COMENIUS UNIVERSITY, BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY
Univerzita Komenského
v Bratislave

universität wien

THE SOUND OF MOVEMENT: DOES MUSICAL STRUCTURE GUIDE LINDY HOP DANCERS

MASTER'S THESIS

2026

BSC. SARAH MARIE WINGERT

COMENIUS UNIVERSITY, BRATISLAVA

FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS

THE SOUND OF MOVEMENT: DOES MUSICAL STRUCTURE GUIDE LINDY HOP DANCERS

MASTER'S THESIS

Study program:	Cognitive Science
Field of study:	2503, Cognitive Science
Supervising department:	Department of Applied Informatics
Supervisor:	RNDr. Kristína Malinovská, PhD.
Consultant:	Anja-Xiaoxing Cui, PhD.

Bratislava, 2026

Bsc. Sarah Marie Wingert

20507989

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname: Sarah Marie Wingert
Study programme: Cognitive Science (Single degree study, master II. deg., full time form)
Field of Study: Computer Science
Type of Thesis: Diploma Thesis
Language of Thesis: English
Secondary language: Slovak

Title: The Sound of Movement: Does musical structure guide Lindy Hop dancers

Annotation: Music and dancing have been an integral part of human society since ancient times and are intertwined with many different cognitive faculties [1, 2]. Dancing with a partner, as in Lindy Hop, requires the complex anticipation of another's intentions based on bodily cues, musical timing, and shared social context [3]. The required cognitive processing can either be facilitated or impeded by music. The aim of this thesis is to investigate how different sections of a jazz ensemble influence the artistry of Lindy Hop dancers using motion capture data via a mixture of phenomenological and quantitative approaches. This entails labeling the data with the guidance of an experienced dancer. This labeling is to reveal the abstract structure of the dance to be quantitatively analyzed.

Aim:
1. Explore the inner structure of Lindy Hop and its cognitive counterparts.
2. Explore the influence of parts of the jazz ensemble on the artistry of dance.
3. Identify which ensemble parts are essential for which aspects of dance artistry.

Literature:
[1] M. Pearce and M. Rohrmeier, "Music Cognition and the Cognitive Sciences," Topics in Cognitive Science, vol. 4, no. 4, pp. 468–484, Oct. 2012.
[2] B. Burger and P. Toiviainen, "Embodiment in Electronic Dance Music: Effects of musical content and structure on body movement," Musicae Scientiae, vol. 24, no. 2, pp. 186–205, Aug. 2018.
[3] J. Phillips-Silver, C. A. Aktipis, and G. A. Bryant, "The Ecology of Entrainment: Foundations of Coordinated Rhythmic Movement," Music Perception, vol. 28, no. 1, pp. 3–14, Sep. 2010.

Supervisor: RNDr. Kristína Malinovská, PhD.
Department: FMFI.KAI - Department of Applied Informatics
Head of department: doc. RNDr. Tatiana Jajcayová, PhD.

Assigned: 23.05.2025

Approved: 19.06.2025 prof. Ing. Igor Farkaš, Dr.
Guarantor of Study Programme

....
Student Supervisor

20507989

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta:	Sarah Marie Wingert
Študijný program:	kognitívna veda (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor:	informatika
Typ záverečnej práce:	diplomová
Jazyk záverečnej práce:	anglický
Sekundárny jazyk:	slovenský

Názov: The Sound of Movement: Does musical structure guide Lindy Hop dancers
Zvuk pohybu: riadia sa tanečníci Lindy Hopu hudobnou štruktúrou

Anotácia: Hudba a tanec sú neoddeliteľnou súčasťou spoločnosti a sú prepojené s kognitívnymi kapacitami [1, 2]. Tanec s partnerom nutne zahŕňa komplexné predpovedanie úmyslov partnera na základe telesných signálov, hudobného načasovania a sociálneho kontextu [3]. Kognitívne procesy potrebné pre párové tancovanie sú ovplyvnené hudbou. Cieľom DP je skúmať, ako rôzne sekcie jazzovej skladby ovplyvňujú umelecké výkony tanečníkov Lindy Hopu s využitím nasnímaných pohybových dát prostredníctvom kombinácie fenomenologických a kvantitatívnych prístupov. Zahŕňa anotovanie údajov pomocou metódy navrhnutej v spolupráci so skúseným tanečníkom a následné odhalenie abstraktnej štruktúry tanca pomocou kvantitatívnej analýzy.

Cieľ:
1. Preskúmajte vnútornú štruktúru Lindy Hopu a jeho kognitívnych aspektov.
2. Preskúmať vplyv rôznych častí jazzovej skladby na umeleckosť tanca.
3. Určiť, ktoré časti skladby ovplyvňujú vybrané aspekty tanečného umenia.

Literatúra:
[1] M. Pearce and M. Rohrmeier, "Music Cognition and the Cognitive Sciences," Topics in Cognitive Science, vol. 4, no. 4, pp. 468–484, Oct. 2012.
[2] B. Burger and P. Toiviainen, "Embodiment in Electronic Dance Music: Effects of musical content and structure on body movement," Musicae Scientiae, vol. 24, no. 2, pp. 186–205, Aug. 2018.
[3] J. Phillips-Silver, C. A. Aktipis, and G. A. Bryant, "The Ecology of Entrainment: Foundations of Coordinated Rhythmic Movement," Music Perception, vol. 28, no. 1, pp. 3–14, Sep. 2010.

Vedúci:	RNDr. Kristína Malinovská, PhD.
Katedra:	FMFI.KAI - Katedra aplikovanej informatiky
Vedúci katedry:	doc. RNDr. Tatiana Jajcayová, PhD.
Dátum zadania:	23.05.2025
Dátum schválenia:	19.06.2025

prof. Ing. Igor Farkaš, Dr.
garant študijného programu

....
študent

....
vedúci práce

Declaration

I, Sarah Marie Wingert, hereby declare that the Master's thesis with the title "The Sound of Movement: Does musical structure guide Lindy Hop dancers" is my own original work and the result of my own research. Formulations and ideas taken from other sources are cited as such.

Karlsruhe, 10.01.2026 ______________________________

Bsc. Sarah Marie Wingert

Acknowledgements

I would like to thank all those who have enabled me to write this Master's thesis.

First and foremost, I give my thanks to Sumner Williams, who has not only agreed to share his experimental data with me but also spent many hours discussing the different concepts and cognitive representations in Lindy Hop from the perspective of an experienced dancer and teacher. Without him, I would not have had the domain knowledge necessary to complete this thesis.

I would also like to thank Anja Xiaoxing-Cui for guiding me through the development of my hypotheses and advising me on the statistical methods used. Her feedback pointed me to potential issues with my hypotheses before they arose and allowed me to react early on. I also give my thanks to Kristína Malinovská for her guidance and feedback on my thesis.

I also owe thanks to the MediaLab, especially Christoph Reuter and Jörg Mühlhans, who introduced us to the MotionCapture technology and let us use their equipment for the recordings. I would also like to thank them and the University of Vienna for providing the software for viewing Motion-Capture files, as well as the host server for the Lindy Hop website. An additional thanks to Christoph Reuter for his help in explaining the meaning of the audio features to me in terms that the musician in me could understand.

Last but not least, I want to thank my friends Marvin Marz and Henrike Arnsberg for accompanying me through the process of writing my thesis and for listening to all my grumbling, as well as my nephew Tobias Kathmann for giving feedback on my presentations. I would also like to thank Marvin Marz and Johannes Bechberger for their feedback on my writing.

Abstrakt

Hudba a tanec sú neoddeliteľnou súčasťou kultúr sveta. Slúžia na náboženské, spoločenské a emocionálne účely a sú prepojené s rôznorodou škálou kognitívnych funkcií. Je známe, že tanec priamo reflektuje hudbu, na ktorú sa tancuje. Štúdie potvrdzujú, že hudba s vysokým podielom synkop a zložité pohyby môžu znížiť presnosť načasovania tanečníkov. To naznačuje problém limitovanosti kognitívnou kapacitou, z čoho vyvstáva otázka, ako tanečníci využívajú znalosť hudobného diela pre sledovanie rytmu. Táto diplomová práca skúma vplyv jednoznačnosti rytmu na tanec v kontexte improvizovaného Lindy dance Hop pri 12 rôznych podnetoch, ktoré tvoria tri swingové skladby, rozdelené na kategórie *celá skladba*, *pozadie*, *rytmus* a *sólo*. Podnety boli prezentované v dvoch etapách, pričom v každej boli štyria tanečníci. Všetci tancovali spolu v oboch roliach (nasledovník a vedúci), čo viedlo k celkovému počtu 24 párov. Pohyby tanečníkov boli zaznamenané pomocou MotionCapture a vo všetkých klipoch boli označené signály vedúceho a pohyby *na frázu*, ako aj rôznorodosť vykonaných pohybov. Dáta boli analyzované pomocou lineárnych modelov. Prediktormi bola napr. vybraná skladba, ale aj kvantitatívne merania, napr. rytmická nepravidelnosť. Signály vykonané na frázu vykazovali významnú koreláciu s jasnosťou rytmu. To naznačuje, že jasnejší rytmus umožňuje tanečníkom tancovať aktívnejšie. Najväčšia variabilita sa odvíjala od konkrétnej dvojice tancujúcich, čo naznačuje, že práve kombinácia ľudí mala oveľa väčší vplyv na ich tanec ako hudba sama.

Keywords: Vtelené vnímanie hudby, vtelený výskum, snímanie pohybu, Lindy Hop, teória pozornosti „bottleneck“

Abstract

Music and dance are an integral part of cultures worldwide. They serve a variety of religious, social, and emotional purposes and are intertwined with diverse cognitive functions. Dance has been shown to directly reflect music. Studies also show that highly syncopated music and complex movements can both reduce dancers' timing accuracy. This indicates limited cognitive resources for performing the dance and thus raises the question how dancers rely on different parts of an ensemble to track the beat in music. This thesis examines the influence of beat clarity on dancing in improvised Lindy Hop dance to 12 different stimuli. Three Swing songs were separated into the *Full Ensemble*, *Background*, *Rhythm*, and *Soloist* sections (the Conditions). The stimuli were presented in two different sessions to four dancers in each, who all danced with each other in both roles (follower and leader), resulting in a total of 24 dyads. MotionCapture data was recorded, and all clips labeled for Leader Cues and Moves per phrase, as well as Richness of moves performed. Linear Models were run with Condition (and Song) or quantitative measures, such as Rhythmic Irregularity, as predictors. Cues *per phrase* showed significant correlation with beat clarity. This indicates that a clearer beat enables dancers to dance more actively. The greatest amount of variance was explained by the dancing dyad, indicating that the people dancing had a much greater effect than the music itself.

Keywords: Embodied Music Cognition, Embodied Research, Motion Capture, Lindy Hop, Bottleneck Theory of Attention

Contents

List of Figures

List of Tables

List of Equations

Chapter 1

Introduction

Music and dance have been an integral part of human society since ancient times and may play an adaptive role in human cognitive development [7, 8]. They are found across all cultures and are therefore a universal human trait. Additionally, they serve many different functions, socially, culturally, and cognitively. Music serves as a form of social bonding and emotional self-regulation, is present in many religious contexts, and is also accompanied by an aesthetic experience [7]. Dance can also serve social bonding and emotional self-regulation, as it is more enjoyable to dance together than alone [9].

Music can be understood as inherently embodied. Multiple studies show that the movement of participants not only reflects the beat of the music but also other aspects like melody [8], harmony, the rhythm of the music itself, and the listener's mood. It has been shown that the rhythm perception is strongly modulated by body movement as well [10]. But the body can also serve as an external timekeeper, as multiple studies suggest [11, 12].

Music and dance are compelling topics in Cognitive Science, as they require attentional processes [13] alongside complex perceptual, cognitive, and emotional processes [7]. Humans are also the only species capable of beat induction in music, even predicting it, before it sounds. While other species can perceive beat, they are not able to anticipate future beats or recognize rhythm [14, 15]. Even algorithms struggle to accurately determine a beat, making this a unique human characteristic.

Dance becomes even more complex in coupled dances, which require coordinated rhythmic movement between partners, an elaborate cognitive process [16]. These dances also highlight interpersonal communication and leading dynamics, as they usually have a fixed leader and follower, and the leader must plan their actions in advance so they can signal them to the follower [17]. This requires much more deliberate motor planning than purely relying on instinct, as it is not enough to react to the music in the moment.

It is no surprise, therefore, that different demands, such as processing music and planning future motor actions, can compete for limited cognitive resources. Attention-demanding tasks (such as highly complex movements) can reduce timing accuracy [13]. At the same time, highly syncopated rhythms can also worsen the timing [18].

This thesis aims to investigate how dancers improvise in response to different stimuli and what role the mentioned cognitive bottleneck may play. For that, we chose the dance style Lindy Hop, as much previous research has focused on predominantly Western dance styles, and this dance, which is based in African-American Jazz culture, may provide a new perspective on previously conducted studies [19]. The dance is also marked by a high kinetic energy and a high level of improvisation, making it much less structured than traditional ballroom dances [20]. As Lindy Hop is traditionally danced with frequent changes of dance partners, Lindy Hop dancers are practised in dancing with multiple other participants as required in the experimental setup.

In this thesis, we aim to extrapolate the underlying schemata of the dance and make quantitative comparisons based on them, rather than simply analysing physical properties such as overall movement. For this purpose, we collected MotionCapture recordings during experimental sessions in which participants were exposed to 12 different stimuli. We use a previously developed schema to label moves and identify cues. We also combine this with the author's embodied approach, taking Lindy Hop classes over the span of nine months, to ensure the labels are based on a thorough understanding of the dance.

The research questions are:

> RQ1: What part of the jazz ensemble facilitates artistry? Artistry translates to both the diversity of moves and the complexity of moves.
>
> RQ2: How commonly are moves of each family danced? Is there a relationship between ensemble part playing and an increased occurrence of a specific family?

The previous research can be interpreted in several ways. It is possible that a clear beat is the most crucial factor for dancing, as it leaves more cognitive capacity for the movement execution. It might also be that melody is the most important, as it may enhance the creativity needed to form ideas for movement. Lastly, dancers may also require both beat and melody to have the creativity and the cognitive capacity to perform moves.

Based on this, the hypotheses are:

- H0: There are no effects of the jazz ensemble on the dance style.
- H1: Melody encourages artistic dancing.
- H2: Beat encourages artistic dancing.

- H3: Beat and melody must interact for artistic dancing.

Artistic dancing is quantified in four different ways:

- The number of cues performed
- The number of moves performed
- The diversity of moves performed (moves from many different families)
- The complexity of moves performed (amount of cues per move)

These hypotheses are tested using linear mixed models on data from two experimental sessions, each with four participants. As each participant danced with each other within one session, there is a total of 24 dyads dancing to 12 different stimuli, resulting in 288 Motion-Capture clips.

Section 2 provides a review of previous research on music and dance. It introduces *Embodied Music Cognition* (Chapter 2.1), the main paradigm guiding this research. We then look at the attentional and cognitive processes that have been associated with music or dance, but also at aspects of social cognition (Chapter 2.2). An outline on *beat induction* and interpersonal synchronization is given and the necessary environment and potential origins are discussed (Chapter 2.3). Chapter 2.4 describes previous research that investigated the quantitative effects of music and social context on dance. We then discuss the choice of the dance style of Lindy Hop for this research and the reasons behind it (Chapter 2.5), as well as give a general overview of the dance (Chapter 2.6).

Section 3 lays out the methodology of this research. The aim of the study, the research question, and the possible hypotheses are laid out (Chapter 3.1). It is described how the stimuli were produced (Chapter 3.2). Assumptions are statistically tested, and the implications for the research are discussed. It then describes the requirements and the recruiting process of the participants (Chapter 3.3), the general process of each experimental setup (Chapter 3.4), and the setup of the MotionCapture hardware and the process used for cleaning the collected data (Chapter 3.5). Chapter 3.6 explains how the inherent grammar of Lindy Hop was extracted, which is used in this thesis's labelling process. A detailed overview of the labelling process is given in Chapter 3.7.

Section 4 presents the study's results. It outlines the demographics of the participants (Chapter 4.1), examines the inter-rater-reliability analysis (Chapter 4.2) and checks for differences in variance between and within leaders (Chapter 4.3). It then describes the overall distributions of move families as well as potential differences between conditions (Chapter 4.4), tests for normality of the data and outlines general descriptives (Chapter 4.5) and finally tests the hypotheses with Linear Mixed Models (Chapter 4.6).

Section 5 looks at the results in light of the defined hypotheses (Chapter 5.1) and dives into the importance of this study and the implications that can be drawn (Chapter 5.2). It also contains an embodied report of the author on dancing in different social cultures (Chapter 5.3) and discusses limitations of the study (Chapter 5.4). In Chapter 5.5 possible endeavours for further research are outlined.

Section 6 provides a final overview of the findings found in this work and what they mean for future research.

Chapter 2

Theoretical Framework and Previous Research

This section explores previous research and discusses how our assumptions for this research develop out of previous findings.

Chapter 2.1 examines the historical and cross-cultural link between music and dance through the lens of *Embodied Music Cognition*. It discusses multiple studies that demonstrate how movement reflects music on multiple levels, including beat, melody, timbre, and emotion, and elaborates on various arguments for how movement influences music perception. Critiques of *Embodied Music Cognition* are mentioned, and the most common theories in the field are described.

Chapter 2.2 examines the importance of music and dance to the Cognitive Sciences. It reviews how music and dance engage a multitude of perceptional, attentional, emotional, and social systems. It also describes how dance expertise and training have been found to influence brain plasticity. We discuss the processes necessary for improvisation and the positive side effects that dancing may have. Lastly, this chapter introduces the cognitive demands of dancing and the concept of a cognitive bottleneck, which is essential to one of our hypotheses.

In Chapter 2.3 we look at the ability of humans to synchronize with a beat and with each other, both essential abilities for dancing together. This chapter investigates what musical features are relevant for *beat induction* and how *beat induction* might develop in infancy. It also reviews how other animals and computers perform in the field of beat induction and how culture and our embodiment may influence our ability to perceive a beat despite noise, and what meter we perceive. It then examines the cues dancers can use to synchronize with one another. We discuss the importance of social context and shared intent, and discuss the special aspects of Lindy Hop in that regard. Lastly, the cognitive bottleneck theory, as proposed in this research, is further specified in the context of Lindy Hop.

Chapter 2.4 describes previous research on the components of the music and the social context that influence dancing. It discusses the influence of different rhythmic elements and melodic instruments, and how a clear beat and melody influence the dancing in measures of overall movement and movement speed. We review how different metrical hierarchies are reflected in different body movements and how a person's personality or the general genre might affect the overall activity.

In Chapter 2.5 we discuss the choice of the dance style of Lindy Hop for this research and the reasons behind it. This chapter discusses the lack of research on non-Western music and how the origins of Lindy Hop in African-American culture offer new insights. It introduces the concept of groove and the aspects of Lindy Hop's nature that make it suitable for research.

Chapter 2.6 gives a general overview of the dance. It briefly describes its origin and cultural development, then delves into the general semantics of the dance. It introduces the two main positions in Lindy Hop Open and Closed, as well as the 6-count and 8-count patterns. This chapter also discusses variations of Lindy Hop and sets the frame for the version danced in this research.

2.1 Music and Dance: Embodied Music Cognition

Music and Dance have always been closely linked [10], and dance can be found across many cultures worldwide as a prime form of human expression [9]. In fact, many researchers believe that dance and music have evolved together and that music gets perceived through the movement of the body. This is called *Embodied Music Cognition* [8].

Embodied Music Cognition holds that moving our bodies strongly influences how we perceive, feel, experience, and comprehend music. This means we do not perceive music as a passive observer; we need to actively engage with it to fully understand it. One classic example is the Samba dance, whose musical structure is ambiguous and hard for Western listeners to comprehend. Only through the structure of the dance movements does the rhythm become "disambiguated". If, as the paradigm suggests, our perception of music is bound to our bodies interacting with the environment, then our understanding of music is generally constrained by our own physical and cultural limitations [21].

Many studies have shown a strong effect of music on bodily movement. It has been found that a movement is easier to perform if it follows the dynamic of the music itself (e.g., down movement on the down beat) [22]. It has also been found that movement reflects not only the direct beat of the music but also its melody, harmony, rhythm, and timbre. The body of dancers can reflect multiple levels of metrical hierarchy simultaneously [16], e.g., bobbing the head to the beat while reflecting the whole bar with horizontal movement [5]. In addition to the direct influence of music on dance, the emotional state of dancers also shapes their movement [12]. Not only that, but observers of a dance can also induce the felt emotions based on pure observation of the body movements [11]. Interestingly, a study found that music-induced body movement correlated with verbal descriptors of how the music was perceived. This indicates that movement to music can be used to infer the dancer's perception and cognitive representation of the music [12].

Despite these studies linking music to movement, it is important to note that the *Embodied Music Cognition* paradigm has also been criticized. Some claim it is formulated too vaguely, and music influencing the body does not prove that the perception of music relies on the body. Others criticize the lack of biological groundwork on why we perceive music the way we do [23].

Some studies can partially alleviate those critiques by finding clear influences of the bodily movement on the music perception. For example, when participants in a study were asked to solve cognitively demanding tasks while playing either *staccato* (separating notes) or *legato* (connected notes), the timing of their beat was disrupted by the cognitive task in the case of *staccato* but not in *legato*. The researchers concluded that this might indicate the use of the body as an additional time-keeper, outsourcing cognitive work. In the *legato* condition, the musicians' arm could move at a steady pace, matching the next beat on its own, without the need for additional cognitive maintenance, while this was not possible in the *staccato* condition due to the disconnect between notes [11].

Other studies also directly show that movement influences music perception, as it can improve time perception and help maintain a beat [12]. It can modulate rhythmic interpretation, sometimes even *metric* interpretation (e.g., Samba) [10]. It might also pull attention to certain musical cues [11].

There are also attempts to address the lack of biological groundwork. Some theories suggest that music perception is based on other bodily functions. For example, the pulse of music might be based on periodic activities that all humans engage in, such as breathing, walking, or our heartbeat [19]. There was even found a socalled "eigen-frequency" of 2 Hz that could be understood as our anchor for judging tempi as fast or slow [24]. This eigen-frequency has been confirmed in numerous studies on walking [11].

As head movement has been found to strongly correlate with the beat of music [5], some researchers also suggest that we might actively use our *vestibular system* in the ear to sync up with the beat of the music by stimulating it with our head movements [25]. This is also supported by the heavy use of bass drums in contemporary music and the finding that participants were better entrained to the tempo after an increase in bass drum volume, as the low vibrations of bass drums might directly affect the *vestibular system* [16].

In essence, music strongly influences movement, and movement at least partially influences the perception of music. While the paradigm of *Embodied Music Cognition*, like the paradigm of Embodied Cognition in general, has its critics, it seems at least evident that the cognitive processing of music is not independent of the bodily movements and that these movements may shed light on the cognitive processes underlying them.

2.2 The Cognitive and Attentional Processes in Music and Dance

Music is a compelling topic in Cognitive Sciences because it is not only found across all cultures and is therefore a universal human trait, but it also serves many different functions. Music is used as a form of social bonding, for emotional self-regulation, in the mother-infant interaction, in some healing methods, and in religious rituals. Even apart from those social contexts, music is also an aesthetic experience [7]. Music can also help manage pain, enhance brain function, aid memory recovery in neurodegenerative disorders, and support neural development in premature babies [26].

Music and dance recruit many different perceptual, cognitive, and emotional areas, but dance training also changes how the brain functions over time. Music perception alone already involves processing complex and parallel temporal information, combining both local (e.g., the currently played note) and hierarchical structures (e.g., the key we are in). It involves a wide range of cognitive functions, including auditory scene analysis, streaming, attention, learning and memory, the formation of expectations, the integration across multiple modes (e.g., music and dance), recognition, syntactic processing, emotional processing, and, in some contexts, social cognition. Sound is not processed in isolation from the environment: While listening, our brains actively try to locate the source of a sound, and they might even split processing into separate streams for different sources that are perceived separately, such as multiple instruments in an orchestra [7].

