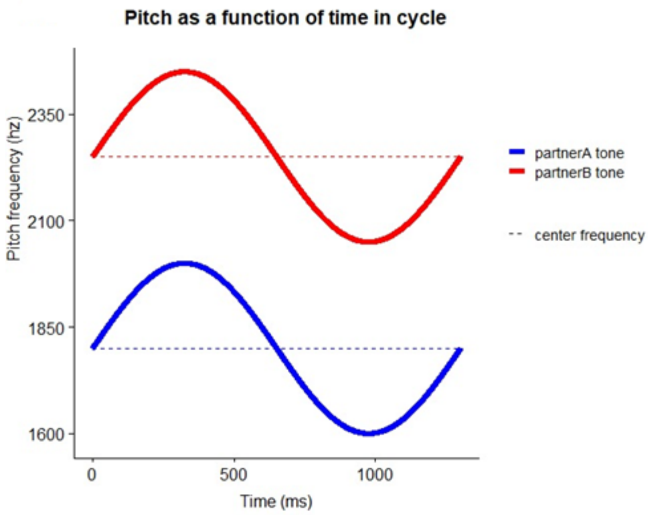
The rotated values were subsequently transformed using a sine function. This way, one full cycle of finger movement around the sensor corresponded to a sinusoid with a range of -1 to 1 (see Subfigure 2.6 (b)). These sinusoidal locations of touch were sonified by linearly mapping the values -1 to 1 onto a pitch frequency range. This way, partners' finger motion produced an FM tone corresponding to their touch position along the y-axis of the sensor. Different pitch frequency ranges were used for each partner in a given pair so that they were able to distinguish their own auditory feedback from their partner's. Specifically, each partner was assigned a unique center frequency, which was the frequency produced when their finger was at the 0° location on the sensor, and in the range of normal hearing. The center frequencies were selected to be a consonant interval, namely a Major 3rd, apart and corresponded to 1800 Hz (always assigned to Partner A, see Subfigure 2.4 (a)) and 2250 Hz (always assigned to Partner B, see Subfigure 2.4 (a)) respectively. The range of pitches produced by moving form 0° to 360° on the ring-shaped sensor corresponded to +/- 200 Hz around each partner's center frequency, where the highest pitch was produced at 90° and the lowest pitch at 270°. When the partner started at the 0° point and moved their finger towards the 90° point, the pitch frequency increased up to 2000/2450 Hz. When they moved their finger down, the pitch frequency decreased, resulting in the lowest value at 270° which corresponded to a pitch frequency of 1600/2050 Hz respectively. Moving the finger to the 360° point, which also corresponded to the 0° point, the pitch frequency increased again to 1800/2250 Hz. Subfigure 2.4 (c) displays the two different pitch frequency ranges used. This pitch range was selected based on results from internal pilots to ensure that partners could clearly hear a continuously changing sound corresponding to their finger movements.




**(a)** Touch positions between 0° and 360°.




**(b)** Sine values between -1 and 1.




**(c)** Two different frequency ranges.

**Figure 2.6.:** For all figures the x-axis reports time in ms. Subfigure (a) shows the original touch locations on the ring-shaped Trill sensor between 0° and 360°. In (b) these touch locations were converted to sine values between -1 and 1. These values correspond to the touch location along the y-axis of the sensor, where -1 is the lowest point (270°) and 1 the highest point (90°). Subfigure (c) shows the values after they had been linearly mapped to two different frequency ranges. The range for both was ± 200 Hz. The red function shows the higher frequency centered at 2250 Hz (indicated by the dashed red line), the blue one the lower frequency centered at 1800 Hz (indicated by the dashed blue line). It should be noted that these subfigures show simulated data and do not reflect the actual movement of a partner on the sensor.

At the start of each experiment trial, partners heard a pacing cue that simulated the sound of 8 regular cycles of finger movement around the sensor. The pacing signal was constructed through the same process as the sounds produced by touching the sensor, using simulated touch locations as input. Specifically, the pacing signal tempo was first determined by the most comfortable tempo in Beats Per Minute (BPM), which was based on internal pilots, that divided evenly into the audio sampling rate of 44.1 kHz. The selected tempo from these internal pilots was 46.14258 BPM, corresponding to an Inter-Beat Interval (IBI) of 1300.317 milliseconds (ms), as described in Equation (2.1). The resulting number of samples per cycle with this IBI is calculated in Equation (2.2) by taking into account the audio sampling rate of 44.1 kHz.

$$46.14258\,\frac{\text{beats}}{\text{min}} \implies \frac{1}{46.14258}\,\frac{60 \cdot \text{s}}{\text{beat}} \approx 1.300317\,\frac{\text{s}}{\text{beat}} = 1300.317\,\frac{\text{ms}}{\text{beat}}, \tag{2.1}$$

$$44.1\,\text{kHz} \cdot 1300.317\,\text{ms} = 44100\,\frac{\text{samples}}{\text{s}} \cdot 1.300317\,\text{s} \approx 57344\,\frac{\text{samples}}{\text{cycle}}. \tag{2.2}$$

The IBI corresponds to the inter-cycle interval (i.e. one cycle of the finger around the sensor). To simulate the stimulus tone produced by moving one's finger around the sensor at the selected inter-cycle interval, the total number of audio samples was divided by the number of unique finger positions on the sensor (see Equation (2.3)). The choice of tempo resulted in sustaining each frequency in the FM tone for 16 audio samples (see Equation (2.3)).

$$\frac{57344\,\text{samples}}{3584\,\text{unique positions}} = 16\,\text{audio samples}. \tag{2.3}$$

#### 2.3.3.2. Audio-Visual Task

For the Audio-Visual Perception Task, partners first heard an auditory signal and then saw a visual stimulus (see for details Subsection 2.4.2). The auditory stimulus consisted of four full cycles around the circle and a partial fifth cycle. Details about the procedure can be found in Subsection 2.4.2. For the visual stimulus an image of a touch sensor, see Figure 2.7, was used (see for details Subsection 2.4.2).

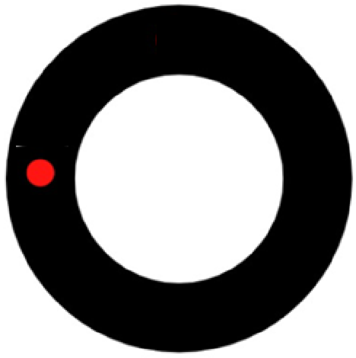


**Figure 2.7.:** An example of the visual stimulus used in the Audio-Visual Perception Task.

### 2.3.4. Experimental Design

The experiment comprised a within-subject design, with a single independent variable, Auditory Feedback. Auditory feedback was manipulated across 4 conditions: the (i) Uncoupled Condition, the (ii) Unidirectional 1 Condition, the (iii) Unidirectional 2 Condition and the (iv) Bidirectional Condition. The conditions are explained in Subsection 2.4.3. The conditions were counterbalanced between pairs and there were five trials per condition.

## 2.4. Tasks & procedure

Participants completed the experiment in pairs. Each of the partners was placed in a separate room that was equipped with a computer monitor, a computer mouse, a computer keyboard, a touch sensor, headphones (model ATH-M50X, Audio-Technica) and a camera (see Subfigure 2.4 (b) in Subsection 2.3.2). Partners were informed that the camera was used only so that the experimenter could monitor and communicate with them, not for recording. The partners read and signed an information sheet and consent form, a demographic questionnaire (see Appendix A), a reimbursement sheet and a COVID-19 hygiene protocol.

The experiment comprised 4 phases: (1) Exploration, (2) Practice, (3) Audio-Visual Perception Task, (4) Auditory-Motor Synchronization Task. Figure 2.11 shows an overview of the 4 conditions of phase (4). Partners completed the first three phases individually and the fourth task, namely the Joint Task/Auditory-Motor Synchronization Task, together.

Each phase is described in detail below.

### 2.4.1. Exploration and Practice Task

Figure 2.8 shows one of the sensors used for the tasks. Partners first completed an Exploration Task. In this phase, each partner was shown the touch sensor and the sound-space mapping (i.e. how finger positions map to tones) was explained as follows (instructions for Partner A):

> *"This touch sensor will produce a pitch based on where your finger is on the circle. When your finger is at 0°, you will produce a pitch frequency of 1800 Hz. As you move your finger towards 90°, the pitch will increase until you reach 2000 Hz, then as you move your finger down to 180°, the pitch will decrease back to 1800 Hz. As you move towards 270°, the pitch will decrease until you reach 1600 Hz, at which point it will increase again as you move towards 360°, at which point it will return to 1800 Hz."*

Depending in which room the partner was seated, they were informed that a touch at the 0°-point (on the left of the sensor) would result in a pitch frequency of 1800 Hz (Partner A) or 2250 Hz (Partner B). The partners were asked to always use their right index fingers and move their fingers in a continuous manner in a clockwise direction around the sensors without lifting up their fingers within one trial.

In the Exploration Task, partners had the opportunity to explore the sensors with their right index fingers and learn the sound-space mapping. They were able to see the frequency (in Hertz, Hz) of the tone that they produced as an output on the computer screen that was placed next to them. Partners were asked to first carefully identify where they needed to place their index finger in order to produce the frequency associated with the 0°-point (i.e. 1800 Hz or 2250 Hz) on the circle. A sticker was placed next to the 0°-point for the Individual Tasks as a guide. Partners were informed that they should try to memorize the 0°-point, as it would be removed for the Joint Task. A picture of one of the used sensors can be found in Figure 2.8. Partners should adjust their fingers around this sticker until they produced precisely the center frequency.

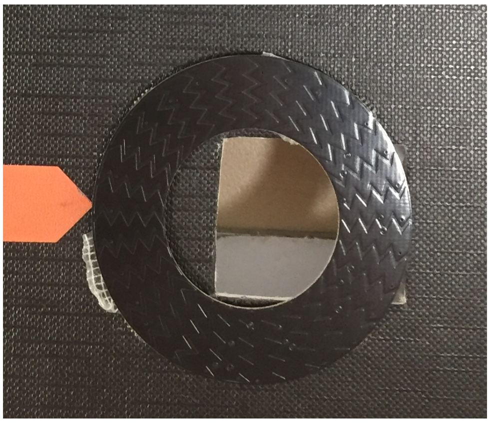


**Figure 2.8.:** One of the sensors used for the production tasks with the orange sticker indicating the 0°-point.

After this, partners should start to move their finger in a continuous manner in a clockwise direction around the sensor. They had 5 trials to explore the sensor, each one minute long. If partners felt that they needed extra trials to familiarize themselves with the sound-space mapping, then additional trials were be added.

Next, the Practice Task was initialized. Here, partners learned to move their finger rhythmically around the sensor at a rate that was first indicated through a pacing signal (see Subsection 2.3.3.1). In this phase, the computer screen was turned off so that partners did not have a screen print-out of the frequencies associated with the sounds that they produced. Each trial started with a pacing signal that sounded like 8 cycles around the circle, lasting 8 x 1.3 s (approximately 10 s in total). Partners were asked to start to move their finger around the sensor at the indicated rate as soon as the pacing signal stopped. They were informed that their goal was to produce a steady rate of movement at the cued rate and to keep moving their finger continuously in a clockwise direction until the sound stopped. The production phase per trial lasted approximately 30 s, resulting in a length of 40 s per trial. Three trials were performed with the possibility to add more if needed.

### 2.4.2. Audio-Visual Perception Task

The third phase of the experiment consisted of the Audio-Visual Perception Task which was also performed individually and in parallel. Figure 2.9 shows the procedure of the Audio-Visual Perception Task. The goal of this phase was to test how well partners had learned the sound-space mapping, in order to make sure that they understood the connection between a location and an FM tone, which was important for a successful completion of the Joint Task. Partners were informed that their task was to indicate which of two sensors showed a red dot at the correct location where the auditory stimulus stopped (instructions for Partner A):

> *"In this task, you will hear a sound that sounds like the pacing signal in the previous task: this is the sound that would be produced if you moved your finger continuously around the circle at a steady rate, starting from 0°. Sometimes the sound will be in the octave you practiced with, and sometimes it will be one octave higher.*
> *While hearing this sound, you will see an image of the circle on the screen and you will imagine moving your finger to produce the sound. It is important that you only imagine this movement, please, do not move your finger above or on the screen or the sensor.*
> *You will hear several full trips around the circle, and then in the final trip the sound will stop somewhere mid-cycle.*
> *When the tone stops, you will see two sensors appear on the screen; one will show a red dot at the correct location on the circle where the sound stopped, and one will show a red dot at an incorrect location.*
> *Your task is to indicate which dot indicates the correct location where your finger would have been when the sound stopped."*

Figure 2.9 shows the timeline and steps of the Audio-Visual Perception Task. The auditory stimulus could be in the octave partners had practiced in (so the octave associated with their sensor), or in the octave associated with their partner (so higher or lower than theirs). The reason for this was so that they would get accustomed to both frequency ranges that they would encounter in the Joint Task. In this task, the computer screen, the computer mouse and the computer keyboard were used. In addition, the touch sensor was covered with a black paper so that they were not able to use the sensor as a visual aid.

