Principles of action planning in music production: evidence from fMRI and EEG studies

DISSERTATION

zur Erlangung des akademischen Grades
Dr. rer. med.

an der Medizinischen Fakultät
der Universität Leipzig

eingereicht von
Roberta Bianco
geboren am 20. April 1986 in Novi Ligure, Italien

angefertigt am
Max-Planck-Institut für Kognitions- und Neurowissenschaften, Leipzig

Betreuer
Prof. Arno Villringer
Dr. Daniela Sammler

Beschluss über die Verleihung des Doktorgrades vom: 15.08.2017

To Prof. G. Santarelli and Paola Fre

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my amazing supervisors, Prof. Arno Villringer for his trust, and Dr. Daniela Sammler for her precious and thorough guidance through all the steps of my scientific and personal growth. I particularly wish to thank my trans-oceanic collaborators Prof. Peter Keller and Dr. Giacomo Novembre for their pioneer work on "the pianist's hand" and their energetic intellectual support. I am also very grateful to Prof. Dr. Angela Friederici for her thoughtful comments on the work gathered in the present thesis.

I sincerely acknowledge the essential support of Seung-Goo Kim for his last-minute methodological advices, of Sven Gutekunst, Joeran Lepsien and Frank Burkhardt for their technical support, of Heike Boethel, Ulrike Barth, Anke Kummer, Mandy Jochemko and the students of our lab for their help in recruitment and data acquisition (Florian Scharf, Jasper Poehls, Corinna Eidner, Elias Matin, Katharina Roscher). Testing more than 100 expert pianists would not have been possible without their help. To all the pianists who took part to my experiments I am also very grateful.

A very special thank goes to my beloved friend Bruno Pace that played all my stimuli and always enlightened me with his profound knowledge of music and life. I further thank my colleagues and friends of the Institute that accompanied me with intense and valuable discussions about science and beyond, Mauricio Martins, Tomas Goucha, Emiliano Zaccarella, Nele Helbernd, Andrea Reiter, Lorenz Deserno, Sofie Volk, Martin Panitz, Charlotte Grosse Wiesmann, Stan van der Burght, and those that left already Hame Park, Dimitrije Markovic and Jelle Bruineberg. I thank Bene and the Brass Banditen for the music we did together in Leipzig, and Jo and Ida for their daily unconditional support.

For anything else one can imagine I thank my best friends of ever, Gabri and my familiy.

Contents

Part I

Framework and Summary

Chapter 1

Principles of musical action planning

> We don't play piano with our
> fingers but with our mind.
>
> *Glenn Gould*

1.1 Theoretical background

Music is an activity that humans have practiced for at least 35,000 years (Conard, Malina, & Münzel, 2009)[1]. Across ages and cultures, humans display in fact natural perceptual musical abilities (Dalla Bella, Giguère, & Peretz, 2007; Trehub, 2003), whereas to create and produce music has often been regarded as a marvellous skill of exceptional individuals, at least in Western culture (Pressing, 1984). Musicians not only have deep abstract musical understanding, but they use this knowledge to inform and adapt their performance in a fluent and intelligent fashion (i.e., depending on their communicative intentions, the audience, the conductor or the co-performers) (Clarke, 2001). From jazz improvisation to rehearsed classical concert interpretations, musicians have to translate musical ideas into appropriately ordered musical elements, whose motor parameters need to be optimally specified for each single act to gain the necessary fluency. They may thus rely on their acquired knowledge of musical structures, e.g., harmonic, to mentally represent and order structurally related musical elements in a sequence. The acquisition of

---

[1] The citation style used in this chapter refers to the rules and conventions established by the American Psychological Association (APA). However, the following chapters' styles adhere to the specific rules of the journals where they were published.

this ability might be a key step during musical training, that allows the detachment from fixed sequential movements towards their creative combination in potentially infinite numbers of structurally coherent sequences. Therefore, as shown in Figure 1.1, at any time during the unfolding musical context, musicians can generate structure-based predictions about forthcoming (i) harmonic sound (Koelsch, 2005) and (ii) the action used to produce that sound, e.g., the next chord, the "what" (Sammler, Novembre, Koelsch, & Keller, 2013; Novembre & Keller, 2011). Then, lower-level (iii) movement parameters, e.g., which fingering to use, need to be specified for the execution of each single act (Verwey, Shea, & Wright, 2015).
This thesis conjoins a series of empirical studies that tackle a common research question from different angles and with different methodological approaches. This common question can be framed as: How does the musician's brain process abstract musical, auditory and motor information to create meaningful output and to meet specific performance-demands?

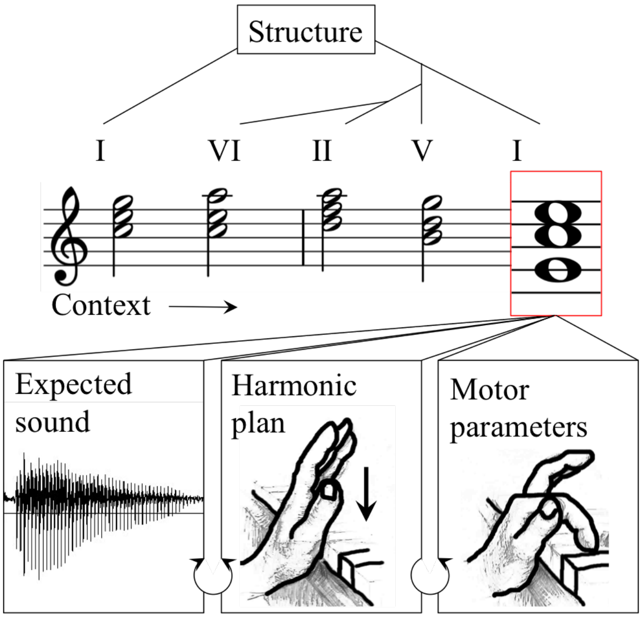


Figure 1.1: Features of musical actions scrutinised in this thesis: throughout the unfolding musical context and based on their long-term knowledge of musical structure, e.g., harmony, musicians generate structure-based predictions about the appropriate order of chords within a sequence, that determines motor (i.e., Harmonic plan) and auditory (i.e., Expected sound) expectations of which chord comes next, i.e., the "what". Motor parameters, i.e., the "how", are specified step by step at the level of single act implementation and should be neurally and cognitively distinguishable from the Harmonic plan.

.

Although important ground work on action control extensively accounts for the complexity inherent to simple movements such as clinking glasses (Wolpert, Diedrichsen, & Flanagan, 2011) or fixed sequences of movements (e.g., serial reaction time task) (Hikosaka, Nakamura, Sakai, & Nakahara, 2002; Penhune & Steele, 2012), still little is known about how the brain generates more complex actions that require flexible integration of single acts into larger coherent sequences that were not rehearsed beforehand (i.e., like in playing music, speaking or drawing) (Wulf & Shea, 2002; Hommel, 2016; MacKay, 1982). It is even less explored how the core processes of action planning may differ between differential abilities required for instance to play a concerto or to spontaneously improvise, to recite a poem or to chat freely on the phone. By restricting the discussion to expert pianists in the "Western tonal" tradition, the present thesis provides an empirical approach to investigate complex sequential action planning based on acquired structure-knowledge in the natural scenario of music production.

Before presenting the conceptual framework and the empirical data of this thesis, I will present existing theoretical and empirical work on the role of harmony in generation of auditory expectations and how these can be linked to generation of corresponding motor images through mechanisms of audio-motor coupling in musicians' brain (as illustrated in Figure 1.1). Further, I will present action control theories focusing on action organisation at multiple levels of processing and how these levels can pertain to musical actions (Figure 1.1) and possibly be shaped in musicians who master different music-genre (jazz or classical).

Knowledge of Harmonic structure. Structurally fundamental events upon which generalisation and predictive processes may operate are framed by theories of "musical syntax" and can include rhythmic, melodic and harmonic dimensions of music (Large & Palmer, 2002; Rohrmeier & Koelsch, 2012; Bharucha & Krumhansl, 1983). The studies of this thesis are all tailored to the investigation of harmonic structures as trigger for expectations of forthcoming harmonic chords in unfolding sequences. Therefore, for simplicity, the term "structure-based" will refer in this thesis to those cognitive processes that are based on contextual harmonic information and long-term knowledge of harmonic regularities. Harmony defines the arrangement of chords into well-formed sequences, so that a finite set of chords can be differently combined into larger and structurally coherent phrases (Swain, 1995; Lerdahl & Jackendoff, 1983). This coherence results in local and non-local dependencies between elements as defined by harmonic rules (Rohrmeier, 2011).

In a typical instance, a sequence starts and ends with a reference chord to which only some chords are expected to move to, whereas others rarely do. Psychologically, listeners who have been sufficiently exposed to the prevailing musical system (Krumhansl, 1983; Lerdahl & Jackendoff, 1983) predict and perceive these dependencies as tension-resolution patterns, in that they can cognitively link current auditory items to past events and generate predictions on forthcoming events (Patel, 2003; Tillmann, 2012). Therefore, harmonic structures may provide an ideal scenario to test cognitive processes like generalisation of structural knowledge to integrate discrete units, e.g., chords, into meaningful temporally evolving sequences, not only in perception (Koelsch & Siebel, 2005), but also in production (Sammler et al., 2013; Novembre & Keller, 2011). Music production of unrehearsed chord sequences that however follow musicians' general long-term knowledge of harmony will be the main framework of this thesis.

Expected sound and Action plan. A first important aspect of musical actions is that they are often linked with their expected sound, e.g., musicians form their movements in accordance with the desired sound, i.e., the auditory effect.

A number of recent findings have started outlining important evidence for the cognitive and neural principles of motor learning/control based on the intended auditory effect (Brown, Zatorre, & Penhune, 2015; Münte, Altenmüller, & Jäncke, 2002; Zatorre, Chen, & Penhune, 2007). One general mechanism thought to underlie these processes is a strong link between musical actions and their ensuing effects, the so called auditory-motor loop. Specifically, the representation of a certain effect can guide the action necessary to produce it, and the execution of a certain action generates forward predictions of the expected sensorial effect (Prinz, 2002; Hommel, Müsseler, Aschersleben, & Prinz, 2001; Wolpert et al., 2011). Within the general framework of predictive coding (Friston, 2010), this mechanism could explain how musicians anticipate future sensorial and motor events during performance. Behavioural and neuroimaging studies on simple or rehearsed musical patterns (note intervals or memorised melodies) have consistently shown auditory-motor coupling when musicians initiate (Keller & Koch, 2008), adjust their actions (Pfordresher, 2006; Zarate, Wood, & Zatorre, 2010), and predict forthcoming events during rehearsed performance (Maidhof, Vavatzanidis, Prinz, Rieger, & Koelsch, 2009). It has also been shown that this loop can be contingently decoupled to suppress unhelpful auditory information (e.g., imagery or feedback) in certain performance conditions (Finney & Palmer, 2003; Pfordresher, 2012; van der Steen, Molendijk, Altenmueller, & Furuya, 2014). Neurally, it has been suggested that audio-motor coupling may rely on general-purpose auditory pathways along the dorsal stream (Rauschecker & Scott, 2009; Brown et al., 2015) which serves sensorimotor transformation of sound to motor representations. However, beyond sound to action transformation, the dorsal stream may be involved also in other types of transforms across modalities, e.g., visual to motor (Gallivan & Culham, 2015). It is thus conceivable that a similar neural organisation along the dorsal stream may be relevant also for visually-driven music production (e.g., during score reading) in which, throughout the imitation of chord sequences, musicians can generate high-level predictions about forthcoming events based on their long-term knowledge of harmony and even in absence of sound (Novembre & Keller, 2011; Sammler et al., 2013).

The first goal of this thesis (see Chapter 2) was therefore to neurally explore whether motor performance of expert pianists can be informed by their knowledge of harmony regardless of auditory feedback, thus whether *action* and *auditory* streams of musical structure processing could be distinguished from each other and independently recruited.

Action hierarchy: Structure-based plan and Motor parameters. A second important aspect of musical actions is that musicians not only have to generate (i) sequences of movements in an appropriate order, e.g., according to harmonic structure, but also (ii) optimally select movement parameters of each single act to reach maximal fluency of performance and to convey a particular interpretation (Pressing, 1987; Chaffin & Logan, 2006; Repp, 2000).

