Feature derivation and classification of auditory-motor coupling dynamics in healthy and neurologically impaired adults

Muhammad Bergas Nur Fayyad[1], **Joeri R. Verbiest**[2,3,4,5], **Anna Ivanova**[6], **Mario Manto**[7,8], **Lousin Moumdjian**[2,5,9] *

1 IQ Health Department, Section Biostatistics, Radboud University Medical Center, Nijmegen, Netherlands, **2** Faculty of Rehabilitation Sciences, REVAL Rehabilitation Research Center, Hasselt University, Hasselt, Belgium, **3** Biomedical Research Institute (BIOMED), Hasselt University, Diepenbeek, Belgium, **4** Data Science Institute (DSI), Hasselt University, Diepenbeek, Belgium, **5** University Multiple Sclerosis Center (UMSC), Hasselt-Pelt, Belgium, **6** Center for Statistics (CenStat), Hasselt University, Diepenbeek, Belgium, **7** Service de Neurologie, CHU-Charleroi, Charleroi, Belgium, **8** Service des Neurosciences, University of Mons, Mons, Belgium, **9** IPEM Institute for Systematic Musicology, Ghent University, Ghent, Belgium

* Lousin.moumdjian@uhasselt.be

Abstract

The use of auditory stimuli in rehabilitation to target walking has been evidenced in persons with neurological conditions. The methodologies focus on the synchronisation of persons' steps to auditory stimuli showing that the type of stimuli and tempi significantly affect the synchronisation. However, the dynamic of the interaction over time between the motor system and the auditory stimuli, i.e., when steps are aligned (termed as locking) and not aligned (termed as unlocking) to the beat of the stimuli, remains unclear. Quantifying these dynamics would assist in the development of personalised rehabilitation. Nevertheless, it is currently challenging given the variability of responses per individual over time. We propose a methodological solution to quantify the dynamics of the step-to-beat coupling over time within an experimental paradigm where healthy (n = 7) and neurological impaired (n = 6) participants walk three minutes to music and metronomes at various tempi. We applied window partitioning within the time series to account for the changing pattern. To classify data into locked and unlocked events, features of fluctuation and trend were derived on which two statistical tests (circular statistical test and slope test) were done, respectively. Based on the ground truth, the performance of our proposed method yielded high accuracy (91%), precision (90%) and recall (97%). The standard deviation of the inter-step intervals was then modelled across the label and experimental factors. The proposed method is suitable for quantifying fine-grained observation of the dynamics of auditory-motor coupling in adult healthy and neurological impaired participants, with the potential of designing personalised rehabilitation.

Introduction

The use of auditory stimuli in rehabilitation to target walking has been evidenced in persons with neurological conditions, for example, in persons with Parkinson's disease [1], stroke [2] and multiple sclerosis [3]. These methodologies are based on sensory-motor synchronisation, where persons are requested to synchronise their steps in time to the beats of music or metronomes. Changes in walking parameters (such as spatiotemporal parameters, e.g., cadence, speed, and stride length) in these studies evaluate the effect of the application of these techniques on the study participants. On the other hand, measuring the synchronisation consistency and accuracy of the step relative to the beat when engaging in such a task also provides abundant information in addition to the changes and dynamics of the walking itself [4].

In a study on neurological patients, our research group previously showed that synchronisation to auditory stimuli significantly differs when walking to music compared to metronomes for different auditory tempi [3,5]. The results obtained per trial of three minutes were summarised into one average value per trial—as commonly done to report on central tendencies in statistical reporting [6]. Albeit these results are comprehensive and provide valuable clinical insights, they still fail to report on the dynamics of the interaction over time. To elaborate, that is the investigation of the dynamics of the interaction occurring between the motor system (intrinsic oscillator) and the auditory stimuli (and extrinsic oscillator) when engaging sensorimotor synchronisation. This study aims to address this gap by quantifying the dynamics of auditory-motor interactions over time, rather than solely focusing on statistical metrics such as averages. Understanding these dynamical processes over time can advance both fundamental and clinical applications. For instance, a clinician, with the help of assistive technologies, can intervene and adjust the components of the auditory-motor paradigm in real time when suboptimal dynamics are detected, enabling in this way a more personalised and effective intervention.

However, quantifying these dynamics over time across participants and trials is challenging given the differences, not only across participants but also within participants [7]. We propose a methodology to tackle this challenge during a task when persons are instructed to synchronise their steps during walking to auditory stimuli (music and metronomes). The relative phase angle (rPA), a measure of an individual's step and the closest beat in the auditory stimuli [4], is an outcome to quantify step-to-beat synchronisation. If one applies a classification logic [8], two possibilities arise: a) participants are locked in phase (i.e., synchronising: their steps are aligned with the beat), and b) participants are not locked in phase (i.e., not synchronising: their steps are not aligned with the beat).

Given the challenges mentioned above, we proposed a methodological framework that takes into account the time domain of the signal. It also avoids specifying arbitrary thresholds for mean and variance in the classification of periods of locking and unlocking. The method includes partitioning the trial into time windows and classifying each window into locked and unlocked states by a systematic statistical evaluation. Afterwards, the time series specific to the walking (gait interval variability) can be merged into the partitioned windows. We hypothesised that this methodological framework will enable the investigation of auditory-motor coupling dynamics over time thereby allowing answering the following research questions. Does locking or unlocking over time within a trial affect the variability of inter-step interval across stimuli and tempi. Answers to such questions can guide the rehabilitation community in developing targeted and individualised rehabilitation plans. As well as contribute to the advancement of assistive technologies, enabling the implementation of personalised rehabilitation strategies that optimise auditory-motor dynamics. We applied this methodology in adult populations with and without neurological impairments.

Methods

Experimental procedure and data

This experiment is part of a larger-scale project, medical ethical approval (B1152021000003) and clinical trial registration (NCT04887753) investigating sensorimotor synchronisation in persons with cerebellar impairments. The recruitment period for the large-scale project started in 29th of April 2021, and ended on the 3rd of May 2024. Recruited participants provided a written consent for study participation by signing the informed consent documentation approved by the medical ethical committee. The following inclusion criteria was used for cerebellar-impaired persons: the presence of a cerebellar impairment diagnosed by a neurologist evidence from MRI imaging (presence of a lesion and/or degeneration), or a minimum score of 1 on the Scale of Assessment and Rating of Ataxia. Participants were excluded if they exhibited cognitive impairment impeding understating the instructions, uncorrected hearing impairment, beat amusia, or pregnancy.

Each participant was equipped with an interactive music player called the D-jogger [9]. The D-jogger consists of headphones (Sennheiser RS 127–8 headphones), inertial measurement unit sensors (NGIMU,x-io technologies limited) strapped to the ankles, and a laptop ((Dell Latitude laptop, Core i5-1145, Windows 10 Pro, ASIO low-latency soundcard).) with a custom-made software. The D-jogger provided auditory stimuli at any given tempo by changing the frequency of the musical beats and metronome ticks and logged all auditory and gait data for calculating outcome measures of gait-music synchronisation discussed in the following paragraphs.

The D-jogger technology has been applied in previous studies to provide and quantify step-beat synchronisation during walking and/or running on adult healthy controls [9–11] and neurological populations [3,5,12–14] as well as in children with developmental coordination disorder [15]. The NGIMU sensors have been validated with gold standard gait analysis methods [16].