Partners were not informed about the exact amount of full cycles that they would hear, but they heard a fixed number of full and partial cycles (4 full cycles, followed by 1 partial cycle, starting at the 0°-point). In the fifth and final cycle, the sound stopped somewhere mid-cycle. Partners then saw two sensors appearing on the screen next to each other. One of them showed a red dot at the location associated with the frequency that occurred when the auditory stimulus stopped (correct location), one showed the red dot at a different (incorrect) location. Partners were asked to indicate the correct location.

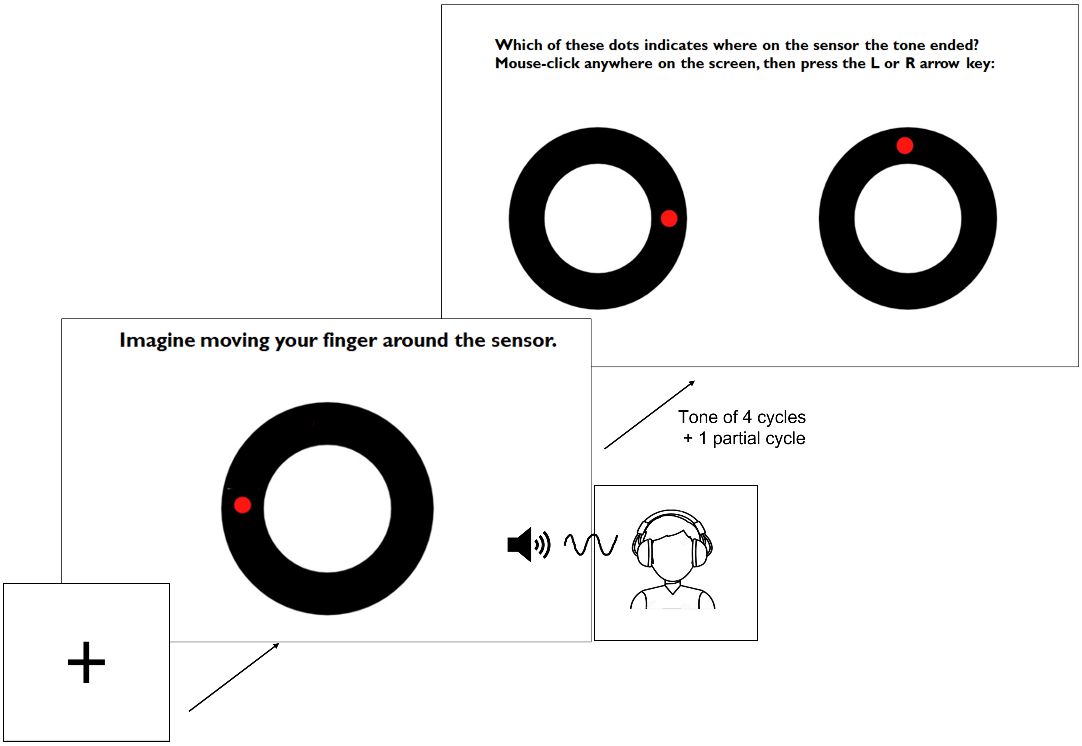
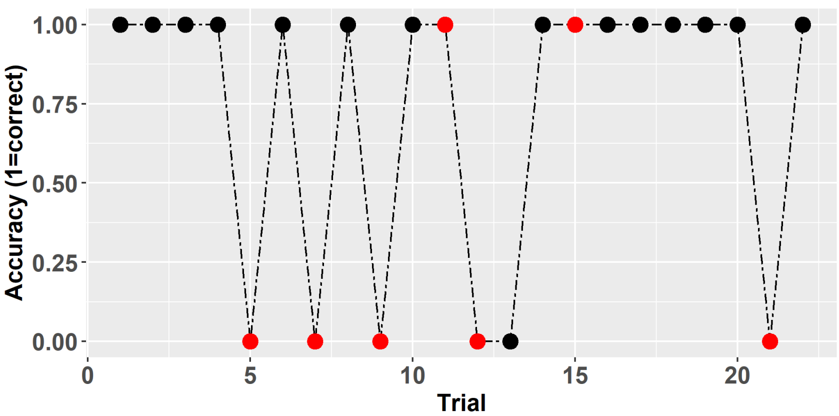


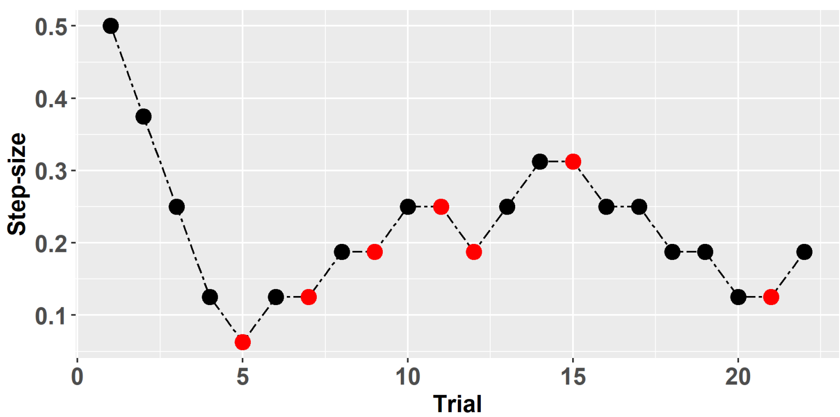

**Figure 2.9.:** The timeline of the Audio-Visual Perception Task, showing the fixation cross and the two different screens the partners saw. While they saw an image of a touch sensor they heard an auditory stimulus of 4 full cycles and one partial cycle before the sound stopped. On the second screen they were prompted with two sensors, each showing a red dot. One of them on the location associated with the frequency of the tone (correct) and one at a different (incorrect) location. Partners had to choose the sensor that correctly displayed the phase on the circle where the tone ended.

At least two practice trials were performed so that partners could get familiarized with the task before the main task was initialized.

The task followed a "1 Up - 2 Down" staircase procedure (Cornsweet, 1962). First, the two answer options were a half circle apart from one another. If the partner gave a correct answer, the difference between the correct and incorrect location was reduced by two step sizes on the sensor, where one step size corresponds to 1/16-th of the circle, namely 0.0625. As soon as the partner gave a wrong answer for the first time, the difference between the correct and incorrect location was again increased by one step size. From now onwards it was necessary that the partner gave two correct answers in a row before the difference between the correct and incorrect location was reduced, now only by one step size. This procedure is shown in Figure 2.10. Subfigure (a) shows the accuracy score of one partner which switches between 0 (incorrect) and 1 (correct). Subfigure (b) shows the change in the difference between the location of the two answer options. Whenever the difference between the two answer options is either increased after it had been decreased before or decreased after it had been increased before, one speaks of a reversal. These reversals are indicated in red in Figure 2.10. Seven reversals or eight wrong answers in a row were needed to finish the task, therefore the time needed for this phase of the experiment varied between partners.




**(a)** Accuracy scores.




**(b)** Difference in location between the correct and incorrect locations.

**Figure 2.10.:** Figures reporting on the results of the Audio-Visual Perception Task for one partner. For all figures the x-axis reports the number of trails. Subfigure (a) shows the accuracy score of 0 (incorrect) and 1 (correct) of the answer given. In (b) the y-axis shows the difference in location between the correct and incorrect location on the sensor. The higher the value on the y-axis the further apart the two answers are. Red dots indicate reversals.

### 2.4.3. Auditory-Motor Synchronization Task

The Auditory-Motor Synchronization Task was the Joint Phase of the experiment and consisted of 4 different Auditory Conditions: (i) Uncoupled, (ii) Unidirectional 1, (iii)

Unidirectional 2, (iv) Bidirectional. The 4 different conditions are reported in Figure 2.11, where Subfigure (a) and (b) show the same information in two different ways. In the Uncoupled Condition, partners were only able to hear themselves and not their partner (purple condition in Figure 2.11). This condition was run at the beginning of the Joint Task and at the very end. The order of the other conditions were counterbalanced across pairs. In the Unidirectional 1 Condition (blue), both partners heard only Partner A. For Unidirectional 2 (blue), both partners only heard Partner B. In the Bidirectional Condition (shown in orange in Figure 2.11), both partners were only able to hear their partner but not themselves. Before each condition, the partners were informed who they would be hearing (i.e. themselves or their partner) but did not know who their partner would be hearing.

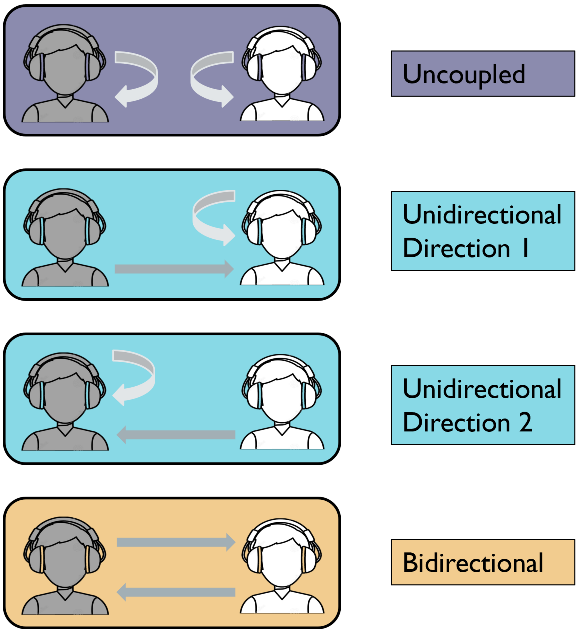
For each trial, partners heard a pacing signal that sounded like 8 cycles around the circle. Partners were aware that the pacing signal had the same rate for both of them and they were instructed as follows:

> *"When the pacing signal stops, you can start moving your finger around the sensor at the rate indicated by the signal. On some trials you will hear yourself, and on some trials you will hear your partner.*
> *On trials where you hear only yourself, your goal is to produce a steady rate of movement at the cued rate. On trials where you hear only your partner, your goal is to produce a steady rate of movement at the cued rate, and also to synchronize your movements with the sound of your partner's movements. Keep moving your finger until the sound stops.*
> *As soon as the pacing signal stops, start moving your finger around the sensor. It is important that you keep moving your finger without lifting it up until the sound stops completely."*

Five trials were run for each condition, each lasting 40 s of which approximately 10 s were the pacing signal. For this phase of the experiment, the stickers indicating the 0°-point were removed so that there was no visual anchoring point for partners' movements during rhythm production.

| | **A Hears oneself** | **A Hears partner** |
|---|---|---|
| **B Hears oneself** | Uncoupled | Unidirectional coupling |
| **B Hears partner** | Unidirectional coupling | Bidirectional coupling |

**(a)** The four different Coupling Conditions in a table.




**(b)** The four different Coupling Conditions shown through directionality.

**Figure 2.11.:** The four different conditions in the Joint Task, conceptually replicating Konvalinka et al., 2010. Subfigure (a) shows the information in a table format, while subfigure (b) explains the conditions through directionality with arrows. In the Uncoupled (purple) Condition, partners could only hear themselves. In the Unidirectional 1 Condition (blue), Partner A heard themselves and Partner B also heard Partner A. In the Unidirectional 2 Condition (blue) both heard Partner B but not Partner A. In the Bidirectional Condition (orange) both partners could only hear their partner but not themselves.

# 3. Data processing & analysis

All data processing, data analyses and the plots were performed and created with R (R Core Team, 2020) (Version 3.6.0 and Version 4.1.1.) in RStudio (RStudio Team, 2020) (Version 1.3.959 and Version 1.4.1717).