According to theories of action control outside the music domain, complex sequential behaviours involve multilayer hierarchical control processes ranging from (i) the appropriate ordering of discrete acts into sequences to (ii) the specification of optimal movement parameters of each single act within the sequence. Despite consistent research on these two levels of action planning, they are often investigated in their own rights rather than in integrated scenarios of complex sequential actions. On the one hand, theories address the planning of *complex sequential actions* with emphasis on how their internal hierarchical structure determines the appropriate ordering of the constituent acts (Koechlin & Summerfield, 2007; Lashley, 1951; Uithol, van Rooij, Bekkering, & Haselager, 2012; Schmidt, 1975). The cognitive reality of such structure-based planning has been evidenced in speech (Dell, 1986) and music performance (Palmer & Pfordresher, 2003; Palmer & van de Sande, 1995) in that production errors are not arbitrary: for example, substitution errors are more likely to occur between two non-adjacent structurally related elements (e.g., nouns) than between two adjacent but unrelated elements (e.g., "writing a mother to my letter" instead of "writing a letter to my mother'). Another set of theories addresses the hierarchical implementation of *single acts* by discriminating between the planning of the goal (the "what") and the subsequent specification of movement parameters (the "how") (Grafton & Hamilton, 2007; Schmidt, 1975; Verwey et al., 2015). Evidence for this distinction comes from behavioural (for review see Wohlschläger, Gattis, & Bekkering, 2003; Rosenbaum, Cohen, Jax, Weiss, & van der Wel, 2007) and electrophysiological studies (Cattaneo et al., 2007; Fogassi et al., 2005) showing that high level action goals (e.g., to drink from a cup or to place it away) determine the choice of optimal movement parameters (e.g., the hand position to grasp the cup by the handle or the rim, respectively). Musical actions can be conceptualised as encompassing both these frameworks: At higher levels, structure-building processes determine the order of chords to play in the *action sequence*, hence the "what" of consecutive acts (Sammler et al., 2013); at lower levels, finger movement parameters, the "how", are specified upon execution of each *single act* of the sequence (Novembre & Keller, 2011).

The second goal of this thesis (see Chapter 3) was to zoom into the multilayer organisation of action planning to dissociate neural signatures of these two levels in the integrated scenario of complex musical actions.

Focus on Structure-based plan or Motor parameters. A third important aspect of musical actions is that they are not rigid entities and need to be variably organised to meet the demands of specific performance conditions. For example, playing a classical concerto or jazz improvisation makes different demands on the performer and requires focus on different levels of music production (Johnson-Laird, 2002).

Despite similar structural knowledge of Western tonal harmony (Johnson-Laird, 2002), jazz pianists are in fact required to focus more on their *structure-generative* abilities — to create harmonic sequences in real-time and tune "mistakes" into viable music (Pressing, 1987; Beaty, 2015), whereas classical pianists focus more on their *structure-interpretative* abilities — to rapidly decode (harmonic) structural cues of the given piece around which expressive strategies (e.g., touch, tempo, choice of fingerings) can then be optimised (Clarke, 2001). Therefore, the action control hierarchy should not be regarded as fixed, but the constituent layers may be modulated in performers with different action focus in their daily practice. Structure-generative focus may particularly train flexibility at the level of structure-based planning of the sequence in response to harmonic novelties — based on jazz musicians' rich knowledge of the range and probability of alternative events that might occur and to which they usually has to respond more or less on the fly; Structure-interpretative focus in turn may particularly boost sensitivity at the level of the movement parameter specification of single acts — based on classical musicians' intense practice of fingering techniques (Gellrich & Parncutt, 1998; Parncutt, 2014), and their ability to rapidly select a certain manner of execution to adequately emphasise structural elements of the piece.

The third goal of this thesis (see Chapter 4) was therefore to explore whether and how musicians adaptively tune these levels of action control to optimally master the specific demands of their genre, i.e., jazz or classical.

1.2 Towards a testable framework

In all studies presented here, pianists were tested with typical chord progressions of Western tonal harmony, in which both classical and jazz tradition are rooted (Johnson-Laird, 2002). To let them focus on the harmonic structure, sequences were kept constant in timing and only differed in tonalities and melodic contour. Importantly, pianists were not exposed to the sequences beforehand, so that they could not rely on retrieval of memorised chunks to process the sequences, but on the combinatorial power of their structural knowledge of music.

The crucial point that leads to our experimental approach is that harmony not only defines the sequence of sounds but also co-determines how single musical acts are sequentially ordered. While it is well established that acquired knowledge of harmony leads to auditory expectations of forthcoming chords (Koelsch & Siebel, 2005; Tillmann, 2012), only recent studies showed that it also benefits musicians' action planning during execution of chord progressions, even when auditory feedback is eliminated (Novembre & Keller, 2011; Sammler et al., 2013). The emergence of action planning based on harmonic knowledge in pianists obviously requires a minimum time of motoric exposure that was assumed to be about 10,000 hours of piano training (Levitin, 2006).

We reasoned that once such knowledge is acquired through intense training, auditory information (e.g., feedback or imagery) may not constitute a precondition for expert pianists to generate structure-based prediction in production. In fact, although there is evidence for audio-motor loop in the musicians' brain (Zatorre et al., 2007), it is also known that expert pianists may contingently decouple their movements from the auditory feedback (Finney & Palmer, 2003; Pfordresher, 2012; van der Steen et al., 2014). Thus, *auditory* and *motor* information about structural regularities may rely on systems that are coupled or decoupled depending on the performance demands. Crucially, to isolate the contribution of structure-based predictions in production from the auditory system, usually co-involved during real performance, pianists were asked to execute chord progressions in which auditory feed-back was eliminated (as in Novembre & Keller, 2011; Sammler et al., 2013). To isolate structure-based predictions in perception, in a separate task they listened to the same chord progressions without playing. In this way (see Chapter 2), we aimed to dissociate predictions based on pianists' knowledge of harmony in perception and production.

In a second step, we reasoned that once structure-based plans about the harmonic relations between discrete acts are available to the motor system, they should be distinguishable from lower-level processes, such as motor-parameter specification of single acts, in accordance with hierarchical planning models (Rosenbaum et al., 2007; Schmidt, 1975; Verwey et al., 2015; Vallacher & Wegner, 1987; Wohlschläger et al., 2003). To test the multilayer organisation of musical actions, we tested pianists with a production task that inclusively addressed two levels of action control hierarchy: (i) structure-based planning of the *action sequence* according to Western tonal harmony, and (ii) movement-parameter specification of *single acts* in terms of fingering applied to perform (see Chapter 3).

At last, we compared two groups of equally expert musicians that were specialised in different genres i.e., jazz or classical. We assumed that both groups had acquired structural knowledge of Western tonal harmony, in which both classical and jazz tradition are rooted, and reached proficient levels of fine motor control. However, jazz pianists are specialised to assemble musical elements into higher-order structures on the fly — *structure-generative* abilities (Pressing, 1987; Beaty, 2015), whereas classical pianists are specialised to implement a vast range of expressive strategies via projection of fine movements parameters (i.e., tempo, touch, choice of particular fingering) onto given musical structures — *structure-interpretative* abilities (Clarke, 2001; Shaffer, 1984). Therefore, on top of a similar hierarchical core structure of action planning — here investigated with regards to structure-based planing and motor parameter specification — specialised musicians may differentially boost one or the other action control process, inadvertently and despite similar instruction and task (see Chapter 4).

Figure 1.2 summarises the framework that guided the three experiments gathered in this thesis. Three hypotheses were investigated. First (Study I, Chapter 2), by means of task-based functional magnetic resonance imaging (fMRI) and functional connectivity at rest, we examined whether structure-based predictions in action can be dissociated from auditory information processing, usually coupled in musicians' brain. Second (Study II, Chapter 3), by means of electroencephalography (EEG), we zoomed into fine-grained aspects of action control to find evidence for dissociable high-level structure-based planning of the *action sequence* and low-level specification of fingering applied to each *single act*, in a integrated scenario of complex musical actions. Finally (Study III, Chapter 4), we explored with EEG whether and how musicians adaptively tune these levels of action control to optimally master the specific demands of their genre, i.e., jazz or classical.

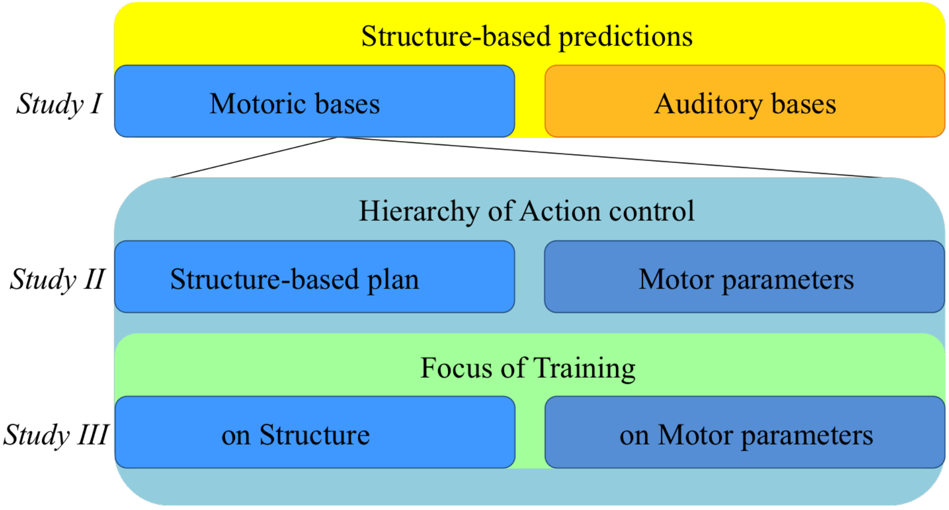


Figure 1.2: Framework adopted to dissect the complexity of musical action planning based on long-term knowledge of harmonic structures. The studies related to this framework are those reported in this thesis.

.

1.3 Summary of the empirical evidence

The three experiments constituting the empirical part of this thesis (see Part II) are presented in this section.

Study I (Bianco, et al., *in press*) presents one fMRI study that examined whether structure-based predictions can be motorically represented in musicians' brains regardless of auditory information and seeks evidence for dissociable contributions of motor and auditory processing streams to production without sound or listening without performing (see Chapter 2). To do so, we combined functional connectivity at rest with task-based functional magnetic resonance imaging (fMRI) while pianists imitated chord sequences without sound or while they listened to the chord sequences without acting.

This work was inspired by the grounded knowledge that tonal music triggers auditory predictive processes based on structural properties of harmony (Koelsch & Siebel, 2005; Tillmann, 2012; Patel, 2003), and by previous work showing initial behavioural and electrophysiological evidence for structure-based predictions while expert pianists executed silent chord sequences (Novembre & Keller, 2011; Sammler et al., 2013). Neural hypotheses on harmonic processing during listening (Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005; Maess, Koelsch, Gunter, & Friederici, 2001; Sammler, Koelsch, & Friederici, 2011) point to a crosstalk between frontal computational (i.e., the inferior frontal gyrus, IFG) and task-relevant posterior regions (i.e., posterior superior temporal gyrus, pSTG) to parse structural dependencies and predict forthcoming auditory information. However, the neural correlates of structure-based predictions in production, e.g., a crosstalk between frontal computational regions and action-relevant posterior regions, whether they involve auditory imagery-related areas and their comparison with the regions subserving analogous harmonic processing during listening remain unexplored.

To fill this gap, twenty-nine classical pianists were presented with unrehearsed harmonically congruent or incongruent chord progressions, either as actions (photos of a hand playing chords) that they were required to imitate without sound, or in an auditory format that they listened to without playing. In the task-based fMRI analysis, harmonically incongruent and congruent chords were contrasted separately for the action and audio tasks. Then, to dissociate action and audio networks for harmonic processing, activation peaks in the IFG clusters of the action and audio tasks were used as seeds for functional connectivity analysis on resting-state data.