While the sound is filtered for emotional and syntactic information, we also simultaneously need to determine the *meter* via *beat induction* (see Ch. 2.3), so that we can coordinate our own bodies rhythmically with it. Studies have shown that this also recruits motor areas for predicting the next beat, indicating that perception and action are interrelated. This also influences dance performed to the music [7]. We must consider these complex processes when investigating dance.

There have also been studies of cognitive processes in dance. For example, it was found that we not only perceive individual body parts but also their configuration, meaning that when we observe dance, we may perceive the entire body as a whole [9].

Another finding was that, unlike athletics, dancing requires not only physical capabilities, but also artistic ones [13]. This is especially the case in improvisation, a special kind of motor performance, where dancers do not follow a rigid routine but are constrained only by the inherent grammar of the dance and the music they improvise to. Improvising to music requires dancers to process incoming sensory information, evaluate possible actions that fit, and finally choose one and execute it [27]. The extent to which brain areas for creative thinking are active during improvisation depends on the dance style, with less-constrained styles eliciting greater activity [9].

Dance has also been shown to be a good source for studies on motor performance and expertise. Previous research has shown that activation in the mirror system was higher when people observed dances they had experience with, indicating that our experience shapes our movement repertoire. Intensive dance training also decreases brain matter volume in motor areas and increases sensitivity to subtle changes in movement [9].

Dancing also has many positive side effects. Dancing also has many positive side effects. It can slow cognitive deterioration, and dancing with a partner leads to greater pleasure than dancing alone [9]. It also enhances sensorimotor control, improving balance both when balancing statically and dynamically (during movement). It improves proprioception, enabling people to rely more on their bodily signals when balancing dynamically, rather than relying on visual cues. Studies have also found that dance training causes synergies, in which a single neuronal activity becomes responsible for an entire group of muscles, reducing tension and improving movement accuracy [13].

An important issue that needs to be pressed here is that movement combinations often require attentional processes and cannot be studied independently of them. Attention-demanding tasks (such as highly complex movements) can reduce timing accuracy. This can be mitigated through physical contact between the dancers and by the practice of that movement [13]. The timing can also worsen if the rhythms are highly syncopated, as this is an additional cognitive demand [18].

This struggle for limited cognitive resources is a key prerequisite for the following thesis. Our assumption that different musical stimuli will change the number and complexity of moves danced is inherently based on the idea, supported by those studies, that greater processing power needed for the music will leave fewer attentional resources for the planning and execution of body movement. This puts this study at an interesting intersection, investigating how the complexity of the music directly influences free dancing.

2.3 Beat Induction and Coordinated Rhythmic Movement

To dance, be it alone or together, one must be able to perceive a beat and a rhythm in music. Most humans can easily tap to a beat [28], and even infants have been found to perceive a beat and most likely also rhythm [29]. Even some animals seem able to synchronize to a beat, though their skills are not as thorough as humans' [14]. Humans can determine the beat without knowing tempo or *metrical system*, regardless of possible noise (changes in tempo or timing), and they can cope with *syncopation*, strong notes that conflict with the metrical schema of the piece. In fact, they tend to even perceive *syncopation* as interesting [15].

This makes it especially interesting that computer algorithms have been struggling for years to accurately determine a beat [15] and that there has been extensive research on the best algorithms [28]. Naturally, this begs the question: Why are we so capable at something that computational algorithms struggle to achieve?

There are different theories on how the ability to induce a beat develops. As it forms in the first months of life, some researchers assume that children learn to associate bodily movement with auditory cues by being rocked to music by their parents [30]. This relates to the theory that we use our vestibular system to synchronize with music, as mentioned in Ch. 2.1. There is definitely also some influence of past experiences, as trained musicians can tap to a beat more accurately and synchronize faster [28]. It has also been found that the ability to keep the beat to a highly syncopated stimulus can be trained by moving to the music [25]. That we are shaped by our experiences also becomes clear when we look beyond the direct beat onto the perception of rhythm and *meter*. The *meter* is cognitively interpreted based on previous experience [29]. A person who grew up with classical music might interpret a piece with a wildly different *meter* than someone who grew up with traditional African polyrhythms.

Several researchers have investigated which musical features are relevant to *beat induction* (the ability to align to a beat) in humans. *Beat induction* is generally based on periodically appearing musical events, but it is robust to slight deviations in event timing. The perception of those metric accents can also be influenced through so-called "phenomenal" accents. These include sudden changes in the dynamics or timbre, long notes, pitch leaps, as well as harmonic changes. For example, changes in the pitch of low tones in a Ragtime did influence the perception of a downbeat [31]. This indicates that even though of less effect, even the melody of a piece contains some information about the beat.

Our ability to synchronize to a complex rhythm is the foundation for *complex coordinated rhythmic movement*, such as dancing with a partner [30]. This does not only require us to align our motor actions with the external musical rhythm [16], but also with the rhythm provided by our partner. The coordination between partners can be improved by reading each other's intentions and predicting future actions based on them. Such can be inferred from eye gaze, vocal signals, body posture, or even explicit instructions (e.g., the leader pushing the follower backward).

It is also advantageous if the dancers share a social context [30]. An example of the importance of that can be read in Ch. 5.3, where we describe the difficulties of dancing Lindy Hop with someone who had learned the dance in a different country and social context. It is not unlikely that dancers are even required to form a kind of shared intentionality that exceeds the individual to collectively dance as a unit [30].

It is important to note that in an improvisational dance like Lindy Hop, the dancers need to communicate and make decisions on the spot. This requires not only the anticipation of the other's intents but also constant attention and fast adaptation. The success in coordinated dancing depends on both familiarity with the other dancer and one's ability to respond to errors and adapt to reestablish correct timing [17]. One good example is when one performs a complex move to an ambiguous rhythmic stimulus and loses track of the beat. Everyone who dances once in a while should know the moment of hesitation that follows, as they attempt to reorient to the beat and the partner and continue in sync.

What must also be noted for dances with a clear leader and follower is that the leader must plan in advance what they wish to do, so that they can give the physical cues on time [17]. After all, the follower needs to know what they ought to do before they actually need to do it. Our second assumption is that planning in advance requires attentional resources, as the leader must envision complex movements before they occur. Therefore, the amount of cues provided by the leader also informs how we measure the complexity of moves.

2.4 Quantitative Effects of Musical Features and Social Factors on Dance

There has been extensive prior research on how musical stimuli and social context relate directly to quantitative measures of dancing. The most prominent feature, with the greatest impact, is the bass drum. A louder bass drum led to more energetic movement in some studies [16], to more movement overall in others [32], and to faster movement in another study [5]. It was also found that low-frequency rhythmic elements (e.g., the bass drum) increased the urge to move [8]. The presence of a bass drum also tends to increase listeners' spontaneous movement [5].

Higher-frequency components and percussive elements (e.g., the hi-hat) were associated with greater overall movement, whereas the bass drum affected movement speed. High-frequency rhythmic elements were also associated with head and foot movements. The attack length of instruments negatively correlated with head and hand movement [8]. As melodic instruments tend to have a longer attack length and rhythmic elements a shorter one, this effectively means that more rhythmic elements meant more movement.

When looking at the clarity of beat in general, studies found that participants increased local movement with a clear beat and moved slower and more on the spot with a strong rhythm [14]. People also tend to "wander around" more if there is little change in the sub-band 2 frequency (low bass guitar, kick drum). When the beat is not clear, there is more hand movement, and it is more accelerated and "jerky" [32]. Pulse clarity also facilitates spontaneous movement, as the beat is easier to detect [5].

In general, vertical movement (up and down) is most often synchronized on the tactus level (on each beat), while horizontal movement (swinging sideways) is more often related to higher metrical levels, like a whole bar. Strong low-frequency rhythmic structures (e.g., bass drum) only correlate horizontally with the bar level and vertically with the tactus (per beat). High-frequency flux (the hi-hat) encourages people to move side-ways. The more a stimulus is in a major key, the faster participants move their heads, and the clearer the key is, the wider the hands are apart [32].

There have also been some studies on social factors influencing dancing. For example, it has been found that female participants moved more actively overall to electronic dance music (EDM) than male participants [16]. The personality of people also influences their movement [14] with extraverted people having greater global and local body movements [25]. Participants also tend to fit the beat more closely and move more actively when they are dancing together in a social context. They also entrain more with each other in that setting [16].

To summarize, a clear beat and a strong bass drum encourage more movement. The beat is mostly represented in vertical movement, with horizontal movement reflecting higher hierarchical metrics, and a lack of direction in the lower frequency band leads to more "wandering around". Some studies investigated EDM [16, 8], while others investigated popular music across different genres [32, 5]. There is little research specifically on jazz music. However, a comparative study found that jazz generally has less movement across all tested parameters than EDM [8]. This shows why further research into the influence of musical stimuli from the genre of jazz music is important.

2.5 The Motivation for Investigating Lindy Hop

There are multiple motivations for investigating the Lindy Hop dance style specifically. The most obvious one is the heavy leaning towards Western music and dances in research. There is reason to believe that cultures influence what role the body plays and that this will also influence music styles.

For example, the musical grammar developed based on Western tonal music lacks an accurate description of the concept of groove [19], as the idea of groove originated in African-American Swing music. Groove is described as a subjective experience that encourages spontaneous movement and feelings of enjoyment. Music is perceived as "groovy" when it is marked by a salient, low-pitched beat, moderate rhythmic complexity, and a medium tempo of about 120 *beats per minute (BPM)* [25]. As the origin of groove, Jazz music offers an interesting alternative perspective on music and dance, thanks to its microrhythmic structure that is not found in Western music [19].

Lindy Hop, in particular, is of interest to us because it developed as a partner dance in the 1930s in the African-American working community [33] and is characterized by much higher kinetic energy than traditional dances like the Fox-Trot or One Step. As it uses the entire body, it is physically more demanding than earlier social dances. It is also a dance that heavily relies on improvisation [20] and is constantly developing.

What also makes it interesting is that it is seen as a cultural and social dance, marked by its unstructured nature and being danced for leisure rather than competitively [33]. As the dance partner is frequently changed, it requires a certain flexibility in reading other people's cues and intentions due to the lack of fixation on one specific person's style.

In summary, it is a great topic for research, as there is currently a lack of studies of non-Western dance styles. It also enables us to study interpersonal communication, as people are not biased toward one specific dance partner, and the high level of kinetic energy and the highly improvisational nature of the dance require clear and active communication. As Lindy Hop tends to borrow from other dance styles quite heavily, no two dancers have the exact same social context and might sometimes need to rely on intuition to follow each other's signals. The high groove of Swing music also encourages the dancers to move spontaneously and naturally, rather than to a fixed routine.

2.6 An Introduction to Lindy Hop

Lindy Hop is a *vernacular Swing Dance* that developed in the 1930s in the African-American working community [33]. During the revivalist movement in the 1990s, it shifted to a dance predominantly danced by the white middle class [34]. The dance usually has a follower and a leader; however, there are specific courses where the dancers learn to dance in both roles [35], and it is common to fill in the follower's role at events to be more flexible.

It generally alternates between Open and Closed position. The Open position is marked by much more improvisational freedom (especially on the follower's side), while the Closed position is more restrictive on the follower's improvisation [10]. It is important to note, however, that Lindy Hop is a very diverse dance. It does not just have one Open and one Closed position. There are many other positions that the dancers can connect in, and these positions can be seen as either on the open side (high improvisational freedom) or on the closed side (low improvisational freedom). In previous research, we categorized all identified positions into the categories Open, Closed, and Unspecified. More on that can be found in Ch. 3.6.

The general base steps in Lindy Hop include a *6-count* and an *8-count* pattern (counted in 8ths). The majority of the move repertoire is based on a *6-count* pattern, which contrasts with the 4/4 rhythm of the music itself. This leads to much longer cycles of sections due to the shift of emphasis between bars [10, 35]. However, some moves in Lindy Hop may be shortened to 4 counts or extended to 8 counts, and it is not uncommon for Lindy Hoppers to improvise footwork variations in spot to fill in the missing amount of counts before the end of a phrase [35]. There might also be some cultural differences in the preference for *6-count* vs. *8-count*, as described in Ch. 5.3.

There are different variations of Lindy Hop, like the Charleston Lindy or the "Fast Lindy" (not an actual name), which primarily uses the *8-count* pattern and is danced to very fast songs. For the purpose of this research, we asked participants to dance only standard Lindy, not Charleston Lindy, to keep the different rounds comparable. None of our stimuli was fast-paced enough to encourage Fast Lindy.

Chapter 3

Methodology

This section lays out the methodology of this research. In Chapter 3.1 the aim of the study, the research question, and the possible hypotheses are laid out. Chapter 3.2 describes the stimuli in more detail and statistically tests assumptions. It also discusses the implications for the research. Chapter 3.3 describes the requirements and the recruiting process of the participants. Chapter 3.4 describes the stimuli used and the general process of each experimental setup, while Chapter 3.5 describes the setup of the MotionCapture hardware and the process used for cleaning the collected data. Chapter 3.6 dives into some previous work that was done as preparation for this thesis. It explains how the inherent grammar of Lindy Hop was extracted and provides a general overview. This grammar is used in this thesis's labeling process. In Chapter 3.7 a detailed overview is given of the labeling process of the author themselves, as well as the instructions given to the external rater. Chapter 3.8 explains the statistical methods used to test the hypotheses.

3.1 Research Questions and Hypotheses

The aim of this study is to investigate how the different parts of the jazz ensemble influence the dance style. We specifically aim to investigate its influence on the planning and execution of dance moves in the leaders' mind. Many previous studies have already assessed that musical content does have an effect on quantitative measures of dance like overall movement speed or globality of movement (see Ch. 2.4). In this thesis, we aim to extrapolate the underlying schemata of the dance and make quantitative comparisons based on them. The research questions are:

> RQ1: What part of the jazz ensemble facilitates artistry? Artistry translates to both the diversity of moves and the complexity of moves.

> RQ2: How commonly are moves of each family danced? Is there a relationship between ensemble part playing and an increased occurrence of a specific family?

To quantify these research questions, we used stimuli containing either a full ensemble, only the soloist, background, or only the rhythm section (for details see Ch. 3.2). In each clip of participants that we recorded, we labeled move types, move counts, and cue counts to reflect the diversity and complexity of moves.

There are different possible ways the previous research can be interpreted to make assumptions about the influence the different ensemble parts will have on the dancing. Through the lens of cognitive capacity alone, a clear beat might be the most crucial part for dance artistry, as it leaves the dancers with more cognitive capacity to plan movements. However, if we see dance as an artistic expression of music and consider that it has been shown to also reflect the melody, harmony, and timbre it is danced to (see Ch. 2.1), one might also assume that the dancers cannot be struck with inspiration if there is no melody to follow. Last but not least, there is the third option: dancers require both a melody and a beat to express themselves artistically. Based on this, the hypotheses are:

- H0: There are no effects of the jazz ensemble on the dance style.
- H1: Melody encourages artistic dancing.
- H2: Beat encourages artistic dancing.
- H3: Beat and melody must interact for artistic dancing.

Artistic dancing is quantified in four different ways:

- The number of cues performed
- The number of moves performed
- The diversity of moves performed (moves from many different families)
- The complexity of moves performed (amount of cues per move)

The diversity is quantified as normalized richness with the following formula:

$$Richness(R) = \frac{\sum uniquefamilies}{\sum moves} \tag{1}$$

Naturally, it is also possible that different aspects of artistic dancing are facilitated by different aspects of jazz music. For example, it makes sense to assume that more complex moves require a clear beat, as less intuitive movements or those that progress over a longer period require more cognitive effort, and the dancers need to have the capacity for that first [13]. Dancers might avoid complex moves when they need to focus on keeping the beat. Likewise, it could be assumed that the diversity of moves might depend more on creativity and therefore be more highly related to the melody.

3.2 The Stimuli

The stimuli were taken from the three Swing songs "A Smooth One" [36], "Doodlin" [37], and "Tatoe Pie" [38]. A 31-33s excerpt was taken out of each song. All sections of the jazz ensemble played in the part chosen. With lalal.ai [39] four conditions were extracted:

1. The full ensemble
2. Only the soloist
3. Only the background instrumentals
4. Only the rhythms section

The stimuli were selected based on lalal.ai's ability to separate them into sections. Since the program struggled to separate winds from each other (e.g., trumpet and saxophone), but could split them from other instrument types, the songs had to use instruments with sufficiently differing spectra for successful separation. Additionally, the audio software Melodyne [40] was used to remove any individual notes from silenced sections that were audible after the AI extraction.

Each stimulus consisted of two *8-counts* of the full ensemble before the separate sections played (see Fig. 1). This was meant as time for the dancers to find the beat and groove into the song. Those initial *8-counts* were discarded and not used in the analysis. Due to the different tempo of the songs, they contained different numbers of beats. They were all cropped at musically meaningful moments, e.g., after two phrases or a chorus. Song 1 "Doodlin" [37] had 48 beats after the initial two *8-counts* of the full ensemble, while the other two songs ("A Smooth One" [36] and "Tatoe Pie" [38]) had 64 beats. Since "Tatoe Pie" and "A Smooth One" both consist of two phrases each, the final analysis is done on Cues per phrase, with "Doodlin" being counted as 1.5 phrases and the other two being counted as two phrases. One phrase consists of 32 beats.

Our initial assumptions based on phenomenological experience (of the researcher) were that the beat is most clear in

Full Ensemble > Rhythm > Background > Soloist,

and that there is the most melody in

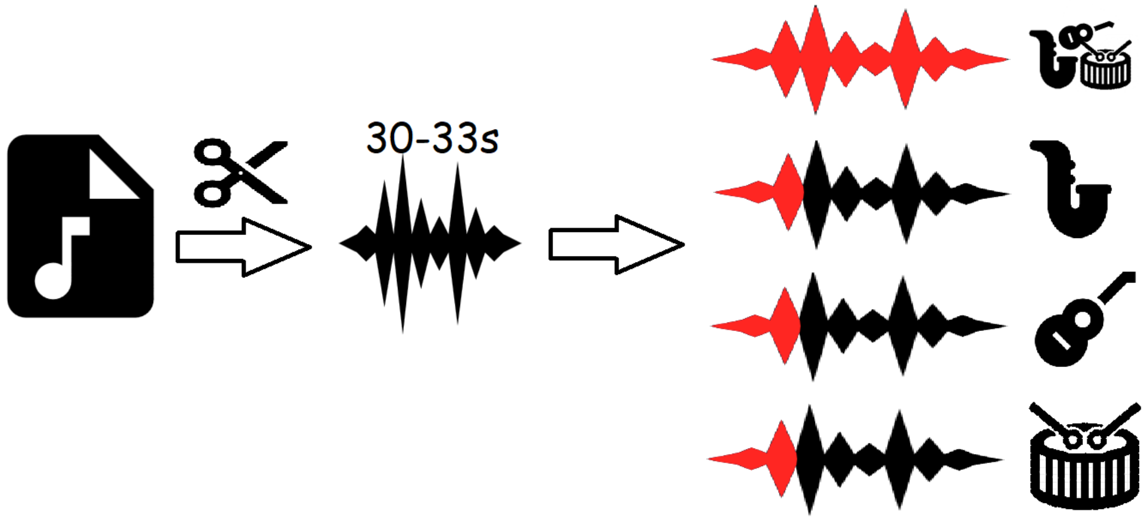


Figure 1: A graphic explaining the extraction process of the twelve stimuli. Three songs were cropped to 30-33s and then four stimuli were extracted by splitting them into different sections with two 8-counts of the full ensemble playing before each.

Full Ensemble > Soloist > Background > Rhythm

It is important to note that all stimuli were jazz pieces. Therefore, both the background and the soloist might use syncopation, which reduces beat clarity.

To test these assumptions, audio features were extracted from all stimuli using the PADMEA software (Perceptual Audio Dimension Modelling and Extraction Application), developed in the musicology department in Vienna, based on [41, 42]. Linear Mixed Models were run for features where the Song was found to have a random effect, whereas Linear (simple) Models were used for features where the Song showed no to minimal effect (Intraclass-Correlation (ICC) < 0.1).

For the beat clarity, the features *rhythmic irregularity*, *tempo variation*, *beats loudness*, *danceability*, *percussion harm(onic) ratio*, *dynamic complexity*, and *tempogram stability* were tested. For the quantitative measure of how melodic a stimulus was, the features *chroma entropy harmonic*, *HPCP (Harmonic Pitch Profile) crest*, and *HPCP entropy* were investigated.

The best indicator for beat clarity was found to be *Tempo variation* with a significant model fit in a Linear Mixed Model ($p < 0.01$) with significant differences between the Full Ensemble-Background, Full Ensemble-Soloist, Background-Rhythm, and Rhythm-Soloist (see Table 1). The Full ensemble and the Rhythm section had low Tempo variation, while the Soloist and Background sections had high Tempo variation. Since the song's tempo was the same across all stimuli, high variation indicates lower beat clarity, as it suggests the beat tracker was unable to accurately detect the beats. The same pattern was found in *Rhythmic irregularity* (Linear Model) as well, even though the differences were non-significant after bonferroni-correction (see Fig. 2).

An additional indicator for beat clarity was *beat loudness*, which was significantly louder in the rhythms condition than in all other conditions (Full Ensemble: $p = 0.009, CI = -0.09626$ to -0.0373; Background: $p < 0.001, CI = -0.12368$ to -0.0647; Soloist: $p < 0.001, CI = -0.1281$ to -0.06910). However, its meaningfulness is limited, as the beats are not necessarily less clear simply because they do not stand out as strongly from the melody, as is the case in the Full Ensemble.

Comparison					**95% CI**				
Type	**vs**	**Type**	**Difference**	**SE**	**Lower**	**Upper**	**t**	**df**	**p**
Full ens.	-	Backgr.	-20.59	3.10	-28.169	-13.01	-6.64	6.00	0.003
Full ens.	-	Rhythm	5.11	3.10	-2.467	12.70	1.65	6.00	0.899
Full ens.	-	Solo.	-13.48	3.10	-21.060	-5.90	-4.35	6.00	0.029
Backgr.	-	Rhythm	25.70	3.10	18.120	33.28	8.30	6.00	<.001
Backgr.	-	Solo.	7.11	3.10	-0.473	14.69	2.29	6.00	0.369
Rhythm	-	Solo.	-18.59	3.10	-26.175	-11.01	-6.00	6.00	0.006

Table 1: The post-hoc tests for *tempo variation* with bonferroni-correction applied to the p-values [1, 2, 3, 4].

These results confirm that the Soloist and Background sections have a lower beat clarity than the Rhythm section. They show no significant differences between the Full Ensemble and the Rhythm section (in *Tempo variation* and *Rhythmic irregularity*). An unexpected result was the low beat clarity in the Background section. When investigating the stimuli again, it was noted that some background stimuli always started on the up-beat, while others had extended pauses. This might have confused the beat tracker, leading to an inflated value.

When looking at the melody, the *HPCP crest* (Linear Model) was found to be the best indicator. There were significant differences between Rhythm and all other conditions (Full Ensemble: $p = 0.001, CI = 4.682$ to 9.992; Background: $p = 0.018, CI = 2.172$ to 7.483; Soloist: $p = 0.008, CI = 2.877$ to 8.187). The difference between Full Ensemble and Background was non-significant ($p = 0.366$), and there were no significant differences between Full Ensemble and Soloist ($p = 0.933$). However, a trend towards less HPCP crest (melody) in the Background and Soloist conditions than the Full Ensemble condition can be seen in the plot (see Fig. 3). Because this analysis exhaustively covers the stimuli used in the study (e.g., we are not sampling from a larger stimulus population), significance values are of less importance and it can be argued that this trend speaks for our assumption, however this limits the conclusions that can be drawn based on melody.