## 3.1. Audio-Visual Perception Task

A qualitative analysis of the Audio-Visual Task was performed. The relative accuracy and perceptual threshold per partner was determined.
The relative accuracy per partner per pair was calculated by dividing the number of correct answers by the total number of prompts.
The perceptual threshold was calculated as the median option location distance of the last 6 reversals of the Audio-Visual Task, which corresponds to the median across all option location distance as the task ended with the $7^{th}$ reversal. The term option location distance indicates the difference between the correct and incorrect location shown as options in the task (see for more details on the task procedure Subsection 2.4.2).

## 3.2. Auditory-Motor Synchronization Task

### 3.2.1. Data pre-processing

The data extracted by the custom software compromised of each partner's touch location at every sample, the corresponding sine value and the pitch frequency. Samples for each partner were only recorded when a new touch location value was detected, indicating an update in the person's finger location. This procedure resulted in an irregular sampling rate, with a break of approximately 5 ms between samples (see Subsection 2.3.1.1 for more details). Therefore, data pre-processing needed to be performed so that partners' samples were temporally aligned.

Specifically, a time vector of 1 ms resolution was constructed by cubic spline interpolation of the sine and frequency values (R Core Team and contributors worldwide, 2021a). Then, the data was down-sampled to 5 ms, based on the most common inter-read interval across partners (see Subsection 2.3.1.1 and Figure 2.3). All data segments where partners were not both continuously moving their fingers around the sensor were marked for rejection from subsequent analyses (i.e. gaps larger than 20% of the prescribed tempo and the beginning and ends of trials where one partner often started/stopped before the other). In addition, it was checked that whenever one partner had no read-out values at some other location in a trial, the read-out value of the other person was also removed from analysis.

Lastly, the data was smoothed with a moving average window of ± 50 ms. This corresponds to a low-pass filter of 10 Hz, which is outside of the frequency range of interest while still large enough to filter out undesired noise in the data.

#### 3.2.1.1. Dependent Variables

**Continuous relative phase.**

Continuous relative phase was measured to assess the absolute temporal offset between partners' finger movements over time. To compute relative phase, first each partner's instantaneous phase was computed with the `atan2` function (R-package `raster`) (Hijmans, 2021). This function takes the sinusoidal function of touch locations as an input and computes the instantaneous phase based on the information in which part of the function the touch location is located (e.g. maps a sine value of 0 to 0° or 180°, based on whether the value is located on an upward or downward flank of the sine function). As this function reports values between 180° and 360° as negative angles (e.g. -90° for 270°), these negative phases were mapped onto the corresponding positive angles (so 270° for -90°), resulting in possible values between 0° to 360°.

The relative phase between partners within one pair was subsequently calculated as the minimal distance between the partners' phase angles. For this, Equation (3.1) was used, where $l$ is the larger phase value of the partners' phase values and $s$ the smaller phase value. The min()-function returns the value of the two inputs that has the smaller value. This was done in contrast to always simply taking the difference between partners' phase angles (i.e. Partner A - Partner B), in order to account for situations in which one partner just crossed the 0°-point, while the other has not. For example, Partner A's finger could be at 4° (just passed the 0°-point) and Partner B at 356° (shortly before the 0°-point). Simply subtracting partners' phase angles would yield 352°, even though they are only 8° apart in space. This is especially relevant from the perspective of the task, in which the difference in sound change partners hear corresponds to 8° and not 352°. Calculating the relative phase with Formula (**??**) made sure that the true spatial proximity of partners' instantaneous phases was reflected in relative phase measures.

$$\text{relative phase} = \min(l - s, 360 + s - l). \tag{3.1}$$

This method of computing relative phase constrained the range of possible values between 0° and 180°. Relative phase was averaged across trials within each condition. For all analyses, the Uncoupled Condition before and after the coupled conditions were collapsed. Small relative phase values indicate high synchrony and large relative phase values indicate low synchrony between partners.

The mean relative phase values on a pair-level for each Auditory Feedback Condition were calculated by averaging across trials per condition, meaning that all analyses were run on the mean relative phase values averaged across trials. Again, small mean relative phase values indicate high synchrony and large mean relative phase values indicate low synchrony between partners.

**Continuous relative phase variability.**

Continuous relative phase standard deviation was calculated to assess the variability of relative phase per feedback condition. To compute relative phase variability, the standard deviation within one pair and feedback condition was calculated per trial with the R-package `circular` (Lund et al., 2017). Then the mean of the standard deviations across trials was taken, resulting in the variability of relative phase (i.e. the standard deviation of relative phase) per pair and feedback condition. Visually, the variability is often reported through the length of the arrows which show the mean relative phase location in a circle. Therefore, the standard deviation is also called "resultant length".

**Cross-correlations.**

In order to be able to report on the directionality of the relative phase – which itself is an unsigned measure – cross-correlations of the movements of the participants' instantaneous phase angles were performed. This way, the effects of Auditory Feedback Condition on leader-follower behavior was analyzed. Figure 3.1 and Figure 3.2 help to visualize the cross-correlation of two periodic functions.

A cross-correlation takes two data sets as inputs and lags one of them in time relative to the other. The correlation between the shifted and non-shifted data vector is then calculated and reported. This is done for several lags (i.e. different sizes of shifts), resulting in a distribution showing the correlation at the different lags. This can then be used to determine the lag for which the correlation is highest.

Finger motion around a ring-shaped sensor results in a sinusoidal, and therefore periodic, movement. The cross-correlation of two periodic functions is periodic itself. This becomes clear with Figure 3.1 which shows a sine function (orange) and the same sine function shifted by 180° (turquoise). A shift of 180° will result in a perfect negative correlation. Therefore, shifts from 0° to 360° will naturally result in a periodicity of the cross-correlation function of two data sets sampled from sine curves between negative and positive correlations.

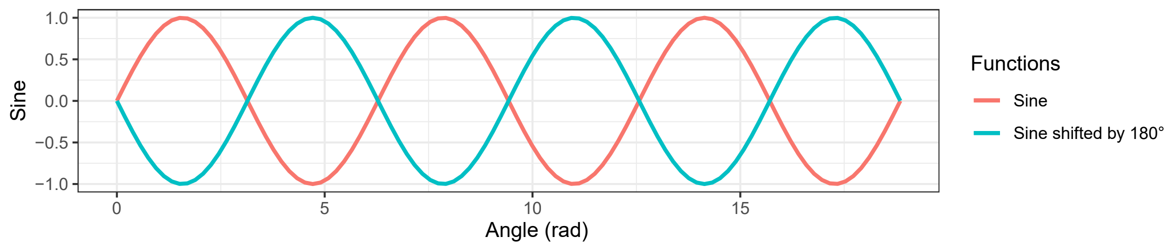
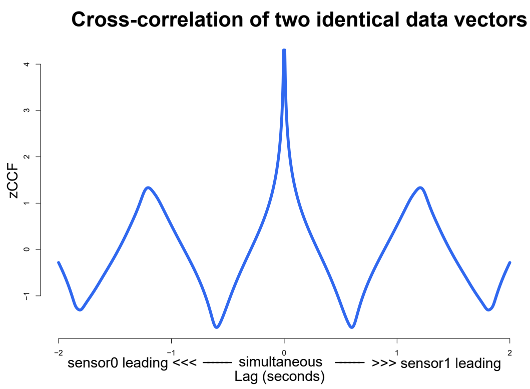
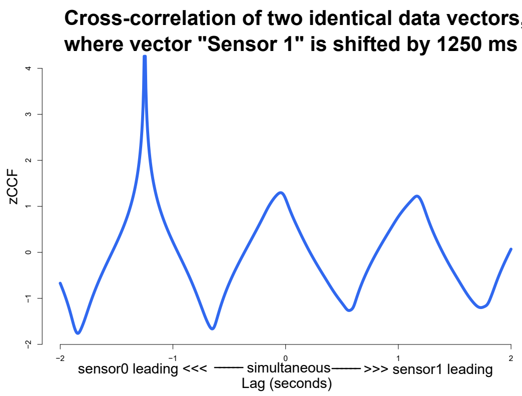



**Figure 3.1.:** Two sine functions to illustrate the periodicity of their cross-correlation function. A sine function (orange) is shown together with the same function shifted by 180° (turquoise). The correlation between them would be perfectly negative. Shifts from 0° to 360° will therefore result in a periodicity of the cross-correlation function.

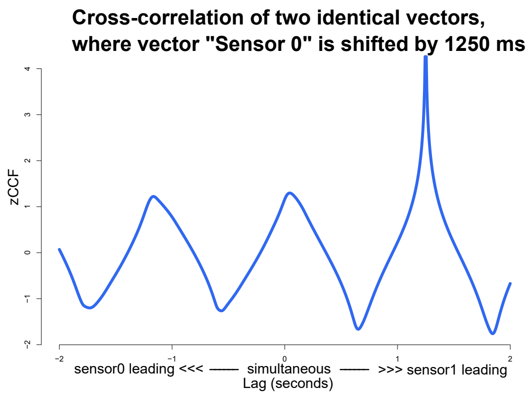
Figure 3.2 qualitatively shows the process of a cross-correlation. For illustration purposes, a data set of values has been duplicated, meaning that the data for "Sensor 0" in this example is identical to the data of "Sensor 1". Subfigure (a) displays a cross-correlation between the identical data sets (i.e. an auto-correlation). The peak cross-correlation can be seen at lag 0. If the data set of "Sensor 1" is lagged by 1250 ms, corresponding to one cycle on the ring-shaped sensor, the peak of the cross-correlation shifts to -1.25 s, as "Sensor 0" is now leading (see Subfigure 3.2 (b)). A shift of the same size of the data set of "Sensor 0" results in a peak of the cross-correlation at around +1.25 s, showing that "Sensor 1" is leading (see Subfigure 3.2 (c)).
This example shows how it is possible to extract any existing lags between the two sensor data series (i.e. lags between partners' movements). The two data sets used in the example above are identical and only lagged to one another. However, the data from the partners in the experiment are not identical, which is why the cross-correlations are not as clear as in the example above.




**(a)** Auto-correlation.




**(b)** Cross-correlation, where "Sensor 1" is shifted by 1250 ms.




**(c)** Cross-correlation, where "Sensor 0" is shifted by 1250 ms.

**Figure 3.2.:** A qualitative explanation of cross-correlation between two identical vectors (labeled "Sensor 0" and "Sensor 1"). (a) shows a cross-correlation between these identical vectors with a peak at 0 s lag. (b) displays a cross-correlation with the data set of "Sensor 1" shifted by 1250 ms, which corresponds to one full cycle on the ring-shaped sensor, resulting in a peak at a lag of -1.25 s. "Sensor 0" is leading in this scenario. (c) shows the cross-correlation with the data set of "Sensor 0" shifted by 1250 ms and a peak at 1.25 s, resulting in "Sensor 1" leading.

Both non-windowed and windowed cross-correlations were computed. The former uses the full data set for computing the cross-correlation. It therefore gives an impression of the overall leader-follower dynamics. The latter uses predefined window sizes of data for computing cross-correlation and is therefore more sensitive to fluctuations in leader-follower dynamics within a time series.

**Non-windowed cross-correlations.** The non-windowed cross-correlation was calculated with the `ccf` function of the package `stats` in R (R Core Team and contributors world-

wide, 2021b).
The calculation of the non-windowed cross-correlation was performed in such a way that a negative value of the resulting maximum peak lag indicates that Sensor 0 (Partner A) is leading Sensor 1 (Partner B) (The Pennsylvania State University, 2021).

**Peak lags.** Partners' cross-correlations were computed, and the lag associated with the peak cross-correlation was identified for each Pair and Condition per trial. The mean across trials of these peak lags was taken.

**Variability of peak lag.** The variability of peak lags, namely the standard deviation of mean peak lags across trials per Pair and Condition, was calculated.

**Windowed cross-correlations.** The windowed cross-correlations were calculated with lags between -0.5 and 0.5 s and window sizes of 2 s with 0.25 s overlap between neighboring windows. For this the R-package `rMEA` (Kleinbub & Ramseyer, 2020) was used. It should be noted that the package is typically used for motion energy analysis and possibly used for the first time for non-energy motion.

**Peak lags.** Partners' windowed lag cross-correlations between lags -0.5 and 0.5 s (corresponding to approximately half a cycle, which is the approximate maximum offset in the Uncoupled Condition) were computed, and the lag associated with the peak cross-correlation was identified for each Pair and Condition and trial. As the windowed cross-correlation results in more outputs per trial than the non-windowed cross-correlation, the median per trial was taken and then averaged across trials.