This experimental approach provided evidence for dissociable dorsal motor and ventral auditory networks for sequential structure processing in music production and perception: we found distinct posterior-dorsal (Brodmann area: BA44/vBA6) and anterior ventral (BA44/45) sub-regions in right IFG that were respectively interconnected with superior parietal lobe (SPL, BA7) for *action* sequence processing and posterior superior temporal areas (pSTG, BA22) for *audio* sequence processing. The differential contribution of posterior motor- and auditory-related areas was interpreted as motoric and auditory store of long-term knowledge of harmonic regularities; the involvement of distinct dorsal and ventral right IFG sub-regions reflected sensitivity to both stimulus- or task-demands in parsing structural dependencies, in line with models of prefrontal cortex organisation (Fuster, 2001; Badre & D'Esposito, 2009; Friederici, 2011) and dual stream models of visual-spatial and auditory processing (Kravitz, Saleem, Baker, & Mishkin, 2011; Rauschecker, 2011). Within the predictive coding framework (Friston, 2010), these data suggest that internal structure-based predictions from frontal areas may progressively inform lower-level modality specific systems of knowledge along dorsal and ventral streams. Furthermore, these networks were recruited during processing of unrehearsed chord progressions, suggesting that structural knowledge provides the brain with a generalisable scaffold to generate and revise predictions about forthcoming events. Last but not least, the absence of auditory activation in the action task demonstrates that pianists flexibly decoupled motor from auditory images of the forthcoming chord in the sequence. Therefore, this study provides the first evidence that structural knowledge of music can be grounded in the visual-motor control system (i.e., frontal and parietal areas along the dorsal stream).

Study II (Bianco et al., 2016) was motivated by the question whether the neural signals elicited by the execution of harmonic violations were specific to structure processing and revision of harmonic predictions (Sammler et al., 2013) or whether they were general signatures of unusual motor patterns (see Chapter 3). Therefore, we zoomed into musicians' action planning to distinguish electrophysiological markers of high-level structure-based planning from lower-level specification of movement parameters during complex musical actions.

Previous research on silent music production (Novembre & Keller, 2011; Sammler et al., 2013) focused on high-level structure-based planning and showed that pianists — while imitating silent videos of a hand playing standard chord progressions — mentally construct the harmonic structure of the sequence based on their long-term knowledge of harmony and the continuously unfolding musical context (in line with music generative models by Palmer & Pfordresher, 2003; Clarke, 2001). Pianists planned ahead harmonically predictable chords (i.e., greater accuracy and faster response times), while they showed behavioural costs and a "reprogramming" negativity in event-related potentials (ERPs) during execution of incongruent chords that mismatched the anticipated harmonic plan. These effects were stronger when target chords were preceded by a long rather than short musical context, demonstrating their specificity to high-level plans built on the harmonic structure emerging from the unfolding context. However, two questions remained to be answered: (i) whether high levels of action pertaining to the appropriate ordering of chords are distinguishable from lower levels pertaining to the choice of fingers conventionally associated with standard progressions (Gellrich & Parncutt, 1998; Sloboda, Clarke, Parncutt, & Raekallio, 1998), and (ii) whether structural planning of the action sequence, which determines which chord to play ("what"), is prioritised over specification of movement parameters of single acts ("how"), in line with models of hierarchical action organisation (Grafton & Hamilton, 2007; Vallacher & Wegner, 1987; Wohlschläger et al., 2003; Uithol et al., 2012).

To answer these questions, twenty-six expert classical pianists were asked to execute chord progressions on a mute keyboard by copying a model hand shown in sequences of photos. We used photos, rather than videos (Novembre & Keller, 2011; Sammler et al., 2013) to obtain precise onset times at target chord presentation. We manipulated the harmonic structure of the chord sequences by rendering the final chord harmonically (in)congruent with the preceding musical context (congruent/incongruent Harmony). We further manipulated the predictability of the final chord by placing it at the end of 5- or 2-chord sequences (long/short Context). These manipulations were assumed to address high levels of structure-based planning, particularly in the long rather than in the short context that provides more information for structure building (interaction of Harmony × Context). To address low levels of single-act parameter specification, we manipulated the last chord not only in terms of which chord to play (congruent/incongruent Harmony), but also in terms of fingering used for the execution of the final chord (correct/incorrect Manner). The execution of the manner violation was assumed to yield distinct signals from the harmonic violation and to not depend on the structural information provided by the preceding context (main effect of Manner, but no interaction of Manner × Context). Behavioural results yielded strong context-dependent priming effects on the response times and key errors associated with the final chord (replicating Novembre & Keller, 2011; Sammler et al., 2013). The novel finding was that the manner of execution was not influenced by the context, unless the underlying chord was harmonically congruent, showing priority of higher-level structure-based plans that in turn can prime selection of finger parameters of single acts. Also, ERP results showed different electrophysiological signals related to execution of harmonic violations (late "reprogramming" negativity, as in Sammler et al., 2013) vs. manner violations (late positivity), distinguishing neural signatures at two levels of the hierarchical organisation of complex sequential actions. This study adds new insights to the notion of multilayer action control processes (Lashley, 1951; Rosenbaum et al., 2007; Schmidt, 1975; Verwey et al., 2015), in that it provides neural evidence for the dissociation between high-level structure-based planning of musical sequences (i.e., chords within a larger harmonic context) and subsequent lower-level parameter specification of single acts (i.e., fingering applied to perform the chord) in an integrated scenario of skilled sequential actions.

Study III (see Chapter 4) adopts the paradigm summarised above, and examined plasticity of the multilayer action control processes in musicians that experience different demands in their daily practice. Musicians not only have deep abstract knowledge of music, but in fact they use this knowledge to inform and flexibly adapt their performance depending on current or long-term demands. Differential abilities required by performance may be thus reflected at different levels of the action hierarchy.

Despite similar knowledge structure, i.e.,"Western tonal harmony", jazz and classical pianists experience different demands in their performance: Jazz more than classical musicians adopt a *structure-generative* focus when playing that may particularly boost high rather than low levels of musical action planning; in contrast, classical more than jazz musicians adopt a *structure-interpretative* focus when playing that may particularly train the musician's sensitivity to adjust low-level movement parameters of the action plan to given structural cues (Clarke, 2001).

Fifteen classical (with no improvisation experience) vs. fifteen jazz pianists (with a minimum of 2 years of improvisation training) — who were comparable in total accumulated hours of piano training across their life (ca. 11500 hours) — were tested with the same production task that (as in the Study II) inclusively addressed: (i) structure-based planning of the action sequence and (ii) parameter specification of single acts. Effects of musical training style on structure-based planning processes should surface as Group × Harmony × Context interactions. Conversely, pianists' differential sensitivity to movement parameter specification of single acts should be reflected by interactions of Group × Manner. Noteworthy, the execution of exactly the same task that was similarly unfamiliar to both groups (photos of an hand playing chords) ruled out the contribution of stimulus and task familiarity.

Combined behavioural, ERP and time frequency EEG measures, results showed indices of hierarchical action control at the levels of the sequence and the single act in both groups in line with (Bianco et al., 2016), but also crucial between-groups differences. Specifically, (i) at sequence level, classical pianists showed indices of cognitive effort to revise their structure-based plans in case of harmonically unexpected chords (context-influence on slowing of response times and on increase of right-frontal theta and decrease of late alpha power). Jazz pianists in turn showed indices of a greater readiness and flexibility to respond to harmonic violations: none of the effects listed for classical pianists and an earlier onset of the context-dependent "reprogramming" negativity than in classical pianists (Sammler et al., 2013; Bianco et al., 2016). This set of results suggests that the structure-generative tendency and readiness for change in jazz pianists may be grounded in the simultaneous pre-activation of several possible harmonic continuations, the most likely of which has right of way but can be rapidly cancelled and turned into one of the other alternatives (Cisek, 2006). Conversely, the greater cognitive effort in case of structural-harmonic deception in classical pianists may be tied to their structure-interpretative focus that requires to narrow down harmonic possibilities in order to rapidly proceed to expressive stages of action planning (the "how", see below).

At (ii) single act level, classical compared to jazz pianists showed in fact greater accuracy to set movement parameters in both contexts (in terms of fingering errors, power changes in theta and alpha band associated with visuo-motor translation processes), suggesting that classical pianists' expressive tendency may have enhanced their sensorimotor preparedness to rapidly perceive and react to observed erroneous finger movements (Candidi, Maria Sacheli, Mega, & Aglioti, 2014; Fagioli, Ferlazzo, & Hommel, 2007; Hommel, 2010). Moreover, classical pianists showed less fingering accuracy in the long than in the short context during imitation of unconventional manner on top of structurally congruent chords (replicating Bianco et al., 2016), suggesting that structure-interpretative focus may lead pianists to rapidly project optimal movement parameters as soon as the structure-based plan has emerged from the context.

Altogether, these findings indicate that classical and jazz pianists give different weights to the hierarchical levels of action planning, intuitively and despite identical instruction and material: While habitual action focus on structure generation in jazz may inadvertently bind resources to high levels of planning, even in a task that did not require creative improvisation, the building of solid associations between finger configurations and frequent structural figures (e.g., scales, intervals, cadences) in classical training (Clarke, Parncutt, Raekallio, & Sloboda, 1997; Gellrich & Parncutt, 1998; Parncutt, 2014) may spread focus from high to low levels to ease rapid selection of expressive features that adequately emphasise structural elements and shade interpretative performance. Therefore, this study demonstrates that the multilayer organisation of musical actions is not a rigid entity and can be variably tuned to meet the demands of specific performance conditions. Remarkably, it nourishes the idea that representation of musical structures benefits motor control in skilled sequential behaviours on top of which performers may project their preferences or tendencies developed in their practiced style, making exceptional performance even more exceptional.

1.4 General discussion

The present thesis has set up a framework for studying the neurocognitive bases of music production based on harmonic knowledge internalised by pianists through intense motor practice (> 10,000 hours of piano playing). We showed that harmony-based predictions can be motorically represented regardless of auditory feedback, they contribute to the action control hierarchy at the level of appropriate action sequencing and, together with lower levels of the hierarchy, they can be modulated by external factors like preferences or tendencies developed through practice. The main tenets are summarised in Figure 1.3 and listed below.

First, neuroanatomical and electrophysiological correlates of structure processing in action can be observed in classical musicians' brains, even in absence of auditory feedback: these correlates were obtained by contrasting neural responses evoked during execution of structurally incongruent minus congruent chord progressions and included a fronto-parietal network (right BA44 and BA7 along the dorsal stream in fMRI, Study I), a late "reprogramming" negativity (Study II and III) as well as conflict-related right-frontal theta power increase followed by right-central alpha suppression in EEG (Study III). These neural effects were linked with behavioural advantage/disadvantage during execution of harmonically congruent/incongruent chords, respectively. Thus, the execution of harmonically (in)congruent progressions in absence of sound provides insights into the neural bases of structure-based predictions and reassessment in musical sequential actions (Palmer & Pfordresher,