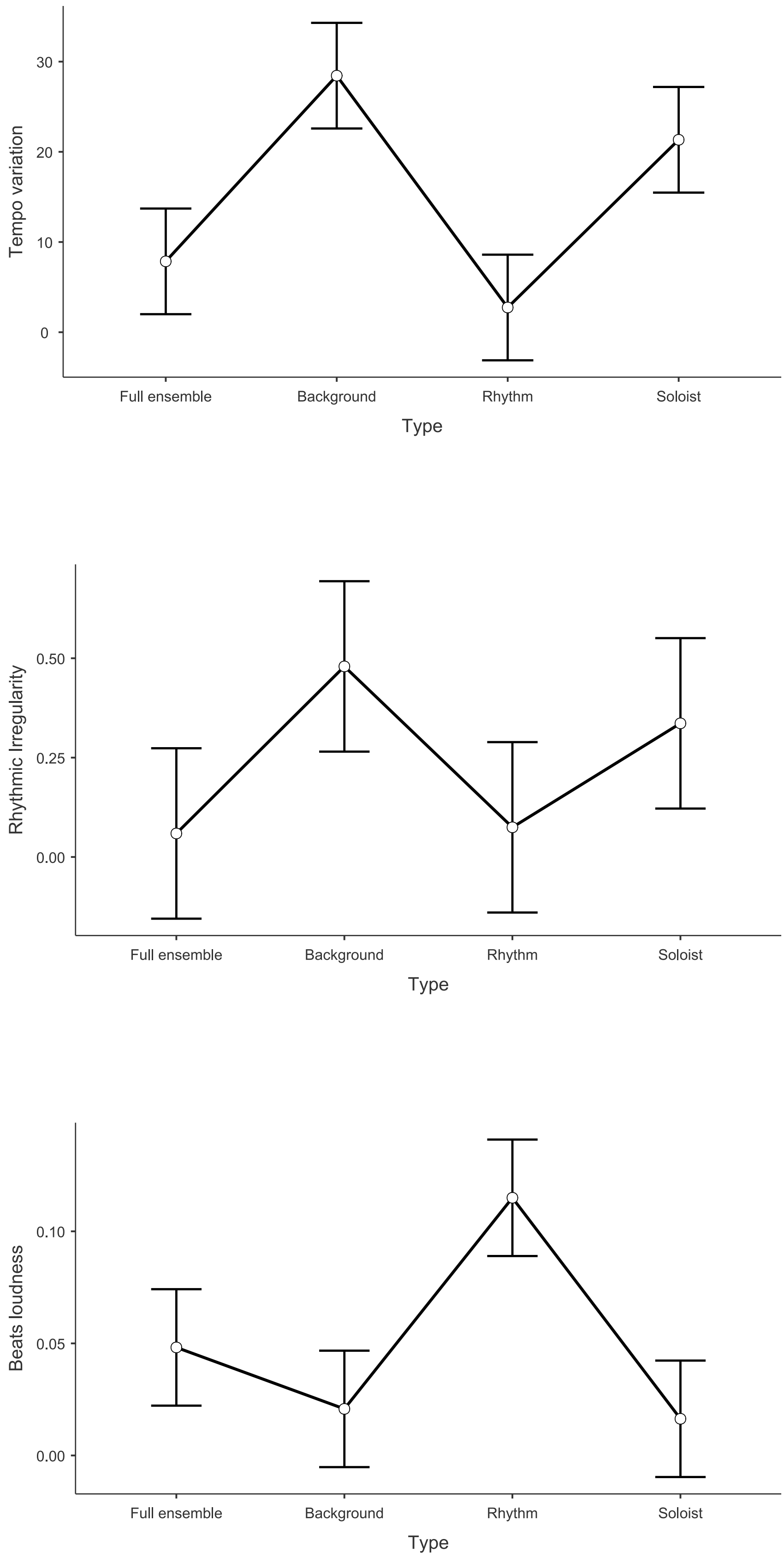


Figure 2: *Tempo variation*, *Rhythmic irregularity* and *Beats loudness* means and CI for each Condition[1, 2, 3, 4].

As a supporting parameter, *HPCP entropy* was investigated and led to similar results. There were significant differences between background and rhythm ($p = 0.001, CI = -0.882$ to -0.413) and soloist and rhythm ($p < 0.001, CI = -1.137$ to -0.668). The Full ensemble was found to have high entropy, due to the bias towards higher entropy in the Full ensemble (more instruments mean more entropy). We cannot count this as meaningful information about how melodic it is. The *HPCP crest* is the more reliable feature for this condition (see Fig. 3).

The updated assumptions after the assumption check would therefore be:

The beat is most clear in

> Full Ensemble/Rhythm > Background/Soloist,

and there is the most melody in

> Full Ensemble > Soloist/Background > Rhythm

3.3 Study Population

The study was performed with 8 participants from Viennese Jazz dance schools. They were recruited directly by Sumner Williams. All participants had to have at least 3 years of experience in Lindy Hop, be competitive dancers, or be teachers in their Lindy Hop scene. They also had to be able to dance both the lead and the follow roles.

The study was conducted in 2 sessions, with 4 participants per session. The participants were relatively familiar with one another due to their involvement in the same social scene (Vienna Lindy Hop socials). However, training, competition, or teaching partners were intentionally separated and did not dance together in the same sessions to eliminate excessive familiarity bias.

3.4 Experimental Setup

Within each session were two participants who had primarily learned to dance as leaders (primary leaders) and two who had primarily learned to dance as followers (primary followers). They all danced with everyone else in that session, both as leaders and followers, resulting in 12 dyads. All dyads (pairs of leader and follower) had to dance to all 12 stimuli.

This means that each individual person danced to all of them six times. The order of the stimuli was randomized within each dyad so that the upcoming stimulus could not be predicted. However, it is likely that there were some order effects due to the dancers growing increasingly familiar with the stimuli.

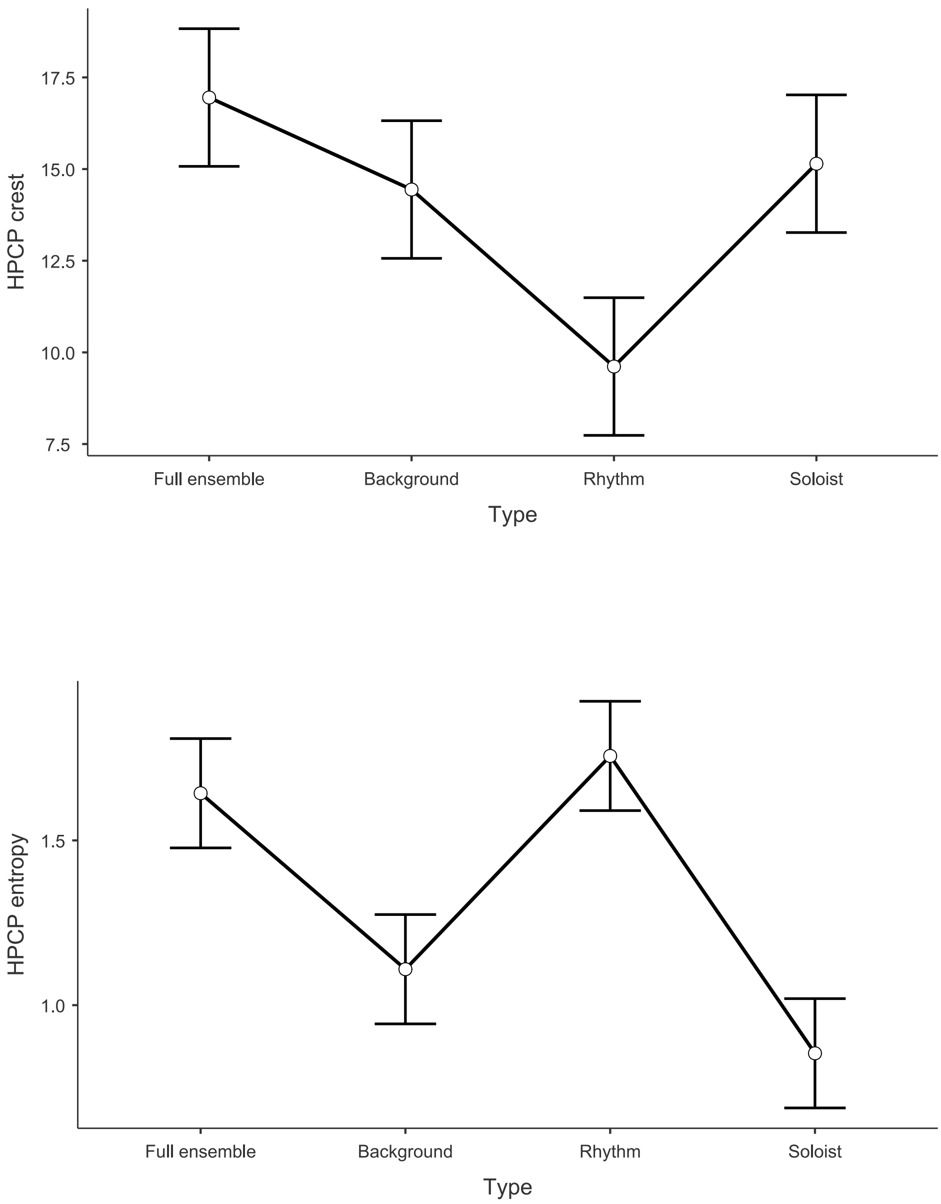


Figure 3: *HPCP crest* and *HPCP entropy* plotted against Condition [1, 2, 3, 4].

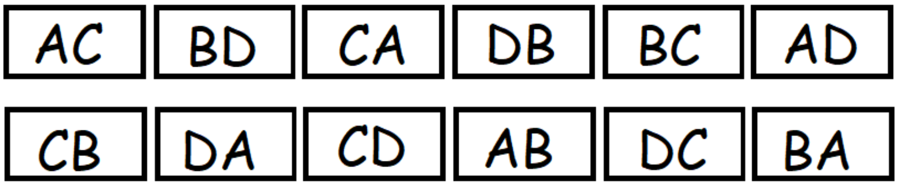


Figure 4: If A and B were primary leaders and C and D primary followers, this was a sample order of how they danced in the setup.

The participants were instructed to dance naturally within the demarcated area. Each dyad took about 20-30 minutes dancing. The order of dyads was chosen to minimize the number of times participants danced two times in a row, so that they usually had 20 minutes of rest in between. If they danced twice in a row, they were offered some resting time of at least five minutes (see Fig. 4 for reference).

The order of dyads and songs, as well as retakes, was recorded. Retakes were done when a marker fell off during the recording. As the dancers grew familiar with the music across multiple dyads, retakes were not as severe either way.

In the first session, 15 clips required retakes, and three clips required two retakes. This means 10.4% of the clips had retakes. In the second session, only nine clips (6.25%) required retakes due to the identification of markers at risk and improved procedure. In the second session, none of the retakes happened within the first two dyads, meaning that at the point of retakes, the stimulus had already been heard multiple times.

3.5 MotionCapture Setup and Prepocessing

All participants were wearing MotionCapture suits with 28 markers. They had four head markers, four hip markers, markers on each shoulder and elbow, two markers per wrist, one on the middle finger, one on the sternum and spine each, as well as on the side of each knee and the ankle. All participants wore normal shoes to keep the dancing natural. On those shoes, one marker was attached to the heel and one above the big toe (see Fig. 5).

The markers were tracked by 14 infrared cameras surrounding the dance area and one video camera without sound. An additional recording was made from the experimenter's position, using a phone to record both audio and sound. At the beginning of each recording, the experimenter produced a visual and an auditory cue by hitting two blocks and two sides of a filmboard with attached markers, respectively.

The data was recorded at 100 frames per second (fps) and cleaned with Qualysis Track Manager (QTM). Gaps were filled with either a polynomial function (below ten frames) or

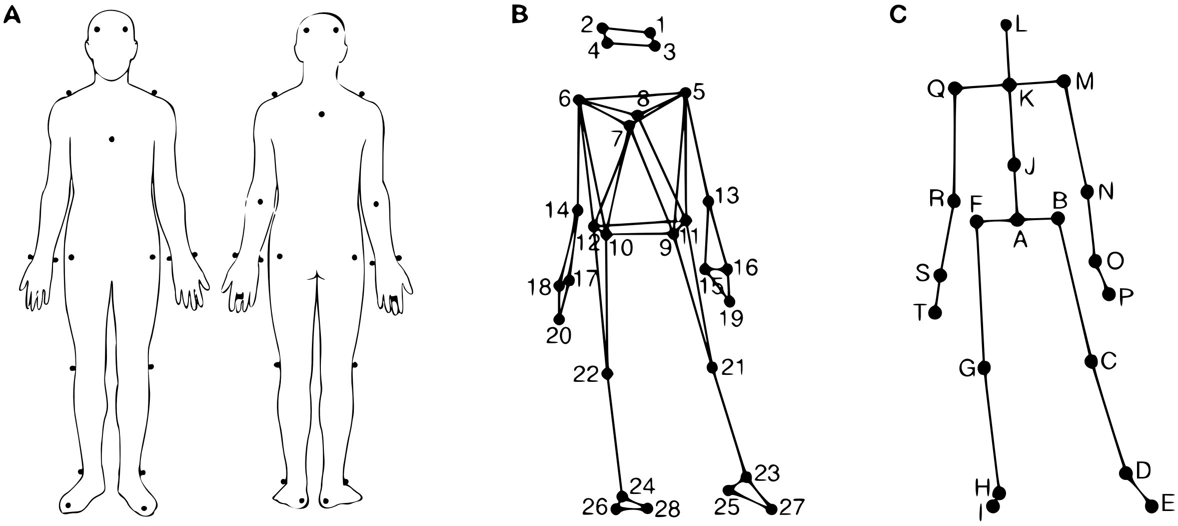


Figure 5: The positions of the markers A) on the participants suits and B) their final representation in QTM [5]

a relational function (above ten frames). Spikes were only smoothed if they displayed signs of unnatural marker jumping, and with a threshold of $150m/s^2$.

Based on the cue visible in the data and the start of music in the audio, the start of the downbeat after the initial two 8-counts was determined in the files. They were cropped to the duration of the pure stimulus to ensure that only the moves performed within the actual stimulus were considered. It also ensured that the partners were already properly synchronized at the point of observation.

The dance area was approximately 4 x 5m, providing the dancers with sufficient space for natural dancing, as the space used in Lindy Hop socials is often even smaller.

3.6 A Grammar for Lindy Hop

For data labeling, a reference was required. In this work, this was a definition of the inherent grammar of the dance style. As part of a previous project, the author of this work took two beginner and one intermediate Lindy Hop courses to learn the dance's basic mechanics.

We then focused on the files in which each dancer danced their primary role. In each session, the primary leaders were A and B, while the primary followers were C and D. Therefore, the files analyzed were of the pairings AC, AD, BC, and BD.

The extraction was limited to those pairings to reduce the amount of variations produced by miscommunication and errors, and to ensure that the performed dances adhered to a certain level of structure that comes with experience. This way, we excluded files from dancers who were experienced in the dance in general but less experienced in their respective roles. Another aspect was to avoid the primary leader - primary leader and primary follower - primary follower pairings in the extraction, since they may be subject to different communication dynamics than regular primary leader-primary follower pairings.

The author manually investigated all extracted files and saved the first occurrence of each potential move, whether it was known to them or not. The resulting excerpts were then identified with the colleague Sumner Williams, who has 11 years of experience as a leader himself and 7 1/2 years as a Lindy Hop instructor. Based on the observed moves and the positions they progressed through, a schema was developed that attempts to reflect the grammar inherent to Lindy Hop.

Based on the initial extraction, we identified 15 positions and 11 move families (see Fig. 6). Three additional positions were added to a second version later for the labeling process. Two of them were necessary due to the separation of the Frankie's 6 procedure into individual moves, one was likely caused by the dancers dancing in the opposite of their primary roles (Reverse Under the Wing), as it is the same as the position Under the Wing with swapped roles.

The grammar in both versions was published on websites [43, 44]. All mentioned positions can be viewed there as 3D images.

The positions were classified into Open, Closed, and Unspecified. Closed positions were further distinguished into Closed Framed (the leader holds the follower in a frame), with its subcategory Closed Facing (the leader is facing the follower), and Closed Side-by-Side. The categorizations occurred based on the following functional aspects:

- In **Open** positions, the leader and follower are only connected by one hand, and these positions have a high improvisational freedom for the follower, while **Closed** positions are marked by a low improvisational freedom for the follower and more connection between follower and leader.

- In **Closed Framed** positions, moves from the Tuck family are very commonly used to go back into Open positions, while such moves are more uncommon in the Open positions and almost impossible in Side-by-Side positions.

- Only in **Closed Facing** positions are moves from the Swing Out Family possible. Only a very specific kind of Swing Out can be performed by a move outside this category: In the position Cross Both Hands, the Cross Hand Double Loop can be performed, which makes this position a grey area between Closed Facing and Unspecified. However, there were no instances of this move during the natural dancing in the experiment, so it was placed in the Unspecified category.

Figure 6: The positions in the multidimensional schema of how they relate to each other

- In **Closed Side-by-Side** positions, the connection between Follower and Leader is very tight. The range of moves possible is very limited.

The moves were categorized into the eleven families based on common mechanics. The families **Free Spin**, **Underarm Turn**, **Swing Out Family**, and **Tuck Family** all consist of moves that contain the name-giving mechanic. **Solo Jazz Moves** is a collection of footwork variations that can be performed individually, but also as an additional variation on already performed moves. These moves are usually synchronized by visual cues and do not contain physical lead cues. **Walks** are all moves in which the dancers walk in a given direction. **Bring Ins** and **Spare Tires** are individual moves that were categorized as families because multiple variations exist. **Procedures** contains not individual moves, but entire procedures (consisting of multiple moves). In the case of the **Circle** (and **Roll In Circle**), these were still labeled as one move in the data as they are tought in one chunk, however **Frankies 6** and **Under the Wing** were labeled separated into their individual moves as there was much more variation in the actual moves executed and the data did not support the idea of it being chunked together in executive planning. The families **Other looping moves** and **Others** are both families for all moves that could not be further assigned to any other family.

The grammar distinguishes between looping (usually **Solo Jazz Moves**, **Walks**, and **Other looping moves**) and non-looping moves (the rest, most of the time). A looping move is defined as a move that does not leave position. For example, when the dancers walk forward together, the way they hold onto each other does not change. They could continue walking indefinitely, since looping moves always return to their starting point.

3.7 The Labeling of the Data

Based on the grammar developed in Ch. 3.6, each MotionCapture clip was labeled by the author. Additionally, one clip per stimulus was selected by randomly sampling twelve out of the 24 dyads without replacement. These twelve clips were then randomly assigned letters and given to an independent external rater teaching at a Lindy Hop dance school in Vienna, separate from the one the author visited.

The external labeler is 46, female, has been dancing for 15 years, and taught as a follower for 14 years. The external rater was given no information about the stimuli, and the MotionCapture recordings lacked audio to ensure a blind approach. They were informed that the letters were randomly assigned and held no meaning. In Ch. 4.2 the results of the interrater-reliability are calculated to test if there was systematic error in the labeling process.

This thesis aims to investigate the cognitive planning of dance and its execution. Therefore, it is important to distinguish between analog and digital lead cues. The idea of analog vs. digital lead cues can be explained like this: imagine you were dancing with your partner, and the cue did *not* occur. How would you react? If we, for example, were doing the base steps of Lindy Hop, and the cue to do the rock-step did not occur, most people would still do it automatically, because they know the base-step pattern of Lindy Hop, and it would require an active different movement of the leader to make the followers stop it. This means that the cue to move back on the rock step is an analogue cue. It is mostly automatic, on both dancers' sides, and therefore unlikely to require any additional processing.

A digital lead cue, however, is any cue that provides new information to the follower, namely, by initiating a move that deviates from the default step pattern. Sometimes, the digital cues might not accurately reflect the complexity of a move. Since we use them as a measure of complexity, there are some rules that might restrict the number of cues counted for specific looping moves or caps applied after a certain number of repetitions (e.g., for Tuck Turns, cues after the 6th Tuck are no longer counted, as it is seen as becoming automatic). To avoid bias through those rules, the actual digital lead cues were recorded in these cases to test if it biases the results.

One file was created per clip. Per move, one line was added into the file containing the columns "Start Frame", "Position", "Closed/Open", "Digital Lead Cues", "Moves", "Looping?", "Type of move", "Type simplified", "Actual Digital Lead Cues", "Frame Start Move".

The following is a rough transcription of the description given to the external rater on how to fill those columns:

- Position: The position in which the dancers are before the move begins. The different positions can be viewed on the website. During labelling, there is the additional position "No connection" not present on the website. This label shall be used if the dancers dance without touching at all for an extended time and falls under the category Unspecified.

  Attention: If the first move is cut off because it begins before the cropped area, the position at which the move initially started shall be entered (raters were able to move the time-bounds for that purpose).

- Closed/Open: Whether the position belongs in the Open, Closed, or Unspecified categories as explained in Ch. 3.6.

- Digital Lead Cues: The digital lead cues counted in the move. The raters could refer to the explanation (similar to the one provided above) and the overview [45] for guidance.

- Moves: The move count of this move. This is usually just 1. There are a few exceptions mentioned in the reference sheet [45] that had to be labeled but are not considered a (separate) move and therefore have a move count of 0, so they can be filtered out in the statistical analysis.

- Looping?: Specifies whether a move is looping (dancers do not leave their position, e.g., sugar pushes) or non-looping (e.g., underarm turn). In looping moves, often only the digital cues for the first repetition of the move are counted (usually 1), as the pattern repeats, and the additional cues do not reflect complexity. The specific rules are available in the reference sheet [45].

- Type of move and Type simplified: The move that the dancers are performing. In "Type of Move", more precise statements are possible (e.g., variation, twisting, etc.) while in "Type simplified" a clear category should be assigned. All possible designations are listed in the reference sheet [45]. The data in "Type simplified" is later automatically processed to add the move family to the files.

- Actual DLCs (ADLCs): Optional field for the actual digital lead cues performed, if the digital lead cues should be counted to a limited extent in the complexity measurement.

- Frame Start Move: Frame of the first cue of the move. Often the rock-step, but it might also happen before it. This is not used further in the analysis due to the limited accuracy of human judgment in a 100 fps recording, but it serves as a guide to investigate possible discrepancies between labels.

  Exception: If a move starts before the cropped area, the start frame is also the first frame of the cropped area.

If moves are cut off, because they start before the start boundary or end after the end boundary, the following rules apply:

- If at least one cue of the move is performed within the boundaries, the move and all cues within the boundaries are counted.

- If none of the cues are performed within the boundaries, the Digital Lead Cues and Moves fields are left blank, and only the type of move is entered in the corresponding fields (if detectable).

- If the move is not clearly recognisable because it was not completely recorded, the "Type simplified" field is left blank, but the cues are still counted, if available, and it counts as one move (if cue count > 0).

The external rater was provided with a tutorial in German (see Appendix A), a YouTube explanation with a sample file not contained in the clips they had to label, a link to version 2 of the website, as well as access to the reference sheet of the count rules for standard versions of the moves.

For the creation of this reference sheet, summaries were generated for each move after complete labelling by the author. Irregularities in the counted cues were investigated and corrected if found to be illegitimate. Possible reasons for such occurrences include tiredness, not paying attention for a minute, or developing a different understanding of the dynamics over the course of the labelling. The general rule for the default form of each move was derived from the summaries after cleanup, and moves with ambiguous labelling (the rule changed over time) were discussed with Sumner Williams to determine the most accurate count method and corrected, respectively, to ensure maximum intrarater reliability. The resulting general rules were then provided to the external rater.

3.8 Statistical Analysis

Due to the special research method used in this thesis, we need to test whether manual labels applied by dancers are a reliable measure for analyzing the effects of the stimuli on dance. To do that, we use Krippendorff's alpha to calculate the interrater-reliability between two independent raters for a portion of the clips. Reliability is calculated per clip for Cue count, Move count, and Move family at the label level. Additionally, averages across all clips are calculated, and a general alpha for the Cue count sum and Move count sum of each clip is determined.

We also need to test whether dyads with the same leader show a similar dance style or whether the follower has too much influence on the dance style. For that, we use a Monte Carlo simulation with 10 000 repetitions to test if dyads with the same leader on average have a lower variance in Cue sum, Move sum, and Richness than dyads with differing leaders. The simulation is run by randomly selecting one stimulus from the 12 stimuli and three dyads, and calculating the variance between them. This indicates whether the dyad or the leader should be used as a clustering method in the Linear Mixed Models (LMMs).