**Variability of peak lags.** The variability of peak lags was also assessed, defined as the standard deviation of peak lags per trial within each Pair and Condition. These standard deviations were then averaged across trials.

### 3.2.2. Analysis

#### 3.2.2.1. Tests for effects of Auditory Feedback Condition on mean relative phase angle

**Mean relative phase values.** The effects of Auditory Feedback Condition on mean relative phase angles were estimated by calculating a one-way repeated measures ANOVA on relative phase for each Pair and Condition. For this, the R-library *afex* was used

(Singmann et al., 2021). Due to the small sample size (n = 5 pairs) and the fact that values were restricted between 0° and 180° (i.e. making an assumption of a normal distribution questionable), a non-parametric equivalent to the repeated measures ANOVA was additionally computed on relative phase for each Pair and Condition, namely the Friedman's Test (DataNovia, 2021). The R-package `rstatix` was used for this (Kassambara, 2021).

**Comparison with mean of chance distributions.** For comparison with the observed mean phase angles, a chance distribution of mean phase angles was calculated. For this, 1000 uniform phase distributions were simulated with the `runif` function (R-package `stats`) (R Core Team and contributors worldwide, 2021b) and the mean was calculated for each. The grand average of these means was estimated for comparison with the observed grand mean averages.

#### 3.2.2.2. Tests for effects of Auditory Feedback Condition on variability of relative phase angles

The effects of Auditory Feedback on variability of relative phase angles were assessed by computing a one-way repeated measures ANOVA on mean resultant length (i.e. standard deviation) for each Pair and Condition. For this, the R-library *afex* was used (Singmann et al., 2021).

#### 3.2.2.3. Tests for effects of Auditory Feedback Condition on leader-follower behavior

The analyses described below were separately performed for the results of the non-windowed and windowed cross-correlations.

**Peak lags.** The mean lags were submitted to a one-way repeated measures ANOVA (R-package `afex` (Singmann et al., 2021)). Because of the small sample size, a Friedman's Test (non-parametric equivalent to the repeated measures ANOVA) was also computed on the peak lags for each Pair and Condition to ensure soundness of the analyses (DataNovia, 2021). The R-package `rstatix` was used (Kassambara, 2021).

**Variability of peak lag.** Standard deviation values were submitted to a one-way repeated measures ANOVA (R-package `afex` (Singmann et al., 2021)). In addition, the non-parametric Friedman's Test was computed on variability of peak lag per Pair and Condition (DataNovia, 2021) (R-package `rstatix` (Kassambara, 2021)).

# 4. Results

## 4.1. Audio-Visual Perception Task

Table 4.1 reports on the results of the Audio-Visual Perception Task, which was performed individually by each partner. The relative accuracy indicates the percentage of correct answers given. The perceptual threshold is the median option location distance of the last 6 reversals of the Audio-Visual Task.

| Pair | Partner | Relative accuracy | Perceptual threshold |
|---|---|---|---|
| **Pair 1** | A | 72.73 % | 0.1875 |
| | B | 80.95 % | 0.2500 |
| **Pair 2** | A | 75.00 % | 0.2500 |
| | B | 75.00 % | 0.1250 |
| **Pair 3** | A | 84.21 % | 0.1250 |
| | B | 71.43 % | 0.2500 |
| **Pair 4** | A | 50.00 % | 0.4375 |
| | B | 55.56 % | 0.4375 |
| **Pair 5** | A | 68.42 % | 0.1875 |
| | B | 68.42 % | 0.1875 |

**Table 4.1.:** Accuracy scores of the Audio-Visual Task of each Partner per Pair as well as each partner's perceptual threshold, which corresponds to the median option location distance across the last 6 reversals.

The presented results are purely descriptive at the moment. A higher sample size would be needed for further analyses (e.g. how the perceptual threshold and relative accuracy correlate with the task performance of each partner and pair).

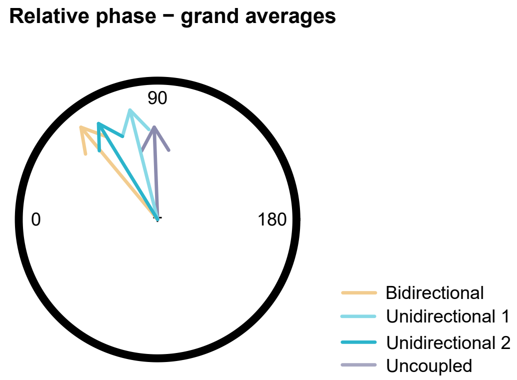
## 4.2. Auditory-Motor Synchronization Task

The used significance level was 0.05. P-values between 0.05 and 0.1 are reported as marginal effects.

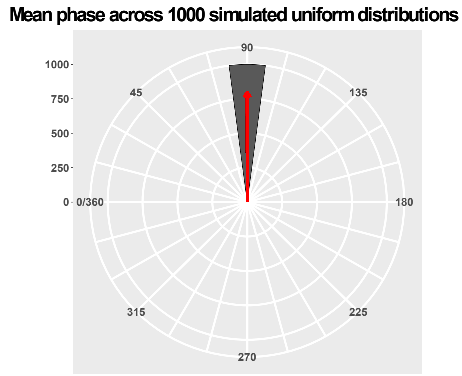
### 4.2.1. Tests for effects of Auditory Feedback Condition on mean relative phase angle

Figure 4.1 shows mean relative phase values for each Auditory Feedback Condition. The results are presented in a circle, arrows represent the mean relative phase and the length of the arrows shows the standard deviation, where a longer arrow represents a lower mean standard deviation. Subfigure 4.1 (a) shows mean relative phase values averaged across pairs, (c) displays these values on a pair-level for each of the 5 pairs. For almost all pairs and the grand average, the Bidirectional Condition (yellow) yields one of the lowest mean relative phase values (i.e. it is the closest to the 0°-point). Looking at the pair-level plots (c), one can see that the Uncoupled Condition (purple) is always close to 90° and that it is very different to the Bidirectional Condition, except for Pair 3.

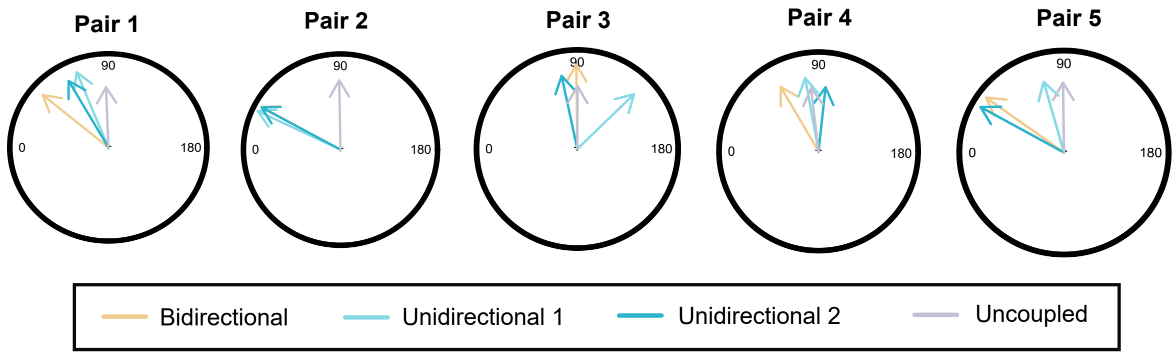
For a comparison with observed mean relative phase angles, Subfigure 4.1 (b) shows a chance distribution of mean relative phase angles which were drawn from a uniform phase distribution. The grand average of these means is indicated by the red arrow. It corresponds to the observed mean relative phase values of the Uncoupled Condition.




**(a)** Mean relative phase values averaged across pairs.




**(b)** Chance distribution of mean relative phase angles.




**(c)** Mean relative phase values on a pair-level.

**Figure 4.1.:** Mean relative phase values for each Auditory Feedback Condition are shown. (a) displays the grand average across pairs and (b) visualizes a chance distribution of mean relative phase angles which were drawn from a uniform phase distribution. The red arrow shows the overall mean. Subfigure (c) shows the mean relative phase values on a pair-level for all 5 pairs. The standard deviation for the grand average was estimated by taking the mean of the standard deviations per feedback condition and pair. The standard deviations per feedback condition and pair stemmed from averaging across trials. The length of the arrows correspond to the mean of the standard deviations, with a longer arrow representing a lower mean standard deviation.

### Mean relative phase values.

Figure 4.2 shows the mean relative phase values per Condition. Mean relative phase angles for each Pair and Condition were submitted to a one-way repeated measures ANOVA, with Auditory Feedback as the independent variable. Findings revealed a main effect of Auditory Feedback Condition on mean relative phase angle, $F(3{,}12) = 3.516$, $p = 0.049$. Subsequent follow-up Tukey tests (corrected for multiple comparisons) were conducted to

assess differences in mean relative phase between pairs of Auditory Feedback Conditions. These tests revealed marginal differences between conditions Bidirectional and Uncoupled, $p = 0.0941$. No further significant differences were observed (all $ps > 0.1$).
The non-parametric equivalent of the repeated measures ANOVA, the Friedman's Test, did not confirm a main effect of Auditory Feedback Condition on mean relative phase angle, $W = 0.264$ (small effect).

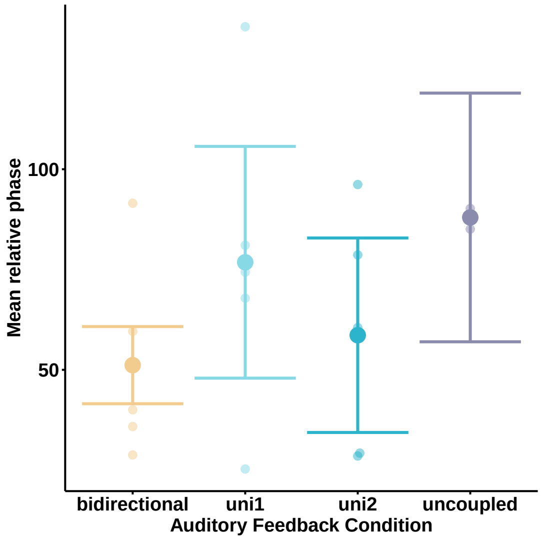



**Figure 4.2.:** Mean relative phase across pairs for each Auditory Feedback Condition: (i) Bidirectional, (ii) Unidirectional 1, and (iii) Unidirectional 2, (iv) Uncoupled. Smaller points represent pair mean relative phases, larger points the grand average.

### 4.2.2. Tests for effects of Auditory Feedback Condition on variability of relative phase angles

Figure 4.3 shows the mean resultant lengths per Condition. The effects of Auditory Feedback on variability of relative phase angles were assessed by computing a one-way repeated measures ANOVA on mean resultant length for each Pair and Condition. Findings revealed a significant main effect of Auditory Feedback Condition on variability of relative phase angles, $F(3,12) = 11.333$, $p = 0.0008$. Subsequent follow-up Tukey tests (corrected for multiple comparisons) were conducted to assess differences in variability of relative phase angles between pairs of Auditory Feedback Conditions. These tests revealed significant differences between conditions Bidirectional and Uncoupled, $p = 0.003$, Uncoupled and Unidirectional 1, $p = 0.045$, as well as a marginal effect between Uncoupled and Unidirectional 2, $p = 0.063$. No further significant differences were observed (all $ps > 0.1$).

Results were confirmed by a non-parametric equivalent of the repeated measures ANOVA (Friedman's Test), $W = 0.856$ (large effect).

