UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Quantifying Emergent, Dynamic Tonal Coordination in Collaborative MusicalImprovisation

Permalink
https://escholarship.org/uc/item/3mb046fk

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 42(0)

Authors
Setzler, Matt
Goldstone, Rob

Publication Date
2020

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Quantifying Emergent, Dynamic Tonal Coordination in Collaborative Musical Improvisation

Matt Setzler (msetzler@indiana.edu)
Cognitive Science Program, Indiana University
1900 E. 10th St. Bloomington, IN 47406 USA

Rob Goldstone (rgoldsto@indiana.edu)
Department of Psychological and Brain Sciences, Indiana University
1101 E 10th St. Bloomington, IN 47405 USA

Abstract

Groups of interacting individuals often coordinate in service of abstract goals, such as the alignment of mental representations in conversation, or the generation of new ideas in group brainstorming sessions. What are the mechanisms and dynamics of abstract coordination? This study examines coordination in a sophisticated paragon domain: collaboratively improvising jazz musicians. Remarkably, freely improvising jazz ensembles collectively produce coherent tonal structure (i.e. melody and harmony) in real time performance without previously established harmonic forms. We investigate how tonal structure emerges out of interacting musicians, and how this structure is constrained by underlying patterns of coordination. Dyads of professional jazz pianists were recorded improvising in two conditions of interaction: a 'coupled' condition in which they could mutually adapt to one another, and an 'overdubbed' condition which precluded mutual adaptation. Using a computational model of musical tonality, we show that this manipulation effected the directed flow of tonal information amongst pianists, who could mutually adapt to one another's notes in coupled trials, but not in overdubbed trials. Consequently, musicians were better able to harmonize with one another in coupled trials, and this ability increased throughout the course of improvised performance. We present these results and discuss their implications for music technology and joint action research more generally.

Keywords: joint action; time series modeling; musical improvisation; tonal consonance

Introduction

Groups of interacting individuals often coordinate in service of abstract goals, such as the alignment of mental representations in conversation, or the generation of new ideas in group brainstorming sessions (Garrod & Pickering, 2009; Paulus, Levine, Brown, Minai, & Doboli, 2010). While some efforts have been made to formalize and experimentally study collective behavior in such abstract spaces (Goldstone & Gureckis, 2009; Frey & Goldstone, 2016; Paulus et al., 2010), most work on human interaction has focused on sensorimotor coupling (e.g. entrainment, synchronization, mimicry) (Shockley, Richardson, & Dale, 2009; Richardson, Dale, & Kirkham, 2007; Richardson & Dale, 2005), so the individual mechanisms and group-level dynamics of high-level coordination remain poorly understood. The current study addresses this gap by examining sophisticated musical coordination in improvising jazz ensembles. Remarkably, freely improvising jazz musicians collectively produce coherent tonal structure (i.e. melody and harmony) in real time performance, without previously established harmonic forms or key signatures. Advances in music information retrieval have made it possible to formalize and quantify such tonal musical structure (Chew et al., 2014). Improvised jazz thus offers a remarkably sophisticated, yet computationally tractable paragon domain to study the basic properties and limits of our ability to coordinate and collectively produce high-level, abstract information.

Musicians respond and adapt to one another when they play together. These interactions are mediated by organizational structures which can vary depending on genre, performance/recording context and personnel. For example, orchestras are hierarchically organized with prescribed leader-follower roles fixed throughout a performance, whereas free improvising jazz ensembles are typically more characterized by feedback loops of mutual influence (D'Ausilio et al., 2012; Borgo, 2005). Ensemble performance research has shown that these underlying patterns of coordination are reflected in synchrony and entrainment of ensemble members (Keller, 2014; Rasch, 1979). For example, small temporal asynchronies in co-performer note onsets have been shown to reflect leader-follower roles and the degree to which musicians mutually adapt to one another (Keller & Appel, 2010; Goebl & Palmer, 2009; Demos, Carter, Wanderley, & Palmer, 2017), and postural sway couplings reflect leader-follower relations in ensembles (Chang, Livingstone, Bosnyak, & Trainor, 2017).

