

**This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.**

**Author(s):** Burger, Birgitta; London, Justin; Thompson, Marc; Toiviainen, Petri

**Title:** Synchronization to metrical levels in music depends on low-frequency spectral components and tempo

**Year:** 2018

**Version:** Accepted version (Final draft)

**Copyright:** © Springer-Verlag GmbH, 2017

**Rights:** In Copyright

**Rights url:** <http://rightsstatements.org/page/InC/1.0/?language=en>

**Please cite the original version:**

Burger, B., London, J., Thompson, M., & Toiviainen, P. (2018). Synchronization to metrical levels in music depends on low-frequency spectral components and tempo. *Psychological Research*, 82(6), 1195-1211. <https://doi.org/10.1007/s00426-017-0894-2>

# **Synchronization to metrical levels in music depends on low frequency spectral components and tempo**

Birgitta Burger<sup>1\*</sup>, Justin London<sup>2</sup>, Marc R. Thompson<sup>1</sup>, and Petri Toiviainen<sup>1</sup>

<sup>1</sup> Finnish Centre for Interdisciplinary Music Research, Department of Music, Art and Culture Studies, P.O. Box 35, FI-40014 University of Jyväskylä, Finland

<sup>2</sup> Department of Music, Carleton College, 1 North College Street, Northfield, MN 55057, United States

\* Corresponding author:

email address: birgitta.burger@jyu.fi

telephone: +358 40 805 4301

Acknowledgments: We wish to thank two anonymous reviewers for their thorough and valuable comments on the manuscript. We further thank Mikko Leimu for technical assistance during the mocap recordings and Emily Carlson for proofreading the manuscript. This research was supported by an Academy of Finland grant (project “Dynamics of Music Cognition,” project numbers 272250, 274037) to authors BB, MT, and PT, and by a Finnish Core Fulbright Scholar grant to author JL.

## **Abstract**

Previous studies have found relationships between music-induced movement and musical characteristics on more general levels, such as tempo or pulse clarity. This study focused on synchronization abilities to music of finely-varying tempi and varying degrees of low frequency spectral change/flux. Excerpts from six classic Motown/R&B songs at three different tempos (105, 115, and 130 BPM) were used as stimuli in this experiment. Each was then time-stretched by a factor of 5% with regards to the original tempo, yielding a total of 12 stimuli that were presented to 30 participants. Participants were asked to move along with the stimuli while being recorded with an optical motion capture system. Synchronization analysis was performed relative to the beat and the bar level of the music and four body parts. Results suggest that participants synchronized different body parts to specific metrical levels; in particular, vertical movements of hip and feet were synchronized to the beat level when the music contained large amounts of low frequency spectral flux and had a slower tempo, while synchronization of head and hands was more tightly coupled to the weak flux stimuli at the bar level. Synchronization was generally more tightly coupled to the slower versions of the same stimuli, while synchronization showed an inverted u-shape effect at the bar level as tempo increased. These results indicate complex relationships between musical characteristics, in particular regarding metrical and temporal structure, and our ability to synchronize and entrain to such musical stimuli.

## **Keywords**

Music-induced movement, motion capture, synchronization, entrainment, nonverbal behavior

## 1 Introduction

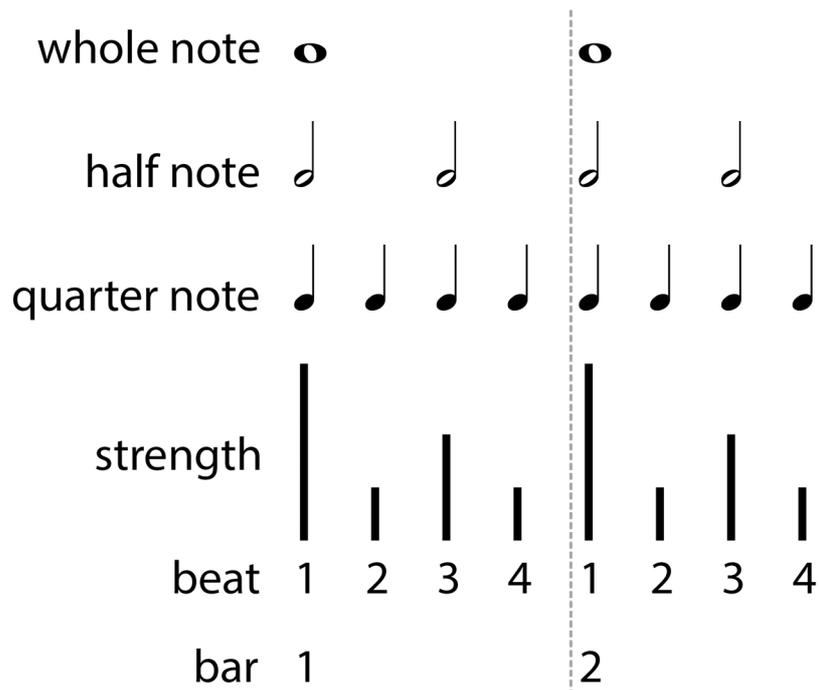
One of the striking characteristics of music is that it has the capacity to induce movements in humans, and indeed, it is often difficult to avoid overt movement when listening to music (Keller & Rieger, 2009; Lesaffre et al., 2008). Such movement responses during music listening usually happen spontaneously and can range from tapping or nodding along with the music to full-body dance movements. Typically, music-induced movement is both organized by and coordinated with the music in some fashion; for instance, people mimic instrumentalists' gestures or entrain with the beat of the music (Godøy, Haga, & Jensenuis, 2006; Leman & Godøy 2010).

Such involvement of the body contributes to the notion of embodied (music) cognition, which claims that human cognition and intelligent behavior is not merely based on passive perception, but requires goal-directed interaction between mind/brain, sensorimotor capabilities, body, and environment (e.g., Varela, Thompson, & Rosch, 1991). Music (or musical involvement) can thus be seen as linking our perception of it to our body movement, so that our bodily movements reflect, imitate, or help us to parse and understand the structure of music (Leman, 2007). Synchronizing to music could therefore be understood as a form of corporeal imitation: "Spontaneous movements [to music] may be closely related to predictions of local bursts of energy in the musical audio stream, in particular to the beat and the rhythm patterns" (Leman, 2007, p. 96).

Leman (2007) suggests several (co-existing) components or concepts of corporeal articulations, which differ in the degree of musical involvement and in the kind of action-perception couplings employed. *Synchronization* forms the fundamental component, as synchronizing to a periodic stimulus is easy and spontaneous. *Inductive resonance* refers to the use of movements for active control, imitation, and

prediction of beat-related features in the music as the first step in engaging with the music, in particular when the musical structure becomes more complex. *Embodied Attuning*, concerns the linkage of body movement to musical features more complex than the basic beat, such as melody, harmony, rhythm, tonality, or timbre. Finally, *Empathy* is seen as the component that links musical features to expressivity and social function of music, including emotional expression.

Musical *beats*, as the basic rhythmic element of music, usually occur at regular temporal intervals (typically notated at the quarter note level). They give rise to a percept of a pulse in the music and a subjective sense of periodicity. By subdividing this basic pulse into smaller units (e.g., eighths or sixteenth notes), as well as grouping pulses into larger cycles (e.g., half bar or bar levels), a metrical grid is created, with subjectively stronger and weaker events (London, 2000). For instance, in a metrical structure with a bar consisting of four beats (quadruple meter), the “one” of each bar has the strongest emphasis, followed by the “three”. Typically, the beat level and the bar level are the most essential metrical levels. A graphical representation of the metrical hierarchy is displayed in Figure 1.



**Fig. 1** Metrical hierarchy indicating the relationships between bars, beats, beat strengths, and musical notation

The theory of *dynamic attending* (e.g., Jones, 1976; Drake, Jones, & Baruch, 2000) proposes that humans, when listening to a complex auditory sequence, spontaneously focus on events occurring at a medium rate (called *referent level*, mostly correspondent to the tactus/beat level, see Jones and Boltz, 1989) and get entrained or *attuned* to this periodicity, but are also able to shift their attention to events happening at either higher or lower metrical levels and attune to these instead. This process is called *focal attending* and rests on *expectancy schemes* (corresponding to the metrical structure), allowing anticipation of successive events that belong to one or more metrical levels. Humans generally perceive and process these metrical structures easily, predicting their temporal structure and spontaneously adjusting their motor output to the sensory input (Fraisse, 1982).

The human capacity to spontaneously synchronize to rhythmic structures has been largely investigated by utilizing various finger tapping paradigms (for reviews on past and current research, see Repp (2005) and Repp and Su (2013)). Ranging from tapping to metronomes to beat-finding in complex music (e.g., Drake et al., 2000; Keller & Repp, 2004; Large, Fink, & Kelso, 2002; Snyder & Krumhansl, 2001; Toiviainen & Snyder, 2003), past research proposes that humans are able to entrain to (musical) beats spontaneously and accurately when the beat period is between 300-900 ms (Fraisse, 1982; Parncutt, 1994; van Noorden & Moelants, 1999), with a preference for beats in the 500-600 ms range. Additionally, spontaneous tapping experiments found that the majority of participants tapped at a rate of around 500-600 ms (Fraisse, 1982, Repp 2005). These findings suggest that a tempo around 500 ms – 120 beats per minute (BPM) – is a rate at which beat induction is optimal and most natural (Fraisse, 1982; Moelants, 2002). Styns, van Noorden, Moelants, and Leman (2007) found further support for a preferred tempo at around 110-120 bpm in a study investigating walking to music.

In a tempo perception study, Drake et al. (1999) found that tapping behavior was influenced by the event density of the rhythmic structure and the tapper's musical background. London (2011) also found that rhythmic stimuli with the same BPM, but different event densities, tended to be judged as being of different tempi. A tempo judgment study, conducted by London, Burger, Thompson, and Toiviainen (2016), including musical stimuli at different tempi being time-stretched by  $\pm 5\%$ , found that sped-up stimuli were rated faster than slowed-down stimuli, even though the actual tempo of a sped-up stimulus was the same or even slower than the slowed-down version of another stimulus. These results suggest that tempo judgments for real

music seem to work differently and are more complicated than judgments of simpler stimuli such as metronomes.