Each participant was asked to walk to different trials of 3 minutes each. The trials consisted of two blocks of stimuli (music and metronomes) at seven tempi: -12%, -8%, -4%, 0, +4%, +8%, and +12% of baseline comfortable cadence. Thus, each participant underwent 14 different trials.

Outcome measures

As seen in Fig 1, two time series measurements are processed: the inter-step interval and relative phase angle. The measurements were derived for each participant and each trial.

A. The inter-step interval. The inter-step interval describes the participant's walking pattern and was derived from the inertial measurement unit sensors (NGIMU, UK) recordings of the step-time. Step_j denotes the time at which the j-th step occurs. We define the time variable t_j which denotes the timing of the j-th step ($t_j = \text{Step}_j$). The inter-step interval δ_j at time t_j is calculated by subtracting the $(j+1)$-th step with the j-th step.

$$\delta_j = \text{Step}_{j+1} - \text{Step}_j, \quad j = 1, 2, \ldots, J.$$

Note that the number of observations of the inter-step interval is $J-1$ since there is no step occurring after the last step Step_j.

B. Relative phase angle (rPA). The relative phase angle (rPA), measured in degrees, describes the participant's interaction with the auditory stimuli. It quantifies the difference between the step and the closest auditory beat (i.e., measures the timing of the footfall relative to the closest beat), ranging from -180 to +180. This is either a negative (footfall before the
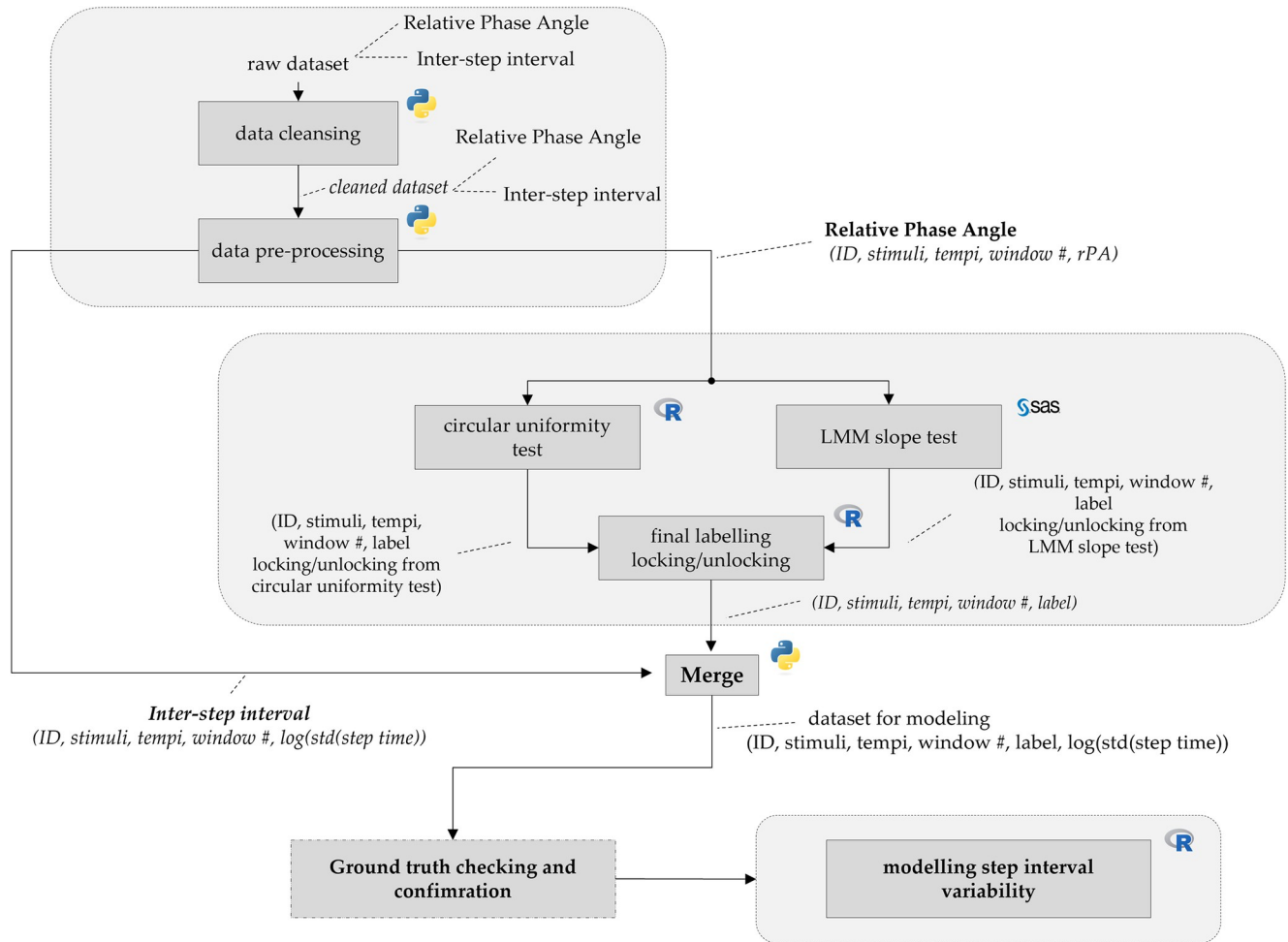
Fig 1. Pipeline of the analysis.

beat) or a positive (footfall after the beat) angle. The rPA is calculated as follows:

$$r_j = 360 * \left(\frac{\text{Step}_j - \text{Beat}_j}{\text{Beat}_{j+1} - \text{Beat}_j} \right), \;\; j = 1, 2, \dots, J,$$

where Beat$_j$ is the time of the beat prior to Step$_j$. One can calculate the average rPA using circular statistics [17].

Data pre-processing

As shown in Fig 1, each outcome measure underwent a separate processing step. We carefully examined the time series of the inter-step intervals to identify any technical sensor glitches or faulty observations. To accomplish this, we analysed data from the inertial sensors by visually assessing the gyroscopic sensor data to detect distinct points corresponding to gait events such as mid-swing and initial contact. A very large and very small inter-step interval—such as three seconds and 0.1 second—are likely to be outliers and invalid. Generally, there is no fixed threshold to determine whether the inter-step interval is invalid. For this reason, observations larger than (median +3sd) or smaller than (median −3sd), where sd denotes the standard deviation, were considered outliers and removed from the analysis [18].

Next, the inter-step interval and the rPA time series were verified and adjusted to contain 180 seconds of recording by removing the observations above the 180-second mark. Next, the time series were divided into nine windows, each spanning 20 seconds. On average, these 20-second intervals included 20–30 steps.

rPA time series. In this section, we describe the methodological solution where we process the quantification of synchronisation over time. The method allows the capture of the dynamic behaviour of synchronisation that varies across and within participants and trials. The proposed solution was to label each partitioned window within the time series with two labels: locked (good synchronisation occurring) or unlocked (i.e., synchronisation not occurring). To perform the labelling, the fluctuation and the trend of rPA are evaluated in two different manners separately with an interchangeable order. The fluctuation and the trend of rPA evaluation give label results after which we combine both with an "if else" logic for the final labels. Below, we provide the details:

a. Evaluation of fluctuation (circular uniformity test)

In evaluating the fluctuation, one may consider calculating the variability (e.g., standard deviation) of the rPA and determine that high variability indicates an unlocking phase and vice versa. However, this approach requires an arbitrary threshold of variability, which cannot be specified by the highly variable individual observations of the participants across trials.