Improvised music is of particular interest, because the influence of coordination extends beyond sensorimotor coordination and into the music's formal architecture which is freely evolving over time in its rhythm, melody, harmony, and texture. These components of musical structure are fixed in composed music, but collectively generated in real time improvised performance in jazz ensembles. These structural features are thus constrained by and presumably reflective of the interactions and coordination patterns jazz musicians spontaneously engage in when they play together. Is this something we can quantify and empirically test? Developing musically relevant measures of improvised coordination can inform the development of artificial generative/interactive music systems (A. Roberts et al., 2019) and benefit music pedagogy by automating assessment of ensemble performance. More broadly, it offers an important extension to JA research, which as mentioned earlier has focused more on sensorimotor coupling and less on high-level, abstract coordination.

©2020 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

Measuring aspects of music that are structurally deep, nuanced and psychologically resonant is computationally tractable thanks to centuries of music theory and decades of Music Information Retrieval. Here we focus on *tonal consonance*, which refers to how different combinations of notes sound on a continuum from dissonant/unstable to consonant/stable. Tonal music is essentially a dynamic interplay between these extremes. Remarkably, freely improvising jazz musicians spontaneously generate coherent tonal structure without prior harmonic forms. We adapt a previously established tonal model – the Tonal Spiral Array (Chew et al., 2014; Chew, 2005; Herremans, Chew, et al., 2016) – to operationalize a measure of tonal consonance. Time series of these features are extracted from individual and collective musical streams from collaboratively improvising jazz musicians playing in experimentally manipulated conditions of interaction.

This study builds on previous empirical studies of coordination in collaboratively improvising musicians. In a notable example, Aucouturier and colleagues showed that experimentally manipulated social attitudes (e.g. dominant, caring) are sonically encoded and perceivable in the music produced by co-improvising musicians, and that these attitudes influence temporal and spectral coordination (Aucouturier & Canonne, 2017). Another study applied nonlinear time series analyses to the body movements and notes of interacting jazz musicians to show that improvised musical coordination is shaped by musical context (e.g. playing with a drone versus a swing backing track) (Walton et al., 2018). These findings lay an important foundation to the study of joint action in improvised music, but since the analyses did not incorporate music theory, the findings are limited to sensorimotor and temporal/spectral coordination, and do not extend to more sophisticated musical phenomena such as the emergence of tonal structure.

In this work we directly manipulate interaction of improvising musicians to examine how different underlying patterns of coordination constrain the exchange and emergence of tonal musical information. Dyads of professional jazz musicians freely improvised in two conditions of interaction: (1) a *coupled* condition, in which pianists improvised simultaneously and (2) a *one-way* condition, in which a single pianist improvised along with a recording of another pianist taken from a previous *coupled* performance. Musicians could mutually adapt to one another in *coupled* trials, but the *one-way* condition enforced an asymmetric causal influence (i.e. from recording to musician), as in the popular recording technique of 'overdubbing'. These conditions allowed us to isolate the effects of mutual coupling by contrasting music produced in two naturalistic musical settings. Tonal consonance time series were extracted from individual and merged musical streams (recorded as isolated MIDI[1] tracks for each musician) produced in each condition. We find that interaction condition systematically altered the coordinated musical behavior of dyads, who produced more consonant tonal structure, which evolved dynamically throughout improvised pieces, in *coupled* trials as compared to *one-way*. These results are presented and discussed in terms of their implications for music technology and joint action research more generally.

Methods

Participants

28 professional jazz pianists (25 male, 3 female) from the New York City music scene participated in this study. Participant age ranged from 21-37. On average participants had over 22 years experience playing piano (sd=5.2) and 15 years experience improvising (sd=4.6). All participants received formal training in piano performance and jazz studies at elite conservatories. None of our subjects had prior experience performing with one another.