Neurobiological studies indicate links between rhythmic (and beat) components of music and movement, as several connections between auditory and motor systems in the brain have been observed (for overviews see Zatorre, Chen, & Penhune, 2007; Patel & Iversen, 2014). Grahn and Brett (2007) postulated that beat reproduction is mediated by motor areas, and Grahn and Rowe (2009) observed activity in the motor system even without actual movement while listening to music. Chen, Penhune, and Zatorre (2009) ran a study including three conditions: passive listening, anticipation of tapping, and actual tapping. They reported activity in different motor areas for all three conditions. Stupacher, Hove, Novembre, Schütz-Bosbach, and Keller (2013) furthermore established links between perceived groove<sup>1</sup> and motor activity in the brain. Behavioral studies have also suggested links between movement/body and rhythm/beat aspects in music. Phillips-Silver and Trainor (2008) showed that head movements could bias metrical encoding of rhythm and meter perception. Trainor, Gao, Lei, Lehtovaara, and Harris (2009) discovered that galvanic stimulation of the vestibular system could be used to disambiguate an ambiguous metric pattern. Moreover, Todd, O'Boyle, and Lee (1999) claimed that pulse perception inevitably requires motor system activity, since pulse is an inherently sensorimotor phenomenon. Collectively, these studies suggest that there is a predisposition for movement when listening to music and that humans furthermore prefer music that facilitates synchronization or entrainment and respond to it with movement (Madison, Gouyon, Ullén, & Hörnström, 2011).

---

<sup>1</sup> Groove is commonly described as “the urge to move in response to music, combined with the positive affect associated with the coupling of sensory and motor processes” (Janata, Tomic, & Haberman, 2012, p. 54).

Despite the considerable amount of literature on beat perception, tapping, and synchronization, far fewer studies have investigated whole-body movement and synchronization. Janata, Tomic, and Haberman (2012) found that participants, when asked to tap to musical stimuli, not only moved the finger/hand, but also other body parts, such as feet and head. Additionally, the more “natural” the tapping condition (isochronous versus free tapping), the more movement was exhibited. This suggests a proclivity towards active bodily responses to music, instead of merely passive listening. Zentner and Eerola (2010) explored infants' capabilities to corporeally synchronize with musical stimuli, finding that infants exhibited more rhythmic movement to music and metronome stimuli than to speech. This could suggest a predisposition for rhythmic movement to music and other metrically regular sounds. Similarly, toddlers were found to synchronize to music with three main types of periodic movement being at times synchronized with the musical pulse (Eerola, Luck, & Toiviainen, 2006).

For music-induced movements in adults, Toiviainen, Luck, and Thompson (2010) showed that eigenmovements (i.e., the Principal Components obtained from a participant-level Principal Component Analysis on marker position data) were synchronized with different metrical levels of the stimulus. Naveda and Leman (2010) studied how the metric hierarchy in Samba and Charleston is represented in repetitive gestures of professional dancers. Van Dyck et al. (2013) found that listeners responded with more spontaneous movements to the increasing presence of the bass drum. Burger, Thompson, Saarikallio, Luck, and Toiviainen (2013) were further able to show that beat- and rhythm-related musical characteristics, such as pulse clarity and spectral flux in low and high frequency ranges, influenced participants' movements to music; strong low frequency spectral flux, for instance, resulted in

increased speed of movement, whereas tempo failed to exhibit any relationship to movement features. Burger, Thompson, Luck, Saarikallio, and Toiviainen (2014) studied period locking of music-induced movement to musical stimuli and found that strong low frequency spectral flux increased period-locking to the beat and the bar level of the music. Moreover, Luck and Toiviainen (2006) as well as Luck and Sloboda (2008, 2009) found that ensembles synchronize to maximal deceleration of the baton of the conductor, and that acceleration was the best predictor of the beat location along the movement trajectory.

Low frequency content of music could play a crucial role in motor attunement and sensory-motor synchronization. Hove, Marie, Bruce, and Trainor (2014) found that time perception is better for lower musical pitch ranges and showed that tapping synchronization was better in lower-pitch sequences. Furthermore, Todd, Rosengren, and Colebatch (2008) found that the human vestibular system is sensitive to low frequency vibration. Stupacher, Hove, and Janata (2016) could predict perceived groove ratings from audio features, such as low frequency spectral content, and found higher groove ratings as well as more accurate tapping behavior with lower frequencies of bass instruments. Strong low frequency cues have also been found to relate to rhythmic structure, rhythmic strength and the propensity to move (Burger, Ahokas, Keipi, & Toiviainen, 2013).

In the studies mentioned above (Burger, Thompson, et al., 2013, Burger et al., 2014, Stupacher et al., 2016), low frequency spectral content has been quantified in terms of spectral flux (Alluri & Toiviainen, 2010). Spectral flux indicates how much the acoustic energy in a given auditory spectrum changes over time across frequencies by calculating the distance of the spectra between two consecutive time points over the course of the stimuli. Large amounts of flux in the low frequency bands are, for

instance, produced by strong rhythmic elements performed by instruments such as kick drum or bass guitar. Dance-music genres such as Techno, as well as Hard Rock or Heavy Metal, usually contain large amounts of low frequency spectral flux due to the strong presence of these low frequency instruments, whereas “softer” music (e.g., Ambient or Folk) would contain fewer strong low frequency components and therefore smaller amounts of low frequency spectral flux. In addition to the low frequency content, strong low frequency spectral flux is produced by high event density and sharp attacks (steep attack slopes).

The present work explores synchronization in music-induced movement, in particular the effects of timing, tempo, and low frequency spectral content. Previous studies have shown relationships between rhythm- and beat-related musical characteristics and movement features (Burger, Thompson, et al., 2013), specifically movement periodicities and period-locking in free movement as well as phase-locking in constrained motion (Burger et al., 2014), and eigenmovements being synchronized to different metrical levels (Toiviainen et al., 2010). The present study aims to further explore the phase-locking/synchronization ability of humans in quasi-spontaneous, music-induced movement, with particular focus on the movements of different body parts. An accurate phase-locking measure has been developed in Burger et al. (2014), which will be refined here with regards to complex whole-body movement data in three dimensions for use in statistical analysis.

The present study will systematically investigate synchronization ability regarding low frequency spectral flux, as well as absolute and relative tempo, using stimuli containing either high or low amounts of low frequency spectral flux to investigate the influence of spectral flux on synchronization. Moreover, this study aims to control for the effect of tempo. In Burger, Thompson, et al. (2013), no

relationship of tempo was found regarding musical characteristics, whereas Burger et al. (2014) found an interaction between tempo and metrical levels with regards to period-locking. This study will categorize stimulus tempo into three core “absolute” tempo levels, 105, 115, and 130 BPM, which together cover the slower, middle and faster ends of the preferred tempo range as well as the range of maximum tempo stability and accuracy (Fraisse, 1982; Moelants, 2002), to systematically investigate the role of overall tempo on synchronization ability.

To study relative tempo perception and embodiment, a time-stretch factor is further entered into the present study, so that each stimulus is presented at +5% and –5% of the core tempo as well as at its original tempo<sup>2</sup>. Including both absolute and time-stretched tempo differences offers a fine-grained sampling of the range of tempi and larger set of stimuli while allowing for precise control and stability of other stimulus characteristics, in particular the differences in low frequency spectral flux, as well as for looking at “local” (time-stretched) versus “global” (song-to-song) differences in tempo and their effects on movement behaviors. Furthermore, the different tempi are expected to reduce fatigue and carryover effects in the sense that participants would get attuned to one tempo without paying attention to other musical characteristics.

Based on the literature reviewed above, we formulated the following hypotheses:

Hypothesis 1) Synchronization accuracy differs between metrical levels, in particular regarding the body parts. Hip and foot movement is assumed to be more synchronized to the beat/tactus level (i.e., embodying the basic rhythmic structure), whereas hand movement in particular should synchronize to the bar level, as they

---

<sup>2</sup> The performances to the original tempo stimuli belonged to a different experiment condition and are therefore not included in this analysis, as this analysis shall give insights into effects of relative and absolute tempo differences. Moreover, these stimuli were subsequently used in a perceptual experiment investigating the abilities to judge tempo based on point-light dance animations (London et al., 2016).

have higher degrees of freedom and thus could express movements related to a longer time span (Burger, Thompson, et al., 2013, Burger et al., 2014; Jones, 1976; Leman, 2007).

Hypothesis 2) Low frequency spectral flux influences synchronization ability such that stronger low frequency flux results in more accurate synchronization, independently of metrical level or absolute tempo (Burger et al., 2014; Hove et al., 2014; Stupacher et al., 2016).

Hypothesis 3) Time-stretching the stimuli ( $\pm 5\%$ ) affects synchronization relative to the preferred tempo (110/120 BPM / Fraise, 1982; Moelants, 2002); in case of the slow core tempo (105 BPM), the sped-up version would be better synchronized than the slowed-down version, while in case of the fast core tempo (135 BPM), the slowed-down version would result in better synchronization than the sped-up version (i.e., there would be an inverted u-shape relationship between relative tempo and synchronization ability).

Hypothesis 4) Taken that the three absolute stimuli tempi are all within the preferred tempo range, we assumed no significant overall differences in synchronization on the song level. This hypothesis would be in line with results by Burger, Thompson, et al. (2013) that overall tempo within this range does not affect synchronization.

## **2 Method**

### **2.1 Participants**

Thirty participants took part in the study (15 female, 15 male, average age: 28.2, SD of age: 4.4). Participants were university students of 15 different nationalities. Four participants had received professional music education. Twenty-two participants had

undergone music education as children or adults, of which 13 were still actively playing an instrument or singing. Fourteen participants had taken dance lessons of various styles. Participation was rewarded with a movie ticket (value  $\approx$  10 Euros).

## **2.2 Stimuli**

The stimulus material consisted of the first 35 seconds from six classic Motown/Rhythm and Blues (R&B) songs from the mid 1960s to the early 1970s (see Table 1). The songs were chosen for their danceability (related to the notion of groove, e.g., Janata et al., 2012), their homogeneity (being of the same music genre and similar in time of release) and their ubiquity in popular music culture (all are considered R&B classics). All songs employed simple duple meters with light to moderate amounts of swing; while the duplet divisions of the beat were not played perfectly straight, there were no overt triplet divisions of the beat (i.e., no “shuffle” rhythms).

Furthermore, the songs were divided into three distinct core tempo groups of around 105, 115, and 130 BPM and slightly time-stretched (using Audacity ver. 2.0.5) to match these BPM rates (see Table 1 for the exact original and final tempi). This tempo range was chosen as it fell within the preferred beat rate for spontaneous tempo (Fraisse, 1982), and could thus be expected to afford a comfortable range of movement.