Therefore, circular statistical test [17] is applied since rPA can be seen and analysed as circular data. Circular statistics deals with data that has no true zero where the beginning and the end of the measurement coincide. The distribution of data is different from the linear case. An example of circular data is the von Mises distribution which is comparable to the normal distribution in the linear case. We refer to Jammalamadaka & SenGupta (2001) [19] for further explanation on circular statistics.

The rPA with slight fluctuation is equivalent to circular data concentrated towards a mean direction. On the other hand, high fluctuation implies that the observations are spread over a circle. Thus, the circular uniformity test on the RPA was performed to evaluate the fluctuation. The null hypothesis is that the data is uniformly distributed, which implies high fluctuation, and the alternative is that the data is concentrated without any preferred mean direction, which implies low fluctuation. Therefore, if the null hypothesis is rejected, we label the window as the locking phase.

b. Evaluation of trend (window-specific slope test)

One approach to evaluate the trend is to obtain the slope of the rPA over time in each window. Since no general threshold existed, an average trend (slope) over all windows in each trial had to be calculated. We approached this evaluation using a linear mixed model (LMM). LMM is often used in longitudinal studies involving repeated measurements that feature two sources of variation, i.e., between and within individuals. LMM accounts for both sources of variation by the inclusion of a random effect. We refer to Verbeke & Molenberghs (2009) [20] for a detailed explanation on LMM.

The model was fitted with the rPA against the time relative to the starting point of the window and including window-specific random intercept and slope. Then, the estimate and the standard error for every window-specific deviation from the overall slope was obtained using best linear unbiased predictor (BLUP) [21]. Finally, based on the estimate and the standard error, we tested the significance of each window-specific deviation from the overall slope. We labelled the window as the unlocking phase in each trial if the window-specific slope deviated significantly from the overall slope.

c. Final labelling of each partitioned window

We combined the results of labels from the above evaluations: a. the circular uniformity test and b. the window-specific slope test.

- A final label of 'locked' was given to the window where both evaluations indicated 'locked.'. This indicated a stable rPA derived from low fluctuation and flat trend features.

- Otherwise, the window was given a final label of 'unlocked'. This indicated unstable rPA derived from high fluctuation and non-flat (slope) trend features.

Ground truth verification of the final attributed labels was conducted by computing confusion matrix [22]. The matrix was scored into its four components as follows:

- A true positive (TP) was defined when the label indicated 'locked' and was coupled with the outcome "correct".

- A true negative (TN) was defined when the label indicated 'unlocked' and was coupled with the outcome "correct".

- A false positive (FP) was defined when the label indicated 'locked' and was coupled with the outcome "incorrect".

- A false negative (FN) was defined when the label indicated 'unlocked' and was coupled with the outcome "incorrect".

After that, the values were counted per outcomes (TP, TN, FP, FN) per partitioned window, and a sum of all was computed; this was used to calculate the recall (TP/TP+FN), precision (TP/TP+FP) and accuracy (TP+TN/Total) of the labels. Two independent reviewers constructed the confusion matrix, and a third reviewer was consulted in case discrepancies arose. The reliability of the ratings was assessed using Cohen's kappa, which measures the agreement between two initial reviewers. Kappa value of 0.60–0.79 indicates a moderate agreement and kappa value of 0.80–0.90 indicates a strong agreement [23].

Merging datasets of inter-step interval and RPA

As seen in Fig 1, at this stage, both time series feature nine labelled partitioned windows merged to create a dataset for modelling. Since window partitioning was done in the same manner for both datasets, merging the datasets was essentially transferring the information of the labels from rPA dataset to the inter-step interval dataset. This results in a dataset containing the outcome measure of inter-step interval together with the label 'locked' or 'unlocked', as well as the subject identifier (ID) and experimental factors (i.e., stimuli and tempi).

Modelling inter-step-interval variability

Modelling inter-step interval variability is the final step of the proposed methodology, which leads to applying a statistical model to answer the research question: *"Does locking or unlocking within a trial affect the variability of inter-step interval across stimuli and tempi."* The variability of inter-step interval over time is modelled as a function of time (i.e., window), factors (i.e., stimuli and tempi), and accounting for the different labels (locked/unlocked). To model the variability in steps, we calculated the standard deviation of the inter-step interval in each window. The standard deviation was then used as the response variable for the modelling. Thus, we have a repeated outcome measure of standard deviation for each trial.

Let δ_{ijklm} denote the inter-step interval for subject i, measured at time t_{ijk} in window ω_k with Tempi$_l$ and Stimuli$_m$. For each window, the standard deviation was calculated, given by

$$S_{iklm} = \sqrt{\frac{\sum_{j=1}^{J_k} \left(\delta_{ijklm} - \mu_{iklm}\right)^2}{J_k}}$$

where μ_{iklm} is the mean of the inter-step interval and J_k is the number of steps in k-th window. From the equation, it can be seen that the standard deviation does not include the index j and the information of time is no longer carried by t$_{ijk}$. Instead, window ω_k is the only variable that carries the time information. Now, the standard deviation S_{iklm} can be seen as a repeated measure over the window ω_k, where the variable window ω_k is now taken as an ordinal variable. In modelling, the ordinal variable of the window is denoted by Window$_k$ which takes integer values from 1 to 9.

The standard deviation can be modelled as Gaussian data and non-Gaussian data with gamma distribution since it can take any non-negative value. Therefore, the modelling was done using LMM with possibly transformed data. The variance-covariance matrix is set to be unstructured [21].

Software

The software used in this study were R version 4.3.2, SAS 9.4, and Python 3. As illustrated in Fig 1, data processing and merging of datasets were conducted in Python. The estimation and the test of the components of random effects using BLUP in LMM was performed in SAS software using PROC MIXED [24] by adding 'solution' in the 'random' statement. The circular uniformity test was done in R using the function circ.range(., test = TRUE) in CircStats package [25]. Fitting LMM was done in R using lmer() function in lme4 package [26].

Results

1.1 Participants

In this methodological study, we present data collected on seven healthy controls (age: mean 61 ± 9 years, sex: 5 female, 2 male) and six patients with cerebellar impairments (age: mean 58 ± 17 years, sex: 3 female, 3 male). Table 1 presents the clinical information of each patient, including their diagnosis, year of diagnosis, MRI results, and scores from the Scale for the Assessment and Rating of Ataxia [27]. Assistive devices were not used by the participants.

1.2 Step-interval time series

Processing and removal of invalid observations from the raw data recordings.

As illustrated in Fig 2, invalid observations shown on the left were processed to obtain a clean inter-step-interval time series as shown in the right.

1.3 Relative phase angle time series

Labelling of the partitioned windows

a. Evaluation of fluctuation (circular uniformity test)

This evaluation was able to catch the phase with high and low fluctuation, although not with full accuracy at all time points. Fig 3, illustrates the time series of rPA during one experimental trial of 180 seconds partitioned into nine windows. The trial has been colour-coded (green and red indicate locked and unlocked window respectively) after undergoing the

Table 1. Descriptive clinical information of each patient, including diagnosis, year of diagnosis, MRI findings, and scores from the Scale for the Assessment and Rating of Ataxia.