Apparatus

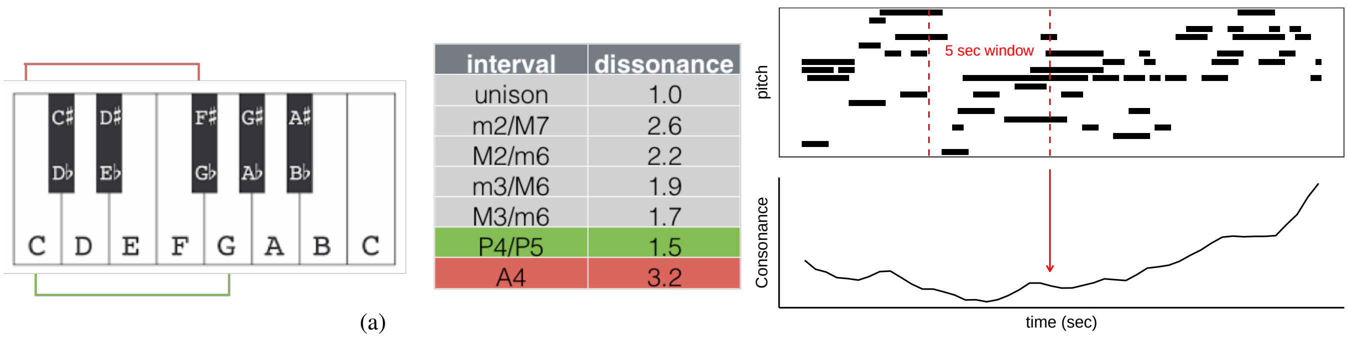
Two MIDI-enabled keyboards were used: a Roland Juno-Di and Nord Electro 2, both of which had 61 semi-weighted keys. Both keyboards were used on every trial (i.e. one-way trials were arranged such that the live pianist played whatever keyboard their ghost partner did not play). Ableton Live 9 Lite (running on a MacBook Air) was used to collect isolated MIDI recordings for each musician. Ableton was also used to synthesize the audio that participants heard, in the fashion of an electric Rhodes. This ensured time alignment of MIDI recordings, and that participants heard the same exact timbre for themselves and their partner irrespective of condition. Participants were recorded at a music studio in Brooklyn, NY. The studio was divided by a curtain such that participants could not see one another. Participants heard to themselves and their partners through Sony CH700N Noise Cancelling headphones. Thus, from the participants' perspective there was no visual or audible indication of their condition on a given trial.

Design and Procedure

Each musician played at least 3 trials in both coupled and one-way conditions. Pairs of participants entered the studio in separate sessions. Each participant played with the same 'live' partner for each of the coupled trials they played in and the same 'ghost' partner for each of their one-way trials. Conditions were interleaved within sessions and counterbalanced across sessions to control for possible order effects. Individual tracks from each *coupled* trial were used to yoke *one-way* trials in the following session, as depicted in Figure 2.

Participants were instructed to improvise a series of short (4-7 minute) duos. These improvisations were 'free', with no accompanying stimuli and no *a priori* musical template or constraints. Other than the suggested time frame, the only instruction musicians were given was to do their best to improvise a compelling piece of music, as they would in a typical

[1] Musical Instrument Digital Interface (MIDI) is a format for representing music on a computer. It symbolically represents the pitch, volume and timing (onset and offset) of musical note sequences.



(a)

(b)

Figure 1: Tonal consonance measure. (a) Every pitch interval is assigned a dissonance rating (perfect fifth and tritone are colored for illustration). Tonal consonance is the negative weighted sum of dissonance levels scaled by how often intervals occur within pitch sets. (b) Consonance time series were computed from music sequences using 5 second sliding window.

performance setting. Participants were told they would be improvising in one of the two conditions (*coupled* or *one-way*), but were not told which condition they were playing in on any given trial. After each trial, participants responded to questionnaires to indicate their subjective experience of the previous improvised performance in terms of: (1) how easy it was to coordinate with their partner (2) how well coordinated they were with their partner (3) quality of the improvised piece and (4) the degree to which they played a supporter or a leader role.

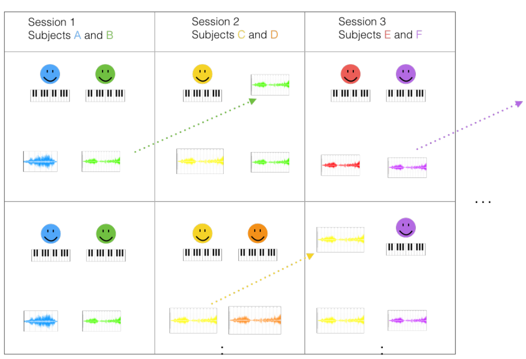


Figure 2: Yoked experimental design. Each participant played a series of *coupled* and *one-way* pieces with the same live and 'ghost' partner. Individual tracks from each *coupled* trial were used to yoke *one-way* trials in the following session.