At each tempo group, one stimulus contained a high amount of low frequency spectral flux, whereas the other had low amount of low frequency spectral flux (as estimated computationally by using the MIRTtoolbox (Lartillot & Toivainen, 2007) in MATLAB). For the calculation, the stimulus is divided into 10 frequency bands, each band containing one octave in the range of 0–22 050 Hz. The sub-band flux is then

calculated for each of these ten bands separately by calculating the Euclidean distances of the spectra for each two consecutive frames of the signal (Alluri & Toiviainen, 2010), using a frame length of 25 ms and an overlap of 50% between successive frames and then averaging the resulting time series of flux values. To assess the low frequency spectral flux, we used sub-band no. 3 (100-200 Hz), as the differences in values were largest. Nevertheless, the other low sub-bands (below sub-band 6) showed similar patterns. Other rhythm-related acoustic features, such as attack length or slope showed more equally distributed patterns, thus differed less across stimuli.

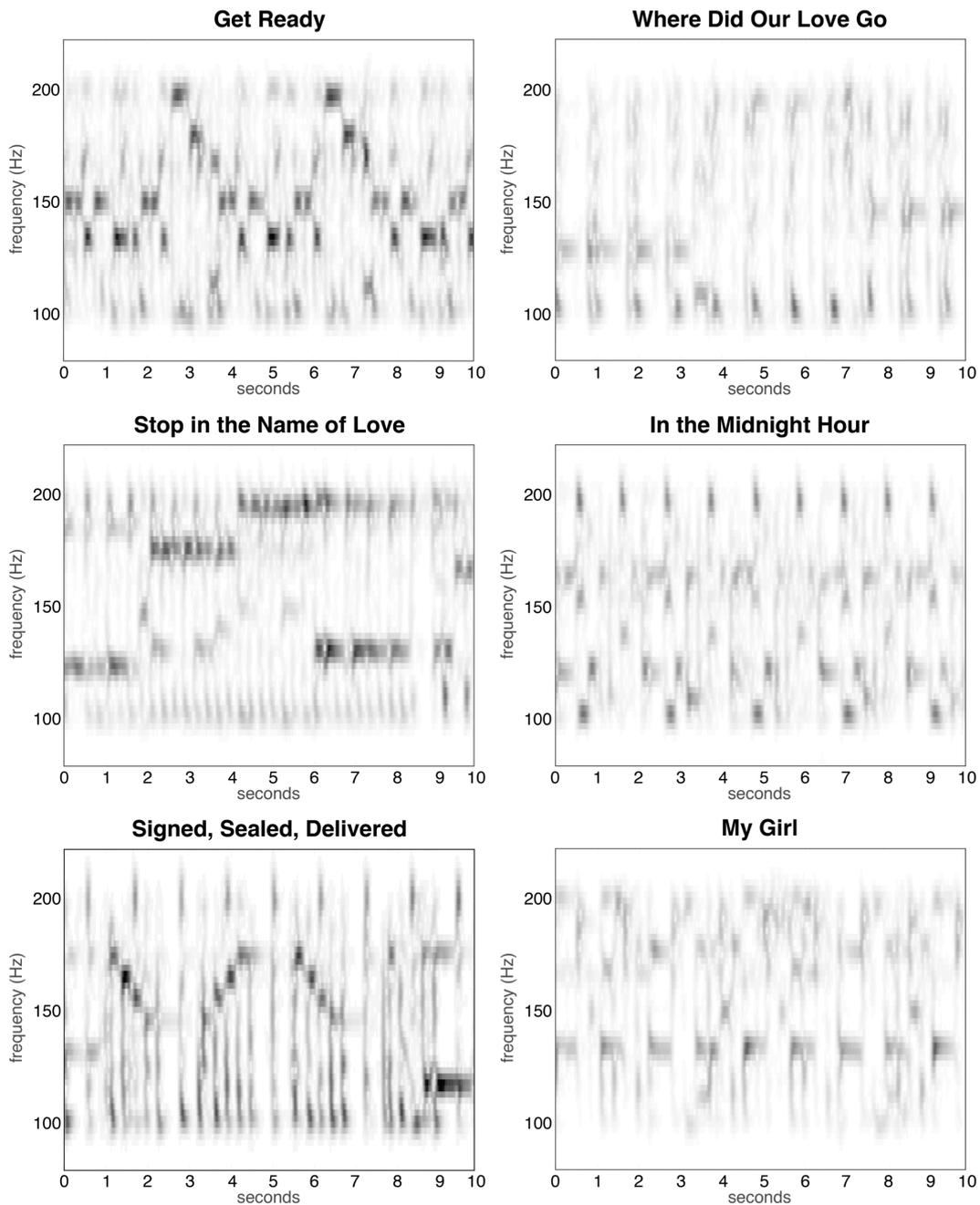
Figure 2 shows spectrograms of the six stimuli to illustrate low frequency spectral flux, whereas Table 1 shows the averaged spectral flux values of sub-band 3 for all six stimuli. As can be seen in the spectrograms, the stimuli containing strong low frequency flux (left side, Fig. 2) have stronger components in the low frequency range (larger amount of dark areas) than the stimuli containing weaker low frequency components (right side, Fig. 2), and they perceptually contain stronger low frequency rhythmic elements (especially kick drum and bass guitar). The loudness of the stimuli was normalized, resulting in similar RMS levels.

Finally, the stimuli were time-stretched a second time to produce tactus rates at  $\pm 5\%$  of the three core rates, resulting in a slow and a fast version of each song, each slightly shorter or longer than the original. The stimulus length was chosen to keep the experiment sufficiently short while having stimuli that were long enough to induce movement.

**Table 1** Overview of the six Motown songs used in the experiment

<b>Artist</b>	<b>Song</b>	<b>Orig.</b>	<b>Core</b>	<b>Slow</b>	<b>Fast</b>	<b>Flux</b>	<b>Mean</b>
---------------	-------------	--------------	-------------	-------------	-------------	-------------	-------------

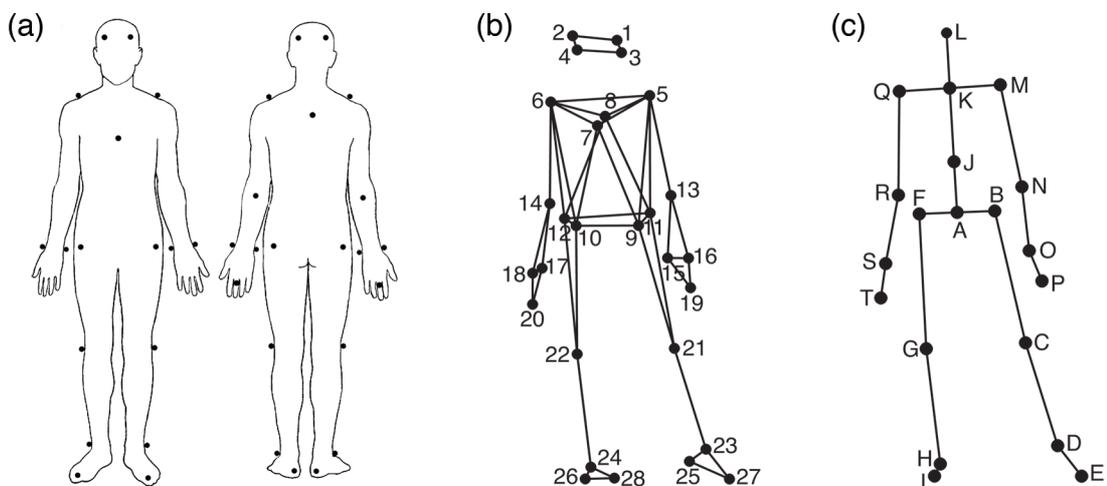
		<b>BPM</b>	<b>BPM</b>	<b>BPM</b>	<b>BMP</b>	<b>level</b>	<b>Flux</b>
Temptations	Get Ready	134.5	130	123.5	136.5	strong	41.12
Supremes	Where Did Our Love Go?	133	130	123.5	136.5	weak	13.38
Supremes	Stop in the Name of Love	117	115	109.25	120.75	strong	23.07
Wilson Pickett	In the Midnight Hour	113	115	109.25	120.75	weak	7.52
Stevie Wonder	Signed, Sealed, Delivered	105.5	105	99.75	110.25	strong	26.09
Temptations	My Girl	103	105	99.75	110.25	weak	9.90



**Fig. 2** 100-200 Hz Spectrograms of the six stimuli (10 second excerpts starting from the first down beat of the music). The left side panels display the stimuli containing strong flux levels, and the right side panels display stimuli containing low flux levels

### 2.3 Apparatus

Participants' movements were recorded using an eight-camera optical motion capture system (Qualisys Oqus 5+, [www.qualisys.com](http://www.qualisys.com)), tracking, at a frame rate of 120 Hz, the three-dimensional positions of 28 reflective markers attached to each participant. The locations of the markers can be seen in Figure 3A and 3B. The location of the markers were as follows (L = left, R = right, F = front, B = back): 1: LF head; 2: RF head; 3: LB head; 4: RB head; 5: L shoulder; 6: R shoulder; 7: sternum; 8: spine (T5); 9: LF hip; 10: RF hip; 11: LB hip; 12: RB hip; 13: L elbow; 14: R elbow; 15: L wrist/radius; 16: L wrist/ulna; 17: R wrist/radius; 18: R wrist/ulna; 19: L middle finger; 20: R middle finger; 21: L knee; 22: R knee; 23: L ankle; 24: R ankle; 25: L heel; 26: R heel; 27: L big toe; 28: R big toe. The musical stimuli were played back via a pair of Genelec 8030A loudspeakers using a Max patch ([www.cycling74.com](http://www.cycling74.com)) running on an Apple computer. The direct (line-in) audio signal of the playback and the synchronization pulse transmitted by the Qualisys cameras when recording, were recorded using ProTools software (<http://www.avid.com/pro-tools>) in order to synchronize the motion capture data with the musical stimulus afterwards. Additionally, a video camera was used to record the sessions for reference purposes.



**Fig. 3** Marker and joint locations. (a) Anterior and posterior view of the marker placement on the participants' bodies; (b) Anterior view of the marker locations as stick figure illustration; (c) Anterior view of the locations of the secondary markers/joints used in the analysis

## **2.4 Procedure**

Participants were recorded individually while being asked to imagine being in a social setting such as a club or disco. The six Motown songs were presented in random order for each participant in blocks including both versions of each particular stimulus, in an order that was counterbalanced among the participants. Participants were asked to dance freely and were further advised to remain synchronized to the music and stay in the capture area marked on the floor (appr. 3 x 4 m). They were free to rest whenever they wished during the experiment; experimental trials took an average of 45 minutes.