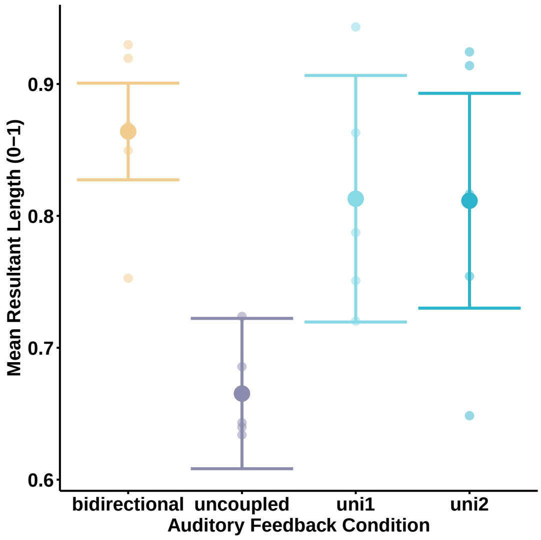
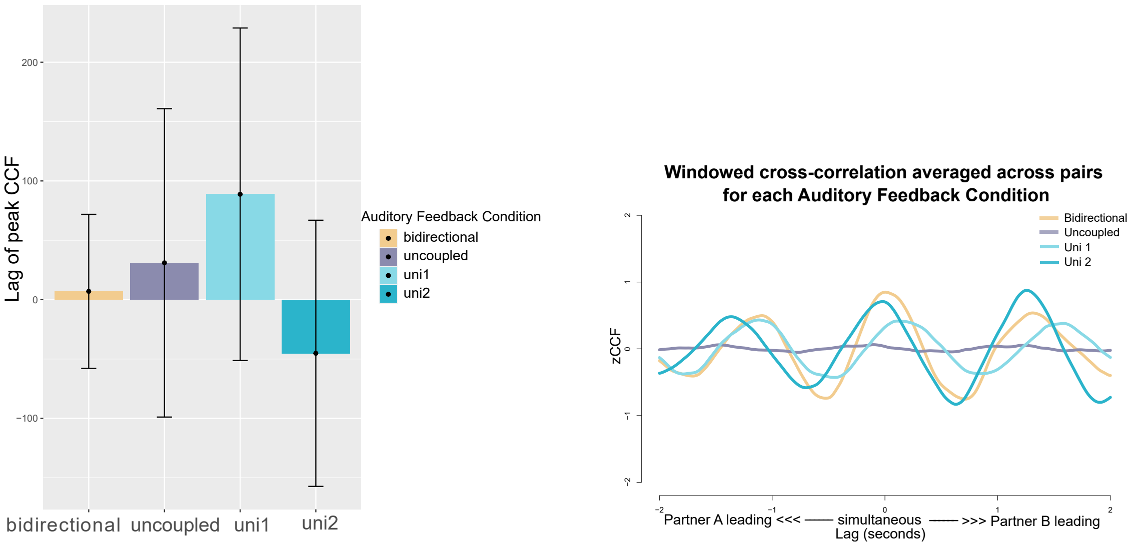



**Figure 4.3.:** Variability of mean relative phase across pairs for each Auditory Feedback Condition: (i) Bidirectional, (ii) Uncoupled, (iii) Unidirectional 1, and (iv) Unidirectional 2. Smaller points represent pair mean relative phases, larger points the grand average.

### 4.2.3. Tests for effects of Auditory Feedback Condition on leader-follower behavior

Subfigure 4.4 (a) shows mean non-windowed cross-correlation lag values of the peak correlation, Subfigure 4.4 (b) displays mean windowed cross-correlation values for lags between -2 and 2 s across pairs for each condition.



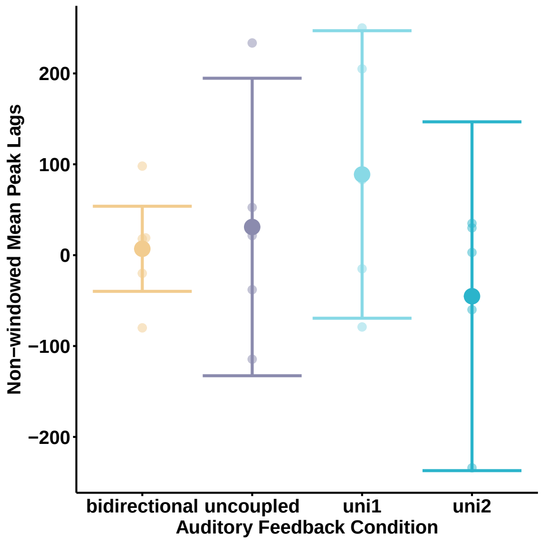

**(a)** Mean non-windowed cross-correlation lag values across pairs.

**(b)** Mean windowed cross-correlation across pairs.

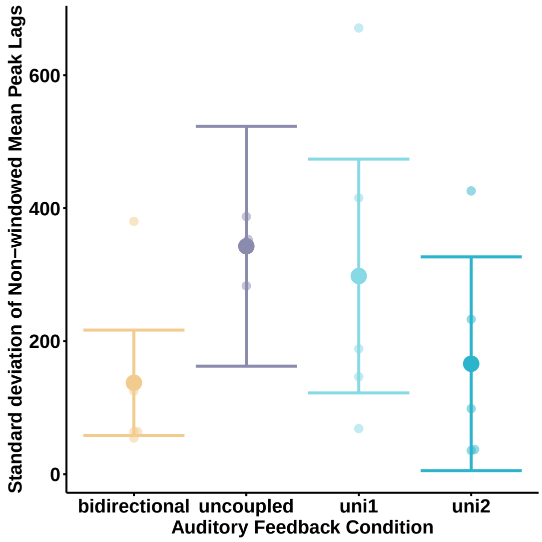
**Figure 4.4.:** Results of (a) non-windowed lag values and windowed (b) cross-correlation functions averaged across pairs for each Auditory Feedback Condition: (i) Bidirectional (orange), (ii) Uncoupled (purple), (iii) Unidirectional 1 (light-blue), and (iv) Unidirectional 2 (blue).

#### 4.2.3.1. Non-windowed cross-correlations

Mean peak lags of non-windowed cross-correlations for each condition can be seen in Subfigure 4.5 (a) and variabilities of peak lags in Subfigure 4.5 (b). Figure 4.6 shows the non-windowed cross-correlation lag values for each pair.

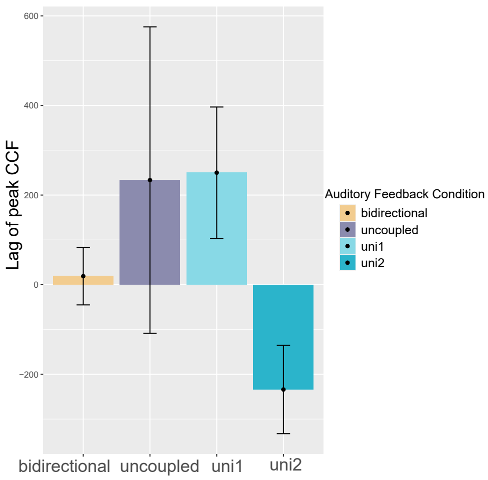



**(a)** Mean peak lags of non-windowed cross-correlation.

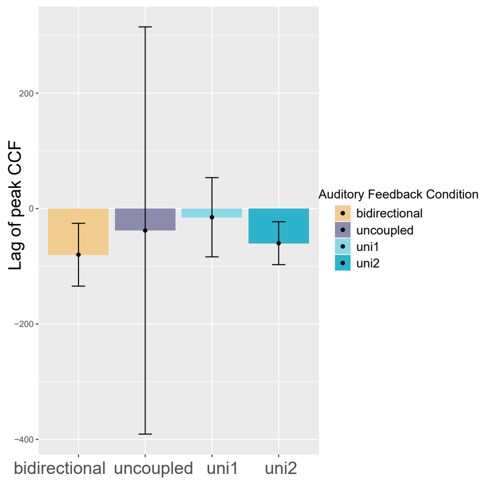



**(b)** Variability of peak lags of non-windowed cross-correlation.

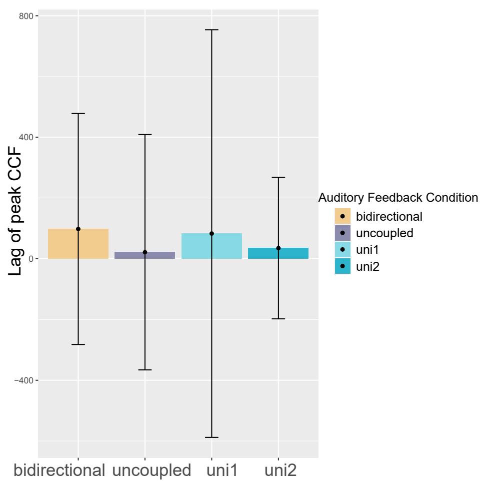
**Figure 4.5.:** Results of non-windowed cross-correlations. (a) shows mean peak lags and (b) variability of mean peak lags across pairs for each Auditory Feedback Condition: (i) Bidirectional, (ii) Uncoupled, (iii) Unidirectional 1, and (iv) Unidirectional 2. Smaller points represent pair non-windowed mean peak lags, larger points the grand average. It can be seen that the all conditions are overlapping.



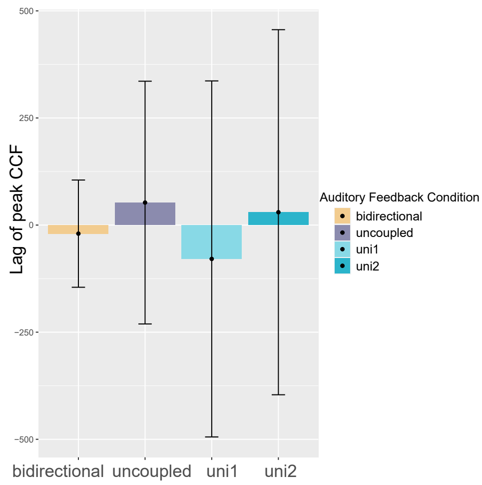

**(a)** Pair 1.



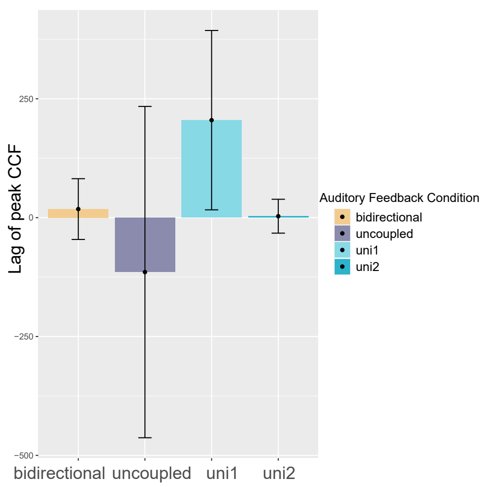

**(b)** Pair 2.




**(c)** Pair 3.




**(d)** Pair 4.




**(e)** Pair 5.

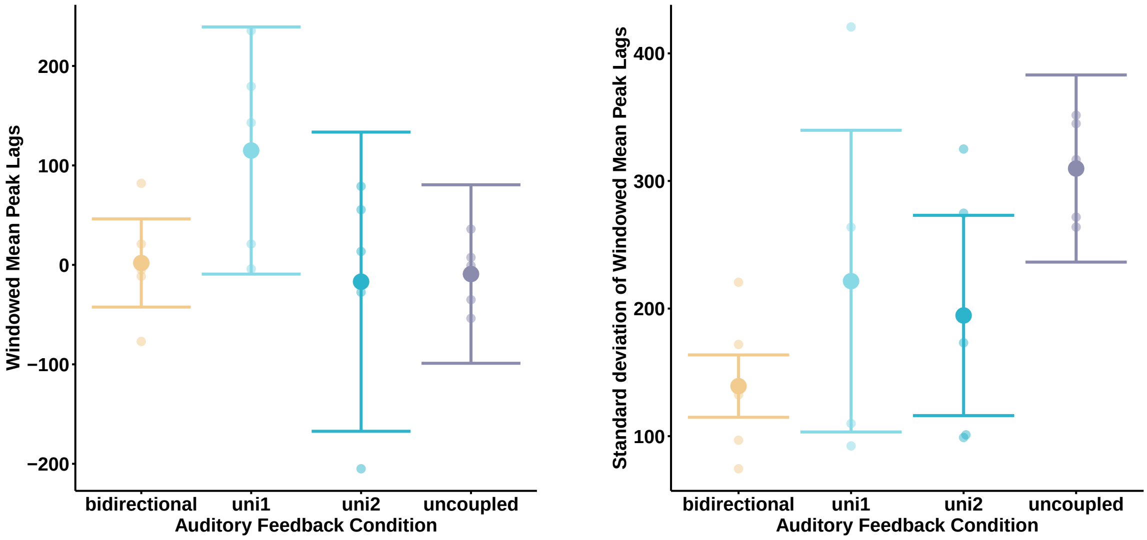
**Figure 4.6.:** Lag peaks of non-windowed cross-correlations per pair. (a) Pair 1, (b) Pair 2, (c) Pair 3, (d), Pair 4 and (e) Pair 5 for each Auditory Feedback Condition: (i) Bidirectional (yellow), (ii) Uncoupled (purple), (iii) Unidirectional 1 (light blue) and (iv) Unidirectional 2 (blue).