Tonal Consonance Measures

Our measure of tonal consonance was adapted from a previous model of tonal structure called the Tonal Spiral Array, which has been validated against listener ratings as well as expert music theory analyses of musical tension (Chew et al., 2014; Herremans et al., 2016). The rationale behind the measure is that certain pairwise pitch intervals are inherently more or less consonant. For example, a tritone (e.g. {C,F#}) is a highly dissonant (i.e. low consonance) interval, whereas a perfect 5th (e.g. {C,G}) is a highly consonant (low dissonant) interval. Accordingly, every interval was assigned a *dissonance* rating, taken from past applications of the Spiral Array model, as indicated in Figure 1A (Herremans et al., 2016). Tonal consonance was then computed as the negative weighted sum of dissonance ratings scaled by how often intervals occurred within a pitch set (a constant of 2 was added to avoid negative values). Table1 shows model ratings for exemplar pitch sets.

Table 1: Consonance ratings of exemplar pitch sets.

Pitch Set	Consonance
{C,E,G} (Cmaj)	.65
{C,Eb,G} (Cmin)	.65
{C,B,G}	.54
{C,E,G,F,A,C} (Cmaj + Fmaj)	.49
{C,B}	.48
{C,E,G,F#,A#,C#} (Cmaj + F#maj)	.13
serial (all 12 pitches)	.09

Consonance time series were computed from music sequences using a 5 second sliding window[2] with 0.2 second hop size, as illustrated in Figure 1B. Three measures of consonance were considered: Individual Consonance (consonance of individual music streams), Combined Consonance (consonance of merged music streams from both players in a dyad) and Emergent Consonance (Combined Consonance minus average Individual Consonance of both musicians in a dyad, plus a constant of 1.8 to avoid negative values). *Emergent consonance* is essentially a measure of tonal coordination, as it captures the consonance arising from the *interaction* of pitches played by the two different musicians. A situation in which each pianist plays consonant notes that clash

[2] All reported analyses were also conducted with 2 and 10 second window sizes and yielded the same results.

with one another would result in low emergent consonance (e.g. {C,E,G} and {F#,A#,C#} are consonant on there own but {C,E,G,F#,A#,C#} is highly dissonant), whereas a situation in which each pianist plays dissonant pitch sets that stabilize one another when sounded together would result in high emergent consonance (e.g. {C,B} and {E,G} have low average consonance but {C,E,G,B} has high consonance because it is tonicized to a Cmaj7 chord).

Results

Directed Flow of Tonal Information

A novel lagged consonance analysis was used to quantify how musicians responded to and harmonized with one another's notes as a function of interaction condition. Lagged consonance was computed by taking individual note sequences of co-performers, shifting them relative to another, and computing Combined and Emergent Consonance time series of the merged pitch collections using a sliding window (5 second sliding window, 2 second hop size). This analysis facilitates the assessment of causal influence and directed flow of tonal information, as it quantifies the degree to which individuals harmonize with the preceding notes of their co-performer. For example, Player A might respond and harmonize with Player B's past notes but not the other way around, which would be reflected in high consonance for B-to-A lags but not for A-to-B lags.

Lagged consonance was computed for every trial in each condition with lags in the range of +/-20 seconds (spaced by increments of 2 seconds). In *one-way* trials, positive lags correspond to evaluating past notes of the ghost recording with future notes of the live musician (ghost-to-live) and vice versa for negative lags (live-to-ghost). The beginnings and endings of pieces (first and last 10%) were discarded to avoid boundary effects. Average Combined and Emergent Consonance were computed at every lag within each trial.

Figure 3 depicts average Emergent Consonance (EC) across the range of lags by condition. EC is essentially symmetric around 0 for *coupled* trials, but is significantly higher for ghost-to-live (positive) lags compared to live-to-ghost (negative) lags within *one-way* trials (paired $t(85)=2.3$, $p<0.05$; mean ghost-to-live compared to mean live-to-ghost EC within each *one-way* trial). This same asymmetry in *one-way* trials (but not *coupled*) trials was found with respect to Combined Consonance (paired $t(85)=2.5$, $p<.01$). These results reflect the underlying causal networks of each condition. Live musicians in *one-way* trials responded to ghost recordings by harmonizing with their past notes (resulting in high ghost-to-live lagged consonance) but ghost recordings could not respond to the notes of the live musicians (resulting in lower, presumably chance-level live-to-ghost lagged consonance). There was no such asymmetry in coupled trials, since both musicians could mutually respond to one another.