## **2.5 Movement data processing**

Using the Motion Capture (MoCap) Toolbox (Burger & Toiviainen, 2013) in MATLAB, movement data of the 28 markers were first trimmed to match the exact duration of the stimuli. Gaps in the data were linearly filled – such gaps happened due to markers being occasionally occluded, but were very short (less than 250 ms), so locations of missing frames could easily be inferred by interpolation.

Following this, the data were transformed into a set of 20 secondary markers – subsequently referred to as joints. The locations of these 20 joints are depicted in Figure 3C. The locations of joints C, D, E, G, H, I, M, N, P, Q, R, and T are identical to the locations of one of the original markers, while the locations of the remaining joints were obtained by averaging the locations of two or more markers; Joint A:

midpoint of the four hip markers (root); B: midpoint of markers 9 and 11 (left hip); F: midpoint of markers 10 and 12 (right hip); J: midpoint of breastbone, spine, and the hip markers (midtorso); K: midpoint of shoulder markers (manubrium), L: midpoint of the four head markers (head); O: midpoint of the two left wrist markers (left wrist); S: mid-point of the two right wrist markers (right wrist).

Subsequently, the data were rotated, so that the hip joints (A, B, and F) were aligned to be parallel to the x-axis on average, and transformed to a local coordinate system with the root (joint A) as the origin. Next, acceleration in three dimensions of six joints (head, left and right hand, left and right foot in the local coordinate system, and hip/root in the global coordinate system) was calculated using numerical differentiation and a Butterworth smoothing filter (second order zero-phase digital filter) and averaged across contralateral joints (in the case of hands and feet), resulting in four body parts times three dimensions. These four body parts represent the movements of the body's center-of-mass and its extremities, as movement data from adjacent/intervening joints are usually highly correlated. Acceleration data was used, as it is more stationary than location data (i.e., the average of the data does not change over time in a windowed analysis, since the acceleration data centers around 0, unlike, for instance, location data), and has been used in earlier synchronization and timing studies (Burger et al., 2014; Luck and Sloboda, 2009; Toiviainen et al., 2010).

Subsequently, a synchronization/phase-locking analysis was performed relative to two metrical levels, the bar level and the beat level. These two were chosen because they have been found to be the two most prominent metrical levels in spontaneous movement (Burger et al., 2014; Toiviainen et al., 2010). Furthermore, the number of combinations between metrical levels and movement descriptors (body parts x dimensions) was limited to medio-lateral (sideways) movement with respect to the bar

level and superior-inferior (vertical) movement with respect to the beat level, which has been previously shown to be the most prominent movement directions at the respective metrical levels (Burger et al., 2014; Toiviainen et al., 2010).

This resulted in altogether eight movement features (four body parts x two metrical levels/dimensions). The decision to focus on medio-lateral movement at bar level and superior-inferior movement at beat level is also data-driven, since an initial periodicity analysis revealed that the other combinations of movement dimensions and metrical levels were only infrequently period-locked to the music, which is a prerequisite for synchronized movement.

In order to address phase locking, the phase of the movement for each respective body part and metrical level was estimated by band-pass filtering the movement data with a zero-phase FFT filter at the frequency corresponding to the beat length of the metrical level, using a bandwidth of 15% of the center frequency and subsequently applying a Hilbert transform. This yielded the movement phase relative to the respective metrical level as a function of time. In a next step, the time points of the musical beats for all six stimuli were manually annotated using SonicVisualizer ([www.sonicvisualiser.org](http://www.sonicvisualiser.org)). In order to compare the movement phase with the beat locations in the music, the phase of the musical beat at the two metrical levels was then estimated by linearly interpolating between the manually annotated beat locations. Next, we trimmed both movement and music data to a length of 20 seconds starting at the first downbeat (approximately after 5 to 7 seconds after the start) of each stimulus to avoid possible artifacts in the beginning and end of each performance due to the calculations and to further account for participants needing time to initially lock into the musical beat.

To assess the relationship between the movement and the musical beat, the difference between the movement and the musical phase was calculated over time. In order to acquire a statistical measure that quantifies the degree of phase locking, or *synchronization ability* to a given musical stimulus, the negentropy (i.e., the additive inverse of the Shannon entropy – Shannon, 1948) was taken from this difference distribution as a measure of synchronization accuracy at the different metrical levels. Entropy measures have been successfully used in studies on neural synchrony (for instance using EEG: Le Van Quyen et al., 2001 or MEG/EMG: Tass et al., 1998) to quantify the strength of the synchronization and derive a measure suitable for statistical approaches. One advantage of using entropy is that it is a robust measure in detecting both in- and anti-phase synchronization due to its non-linearity (unlike other measures such as the mean absolute difference). Furthermore, the measure is assumption- and parameter-free. Shannon entropy is defined as:

$$H(X) = - \sum_{i=1}^n P(x_i) \log P(x_i)$$

with P being the probability mass function. Normalized Shannon Entropy,  $H_n$ , is obtained by dividing H by  $\log(n)$ , resulting in a range between 0 and 1. The normalized negentropy was subsequently calculated as follows:

$$J_n(X) = 1 - H_n(X)$$

Regarding our measure of synchronization ability, a higher negentropy value corresponds more accurate synchronization to the music, whereas a lower negentropy value equals less accurate synchronization.

### 3 Results

In order to assess synchronization ability of participants, in particular the effects of low frequency spectral flux (in the following referred to as *Flux*), core tempo levels (in the following referred to as *Tempo*), and tempo stretching (in the following referred to as *Time-stretch*) of musical stimuli on synchronization behavior, a variety of descriptive and inferential statistical analyses were conducted. Analysis focused on vertical (superior-inferior) movement relative to the beat level of the music and horizontal (medio-lateral) movement relative to the bar level of the music of four different body parts being hips, head, feet, and hands.

### **3.1 Beat level**

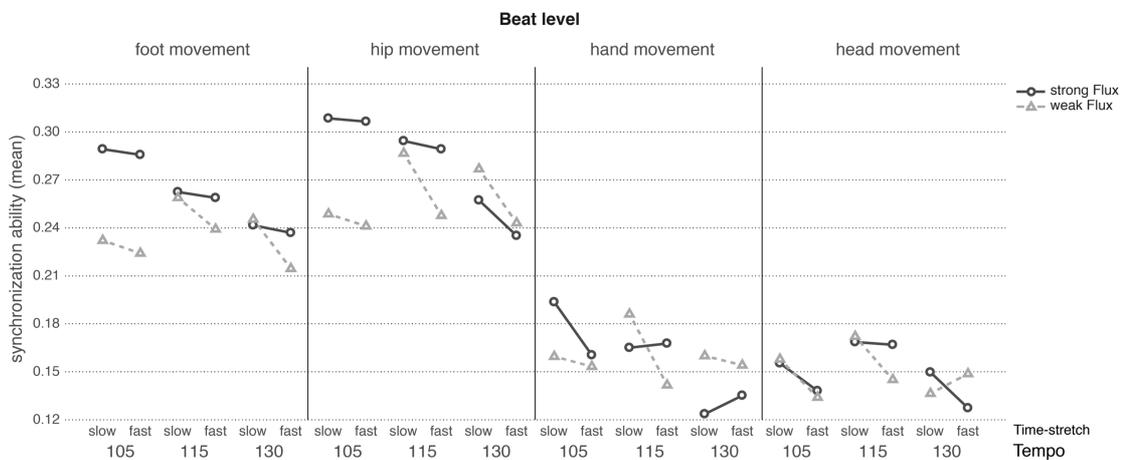
To assess effects of low frequency spectral flux, core tempo levels, and time-stretch at the beat level, separate three-factor repeated measures ANOVAs were conducted (see Table 2). For foot synchronization, the repeated measures ANOVA resulted in a significant main effect for Flux as well as in a significant interaction for Flux x Tempo. The repeated measures ANOVA for hip synchronization resulted in significant main effects for Flux and Time-stretch as well as a significant interaction for Flux x Tempo. For hand synchronization, a significant main effect was found for Tempo as well as a significant interaction for Flux x Tempo, while the repeated measures ANOVA for head synchronization yielded a significant main effect for Tempo. A graphical (descriptive) overview of the three-factor outcomes is given in Figure 4. Figure 4 also shows that synchronization ability was overall higher for foot and hip movement, and thus participants were more accurately synchronized with feet and hips than with hands and head. Three additional trends are visible in Figure 4 (in line with the ANOVA results): firstly, synchronization ability tended to be more accurate for the strong Flux stimuli than for the weak Flux stimuli (in particular for

foot and hip movement); secondly, synchronization ability decreased with faster Tempo, especially for strong Flux stimuli; and thirdly, synchronization ability tended to be more accurate for the slowed-down than the sped-up versions of the stimuli.