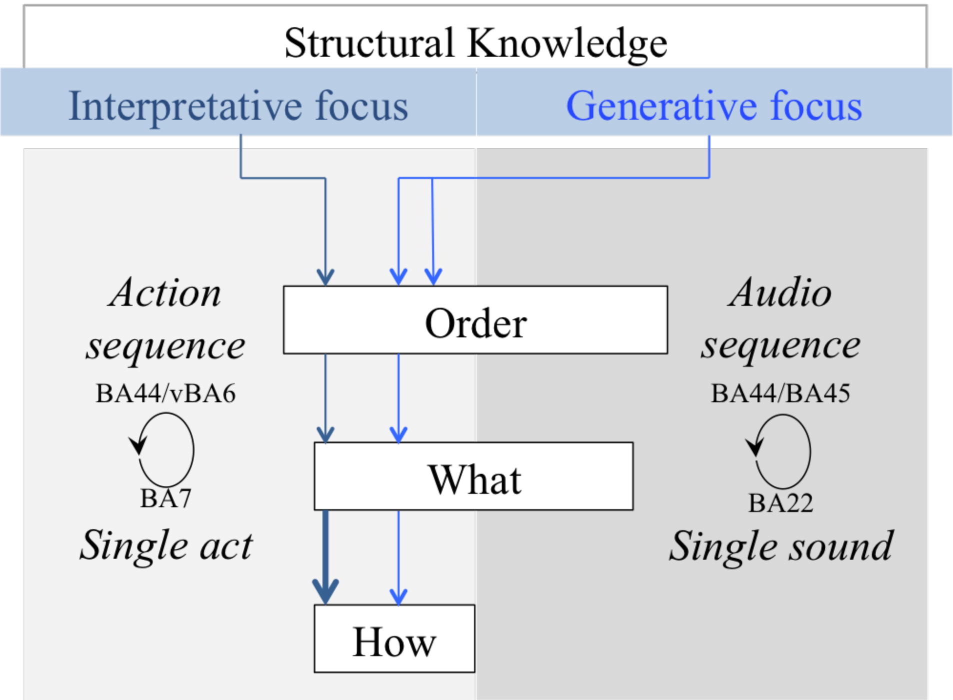


Figure 1.3: Processing of auditory and motoric information in musical actions based on structural knowledge (i.e., harmony), and influence of focus of practice on action planning. Based on their acquired structure-knowledge, expert pianists integrate single units (of sound or act) into higher-level harmonic structures defining the appropriate ordering, and consequently the harmonic identity ("what") of forthcoming chords. Dissociable fronto-parietal (involving BA44/vBA6 and BA7) and fronto-temporal (involving BA44/45 and BA22) networks might be the neural underpinnings of harmonic processing in action and perception, respectively. In production, single acts are implemented at lower-levels of action planning via specification of movement parameters (e.g., fingering, "how") at later stages of motor programming. Focus on structure-generative (jazz) vs. structure-interpretative (classical) abilities during practice influences how these levels (order, what and how) are preferentially used in complex sequential planning: jazz pianists bind resources to high levels of planning, possibly by pre-activation of several harmonic options (multiple arrows), while classical pianists spread focus from high to low levels, via straightforward planning of the most probable order of the action sequence (one arrow) and rapid assignment of more relevant motor parameters of the single act (bold arrow).

.

2003; Lashley, 1951; MacKay, 1982), in that they can be grounded in the visual-motor system. Remarkably, this suggests that, after intense training, acquired knowledge of harmony influences experts' performance at high-levels of the action control hierarchy to meaningfully integrate discrete elements into sequences. This might in turn increase proficiency of performance on top of fine movement optimisation (Penhune & Steele, 2012). The automatisation of high-level/cognitive structures may be a key step during training to reach excellent levels of fluency and flexibility, to free cognitive resources for more communicative interpretative purposes.

Second, harmonic processing in musical actions and auditory perception relied on dissociable fronto-parietal and fronto-temporal neural networks, respectively. The divergence of dorsal (action-task related) and ventral (audio-task related) rIFG peaks and connectivity profiles are in line with dual stream models of the visuo-spatial (Goodale & Milner, 1992; Kravitz et al., 2011) and auditory system (Rauschecker & Scott, 2009; Rauschecker, 2011). According to these models, dorsal portions of IFG (BA44) are interconnected with the parietal and temporal lobe within dorsal processing streams for time-dependent mechanisms that afford transformation between sensory input (visuo-spatial or sound) and motor representations, thereby supporting action. Dorsal stream involvement has been shown previously for goal-related actions (Kravitz et al., 2011), speech production (Hickok & Poeppel, 2007) and singing (Loui, 2015; Zarate, 2013) and is compatible with our fronto-parietal network observed in pianists during musical action imitation. Moreover, since the audio contrast comprised frontal activation extending to dorsal IFG (BA44), it is plausible that also during listening (although without imitation) a dorsal stream of auditory information might have been involved for mapping sound to action simulated by pianists (Zatorre et al., 2007). Ventral IFG (BA45), in turn, is known as endpoint of the auditory ventral stream that, in concert with posterior temporal areas, is classically thought to process pitch information during singing (Zarate, 2013) and to map sound to meaning (Hickok & Poeppel, 2007; Rauschecker & Scott, 2009), compatible with our fronto-temporal network observed during listening. Although musical harmony does not have referential meaning as language, harmonic incongruities do have musical significance to listeners — i.e., intra-musical meaning (Koelsch, 2011) — in that the harmonic context leads towards a target chord that can be classified as more or less appropriate for musical closure. Flexible and proficient music performance is likely to benefit from the dynamic weighting of these dissociable visual-motor and auditory circuits for prediction and motor planning based on internalised knowledge of harmony.

Third, behavioural and electrophysiological signatures of hierarchical organisation of action control processes can be observed in complex sequential behaviours, such as piano playing (Studies II and III) (Koechlin & Summerfield, 2007; Lashley, 1951; Vallacher & Wegner, 1987). During music production, higher-level structure-plans about appropriate action sequencing are constructed throughout the unfolding musical context and reassessed in response to unexpected harmonic events (context influence on response times, key errors, late negativity, theta and alpha power changes elicited by incongruent compared to congruent chords). Therefore, based on internalised structural knowledge, forthcoming acts (what to play) can be determined by the higher structure-plans emerging from the unfolding sequence. Conversely, the set of fingering parameters (how to play) is specified at the later stage of single act execution, being not modulated by the previous context (no context influence on response times, fingering errors, positivity, theta and alpha power changes elicited by incorrect compared to correct fingering). These principles may be generalisable to complex sequential actions and might be pivotal for fluency, flexibility and coordination of other fast and accurate sequential movements, including speech, in which proficiency is reached by the majority of the human population (MacKay, 1982).

The last important point of this thesis is that the multilevel action planning can be tuned in performers that experience different demands in their daily practice, i.e., focus on different aspects of the action (Study III). Behavioural and electrophysiological results from jazz and classical pianists — who are usually required to focus more on their structure-generative or structure-interpretative abilities respectively — demonstrated measurable differences in action control processes at the levels of structure-planning of the action sequence and parameter specification of single acts. This study adds new insights into the notion that focus on the outcome of an action fosters motor skill learning and consolidation (Wulf, Höß, & Prinz, 1998; Wulf, McNevin, & Shea, 2001): we showed that habitual focus on different aspects of the action outcome (i.e., the harmonic discourse of a piece vs. the interpretation conveyed by a particular way of execution) can have differential effects on action control processes during execution of the same task by boosting abilities specifically required in jazz and classical performance. This implies that core motor functions in the action control hierarchy are influenced by contingencies and past experience of the performers, possibly accounting for the differential specialisation in highly skilled performance and, more generally, the exceptionality that characterises individual behaviours, like drawing or speaking.

1.5 Conclusion

The experimental approach adopted in these studies built the basis for future research on complex sequential behaviours by taking music production as a testable example. Musical performance is in fact an ecologically valid complex action, that can be precisely measured in lab conditions while applying several possible types of manipulations (motor and stimulus complexity, sensorial feedback, solo/ensemble, subjective reward) and testing different types of populations (patients/beginners/experts/specialised experts) (D'Ausilio, Novembre, Fadiga, & Keller, 2015). Music, because of its intrinsically structural and combinatorial nature, may allow to better understand how the brain makes use of regularities to generate a possibly infinite number of meaningful behaviours. Moreover, the merged cognitive and motor lines of inquiry in music performance make musicians an ideal test-bed to investigate how cognition interacts with motor control to flexible integrate single acts into larger coherent sequences and to express communicative intentions — phenomena that similarly occur in other human actions, such as speech. These core attributes of music performance can be intentionally or intuitively adjusted to particular conditions under the influence of past experience. Therefore, they can be informative about that variability of complex behaviours that makes each individual exceptional. Finally, music through creation and resolution of expectancies evokes pleasure in listeners and engages reward-related brain regions. Similarly, musicians' dedicated and intense practice may be driven by an inner reward generated when an expected musical idea is satisfactorily translated into action. Future studies on expert musicians may help to understand how certain complex actions with no immediate reward value can be reinforced and reiterated via interaction with high cognitive functions, by eventually leading already sophisticated human behaviours to extraordinary levels of exceptionality.

1.6 Outlook

Inevitably, the presented findings opened new questions that deserve future investigation. Some outstanding questions immediately related to these studies are presented below.

Is there a bidirectional information flow between computational and modality-specific regions during harmonic processing? In both audio and action tasks in Study I, the processing of harmonic violations recruited lower modality-specific and higher computational regions (BA21 and BA44/45 in the audio, as well as BA7 and BA44/BA6 in the action task) between which feedforward and feedback information loops may allow bottom-up and top-down interaction (Friston, 2010). From our EEG investigations in production and previous auditory studies (for review see Koelsch, 2011), we can define early and late stages in which stimulus evaluation and structural revision processes may respectively occur. Evidence for bidirectional and recurrent processing between lower and higher cortical regions may be obtained by means of non-invasive brain stimulation techniques (e.g., transcranial magnetic stimulation, TMS) targeting lower-level nodes of the network at different stages after stimulus onset. Disruption of harmonic processing (measured with overt judgement on harmonic congruency) may reflect initial feedforward information when stimulation occurs at early stages after stimulus onset, while it may reflect revision processes under "top-down" control from higher nodes when stimulation is applied at later stages.

Where and when do the audio and action networks interact in real performance with sound? A relevant next step would be to investigate the neural bases and temporal dynamics of the interaction between the audio and action streams during real performance. A division of labor between parallel ventral and dorsal processing streams has been postulated for the execution of a wide range of higher-order cognitive processes, including visual and auditory perception, language, prosody and singing (Goodale & Milner, 1992; Kravitz et al., 2011; Rauschecker & Scott, 2009; Hickok & Poeppel, 2004; Zarate et al., 2010; Sammler, Grosbras, Anwander, Bestelmeyer, & Belin, 2015). Study I showed that a similar conceptualisation applies to harmonic processing in action and perception. However, a clear division of the ventral and dorsal processing streams is artificial (Cloutman, 2013), resulting from experimental manipulations (listening without playing and playing without listening) which does not reflect the natural environment of a musician. In fact, a successful performance requires at least the complex collaboration and integration of processing between the auditory and visual-motor systems. The neural bases of this interaction and its temporal dynamics can be addressed by means of magnetoencephalography

(MEG) while musicians are required to play (with sound) musical sequences that are harmonically manipulated (congruent/incongruent) either in terms of action, auditory feedback, both, or none of these factors. Divergent temporal signals between the harmonic violations in audio and action format and their source localisation may unveil integration or segregation of the two information types along the dual streams at different processing time points.

Do differently specialised musicians rely on the same action network for harmonic processing? Studies II and III distinguished different levels of action control in sequential planning (the order of units in a sequence, the what and the how). Crucially, Study III showed that a rigid notion of hierarchical planning is too simplistic to explain the variability of music performance. For instance, jazz pianists showed greater flexibility in structure-building/reassessment compared to classical pianists. However, we don't know whether this difference in flexibility is explained by similar but more efficient structure-building computations, or by the contribution of other systems (e.g., auditory imagery which is particularly trained in jazz musicians). By capitalising on the design of Study I on classical pianists, neuroimaging evidence from jazz pianists should unveil whether they rely on the same fronto-parietal network as classical pianists, but have greater task-specific functional connectivity (Pinho, de Manzano, Fransson, Eriksson, & Ullén, 2014) or greater IFG activation for keeping active more harmonic alternatives, or whether other systems, like the auditory system, contribute to speed up the revision of the structure-based plan.

Does different action focus during practice (i.e., structure-generative or structure-interpretative) lead to long- or short-term effects on the way pianists implicitly prioritise one or the other level of action planning? To examine whether the results of Study III reflect either short-term reversible cognitive tuning or to long-term development of a "cognitive style", one can first test the effects of task instruction to focus either on one or the other level of action (the "what" or the "how"). Therefore, if these effects were driven by the long-term tendencies developed by pianists through practice, even when instructed to focus on the one single aspect of the action, groups should still show training-style dependent behavioural and neural responses. If so, long-term brain plasticity accounts could be investigated with between-group comparison of structural-anatomical measures.