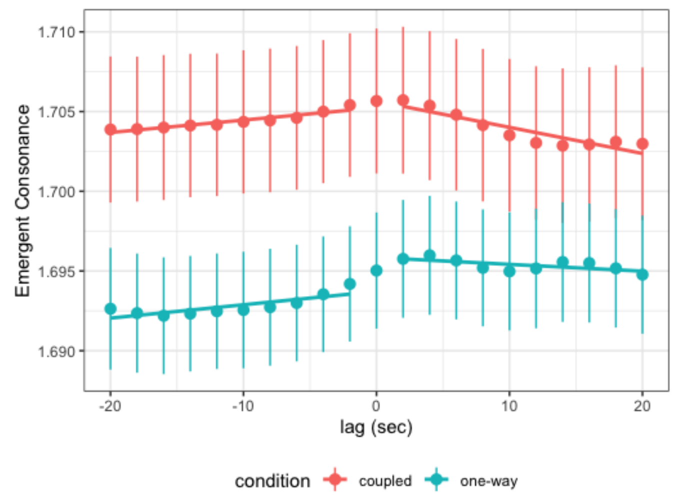


Figure 3: Musicians harmonize with preceding notes of their partner. Mean lagged Emergent Consonance (EC) across all pieces in each condition, error bars denote standard error. Within *one-way* trials, EC is higher at ghost-to-live (positive) lags versus live-to-ghost (negative) lags. EC is symmetric around zero, and higher overall in coupled trials.

Emergence of Group-Level Tonal Structure

How does interaction condition effect dyads' ability to *collectively* produce tonal structure at the group-level? Time series of lag-0 (simultaneous) Combined and Emergent Consonance were computed for every trial in each condition with 5 second sliding window and 0.2 second hop size. Overall there was more EC in *coupled* trials versus *one-way* trials (paired-$t(42)=2.21$, $p<0.05$, difference of means -0.011; paired t-test comparing mean consonance of each *coupled* trial versus that of the correspondingly yoked *one-way* pieces). Thus, mutually adaptive dyads exhibited higher tonal coordination compared to overdubed dyads, as they were better able to harmonize their notes with one another. There was no such effect in terms of *combined consonance*, which is perhaps unsurprising given that a dynamic range of tonal consonance (e.g. tension and release) is generally desired in tonal music.

Is there systematic structure to how Emergent Consonance evolves over the course of improvised pieces? If so, is this temporal structure modulated by interaction condition? To answer these questions we examined EC over normalized time, as the average EC in 50 equispaced time bins for each trial, as depicted in 4 (other binning schemes, not reported, yielded the same results). Hierarchical Bayesian modeling was used to model the time course of EC as a function of condition. A quadratic model of EC as a function of time was fit for every trial, and trial-level parameters were modeled as being drawn from condition-level distributions. This allowed us to isolate how the time course of consonance was modulated by condition by comparing means of condition-level distributions for each term.

This analysis revealed temporal structure that was common to trials in both conditions, but this structure was more exaggerated in *coupled* trials. In both conditions we found a positive linear term (posterior estimate of linear term for coupled trials was 6.4e-3 with 95% CI [4.6e-3,8.1e-3], with Effective Sample Size (ESS)=7890, Gelman-Rubin Statistic=1.00; estimate of average slope for one-way trials was 4.2e-3 with 95% CI [2.9e-3,5.5e-3], ESS=8208, Gelman-Rubin Statistic=1.00) and a negative quadratic term (posterior estimate of quadratic coefficient for coupled trials was -1.1e-4 with [-1.4e-4,-7.6e-5] 95% CI, ESS=9170, Gelma-Rubin Stat=1.00; estimate was -7.7e-5 for one-way trials with [-1.03e-4,-5.1e-5] 95% CI, ESS=8220, Gelman-Rubin Stat=1.00). The positive linear term indicates a general trend for EC to increase throughout an improvised performance, while the negative quadratic term reflects the extent to which a non-monotonic low-high-low structure for EC is found, which combines additively with the linear relation. Interestingly, the linear term was significantly more positive in *coupled* trials versus *one-way* trials (posterior estimate of *difference* between linear slopes i.e. *coupled* minus *one-way* is 2.2e-3 with [1.6e-06,4.3e-03] 95% CI i.e. non-overlapping with zero, ESS=8171, Gelman-Rubin Stat=1.00). These posterior distributions indicate a general tendency for tonal coordination to increase throughout performances, and this increase happens significantly moreso in *coupled* trials.