Afterwards, we explore the second research question, using box plots to display the use of Moves across the different Move families, split by Condition and Overall, and discuss noticeable differences. The data are then tested for normality, and the means and standard deviations are reported.

To select the best linear mixed models for each variable, models with the audio features (Tempo variation and HPCP crest) and with Condition and Song as factors are compared (using AIC, BIC, and Log Likelihood). All models are run with the GamlJ module for Jamovi [1, 3]. Nominal variables (Song and Condition) are entered as factors. Continuous variables (Tempo variation and HPCP crest) as covariates. Dyad is a clustering variable in all models.

Estimations are run using Restricted Maximum Likelihood, and confidence intervals are based on fixed parameters. The formula for all linear models is:

$$\{variable\} \sim 1 + \{\ldots predictor(s)\} + (1|Dyad) \tag{2}$$

In some cases, the interaction of predictors was also tested. Post-hoc tests are performed via pair-wise comparison with Bonferroni correction for p-values. The best models for each variable are selected, and the normality of the residuals is tested using the Shapiro-Wilk test. The best-fitting models are then used to test for significant differences in Cues per phrase, ADLCs per phrase, Moves per phrase, and Richness between Conditions or audio features.

Chapter 4

Results

This section describes the statistical analysis of the data and the results. In Chapter 4.1 the demographics of the participants are outlined. The participants were young adults with a high level of education who were familiar with German-speaking environments and, therefore, have no barriers to participating in social dancing in Austria.

In Chapter 4.2 the results of the interrater-reliability analysis are reported, and the reliability of different variables is assessed based on these values. The reliability between raters was found to vary widely across dyads, with moderate to good agreement in the majority of dyads. The sum of Cues labeled on each clip was found to be reliable, whereas the sum of Moves was found to be unreliable. The labeling procedure of the external rater can be found in Chapter 3.7.

Chapter 4.3 investigates whether there are any differences in the variances for the same stimulus within a leader or between different leaders and uses it to inform the linear models that are discussed in Chapter 4.6. No differences in variance were found for most parameters. Small differences in variance were found for the Moves per phrase, as discussed in detail in this chapter.

Chapter 4.4 outlines the distribution of different move families overall and across conditions, and notes any noticeable differences. The most frequent moves performed were Underarm Turns. Swing Outs were danced twice as frequently in the Rhythm Condition as in the Solo and Background Conditions.

Chapter 4.5 provides general information on the collected data, including the mean per condition, as well as checks for normality. The Full Ensemble Condition was danced with the most cues, *ADLCs*, and Moves per phrase, but also had the lowest Richness. The Song Doodlin [37] had the highest mean on all parameters. Distributions were found to be normal for Cues per phrase for all conditions and for Richness for all but the Background Condition. Moves per phrase and ADLCs were found to have non-normal distributions in some Conditions.

Chapter 4.6 provides an in-depth description of the tested predictors, the selection of the best models, and an analysis of the results. The models selected were:

- Predictor Condition for Cues per phrase (and ADLCs per phrase as control)
- Predictor Tempo variation for Cues per phrase (and ADLCs per phrase as control)
- Predictors Condition and Song for Moves per phrase
- Predictor Tempo variation for Richness

Residuals of error only deviated significantly from normality for the LMMs predicting the ADLCs. However, Moves per phrase also had low p-values (non-significant).

The Condition was found to influence the number of Cues given by about 10%, with the Full Ensemble and Rhythm Conditions being danced with more Cues. Cues per phrase decreased with a higher tempo variation. The analysis with ADLCs confirmed these findings. Moves per phrase were shown to deviate much more based on Song than on Condition, however the difference between Full Ensemble and Background was still significant. There was a significant positive correlation between Tempo variation and Richness.

4.1 Demographics

There was an equal distribution of genders ($M = 50\%, F = 50\%$), with all men leading as a primary role and all women following as their primary role. One of the female participants had almost as much leading experience as her following experience.

The participants were aged 22 to 37 years ($M = 30.4, SD = 4.75$). All participants were from the academic sector, with either a bachelor's degree, a master's degree, or an equivalent matura ($Bch. = 3; Mst. = 2; Mat. = 3$). Seven out of eight participants were Caucasian, with one participant of Asian descent. All participants were from a German-speaking country ($Austria = 7; Germany = 1$).

Their total years of Lindy Hop dancing experience ranged from 1.5 to 13 ($M = 6.50; SD = 4.33$), with equal experience in their primary role. Their experience in the secondary role ranged from none to 12 years ($M = 3.06; SD = 3.79$).

Overall, the demographic consists of young adults with a high level of education who are familiar with German-speaking environments and therefore have no barriers to participating in social dancing in Austria. Due to the sample being predominantly Caucasian and all participants being of German or Austrian nationality, it can be expected that the analysis will reflect the European tradition of Lindy Hop, which may deviate significantly from the original version created by African-American people in the 1930s [33]. This is especially the case since both the author and their colleague are Caucasian as well.

This means we naturally viewed the dance through a European lens because we lacked experience with African-American culture. However, this can easily be accepted, since dance is always evolving, and the revivalist movement of Lindy Hop in the 1990s was predominantly marked by Lindy Hop as a dance for the white middle class, rather than the black working class in which it had been practiced 60 years earlier [34]. However, we need to be aware that the cultural background influences the expression of dance and that the extracted MotionCapture data is not representative of the dance as it was and is practiced by African-American people.

Another point to note is the heavy gender bias in the primary roles of the dancers, with men as primary leaders and women as primary followers. While there is an effort in the Viennese community to break down those gender stereotypes [46], Lindy Hop was traditionally a dance heavily based on those stereotypes [34], and this is still evident in the primary role many dancers choose.

4.2 Interrater-Reliability

Based on the labels provided by the author and the external rater (see Chapter 3.7), the interrater-reliability was calculated using Krippendorff's alpha for each individual MotionCapture clip as well as for the sums of Cues and Moves labelled in each clip. The results are shown in Table 2. The labels were entered into a table for each clip, with each row representing one label.

On each clip, the timestamps were aligned so that the labels applied to the same moment in time. This was necessary because sometimes one rater interpreted movements as a single move, while the other interpreted them as two moves, leading to different amounts of labels. However, this should not influence the reliability of any further labels. Krippendorff's alpha compares the agreement between raters on each label to calculate the final value, so empty labels had to be added if a movement was described by a differing number of labels. Empty labels that were produced through that adaptation defaulted to having 0 Cues, 0 Moves, and the family "Others" in the Move family.

The alpha values on the clips were calculated for assigned Move families, Cue count, and Move count. The average alpha value across all clips was calculated by averaging all alphas. It is important to note that this thesis does not analyze individual labels. The overall sum per clip is of much greater importance. However, these aspects were still considered, as they describe the schema's overall reliability.

The results showed that while most clips had moderate to sometimes even perfect agreement at the individual label level, some parameters occasionally showed low agreement (Move family and Move count on H, all parameters on A, Move count on B and K, Cue count on D). This indicates that there were deviations in the dyads' level of improvisational expression, as "textbook" moves were more easily labelled identically than those with a high level of improvisation.

An important thing to note is that systematic disagreement was found in the move labeling in clip A. This led us to suspect a training effect, with the external rater still learning how to apply the schema during the first clip. On request, the external rater reported that they did label the clips in the order of the letters (A being the first) and that they still felt unpracticed on the first clip, planning to get back to it and redo it at a later time. However, when they looked at it later, they considered the labels fitting well enough.

Due to this likely training effect, the results show the averages both with and without the exclusion of A. Excluding A, the Cue count and Move count per label showed moderate agreement, while the Move family showed almost moderate agreement. Since this thesis investigates the overall sums per clip, the interrater-reliability was also calculated for that.

Sums were calculated for Cues and Moves on each clip (see Table 2), and Krippendorff's alpha was then calculated. The Cue sum showed moderate-to-good (A excluded) to good (A included) agreement, while the Move sum showed poor agreement. This indicates that Cues were stable, while Moves depend heavily on the labeler's perspective on how they chunk the dance moves together. Results on Moves, therefore, need to be considered tentatively as they might only reflect the author's personal bias.

4.3 Variance Within and Between Leaders

To inform the quantitative analysis in Chapter 4.6, we wanted to test whether the same leader (with different followers) had lower variance in Cues, Moves, and Richness for the same stimulus than differing leaders. We wanted to know this to decide whether data sets from the same leader can be considered within-person datasets or whether the follower has too much influence. To test this, we performed a Monte Carlo simulation with 10 000 repetitions.

In each repetition, three out of the 24 dyads were randomly picked, and a random stimulus out of the twelve presented was chosen. The labeled files of the selected dyads dancing to the selected stimulus were loaded, and the variances between those clips were calculated for Cues per phrase, Moves per phrase, and Richness. The results can be seen in Fig. 7 and 8.

Clip	Letter	Move family α	Cue count α	Move count α	Cue Sum E	Cue Sum A	Move Sum E	Move Sum A
BC 1 Jan	I	~0.738*	~0.991**	1***	8	9	6	6
DA 2 Nov	H	0.280	0.	0.250	7	5	9	5
BA 3 Nov	A	~0.346	~0.417	~(-0.211)’	15	15	8	11
DB 4 Nov	C	~0.679*	~0.761*	0.800**	10	13	7	8
BC 5 Nov	B	~0.609	~0.775*	~0.274	15	16	9	10
AD 6 Nov	L	~0.795*	~0.832**	~0.632	15	13	10	9
AB 7 Nov	J	~0.713*	1***	1***	14	14	10	10
BA 8 Jan	G	~0.691*	~0.687*	~0.632	11	12	9	10
CA 9 Jan	E	~0.703*	~0.984**	1***	15	14	10	10
DA 10 Jan	K	~0.813**	~0.8**	~0.432	12	10	12	10
AC 11 Nov	D	~0.667^	~0.440	~0.627	8	11	8	9
CA 12 Nov	F	~0.652	~0.839**	~0.762*	15	13	11	10
Averages including A		~0.640	~0.765*	~0.6		~0.8**		~0.467
Averages excluding A		~0.667^	~0.796*	~0.673*		~0.789*		~0.664

Table 2: The results of the interrater-reliability calculated for each clip as well as averages.
Cue/Move Sum E = Cue/Move Sum external rater

Cue/Move Sum A = Cue/Move Sum Author

For the sums an overall α across all files was calculated and can be seen in the average fields. The individual clips contain the sums, not the α values (as alphas on one value are pointless).

The α values provided for individual clips describe the agreement of the raters across all labels in that clip. Timestamps were aligned so that rows labeled the same move even if previous moves were differing between raters.

*** = perfect agreement

** = satisfactory level of agreement

* = moderate agreement

^ = close enough to general cut-off to still be considered moderate agreement in our context

’ = systematic disagreement The cut-offs are taken from [47] and the table was created with [6].

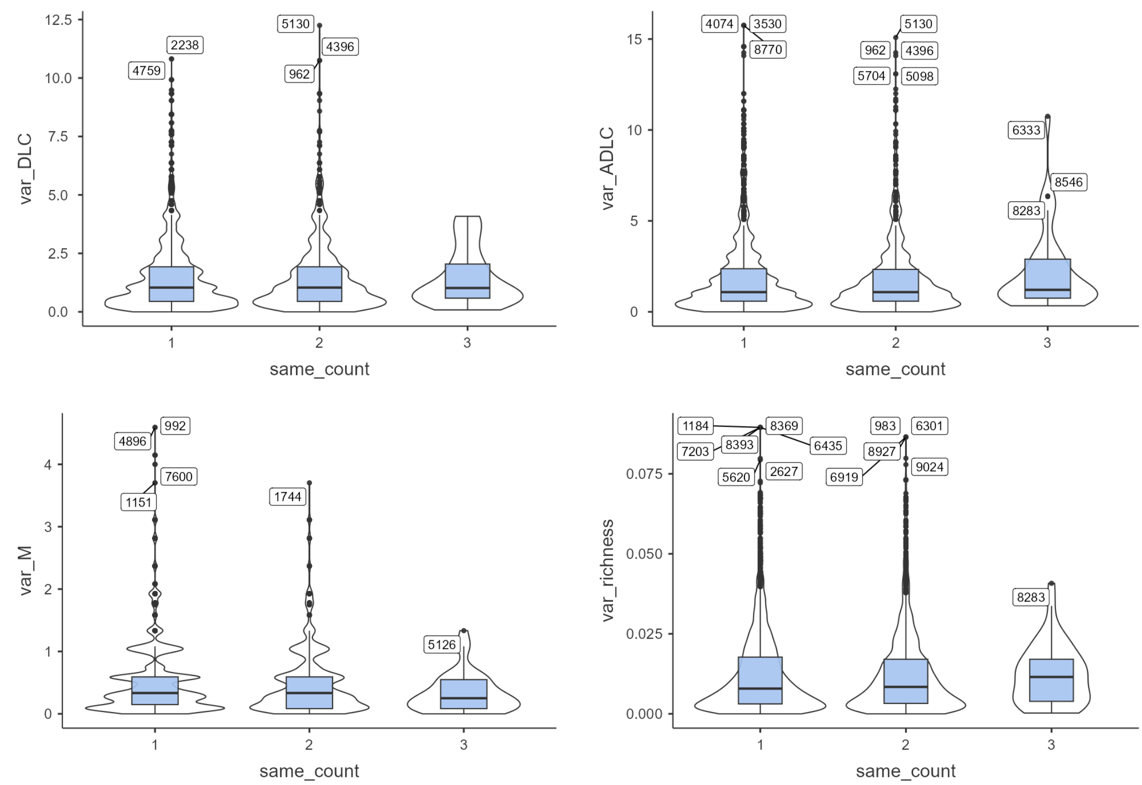


Figure 7: The box- and violinplots for the variances before exclusion of outliers. $same\_count$ is the amount of clips from the same leader, var_DLC = Variance *(Digital Lead) Cues*, var_ADLCs = Variance *Actual DLCs*, var_M = Variance of Moves, $var\_richness$ = Variance of richness [1, 2]

Due to the stochastic nature of random selection, drawing the same leader three times is much less likely than drawing three different leaders (or drawing the same leader twice). Therefore, the resulting groups were N=7541 for three different leaders, N=2421 for two times the same leader, and N=38 for three times the same leader.

There were no significant differences on most parameters (see Table 3). The cut-offs for the outliers were a variance above 10.0 for Cues and Actual Digital Leads Cues (ADLCs), a variance above 2.0 for Moves, and a variance above 0.07 for richness. While the Moves per phrase were significant both before and after excluding the most extreme outliers, after exclusion, it is minimal (1: $CI = 0.452 - 0472$, 2: $CI = 0.441 - 0.476$, 3: $CI = 0.227 - 0.440$), and the difference disappears when further outliers are excluded. However, it is interesting to note that the leader variance does appear lower for Moves for the same leader, as the move sum distribution is also not normally distributed (see Chapter 4.5). This might indicate that clustering this variable by leader might contribute to better results in the final models.

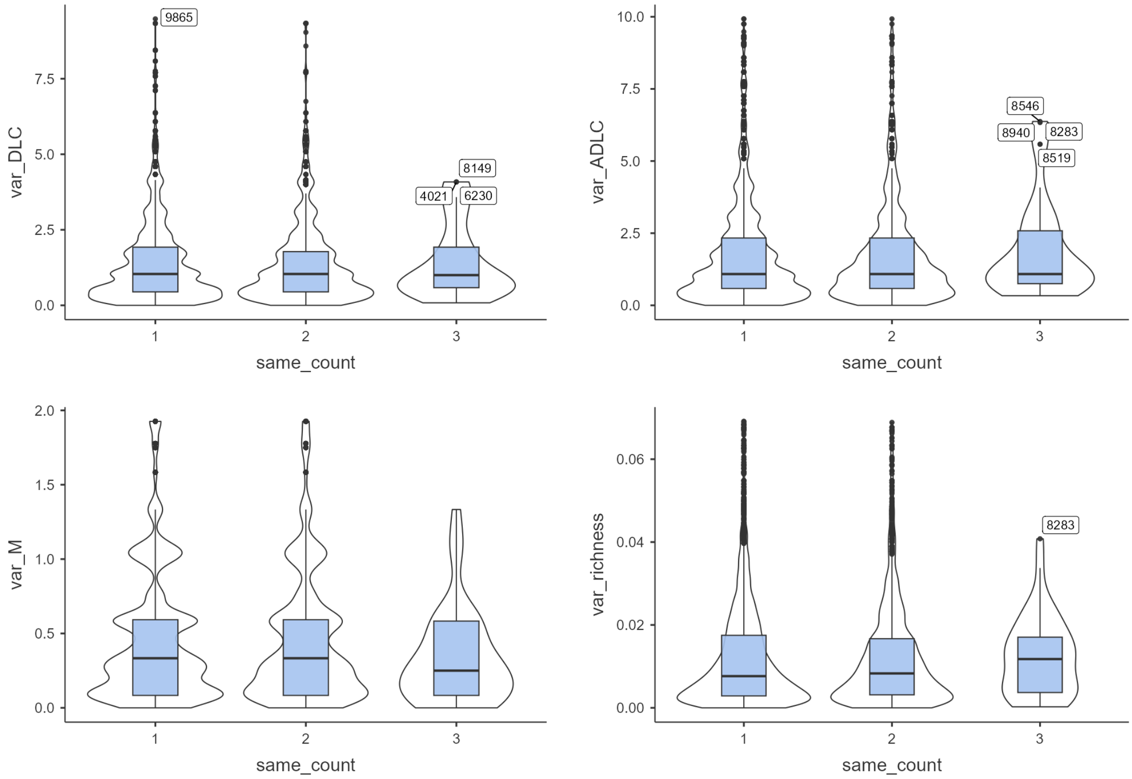


Figure 8: The box- and violinplots for the variances after exclusion of outliers. $same\_count$ is the amount of clips from the same leader, var_DLC = Variance *(Digital Lead) Cues*, var_ADLCs = Variance *Actual DLCs*, var_M = Variance of Moves, $var\_richness$ = Variance of richness [1, 2]

Variance of	**Amount of clips from same leader**	**Estimate**	**Lower bound**	**Upper bound**
	1	1.44	1.41	1.48
Cues	2	1.44	1.38	1.50
	3	1.46	1.06	1.86
	1	1.82	1.78	1.86
Actual DLCs	2	1.83	1.75	1.91
	3	2.18	1.45	2.9
	1	0.508	0.496	0.521
Moves	2	0.492	0.471	0.512
	3	0.327	0.222	0.431
	1	0.0122	0.0119	0.0125
Richness	2	0.0127	0.0122	0.0132
	3	0.0122	0.00910	0.0153

Table 3: The estimated averages (Estimate) and the 95% confidence intervals (Lower bound and Upper bound) for one, two, or three clips from the same leader for all four parameters without outlier exclusion. Generated with [6]

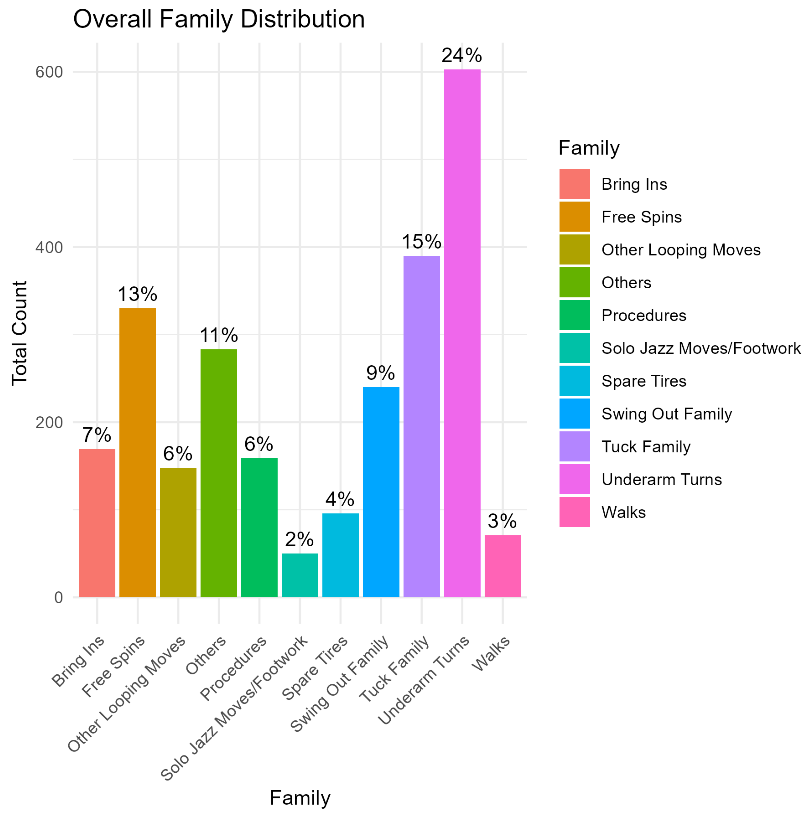


Figure 9: The distribution of move families across all clips.

4.4 The Distributions of Move Families

Based on the Move families labeled by the author, bar plots of the move distribution were produced for each Song, each Condition, and overall to map out the general frequency of moves. It showed that about a quarter (24%) of all moves performed were Underarm Turns, followed by the Tuck Turn family (15%), Free Spins (13%), and Swing Outs (9%). Moves that could not be classified in one of the defined families made up about a tenth (11%) of the entire distribution.

An interesting finding was that the rhythm condition had almost twice as many Swing Outs as the Solo and Background conditions (13% vs. 7%). This might indicate that the dancers were in more of a flow in the Rhythm condition, leading them to repeat many Swing Out variations in a row.

4.5 Descriptives

The data consisted of recordings of 24 dyads, comprising a total of 8 participants, in different combinations. Since all dyads danced to all 12 stimuli, the total dataset spanned 288 data points, with 72 per condition.

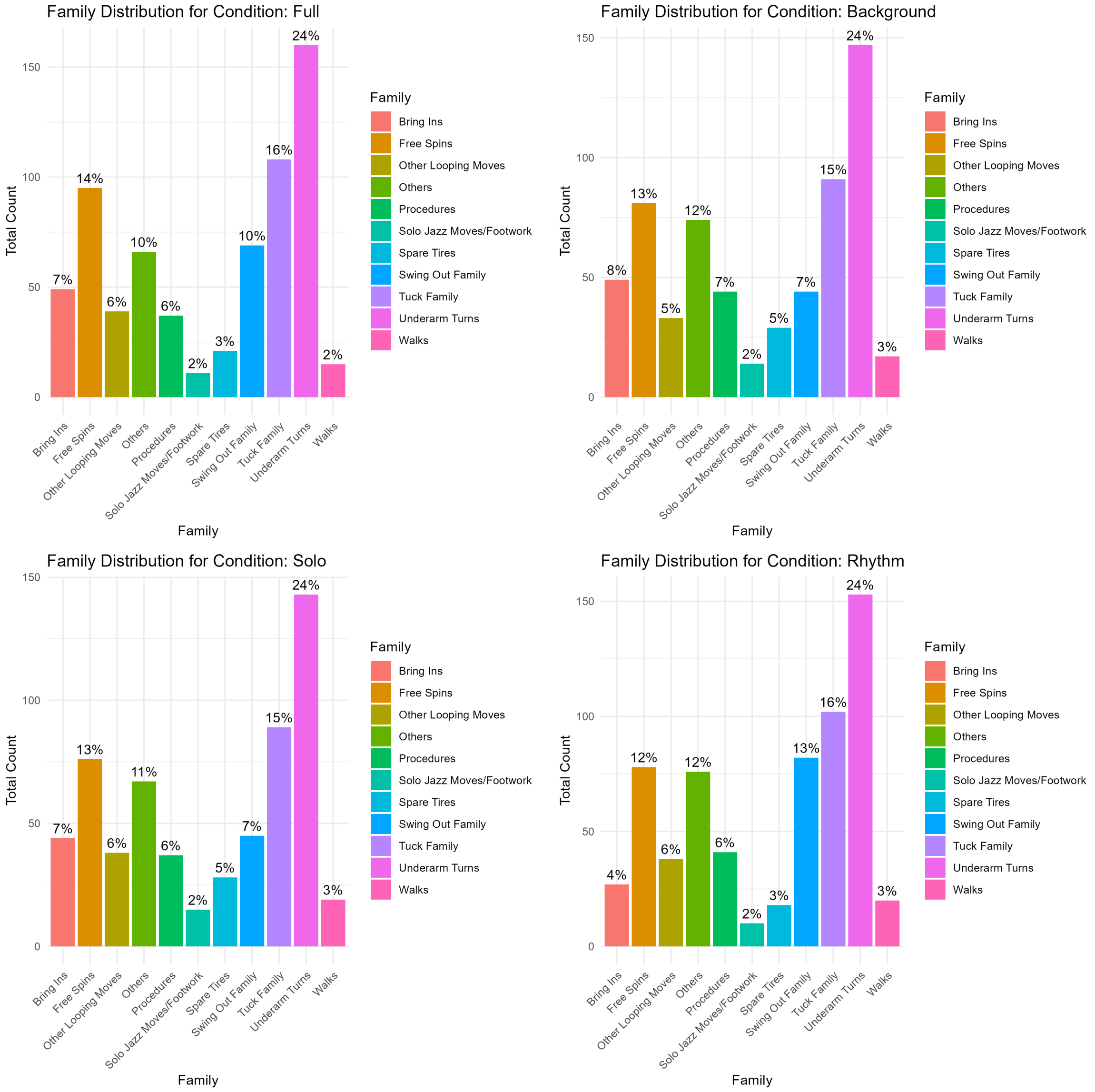


Figure 10: The distribution of move families per condition.