Diagnosis	Year of diagnosis	MRI findings	Gait (0–8)	Stance (0–6)	Sitting (0–4)	Speech disturb-ance (0–6)	Finger chase AVG_L,R (0–4)	Nose-finger test AVG_L, R (0–4)	Fast alternating hand movement AVG_L,R (0–4)	Heel-shin slide AVG_L,R (0–4)
Cerebellar Stroke	2021	Right posterior stroke (posterior inferior cerebellar artery)	1	0	0	0	0	0	0	0
Cerebellar stroke	2021	Multiple lesions of bilateral cerebellar hemisphere	1	0	0	0	0	0.5	0	0.5
Arterior-venous stenosis malformulation	2009	Bilateral posterior cerebellar hemispheric lesions and mild cerebellar atrophy	1	0	0	0	0	0	0	0
Cerebellar stroke	2019	Left hemisphertic cerebellar stroke	2	1	0	1	0.5	0.5	0.5	0.5
Cerebellar stroke	2019	Bilateral posterior inferior cerebellar artery stroke	1	2	0	0	0	0	0	0
Cerebellar stroke	2019	Ischemia in Left posterior cerebellar hemisphere (ASCA)	1	0	0	0	1	0	0	0

https://doi.org/10.1371/journal.pone.0315607.t001

circular uniformity test. Note that high fluctuation of rPA indicates unlocking phase and, conversely, low fluctuation indicates locking phase. In Fig 3, the rPA in window 1, 2, and 8 fluctuate highly between the values around -150 and 150. It suggests that the steps in these windows were unsynchronised, and thus classified as unlocked. In contrast, the rPA in window 4, 5, and 6 fluctuates between the values around -100 and 0 which suggests that the steps were synchronised and thus classified as locked.

The example illustrates that the unlocked phase was correctly identified in window 1 and 2 and correctly identified 'locked' label in windows 3 to 6. However, it failed to identify windows 7 and 8 as 'unlocked'.
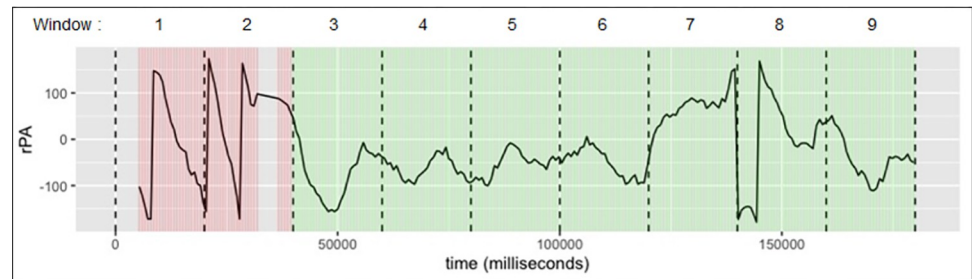
b. Evaluation of trend (window-specific slope test)
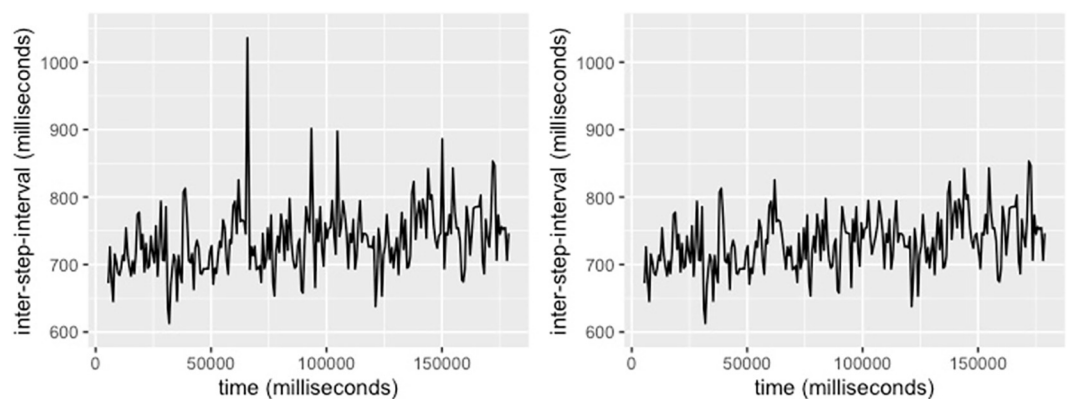
Fig 2. Example of inter-step-interval of a trial, stimuli: Music, tempi: -8% before outlier removal (left) and after outlier removal (right).

https://doi.org/10.1371/journal.pone.0315607.g002

Fig 3. Relative phase angle with colour label for each window based on circular uniformity test: Unlocked (red) and locked (green).

Fig 4, illustrates the time series of rPA during one experimental trial of 180 seconds partitioned into nine windows. The top panel of the figure shows the rPA from window 1 through window 9. The middle panel shows the rPA of nine 20-second windows stacked on each other. The bottom panel shows the label based on the result of LMM slope test. A LMM was fitted with window-specific random intercept and slope on the stacked 20-second windows. Each window-specific slope was tested to determine whether it deviates significantly from the average slope, which is the fixed effect, of the overall 20-second windows. When a significant deviation was absent, the window was labelled 'locked' (as seen in the first eight windows on the bottom panel of Fig 4). When a deviation was present, the window was labelled 'unlocked' (as seen by the last window on the bottom panel of Fig 4).