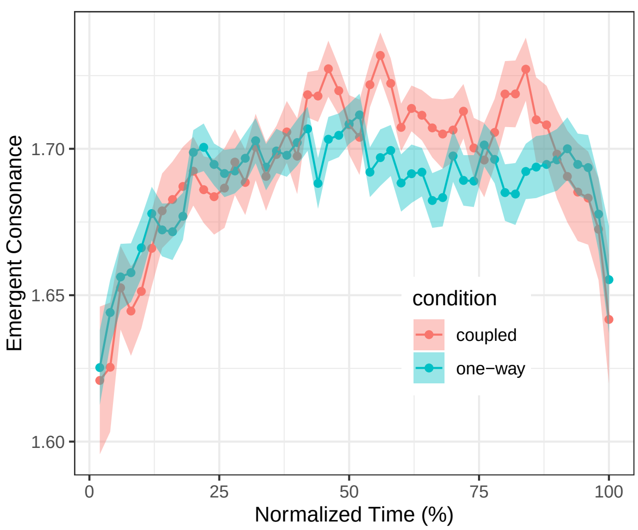


Figure 4: Mean Emergent Consonance (EC) over normalized time across every trial in each condition. Error ribbons denote standard error. EC increases within trials, and this increase is exaggerated in *coupled* trials.

Discussion

This study examined how tonal structure (i.e. harmony) produced by collectively improvising jazz musicians is shaped by underlying patterns of coordination. We recorded dyads of professional jazz pianists improvising in two conditions of interaction: one in which pianists improvised together simultaneously (*coupled*), and an ‘overdubbed’ condition which precluded mutual adaptation (*one-way*). These interaction conditions constrained the directed flow of musical information, which systematically altered musicians’ ability to harmonize their notes with one another to collectively produce sophisticated tonal structure.

We conducted a novel lagged consonance analysis to quantify the directed flow of tonal information from one musician to another. This analysis revealed that musicians harmonize with the past notes of their partner, which occurs mutually in *coupled* trials, but asymmetrically in *one-way* trials because there is only one direction of causal influence (i.e. from recording to live musician) in the dyad. This finding builds on the results of previous studies which have demonstrated that causal influence and leadership roles in performing music ensembles are reflected in the postural sway, note onset asynchronies and temporal coordination of performers (Chang et al., 2017; Keller & Appel, 2010; Aucouturier & Canonne, 2017). In this case we see that causal influence is reflected in the exchange of abstract tonal information in freely improvising musicians.

Our lagged consonance analysis was inspired by the technique of Cross Recurrence Quantification Analysis (CRQA), which has been used extensively in joint action research to study the dynamics of dyadic interaction (Dale & Spivey, 2006; Louwerse, Dale, Bard, & Jeuniaux, 2012; Richardson et al., 2007; Richardson & Dale, 2005). In these applications, CRQA operates on similarity matrices quantifying behavioral matching of interacting individuals over time and across a range of time lags. CRQA has the flexibility to detect non-linear patterns of coordination which can evolve over time, unlike related time series analyses such as Granger causality which assume stationary time series (Marwan, Romano, Thiel, & Kurths, 2007). Our lagged consonance analysis is an extension of CRQA: instead of using a similarity measure to quantify behavioral matching, we used a measure of tonal consonance to the notes played by co-performers over time and across time lags. This allowed us to assess *complementary* coordination (as opposed to mimicry) with a functional, domain specific operalization. Complex, naturalistic human interaction often involves such complementary, functional coordination, and similar lagged analyses can be used with appropriate operalizations of domain-specific coordination in future joint action research.

At the group-level, interaction condition shaped the emergence and dynamics of tonal structure. Mutual coupling enhanced pianists’ tonal coordination, as they achieved greater Emergent Consonance (EC), which indicated an enhanced ability to harmonize their notes with one another. Additionally, EC systematically increased throughout improvised performances in both conditions, but significantly more so in *coupled* trials. This echoes previous findings that successful group coordination can emerge spontaneously, without explicit communication, as individuals interact and work together towards a common goal (M. E. Roberts & Goldstone, 2011). Interestingly, in the case of improvised music, the dynamics of how coordination evolves throughout real-time performance are reflected in the music generated, and may be central to our aesthetic experience of collaboratively improvised music as listeners.