**Table 2** Numerical results of the main effects and interactions of the three-factor repeated measures ANOVAs at the beat level (factors: Flux: low frequency spectral flux; Tempo: core tempo; Time-stretch: tempo stretching of stimuli by  $\pm 5\%$ ). Significant effects ( $p < .05$ ) are indicated in bold

Body part	Effects	<i>F</i>	df	<i>p</i>	$\eta_p^2$
<b>Foot</b>	<b>Flux</b>	<b>11.16</b>	<b>1, 29</b>	<b>.002</b>	<b>.28</b>
	Time-stretch	1.95	1, 29	.173	.06
	Tempo	1.95	2, 58	.151	.06
	Flux x Time-stretch	1.92	1, 29	.176	.06
	<b>Flux x Tempo</b>	<b>3.37</b>	<b>2, 58</b>	<b>.041</b>	<b>.10</b>
	Time-stretch x Tempo	0.23	2, 58	.793	.01
	Flux x Time-stretch x Tempo	0.22	2, 58	.805	.01
<b>Hip</b>	<b>Flux</b>	<b>10.00</b>	<b>1, 29</b>	<b>.004</b>	<b>.26</b>
	<b>Time-stretch</b>	<b>5.46</b>	<b>1, 29</b>	<b>.027</b>	<b>.16</b>
	Tempo	2.58	2, 58	.084	.08
	Flux x Time-stretch	2.59	1, 29	.118	.08
	<b>Flux x Tempo</b>	<b>6.55</b>	<b>2, 58</b>	<b>.003</b>	<b>.18</b>
	Time-stretch x Tempo	1.19	2, 58	.310	.04
	Flux x Time-stretch x Tempo	0.48	2, 58	.622	.02

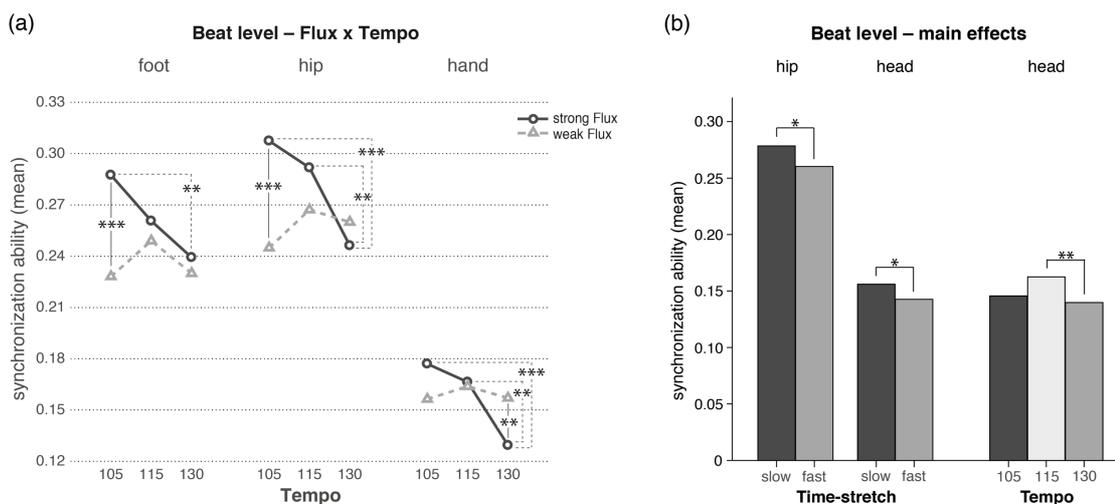
<b>Hand</b>	Flux	0.04	1, 29	.837	.00
	Time-stretch	3.66	1, 29	.066	.11
	<b>Tempo</b>	<b>4.61</b>	<b>2, 58</b>	<b>.014</b>	<b>.14</b>
	Flux x Time-stretch	1.55	1, 29	.224	.05
	<b>Flux x Tempo</b>	<b>5.37</b>	<b>2, 58</b>	<b>.007</b>	<b>.16</b>
	Time-stretch x Tempo	1.99	2, 58	0.145	.06
	Flux x Time-stretch x Tempo	2.88	2, 58	.064	.09
<b>Head</b>	Flux	0.14	1, 29	.711	.01
	<b>Time-stretch</b>	<b>6.57</b>	<b>1, 29</b>	<b>.016</b>	<b>.19</b>
	<b>Tempo</b>	<b>4.18</b>	<b>2, 58</b>	<b>.020</b>	<b>.13</b>
	Flux x Time-stretch	0.01	1, 29	.921	.00
	Flux x Tempo	0.38	2, 58	.689	.01
	Time-stretch x Tempo	0.43	2, 58	.654	.02
	Flux x Time-stretch x Tempo	2.56	2, 58	.086	.08



**Fig. 4** Line plots representing the synchronization ability averages on the beat level per factor as used in the three-factor repeated measures ANOVAs (factors: Flux: low frequency spectral flux; Tempo: core tempo; Time-stretch: tempo stretching of stimuli by  $\pm 5\%$ ). The y-axis indicates the averaged synchronization ability (the higher the value, the better the synchronization between movement and beat level of the music). The x-axis is divided into the slow and fast time-stretched pairs (Time-stretch), grouped by the core tempo levels (Tempo). Strong and weak low frequency spectral flux stimuli are indicated with a dark gray circle/solid line (strong Flux) and a light gray triangle/dashed line (weak Flux)

Following up the results of the three-way repeated measures ANOVAs, we first analyzed the significant interactions followed by the remaining significant main effects. We found significant interactions between Flux and Tempo for foot, hip and head synchronization (for a graphical overview, see Figure 5A). Pairwise comparisons (paired samples t-tests using the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995) to control the false discovery rate set at  $p < .05$ ) between the different Flux and Tempo levels revealed significant differences between weak versus strong Flux for slow Tempo,  $t(29) = 4.41$ ,  $p < .001$ , for foot movement, with strong Flux showing more accurate synchronization than weak Flux. Furthermore for foot movement, there was a significant difference between slow and fast Tempo for strong Flux,  $t(29) = 2.92$ ,  $p = .007$ , with synchronization being more accurate for the slow stimuli than for the fast stimuli. In case of hip movement, significant differences were found between weak and strong Flux for slow Tempo,  $t(29) = 4.55$ ,  $p < .001$ , with more accurate synchronization for the strong Flux stimuli, as well as between slow and fast Tempo and mid and fast Tempo for strong Flux,

$t(29) = 3.91, p < .001$ , and  $t(29) = 3.12, p = .004$ , with the slow and mid tempo being both more accurately synchronized than the fast stimuli. For hand movement, significant differences were found between weak and strong Flux for fast Tempo,  $t(29) = -2.91, p = .007$ , with weak Flux being more accurately synchronized, as well as between slow and fast Tempo and mid and fast Tempo for strong Flux,  $t(29) = 3.99, p < .001$ , and  $t(29) = 3.21, p < .003$ , with the slow and mid Tempo stimuli being more accurately synchronized than the fast stimuli. Furthermore, significant main effects were found for the Time-stretch factor for hip movement as well as for head movement, with more accurate synchronization for the slowed-down stimuli (-5 BPM) than for the sped-up stimuli (+5 BPM) in both cases,  $t(29) = 2.34, p = .027$  and  $t(29) = 2.56, p < .016$ . For head movement, there was furthermore a main effect for Tempo with a significant difference between the mid and the fast Tempo level, with the mid Tempo stimuli showing more accurate synchronization than the fast Tempo stimuli ( $p = .006$ , pair-wise comparisons adjusted with Bonferroni correction). A graphical overview of the main effects is displayed in Figure 5B.



**Fig. 5** Synchronization ability averages (y-axis; the higher the value, the better the synchronization between movement and beat level of the music) for significant

interactions (a) and main effects (b) with \*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$  (factors: Flux: low frequency spectral flux; Tempo: core tempo; Time-stretch: tempo stretching of stimuli by  $\pm 5\%$ ). The remaining differences are non-significant

### 3.2 Bar level

To assess effects of low frequency spectral flux, core tempo levels, and time-stretch at the bar level, separate three-factor repeated measures ANOVAs were ran on each body part (see Table 3). For foot synchronization, a significant main effect was found for Tempo, while for hip synchronization, significant main effects were found for Time-stretch and Tempo. The repeated measures ANOVA for hand synchronization resulted in significant main effects for Flux and Tempo. In case of head synchronization, the ANOVA resulted in significant main effects for Flux and Tempo as well as in a significant interaction for Flux x Tempo. A graphical (descriptive) overview of the three-factor outcomes is given in Figure 6. In contrast to the results at the beat level, synchronization ability was similar across the four body parts at the bar level. Furthermore, three trends are visible from Figure 6 (in line with the ANOVA results): firstly, synchronization ability tended to be more accurate for the weak Flux stimuli than for the strong Flux stimuli; secondly, synchronization ability seemed most accurate at the mid Tempo level; and thirdly, synchronization ability tended to be more accurate for the slowed-down than the sped-up versions of the stimuli.

**Table 3.** Numerical results of the main effects and interactions of the three-factor repeated measures ANOVAs at the bar level (factors: Flux: low frequency spectral

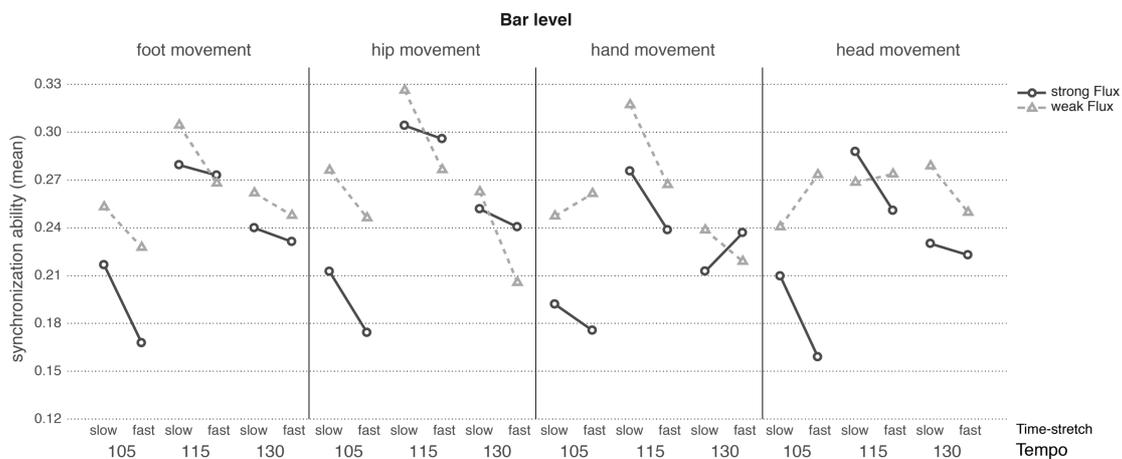
flux; Tempo: core tempo; Time-stretch: tempo stretching of stimuli by  $\pm 5\%$ ).