A combination of both evaluations

Each test captures the locking/unlocking pattern, but not always. With a simple decision rule, the two evaluations were combined. The decision rule was that the window was labelled as unlocked if either of the two evaluations resulted in an unlocked label. Combining both tests improved the performance in capturing the correct labelling. Fig 5 shows a trial where the
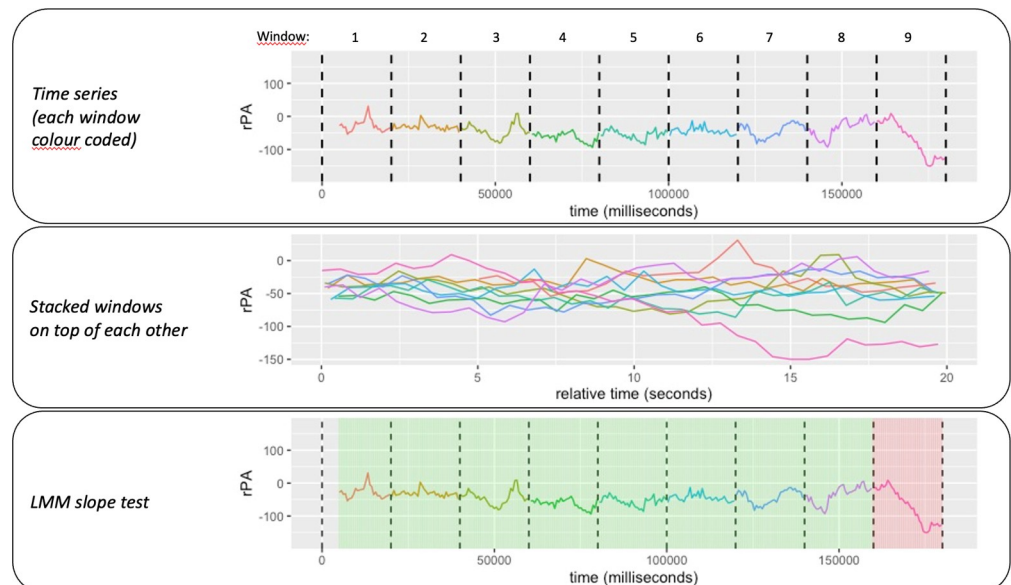
Fig 4. Relative phase angle (rPA) from window 1 to 9 (top panel), rPA in all windows stacked on top of each other (middle panel), and rPA with the locking/unlocking label with green indicating a locking window and red otherwise (bottom panel).
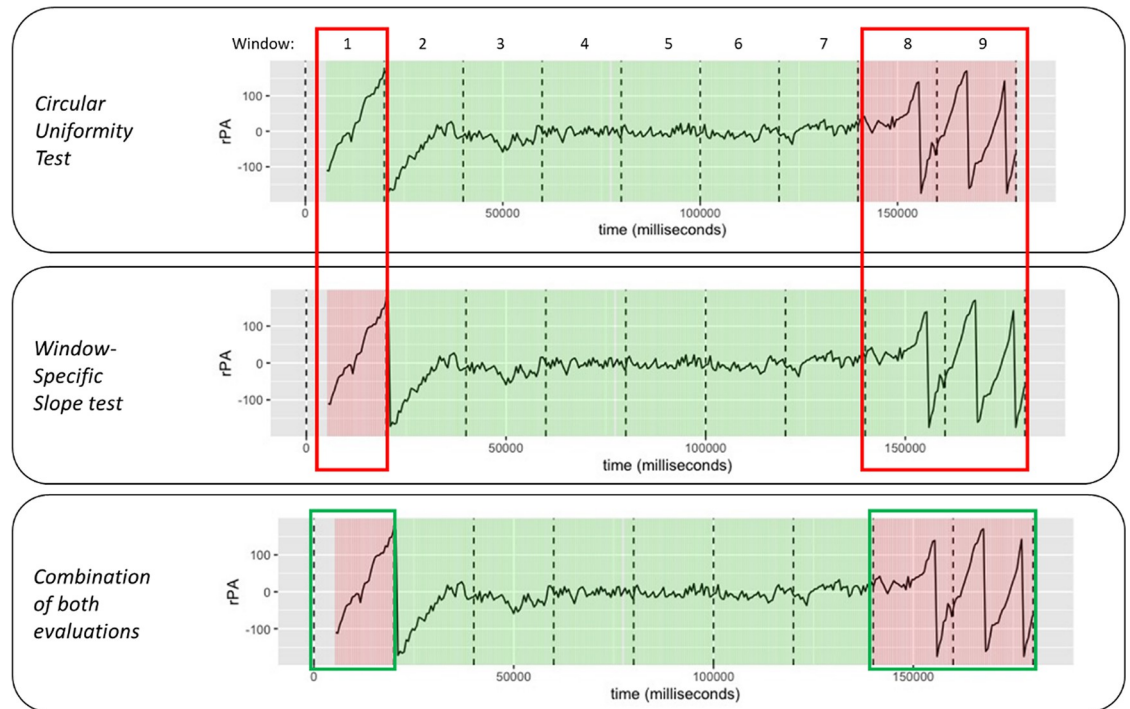
Fig 5. Labelling based on the circular uniformity test (top), slope test (middle), and final label combining both tests (bottom).

https://doi.org/10.1371/journal.pone.0315607.g005

first window was inaccurately labelled as unlocked with the circular uniformity test, but the slope test was able to provide it with the correct 'unlocked' label. The opposite case occurred in the last two windows, where the label 'unlocked' was correctly attributed based on the circular uniformity test but not with the slope test. Nevertheless, the window was attributed to the correct label by combining both pieces of information.

Ground truth verification

The reliability of ratings between the two assessors. The Cohen's Kappa statistics indicated that the initial two raters demonstrated a moderate agreement (Kappa = 0.744). Out of 1,571 total ratings, the raters disagreed on 82 instances, which were subsequently resolved by the third rater, ensuring consensus.

The results of the confusion matrix. The confusion matrix calculations of the combined methodology, accounted for an accuracy of 91%, a precision of 90% and a recall of 97%. Thus, the proposed methodology was deemed valid for statistical modelling in the next steps.

Modelling inter-step interval variability

The inter-step-interval variability was computed in terms of the standard deviation of the inter-step-interval per window for participant and per trial before moving to model fitting. Fig 6 shows that the data was right-skewed. Therefore, the data was log-transformed (bottom panel) and used in the modelling steps described below.

Linear mixed model (LMM). The log standard deviation changes over the window linearly. For this reason, the log standard deviation was modelled as a linear function of the ordinal variable $Window_k$. Additionally, for software convenience (related to the convergence of the model), $Window_k$ was rescaled by dividing it by 10. Since it was of interest to investigate the effect of the locked/unlocked windows and the effect of tempi, the model includes the
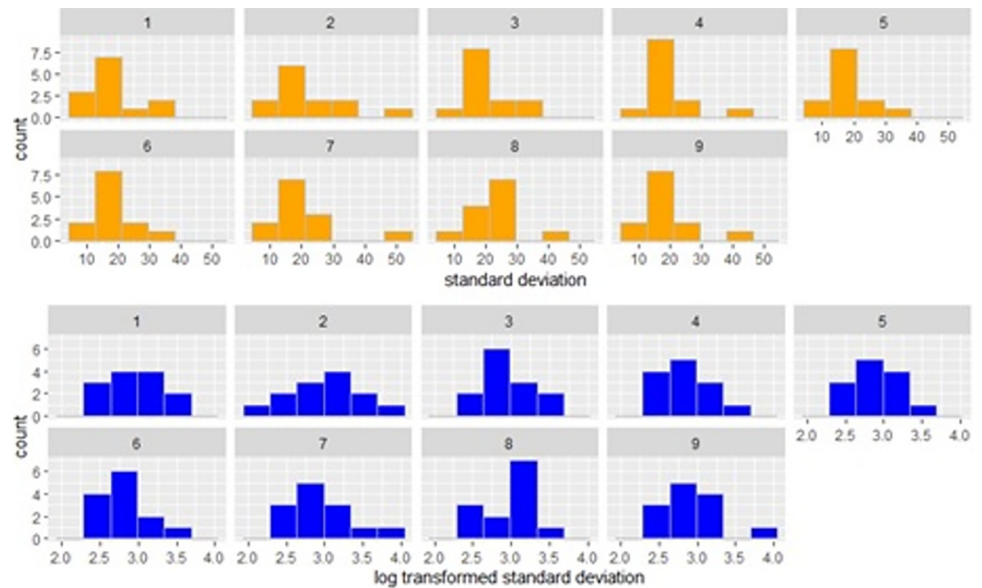
Fig 6. Standard deviation of inter-step-interval from all participants for a specific stimuli and tempi in each window, in original scale (top) and log-transformed (bottom).

interaction of these two variables and the two-way interactions of each of them with the other covariates. All main effect of each covariate was also included.

In modelling, model reduction was done on the covariance structure (i.e., random effects) and the mean structure (i.e., covariates). Note that only the interaction effect was reduced while the main effect of each covariate remained in the model regardless of the insignificance. Eventually, the final model only included random intercept and covariates shown in Table 2.

Based on Table 2, the clinical research questions could be answered, a short summary of such answers is listed below. Note that the answer to specific clinical research questions is not in the scope of the current work.

- There was a significant difference in inter-step interval standard deviation between locked and unlocked windows within Tempi: 0.

- There was a significant effect of Tempi on inter-step interval standard deviation, and the effect was different between the locked and unlocked labels for the two stimuli.