In sum, this study demonstrates that high-level musical structure (i.e. tonality) produced by improvising jazz ensembles is shaped by underlying patterns of coordination. These findings have implications for the popular recording technique of 'overdubbing' – which may alter the music produced by musicians performing without mutual adaptation. Moreover, our measures of tonal coordination in expert improvisers can contribute to the design of generative and interactive AI music systems to make them more human-like and more musical (A. Roberts et al., 2019; Huang et al., 2019). This work also contributes to our understanding of JA more generally; with appropriate operationalizations, it is possible to quantitatively assess how abstract features of emergent group performance are constrained by and reflective of underlying, dynamic patterns of coordination.

References

Aucouturier, J.-J., & Canonne, C. (2017). Musical friends and foes: The social cognition of affiliation and control in improvised interactions. *Cognition*, *161*, 94–108.

Borgo, D. (2005). *Sync or swarm: Improvising music in a complex age*. A&C Black.

Chang, A., Livingstone, S. R., Bosnyak, D. J., & Trainor, L. J. (2017). Body sway reflects leadership in joint music performance. *Proceedings of the National Academy of Sciences*, 201617657.

Chew, E. (2005). Regards on two regards by messiaen: post-tonal music segmentation using pitch context distances in the spiral array. *Journal of New Music Research*, *34*(4), 341–354.

Chew, E., et al. (2014). Mathematical and computational modeling of tonality. *AMC*, *10*, 12.

Dale, R., & Spivey, M. J. (2006). Unraveling the dyad: Using recurrence analysis to explore patterns of syntactic coordination between children and caregivers in conversation. *Language Learning*, *56*(3), 391–430.

D'Ausilio, A., Badino, L., Li, Y., Tokay, S., Craighero, L., Canto, R., … Fadiga, L. (2012). Leadership in orchestra emerges from the causal relationships of movement kinematics. *PLoS one*, *7*(5), e35757.

Demos, A. P., Carter, D. J., Wanderley, M. M., & Palmer, C. (2017). The unresponsive partner: roles of social status, auditory feedback, and animacy in coordination of joint music performance. *Frontiers in psychology*, *8*, 149.

Frey, S., & Goldstone, R. L. (2016). Does successful small-scale coordination help or hinder coordination at larger scales? *Interaction Studies*, *17*(3), 371–389.

Garrod, S., & Pickering, M. J. (2009). Joint action, interactive alignment, and dialog. *Topics in Cognitive Science*, *1*(2), 292–304.

Goebl, W., & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception: An Interdisciplinary Journal*, *26*(5), 427–438.

Goldstone, R. L., & Gureckis, T. M. (2009). Collective behavior. *Topics in cognitive science*, *1*(3), 412–438.

Herremans, D., Chew, E., et al. (2016). Tension ribbons: Quantifying and visualising tonal tension.

Huang, C.-Z. A., Vaswani, A., Uszkoreit, J., Shazeer, N., Simon, I., Hawthorne, C., … Eck, D. (2019). Music transformer: Generating music with long-term structure.. Retrieved from `https://arxiv.org/abs/1809.04281`

Keller, P. E. (2014). Ensemble performance: Interpersonal alignment of musical expression. *Expressiveness in music performance: Empirical approaches across styles and cultures*, 260–282.

Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception: An Interdisciplinary Journal*, *28*(1), 27–46.

Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive science*, *36*(8), 1404–1426.

Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics reports*, *438*(5-6), 237–329.

Paulus, P. B., Levine, D. S., Brown, V., Minai, A. A., & Doboli, S. (2010). Modeling ideational creativity in groups: Connecting cognitive, neural, and computational approaches. *Small Group Research*, *41*(6), 688–724.

Rasch, R. A. (1979). Synchronization in performed ensemble music. *Acta Acustica united with Acustica*, *43*(2), 121–131.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive science*, *29*(6), 1045–1060.

Richardson, D. C., Dale, R., & Kirkham, N. Z. (2007). The art of conversation is coordination. *Psychological science*, *18*(5), 407–413.

Roberts, A., Huang, A., Hawthorne, C., Howcroft, J., Wexler, J., Hong, L., & Dinculescu, M. (2019). Approachable music composition with machine learning at scale. In *Proceedings of the 20th international society for music information retrieval conference (ismir).*

Roberts, M. E., & Goldstone, R. L. (2011). Adaptive group coordination and role differentiation. *PLoS One*, *6*(7).

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science*, *1*(2), 305–319.

Walton, A. E., Washburn, A., Langland-Hassan, P., Chemero, A., Kloos, H., & Richardson, M. J. (2018). Creating time: Social collaboration in music improvisation. *Topics in cognitive science*, *10*(1), 95–119.