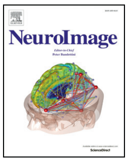
Part II

Experiments

Chapter 2

Neural networks for harmony in perception and action

ARTICLE IN PRESS
YNIMG-13383; No. of pages: 11; 4C: 3, 5, 6
NeuroImage xxx (2016) xxx–xxx
Contents lists available at ScienceDirect
NeuroImage
journal homepage: www.elsevier.com/locate/ynimg



Neural networks for harmonic structure in music perception and action

R. Bianco [a,*], G. Novembre [b], P.E. Keller [b], Seung-Goo Kim [a], F. Scharf [a], A.D. Friederici [a], A. Villringer [a], D. Sammler [a]

[a] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[b] The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, Australia

ARTICLE INFO

Article history:
Received 13 April 2016
Accepted 15 August 2016
Available online xxxx

Keywords:
Music
Harmony
Syntax
IFG
Functional connectivity
Prediction

ABSTRACT

The ability to predict upcoming structured events based on long-term knowledge and contextual priors is a fundamental principle of human cognition. Tonal music triggers predictive processes based on structural properties of harmony, i.e., regularities defining the arrangement of chords into well-formed musical sequences. While the neural architecture of structure-based predictions during music *perception* is well described, little is known about the neural networks for analogous predictions in musical *actions* and how they relate to auditory perception. To fill this gap, expert pianists were presented with harmonically congruent or incongruent chord progressions, either as musical actions (photos of a hand playing chords) that they were required to watch and imitate without sound, or in an auditory format that they listened to without playing. By combining task-based functional magnetic resonance imaging (fMRI) with functional connectivity at rest, we identified distinct sub-regions in right inferior frontal gyrus (rIFG) interconnected with parietal and temporal areas for processing action and audio sequences, respectively. We argue that the differential contribution of parietal and temporal areas is tied to motoric and auditory long-term representations of harmonic regularities that dynamically interact with computations in rIFG. Parsing of the structural dependencies in rIFG is co-determined by both stimulus- or task-demands. In line with contemporary models of prefrontal cortex organization and dual stream models of visual-spatial and auditory processing, we show that the processing of musical harmony is a network capacity with dissociated dorsal and ventral motor and auditory circuits, which both provide the infrastructure for predictive mechanisms optimising action and perception performance.

1. Introduction

The brain shows a fine sensitivity to patterns and regularities that afford the prediction of incoming events in different domains (Tenenbaum et al., 2011). The theory of predictive coding (Friston, 2010) constitutes a unifying framework for human cognition and considers the brain as a "hypothesis tester" with the goal to optimise perception and action by constantly matching incoming sensory inputs with top-down predictions. Within a multi-level cascade of neural processes at different time scales, higher-level predictions act as priors for lower-level processes based on contextual information, previous exposure and acquired long-term knowledge. Recently, predictive coding theory has been used to explain predictions in the action domain (Kilner et al., 2007), as well as in music perception based on priors related to melodic (pitch) content (Pearce et al., 2010), metric structure (Vuust and Witek, 2014), or harmony (Rohrmeier and Koelsch, 2012). The present study takes a comparative stance on predictions in both music perception and action, with a specific focus on Western tonal harmony.

Theoretical accounts refer to harmony as combinatorial arrangement of chords within musical sequences characterized by local and non-local dependencies (Swain, 1995). An instance of these dependencies is that a typical chord progression in Western tonal harmony starts and ends with a reference chord to which some chords are overwhelmingly likely to move to, while they rarely move to others (Tymoczko, 2003). Psychologically, these dependencies are predicted and perceived as tension-resolution patterns by listeners who have been sufficiently exposed to the prevailing musical system (Krumhansl, 1983; Lerdahl and Jackendoff, 1983). Convention in the field of music cognition has that the harmonic principles that govern musical structure are considered as part of a musical "syntax" (Bharucha and Krumhansl, 1983; Koelsch and Siebel, 2005; Patel, 2003), that also includes melodic and/or rhythmic principles of music (Large and Palmer, 2002; Rohrmeier and Koelsch, 2012). Here, we consider "syntax" generally as the knowledge of regularities that control the integration of smaller units into larger musical phrases (Swain, 1995) and thereby support predictions. It is well established that tacit knowledge about structural regularities of music 1) is acquired implicitly (Loui et al., 2009; Rohrmeier and

---

* Corresponding author at: Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstr. 1a, 04103 Leipzig, Germany.
E-mail address: bianco@cbs.mpg.de (R. Bianco).

http://dx.doi.org/10.1016/j.neuroimage.2016.08.025
1053-8119/

Please cite this article as: Bianco, R., et al., Neural networks for harmonic structure in music perception and action, NeuroImage (2016), http://dx.doi.org/10.1016/j.neuroimage.2016.08.025

Rebuschat, 2012; Tillmann et al., 2000), 2) largely shapes our musical competence across different musical systems and cultures (Eerola et al., 2006; Lartillot and Ayari, 2011), and 3) enables listeners to cognitively link current auditory items to past events and to generate predictions on forthcoming events (Patel, 2003; Tillmann, 2012). In the present study we will focus on harmonic regularities and investigate how they govern predictions during (auditory) music perception and (silent) musical actions.

Harmony not only defines the sequence of musical *sounds* but also co-determines the associated chain of musical *actions*. Therefore, the implicit knowledge of harmonic regularities might influence not only *listeners*' predictions, but also musicians' *action* planning during performance (Palmer and van de Sande, 1995). While regularity-based predictions during music *listening* have already been thoroughly investigated (Rohrmeier and Koelsch, 2012; Tillmann, 2012), the neural basis of motor predictions in musical *actions* has not been explored in depth (Maidhof et al., 2009; Ruiz et al., 2009). Recent behavioural (Novembre and Keller, 2011) and electrophysiological studies on music production (Bianco et al., 2016; Sammler et al., 2013b) revealed slower response times, higher number of errors and neural processing costs (a centro-parietal negativity) in expert pianists when asked to silently execute harmonically incongruent compared to congruent chord progressions. These costs were associated with the motor reprogramming of a pre-planned, congruent, action in face of an unexpected incongruity, and were taken as indirect evidence that pianists' action planning was based on musical context and internalised knowledge of harmony. In other words, these findings imply that harmonic structure might implicitly regulate mechanisms of motor control to improve music performance beyond fine movement optimization (Bianco et al., 2016; Novembre and Keller, 2011).

The goal of the present study is to identify the brain areas involved in motor planning based on the regularities of Western tonal harmony, to explore the connectivity between these areas and to compare this network with the neural network sub-serving analogous processes in auditory music perception. The rationale behind this study is that expert pianists have internalised the rules of harmony not only auditorily but also in the hand action domain. Their substantial motor training should enable them to parse harmonic dependencies also in sequences of silent musical actions to facilitate prediction and planning of forthcoming motor acts during performance. This is because the same harmonic structure in sequences of sounds or sequences of actions without sound (i.e., those movements typically employed for producing these sounds) should trigger cognitive processes that are analogous with regard to the structural information. At the same time, processing should differ between perception and action with regard to the associated sensory and memory retrieval processes (i.e., auditory sound vs. motoric act). Here, we sought to isolate and compare the neural networks involved in harmony processing during either perception or (silent) actions, i.e. to probe the potential contribution of auditory and motor prediction of harmony that are otherwise co-occurring during real music production.

Neural hypotheses for musical syntax processing (i.e., harmony) in music *perception* (Koelsch, 2011; Patel, 2003; Tillmann, 2012) posited a special role of frontal computational regions that successively integrate incoming information into higher-order structures by drawing on knowledge about regularities stored in posterior brain regions. Neuroimaging research points to the inferior frontal gyrus (IFG) as the critical computational area that, together with a repository of regularities in posterior auditory regions superior temporal gyrus (STG), affords the prediction of future musical sounds based on the context and listener's long-term music structural knowledge (Kim et al., 2011; Koelsch et al., 2005; Maess et al., 2001; Musso et al., 2015; Sammler et al., 2011; Tillmann et al., 2006). Interestingly, IFG has been associated not only with structural integration and prediction of musical sequences, but also with structuring of complex *actions* (Fuster, 2001; Koechlin and Summerfield, 2007) outside the music domain. Lesions of the left IFG cause impairment in sequencing pictures representing human actions (Fazio et al., 2009), and bilateral IFG are involved in evaluating whether constituent acts belong to the same or separate sub-goals (Farag et al., 2010). Moreover, bilateral IFG activations have been reported during execution of series of motor acts that were organised according to hierarchical action plans (Koechlin and Jubault, 2006). In sum, IFG has become central to hypotheses on processing of structured sequential information in perception and action (Fitch and Martins, 2014; Fiebach and Schubotz, 2006, for various perspectives see Cortex, 2006, vol.2, issue 42), making it conceivable that IFG is also involved in parsing and predicting structural information embedded in musical actions.

What has received less attention than the role of IFG, however, is its interaction with task-relevant posterior systems of knowledge during structural processing. In other words, apart from frequently reported co-activations of IFG and auditory temporal regions during music listening (Koelsch and Siebel, 2005), the characterization of other 'modality-specific regions', e.g., in musical action, and particularly their connectivity with frontal 'computational regions' remains uncertain. In this study, we tested whether pianists' action planning based on knowledge of Western tonal harmony involves (i) IFG in interaction with (ii) posterior visual-motor areas. Furthermore, we (iii) compared the functional connectivity profiles of IFG during the processing of musical actions and auditory sequences that contained similar harmonic violations.

We acquired resting state fMRI data from expert pianists, and then fMRI data during an audio and an action task in which the same harmonic sequences were either auditorily presented or had to be motorically imitated. In the audio task, pianists listened to 5-chord sequences (similar to Koelsch et al., 2005) in which the last chord was either harmonically congruent or incongruent with the preceding musical context. In the action task, in total absence of musical sound, participants were presented with series of photos of a pianist's hand performing the same congruent/incongruent chord progressions on a piano (Bianco et al., 2016). To engage the motor system in the processing of musical actions, pianists had not only to watch the movements, but also to manually reproduce them on a glass-board. The contrasts of incongruent minus congruent chords during listening or imitation were used to functionally segregate modality-specific areas and to isolate frontal computational areas. To demonstrate crosstalk between these regions, we used the latter as seeds in a functional connectivity analysis of the resting state fMRI data.

If harmonic violations of audio sequences activate IFG, then violations of action sequences with the same musical structure should also activate IFG as parser of harmonic regularities and top-down generator of predictions. On the other hand, we expected to find divergent activity in temporal auditory or parietal visual-motor regions associated with item identification and storage of knowledge in their modality-specific format. Finally, by mirroring task-based activation (Smith et al., 2009), the resting-state data should reveal processing streams involved in processing harmonic regularities in music perception and action.

2. Materials and methods

2.1. Participants

29 pianists (17 female) aged 20–32 years (mean age: 24.7, SD = 2.9) took part in the experiment. They had a minimum of 5 years of piano training in classical Western tonal music (range = 5–27 years, mean years of training = 17.2, SD = 4.8) and had started to play the piano at an average age of 7.3 years (SD = 3.08). None of the pianists had training in improvisation or other musical styles. All participants were naïve with regard to the purpose of the study. Written informed consent was obtained from each participant before the study that was approved by the local ethics committee.