**Peak lags.** The one-way repeated measures ANOVA of mean peak lags of non-windowed cross-correlation resulted in a non-significant effect of Auditory Feedback Condition, $F(3{,}12) = 1.049$, $p = 0.406$.
Similarly, the non-parametric Friedman's Test, $W = 0.152$ (small effect), did not give reason to assume an influence of the Auditory Feedback Condition.

**Variability of peak lags.** The variability of peak lags, namely the standard deviation of mean peak lags across trials per Pair and Condition, were submitted to a one-way repeated measures ANOVA, which revealed a marginal effect of Auditory Feedback Condition, $F(3{,}12) = 3.210$, $p = 0.062$. Subsequent Tukey tests (corrected for multiple comparisons) revealed marginal differences between conditions Bidirectional and Uncoupled, $p = 0.068$. No further significant differences were observed (all $p$s $> 0.1$).
The non-parametric Friedman's Test confirmed the obtained effect of Auditory Feedback Condition, $W = 0.52$ (marginally large effect).

#### 4.2.3.2. Windowed cross-correlations

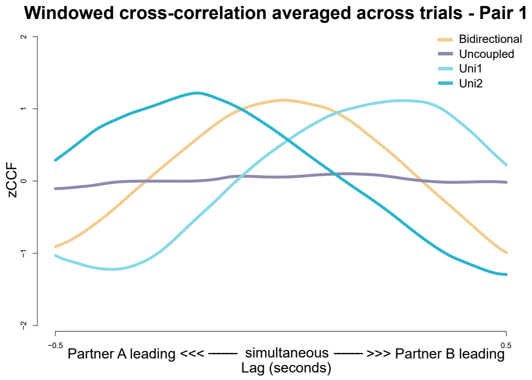
Mean peak lags of windowed cross-correlation for each condition can be seen in Subfigure 4.7 (a) and variabilities of peak lags in Subfigure 4.7 (b). Figure 4.8 shows the windowed cross-correlation lag values for each pair.



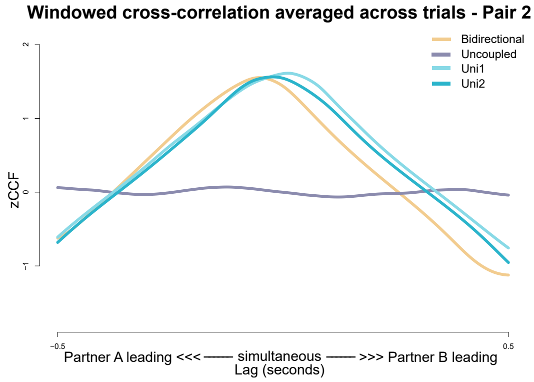

**(a)** Mean peak lags of windowed cross-correlation.

**(b)** Variability of peak lags of windowed cross-correlation.

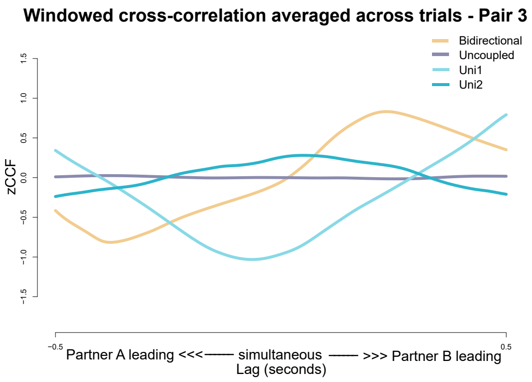
**Figure 4.7.:** Results of windowed cross-correlations. (a) shows mean peak lags and (b) variability of mean peak lags across pairs for each Auditory Feedback Condition: (i) Bidirectional, (ii) Unidirectional 1, and (iii) Unidirectional 2, (iv) Uncoupled. Smaller points represent pair mean relative phases, larger points the grand average.



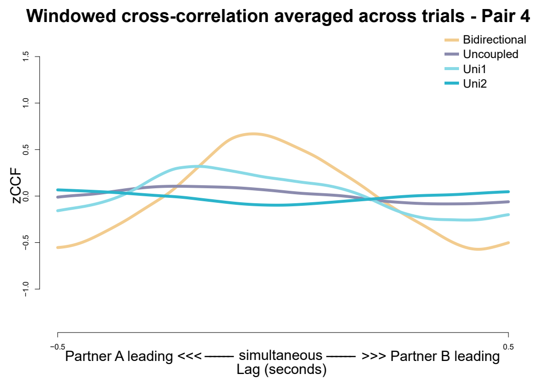

**(a)** Pair 1.



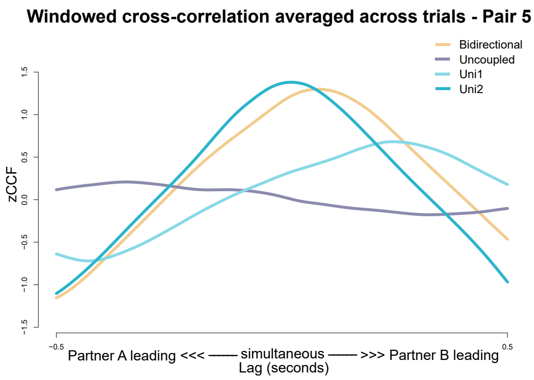

**(b)** Pair 2.




**(c)** Pair 3.




**(d)** Pair 4.




**(e)** Pair 5.

**Figure 4.8.:** Lag peaks of windowed cross-correlations per pair. (a) Pair 1, (b) Pair 2, (c) Pair 3, (d), Pair 4 and (e) Pair 5 for each Auditory Feedback Condition: (i) Bidirectional, (ii) Uncoupled, (iii) Unidirectional 1, and (iv) Unidirectional 2.

**Peak lags.** The lag associated with the peak cross-correlation between lags -0.5 and 0.5 s (i.e. approximately half a cycle, corresponding to the approximate maximum offset in the Uncoupled Condition) was extracted per Pair and Condition. A one-way repeated

measures ANOVA on these lags showed no significant effect of Auditory Feedback Condition, $F(3{,}12) = 2.466$, $p = 0.112$.
Similarly, the non-parametric Friedman's Test, $W = 0.36$ (moderate effect), did not give reason to assume an influence of the Auditory Feedback Condition.

**Variability of peak lags.** Standard deviation values of peak lags across trials within Pairs and Conditions across windows were submitted to a one-way repeated measures ANOVA, showing a significant effect of Auditory Feedback Condition, $F(3{,}12) = 5.973$, $p = 0.010$. Subsequent Tukey tests (corrected for multiple comparisons) revealed significant differences between the conditions Bidirectional and Uncoupled, $p = 0.001$. No further significant differences were observed (all $p$s $> 0.1$).
The results were confirmed by a Friedman's Test, $W = 0.6$ (large effect).

# 5. Discussion

## 5.1. Interpretation and implications of the results

### 5.1.1. Discrete vs. continuous auditory-motor rhythms and the influence of Auditory Feedback Conditions

Discrete and continuous sounds and motions differ from one another and require different timing mechanisms (Huys et al., 2008; Torre & Balasubramaniam, 2009). Research on temporal synchronization between partners in joint action has mainly focused on discrete auditory-motor rhythms (Demos et al., 2019; Goebl & Palmer, 2009; Konvalinka et al., 2010; Loehr et al., 2013; Loehr & Palmer, 2011; Loehr & Vesper, 2015; Zamm et al., 2015; Zelic et al., 2019), looking at the synchronization of partners at the onset of a discrete sound. These studies showed that auditory feedback has an effect on leader-follower behavior. If no auditory information is available, partners do not synchronize or adapt to one another and show an uncoupled behavior (Demos et al., 2019; Goebl & Palmer, 2009; Konvalinka et al., 2010). If partners can hear both, their partner and themselves, they tend to mutually adapt to one another. This results in a bidirectional pattern (Demos et al., 2019; Goebl & Palmer, 2009; Konvalinka et al., 2010; Zamm et al., 2015) and a "hyper-follower" dynamic (Konvalinka et al., 2010). In situations in which one of the auditory feedbacks is removed in a discrete auditory-motor rhythm study, unidirectional adaptation emerges (Demos et al., 2019; Konvalinka et al., 2010). In these cases, the partner whose output is no longer audible to the other partner takes up the role of a follower by anticipating the other's actions, while the other becomes the leader (Demos et al., 2019; Konvalinka et al., 2010).

However, in many situations (e.g. music) auditory-motor information and movement is continuous. The current experiment made a novel contribution to the literature by investigating how partners coordinate the timing of continuous rhythmic movements that produce continuous auditory feedback.

#### 5.1.1.1. Mean relative phase angles and variability of relative phase angles

The present paradigm revealed similar effects on synchronization as observed in discrete coordination paradigms. Specifically, if both partners were able to hear their partner (Bidirectional Condition), the grand average across pairs of mean relative phases showed the smallest value (see Figure 4.1 (a)). This shows that in situations in which the most auditory feedback was available to partners, partners showed the highest synchronization. This is in line with findings from studies with discrete auditory-motor rhythms and motions (Demos et al., 2019; Goebl & Palmer, 2009; Konvalinka et al., 2010; Zamm et al., 2015). It should be noted that not all pairs displayed this effect. Further pairs are required to obtain a more stable estimate.

Removing the feedback from one's partner, so that the partners could only hear themselves, also showed the same pattern as previously observed in the discrete domain (Demos et al., 2019; Goebl & Palmer, 2009; Konvalinka et al., 2010). Pairs displayed uncoupled behavior without any synchronization, resulting in a mean relative phase angle across pairs of around 90° for the Uncoupled Condition (see Figure 4.1 (a)). All pairs showed this behavior with a mean relative phase angle of around 90° for the Uncoupled Condition (see Figure 4.1 (c)). A chance distribution was constructed by taking the mean phase values of 1000 simulated uniform distributions, resulting in a grand average of 90°. This shows that chance-level yields a result of 90°, which was also found for the Uncoupled Condition. It can therefore be concluded that removing the feedback from one's partner for both partners results in synchronization comparable to chance-level, namely no synchronization takes place.

Conditions, in which both partners only heard one partner (i.e. in Unidirectional 1 they heard Partner A, in Unidirectional 2 they heard Partner B), showed grand average mean relative phases located between the Bidirectional and Uncoupled Condition. This shows that in situations in which partners only heard part of the auditory feedback (either themselves or their partner), they were better able to synchronize than in the Uncoupled Condition as the mean relative phase was larger. However, they were not as well synchronized as in the Bidirectional Condition. These findings show that the more auditory feedback was available, the better partners were able to synchronize (i.e. the smaller the mean relative phase). These findings are comparable to results from the discrete auditory domain (Demos et al., 2019; Konvalinka et al., 2010).
Furthermore, results from a one-way repeated measures ANOVA showed a significant ef-

fect of Auditory Feedback Conditions on mean relative phase with a marginal difference between the Bidirectional and Uncoupled Condition, indicating that the synchronization was better in the Bidirectional Condition. However, it should be noted that a non-parametric equivalent was not able to confirm this significant finding.

Also, a one-way repeated measures ANOVA on mean resultant length (e.g. variability) of relative phase angles revealed a significant effect of Auditory Feedback Conditions (confirmed by non-parametric Friedman's Test) with significant differences between Bidirectional and Uncoupled Conditions as well as marginal differences between Uncoupled and Unidirectional 1 and Uncoupled and Unidirectional 2 Conditions. These findings show that the stability of synchrony (i.e. variability) differed between conditions. Partners showed the highest stability in synchrony in the Bidirectional Condition and the highest variability (i.e. least stability) in the Uncoupled Condition, meaning that they were able to stay synchronized the longest in the Bidirectional Condition. The analysis further revealed marginal differences in stability of synchrony between the Unidirectional and Uncoupled Condition. Partners were able to stay synchronized for a longer time period in the Unidirectional Conditions than in the Uncoupled Condition. These results show that more auditory feedback yields greater stability of synchrony between partners.

#### 5.1.1.2. Leader-follower behavior

To assess patterns of leader-follower behavior, cross-correlations (non-windowed and windowed), were calculated.