Discussion

We proposed a methodology to quantify steps-to-beat coupling dynamics of participants over time. This is relevant given the individual differences present during such tasks over time, and more so in neurological populations.

Methodologically, we have quantified the phase in which the subjects were locking or unlocking their step to the beat of stimuli with the proposed method. This approach is different than other studies applying methodologies deriving features of the complete time series [28,29], or specific to gait datasets, features extracted by the means segmentation and feature extraction derived from video recordings of whole body silhouettes [30,31]. Here, the approach we selected was a window partitioning approach -similar to the window segmentation [32]—allowing us to quantify the phase over time for each trial by deriving features in

Table 2. Parameter estimate of Linear mixed model.

Parameter	Estimate	Std. error	t	p-value
(Intercept)	2.9725	0.0726	40.964	0.000
Tempi: n12	0.1601	0.0378	4.241	0.000
Tempi: n4	0.0125	0.0377	0.333	0.739
Tempi: n8	0.1276	0.0386	3.305	0.001
Tempi: p12	-0.1970	0.0425	-4.637	0.000
Tempi: p4	-0.0562	0.0399	-1.406	0.160
Tempi: p8	-0.1241	0.0431	-2.876	0.004
Stimuli: music	0.2254	0.0383	5.888	0.000
Window_lab_resc	0.0000	0.0283	-0.001	0.999
Lock_lab: unlocked	0.0594	0.0479	1.240	0.215
Tempi: n12 * Stimuli: music	-0.1945	0.0551	-3.533	0.000
Tempi: n4 * Stimuli: music	-0.0200	0.0526	-0.380	0.704
Tempi: n8 * Stimuli: music	-0.1196	0.0562	-2.128	0.034
Tempi: p12 * Stimuli: music	0.0016	0.0542	0.030	0.976
Tempi: p4 * Stimulimusic	-0.1282	0.0558	-2.296	0.022
Tempi: p8 * Stimulimusic	-0.0432	0.0562	-0.769	0.442
Tempi: n12 * Lock_lab: unlocked	0.0518	0.0662	0.783	0.434
Tempi: n4 * Lock_lab: unlocked	0.0162	0.0647	0.250	0.803
Tempi: n8 * Lock_lab: unlocked	-0.0304	0.0699	-0.436	0.663
Tempi: p12 * Lock_lab: unlocked	0.2790	0.0619	4.506	0.000
Tempi: p4 * Lock_lab: unlocked	0.0703	0.0658	1.069	0.285
Tempi: p8 * Lock_lab: unlocked	0.2064	0.0656	3.147	0.002

each window. The features could be any measurements that tell us the characteristics of the window. Here, we have derived the slope and the circular statistic in each window to represent the trend and the fluctuation, respectively. These features have been evaluated separately and then combined with a simple decision rule. The features were able to distinguish between the locked and the unlocked window.

The window length of 20 seconds was opted to partition the time series for the method to be more flexible in capturing the dynamic of the synchronisation over time. We do however acknowledge the limitation of this choice, as this may compromise the number of observations of step per window. In this study, given that the observations where of gait data, sufficient number of steps were present within each time window for the application of the circular uniformity and slope tests.

After applying our proposed method and having the label of locking/unlocking in each window (by the combined methodology), the inter-step interval variability has been modelled by an LMM as a function of the locking and unlocking phase, tempi, stimuli, and time represented by the window. This methodology allows us to answer the clinical research question—although out of scope of this work—about the differences in inter-step interval variability between the locking and unlocking phase across groups, tempi and stimuli. It thus provides a more fine-grained understanding on the dynamics of the coupling, which could lead to personalised use of rhythm-based interventions in neurological populations [33].

Regarding the final choice to model the inter-step-interval variability, besides LMM, one can also consider using Generalized Linear Mixed Model (GLMM) or Generalized Estimating Equation (GEE) [34]. These models take into consideration that the observation follows a non-Gaussian distribution. In this study, the response variable was the standard deviation of

the inter-step interval. Since it can take value from zero to infinity and has a right-skewed distribution, gamma distribution can be used in GLMM and GEE with log link function. The choice of the model for future analysis depends on the interpretation of interest, as well as the software availabilities and capabilities. In general, for population average interpretation, GEE is practically less software-intensive than GLMM. GLMM requires additional computation to obtain the marginal estimates. On the other hand, GLMM also provides subject-specific interpretation, which can be used to investigate each individual further, e.g., by detecting a subject with an extreme pattern. While the LMM (as was applied here) can also be used for marginal and subject-specific interpretation.

The methodology proposed in this work complements existing metrics of auditory-motor synchronisation which primarily measures synchronisation precision and accuracy [4,35]. While the latter metrics remain crucial for capturing important changes across trials (especially when complemented with the gait spatiotemporal parameters) as shown by studies on Parkinson's [1,36], multiple sclerosis [3,12,14,37,38] and stroke, our approach offers a more fine-grained analysis. It allows us to quantify changes in dynamics, distinguishing between periods of locked (well-synchronised), and unlocked (poorly synchronised) states. This approach has the potential to enable continuous monitoring and classification. A future step could involve the modification of these dynamics through biofeedback, using auditory signals to adjust behaviour according to clinical needs. These auditory-biofeedback strategies have already been applied in the field of sport science to reduce injuries or enhance performance [39,40] and can now complement personalised music therapy or personalised auditory feedback in rehabilitation [33,41].

The method holds broad potential and warrants future investigations, as it could be applied to quantify auditory-motor dynamics not only in neurological populations with motor impairments, but also in those with cognitive or attentional deficits. For example, these investigations can be extended into the field of paediatric rehabilitation for children with attentional deficits, such as ADHD, autism and developmental coordination disorder and other related pathologies, as auditory-motor coupling strategies have demonstrated effects in these populations [15,42–44]. Given the positive attitude towards technology in healthcare across continents [45–48], this approach has the potential to be effectively integrated into clinical practice.

The sample size justification in this study is inherently tied to the proposed methodology. The proposed methodology can be impacted by the presence of different gait patterns or responses on the rPA, rather than sample size. This is because the classification takes as input the number of steps per defined window and accordingly classifies the window to 'locked' or 'unlocked' based on the set of pre-defined features. In this context, the important factor of the methodology is the number of observations available per time-window rather than number of participants included in the sample. With our sample, we can confirm that first, the majority of the included patients were diagnosed with cerebellar stroke, and had similar gait characteristics (SARA gait sub-score, median 1) as seen in the Table 1. Second, in the included sample, each time-window contained a median of 33 steps, and was sufficient for the classification to be accurate and precise, as indicated by the results of the ground truth verification. Thus our proposed classification method performs very well and is feasible to apply in both populations.

However, when testing other neurological populations with varied gait patterns, the window length may have to be adapted in order to ensure appropriate number of observations as well as considering the adaptation features for the evaluation of rPA in each window. Classification approaches in other fields such as neuroimaging [49] and biomechanics [50] rely on appropriate multi-feature selection to enhance the accuracy and overall success of the classification algorithms. Future work is thus warranted to identify additional relevant features that

could increase the accuracy of time-window labelling within our proposed classification algorithm.

Noteworthy, is that the scope of this study was the proposed methodological framework. Given its successful implementation, future work can apply this to investigate relevant clinical research questions, taking into account clinical factors (such as medication use, motor impairment, etc.). Moving forward, sample size calculations would be necessary to determine statistical power when modelling the data after its classification. This will ensure sensitivity to infer the effects of interest, providing robust conclusions to be drawn in response to the clinical questions.