Significant effects ( $p < .05$ ) are indicated in bold

<b>Body part</b>	<b>Effects</b>	<b><i>F</i></b>	<b><i>df</i></b>	<b><i>p</i></b>	<b><math>\eta_p^2</math></b>
<b>Foot</b>	Flux	4.19	1, 29	.050	.13
	Time-stretch	2.71	1, 29	.111	.09
	<b>Tempo</b>	<b>12.10</b>	<b>2, 58</b>	<b>&lt; .001</b>	<b>.29</b>
	Flux x Time-stretch	0.04	1, 29	.850	.00
	Flux x Tempo	0.98	1.60, 46.43*	.367	.03
	Time-stretch x Tempo	0.44	2, 58	.648	.02
	Flux x Time-stretch x Tempo	0.44	2, 58	.647	.02
<b>Hip</b>	Flux	2.20	1, 29	.149	.07
	<b>Time-stretch</b>	<b>8.11</b>	<b>1, 29</b>	<b>.008</b>	<b>.22</b>
	<b>Tempo</b>	<b>9.62</b>	<b>2, 58</b>	<b>&lt; .001</b>	<b>.25</b>
	Flux x Time-stretch	1.56	1, 29	.222	.05
	Flux x Tempo	3.15	2, 58	.050	.10
	Time-stretch x Tempo	0.02	2, 58	.982	.00
	Flux x Time-stretch x Tempo	0.41	2, 58	.664	.01
<b>Hand</b>	<b>Flux</b>	<b>7.76</b>	<b>1, 29</b>	<b>.009</b>	<b>.21</b>
	Time-stretch	1.20	1, 29	.282	.04
	<b>Tempo</b>	<b>6.19</b>	<b>2, 58</b>	<b>.004</b>	<b>.18</b>
	Flux x Time-stretch	0.15	1, 29	.706	.01
	Flux x Tempo	3.09	2, 58	.053	.09
	Time-stretch x Tempo	1.41	2, 58	.252	.05

	Flux x Time-stretch x	0.76	2, 58	.471	.03
	Tempo				
<b>Head</b>	<b>Flux</b>	<b>7.44</b>	<b>1, 29</b>	<b>.011</b>	<b>.20</b>
	Time-stretch	1.71	1, 29	.201	.06
	<b>Tempo</b>	<b>5.33</b>	<b>2, 58</b>	<b>.007</b>	<b>.16</b>
	Flux x Time-stretch	1.77	1, 29	.193	.06
	<b>Flux x Tempo</b>	<b>4.14</b>	<b>2, 58</b>	<b>.021</b>	<b>.13</b>
	Time-stretch x Tempo	0.06	2, 58	.946	.00
	Flux x Time-stretch x	1.34	2, 58	.271	.04
	Tempo				

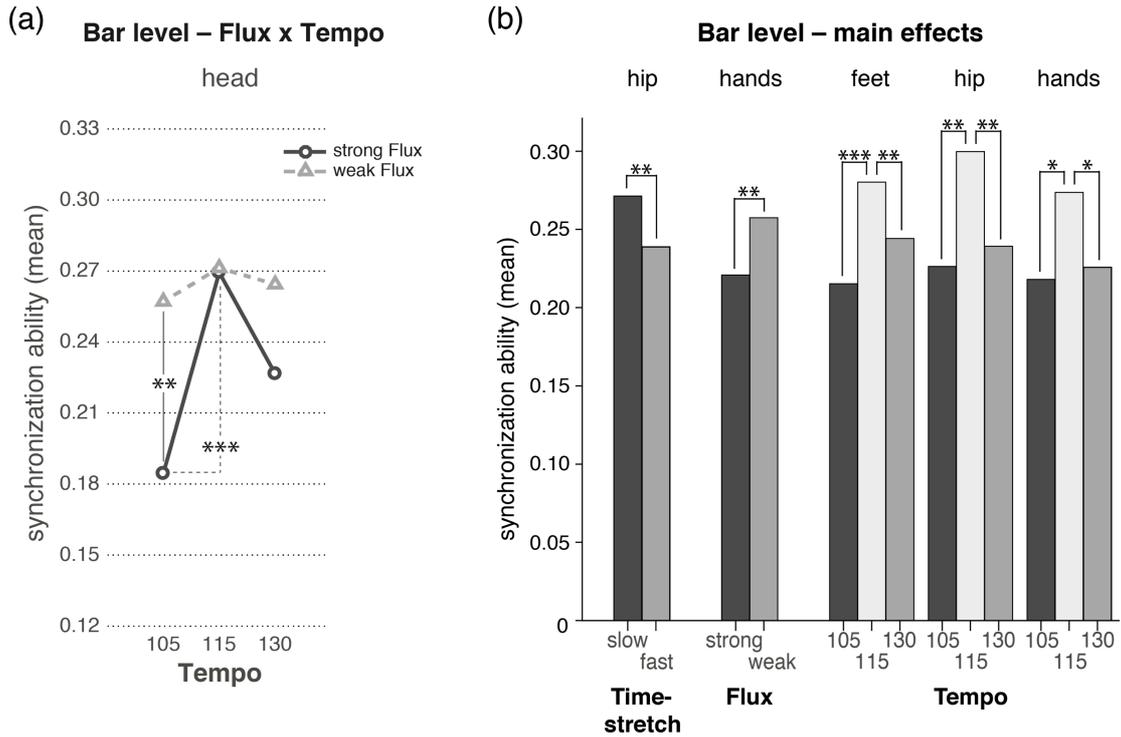
\* Greenhouse-Geissner correction applied due to violation of sphericity assumption.



**Fig. 6.** Line plots representing the synchronization ability averages per factor on the bar level as used in the three-factor repeated measures ANOVAs (factors: Flux: low frequency spectral flux; Tempo: core tempo; Time-stretch: tempo stretching of stimuli by  $\pm 5\%$ ). The y-axis indicates the averaged synchronization ability (the higher the value, the better the synchronization between movement and bar level of the music). The x-axis is divided into the slow and fast time-stretched pairs (Time-stretch),

grouped by the core tempo levels (Tempo). Strong and weak low frequency spectral flux stimuli are indicated with a dark gray circle/solid line (strong Flux) and a light gray triangle/dashed line (weak Flux)

When further investigating the significant interaction Flux x Tempo for head movement, the pairwise comparison (paired samples t-tests using the Benjamini-Hochberg procedure to control the false discovery rate set at  $p < .05$ ) indicated a significant difference between strong versus weak Flux at the slow Tempo level,  $t(29) = -3.11$ ,  $p = .003$ , with weak Flux being more accurately synchronized than strong Flux, and a significant difference between slow and mid Tempo for strong Flux,  $t(29) = -4.15$ ,  $p < .001$ , with more accurate synchronization at mid Tempo than at slow Tempo (see Figure 7A). Moreover, following up the significant main effect for Time-stretch for hip movement, participants synchronized more accurately to the slowed-down stimuli (-5 BPM) than to the sped-up ones (+5 BPM),  $t(29) = 2.85$ ,  $p = .008$ . In case of the significant main effect for Flux for hand movement, the subsequent analysis indicated more accurate synchronization for the weak Flux stimuli compared to the strong Flux ones,  $t(29) = -2.79$ ,  $p = .009$ . For the three significant main effects for Tempo (foot, hip and hand movement), pairwise comparisons (using Bonferroni correction) indicated significant differences between the slow and the mid Tempo (foot:  $p < .001$ , hip:  $p = .002$ , hand:  $p = .013$ ) as well as between the mid and the fast Tempo (foot:  $p = .009$ , hip:  $p = .005$ , hand:  $p = .018$ ), with the mid Tempo being most accurately synchronized (see Figure 7B for a graphical overview).



**Fig. 7** Synchronization ability averages (y-axis; the higher the value, the better the synchronization between movement and bar level of the music) for the significant interaction (A) and main effects (B) with \*\*\*  $p < .001$ , \*\*  $p < .01$ , \*  $p < .05$  (factors: Flux: low frequency spectral flux; Tempo: core tempo; Time-stretch: tempo stretching of stimuli by  $\pm 5\%$ ). The remaining differences are non-significant

## 4 Discussion

We investigated music-induced synchronization ability in relation to the amount of spectral flux contained in the low frequency range of the stimuli, the amount of time-stretch of the stimuli, and the core tempo level of the stimuli. In particular, synchronization was investigated with respect to four body parts (hip, head, feet, and hands) and two metrical levels (beat and bar level). Synchronization ability was defined as the negentropy of the phase difference distribution between the movement (for beat level superior-inferior acceleration and for bar level medio-lateral

acceleration of the respective body parts) and the music, relative to the metrical levels.

Overall, participants synchronized different body parts with respect to the two metrical levels; in particular, the feet and hip were more accurately synchronized to the beat level than the hands and head, whereas synchronization to the bar level was more uniform across body parts. Furthermore, low frequency spectral flux, time-stretching, and core tempo levels influenced synchronization ability in distinct ways. At the beat level, (vertical) synchronization ability tended to be higher for stimuli containing strong low frequency spectral flux than for stimuli containing less low frequency spectral flux, in particular for feet and hip movement and at slower tempi. At the bar level, however, participants synchronized their (sideways) movements more accurately to stimuli containing weak low frequency spectral flux compared to stimuli containing strong low frequency spectral flux, especially in case of hand movements. Moreover, participants tended to be better synchronized to the slower (-5 BPM) rather than the faster (+5 BPM) versions of the same stimuli at both metrical levels, in particular in hip and head movements at the beat level and hip movement at the bar level.

Regarding the core tempo levels, synchronization ability at the beat level to the stimuli containing strong low frequency spectral flux decreased the faster the tempo became, whereas synchronization to the stimuli containing little low frequency spectral flux showed a tendency towards a (non-significant) inverted u-shape relationship with most accurate synchronization at the mid tempo level (foot and hip movement). At the bar level, participants were overall more synchronized to the mid tempo level (115 BPM) than to both slow (105 BPM) and fast (130 BPM) tempo levels, constituting a significant inverted u-shape relationship between the three core

tempo levels. Furthermore, core tempo was interacting with flux for participants' head movements: when moving to stimuli containing strong low frequency spectral flux, participants were more accurately synchronized to the mid tempo level compared to the slow tempo level (as well as to the fast, although non-significantly) They were also more synchronized when moving to weak low frequency spectral flux stimuli compared to strong flux stimuli at the slow tempo.

More broadly, we found that synchronization ability at the beat level resulted in higher average values for hip and foot movement than for head and hand movement, whereas at the bar level it led to similar average values for all four body parts. Regarding the beat level, this result supports Hypothesis 1, suggesting that attunement to the beat level of the music can be specifically related to the core body and the feet, as, for instance, performing footsteps in time with the beat would resonate in the torso movement. These results are in line with findings by Burger et al. (2014) and Toiviainen et al. (2010) that vertical (bouncing) movement is related to the beat level. This might be related to biomechanical properties and constraints, as it requires less effort to bounce to the beat level than, for instance, to sway from side to side. Regarding the bar level, results fail to support Hypothesis 1 (i.e., head and hand movements being more synchronized with the music), since synchronization ability was found to be similar across the four body parts. Thus, our results could suggest that, instead of being restricted to head and hands, the whole body exhibits more complex synchronized movements that unfold over a longer period of time.