**Non-windowed cross-correlation.** The non-windowed cross-correlation showed a mean lag peak close to 0 for the Bidirectional Condition (see Figure 4.4 (a)), which is in line with results from studies with discrete auditory-motor rhythms (Demos et al., 2019; Konvalinka et al., 2010), which found that partners mutually adapt to one another and show a "hyper-follower" dynamic. A lag peak close to lag 0 indicates that they had the highest correlation when no lag in time was present. Non of the partners was leading the other and no leader-follower dynamic emerges, suggesting that partners were mutually adapting to one another in the Bidirectional Condition.

For Unidirectional 1 (both hear Partner A) the cross-correlation yielded mean positive lags, while Unidirectional 2 (both hear Partner B) showed mean negative results. The used `ccf` function was used in such a way that a negative lag means that Partner A is leading Partner B. In contrast, a positive lag indicates that Partner B is leading Partner A. Therefore, the non-windowed cross-correlation showed that the positive lag in the Unidirectional 1 Condition (both hear Partner A) indicates that Partner B is leading Partner A, despite the fact that no auditory feedback of Partner B was present. The Unidirectional 2 Condition showed a negative lag in non-windowed cross-correlation, therefore, Partner A was leading Partner B. This shows the same behavior as for Unidirectional 1, namely that the partner whose feedback had been removed took up the leader role. This behavior reflects mean negative asynchrony (sometimes also called anticipation tendency), which refers to the phenomenon that people tend to act ahead of an unresponsive metronome when trying to synchronize with it (e.g. tap ahead of the metronome by a few milliseconds) (Repp, 2005; Repp & Su, 2013). The reported leader-follower dynamics are in line with these findings of anticipatory behavior with an unresponsive metronome (Repp, 2005; Repp & Doggett, 2007; Repp & Su, 2013; Washburn et al., 2019). In the Unidirectional Condition the partner, whose feedback has been removed, needs to synchronize with a partner, who only hears their own output, i.e. an unresponsive partner. The results indicate that in these situations partners tend to anticipate the movement of their unresponsive partner in the same way as in situations with a metronome, which is also unresponsive, by taking up a leader role. One explanation for these results could be that partners treat their unresponsive partner like a metronome. It should be noted that not all pairs displayed these effects. Further pairs are required to obtain a more stable estimate.

A marginal difference in variability of peak lags between Bidirectional and Uncoupled Condition indicates that partners showed a higher stability of peak lags in the Bidirectional than in the Uncoupled Condition. More specifically, partners showed a stable pattern on lag 0, meaning that they showed a stable dynamic in which non of the partners was clearly leading or following but they were mutually adapting to one another without much variability. In contrast, partners did not show a stable peak lag pattern in the Uncoupled Condition, meaning that they showed no stable behavior of adapting to one another. These results show that the amount of auditory feedback influences the stability of adaptive behavior in such a way that more auditory feedback shows less variability in the adaptation dynamics.

**Windowed cross-correlation.** The same leader-follower dynamics as extracted by the non-windowed cross-correlation were confirmed by windowed cross-correlation (see Figure 4.4), suggesting stability of the pattern across the duration of a trial.

Only low correlations were observed at any specific lag for the Uncoupled Condition. This suggests a chance-level in synchronization for the Uncoupled Condition as all lags between partners are equally likely to get established between them. A high positive correlation for lag 0 and any lags which are multiples of around 1.3 s (one full circle) were observed for the Bidirectional Condition. This shows that partners synchronized their movements based on the audible pitch frequencies with one another. They were not taking up follower or leader roles but rather mutually adapting to one another, which is in line with results from studies with discrete auditory-motor rhythms (Demos et al., 2019; Konvalinka et al., 2010), which found that partners show a "hyper-follower" dynamic and mutually adapt to one another.

In line with the results of the non-windowed cross-correlation, the Unidirectional 1 Condition showed that Partner B was leading Partner A, while in the Unidirectional 2 Condition Partner A was leading Partner B. This shows again the pattern that the partner whose feedback had been removed took up the leader role, while the partner whose output was still audible took up the follower role. This leader-follower behavior is in line with results from studies with unresponsive metronomes, which show that people tend to show anticipatory behavior (i.e. mean negative asynchrony) (Repp, 2005; Repp & Doggett, 2007; Repp & Su, 2013; Washburn et al., 2019). As stated above, an interpretation of this behavior could be that partners treat their unresponsive partner like a metronome. It should be noted that not all pairs displayed these effects. Further pairs are required to obtain a more stable estimate.

A significant difference in variability of peak lags between Bidirectional and Uncoupled Condition was revealed for the windowed cross-correlation. This indicates that partners showed a more stable adaptation dynamic in the Bidirectional than in the Uncoupled Condition. As the peak lag was around lag 0 in the Bidirectional Condition, this indicates that partners were mutually adapting to one another without much variability across the duration of the trials. In contrast, partners did not show a stable peak lag pattern in the Uncoupled Condition, which means that no stable behavior of adaptation was established. These results confirm the ones obtained with the non-windowed cross-correlation, namely that the amount of auditory feedback influences the stability of adaptive behavior in such a way that more auditory feedback shows less variability in the adaptation dynamics.

### 5.1.2. Influence of musical training

One aspect to consider is the influence of musical training, which was recorded for each participant and then qualitatively compared to the findings.

#### 5.1.2.1. Mean relative phase angles and variability of relative phase angles

The differences in years of musical training within a pair ranged from 1 to 11 years (see Table 2.2). However, no connection between differences in years of musical training and mean relative phase angles could be qualitatively established (see Figure 4.1). In addition, no qualitative pattern emerged when looking at the total number of years of musical training within a pair.

When comparing the conditions Unidirectional 1 and Unidirectional 2, another pattern emerges. If the feedback of the partner with more years of musical training could be heard, pairs performed worse in this condition than in the Unidirectional Condition in which they heard the partner with less musical experience. This suggests, that partners with more musical training were better at adapting to their partner than those with less years of musical training. However, this pattern needs to be quantitatively confirmed with a larger sample.

#### 5.1.2.2. Leader-follower behavior

**Cross-correlations.** Non-windowed and windowed cross-correlation showed the same patterns. Pairs with few years of musical training for both partners qualitatively showed leader-follower patterns different to the other pairs (see Figure 4.6 and Figure 4.8). Further, also cross-correlations showed that in the Unidirectional Condition in which the partner with more musical training heard the partner with less musical training, pairs performed better (i.e. higher correlation values). One possible interpretation is that partners with more musical training were better at adapting to their partner than those with less years of musical training. This pattern needs to be quantitatively confirmed with a larger sample.

#### 5.1.2.3. Individual Differences

A caveat for interpretation are individual differences between participants and pairs. Participants showed varying results in the accuracy scores and threshold values, indicating that not all participants learned the sound-space mapping equally well (see Table 4.1). In

addition, participants had differing years of musical education (see Table 2.1). Comparing the years of musical training of each participant with the accuracy scores and threshold values of the Audio-Visual Perception Task, showed that there was a tendency for a better performance in the Audio-Visual Perception Task if the participant had more years of musical training and a rather low score for participants with none or only 1-2 years of musical training (compare Table 2.2 and Table 4.1). These individual differences can have an impact on the interpretations of the results.
Further, these individual differences lead to differences between pairs. It should be noted that not all pairs learned the sound-space mapping equally well, which should be kept in mind when interpreting data per pair.

## 5.2. Future directions

As the number of pairs in this study was limited to 5, further research with a larger sample size is needed to investigate the observed patterns further.
In addition, a higher sample size could also give further statistical insights into the relationship between years of musical training of participants and their ability to adapt to their partner. Differences in years of musical training within a pair could also play a role, which could be further investigated by actively matching participants of similar or different numbers of years of musical education.

Further, with a larger sample size the results of the Audio-Visual Perception Task could be used as a covariate in the analyses as it can be expected that a worse performance in this phase indicates that this participant will have problems adapting in the Joint Task.

In this study, the pitch frequency was not counterbalanced across pairs, meaning that Partner B was always assigned the higher pitch frequency. Another experiment should also counterbalance these pitches to confirm the results.

## 5.3. Summary & Contributions

This experiment looked at synchronization of continuous auditory-motor rhythm between partners, which helps to understand interpersonal communication in social interaction better, specifically in music performance (Goebl & Palmer, 2009).

Synchronization of continuous auditory-motor rhythms with a partner was influenced by auditory feedback coupling. The more auditory feedback partners had available the better they were able to synchronize (i.e. smaller phase offset). In situations in which they heard each other (Bidirectional Condition), they showed the highest synchronization (i.e. smallest phase offset) and a pattern of mutual adaptation of their movements (peak lags at lag 0). If feedback of one partner was removed (Unidirectional Conditions), this partner started to anticipate the movement of the partner whose feedback was still audible. The partner whose feedback had been removed took up a leader role, while the other partner took up a follower role. If partners were unable to hear each other (Uncoupled Condition), they showed no synchronization and a distribution of mean relative phases at chance-level. The amount of auditory feedback information drastically changed the adaptation behavior and ability to synchronize with a partner. Further, stability of synchrony and stability of leader-follower dynamics were influenced by auditory feedback coupling. The more auditory feedback was available to partners, the higher was the stability of synchrony and stability of leader-follower dynamics (i.e. less variability).

The current experiment made a novel contribution to the literature of auditory-motor rhythms synchronization by studying how partners coordinate continuous auditory-motor rhythms with one another. As much of the literature investigating synchronization of auditory-motor rhythms has focused on how partners coordinate discrete movements and sounds (Demos et al., 2019; Goebl & Palmer, 2009; Konvalinka et al., 2010; Loehr et al., 2013; Loehr & Palmer, 2011; Loehr & Vesper, 2015; Zamm et al., 2015; Zelic et al., 2019), this study provides first insights in the continuous auditory-motor rhythm domain. More specifically, it investigated how partners coordinate timing of continuous movements that produce continuous auditory feedback. This is relevant, as discrete and continuous sounds are perceptually distinct and different movements, which underlie different timing mechanisms (Zelic et al., 2019), are needed to produce them (Torre & Balasubramaniam, 2009). Therefore, the observed patterns for discrete coordination do not necessarily give insights into the patterns of how partners synchronize continuous auditory-motor rhythms under different auditory feedback conditions. Further, the results on continuous coordination obtained here are especially important as many natural situations requiring auditory-motor synchronization (e.g. music) involve continuous movements and sounds rather than discrete ones.

The current study shows that the present paradigm revealed similar effects of auditory feedback coupling on synchronization as observed in discrete coordination paradigms. The amount of auditory feedback is crucial for synchronization also for continuous movements and sounds. Full auditory feedback revealed a pattern of mutual adaptation and the highest synchronization, while the least amount of auditory feedback resulted in no synchronization between partners. An anticipatory behavior was observed for situations in which the feedback of only one partner was audible.

On a broader scope, the results of this study yield insights into mechanisms of interpersonal synchronization, successful action coordination and interpersonal communication (e.g. in music performance). Further, findings about sensorimotor synchrony can provide important insights for rehabilitation and treatments of motor coordination diseases, such as Parkinson (Fujii & Wan, 2014).

# Abbreviations

**BPM** Beats Per Minute. 18

**CEU PU** Central European University Private University. 9, 10

**FM** Frequency Modulated. 15, 16, 18, 22

**IBI** Inter-Beat Interval. 18

**SOMBY-Lab** Social Mind and Body Lab. 10

# List of Figures

# List of Tables

# Bibliography


Augmented Instruments Ltd. (2021a). *Bela - About*. Retrieved July 4, 2021, from https://bela.io/about

Augmented Instruments Ltd. (2021b). *Bela project settings*. Retrieved July 4, 2021, from https://learn.bela.io/the-ide/settings/#project-settings

Augmented Instruments Ltd. (2021c). *I2C protocol*. Retrieved August 20, 2021, from https://learn.bela.io/products/trill/all-about-i2c/

Augmented Instruments Ltd. (2021d). *Integrating Trill into your projects*. Retrieved July 19, 2021, from https://learn.bela.io/products/trill/integrating-trill-into-your-projects/

Augmented Instruments Ltd. (2021e). *Trill sensors*. Retrieved July 4, 2021, from https://bela.io/products/trill/

Bauer, A.-K. R., Bleichner, M. G., Jaeger, M., Thorne, J. D., & Debener, S. (2018). Dynamic phase alignment of ongoing auditory cortex oscillations. *NeuroImage*, *167*, 396–407. https://doi.org/10.1016/j.neuroimage.201.11.037

BelaPlatform. (2021a). *Bela versions*. Retrieved July 17, 2021, from https://github.com/BelaPlatform/bela-image-builder/releases

BelaPlatform. (2021b). *Fritzing*. Retrieved August 21, 2021, from https://github.com/BelaPlatform/fritzing

CEU. (2021). *SONA - Research Participation System*. https://ceuparticipate.sona-systems.com/exp_rests1.aspx?p_experiment_id=344

Cornsweet, T. N. (1962). The Staircase-Method in Psychophysics.