The sum of Cues per clip and the sum of Moves per clip were divided by the length of the stimulus in phrases (1.5 for Song 1 "Doodlin" and two for the other songs). A phrase consists of 32 beats. The data was clustered by Condition.

The Full Ensemble condition was, on average, danced with the most Cues, most *Actual Digital Lead Cues (ADLCs)*, and most Moves per phrase, but also had the lowest Richness. The Soloist condition had, on average, the highest diversity of moves (Richness), but was also danced with the least Cues, the least *actual Digital Lead Cues (ADLCs)*, and the least Moves (see Table 5).

For the Songs, Doodlin [37] had the highest means on all variables. Tatoe Pie [38] had the lowest amount of Cues and Richness, while A Smooth One [36] had the lowest Moves per phrase and *ADLCs* (see Table 4). This does not imply that the differences are significant, and this will further be investigated in Chapter 4.6.

Cues per phrase did not significantly deviate from normality in any condition. However, some values were quite low (Full: $p = 0.150$, Background: $p = 0.076$, Rhythm: $p = 0.188$, Solo: $p = 0.563$). Nonetheless, the check for normality of residuals in the linear model showed high normality (Shapiro-Wilk: $p = 0.965$), indicating that the linear model can still be evaluated reliably. The actual DLCs significantly deviated from normality in the Full ($p < 0.001$) and Background ($p = 0.042$) conditions, supporting the decision to cap the Digital Lead Cues, as this improved normality relative to uncapped values.

Moves per phrase significantly deviated from normality on all conditions (Full: $p = 0.03$, Background: $p = 0.014$, Rhythm: $p = 0.001$, Solo: $p = 0.025$). In combination with the low agreement on Moves in the interrater-reliability (see Chapter 4.2), this indicates that the results can only be interpreted tentatively. A Linear Mixed Model could still be used, as the residuals still showed high normality (Shapiro-Wilk: p=0.970).

Cues divided by Moves have been investigated, however they significantly deviated from normality for all conditions (Full: $p = 0.029$; Background: $p < 0.001$; Rhythm: $p = 0.015$; Solo: $p < 0.001$). They also significantly deviated in all tested models ($p < 0.001$), so they were not considered in the quantitative analysis.

The Richness distribution deviated significantly from normality only for the Background condition ($p = 0.007$). The overall low values on normality for the Background condition might be caused by systematic differences in the overall activity in Song 1 ("Doodlin") as observed in Chapter 4.6. All means, standard deviations, and normality tests can also be found in Table 5.

	Song	**Cues per phrase**	**Moves per phrase**	***ADLCs* per phrase**	**Richness**
	A Smooth One	6.72	4.59	7.05	0.607
Mean	Doodlin	7.08	5.20	7.29	0.666
	Tatoe Pie	6.68	4.73	7.10	0.601
	A Smooth One	1.17	0.634	1.34	0.112
Standard deviation	Doodlin	1.38	0.890	1.49	0.150
	Tatoe Pie	1.24	0.628	1.48	0.120

Table 4: The mean and standard deviation of all variables for all Songs [2, 1].

	Condition	Cues per phrase	*ADLCs* per phrase	Moves per phrase	Richness
Mean	Background	6.43	6.75	4.77	0.636
	Full	7.26	7.66	5.10	0.598
	Rhythm	7.23	7.59	4.92	0.617
	Solo	6.39	6.60	4.57	0.649
Standard deviation	Background	1.16	1.40	0.807	0.125
	Full	1.08	1.38	0.721	0.132
	Rhythm	1.13	1.18	0.704	0.128
	Solo	1.43	1.47	0.762	0.137
Shapiro-Wilk W	Background	0.969	0.965	0.957	0.951
	Full	0.974	0.926	0.962	0.978
	Rhythm	0.976	0.977	0.935	0.979
	Solo	0.985	0.986	0.961	0.987
Shapiro-Wilk p	Background	0.076	0.042	0.014	0.007
	Full	0.150	<.001	0.030	0.228
	Rhythm	0.188	0.210	0.001	0.274
	Solo	0.563	0.608	0.025	0.687

Table 5: The mean, standard deviation, and normality test of all variables for all conditions [2, 1].

4.6 The Influence of the Stimuli on the Dance Style

4.6.1 Comparing Different Models for Model Fit

For the linear mixed models, the audio features *Tempo variation* and *HPCP crest* were tested as predictors directly, along with Conditions and Song as additional factors. The predictors were tested individually and in combinations, and the best models were determined. All models were clustered by dyad (random intercept effect). *HPCP crest* was shown to be non-significant for all assessed variables, even when it was the sole predictor in the model (Cues: $p = 0.110$, ADLCs: $p = 0.107$, Moves: $p = 0.842$, Richness: $p = 0.957$).

For Cues and Actual DLCs, the best-performing model used only the Condition as a predictor (Cues: $LogLikelihood = -453; AIC = 918; BIC = 940$, ADLCs: $LogLikelihood = -490; AIC = 991; BIC = 1013$). The model using Condition and Song had a higher Log Likelihood and equal AIC, but a lower BIC due to the higher complexity for Cues, but performed a bit worse for Actual DLCs (Cues: $logLikelihood = -451; AIC = 918; BIC = 948$, ADLCs: $LogLikelihood = -490; AIC = 997; BIC = 1026$).

The other Linear Mixed Models had similar fits, with the model using both Condition and *Tempo variation* as predictors scoring second on Log Likelihood and AIC (Cues: $LogLikelihood = -455; AIC = 924; BIC = 950$, ADLCs: $LogLikelihood = -493; AIC = 999; BIC = 1025$). *Tempo variation* was not a significant predictor when combined with condition ($p = 0.194$), but was highly significant as a sole predictor ($p < 0.001$).

The model with *Tempo variation* as the only predictor explained 7.2% of the variance for Cues with fixed effects ($MarginalR^2 = 0.072$), the Condition alone explained 10.6% of the variance ($MarginalR^2 = 0.106$), and the combination of both explained 11% of the variance ($MarginalR^2 = 0.11$). This, in combination with the non-significance of *Tempo variation* when combined with the Condition indicates a strong overlap between the predictors, suggesting that more than half of the prediction strength the Condition has (6.8%) is due to the *Tempo variation* feature and, therefore, beat clarity. That leaves about 3.8% of the variance explained solely by Condition, not by Tempo variation. For comparison: The clustering by dyads explained 16% of the variance ($ConditionalR^2 = 0.27$ for the model with *Tempo variation* and Condition). Due to these results, we will look at the plots of both models, the Condition and the *Tempo variation*, as some relationships are more visible in the second.

For the Moves per phrase, the model with both Song and Condition as predictors performed best ($Loglikelihood = -302; AIC = 619; BIC = 648$), with both predictors being significant ($p < 0.01$). It accounted for 17.5% of the variance ($MarginalR^2 = 0.175$,) while Condition alone accounted for only 6.3% ($MarginalR^2 = 0.063$). Song alone predicted

11.3%. Adding the interaction between Song and Condition to the model is not significant ($p = 0.5$).

When it comes to Richness of moves, the Condition was interestingly not a significant predictor ($p = 0.095$), and a model with the Condition (as sole or additional predictor) performed worse than the model with only *Tempo variation* as a predictor, which had the best fit ($LogLikelihood = 170; AIC = -333; BIC = -318$). *Tempo variation* explained 2.2% of the variance, while the dyad explained 3.2% ($MarginalR^2 = 0.022; ConditionalR^2 = 0.054$).

Due to the evaluation, the predictors used are only Condition and only *Tempo variation* for Cues per phrase (and ADLCs as control), Condition and Song for Moves per phrase, and only Tempo variation for Richness. The parameter estimates for the primary models can be found in Tables 6 to 9.

4.6.2 Assumption Checks

Only one chosen model showed significant deviations from normality in the distribution of residuals, as assessed by the Shapiro-Wilk test. However, there were some notable differences in p-values between some models.

Both models for Cues had high p-values, indicating that they fit normality well (Condition only: $p = 0.963$; *Tempo variation* only: $p = 0.990$). As anticipated, the models for *ADLCs* had relatively low normality of residuals with *Tempo variation* even significantly deviating (Condition only: $p = 0.1$; *Tempo variation* only: $p = 0.031$). That, again, reinforces the idea that restricting the cue count to meaningful cues was sensible.

For the Moves, the model showed poor fit to normality ($p = 0.177$). While this still did not significantly deviate from normality, it is important to note, as Moves per phrase have already been shown to not be distributed normally (see Chapter 4.5) and have low interrater-reliability (see Chapter 4.2).

The model for Richness showed medium compliance to the assumption of normality ($p = 0.476$).

4.6.3 Results

For the Cues per phrase, the Conditions with Full Ensemble and Rhythm had significantly more cues than the conditions Background and Solo ($p < 0.001$). The *Tempo variation* as sole predictor also significantly predicted Cues per phrase ($p < 0.001$) with an estimate of -0.0316 Cues per phrase per point in *Tempo variation*.

				95% Confidence Intervals				
Names	**Effect**	**Estimate**	**SE**	**Lower**	**Upper**	**df**	**t**	**p**
(Intercept)	(Intercept)	6.8258	0.123	6.583	7.068	23.0	55.408	<.001
Condition 1	Full - Background	0.8264	0.183	0.466	1.187	261.0	4.514	<.001
Condition 2	Rhythm - Background	0.7963	0.183	0.436	1.157	261.0	4.350	<.001
Condition 3	Solo - Background	-0.0417	0.183	-0.402	0.319	261.0	-0.228	0.820

Table 6: The parameter estimates for the predicted variable Cues per phrase with the Condition as the predictor [1, 2].

			95% Confidence Intervals				
Names	**Estimate**	**SE**	**Lower**	**Upper**	**df**	**t**	**p**
(Intercept)	6.8258	0.12319	6.5833	7.0683	23.0	55.41	<.001
TempoVariation	-0.0316	0.00609	-0.0436	-0.0196	263.0	-5.19	<.001

Table 7: The parameter estimates for the predicted variable Cues per phrase with the Tempo variation as predictor [1, 2].

Names	Effect	Estimate	SE	95% Confidence Intervals Lower	Upper	df	t	p
(Intercept)	(Intercept)	4.841	0.0715	4.7006	4.9822	23.0	67.68	<.001
Condition 1	Full - Background	0.333	0.1067	0.1234	0.5433	259.0	3.13	0.002
Condition 2	Rhythm - Background	0.153	0.1067	-0.0572	0.3627	259.0	1.43	0.153
Condition 3	Solo - Background	-0.194	0.1067	-0.4044	0.0155	259.0	-1.82	0.069
Song 1	Doodlin - A Smooth One	0.608	0.0924	0.4258	0.7895	259.0	6.58	<.001
Song 2	Tatoe Pie - A Smooth One	0.135	0.0924	-0.0464	0.3172	259.0	1.47	0.144

Table 8: The parameter estimates for the predicted variable Moves per phrase with the Condition and the Song as predictor [1, 2].

Names	Estimate	SE	95% Confidence Intervals Lower	Upper	df	t	p
(Intercept)	0.62491	0.00890	0.607	0.64243	23.0	70.21	<.001
TempoVariation	0.00181	6.95e-4	4.46e-4	0.00318	263.0	2.61	0.010

Table 9: The parameter estimates for the predicted variable Richness with the Tempo variation as predictor [1, 2].

The difference between Full Ensemble/Rhythm and Background/Solo was between 0.7963 and 0.8681 Cues per phrase. Given that the mean cue count was between 6.60 and 7.66 Cues per phrase (see Chapter 4.5), the Condition makes a difference of about 10%, which means it is not just significant, but also meaningful. The plots for Conditions and *Tempo variation* are shown in Fig. 11.

These results were checked for bias by running the same models on the *actual DLCs*. There too, the Cues per phrase were significantly higher in the Full Ensemble/Rhythm conditions than in the Background and Solo conditions ($p < 0.001$), and the *Tempo variation* (as a sole predictor) also significantly predicted the cues ($p < 0.001$). The effects were slightly higher than for the model with capped cues, with an estimate of -0.0379 cues per phrase per point in *Tempo variation* and between 0.8380 and 1.0579 cues per phrase difference between the conditions. However, as *actual DLCs* have been shown to deviate from normality both in their overall distribution and in the distribution of error residuals, these results must be taken with a grain of salt and can only be used to confirm that capping cues did not bias the results in any meaningful way.

Overall, all four models provide evidence that the leader gave more cues when the beat was clearer.

For Moves, the model showed significant differences between the Full Ensemble and the Background Condition ($p = 0.002$), with more moves in the Full Ensemble Condition and a trend towards more cues in the Background Condition than the Solo Condition, however, not significant ($p = 0.069$). There was a significant difference between the songs Doodlin and A smooth one ($p < 0.001$), with Doodlin being danced with about 0.608 moves per phrase more than A smooth one, about twice as much as the biggest difference between Conditions (Full - Background: $Estimate = 0.333$). This shows that for the Moves per phrase, the Song was of much more importance than the specific Condition (see Fig. 12). However, explicit conclusions are not possible, as the distribution of moves was overall not normal.

The model for Richness showed that *Tempo variation* significantly predicted the Richness of moves ($p = 0.010$) with an estimate of about 0.181% of change in Richness per point in tempo variation (Estimate: 0.00181; transformed into a percentage as Richness can only reach between values of 0 and 1). The plot is shown in Fig 13.

In all models except the Moves per phrase model, the random intercept for the dyad explained more variance than the fixed effects (dyad offsets were additionally shown in Fig. 11 and 13). The exhaustive statistics for all models are in Appendix B.

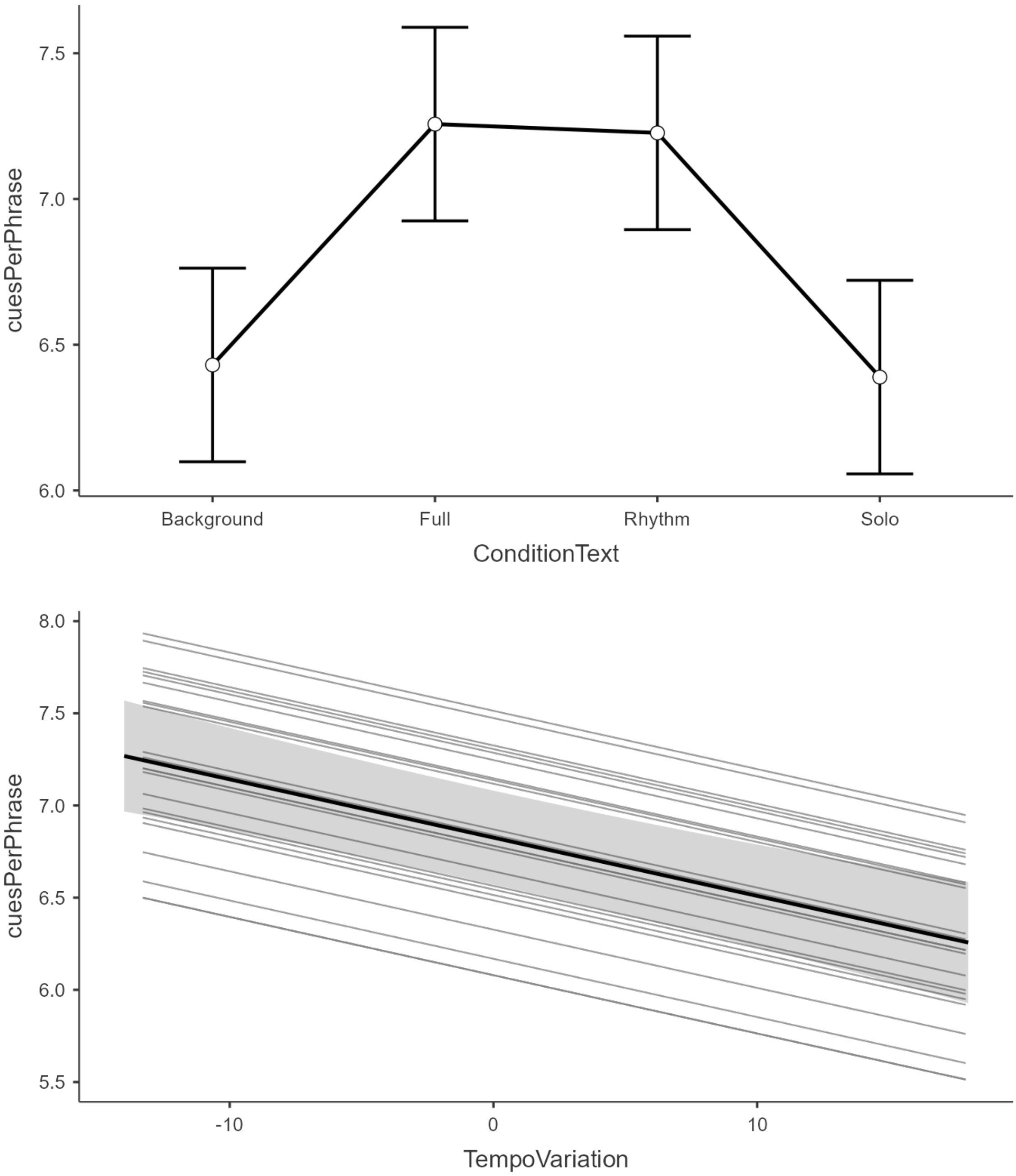


Figure 11: The plots of Cues per phrase per Condition and in relation to *Tempo variation* [1, 2].

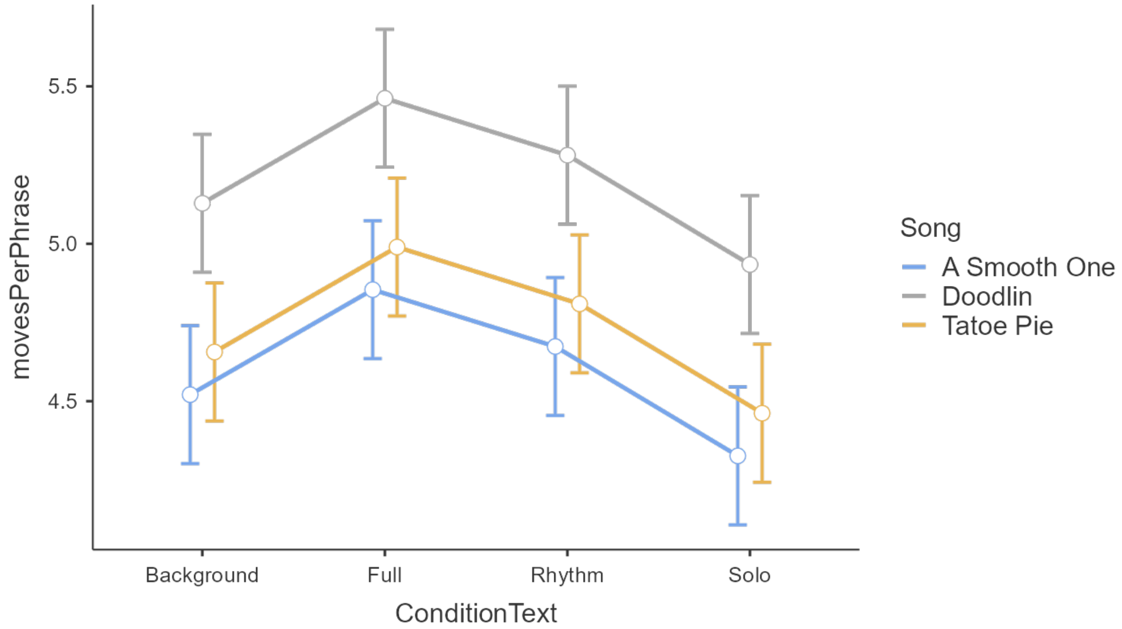


Figure 12: The plot of Moves per phrase per Condition and Song [1, 2].

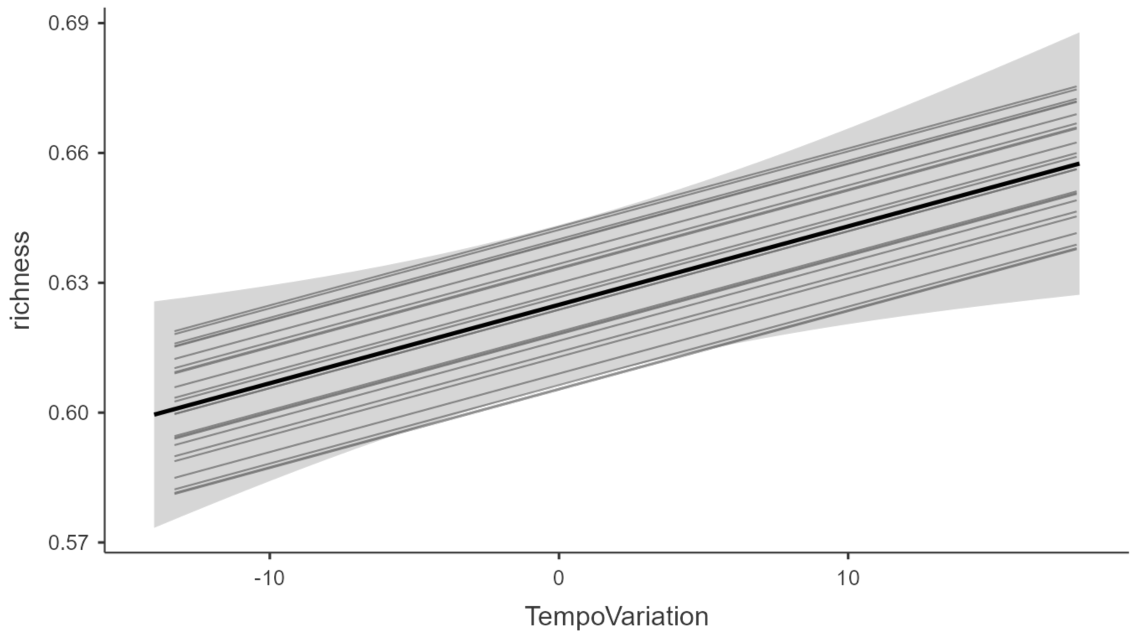


Figure 13: The plot of Richness per *Tempo variation* [1, 2].

Chapter 5

Discussion

This section discusses the results and the implications for future research. Chapter 5.1 discusses the results of the models in light of the defined hypotheses and what they may mean for the influence music has on the cognitive processes affecting the dancing performed. Chapter 5.2 dives into the importance of this study and what implications can be drawn from the results on a more general matter. In Chapter 5.3 the author gives a report of the personal experience in dancing Lindy Hop in different social cultures and what the differences between cultures may mean for the generalizability of the results. Chapter 5.4 expands on those limitations. Chapter 5.5 discusses possible endeavors for further research based on the results of this thesis.