Please cite this article as: Bianco, R., et al., Neural networks for harmonic structure in music perception and action, NeuroImage (2016), http://dx.doi.org/10.1016/j.neuroimage.2016.08.025

2.2. Stimuli

Stimuli (see Fig. 1) consisted of 60 different chord sequences that were presented as piano sounds in the listening task (similar to Koelsch, 2005), and as photos of a hand playing chords on a piano in the action imitation task (Bianco et al., 2016). The sequences were composed of 5 chords according to the rules of classical harmony and had various melodic contours. The first chord always represented the tonic (based on the first degree of the scale in the relevant musical key). The second chord could be tonic, mediant (based on the third scale degree) or subdominant (based on the fourth scale degree). Chords at the third position were subdominant, dominant, or dominant six-four chords, and chords at the fourth position were dominant seventh chords. At the last position, the target chord of each sequence was manipulated in terms of harmonic congruency (CONG), so that the last chord could be either congruent (a Tonic chord typically used to resolve a musical sequence) or incongruent (a Neapolitan chord that sounds normal when played in isolation but constitutes a violation when used at the end of a standard harmonic progression). Both the Tonic and Neapolitan are consonant major chords built on the 1st and lowered 2nd scale degree, respectively (i.e., A for Tonic and Bb for Neapolitan in A-major). Consequently, and due to the relationship of the tonalities within the circle of fifths, the exact same chord that acts as a Tonic in one tonality (e.g., A – #C – E in A-major), acts as a Neapolitan in another tonality (i.e., Bbb (=A) – Db (=#C) – Fb (=E) in Ab-major). We exploited this relationship and presented five sequences from each of six different tonalities (D, E, Bb, Ab, A and Eb major), such that the majority of final chords were presented as both Tonic and Neapolitan across the experiment. Therefore, potential neural differences in processing congruent and incongruent chords cannot be due to chord identity but more likely reflect harmony-related processes. With regard to the stimuli of the action block, the choice of different tonalities further allowed us to balance the visual appearance (i.e., number of black/white keys) and difficulty of execution of the target chord (i.e., movement distance from second last to target chord) in congruent and incongruent conditions (for visual appearance: average of 1.3 ± 0.5 black keys in the congruent and 1.2 ± 0.7 in the incongruent chords; for difficulty: average of 1 ± 0 key distance in the congruent and 1.5 ± 0.2 in the incongruent sequences). All sequences were played with normal fingering that was rated as being similarly conventional for congruent and incongruent endings (see Bianco et al., 2016).

The audio stimuli were created with Logic Pro 8 (Apple Inc.), normalised for loudness (RMS, root mean square) with Adobe Audition CS 6 and had a total duration of 6 s (1 s for each of the first four chords and 2 s for the target chord). In the action block, the same chord sequences were presented as photo series showing a male pianist's right hand pressing three keys forming each chord on a piano in conventional fingering (Yamaha Clavinova CLP150, Yamaha Music Europe GmbH, Rellingen, Germany). Red circles were superimposed on top of each pressed key (cf. Bianco et al., 2016) for the whole duration of the photo to facilitate the recognition of the pressed keys. Each photo was presented for 2 s (total sequence duration: 10 s).

2.3. Procedure

The experimental session started with 14 min resting state fMRI data acquisition in which participants were instructed to keep their eyes open and not to fall asleep. To prevent any task-related bias in the measures of functional connectivity, pianists were asked not to practice piano on the scanning day. Thereafter, the task session started and lasted for approximately 25 min.

In the scanner, participants were required to imitate musical actions or to listen to musical sequences in two separate blocks with a counterbalanced order across the group (Fig. 1). Stimulus presentation was controlled in an event-related design with Presentation software (version 14.9, Neurobehavioural Systems, Inc.). In both blocks, congruent and incongruent sequences were intermixed in a way that no >3 sequences of the same condition followed each other. The inter-trial interval (ITI) ranged from 3 to 9 s and during this period participants saw a black screen. During action imitation, no sound was played. Pianists were asked to watch the performing hand in the photos and to simultaneously copy the presented hand postures on a 5×15 cm^2 glass-board with their right hand (Fig. 1, left panel). To motivate participants to follow the sequence accurately, they were told that their performance was monitored with a camera (MR-compatible camera, 12M camera, MRC Systems, Heidelberg Germany). Since it is a common way for pianists to mentally practice by motorically simulating piano performance, the playing along was meant to maximally involve the motor system during the processing of musical actions. In the audio block, pianists were asked to carefully listen to the sequences without playing along (Fig. 1, right panel).

Only to ensure that participants paid attention to the stimuli and to assess their awareness of the violations, 10 trials (1/6 of the trials) in both tasks were followed by a prompt that asked participants to judge the harmonic correctness of the last presented sequence. The judgement required a button response performed with the index or middle finger of the left hand (key assignment was counterbalanced across participants). These judgement trials were equally distributed over congruent and incongruent trials, and required 50% yes and 50% no responses.

After the scanning session, participants filled out a questionnaire to assess the degree to which they had imagined the sound of the chord sequences during the action imitation task and the movements to produce the sequences during the listening task.

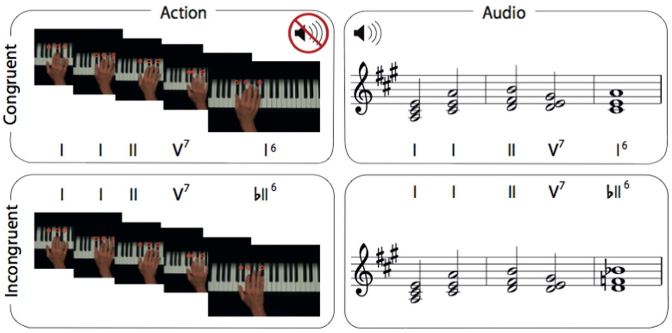


Fig. 1. Experimental design: expert pianists were presented with harmonically congruent or incongruent chord progressions, presented either as muted musical actions (photos of a hand playing chords) that they were required to imitate on a glass-board (left panel), or in an auditory format that they listened to (right panel).

Please cite this article as: Bianco, R., et al., Neural networks for harmonic structure in music perception and action, NeuroImage (2016), http://dx.doi.org/10.1016/j.neuroimage.2016.08.025

2.4. Data acquisition

The experiment was carried out on a 3.0-Tesla Siemens TIM Trio whole body magnetic resonance scanner (Siemens AG, Erlangen, Germany) using a 32-radiofrequency-channel head coil. Functional magnetic resonance images were acquired using a T2*-weighted 2D echo planar imaging (EPI) sequence. During 14 min of acquisition ($TE = 36.5$ ms, $TR = 1400$ ms) at rest (eyes open, instructed not to fall asleep) 410 volumes were acquired with a square FOV of 64 axial slices of 2.3 mm thickness and no gap ($2.3 \times 2.3 \times 2.3$ mm^3 voxel size) with a flip angle of 69°. Functional images during the two tasks were acquired using an EPI sequence with $TE = 30$ ms and $TR = 2000$ ms. 456 and 377 volumes were acquired in the action and audio block, respectively, with a square FOV of 210 mm, with 37 interleaved slices of 3.2 mm thickness and 15% gap ($3 \times 3 \times 3.68$ mm^3 voxel size) aligned to the AC-PC plane, and a flip angle of 77°. For anatomical registration, high-resolution T1-weighted images were acquired using a 3D MP2RAGE sequence ($TI_1 = 700$ ms, $TI_2 = 2500$ ms, $TE = 2.03$ ms, $TR = 5000$ ms) with a matrix size of $240 \times 256 \times 176$, with 1 mm isotropic voxel size, flip angle$_1$ of 4°, flip angle$_2$ of 8°, and GRAPPA acceleration factor of 3.

2.5. Data analysis

2.5.1. Task-based fMRI

fMRI data of 29 participants were analysed with statistical parametric mapping (SPM8; Welcome Trust Centre for Neuroimaging; http://www.fil.ion.ucl.ac.uk/spm/software/spm8/) using standard spatial pre-processing procedures. These consisted of: slice time correction (by means of cubic spline interpolation method), spatial realignment, co-registration of functional and anatomical data (uniform tissue-contrast image masked with the 2nd inversion image from the MP2RAGE sequence), spatial normalisation into the MNI (Montreal Neurological Institute) stereotactic space, that included resampling to $2 \times 2 \times 2$ mm voxel size. Finally, data were spatially low-pass filtered using a 3D Gaussian kernel with full-width at half-maximum (FWHM) of 8 mm and temporally high-pass filtered with a cut-off of 1/128 Hz to eliminate low-frequency drifts.

Statistical parametric maps for the whole brain data were generated in the context of the general linear model (GLM) separately for the action imitation and the listening task. The evoked hemodynamic response to the onset of the final chord was modelled for the congruent and incongruent conditions as boxcars convolved with a hemodynamic response function (HRF). To this design, we added estimated motion realignment parameters as covariates of no interest to regress out residual motion artefacts and increase statistical sensitivity. To identify hemodynamic responses related to the processing of harmonic violations, we computed the first level contrast CONG (i.e., incongruent > congruent chords), separately for the action imitation and the listening task. For random effects group analyses, the resulting contrast images were submitted to one-sample *t*-tests. Additionally, to identify areas that are modality-specific to either action or audio representation of the harmonic structure, we compared the CONG contrasts of the two tasks by means of paired *t*-tests. We controlled family-wise error rate (FWER) of clusters below 0.05 with a cluster-forming height-threshold of 0.001. Anatomical labels are based on Harvard-Oxford cortical structural atlas implemented in FSL (http://neuro.debian.net/pkgs/fsl-harvard-oxford-atlases.html).

2.5.2. Resting-state fMRI

In order to investigate intrinsic connectivity of the peak regions from task-based fMRI datasets (Bressler and Menon, 2010), independent resting state fMRI datasets were obtained from 28 of the pianists that participated in the task-fMRI session (one r-fMRI data set was not acquired due to technical problems). The pre-processing of the resting state data (realignment, unwarping, slice-timing correction) was done using SPM8 by means of DPARSF (http://rfmri.org/DPARSF) SPM-based toolboxes. We applied a GLM to regress out non-neuronal signal changes due to physiological noise and, most importantly, head motions. The regressors included six rigid-body motion parameters, five principle components extracted by the “anatomical CompCor” (Behzadi et al., 2007) (i.e., signal from white matter and cerebral fluid masks defined from anatomical scans), and finally global signal (Power et al., 2015). Thereafter, band-pass-filtering (0.009 and 0.08 Hz), spatial normalisation of functional data into MNI stereotactic space (with resampling to $2 \times 2 \times 2$ mm^3 resolution), and finally a minimal spatial smoothing with the FWHM of 3 mm were applied to the residual time-series.

Resting-state functional connectivity (RSFC) was defined by Pearson's correlations between a time-series of a seed region and time-series of whole brain voxels. Spherical seed regions (5 mm radius) were centred in IFG at the peak coordinates of the CONG contrasts obtained in the task-based analyses of the action imitation and the listening task. In order to match the smoothness of noise in task-based and resting-state analyses, the correlation maps were further smoothed with the FWHM of 2 mm, resulting in an effective FWHM of about 8 mm.

Voxel-wise paired t-tests were performed to identify differences between the two seed-based correlation maps (i.e., action and audio seeds). The normality assumption based upon the difference between the two correlation coefficients across subjects was fulfilled, as confirmed by Kolmogorov-Smirnov tests. We controlled FWER of clusters below 0.05 with a cluster-forming height-threshold of 0.001 in all reported results. Harvard-Oxford cortical structural atlas was used to assign anatomical labels.

3. Results

3.1. Behavioural

To ensure that participants paid attention to the stimuli in both modalities and that they were generally able to recognise the harmonic structure underlying the sequences, they were required to overtly judge harmonic congruency in 1/6 of the trials. They performed significantly above chance level in these explicit judgments both in the action (mean ± *SD*: 68.96 ± 27.06% correct, $p < 0.001$) and in the audio block (mean ± *SD:* 91.03 ± 16.40% correct, $p < 0.001$), as tested with one-sample *t*-tests against 50% chance level. Action block performance was lower than audio block performance ($t(28) = 4.704$, $p < 0.001$), partly due to a response bias towards “congruent” answers in the action block (32% of incongruent trials misclassified as congruent vs. 15% of congruent trials misclassified as incongruent: $t(28) = -2.95$, $p = 0.007$). These differences in explicit judgement might indicate that pianists were less consciously aware of the harmonic violations during action imitation than during passive listening, possibly because it is more taxing to copy sequences of actions on-line than to just listen. This may have led pianists to focus on the motor-executive task rather than harmonic relationships in the action block (see Discussion).

3.2. fMRI

3.2.1. Fronto-parietal vs. fronto-temporal areas for musical action vs. perception

In the action imitation task, the CONG contrast (incongruent vs. congruent chords) yielded larger hemodynamic responses in frontal and parietal areas, comprising the dorsal portion of rIFG (BA44) bordering precentral sulcus, and bilateral clusters extending from superior parietal cortex (SPL: BA7) to the inferior parietal and middle occipital gyrus (MOG: BA19) (Fig. 2 left-upper panel, Table 1).