Conclusion

Our proposed classification methodology demonstrated high accuracy, precision and recall in classifying and quantifying the dynamics of step-to-beat coupling over time. The method has the potential to contribute to the monitoring and assessment of individual differences step-to-beat dynamics over time in adult populations, both with and without neurological impairments. Future studies are warranted to validate this approach in larger cohorts across a range of neurological conditions. Ultimately, this methodology can facilitate the development of assistive technologies for clinicians, enabling real-time interventions to adjust the auditory-motor paradigm when suboptimal dynamics are detected, thereby fostering more personalised and effective approach to rehabilitation.

Author Contributions

Conceptualization: Muhammad Bergas Nur Fayyad, Lousin Moumdjian.

Data curation: Muhammad Bergas Nur Fayyad, Joeri R. Verbiest, Lousin Moumdjian.

Formal analysis: Muhammad Bergas Nur Fayyad.

Funding acquisition: Mario Manto, Lousin Moumdjian.

Investigation: Lousin Moumdjian.

Methodology: Muhammad Bergas Nur Fayyad, Joeri R. Verbiest, Anna Ivanova, Lousin Moumdjian.

Project administration: Joeri R. Verbiest, Anna Ivanova, Mario Manto, Lousin Moumdjian.

Resources: Mario Manto, Lousin Moumdjian.

Software: Muhammad Bergas Nur Fayyad, Anna Ivanova.

Supervision: Joeri R. Verbiest, Anna Ivanova, Mario Manto, Lousin Moumdjian.

Validation: Joeri R. Verbiest, Anna Ivanova, Lousin Moumdjian.

Writing – original draft: Muhammad Bergas Nur Fayyad, Lousin Moumdjian.

Writing – review & editing: Muhammad Bergas Nur Fayyad, Joeri R. Verbiest, Anna Ivanova, Mario Manto, Lousin Moumdjian.

References

1. Ghai S, Ghai I, Schmitz G, Effenberg AO. Effect of rhythmic auditory cueing on parkinsonian gait: A systematic review and meta-analysis. Sci Rep. 2018; 8(1):506. https://doi.org/10.1038/s41598-017-16232-5 PMID: 29323122

2. Hutchinson K, Sloutsky R, Collimore A, Adams B, Harris B, Ellis TD, et al. A Music-Based Digital Therapeutic: Proof-of-Concept Automation of a Progressive and Individualized Rhythm-Based Walking

Training Program After Stroke. Neurorehabilitation and neural repair. 2020; 34(11):986–96. https://doi.org/10.1177/1545968320961114 PMID: 33040685

3. Moumdjian L, Moens B, Maes PJ, Van Nieuwenhoven J, Van Wijmeersch B, Leman M, et al. Walking to Music and Metronome at Various Tempi in Persons With Multiple Sclerosis: A Basis for Rehabilitation. Neurorehabilitation and neural repair. 2019; 33(6):464–75. https://doi.org/10.1177/1545968319847962 PMID: 31079541

4. Moumdjian L, Buhmann J, Willems I, Feys P, Leman M. Entrainment and Synchronization to Auditory Stimuli During Walking in Healthy and Neurological Populations: A Methodological Systematic Review. Frontiers in human neuroscience. 2018; 12:263. https://doi.org/10.3389/fnhum.2018.00263 PMID: 29997491

5. Moumdjian L, Maes PJ, Dalla Bella S, Decker LM, Moens B, Feys P, et al. Detrended fluctuation analysis of gait dynamics when entraining to music and metronomes at different tempi in persons with multiple sclerosis. Sci Rep. 2020; 10(1):12934. https://doi.org/10.1038/s41598-020-69667-8 PMID: 32737347

6. Cooksey RW. Descriptive Statistics for Summarising Data. Illustrating Statistical Procedures: Finding Meaning in Quantitative Data Singapore: Springer; 2020. p. 61–139.

7. Miyata K, Yamamoto T, Fukunaga M, Sugawara S, Sadato N. Neural correlates with individual differences in temporal prediction during auditory-motor synchronization. Cereb Cortex Commun. 2022; 3 (2):tgac014. https://doi.org/10.1093/texcom/tgac014 PMID: 35529518

8. Scitovski R, Sabo K., Martínez-Álvarez F., Ungar S. Cluster Analysis and Applications: Springer Nature; 2021.

9. Moens B, Muller C, van Noorden L, Franek M, Celie B, Boone J, et al. Encouraging spontaneous synchronisation with D-Jogger, an adaptive music player that aligns movement and music. PloS one. 2014; 9(12):e114234. https://doi.org/10.1371/journal.pone.0114234 PMID: 25489742

10. Moens B. D-Jogger: An interactive music system for gait synchronisation. With applications for sports and rehabilitation [PhD thesis]. Gent Gent University; 2018.

11. Moens B, Leman M. Alignment strategies for the entrainment of music and movement rhythms. Ann N Y Acad Sci. 2015; 1337:86–93. https://doi.org/10.1111/nyas.12647 PMID: 25773621

12. Moumdjian L, Moens B, Maes PJ, Van Geel F, Ilsbroukx S, Borgers S, et al. Continuous 12 min walking to music, metronomes and in silence: Auditory-motor coupling and its effects on perceived fatigue, motivation and gait in persons with multiple sclerosis. Mult Scler Relat Disord. 2019; 35:92–9.

13. Moumdjian L, Moens B, Van Wijmeersh B, Leman M, Feys P. Application of step and beat alignment approaches and its effect on gait in progressive multiple sclerosis with severe cerebellar ataxia: A proof of concept case study. Multiple sclerosis. 2022; 28(3):492–5. https://doi.org/10.1177/13524585211054000 PMID: 34726562

14. Moumdjian L, Moens B, Vanzeir E, De Klerck B, Feys P, Leman M. A model of different cognitive processes during spontaneous and intentional coupling to music in multiple sclerosis. Ann N Y Acad Sci. 2019; 1445(1):27–38. https://doi.org/10.1111/nyas.14023 PMID: 30865313

15. Goetschalckx M, Feys P, Rameckers E, Moens B, Leman M, Moumdjian L. Auditory-motor synchronization in developmental coordination disorder: Effects on interlimb coordination during walking and running. Ann N Y Acad Sci. 2024. https://doi.org/10.1111/nyas.15228 PMID: 39320072

16. Pierleoni P, Pinti F, Belli A, Palma L. A dataset for wearable sensors validation in gait analysis. Data Brief. 2020; 31:105918. https://doi.org/10.1016/j.dib.2020.105918 PMID: 32637511

17. Berens P. CircStat: A MATLAB Toolbox for Circular Statistics. Journal of Statistical Software. 2009; 31 (10).

18. Phinyomark A, Larracy R, Scheme E. Fractal Analysis of Human Gait Variability via Stride Interval Time Series. Front Physiol. 2020; 11:333. https://doi.org/10.3389/fphys.2020.00333 PMID: 32351405

19. Jammalamadaka SR, SenGupta A. Topics in Circular Statistics: World Scientific, River Edge, N.J.; 2001.

20. Molenberghs G, Verbeke G. Linear Mixed Models for Longitudinal Data: Springer New York, NY; 2000.

21. Verbeke G. M, G. Linear Mixed Models in Practice: a SAS-Oriented Approach. New York Springer 1997.

22. Ting KM. Confusion Matrix. Encyclopedia of Machine Learning. Boston: Springer; 2011.

23. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb). 2012; 22(3):276–82. https://doi.org/10.1016/j.jocd.2012.03.005 PMID: 23092060

24. SAS (SAS Institute I. SAS/STAT® 9.4: User's Guide. 2004.

25. S-plus original by Ulric Lund RpbCA. Circular Statistics, from "Topics in Circular Statistics" (2001): GPL-2; 2018-07-01 [Version 0.2–6:[Available from: https://cran.r-project.org/web/packages/CircStats/index.html.