Our results regarding beat- versus bar-level synchronization serve as a refinement of previous results in terms of different body parts being differently coupled to the metrical structure of the music. These results are also in line with Leman's (2007) concept of *synchronization*, which suggest that bodily movement, in our case vertical

bouncing motion, is used to follow and embody the basic beat structure as an initial and simple way to engage with music. These results also support the theory of *dynamic attending* (Jones, 1976; Jones & Boltz, 1989; Drake et al., 2000) in that humans are able to perceive and attune to different metrical levels simultaneously or shift in-between. Participants were attuned to the referent level, the beat period, but were also perceiving the metrical hierarchy (other metrical levels, i.e., the bar level) in the music (maybe a more complex rhythmic structure or weaker low frequency spectral flux). This could have either made them shift attention and synchronize to another, more appropriate, metrical level, or even embody different metrical levels simultaneously by using different body parts and movement directions.

Regarding Hypothesis 2 – that larger amounts of low frequency spectral flux increase synchronization ability – results indicate that synchronization might not solely depend on the amount of low frequency spectral flux, but rather on the combination of the amount of flux and the stimulus tempo. For the slower tempo levels, results show that strong low frequency spectral flux increases synchronization ability for hip and foot movement at the beat level, which would support Hypothesis 2. However, synchronization ability decreases at faster tempi and resembles the values found for synchronization to the weak low frequency spectral flux stimuli. This might happen because relatively complex metrical structures in stimuli with large amounts of low frequency spectral flux could be easier to perceive at slower tempi than at faster tempi, so participants could have been able to synchronize more accurately and embody the beat structure in their hip (bouncing) and feet (tapping, stepping) movement at slower tempi.

These results can further be seen in light of Leman's (2007) concept of *inductive resonance*: spontaneous body synchronization as a way to embody fundamental

musical features such as the basic rhythmic structure, with an active imitation and prediction thereof. Low frequency spectral flux, potentially in combination with tempo, could modulate such imitation and attunement, since it provides well-defined temporal anchor points due to the strong low frequency content. This outcome could also support Hove et al.'s (2014) finding that time perception is better for lower musical frequency ranges, as our participants more readily attuned to the stimuli containing a higher amount of low frequency content. Additionally, the results are in line with Stupacher et al. (2016), and suggest that strong low frequency spectral flux stimulates (synchronized) movement, as these stimuli could be perceived as "groovier" than stimuli weak in low frequency spectral flux.

These results also refine previous results on the relationship between low frequency spectral flux and general music-induced movement (Burger, Thompson, et al., 2013) in relation to synchronization behavior. Whereas the previous study found that low frequency spectral flux was associated with head speed (i.e., the head moving faster as the amount of low frequency spectral flux increased), the present study suggests that the head exhibits less synchronized movement to the beat level with increased flux. While head synchronization was more accurate for the stimuli containing strong low frequency spectral flux at slow tempi and more accurate for weak flux stimuli at fast tempi, the overall synchronization ability was lower compared to the feet and hip movements. Therefore, the head might not have been the foremost body part that participants used to synchronize to the beat of the music.

When considering synchronization at bar level in relation to low frequency spectral flux, Hypothesis 2 was not supported. In contrast to the beat level findings, participants were generally more tightly synchronized to stimuli containing little low frequency spectral flux, in particular for hand movement. Thus, in cases where the

surface rhythmic structure is less clear, participants might have used (and needed) larger and more complex movements to entrain to the music. Sideways (swaying) movements are well suited for this purpose, being flexible enough for large movement unfolding over a longer time frame (i.e., a bar or half-bar). In particular, movements of hands would have the required degrees of freedom to embody such larger musical structures. Music with ambiguous rhythmic structures containing small amounts of low frequency spectral flux may require higher-level feature processing that is less related to the basic beat structure, but more to ‘musical’ characteristics, such as timbral evolution or rhythmic complexity (cf. Burger, Thompson, et al., 2013).

With regards to embodied music cognition, *embodied attuning* (Leman, 2007) could provide a theoretical framework for this relation of longer, more complex musical and rhythmical structures yielding longer, more complex movement sequences that led participants to attune to higher metrical levels than the beat level. Participants might have required their whole body, and in particular their hands and arms, to parse musical structure that evolved over a longer time span and contained less low frequency content, but more higher-frequency content (e.g., timbral characteristics of the music).

For head movement at the bar level, flux level interacted with core tempo in the ways that stimuli containing weak low frequency spectral flux were more accurately synchronized at slow tempi than stimuli with strong low frequency spectral flux. Furthermore, synchronization was more accurate at the mid tempo level compared to the slow tempo level in case of strong low frequency spectral flux. However, this result might rather be related to characteristics of the chosen stimuli than a generalizable result.

The relationship between the stimuli's amount of spectral flux and emergent synchronization to the two different metrical levels is noteworthy. It indicates a rather complex and hierarchical relationship between the musical characteristics that provide different synchronization cues and our ability to entrain to such musical stimuli and move with them in synchronized manners. Our results further suggest that our initial Hypothesis 2 was too general and requires more differentiation regarding particular behaviors at different metrical levels: synchronization to the beat level increases when basic rhythmic structures are strong and easily perceivable (strong flux, slower tempi), while synchronization moves to higher metrical levels in case of less pronounced rhythmic structures present in the low frequencies (weak flux). These results imply that synchronization might be modulated by the amount of low frequency spectral flux and therefore by the clarity and salience of the surface (low frequency) rhythmic and metrical structure.

Our results further show that participants were more synchronized to the slowed-down than to the sped-up versions of the stimuli, with significant differences for hip, head, and hand movement at the beat level, and for hip movement at the bar level. This result is counter to Hypothesis 3, as the results suggest that time-stretching does not affect synchronization relative to 110/120 BPM. Therefore, it might rather be a relative, within-stimulus effect rather than a global, across-stimulus effect that could be related to some kind of (implicit) relative tempo memory or anchoring for each song (London et al., 2016). It might have been easier to attune to the slower version of each song, as participants would have had more time to predict and attune to the relative timings.

Regarding the core tempo levels (105, 115, and 130 BPM), the analysis returned interesting results. Hypothesis 4 was not supported for both the beat and the bar level,

since we found significant differences for the different core tempo levels. At the beat level, core tempo interacted with the low frequency spectral flux for foot, hip, and hand movement, and showed a significant main effect for head movement, and thus suggests that synchronization behavior might differ from general spontaneous music-induced movement (Burger, Thompson, et al., 2013) or period-locking at beat level (Burger et al., 2014) that showed no significant effect of tempo. However, these previous analyses have not looked at possible interactions with low frequency spectral flux, which might be a crucial cue for bodily responses to music.

At the bar level, however, we found an inverted u-shape pattern for foot, hip, and hand movement with most accurate synchronization to the mid tempo level. Thus, the synchronization to the downbeat in each bar has been much more accurate in the mid tempo than in the other tempi. The mid (core) tempo and the tempo range after the time-stretch (109-120 BPM) are well located within the preferred tempo range of 110/120 BPM (Fraisse, 1982; Parncutt, 1994), which could suggest that preferred tempo has a larger effect on synchronization ability to longer time spans (such as movement over a musical bar) than to the lower metrical levels (i.e., the beat level). It is likely that such a relationship between preferred tempo and tempo levels did not occur for the beat level, since the tempo differences – between 99.75 and 136.5 BPM – were not large enough for participants to have divergent synchronization responses. However, when relating to the bar level, the temporal differences increased and could therefore have an impact on synchronization behavior. This outcome is ambiguous in relation to results found in Burger et al. (2014), that participants switched from being period-locked to a lower metrical level with slower tempi being period-locked to a higher metrical level with faster tempi. As such, this requires continuous refinement in future studies, incorporating a wider selection of tempi further diverging from the

preferred tempo range, for example as low as 60 BMP and as high as 180 BPM, as the tempo range might have been too small to confirm the previous result.

In general, the results could be solely related to the characteristics of Motown music, since we specifically chose this kind of music to provide our participants with music conducive to dancing. With this restriction, we aimed at selecting stimuli that differed in certain musical characteristics (low frequency spectral flux, tempo), while staying within one musical style. However, a wider range of dance music styles and a larger number of stimuli, systematically varying in low frequency spectral flux and tempo, should be tested in future experiments. This would furthermore offer opportunities to generalize the resulting interaction effects between flux and tempo at the beat level and the other effects. Observations are now only based on one stimulus per factor level, so results might be related to an effect of stimulus or musical genre. By increasing the number of stimuli, also other statistical approaches than ANOVAs could be possible.

While using real music offers most ecological validity for a dance study in a lab environment, it also introduced some lack of control. Despite selecting music from one genre, being recorded and instrumented in similar ways, the stimuli were still different pieces of music and differed both in musical characteristics as well as in subjective characteristics, such as familiarity and likeability. To overcome parts of this challenge, musical stimuli could be custom-made and musical features, such as the low frequency content could be systematically manipulated. However, stimuli might then sound rather similar (and/or fabricated) and the advantage of using music that was originally intended to make people dance is lost. As a compromise, using existing (popular) music in combination with computational extraction and/or

musicological analysis of music characteristics could be suitable way to provide sufficient control of stimuli.

With recording participants individually, we gained insights into their personal characteristic responses to music. However, dancing is an activity often done collectively in groups (Brown, Merker, & Wallin, 2000; Dunbar, 2012). This aspect of sociality is important and should not be neglected. Therefore, future studies will be conducted in group settings to investigate how interpersonal coordination might differ from moving individually, ultimately increasing the ecological validity of our results.

## **Conclusion**

By investigating the effect of low frequency spectral flux and tempo on music-induced synchronization, we revealed complex, hierarchical relationships between characteristics of the rhythmic and temporal structure and our ability to synchronize and entrain to such musical stimuli. Strong low frequency spectral flux was found to result in tighter synchronization at slower tempi at the beat level, whereas it became a less salient cue at faster tempi. At the bar level, weak low frequency spectral flux showed generally tighter synchronization, with a peak at the mid tempo level. In conclusion, real music presents a rich and complex set of affordances for rhythmic synchronization, and careful analysis of how we move to such music can reveal what those affordances are. Besides further attempts to generalize the results, more insights into underlying mechanisms of how humans synchronize to complex musical stimuli could be provided by neuroscience approaches (e.g., combining motion capture with EEG).

## **Ethical standards**

All participants gave their informed consent prior to their inclusion in the study and they were free to discontinue the experiment at any point. Ethical permission for this study was not needed, which was according to the guidelines stated by the university ethical board. Details that might disclose the identity of the participants have been omitted. Furthermore, the manuscript does not contain clinical studies or patient data.