5.1 The Cognitive Aspects: An Interpretation of the Results

Chapter 4.3 showed that there was no difference in variance between dyads with the same vs. different leaders. This indicates that followers play an important role in the dance style as well, and it emerges in the interaction between the two dancers rather than being a sum of their individual dance styles. This reinforces the idea of interpersonal synchronization as a synergy, creating more than the sum of the individual persons' dances.

As mentioned in Chapter 4.6 *HPCP crest* was shown to be non-significant for all assessed variables, even when it was the sole predictor in a model. As *HPCP crest* was the main predictor reflecting melody, this might indicate that melody does not play a major role in leading behaviour. This is especially interesting, as previous research has found that dance also reflects melody. While it is little surprise that the number of Cues is not directly influenced by it, one might have expected it to at least influence the diversity of moves performed.

There are multiple explanations for this surprising result. On one hand, HPCP crest might not be a sufficient reflection of melody in the stimuli, as it also showed no significant differences between most Conditions when tested in Chapter 3.2. However, the Conditions as a predictor did not predict the Richness better than Tempo variation on its own. One possible explanation is that HPCP crest accurately reflected the melody in the Conditions, and there were simply no significant differences between Full Ensemble, Background, and Soloist in the amount of melody. This certainly makes sense for the comparison between Full Ensemble and Soloist. However, it is hard to imagine, from a phenomenological perspective, that the Background alone would be perceived as equally melodic as the Full Ensemble. The other possibility is that melody is reflected in other aspects of the dance, such as fluidity of hand movement or foot movement, that are not associated with the amount or diversity of Cues or Moves.

The two models for Cues per phrase, with Condition and Tempo variation as predictors, respectively, had high overlap in explanatory power. About 3.8% of the variance was explained solely by Condition, indicating that more than just beat clarity influences the lead style. However, this accounted for only a small percentage of the variance, suggesting that beat clarity is at least a major influence.

Interestingly, the dyad as a random effect explained about 16% of the variance in both models. This supports previous research that the personality and gender of the dancers may influence how actively they dance. Some dyads consisted of two men or two women. It would be interesting to test whether there were systematic differences between those. However, this is not possible within this research due to the gender bias in the primary role.

The Moves per phrase were an interesting matter, since the distribution was not normal and the interrater-reliability was low. The results of the models indicate that the Song explained much more variance than the Condition. These differences could not be explained by Tempo variation or the HPCP crest, indicating that other aspects of the Songs may cause these changes. The Moves per phrase are also interesting in another aspect, as the fixed effects from the dyad explained less variance than Song and Condition together, as if it played less of a role there. However, due to the lack of normality and the lower interrater-reliability this might just be an artifact.

Given the results discussed in Chapter 4.6, we find support for the hypothesis that a clear beat is the most relevant:

> H2: Beat encourages artistic dancing.

There is no evidence that melody plays any relevant part in the overall amount of cues performed, nor does the interaction of rhythm and melody. While the analysis of the move count points in the same direction, with higher counts for conditions with a clearer beat, it cannot be taken on its own, given the lack of normality.

These results support the notion that cognitive processing is limited and that less clarity in the beat hinders the leader's ability to plan ahead and perform more complex movements, leading them to use fewer cues overall. This aligns with previous research findings that timing worsens on complex movement and syncopation. However, since the Conditions explained the variance better than just the Tempo variation, there might also be other factors contributing to this difference in Cue count. One possible explanation is that dancers moved more in general when low-frequency elements (like kick-drum) were present, as some previous studies showed that the overall urge to move increased when low-frequency elements were louder, independent of beat clarity [8].

As seen in the distributions, there was a higher distribution of Swing Outs in the Full Ensemble and especially the Rhythm condition than in the Soloist and Background conditions. This also indicates that a clear beat helps dancers get into a flow where they repeat swinging moves (that belong to the family of Swing Outs) multiple times in a row. Ironically, this flow hinders the diversity of moves as the focus on Swing Outs prevents the dancers from performing other moves.

When we look at the Richness of moves, it seems that the lack of a clear beat and the associated flow lead to greater diversity. However, this is not necessarily the case. If we look at the overall distributions, there are about five different move families that make up the majority of the distribution (72%). Given that the average diversity of moves was at a minimum 59.8% for the Full Ensemble and at a maximum 64.9% for the Soloist condition, it is highly possible that this is just a side effect of the fact that fewer moves were performed overall in the latter. In fact, the effect disappears when the model is run on the absolute number of move families per clip rather than the Richness variable.

5.2 Implications of the Study

This study not only added to the research on Lindy Hop, which is so far pretty limited, but also took an entirely new approach. While most studies directly examined physical measures, such as movement of different body parts or the locality of movement in relation to music, limiting conclusions about cognitive processes, this thesis combined an embodied approach with a systematic analysis and quantitative comparisons.

The author learned the dance in question themselves from scratch and observed their cognitive structure while dancing throughout the learning process of multiple courses, representing the embodied aspect. In discussion with an experienced Lindy Hop dancer and research colleague, they developed a systematic grammar used for labeling the Cues and Moves in the data, rather than mere physical properties. The labeled amount of Cues proved stable and reliable across different, independent raters from separate dancing schools.

Through this combination of methods, it is possible to provide tentative evidence for the cognitive processes underlying the planning and execution of motor movements. Previous studies have shown that complex movement and high syncopation reduce timing accuracy [13, 18], competing for cognitive resources. This thesis now shows that the clarity of a beat also directly affects how actively the leader leads, indicating that the increased cognitive resources required for beat tracking result in fewer cues being given.

5.3 An Embodied Report on the Cultural Variability of Lindy Hop

In recent years, a new form of research has formed, called Embodied Research [48]. It describes the idea that research is not performed as an outside observer, but with and through the body one has. Corinne Jola stresses the importance of scientists being embodied in their field of study and participating in it actively [49].

During the process of this thesis, I took this embodied approach myself and learned Lindy Hop from scratch. All participants of the experiment were also located in Vienna and frequently dancing on socials in that area. Due to my moving to Karlsruhe during my thesis, I got to attend socials there as well and experience the cultural shift between the two towns, which I want to expand on in this section.

It is important to understand that Lindy Hop is a highly improvisational and social dance that frequently borrows from other dance styles (as mentioned in Ch. 2.6). This means it is much more variable than more rigid ballroom dances, shaped by the local community's cultural norms.

I got to experience this when I first attended a Social Dance in Karlsruhe and danced my first dance with switching roles. My dance partner led moves I had never seen before, despite having labeled extensive MotionCapture data, and I led moves he had never heard of. One instance in which that was especially apparent was in the Classic Closed position of Lindy Hop. In Viennese dance schools, we're taught to do a *6-count* by default, with *8-counts* being more of a variation. My dance partner in Karlsruhe, however, had learned to dance *8-counts* by default, with *6-counts* being more rare. This led to a lot of confusion in our initial dance as I did not lead the back step in a *6-count* (as it was the default), but the forward step in an *8-count*, and he did the opposite. Only after discussing our different ideas about the default base step in the Closed position were we able to coordinate our dancing much better.

This anecdote shows one important aspect of the present research: it reflects Lindy Hop as it is danced in Vienna. When we look at the cues given by the leaders, and we count a cue for the *8-count* base steps, but not for the *6-count* base steps, then that happens because the *6-*

count is the default in Vienna and does not require an extra cue, while the *8-count* does. This coding works because all of our dancers were from the local Viennese community, but we must be aware that different local communities may have different cognitive maps of Lindy Hop, and that this might mean the results of different musical stimuli manifest differently in those communities. This work can only make statements about how the complexity of moves is affected, but not which moves are considered how complex in each community.

5.4 Limitations of this Study

As this experiment had only 8 participants, only 24 dyads could be investigated. As the dyads explained most of the variance, it would be interesting to see how the results may change with more participants. Because each stimulus was only about 30 seconds long, the time to get used to the music was also limited. It might be that dancers become more creative in their movements the longer they dance to a song, using only their most standard repertoire in the initial bars.

Another limitation is that the study was conducted in German-speaking countries in Europe, with mostly Caucasian, middle-class participants, and therefore does not reflect Lindy Hop in its traditional Black working-class culture but only the revivalist version.

Also, the labeling was performed manually on a qualitative basis rather than quantitatively, with e.g., self-supervised machine learning. On the one hand, this is an advantage because machines tend to find patterns even when they do not mean anything, and because the same move in different positions must be abstracted into a single cluster. On the other hand, this also means that mistakes can be made and that labels may not perfectly reflect the danced moves. This showed especially when it comes to labeling moves, where the interrater-reliability was low, indicating that those labels might not be too accurate within the same rater as well. As the cues proved to be fairly reliably labeled between raters, the discussion therefore focused on them.

Last but not least, this thesis aimed at investigating abstract components of dance (cues and moves) that may be used to induce conclusions about the cognitive processes underlying them, however there were no direct measurements of mental activity like EEG or fMRIs, so it is important to note that these ideas are founded in the embodied perception of the author and the idea that the leaders capacity to plan moves in advance directly influences the amount of cues performed.

5.5 Further Research

This thesis opens up interesting avenues for further research.

The most obvious one is how the results might change if more dancers and dyads were involved. How might the results be influenced by the personality of the dancers or their social context? What differences might be found between dancers of different countries or ethnicities, especially comparing dancers that were raised in Black culture with those that dance detached from the culture that gave rise to it?

However, it would also be very interesting to test how the dancers compensate for reduced lead cues. Do they fallback to more local movement, performing footwork variations, or more improvisational movement until they have planned their next movements and are ready to perform them? Does the follower increasingly contribute to the dance, providing movement ideas or hinting at the leader to compensate for his lack of leading? Do dancers tend to create the beat and rhythm themselves when it is lacking?

A third research endeavor might be to determine how the grammar could be improved to ensure greater reliability in the labels for moves. Or whether the differences between raters are an inherent necessity of the high improvisational freedom and wide differences between differing social contexts and dance schools. It would also be interesting to investigate which other abstract measures could be derived from the MotionCapture data that could provide insight into the cognitive processes behind.

Chapter 6

Conclusion

This thesis offered a new perspective by investigating the understudied dance style, Lindy Hop, through an Embodied Research approach. Previous research had already shown that timing to music worsened both for complicated rhythms and complex movements. However, those studies did not allow the dancers to freely change their movements.

This study addressed this research gap by investigating how freely improvised movement changed across different musical stimuli. 24 dyads danced to 12 stimuli freely while being recorded in MotionCapture suits. The stimuli were derived from three songs, each split into the Conditions Full Ensemble, Background, Soloist, and Rhythm. Assumptions about beat clarity and melody were tested using audio features extracted with the PADMEA software.

Labels were applied to the MotionCapture recordings based on a previously developed grammar. The reliability of the labels was tested using duplicate labels for 12 of the 288 clips, with one set provided by an external rater. Linear mixed models, with audio features and Condition and Song as predictors, were tested, and the best models were selected.

The study has shown that a clear beat encourages leaders to provide more lead cues and lead more actively. That aligns with the previous research showing that complex movement and high syncopation can both reduce timing accuracy. As the dancers were able to improvise freely in this study rather than follow a fixed routine, it is likely that they adjusted the number of complex movements to their ability to time to the beat, which depended on its clarity. It supports the idea of a bottleneck in cognitive processing, with different aspects of dancing taking up different amounts of resources depending on how much is needed for the unchangeable aspect (the music).

There might also be other explanations for the results, such as the increased urge to move with a stronger base drum, which was found in previous research. This study found no effect of melody on the number of cues performed, indicating that it might not be as relevant a factor in Jazz dance as in Western dance styles. However, the lack of a significant effect does not confirm the absence of an effect; the study population might also have been too small.

This study is limited to Caucasian residents of Vienna. Lindy Hop is a very diverse dance, and there are many differences between cultural groups based on other Swing dance styles taught in their respective cultures, since they influence one another. More research in different social groups, both culturally and ethnically, is needed to confirm these findings.

It is also worth noting that the biggest influence on Cue count was the dyad dancing, rather than the stimuli. Therefore, it is a relevant endeavor to test whether these differences can be predicted by personality measures or by the match between partners' average tapping speeds.

What can be said with some confidence, however, is that beat clarity exerts a stronger influence on Jazz dance than melody, supporting hypothesis H2.

Bibliography

[1] “Jamovi - open statistical software for the desktop and cloud,” https://www.jamovi.org., accessed: 2025-10-9.

[2] “The comprehensive R archive network,” https://cran.r-project.org., accessed: 2025-10-9.

[3] M. Gallucci, “GAMLj,” https://gamlj.github.io/., accessed: 2025-10-9.

[4] Patil, *Extracting, Computing and Exploring the Parameters of Statistical Models using R*. CRAN. link, 2020.

[5] B. Burger, M. R. Thompson, G. Luck, S. H. Saarikallio, and P. Toiviainen, “Hunting for the beat in the body: on period and phase locking in music-induced movement,” *Frontiers in human neuroscience*, vol. 8, p. 903, 2014.

[6] “Tables generator,” https://www.tablesgenerator.com/, accessed: 2025-11-07.

[7] M. Pearce and M. Rohrmeier, “Music cognition and the cognitive sciences,” *Topics in cognitive science*, vol. 4, no. 4, pp. 468–484, 2012.

[8] B. Burger and P. Toiviainen, “Embodiment in electronic dance music: Effects of musical content and structure on body movement,” *Musicae Scientiae*, vol. 24, no. 2, pp. 186–205, 2020.

[9] V. Sevdalis and P. E. Keller, “Captured by motion: Dance, action understanding, and social cognition,” *Brain and cognition*, vol. 77, no. 2, pp. 231–236, 2011.

[10] W. T. Fitch, “Dance, music, meter and groove: a forgotten partnership,” *Frontiers in human neuroscience*, vol. 10, p. 64, 2016.

[11] M. Leman and P.-J. Maes, “The role of embodiment in the perception of music,” *Empirical musicology review*, vol. 9, no. 3-4, 2015.

[12] P.-J. Maes, M. Leman, C. Palmer, and M. M. Wanderley, “Action-based effects on music perception,” *Frontiers in psychology*, vol. 4, p. 1008, 2014.

[13] B. Bläsing, B. Calvo-Merino, E. S. Cross, C. Jola, J. Honisch, and C. J. Stevens, "Neurocognitive control in dance perception and performance," *Acta Psychologica*, vol. 139, no. 2, pp. 300–308, 2012. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0001691811002320

[14] B. H. Repp and Y.-H. Su, "Sensorimotor synchronization: a review of recent research (2006–2012)," *Psychonomic bulletin review*, vol. 20, pp. 403–452, 2013.

[15] S. Dixon, "Beat induction and rhythm recognition," in *Australian joint conference on artificial intelligence*. Springer, 1997, pp. 311–320.

[16] E. Van Dyck, D. Moelants, M. Demey, A. Deweppe, P. Coussement, and M. Leman, "The impact of the bass drum on human dance movement," *Music Perception: An Interdisciplinary Journal*, vol. 30, no. 4, pp. 349–359, 2012.

[17] P. E. Keller, G. Novembre, and M. J. Hove, "Rhythm in joint action: psychological and neurophysiological mechanisms for real-time interpersonal coordination," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 369, no. 1658, p. 20130394, 2014.

[18] C. I. Karageorghis, L. P. Lyne, M. Bigliassi, and P. Vuust, "Effects of auditory rhythm on movement accuracy in dance performance," *Human Movement Science*, vol. 67, p. 102511, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167945719302623

[19] V. Iyer, "Embodied mind, situated cognition, and expressive microtiming in african-american music," *Music perception*, vol. 19, no. 3, pp. 387–414, 2002.

[20] H. Spring, "Swing and the lindy hop: dance, venue, media, and tradition," *American Music*, pp. 183–207, 1997.

[21] M. Leman, P.-J. Maes, L. Nijs, and E. Van Dyck, "What is embodied music cognition?" in *Springer handbook of systematic musicology*. Springer, 2018, pp. 747–760.

[22] A. Miura, K. Kudo, T. Ohtsuki, and H. Kanehisa, "Coordination modes in sensorimotor synchronization of whole-body movement: a study of street dancers and non-dancers," *Human movement science*, vol. 30, no. 6, pp. 1260–1271, 2011.

[23] J. R. Matyja, "Embodied music cognition: Trouble ahead, trouble behind," *Frontiers in psychology*, vol. 7, p. 1891, 2016.

[24] L. Van Noorden and D. Moelants, "Resonance in the perception of musical pulse," *Journal of New Music Research*, vol. 28, no. 1, pp. 43–66, 1999.

[25] S. R. O'Connell, J. E. Nave-Blodgett, G. E. Wilson, E. E. Hannon, and J. S. Snyder, "Elements of musical and dance sophistication predict musical groove perception," *Frontiers in Psychology*, vol. 13, p. 998321, 2022.

[26] S. Castillo-Pérez, V. Gómez-Pérez, M. C. Velasco, E. Pérez-Campos, and M.-A. Mayoral, "Effects of music therapy on depression compared with psychotherapy," *The Arts in psychotherapy*, vol. 37, no. 5, pp. 387–390, 2010.

[27] J. Pressing, "Cognitive processes in improvisation," in *Cognitive Processes in the Perception of Art*, ser. Advances in Psychology, W. R. Crozier and A. J. Chapman, Eds. North-Holland, 1984, vol. 19, pp. 345–363. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0166411508623584

[28] S. Dixon, "Automatic extraction of tempo and beat from expressive performances," *Journal of New Music Research*, vol. 30, no. 1, pp. 39–58, 2001.

[29] H. Honing, "Without it no music: beat induction as a fundamental musical trait," *Annals of the New York Academy of Sciences*, vol. 1252, no. 1, pp. 85–91, 2012.

[30] J. Phillips-Silver, C. A. Aktipis, and G. A. Bryant, "The ecology of entrainment: Foundations of coordinated rhythmic movement," *Music Perception*, vol. 28, no. 1, pp. 3–14, 09 2010. [Online]. Available: https://doi.org/10.1525/mp.2010.28.1.3

[31] F. Gouyon, G. Widmer, X. Serra, and A. Flexer, "Acoustic cues to beat induction: A machine learning perspective," *Music Perception*, vol. 24, no. 2, pp. 177–188, 2006.

[32] B. Burger, M. R. Thompson, S. Saarikallio, G. Luck, and P. Toiviainen, "Influence of musical features on characteristics of music-induced movements," in *Proceedings of the 11th International Conference on Music Perception and Cognition (ICMPC)*, 2010, pp. 425–428.

[33] S. Carroll, "The lindy binge: the social and cultural functions of lindy exchanges," *Continuum*, vol. 20, no. 4, pp. 447–456, 2006. [Online]. Available: https://doi.org/10.1080/10304310600987262

[34] K. Unruh, "May we have this dance?: Cultural ownership of the lindy hop from the swing era to today," *Atlantic Studies*, vol. 17, no. 1, pp. 40–64, 2020. [Online]. Available: https://doi.org/10.1080/14788810.2019.1698241

[35] IG HOP, "Courses lindy hop beginner i and ii and intermediate," 02 2025. [Online]. Available: https://ighop.at/kurse/lindy-hop-kurse/

[36] Hot Swing Sextet, "A smooth one - black market stuff," 2018. [Online]. Available: https://music.youtube.com/watch?v=MXgJ4yq8AZo&si=JuuXUlemU9AgvwAS

[37] H. Silver, "Doodlin - horace silver and the jazz messengers," 2004. [Online]. Available: https://music.youtube.com/watch?v=lQnxYseLmBo&si=1k6o1GY3P9P9U6e-

[38] Hot Swing Sextet, "Tatoe pie - what's your jive?" 2020. [Online]. Available: https://music.youtube.com/watch?v=6erw4Hk1rN8&si=4YHvxJgp-eTIwkfq

[39] O. GmbH, "lalal.ai," accessed on 05.10.2025. [Online]. Available: https://www.lalal.ai/

[40] celemony, "Melodyne 5 studio." [Online]. Available: https://shop.celemony.com/cgi-bin/WebObjects/CelemonyShop.woa

[41] O. Lartillot and P. Toiviainen, "A matlab toolbox for musical feature extraction from audio," in *International conference on digital audio effects*, vol. 237. Bordeaux, 2007, p. 244.

[42] O. Lartillot, P. Toiviainen, and T. Eerola, "A matlab toolbox for music information retrieval," in *Data Analysis, Machine Learning and Applications: Proceedings of the 31st Annual Conference of the Gesellschaft für Klassifikation eV, Albert-Ludwigs-Universität Freiburg, March 7–9, 2007*. Springer, 2008, pp. 261–268.

[43] S. Wingert, "Lindy hop classification," 03 2025. [Online]. Available: https://muwiserver.synology.me:444/Wingert/

[44] ——, "Lindy hop classification," 07 2025. [Online]. Available: https://muwiserver.synology.me:444/Wingert2/

[45] ——, "External raters guidelines," uRL available at https://shorturl.at/y5lLq, accessed 2025-11-21.

[46] IG HOP, "Leader-support-fonds," 03 2025. [Online]. Available: https://ighop.at/project/lsf-20/

[47] G. Marzi, M. Balzano, and D. Marchiori, "K-alpha calculator–krippendorff's alpha calculator: a user-friendly tool for computing krippendorff's alpha inter-rater reliability coefficient," *MethodsX*, vol. 12, p. 102545, 2024.

[48] B. Spatz, "Embodied research: A methodology," *Liminalities*, vol. 13, no. 2, 2017.

[49] C. Jola, "Research and choreography: Merging dance and cognitive neuroscience," in *The Neurocognition of Dance*. Psychology Press, 2010, pp. 213–244.

[50] D. Bogdanov, N. Wack, E. Gómez Gutiérrez, S. Gulati, P. Herrera Boyer, O. Mayor, G. Roma Trepat, J. Salamon, J. R. Zapata González, and X. Serra, "Essentia: An audio analysis library for music information retrieval," 2013.

[51] M. M. e. a. McFee, Brian, "librosa/librosa: 0.11.0," Mar. 2025. [Online]. Available: https://doi.org/10.5281/zenodo.15006942

[52] P. Herrera and S. Streich, "Detrended fluctuation analysis of music signals: Danceability estimation and further semantic characterization," in *Audio Engineering Society Convention 118*. Audio Engineering Society, 2005.

[53] S. Streich *et al.*, *Music complexity: a multi-faceted description of audio content*. Universitat Pompeu Fabra, 2006.

[54] P. Grosche, M. Müller, and F. Kurth, "Cyclic tempogram—a mid-level tempo representation for musicsignals," in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2010, pp. 5522–5525.

Glossary

6-count

A term used in Lindy Hop to describe any move that is performed over the span of six 8ths. The standard footwork is rock-step-slow-slow or rock-step-triple step-triple step. With slow/triple step taking 2 beats each and rock and step taking one beat each.

8-count

A term used in Lindy Hop to describe any more that is performed over the span of eight 8ths. The standard footwork is rock-step-slow-rock-step-slow or rock-step-triple step-rock-step-triple step.

Beat induction

The ability of humans to perceive and synchronize with a beat.

Beats loudness

The loudness of the detected beats as measured by the spectral energy [50]. If a lot of beats are silent, this value is low.

Beats per minute (BPM)

The beats that a piece progresses through within a minute. A common tempo denotation for music.

Chroma entropy harmonic

The entropy that is contained in the chromagramm of the harmonic part of the signal [51, 50]. A chromagram is a projection of the auditory spectrum into the 12 semitones of Western music independent of octave or microtones. High entropy therefore means that a lot of semitones are present at the same time and you cannot identify one specific key.

Complex coordinated rhythmic movement

Any form of movement performed together with another human (coordinated) to a complex rhythm.

Danceability

Estimated danceability of a piece based on the algorithm described in [50, 52].

Dynamic complexity

This feature describes how much fluctuation is in the loudness of the stimulus. Musically spoken a high value of dynamic complexity means that the piece is not equally loud all the time, but has very silent and very loud parts. This might relate to beat clarity as clearly separated notes lead to high dynamic complexity (silence in the breaks, loudness during the note). Mathematically spoken this feature is calculated by the average (absolute) deviation of the globally averaged loudness level [50, 53].