Cypress Semiconductor Corporation. (2015). *CapSense Sigma-Delta Datasheet*. Retrieved July 17, 2021, from https://www.cypress.com/file/124551/download

DataNovia. (2021). *Friedman Test in R*. https://www.datanovia.com/en/lessons/friedman-test-in-r/

Demos, A. P., Layeghi, H., Marcelo M., W., & Caroline, P. (2019). Staying Together: A Bidirectional Delay-Coupled Approach to Joint Action. *Cognitive Science*, *43*, 1–25. https://doi.org/10.1111/cogs.12766

Fujii, S., & Wan, C. Y. (2014). The role of rhythm in speech and language rehabilitation: the SEP hypothesis. *Frontiers in Human Neuroscience*, *8*, 1–15. https://doi.org/10.3389/fnhum.2014.00777

Giuliomoro. (2021). *Trill ring sampling issue*. Retrieved July 4, 2021, from https://forum.bela.io/d/1544-trill-ring-sampling-issue/54

Goebl, W., & Palmer, C. (2009). Synchronization of Timing and Motion Among Performing Musicians. *Music Perception*, *26*, 427–438. https://doi.org/10.1525/MP.2009.26.5.427

Greenberg, D., Baron-Cohen, S., Stillwell, D., Kosinksi, M., & Rentfrow, P. (2015). Musical preferences are linked to cognitive styles. *PLoS One*, 1–22. https://doi.org/10.1371/journal.pone.0131151

Haegens, S., & Golumbic, E. Z. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience and Biobehavioral Reviews*, *86*, 150–165. https://doi.org/10.1016/j.neubiorev.2017.12.002

Heggli, O. A., Cabral, J., Konvalinka, I., Vuust, P., & Kringelbach, M. L. (2019). A Kuramoto model of self other integration across interpersonal synchronization strategies. *PLoS: Computational Biology*, *15*, 1–17. https://doi.org/10.1371/journal.pcbi.1007422

Hijmans, R. J. (2021). *Package 'raster'*. https://cran.r-project.org/web/packages/raster/raster.pdf

Huys, R., Studenka, B. E., Rheaume, N. L., Zelaznik, H. N., & Jirsa, V. K. (2008). Distinct Timing Mechanisms Produce Discrete and Continuous Movements. *PLoS Computational Biology*, *4*, 1–8. https://doi.org/10.1371/journal.pcbi.1000061

Kassambara, A. (2021). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests*. https://cran.r-project.org/web/packages/rstatix/index.html

Kleinbub, J. R., & Ramseyer, F. T. (2020). rMEA: An R package to assess nonverbal synchronization in motion energy analysis time-series. *Psychotherapy Research*. https://doi.org/10.1080/10503307.2020.1844334

Konvalinka, I., Bauer, M., Stahlhut, C., Hansen, L. K., Roepstorff, A., & Frith, C. D. (2014). Frontal alpha oscillations distinguish leaders from followers: Multivariate decoding of mutually interacting brains. *NeuroImage*, *94*, 79–88. https://doi.org/10.1016/j.neuroimage.2014.03.003

Konvalinka, I., & Roepstorff, A. (2012). The two-brain approach: how can mutually interacting brains teach us something about social interaction? *Frontiers in Human Neuroscience*, *6*, 1–10. https://doi.org/10.3389/fnhum.2012.00215

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. *The quarterly jour-*

*nal of experimental psychology*, *63*, 2220–2230. https://doi.org/10.1080/17470218.2010.497843

Loehr, J. D., Kourtis, D., Vesper, C., Sebanz, N., & Knoblich, G. (2013). Monitoring Individual and Joint Action Outcomes in Duet Music Performance. *Journal of Cognitive Neuroscience*, *25*, 1049–1061. https://doi.org/10.1162/jocn_a_00388

Loehr, J. D., & Palmer, C. (2011). Temporal coordination between performing musicians. *The quarterly journal of experimental psychology*, *64*, 2153–2167. https://doi.org/10.1090/17470218.2011.603427

Loehr, J. D., & Vesper, C. (2015). The sound of you and me: Novices represent shared goals in joint action. *The quarterly Journal of Experimental Psychology*, *69*, 535–547. https://doi.org/10.1080/17470218.2015.1061029

Lund, U., Agostinelli, C., Arai, H., Gagliardi, A., Portugues, E. G., Giunchi, D., Irisson, J.-O., Pocernich, M., & Rotolo, F. (2017). *Package 'circular'*. https://cran.r-project.org/web/packages/circular/circular.pdf

Maes, P.-J., Wanderley, M. M., & Palmer, C. (2015). The role of working memory in the temporal control of discrete and continuous movements. *Exp Brain Res*, *233*, 263–273. https://doi.org/10.1007/s00221-014-4108-5

McPherson, A. P., & Zappi, V. (2015). *An Environment for Submillisecond-Latency Audio and Sensor Processing on BeagleBone Black*. Journal of Audio Engineering Society.

Moro, G., Bin, A., Jack, R. H., Heinrichs, C., & McPherson, A. P. (2016). *Making High-Performance Embedded Instruments with Bela and Pure Data*.

Novembre, G., Knoblich, G., Dunne, L., & Keller, P. E. (2017). Interpersonal synchrony enhanced through 20Hz phase-coupled dual brain stimulation. *Social Cognitive and Affective Neuroscience*, 662–670. https://doi.org/10.1093/scan/nsw172

Novembre, G., Ticini, L. F., Schütz-Bosbach, S., & Keller, P. E. (2014). Motor simulation and the coordination of self and other in real-time joint action. *Social Cognitive and Affective Neuroscience*, *9*, 1062–1068. https://doi.org/10.1093/scan/nst086

Palmer, C., Spidle, F., Koopmans, E., & Schubert, P. (2019). Ears, heads, and eyes: When singers synchronise. *Quarterly Journal of Experimental Psychology*, *72*, 2272–2287. https://doi.org/10.1177/1747021819833968

Pikovsky, A., Rosenblum, M., & Kurths, J. (2001). *Synchronization*. Cambridge University Press. https://doi.org/10.1017/CBO9780511755743

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

R Core Team and contributors worldwide. (2021a). *The R Stats Package*. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/stats-package.html

R Core Team and contributors worldwide. (2021b). *The R Stats Package*. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/stats-package.html

Repp, B. H. (2005). Sensorimotor syncrhonization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*, 969–992.

Repp, B. H. (2010). Sensorimotor synchronization and perception of timing: Effects of music training and task experience. *Human Movement Science*, *29*, 200–213. https://doi.org/10.1016/j.humov.2009.08.002

Repp, B. H., & Doggett, R. (2007). Tapping to a Very Slow Beat: A Comparison of Musicians and Nonmusicians. *Music Perception: An Interdisciplinary Journal*, *24*, 367–376. https://doi.org/10.1525/MP.2007.24.4.367

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin & Review*, *20*, 403–452. https://doi.org/10.3758/s13423-012-0371-2

Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R., & Schmidt, R. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human Movement Science*, 1–25. https://doi.org/1016/j.humov.2007.07.002

Richardson, M. J., Marsh, K. L., & Schmidt, R. (2005). Effects of Visual and Verbal Interaction on Unintentional Interpersonal Coordination. *Journal of Experimental Psychology*, *31*, 62–79. https://doi.org/10.1037/0096-1523.31.1.62

RStudio Team. (2020). *RStudio: Integrated Development Environment for R*. RStudio, PBC. Boston, MA. http://www.rstudio.com/

Schuster, H., & Wagner, P. (1989). Mutual Entrainment of Two Limit Cycle Oscillators with Time Delayed Coupling. *Progress of Theoretical Physics*, *81*, 939–945.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: bodies and minds moving together. *TRENDS in Cognitive Science*, *10*, 70–76. https://doi.org/10.1016/j.tics.2005.12.009

Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M. S., Højsgaard, S., Fox, J., Lawrence, M. A., Mertens, U., Love, J., Lenth, R., & Christensen, R. H. B. (2021). *Package 'afex'*. https://cran.r-project.org/web/packages/afex/afex.pdf

Strogatz, S. H., & Stewart, I. (1993). Coupled Oscillators and Biological Synchronization. *Scientific American*, 102–109.

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct Neural Substrates of Duration-Based and Beat-Based Auditory Timing. *The Journal of Neuroscience*, *31*, 3805–3812. https://doi.org/10.1523/JNEUROSCI.5561-10.2011

The Pennsylvania State University. (2021). *Cross Correlation Functions and Lagged Regressions*. https://online.stat.psu.edu/stat510/lesson/8/8.2

Torre, K., & Balasubramaniam, R. (2009). Two different processes for sensorimotor synchronization in continuous and discontinuous rhythmic movements. *Experimental Brain Research*, *199*, 157–166. https://doi.org/10.1007/s00221-009-1991-2

van der Wel, R. P., Knoblich, G., & Sebanz, N. (2011). Let the Force Be With Us: Dyads Exploit Haptic Coupling for Coordination. *Journal of Experimental Psychology*, *37*, 1420–1431. https://doi.org/10.1037/a0022337

van Gelder, T. (1998). The dynamical hypothesis in Cognitive Science. *Behavioral and brain sciences*, *21*, 615–665.

Washburn, A., Román, I., Huberth, M., Gang, N., Dauer, T., Reid, W., Nanou, C., Wright, M., & Fujioka, T. (2019). Musical Role Asymmetries in Piano Duet Performance Influence Alpha-Band Neural Oscillation and Behavioral Synchronization. *Frontiers in Neuroscience*, *13*, 1–18. https://doi.org/10.3389/fnins.2019.1088

Zamm, A., Pfordresher, P. Q., & Palmer, C. (2015). Temporal coordination in joint music performance: effects of endogenous rhythms and auditory feedback. *Experimental Brain Research*, *233*, 607–615. https://doi.org/10.1007/s00221-014-4140-5

Zelaznik, H. N., Spencer, R. M., Ivry, R. B., Baria, A., Bloom, M., Dolansky, L., Justice, S., Patterson, K., & Whetter, E. (2005). Timing Variability in Circle Drawing and Tapping. Probing the Relationship Between Event and Emergent Timing. *Journal of Motor Behavior*, *37*, 395–403. https://doi.org/10.3200/JMBR.37.5.395-403

Zelic, G., Nijuhuis, P., Charaf, S. A., Keller, P. E., Davis, C., Kim, J., & Varlet, M. (2019). The influence of pacer-movement continuity and pattern matching on auditory-motor synchronisation. *Experimental Brain Research*, *237*, 2705–2713. https://doi.org/10.1007/s00221-019-05625-9

# Appendix

## A. Demographic questionnaire

**To be filled out by experimenter after the study:**

Date:

Participant ID:

**To be filled out by participant:**

Age: _____________

Gender: _____________

First (native) language: _____________

Second language(s): _____________

# years of training on your primary musical instrument(s): _____________

Age at which you started musical training: _____________

# years of training on other instruments/age started/age stopped: ________________

Did you ever play in musical ensembles? (Yes/No): _____________

Dominant Hand (Right/Left/Ambidextrous): _____________

Do you have normal hearing? (Yes/No): _____________

Do you have normal/corrected-normal vision? (Yes/No): _____________

Do you have difficulty moving to a beat? (Yes/No): _____________

Do you have difficulty hearing pitches? (Yes/No): _____________

Do you have absolute pitch? (Yes/No): _____________

If you have absolute pitch, please indicate whether your absolute pitch is for tones produce by all instruments or just a specific instrument: __________

How many hours per week do you spend listening to music?: _____________

Do you have any training in dance or other activities that involve moving to music? (Yes/No): _____________

If yes to the above, what activities do you have training in and how many years of training per activity?: _____________

Do you have any history of psychiatric/neurological diagnosis?

Do you have any history of consuming neuropsychiatric medication?

Did you know your partner in advance of the study? (Yes/No): ______

If so, have you ever played music with your partner? (Yes/No): _________

If yes to the above, please describe in what capacity you have played music (e.g. piano duets, in a band, etc.): ____________