26. Bates D, Mächler M., Bolker B., Walker S. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software. 2015; 67(1):1–48.

27. Schmitz-Hubsch T, du Montcel ST, Baliko L, Berciano J, Boesch S, Depondt C, et al. Scale for the assessment and rating of ataxia: development of a new clinical scale. Neurology. 2006; 66(11):1717–20. https://doi.org/10.1212/01.wnl.0000219042.60538.92 PMID: 16769946

28. Lycke H, Decoster W, Ivanova A, Van Hulle MM, de Jong FI. Discrimination of three basic female voice types in female singing students by voice range profile-derived parameters. Folia Phoniatr Logop. 2012; 64(2):80–6. https://doi.org/10.1159/000337042 PMID: 22507900

29. Lycke H, Ivanova A, Van Hulle MM, Decoster W, de Jong FI. Discrimination of three basic male voice types by voice range profile-derived parameters. Folia Phoniatr Logop. 2013; 65(1):20–4. https://doi.org/10.1159/000350492 PMID: 23711640

30. Sivolobov S. Human gait feature extraction method. Procedia Computer Science. 193(2021):220–7.

31. De Boer N, Abdi, H., Fielding, M., Nahavandi, S., editor Segmentation and Feature Extraction of Human Gait Motion. in DesTech Conference Proceedings; 2016: Creative Commons Attribution License.

32. Guijo-Rubio D, Duran-Rosal AM, Gutierrez PA, Troncoso A, Hervas-Martinez C. Time-Series Clustering Based on the Characterization of Segment Typologies. IEEE Trans Cybern. 2021; 51(11):5409–22. https://doi.org/10.1109/TCYB.2019.2962584 PMID: 31945011

33. Lai-Tan NP M.G., Kawsar F., Deligianni F. Toward Personalized Music-Therapy: A Neurocomputational Modeling Perspective. IEEE Pervasive Computing. 2023; 22(3):27–37.

34. Molenberghs G, Verbeke G. Models for Discrete Longitudinal Data: Springer New York, NY; 2005. XXII, 687 p.

35. Repp BH, Su YH. Sensorimotor synchronization: a review of recent research (2006–2012). Psychon Bull Rev. 2013; 20(3):403–52. https://doi.org/10.3758/s13423-012-0371-2 PMID: 23397235

36. Bella SD, Benoit CE, Farrugia N, Keller PE, Obrig H, Mainka S, et al. Gait improvement via rhythmic stimulation in Parkinson's disease is linked to rhythmic skills. Sci Rep. 2017; 7:42005. https://doi.org/10.1038/srep42005 PMID: 28233776

37. Vanbilsen N, Moumdjian L., Kinnaert F., Degens B., Moens B., Goetschalckx M., et al. Step to the Beat: Auditory-motor coupling during walking to Higher and Lower Tempi with Music and Metronomes in Progressive Multiple Sclerosis: an observational study. Under revision.

38. Kong L, Zhang X, Meng L, Xue H, Zhou W, Meng X, et al. Effects of music therapy intervention on gait disorders in persons with multiple sclerosis: A systematic review of clinical trials. Mult Scler Relat Disord. 2023; 73:104629. https://doi.org/10.1016/j.msard.2023.104629 PMID: 36963169

39. Van den Berghe P, Lorenzoni V, Derie R, Six J, Gerlo J, Leman M, et al. Music-based biofeedback to reduce tibial shock in over-ground running: a proof-of-concept study. Sci Rep. 2021; 11(1):4091. https://doi.org/10.1038/s41598-021-83538-w PMID: 33603028

40. Lorenzoni V, Staley J, Marchant T, Onderdijk KE, Maes PJ, Leman M. The sonic instructor: A music-based biofeedback system for improving weightlifting technique. PloS one. 2019; 14(8):e0220915. https://doi.org/10.1371/journal.pone.0220915 PMID: 31461448

41. Dotov DG, Cochen de Cock V, Geny C, Ihalainen P, Moens B, Leman M, et al. The role of interaction and predictability in the spontaneous entrainment of movement. J Exp Psychol Gen. 2019; 148 (6):1041–57. https://doi.org/10.1037/xge0000609 PMID: 31180717

42. Carnevali L, Valori I, Mason G, Altoe G, Farroni T. Interpersonal motor synchrony in autism: a systematic review and meta-analysis. Frontiers in psychiatry. 2024; 15:1355068. https://doi.org/10.3389/fpsyt.2024.1355068 PMID: 38439792

43. Tryfon A, Foster N.E., Ouimet T., Doyle-Thomas K., Anagnostou E., Sharda M., et al., for NeuroDevNet ASD imaging group. Auditory-motor rhythm synchronization in children with autism spectrum disorder. Research in Autism Spectrum Disorders. 2017; 35:51–61.

44. Kliger Amrani A, Zion Golumbic E. Spontaneous and stimulus-driven rhythmic behaviors in ADHD adults and controls. Neuropsychologia. 2020; 146:107544. https://doi.org/10.1016/j.neuropsychologia.2020.107544 PMID: 32598965

45. Alruwaili MM, Abuadas FH, Alsadi M, Alruwaili AN, Elsayed Ramadan OM, Shaban M, et al. Exploring nurses' awareness and attitudes toward artificial intelligence: Implications for nursing practice. Digit Health. 2024; 10:20552076241271803. https://doi.org/10.1177/20552076241271803 PMID: 39114115

46. Sidelil H, Demissie A, Demissie G.D., Fikade B., Hailegebreal S., Tilahun B. Attitude towards tele rehabilitation-based therapy services and its associated factors among health professional working in

specialized teaching hospitals in Amhara region, Northwest Ethiopia. Informatics in Medicine Unlocked. 2023; 36.

47. Oh S, Hodson T, Kendall M, Shirota C, Kendall E. A technology-focused service within a rehabilitation setting: exploring the experiences and current technology demands of people with disabilities. Disabil Rehabil Assist Technol. 2024; 19(5):2000–9. https://doi.org/10.1080/17483107.2023.2244001 PMID: 37615165

48. Consing AMU, Aguila M.E.R. A survey on attitutes of rehabilitation professionals on the use of immersive gamification technology system in therapy. Philippine Journal of Physical Therapy. 2024; 3(1):31–.

49. Lui YW, Xue Y, Kenul D, Ge Y, Grossman RI, Wang Y. Classification algorithms using multiple MRI features in mild traumatic brain injury. Neurology. 2014; 83(14):1235–40. https://doi.org/10.1212/WNL. 0000000000000834 PMID: 25171930

50. Muller A, Vallee-Marcotte J, Robert-Lachaine X, Mecheri H, Larue C, Corbeil P, et al. A machine-learning method for classifying and analyzing foot placement: Application to manual material handling. J Biomech. 2019; 97:109410. https://doi.org/10.1016/j.jbiomech.2019.109410 PMID: 31648789