Embodied Music Cognition

The idea that the involvement of our bodies strongly influences how we perceive, feel, experience and comprehend music. Our cognition is embodied, meaning that it only exists in that form through our bodily interactions (see Ch. 2.1).

Harmonic Pitch Profile (HPCP)

The Harmonic Pitch Class Profile describes how intensely each of the 12 semitones of Western music are present in the stimulus.

HPCP crest

The ratio of the peak of the HPCP distribution to the overall average. This is a good indicator for melody, as a high value means that some tones appeared more frequently than others, indicating that the tones moved within a common key.

HPCP entropy

The entropy of the *Harmonic Pitch Profile*. A high entropy means that the musical stimulus is spread across all 12 semi-tones and there is no peak at specific tones.

legato

A musical instruction to play notes without any separation in between.

Meter (Music)

A regularly recurring pattern in music, containing information about intended accents, e.g. a 3/4 meter in Waltz music.

Percussion harm(onic) ratio

The (spectral) energy of the percussive elements in ratio to the energy of the harmonic elements [51] This does first and foremost describe how rhythmic vs. how tonal a stimulus is. It does not imply that the rhythmic elements are particularly regular or provide a clear beat. It also does not imply that the tonal elements fit into the same harmony.

Rhythmic irregularity

A measure describing how rhythmic or arhythmic a stimulus is with 0 being perfectly rhythmic and 1 being fully arhythmic. It is calculated by determining the deviation of the average rhythm.

staccato

A musical instruction to separate individual notes clearly.

Syncopation

A rhythmic instance that conflicts with the intended meter of a piece. Examples are notes that are connected across bars or off-beat note onset.

Tempo variation

The standard deviation of the estimated tempo (in *BPM*) as perceived by the analysis software. The software estimates the tempo in 4s windows of the stimulus [51]. As all stimuli of the same song should objectively have the same amount of tempo variation a high variation perceived by the analysis software only in some stimuli indicates difficulties in tracking the beat and therefore low beat clarity.

Tempogram stability

Similar to tempo variation. It looks at the tempo stability in the percussive part of the stimulus, while identifying meters that are 2x the speed of another and clustering. Rhythms using syncopation or trioles are less stable, so this parameter is reflective of the complexity of a rhythm [54, 51].

Vernacular Swing/Jazz dance

A jazz dance that originated in the African-American culture, contrasted against more modern dance forms that have Western origins.

Vestibular system

A sensory system located in the ear that is responsible for our balance and spatial orientation.

Appendix A

The Original Instructions Provided to the External Rater

Lieber Rater,

vielen Dank, dass du dich bereit erklärt hast, einige MotionCapture Daten zu labeln. Das ist eine enorm wertvolle Aufgabe, denn damit können wir prüfen, ob Tänzer verschiedener Schulen die Daten ähnlich bewerten. Das liefert uns Informationen darüber, welche Variablen statistisch tragbar sind und welche individuell so verschieden bewertet werden, dass sie in der statistischen Analyse nicht verwendet werden können. Die Clips, die wir dir gesendet haben, wurden randomisiert mit einem Buchstaben versehen, damit du nicht von anderen Faktoren beeinflusst wirst. Bitte benenne die Dateien entsprechend dem Buchstaben des Clips, damit wir sie später wieder zuordnen können. Das Labeln der Dateien dauert im Schnitt 10-15 Minuten pro Clip. Mit Vorbereitungszeit kannst du also mit einem Zeitaufwand von 4-6h rechnen. Es ist empfehlenswert, das Labeln auf mehrere Tage aufzuteilen, da die Konzentration schnell sinken kann.

Bevor du mit dem Labeln beginnst, möchte ich dich bitten, das Tutorial-Video anzusehen und diese Anleitung aufmerksam zu lesen.

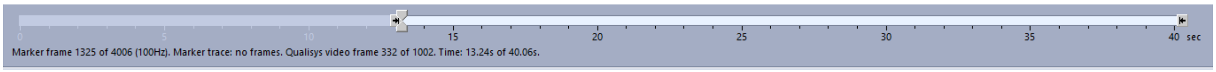
Du erhältst zusammen mit dieser Anleitung 4 weitere Ressourcen:

A) Eine Beispiel-Datei (im .xslx-Format) davon, wie Labels am Ende aussehen können; diese Datei ist von einem Clip, den du nicht labelst, lass dich also nicht vom genauen Inhalt irritieren
B) Eine allgemeine Übersicht für die verschiedenen Moves, die beschreibt, wie viele Cues, Moves und „Actual DLCs" standardmäßig bei dem jeweiligen Move gelabelt werden. Zudem gibt es manchmal erklärende Notizen. Diese Übersicht wird in dieser Anleitung noch genauer erklärt. Du findest sie unter: https://docs.google.com/spreadsheets/d/1xTR2dK3T8-Fyi7A6O9PfX3qCVUmE8ch6QnBeMqZrtTA/edit?usp=sharing
C) Die Lindy-Hop Website. Auf dieser Ressource sind alle Moves nach Familie gruppiert und Videos für jeden Move abgreifbar. Diese können die Standard-Ausführung oder auch Variationen sein. Du findest sie unter folgendem Link: https://muwiserver.synology.me:444/Wingert2/
Solltest du dir einmal nicht sicher sein, welchen Move eine Bezeichnung beschreibt oder wie eine bestimmte Position heißt, dann kannst du sie dir hier anschauen. Wenn es Probleme mit dieser Seite geben sollte, kannst du auch vorübergehend die alte Version unter https://muwiserver.synology.me:444/Wingert/ verwenden. Bitte informiere mich aber zeitnah über Probleme, da auf dieser manche Moves fehlen oder nicht optimale Beispiel-Clips haben.
D) Ein Video-Tutorial, das den Labeling-Prozess beschreibt. Wir empfehlen dieses zu Beginn zu schauen, da eine visuelle Erklärung deutlich leichter zu verstehen ist.
Du findest es hier: https://youtu.be/GFU97hV-SFM

Den Clip öffnen

Die Clips werden dir im QTM-Format gegeben. Das sind MotionCapture-Dateien. Um sie öffnen zu können, musst du dir die Software Qualysis Track Manager installieren. Dazu benötigst du einen Windows-Computer und eine Lizenz. Die Lizenz wurde bereits für deine Email-Adresse angefragt. Nachdem du deinen Account erstellt hast, kannst du die Software auf der Website downloaden und installieren. Auf der Website findest du auch viele Tutorials zur Verwendung der Software, doch die wichtigsten Punkte sind auch hier zusammengefasst.

Wenn du die Datei öffnest (siehe Video), siehst du im unteren Bereich eine Zeitlinie:



Die Dateien sind bereits auf den Bereich zugeschnitten, den du labeln sollst. Notiere dir am besten direkt den Start Frame in der ersten Zeile.

- Bewege den Cursor hin und her, um dich im Video vorwärts oder zurückzubewegen. Achte darauf die Begrenzungen (Pfeil links und rechts) nicht zu verschieben, bevor du Start und Ende des Clips notiert hast. Sollte eine Bewegung abgeschnitten sein, kannst du die Grenzen verschieben, um Beginn/Ende zu beobachten, setze sie danach jedoch wieder zurück.
- Um die Zeitliniengrenzen exakt zu verändern kannst du mit einem Doppelklick darauf ein Menü öffnen und bei First/Last selected frame die richtige Zahl eintragen.
- Wenn du das Video einfach nur abspielen willst, reicht es die Leertaste zu drücken.
- Mit dem Mausrad kannst du scrollen.
- Wenn du die linke Maustaste gedrückt hältst, kannst du um das Tanzpaar rotieren.
- Wenn du die rechte Maustaste gedrückt hältst, kannst du deine Perspektive verschieben.

Wie das Labeln generell funktioniert

Bevor wir uns die Beispiel-Datei genauer anschauen, eine kurze Exkursion zum Sinn der Labels:

Digitale Cues und Actual DLCs

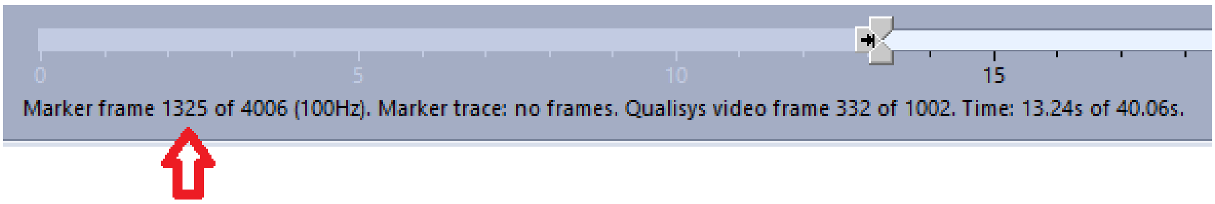
Wir wollen die Komplexität der Moves anhand der Anzahl von Cues messen, die sie (vom Leader) benötigen. Dabei ist es wichtig zwischen digitalen und analogen Cues zu unterscheiden. Stell dir das ganze so vor: Wenn du vor deinem Partner stehst und ihr in der offenen Position Grundschritte macht, dann gibt der Leader dem Follower Cues. Aber was würde der Follower machen, wenn der Leader vergisst einen Cue zu geben? Die automatische Reaktion wäre erst einmal weiter den Grundschritt zu machen. Damit aufzuhören würde aktive Aufmerksamkeit brauchen. Das ist ein analoger Cue. Ein digitaler Cue ist ein Cue, der beeinflusst, wie sich der Follower weiterbewegt. Zum Beispiel wenn man bei Mambo Steps die Richtung ändert.

Da bei manchen Moves die digitalen Cues nicht ganz die Komplexität reflektieren, gibt es manchmal Sonderregeln, was hier gezählt wird und was nicht.

Für die tatsächlichen digitalen Cues gibt es dann das optionale Feld, der „Actual DLCs". Genaueres dazu gibt es jedoch noch in der Übersichtstabelle (Ressource B).

Öffne nun die Beispiel-Datei (Ressource A), die dir geschickt wurde und sieh dir die oberste Spalte an. Verwende diese Datei gerne als Template für deine eigenen Labels. Im Folgenden gehen wir die Spalten einzeln durch, damit du verstehst, was du eintragen musst:

- Start Frame: Dies ist normalerweise der Zeitpunkt, an dem beide Füße etwa in der Grundposition stehen. Den genauen Frame kannst du unten in der Zeitlinie ablesen.
    - Ausnahmen: Bei der ersten oder letzten Zeile (Anfang oder Ende des Clips) entspricht der Start Frame der originalen linken oder rechten Begrenzung der Datei.
    - Wenn der erste digitale Cue (z.B. eine Armführung) vor der Grundposition liegt, kannst du den Start Frame entsprechend früher setzen.



- Position: Hier benennst du die Position, in der sich die Tänzer vor Beginn des Moves befinden. Die verschiedenen Positionen sind auf der Website einsehbar.
  Im Labeling-Prozess gibt es eine weitere Position, die auf der Website nicht erwähnt wird: No connection. Wenn die Tänzer über einen längeren Zeitraum ohne Berührung tanzen, gilt das als diese Position. Sie fällt unter die Kategorie Unspecified (siehe nächster Punkt).
    - Achtung: Wenn der erste Move abgeschnitten ist, weil er vor dem zugeschnittenen Bereich beginnt, dann prüfe, in welcher Position er begonnen wird und trage diese ein. Dafür kannst du die Grenze verschieben und später wieder zurücksetzen. Mehr zu abgeschnittenen Moves weiter unten.
- Closed/Open: In dieser Zeile trägst du ein, ob die Position eine offene (Open), eine geschlossene (Closed) oder etwas anderes (Unspecified) ist. Die Zuordnung von Positionen in diese Kategorien findest du auf der Website.
- Digital Lead Cues: Hier trägst du die digitalen Cues ein, die du im Move zählst. Orientiere dich dabei an der Erklärung und an der Übersicht.
- Moves: Hier werden die Moves gezählt. In aller Regel ist das immer 1. Es gibt wenige Ausnahmen, diese werden in der angehängten Tabelle erläutert.
- Looping?: In diesem Feld gibst du an, ob ein Move Looping oder Non-looping ist. Die Idee ist folgendermaßen: Wenn die Tänzer während dem Move ihre Position nicht verlassen, dann ist es ein Looping Move (z. B. Sugar Pushes), ansonsten nicht (z. B. Underarm Turns). In Looping Moves wird auch oft nur 1 digitaler Cue für das

Beginnen gezählt, da das Muster sich immer wieder wiederholt und die weiteren Cues daher die Komplexität nicht widerspiegeln. Wann das gilt, siehst du in der Übersicht.

- Type of move & Type simplified: In diesen Feldern gibst du an, welchen Move die Tänzer performen. In Type of Move hast du die Möglichkeit etwas genauere Aussagen zu treffen (z.B. Variation, Twisting, etc.), in Type simplified solltest du sie dann einer klaren Kategorie zuordnen. Alle möglichen Bezeichnungen sind in der mitgelieferten Übersicht aufgelistet.
- Actual DLCs: Dieses Feld ist optional und muss nicht immer befüllt werden. Für manche Moves, in denen die digitalen Cues nur eingeschränkt gezählt werden, ist dies die Möglichkeit die tatsächliche Anzahl an Cues einzutragen, die die Bewegung des Followers ändern.
- Frame Start Move: In diese Zeile trägst du den Frame ein, bei dem der erste Cue des Moves gegeben wird. Häufig ist das der Rock-Step, es kann jedoch auch vor oder nach dem Rock-Step passieren.
    - Ausnahme: Wenn ein Move bereits vor dem zugeschnittenen Bereich beginnt, ist der Start Frame auch hier der erste Frame des zugeschnittenen Bereichs.

Manchmal kann es vorkommen, dass Moves abgeschnitten sind. Etwa, weil sie vor der vorderen Grenze beginnen oder nach dem Ende enden. Dann gilt folgende Regel: Wenn mind. 1 Cue des Moves innerhalb der Begrenzungen durchgeführt wird, zählt man den Move und alle Cues, die innerhalb der Begrenzungen liegen. Wenn keiner der Cues innerhalb der Begrenzungen durchgeführt wird, dann lässt man die Felder Digital Lead Cues & Moves leer und schreibt nur die Art des Moves in die entsprechenden Felder (falls feststellbar). Wenn der Move nicht klar erkennbar ist, da er nicht vollständig aufgenommen wurde, lässt man das Feld „Type simplified“ leer, kann aber die Cues innerhalb der Begrenzung trotzdem noch zählen (und bei Moves „1“ eintragen), sofern vorhanden.

ACHTUNG! Wenn du vorherige Dateien als Template nutzt, achte bitte darauf alle Felder zu leeren, bevor du die neuen Werte einträgst, auch optionale!

Frame Start & Frame Start Move: Die Videos wurden mit 100fps aufgenommen, wir haben also eine Genauigkeit von 10ms. Natürlich kann kein Mensch so exakt eintragen, wann eine Position oder ein Cue begonnen hat und das ist uns bewusst. Diese Zeilen spielen eine geringere Rolle in der Analyse und es genügt den ungefähren Wert einzutragen (Ausnahme: Start- und Endbegrenzung).

Verwenden der Übersicht

Öffne nun die Übersicht (Resource B), die im Folgenden erläutert wird:

Uns ist bewusst, dass das Konzept der digitalen Cues sehr komplex ist. Deshalb haben wir dir diese Übersicht zusammen mit der Webseite bereitgestellt. Sie zeigt für jeden Move an, wie viele digitale Cues in der Standardvariante gezählt werden. Variationen können natürlich davon abweichen, aber es ist ein guter Anhaltspunkt dafür, was wie gezählt wird. Lass uns einen Blick auf die verschiedenen Spalten werfen:

- Family: Diese Spalte dient als eine ungefähre Orientierung zu welcher Familie Moves gehören. Unter dieser Familie findest du die Moves auch auf der Website, solltest du dir das Video dazu anschauen wollen.
- Move: Hier steht die Bezeichnung des Moves innerhalb der Familie. Diese Bezeichnung solltest du allgemein beim Labeln verwenden und darunter findest du sie auch auf der Website.
- Start position (Open vs. Closed): Diese Spalte definiert NICHT in welchen Positionen der Move möglich ist, sondern wie Cues je nach Startposition gezählt werden. Meistens spielt die Startposition keine Rolle, in seltenen Fällen gibt es jedoch Unterschiede. Zum Beispiel haben Swing Outs aus der offenen Position einen Cue mehr, da sie das Heranbringen des Followers miteinschließen. Gleichzeitig wird dieses Feld auch manchmal für Besonderheiten (fett gedruckt) verwendet, wie zum Beispiel bei Promenade oder Circle. Dort wird 1 Move in 2 Zeilen aufgeteilt, um den Looping-Teil (verlässt nicht die Position) vom Non-looping Teil abzutrennen (für reine Analyse-Zwecke). Bitte lies dir in solchen Fällen immer die Notizen durch.
- Standard cue count: Diese Spalte beschreibt, wie viele digitale Cues für den Move gezählt werden. Häufig steht dort einfach eine Zahl, bei komplexeren Moves mit mehr Variation kann es jedoch vorkommen, dass dort eine allgemeine Regel steht.
- Standard move count: Diese Spalte beschreibt, wie viele Moves normalerweise gezählt werden. In den allermeisten Fällen ist die Zahl 1. In manchen Zeilen wird jedoch etwas anderes stehen. Bitte pass bei solchen Moves besonders auf.
- Count method for actual DLCs: Wenn die digitalen Cues für einen Move beschnitten werden, steht hier die Regel, wie die tatsächlichen DLCs gezählt werden.
- Standard looping/Not looping: Hier steht, ob ein Move standardmäßig als Looping oder Not looping gesehen wird.
- Notes: Hier stehen etwaige Notizen, die du kennen solltest.

Das war alles: Viel Erfolg!

Appendix B

The Extensive Data of the Linear Mixed Models

Ergebnisse

Mixed Model

Model Info

Info		
Model Type	Mixed Model	Linear Mixed model for continuous y
Model	lmer	cuesPerPhrase ~ 1 + ConditionText + (1 | Dyad)
Distribution	Gaussian	Normal distribution of residuals
Direction	y	Dependend variable scores
Optimizer	bobyqa	
DF method	Satterthwaite	
Sample size	288	
Converged	yes	
Y transform	none	
C.I. method	Wald	

[3]

Model Results

Model Fit

Type	R^2	df	LRT X^2	p
Conditional	0.266	4	57.288	<.001
Marginal	0.106	3	38.887	<.001

[4]

Fixed Effects Omnibus Tests

	F	df	df (res)	p
ConditionText	13.8	3	261	<.001

Parameter Estimates (Fixed coefficients)

				95% Confidence Intervals				
Names	Effect	Estimate	SE	Lower	Upper	df	t	p
(Intercept)	(Intercept)	6.8258	0.123	6.583	7.068	23.0	55.408	<.001
ConditionText1	Full - Background	0.8264	0.183	0.466	1.187	261.0	4.514	<.001
ConditionText2	Rhythm - Background	0.7963	0.183	0.436	1.157	261.0	4.350	<.001
ConditionText3	Solo - Background	-0.0417	0.183	-0.402	0.319	261.0	-0.228	0.820

[5]

Random Components

Groups	Name	Variance	SD	ICC
Dyad	(Intercept)	0.264	0.514	0.179
Residual		1.206	1.098	

Anmerkung. Number of Obs: 288 , Number of groups: Dyad 24

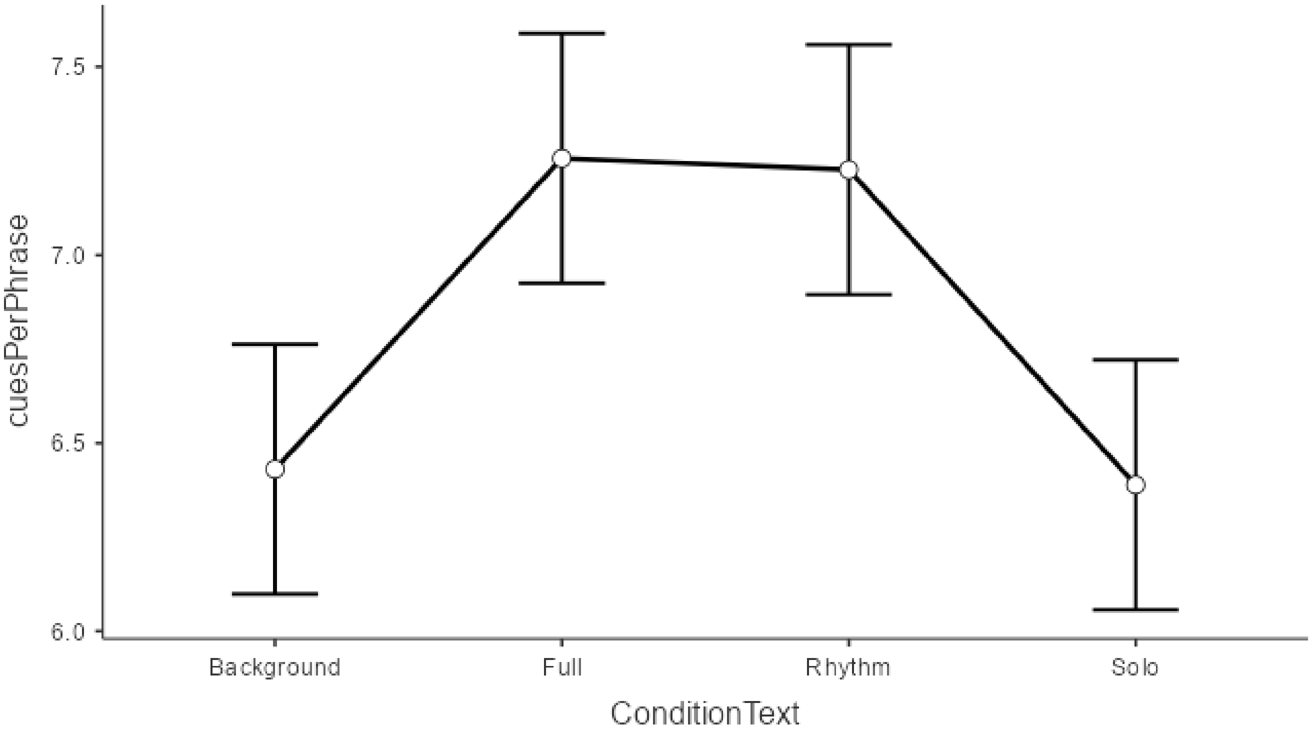
Post Hoc Tests

Post Hoc comparison: ConditionText

Comparison							
ConditionText	**vs**	**ConditionText**	**Difference**	**SE**	**t**	**df**	**$p_{bonferroni}$**
Background	-	Full	-0.8264	0.183	-4.514	261	<.001
Background	-	Rhythm	-0.7963	0.183	-4.350	261	<.001
Background	-	Solo	0.0417	0.183	0.228	261	1.000
Full	-	Rhythm	0.0301	0.183	0.164	261	1.000
Full	-	Solo	0.8681	0.183	4.742	261	<.001
Rhythm	-	Solo	0.8380	0.183	4.578	261	<.001

Results Plots

Plot: cuesPerPhrase ~ ConditionText



Linear Mixed Model

Model Info

Info		
Model Type	Mixed Model	Linear Mixed model for continuous y
Model	lmer	cuesPerPhrase ~ 1 + TempoVariation + (1 | Dyad)
Distribution	Gaussian	Normal distribution of residuals
Direction	y	Dependend variable scores
Optimizer	bobyqa	
DF method	Satterthwaite	
Sample size	288	
Converged	yes	
Y transform	none	
C.I. method	Wald	

Anmerkung. All covariates are centered to the mean

[3]

Model Results

Model Fit

Type	R^2	df	LRT X^2	p
Conditional	0.231	2	44.160	<.001
Marginal	0.072	1	25.759	<.001

[4]

Fixed Effects Omnibus Tests

	F	df	df (res)	p
TempoVariation	27.0	1	263	<.001

Parameter Estimates (Fixed coefficients)