## Literature

- Alluri, V., & Toiviainen, P. (2010). Exploring perceptual and acoustical correlates of polyphonic timbre. *Music Perception*, *27*, 223–242. doi: 10.1525/mp.2010.27.3.223
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)*, *57*, 289–300.
- Brown, S., Merker, B., & Wallin, N. L. (2000). An introduction to evolutionary musicology. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 3–24). Cambridge, MA: MIT Press.
- Burger, B., Ahokas, R., Keipi, A., & Toiviainen, P. (2013). Relationships between spectral flux, perceived rhythmic strength, and the propensity to move. In R. Bresin (Ed.), *Proceedings of the 10th Sound and Music Computing Conference* (pp. 179–184). Stockholm, Sweden: KTH Royal Institute of Technology.
- Burger, B., Thompson, M. R., Saarikallio, S., Luck, G., & Toiviainen, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in Psychology*, *4*:183. doi: 10.3389/fpsyg.2013.00183
- Burger, B., Thompson, M. R., Luck, G., Saarikallio, S., & Toiviainen, P. (2014). Hunting for the beat in the body: On period and phase locking in music-induced movement, *Frontiers in Human Neuroscience*, *8*:903. doi: 10.3389/fnhum.2014.00903
- Burger, B. & Toiviainen, P. (2013). MoCap Toolbox – A Matlab toolbox for computational analysis of movement data. In R. Bresin (Ed.), *Proceedings of the 10th Sound and Music Computing Conference* (pp. 172–178). Stockholm, Sweden: KTH Royal Institute of Technology.
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2009). The role of auditory and pre-motor cortex in sensorimotor transformations. *The neurosciences and music III – Disorders and plasticity: Annals of the New York Academy of Science*, *1169*, 15–34.
- Drake, C., Gros, L., & Penel, A. (1999). How fast is that music? The relation between physical and perceived tempo. In S. W. Yi (Ed.), *Music, mind, and science* (pp. 190–203). Seoul, South Korea: Seoul National University Press.

- Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition*, *77*, 251–288. doi: 10.1016/S0010-0277(00)00106-2
- Dunbar, R. (2012). On the Evolutionary Function of Song and Dance. In N. Bannan (Ed.), *Music, Language, & Human Evolution* (pp. 201–214). Oxford, UK: Oxford University Press.
- Eerola, T., Luck, G., & Toiviainen, P. (2006). An investigation of pre-schoolers' corporeal synchronization with music. In M. Baroni, A. R. Addessi, R. Caterina, & M. Costa (Eds.), *Proceedings of the 9th International Conference on Music Perception and Cognition* (pp. 472–476). Bologna, Italy: University of Bologna.
- Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 149–180). New York, NY: Academic Press. doi: 10.1016/B978-0-12-213562-0.50010-3
- Godøy, R. I., Haga, E., & Jensenius, A. R. (2006). Playing “Air instruments”: mimicry of sound-producing gestures by novices and experts. In S. Gibet, N. Courty, & J.-F. Kamp (Eds.), *Gesture in human-computer interaction and simulation, Lecture Notes in Computer Science, 3881* (pp. 256–267). Berlin; Heidelberg, Germany: Springer.
- Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience*, *19*, 893–906. doi: 10.1162/jocn.2007.19.5.893
- Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience*, *29*, 7540–7548. doi: 10.1523/JNEUROSCI.2018-08.2009
- Hove, M. J., Marie, C., Bruce, I. C., & Trainor, L. J. (2014). Superior time perception for lower musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proceedings of the National Academy of Sciences*, *111*, 10383–10388.
- Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology*, *141*, 54–75. doi: 10.1037/a0024208
- Jones, M. C. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, *83*, 323–355. doi: 10.1037/0033-295X.83.5.323
- Jones, M. R., & Boltz, M. G. (1989). Dynamic attending and responses to time. *Psychological Review*, *96*, 459–491.
- Keller, P. E., & Repp, B. H. (2004). When two limbs are weaker than one: Sensorimotor syncopation with alternating hands. *Journal of Experimental Psychology*, *57*, 1085–1101. doi: 10.1080/02724980343000693
- Keller, P. E., & Rieger, M. (2009). Special Issue – Musical movement and synchronization. *Music Perception*, *26*, 397–400. doi: 10.1525/mp.2009.26.5.397
- Large, E. W., Fink, P., & Kelso, J. A. S. (2002). Tracking simple and complex sequences. *Psychological Research*, *66*, 3–17. doi: 10.1007/s004260100069
- Lartillot, O., & Toiviainen, P. (2007). A Matlab toolbox for musical feature extraction from audio. In *Proceedings of the 10th International Conference on Digital Audio Effects* (pp. 1–8). Bordeaux, France: University of Bordeaux.
- Leman, M. (2007). *Embodied music cognition and mediation technology*. Cambridge, MA: MIT Press.
- Leman, M., & Godøy, R. I. (2010). Why study musical gesture? In R. I. Godøy & M. Leman (Eds.), *Musical gestures. Sound, movement, and meaning* (pp. 3–11). New York, NY: Routledge.

- Lesaffre, M., De Voogdt, L., Leman, M., De Baets, B., De Meyer, H., & Martens, J.-P. (2008). How potential users of music search and retrieval systems describe the semantic quality of music. *Journal of the American Society for Information Science and Technology*, *59*, 695–707. doi: 10.1002/asi.20731.
- Le Van Quyen, M., Foucher, J., Lachaux, J.-P., Rodriguez, E., Lutz, A., Martinerie, J., & Varela, F. J. (2001). Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony. *Journal of Neuroscience Methods*, *111*, 83–98.
- London, J. (2000). Rhythm. *Grove Music Online. Oxford Music Online*. Oxford University Press. <http://www.oxfordmusiconline.com/subscriber/article/grove/music/45963>
- London, J. (2011). Tactus  $\neq$  tempo: some dissociations between attentional focus, motor behavior, and tempo judgment. *Empirical Musicology Review*, *6*, 43–55.
- London, J., Burger, B., Thompson, M. R., & Toiviainen, P. (2016). Speed on the dance floor: auditory and visual cues for musical tempo. *Acta Psychologica*, *164*, 70–80.
- Luck, G., & Sloboda, J. A. (2008). Exploring the spatiotemporal properties of the beat in simple conducting gestures using a synchronization task. *Music Perception*, *25*, 225–239. doi: 10.1525/mp.2008.25.3.225
- Luck, G., & Sloboda, J. A. (2009). Spatio-temporal cues for visually mediated synchronization. *Music Perception*, *26*, 465–473. doi: 10.1525/mp.2009.26. 5.465
- Luck, G., & Toiviainen, P. (2006). Ensemble musicians' synchronization with conductors' gestures: An automated feature-extraction analysis. *Music Perception*, *24*, 189–199. doi: 10.1525/mp.2006.24.2.189
- Madison, G., Gouyon, F., Ullén, F., & Hörnström, K. (2011). Modeling the tendency for music to induce movement in humans: First correlations with low-level audio descriptors across music genres. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 1578–1594.
- Moelants, D. (2002). Preferred tempo reconsidered. In C. Stevens, D. Burnham, G. McPherson, E. Schubert, & J. Renwick (Eds.), *Proceedings of the 7th International Conference on Music Perception and Cognition, Sydney, Australia* (pp. 580–583). Adelaide, Australia: Causal Productions.
- Naveda, L., & Leman, M. (2010). The spatiotemporal representation of dance and music gestures using Topological Gesture Analysis (TGA). *Music Perception*, *28*, 93–112.
- Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, *11*, 409–464.
- Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The action simulation for auditory prediction (ASAP) hypothesis. *Frontiers in Systematic Neuroscience*, *8*:57. doi: 10.3389/fnsys.2014.00057
- Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition*, *67*, 94–102.
- Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, *12*, 969–992.
- Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin & Review*, *20*, 403–452.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x
- Snyder, J., & Krumhansl, C. L. (2001). Tapping to ragtime: Cues to pulse finding. *Music Perception*, *18*, 455–489. Doi: 10.1525/mp.2001.18.4.455

- Stupacher, J., Hove, M. J., & Janata, P. (2016). Audio features underlying perceived groove and sensorimotor synchronization in music. *Music Perception, 33*, 571–589.
- Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., & Keller, P. E. (2013). Musical groove modulates motor cortex excitability: A TMS investigation. *Brain and Cognition, 82*, 127–136.
- Styns, F., van Noorden, L., Moelants, D., & Leman, M. (2007). Walking on music. *Human Movement Science, 26*, 769–785. doi: 10.1016/j.humov.2007.07.007
- Tass, P., Rosenblum, M. G., Weule, J., Kurths, J., Pikovsky, A., Volkmann, J., Schnitzler, A., & Freund, H.-J. (1998). Detection of n:m phase locking from noisy data: Application to magnetoencephalography. *Physical Review Letters, 81*, 3291–3294.
- Todd, N. P. M., O’Boyle, D. J., & Lee, C. S. (1999). A sensory-motor theory of rhythm, time perception and beat induction. *Journal of New Music Research, 28*, 5–28.
- Todd, N. P. M., Rosengren, S. M., & Colebatch, J. G. (2008). Tuning and sensitivity of the human vestibular system to low-frequency vibration. *Neuroscience Letters, 444*, 36–41.
- Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: Hierarchical eigenmodes in music-induced movement. *Music Perception, 28*, 59–70. doi: 10.1525/mp.2010.28.1.59
- Toiviainen, P., & Snyder, J. S. (2003). Tapping to Bach: Resonance-based modeling of pulse. *Music Perception, 21*, 43–80. doi: 10.1525/mp.2003.21.1.43
- Trainor, L. J., Gao, X., Lei, J.-J., Lehtovaara, K., & Harris, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. *Cortex, 45*, 35–43. doi: 10.1016/j.cortex.2007.10.014
- Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., & Leman, M. (2013). The impact of the bass drum on human dance movement. *Music Perception, 30*, 349–359. doi: 10.1525/mp.2013.30.4.349
- Van Noorden, L., & Moelants, D. (1999). Resonance in the perception of musical pulse. *Journal of New Music Research, 28*, 43–66. doi: 10.1076/jnmr.28.1.43.3122
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, MA: MIT Press.
- Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience, 8*, 547–558. doi: 10.1038/nrn2152
- Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Science of the United States of America, 107*, 5768–5773. doi: 10.1073/pnas.1000121107