In the listening task, the same contrast evoked stronger activity in frontal and temporal areas, including right IFG (BA44/45, peak in ventral

Please cite this article as: Bianco, R., et al., Neural networks for harmonic structure in music perception and action, NeuroImage (2016), http://dx.doi.org/10.1016/j.neuroimage.2016.08.025

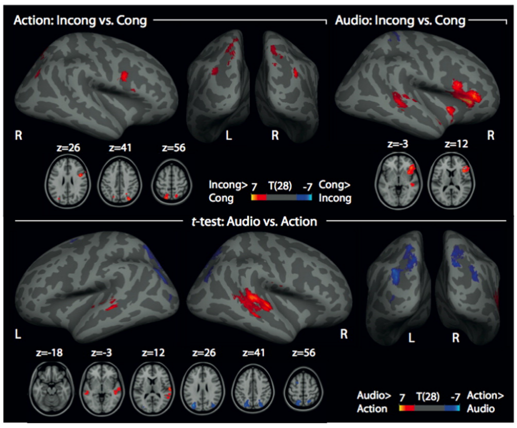


Fig. 2. Harmonic violations elicited activations in fronto-parietal areas during action imitation (upper left panel) and in fronto-temporal areas during listening (upper right panel). Areas involved in structural processing specifically for the action and the audio sequences were identified in bilateral posterior parietal regions (cold colours) and in bilateral temporal regions (hot colours), respectively (lower panel). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

BA45) and the right posterior superior temporal gyrus and sulcus (pSTG/STS: BA22) (Fig. 2 right-upper panel, Table 1).

To identify areas exclusively recruited depending on stimulus format (photos of actions or audio), the CONG contrasts of both tasks were compared using a paired *t*-test (Table 2). Incongruent actions elicited greater activity in bilateral SPL, MOG, and in the left superior/middle frontal gyrus (frontal eye fields, FEF: BA8) (Fig. 2 lower panel, cold colours). Conversely, auditory violations yielded larger BOLD responses in bilateral STS/STG, compared to the action task (Fig. 2 lower panel, hot colours).

To identify areas commonly recruited during both audio and action task, we masked the audio CONG contrast with the action contrast. This analysis yielded a cluster in rIFG (BA44, $x = 64$, $y = 18$, $z = 24$, cluster extent $= 28$ voxels, $Z = 3.53$, $p_{voxel} < 0.001$) that, however, did not survive the cluster-level FWER correction.

Table 1
Congruency effect (incongruent > congruent) in the action imitation and listening tasks.

Region	Hem.	BA	k	x	y	z	Z-value
Action: CONG incongruent > congruent							
Precentral/inferior frontal gyrus	R	44	182	44	6	26	4.29
		44		64	18	24	3.53
		44		54	14	16	3.49
Middle occipital gyrus	R	19	352	40	−80	36	4.66
Superior parietal lobe	R	7P		32	−78	42	4.32
		7P		16	−70	58	3.54
Superior parietal lobe	L	7P	510	−16	−74	58	4.27
		7P		−20	−70	50	3.97
Middle occipital gyrus	L	19		−30	−80	34	3.72
Audio: CONG incongruent > congruent							
Inferior frontal gyrus	R	45	1667	44	34	2	5.12
		45		44	18	16	4.98
		Insula		36	10	−2	4.49
Superior temporal sulcus, post.	R	22	256	48	−32	0	3.92
Superior temporal gyrus, post.	R	22		70	−24	6	3.59
		22		60	−34	8	3.46
Cerebellum (crus II)	L	–	132	−14	−76	−36	4.37

Whole-brain activation cluster sizes (k), MNI coordinates (x, y, z), and Z-scores for the CONG contrast in action imitation and listening tasks ($p_{voxel} < 0.001$; $p_{cluster} < 0.05$, FWE corrected). BA: Brodmann area, Hem.: hemisphere, post.: posterior.

Table 2
t-test between the CONG contrasts in the action imitation and listening task.

Region	Hem.	BA	k	x	y	z	Z-value
Action > audio							
Superior parietal lobe	R	7	806	26	−76	46	4.46
		7		24	−58	44	4.22
Middle occipital gyrus	R	19		34	−80	38	4.08
Middle occipital gyrus	L	19	1436	−32	−76	24	4.99
Superior parietal lobe	L	7		−30	−74	36	4.75
		7		−20	−72	48	4.41
Superior frontal gyrus	L	8	186	−20	2	60	3.97
		8		−22	−2	72	3.60
Audio > action							
Superior-middle temporal gyrus	R	21/22	1100	64	−16	8	4.79
				62	−32	4	4.64
				52	−30	−8	4.20
Superior-middle temporal gyrus	L	21/22	166	−52	−32	−2	4.11
				−52	−16	2	3.91
				−62	−30	6	3.51

Whole-brain activation cluster sizes (k), MNI coordinates (x, y, z), and Z-scores for the paired-samples *t*-test comparison of the action > audio and audio > action CONG contrast ($p_{voxel} < 0.001$; $p_{cluster} < 0.05$, FWE corrected). BA: Brodmann area, Hem.: hemisphere.

3.2.2. Seed-based functional connectivity from the IFG peak maxima

To gather evidence for the communication between IFG and modality-specific areas, we conducted a seed-based functional connectivity analysis on the resting-state fMRI data acquired from the same pianists. The activation peaks in the IFG clusters in the action imitation and audio task were chosen as seed regions, which were located between the right dorsal BA44 and the pre-central sulcus (action-seed) and in right BA45 (audio-seed). The results are depicted in Fig. 3 (upper and middle panels) and show positive functional connectivity (hot colours) between IFG and (amongst others) regions that were functionally specific to the action or audio musical task.

In line with activity in IFG and posterior parietal regions in the action task, the action-seed in IFG (BA44) exhibited positive correlations within a dorsal motor network comprising bilateral parietal cortex, extending from the anterior ventral supramarginal gyrus (BA40) to the posterior superior parietal lobes (BA7) (Table 3). Notably, there were no significant correlations with temporal regions that were specific to the audio modality. A large cluster peaking in bilateral precentral gyrus showed positive correlations with the action-seed, including subclusters in bilateral BA44 extending to insular regions, ventral premotor cortex (BA6), middle frontal gyrus (BA9) bordering the superior frontal gyrus and the inferior portion of the frontal pole (BA10). Medially, the action seed exhibited positive correlations with the right posterior border of the supplementary motor cortex (BA6) and anterior cingulate (BA24). Finally, there were positive correlations with right inferior temporal gyrus at the temporo-occipital junction (BA20), and bilateral occipitotemporal areas (BA37), cerebellum and thalamus.

Consistent with activity in IFG and temporal areas in the audio task, the audio-seed in IFG (BA45) exhibited positive correlations within the auditory network comprising the posterior part of the right superior temporal gyrus (BA22) and left Heschl's gyrus (including BA41/42) (Table 3). Additionally, there were positive correlations with frontal areas in the right hemisphere including orbitofrontal (BA47/11/12) and frontopolar regions (BA10), superior (BA8), middle frontal areas (BA9), and anterior cingulate gyrus (BA24), and in the left hemisphere, including BA45, BA47, BA9, BA10, and BA12. In the parietal cortex, positive correlations were restricted to bilateral anterior ventral supramarginal gyrus (BA40), without extending to more posterior parietal regions. Finally, there were positive correlations with thalamus and right putamen.

Apart from positive correlations, activity in both action and audio seeds was negatively correlated (Fig. 3 upper and middle panels, cold colours) with activity in areas belonging to the default mode network (DMN), namely the cingulate gyrus and the superior portion of bilateral

Please cite this article as: Bianco, R., et al., Neural networks for harmonic structure in music perception and action, NeuroImage (2016), http://dx.doi.org/10.1016/j.neuroimage.2016.08.025

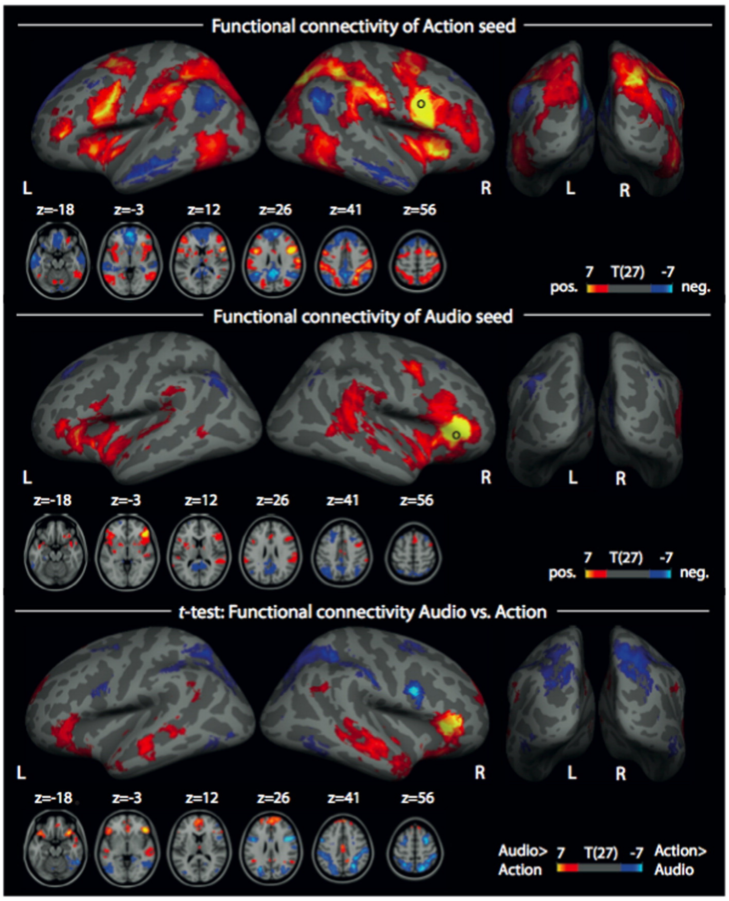


Fig. 3. *Upper and middle panels*: seed-based functional connectivity maps of resting-state data from the action seed in dorsal BA44 and audio seed in BA45, respectively. Seeds are depicted as black circles. Hot and cold colours indicate positive and negative functional connectivity, respectively. Consistent with the task-based activations, the topographical connectivity patterns include posterior parietal regions from the action-seed and temporal regions from the audio-seed. *Lower panel*: *t*-test between the connectivity maps of the action- and audio seed. Correlation values in posterior parietal areas were significantly higher for the action-seed than the audio-seed, whereas temporal regions were more strongly correlated to the audio- than the action-seed. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

lateral occipital cortex extending into angular gyrus. These regions typically show a decrease of activation during attention-demanding tasks and goal-directed behaviours (Uddin et al., 2009). Additionally, negative correlations were found between the action-seed and bilateral anterior middle temporal gyrus and medial prefrontal cortex, anti-correlations that have been associated with highly difficult goal-directed tasks, as could apply in the case of our action-task (McKiernan et al., 2003).

Finally, a paired-samples *t*-test comparing the connectivity maps of the action- and audio-seed (Table 4) confirmed their differential predominant connectivity to parietal and temporal areas, respectively. Specifically, connectivity of the action-seed (compared to audio-seed) was stronger to bilateral posterior parieto-occipital areas, as well as to bilateral cerebellum, right frontal pole, frontal medial cortex and anterior cingulate gyrus, left superior frontal and precentral gyrus. Conversely, the connectivity of the audio-seed (compared to action-seed) was stronger to bilateral superior or middle temporal gyrus, as well as to bilateral cerebellum, right superior frontal gyrus, posterior cingulate and angular gyrus, thalamus, and left frontal operculum.