			95% Confidence Intervals				
Names	**Estimate**	**SE**	**Lower**	**Upper**	**df**	**t**	**p**
(Intercept)	6.8258	0.12319	6.5833	7.0683	23.0	55.41	<.001
TempoVariation	-0.0316	0.00609	-0.0436	-0.0196	263.0	-5.19	<.001

[5]

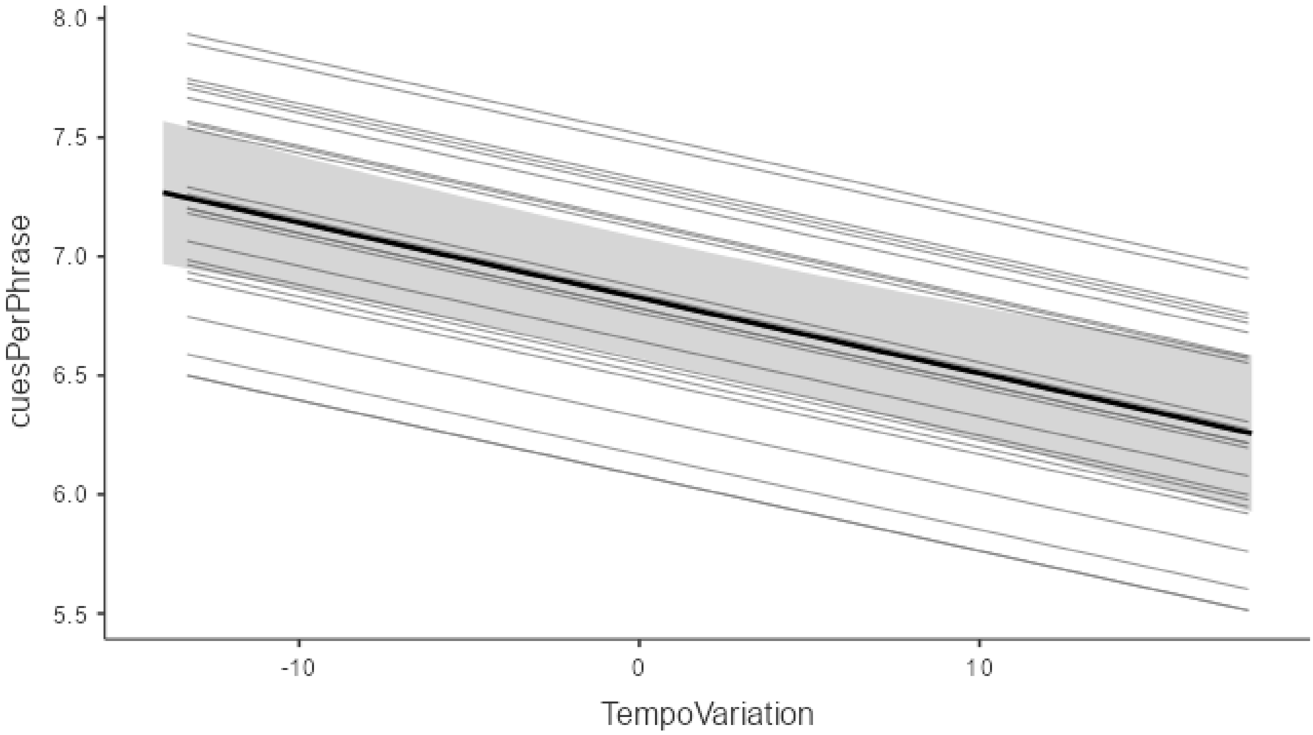
Random Components

Groups	Name	Variance	SD	ICC
Dyad	(Intercept)	0.259	0.509	0.171
Residual		1.258	1.122	

Anmerkung. Number of Obs: 288 , Number of groups: Dyad 24

Results Plots

Plot: cuesPerPhrase ~ TempoVariation



In Plot: cuesPerPhrase ~ TempoVariation random effects are plotted across Dyad

Linear Mixed Model

Model Info

Info		
Model Type	Mixed Model	Linear Mixed model for continuous y
Model	lmer	movesPerPhrase ~ 1 + ConditionText + Song + (1 | Dyad)
Distribution	Gaussian	Normal distribution of residuals
Direction	y	Dependend variable scores
Optimizer	bobyqa	
DF method	Satterthwaite	
Sample size	288	
Converged	yes	
Y transform	none	
C.I. method	Wald	

[3]

Model Results

Model Fit

Type	R^2	df	LRT X^2	p
Conditional	0.322	6	81.248	<.001
Marginal	0.175	5	66.563	<.001

[4]

Fixed Effects Omnibus Tests

	F	df	df (res)	p
ConditionText	8.85	3	259	<.001
Song	23.85	2	259	<.001

Parameter Estimates (Fixed coefficients)

				95% Confidence Intervals				
Names	Effect	Estimate	SE	Lower	Upper	df	t	p
(Intercept)	(Intercept)	4.841	0.0715	4.7006	4.9822	23.0	67.68	<.001
ConditionText1	Full - Background	0.333	0.1067	0.1234	0.5433	259.0	3.13	0.002
ConditionText2	Rhythm - Background	0.153	0.1067	-0.0572	0.3627	259.0	1.43	0.153
ConditionText3	Solo - Background	-0.194	0.1067	-0.4044	0.0155	259.0	-1.82	0.069
Song1	Doodlin - A Smooth One	0.608	0.0924	0.4258	0.7895	259.0	6.58	<.001
Song2	Tatoe Pie - A Smooth One	0.135	0.0924	-0.0464	0.3172	259.0	1.47	0.144

[5]

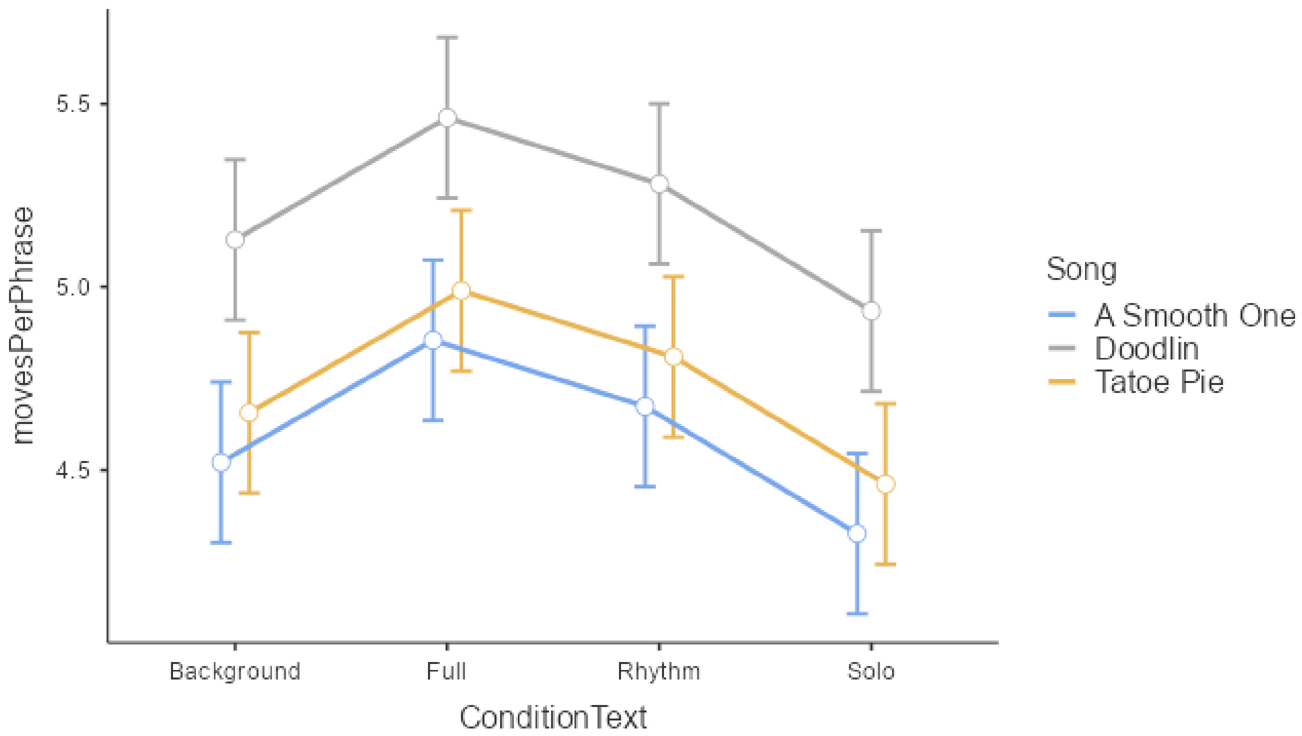
Random Components

Groups	Name	Variance	SD	ICC
Dyad	(Intercept)	0.0887	0.298	0.178
Residual		0.4095	0.640	

Anmerkung. Number of Obs: 288 , Number of groups: Dyad 24

Results Plots

Plot: movesPerPhrase ~ ConditionText ✻ Song



Linear Mixed Model

Model Info

Info		
Model Type	Mixed Model	Linear Mixed model for continuous y
Model	lmer	richness ~ 1 + TempoVariation + (1 | Dyad)
Distribution	Gaussian	Normal distribution of residuals
Direction	y	Dependend variable scores
Optimizer	bobyqa	
DF method	Satterthwaite	
Sample size	288	
Converged	yes	
Y transform	none	
C.I. method	Wald	

Anmerkung. All covariates are centered to the mean

[3]

Model Results

Model Fit

Type	R^2	df	LRT X^2	p
Conditional	0.054	2	7.588	0.023
Marginal	0.022	1	6.757	0.009

[4]

Fixed Effects Omnibus Tests

	F	df	df (res)	p
TempoVariation	6.81	1	263	0.010

Parameter Estimates (Fixed coefficients)

			95% Confidence Intervals				
Names	Estimate	SE	Lower	Upper	df	t	p
(Intercept)	0.62491	0.00890	0.607	0.64243	23.0	70.21	<.001
TempoVariation	0.00181	6.95e-4	4.46e-4	0.00318	263.0	2.61	0.010

[5]

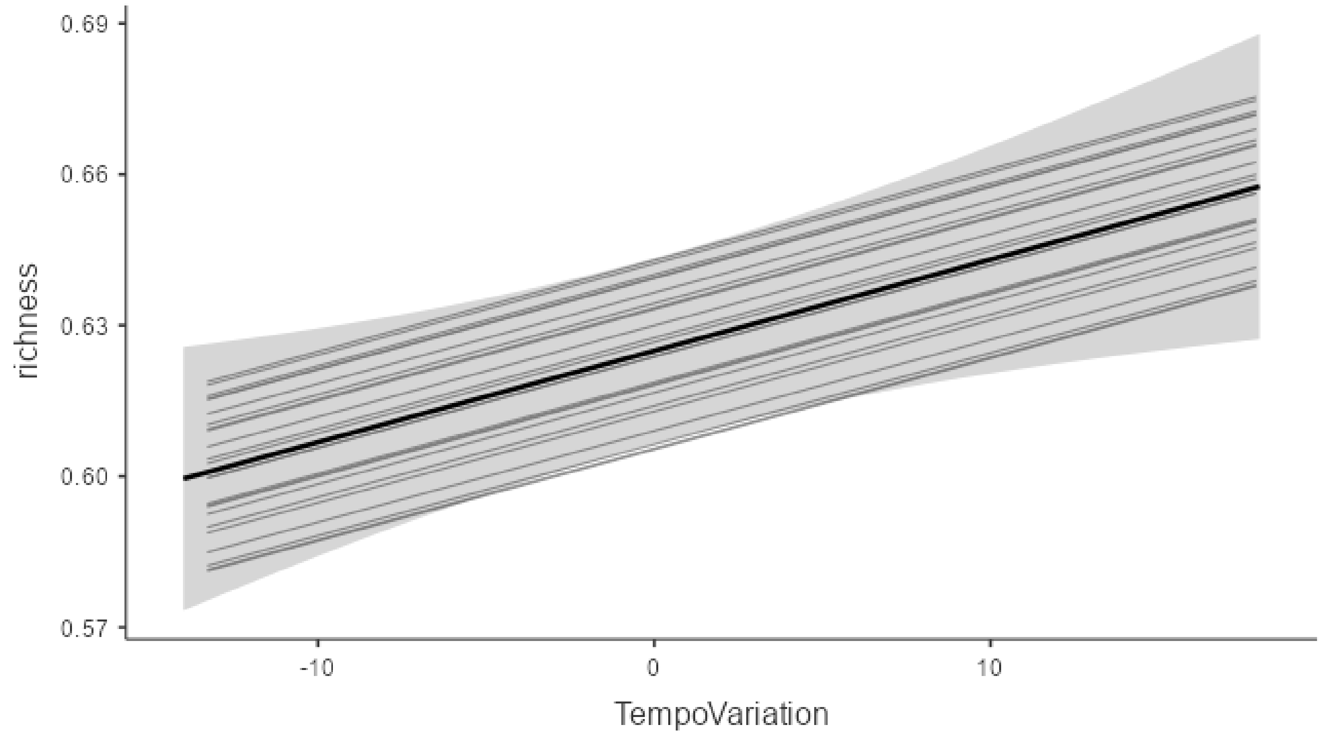
Random Components

Groups	Name	Variance	SD	ICC
Dyad	(Intercept)	5.38e-4	0.0232	0.0318
Residual		0.0164	0.1279	

Anmerkung. Number of Obs: 288 , Number of groups: Dyad 24

Results Plots

Plot: richness ~ TempoVariation



In Plot: richness ~ TempoVariation random effects are plotted across Dyad

Linear Mixed Model

Model Info

Info		
Model Type	Mixed Model	Linear Mixed model for continuous y
Model	lmer	aDLCsPerPhrase ~ 1 + ConditionText + (1 | Dyad)
Distribution	Gaussian	Normal distribution of residuals
Direction	y	Dependend variable scores
Optimizer	bobyqa	
DF method	Satterthwaite	
Sample size	288	
Converged	yes	
Y transform	none	
C.I. method	Wald	

[3]

Model Results

Model Fit

Type	R^2	df	LRT X^2	p
Conditional	0.243	4	52.259	<.001
Marginal	0.109	3	39.001	<.001

[4]

Fixed Effects Omnibus Tests

	F	df	df (res)	p
ConditionText	13.8	3	261	<.001

Parameter Estimates (Fixed coefficients)

				95% Confidence Intervals				
Names	Effect	Estimate	SE	Lower	Upper	df	t	p
(Intercept)	(Intercept)	7.148	0.131	6.890	7.405	23.0	54.569	<.001
ConditionText1	Full - Background	0.910	0.210	0.497	1.322	261.0	4.341	<.001
ConditionText2	Rhythm - Background	0.838	0.210	0.425	1.250	261.0	3.999	<.001
ConditionText3	Solo - Background	-0.148	0.210	-0.561	0.264	261.0	-0.707	0.480

[5]

Random Components

Groups	Name	Variance	SD	ICC
Dyad	(Intercept)	0.280	0.529	0.150
Residual		1.581	1.257	

Anmerkung. Number of Obs: 288 , Number of groups: Dyad 24

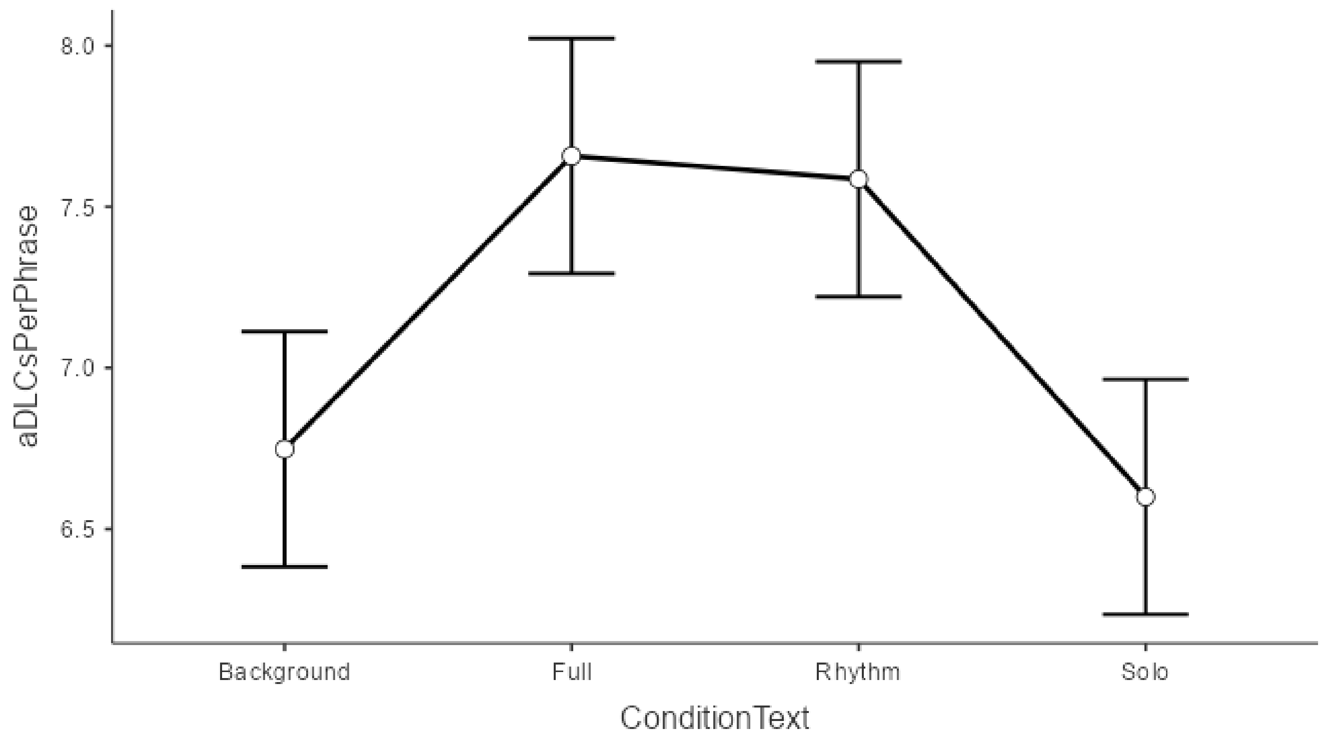
Post Hoc Tests

Post Hoc comparison: ConditionText

Comparison							
ConditionText	**vs**	**ConditionText**	**Difference**	**SE**	**t**	**df**	$p_{bonferroni}$
Background	-	Full	-0.9097	0.210	-4.341	261	<.001
Background	-	Rhythm	-0.8380	0.210	-3.999	261	<.001
Background	-	Solo	0.1481	0.210	0.707	261	1.000
Full	-	Rhythm	0.0718	0.210	0.342	261	1.000
Full	-	Solo	1.0579	0.210	5.048	261	<.001
Rhythm	-	Solo	0.9861	0.210	4.706	261	<.001

Results Plots

Plot: aDLCsPerPhrase ~ ConditionText



Linear Mixed Model

Model Info

Info		
Model Type	Mixed Model	Linear Mixed model for continuous y
Model	lmer	aDLCsPerPhrase ~ 1 + TempoVariation + (1 | Dyad)
Distribution	Gaussian	Normal distribution of residuals
Direction	y	Dependend variable scores
Optimizer	bobyqa	
DF method	Satterthwaite	
Sample size	288	
Converged	yes	
Y transform	none	
C.I. method	Wald	

Anmerkung. All covariates are centered to the mean

[3]

Model Results

Model Fit

Type	R^2	df	LRT X^2	p
Conditional	0.214	2	41.580	<.001
Marginal	0.082	1	28.322	<.001

[4]

Fixed Effects Omnibus Tests

	F	df	df (res)	p
TempoVariation	29.8	1	263	<.001

Parameter Estimates (Fixed coefficients)

			95% Confidence Intervals				
Names	Estimate	SE	Lower	Upper	df	t	p
(Intercept)	7.1476	0.13098	6.8898	7.4054	23.0	54.57	<.001
TempoVariation	-0.0379	0.00694	-0.0515	-0.0242	263.0	-5.46	<.001

[5]

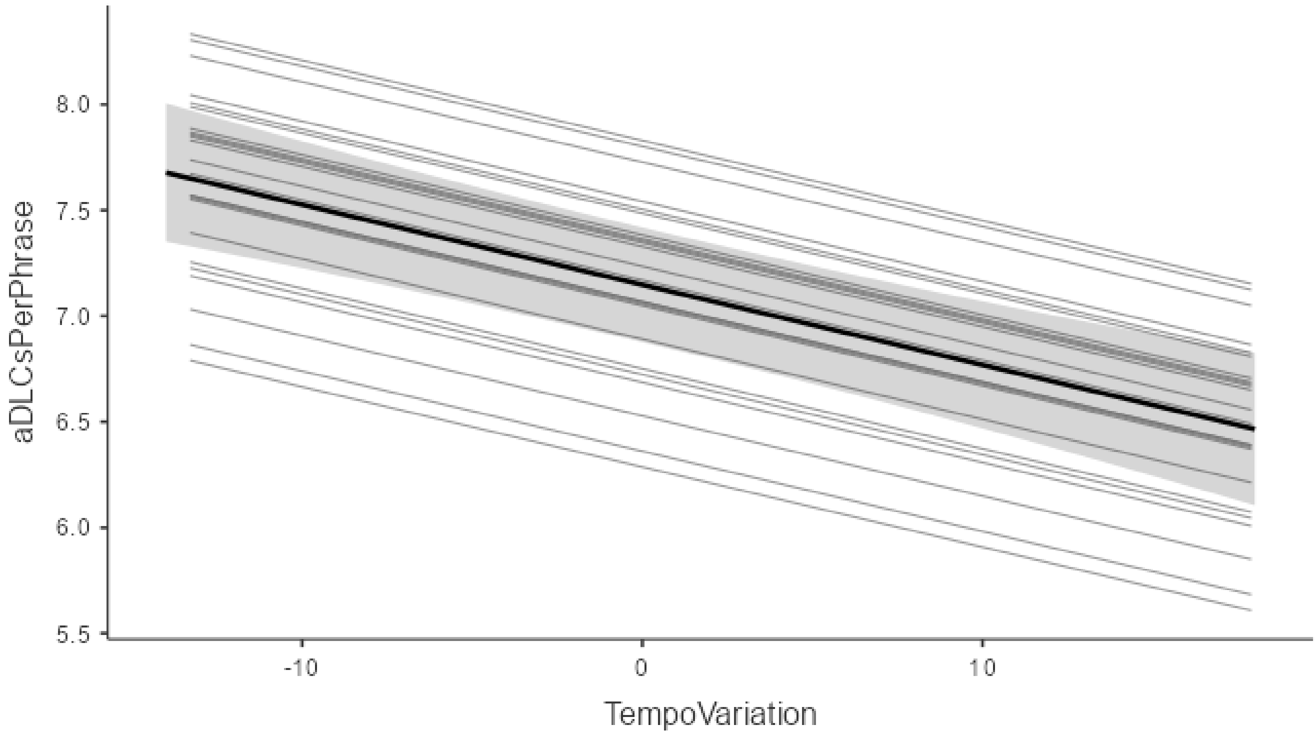
Random Components

Groups	Name	Variance	SD	ICC
Dyad	(Intercept)	0.276	0.525	0.144
Residual		1.634	1.278	

Anmerkung. Number of Obs: 288 , Number of groups: Dyad 24

Results Plots

Plot: aDLCsPerPhrase ~ TempoVariation



In Plot: aDLCsPerPhrase ~ TempoVariation random effects are plotted across Dyad

Referenzen

[1] The jamovi project (2024). *jamovi*. (Version 2.6) [Computer Software]. Retrieved from https://www.jamovi.org.

[2] R Core Team (2024). *R: A Language and environment for statistical computing*. (Version 4.4) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from CRAN snapshot 2024-08-07).

[3] Gallucci, M. (2019). *GAMLj: General analyses for linear models*. [jamovi module]. Retrieved from https://gamlj.github.io/.

[4] Gallucci, M. (2020). *Model goodness of fit in GAMLj*. . link.

[5] Lüdecke, Ben-Shachar, Patil & Makowski (2020). *Extracting, Computing and Exploring the Parameters of Statistical Models using R*. CRAN. link.