4. Discussion

The present study investigated the neural bases of action planning and prediction based on long-term knowledge of harmonic regularities and compared them with those involved in auditory prediction. Functional neuroimaging data of expert pianists were acquired at rest, during imitation of (without sound) or listening to (without imitation) harmonically congruent or incongruent chord sequences presented as photos of musical actions or sounds, respectively. Violations in both musical *actions* and *sounds* recruited distinct sub-regions (BA 44 and BA 45, respectively) in right IFG (rIFG) interconnected with parietal visual-motor and temporal auditory areas, respectively. We propose that motoric and auditory long-term representations of harmonic regularities are likely to account for the differential involvement of parietal and temporal areas that enter into dynamic interactions with computations in rIFG. Moreover, the involvement of rIFG in parsing musical action and sound sequences is sensitive to stimulus properties and task – production or perception – accounting for the divergent peak localizations, in line with prevailing models of general prefrontal cortex organization

Please cite this article as: Bianco, R., et al., Neural networks for harmonic structure in music perception and action, NeuroImage (2016), http://dx.doi.org/10.1016/j.neuroimage.2016.08.025

Table 3
Resting-state functional connectivity from the action and audio seed in right inferior frontal gyrus.

Region	BA	Action-seed					Audio-seed				
		k	x	y	z	Z-value	k	x	y	z	Z-value
Right hemisphere (positive correlations)											
Frontal pole	10	125	26	38	−16	4.94					
Middle frontal gyrus	9						310	42	4	46	5.22
Superior frontal gyrus	8						349	4	18	60	4.79
Supplementary motor cortex	6	734	6	14	52	5.92					
Cingulate gyrus, ant.	24	133	4	6	28	5.76	324	4	32	22	4.86
Supramarginal gyrus/superior parietal lobe	40/7	5418	52	−30	48	6.89					
Superior temporal gyrus, post.	22						119	50	−14	−8	4.33
Superior temporal gyrus, post.	22						1148	52	−30	6	4.72
Middle temporal gyrus	20/21/37	1304	52	−56	−12	5.92					
Putamen	–						36	32	−12	−8	4.52
Thalamus	–	120	8	−14	8	5.84					
Cerebellum (VIIb)	–	350	18	−68	−48	5.36					
Cerebellum (VI)	–	95	8	−70	−22	4.98					
Left hemisphere (positive correlations)											
Frontal pole	10	531	−44	38	8	6.28					
Frontal operculum cortex	45						2198	−38	26	0	6.29
Inferior frontal gyrus	44						50	−46	12	22	4.29
Precentral gyrus	6/44/Ins	2415	−44	6	24	6.89					
Middle frontal gyrus	9	626	−26	0	50	6.15					
Cingulate gyrus, ant.	24						52	−2	−12	42	3.98
Heschl's gyrus (H1 and H2)	41/42						173	−52	−14	4	4.07
Supramarginal gyrus/superior parietal lobe	40/7	3814	−60	−30	42	5.92	143	−66	−38	26	4.21
Middle temporal gyrus	20/21/37	1162	−60	−60	−6	5.60	63	−62	−60	8	4.03
Thalamus	–	116	−12	−14	6	4.45	93	−6	−14	2	5.43
Cerebellum (VIIb)	–	969	−26	−66	−52	6.06					
Cerebellum (VI)	–	95	−22	−62	−28	4.98					
Cerebellum (crus II)	–						58	−16	−78	−34	4.40
Right hemisphere (negative correlations)											
Frontal pole	10	11,860	6	60	22	7.08					
Superior frontal gyrus	8						384	28	30	54	4.97
Middle temporal gyrus	21	2367	60	4	−24	6.88					
Cingulate gyrus, post.	24	5798	10	−50	34	7.66	3682	10	−46	12	5.35
Cerebellum (IX)	–	187	4	−50	44	5.81					
Cerebellum (crus I)	–	1799	26	−88	−30	5.99	326	36	−52	−34	4.99
Left hemisphere (negative correlations)											
Frontal pole	10						105	−20	64	−6	4.37
Superior frontal gyrus	8						601	−20	28	38	5.67
Middle temporal gyrus	21	2367	−62	−24	−12	6.88					
Inferior temporal gyrus, post.	20						70	−60	−44	−14	4.17
Lateral occipital cortex, sup.	39	1764	−40	−50	26	6.27	1268	−36	−66	38	5.57
Hippocampus	–	62	−34	−34	−8	5.01	65	−30	−34	−12	4.07
Cerebellum (crus I)	–	27	−44	−56	−42	4.275					

Results of the whole-brain functional connectivity analysis from IFG activation maxima in action imitation and listening tasks. k: cluster size, MNI coordinates (x, y, z), and Z scores. ($p_{voxel} < 0.001$; $p_{cluster} < 0.05$, FWE corrected), BA: Brodmann area, ant.: anterior, post.: posterior, sup.: superior.

(e.g., Fuster, 2001), and dual stream models of the visuo-spatial (e.g., Goodale and Milner, 1992) and auditory system (e.g., Rauschecker and Scott, 2009). Altogether, our results emphasise dissociable, neural action and audio networks in which modality-specific long-term knowledge and contextual information act as priors for the prediction of forthcoming events. In this respect, predictive coding models (Friston, 2010) may yield a unifying explanatory framework for information processing across both action and perception.

4.1. Musical action

The imitation of incongruent actions elicited activations in fronto-parietal areas (see Table 1), including the right inferior frontal gyrus (IFG: dorsal BA44 extending to the border of the precentral sulcus) and bilateral posterior parietal cortex (pSPL: BA7; MOG: BA19).

This activation pattern resembles the typical dorsal fronto-parietal network for visually guided behaviour that integrates sensory information with action-goals through sensorimotor transformations (Gallivan and Culham, 2015; Kravitz et al., 2011). Accordingly, MOG is known as an area involved in capturing relevant visual-spatial dimensions of objects and visually-guided actions (Lingnau and Downing, 2015). SPL has been associated with high-level aspects of motor behaviour, such as the formation of intentions and early movement plans. These processes are aided by critical operations of multisensory integration and visuomotor transformation in SPL (Andersen and Buneo, 2002). Activations in pSPL have been reported during motor imagery of action-goals and trajectories (Aflalo et al., 2015), attentional spatial remapping/reprogramming of pre-selected actions (O'Reilly et al., 2013), and transformation of spatial target information into corresponding actions (Barany et al., 2014; Schon et al., 2002).

One crucial finding was the recruitment of the rIFG (dorsal BA44) when the final chord, predicted by the harmonic structure of the given musical sequence, was violated. This is consistent with the role of IFG in processing high-level aspects of motor behaviours (Grafton and Hamilton, 2007). Experimental evidence emphasises the role of bilateral IFG in processing hierarchical relationships within action sequences either when judging complex familiar activities (Farag et al., 2010) or when executing abstract hierarchically organised patterns of action sequences (Koechlin and Jubault, 2006). Altogether, these combined results suggest that the right IFG supports the structural integration of

Please cite this article as: Bianco, R., et al., Neural networks for harmonic structure in music perception and action, NeuroImage (2016), http://dx.doi.org/10.1016/j.neuroimage.2016.08.025

Table 4
Comparison of rs-functional connectivity from the action- and audio-seed in the right IFG.

Region	BA	Action > Audio seed					Audio > Action seed				
		k	x	y	z	Z-value	k	x	y	z	Z-value
Right hemisphere											
Frontal pole	10	93	48	42	14	3.56					
Superior frontal gyrus	9						4263	4	56	42	5.28
Frontal medial cortex	11	59	4	44	−18	3.88					
Cingulate gyrus	23/24	52	2	6	30	4.58	200	2	−14	38	4.89
Superior temporal gyrus	22						645	52	−8	−8	5.16
Angular gyrus	40						161	44	−46	32	4.20
Lingual gyrus	27	72	16	−42	−6	4.87					
Infer. temporal gyrus, temp-occ.j.	37	996	52	−56	−14	5.53					
Lateral occipital cortex, sup.	7	4448	24	−68	50	6.03					
Precuneus cortex	17	373	24	−54	18	4.80					
Thalamus	–						50	2	−12	10	4.49
Cerebellum (crus II)	–	153	4	−78	−44	4.24	68	30	−88	−36	3.86
Cerebellum (XI)	–	51	16	−46	−48	4.46					
Left hemisphere											
Frontal operculum cortex	47						976	−40	26	0	5.35
Superior frontal gyrus	8	640	−24	4	52	5.17					
Precentral gyrus	6	549	−52	6	40	5.43					
Middle temporal gyrus (middle)	20						206	−56	−20	−12	4.88
Middle temporal gyrus (post.)	21						64	−54	−38	0	4.38
Temporal occipital fusiform cortex	7	66	−24	−58	−12	4.10					
Lateral occipital cortex, sup.	7	3678	−26	−76	30	5.66					
Lateral occipital cortex, inf.	19	567	−50	−76	−4	4.87					
Cerebellum (crus I/II)	–	114	−6	−76	−40	4.30	168	−24	−76	−34	4.44

Results of the *t*-test between whole-brain functional connectivity from IFG activation maxima in action imitation and listening task. BA: Brodmann area, k: cluster size, MNI coordinates (x, y, z), and Z scores. ($p_{voxel} < 0.001$; $p_{cluster} < 0.05$, FWE corrected). temp-occ.j.: temporo-occipital junction, BA: Brodmann area, post.: posterior, sup.: superior, inf.: inferior.

simple acts into more complex combinatorial action sequences. The greater BOLD response during incongruent (compared to congruent) chords may be due to a mismatch with the predicted musical motor act that leads to higher computational costs during structural integration. Importantly, these findings indirectly show that pianists' knowledge of harmonic regularities transfers to the motor domain and enables them to predict and plan forthcoming musical acts during performance.

The absence of auditory activation in the incongruent vs. congruent contrast suggests that pianists relied more on their action knowledge recalled by the execution of the preceding chords than on auditory mechanisms (Bianco et al., 2016; Novembre and Keller, 2011; Sammler et al., 2013b). Note that this finding does not conflict with the large body of experimental evidence for action-perception coupling in trained musicians (for review, see Novembre and Keller, 2014; Zatorre et al., 2007) Our unusual and taxing imitation task on unrehearsed sequences may have led pianists to focus on the motor part of the task, possibly suppressing unhelpful auditory images (cf. Pfordresher, 2012; van der Steen et al., 2014) (cf. Pfordresher, 2012). Alternatively, auditory feed-forward mechanisms may not discriminate between congruent and incongruent chords such that auditory activations cancelled out.

Overall, these fronto-parietal activations complement and support our previous behavioural (Novembre and Keller, 2011) and EEG studies on expert pianists (Bianco et al., 2016; Sammler et al., 2013a,b): silent production of harmonically incongruent chords elicited response time costs and a centro-parietal negativity that was associated with mechanisms of motor reprogramming of a pre-planned action in face of the violation. The activations of SPL and MOG match and support our interpretation of the posterior negativity as a correlate of the spatial remapping and reprogramming of pre-planned actions, and the activation of IFG lends evidence that these mechanisms stand under frontal control.

Within the predictive-coding framework (Friston, 2010), a bidirectional flow of information can be suggested to occur in the two hemispheres between parietal areas, processing visual-motor inputs, and the IFG, performing structural integration of incoming items. Indeed, the functional connectivity analysis of our resting state data revealed strong positive correlations between right BA44 and, amongst others, bilateral superior parietal lobes, also revealed by the task-based analysis. A fronto-parietal network relying on the route of the dorsal visual stream has been associated with sensorimotor transformation during visually guided action planning (Goodale and Milner, 1992). According to motor control theory, these operations might be supported by “forward models”, through which the expected outcome of an action is compared with actual sensory feedback (Wolpert and Flanagan, 2001). In this framework, posterior parietal regions simultaneously represent potential actions whose pre-selection is biased by the influence of internal models from prefrontal regions (Cisek, 2006). The novel finding is that these internal models may be shaped by the musician's knowledge of harmonic regularities and musical context. We propose that, on the one hand, visual-motor information about the current act is forwarded from posterior regions to the IFG that integrates the items and builds up an internal model of the sequence's harmonic structure. On the other hand, this internal model affords predictions of visual-spatial surface features of the next chord in MOG and may bias the pre-selection/representation of harmonically appropriate forthcoming motor acts in SPL. The generated model would be continually validated/updated via the matching between the expected action and the combined visual and proprioceptive signals from the current input (Wolpert and Flanagan, 2001). Interestingly, the combined findings raise the hypothesis that (musical) action knowledge, internal visual-motor models and fronto-parietal information flow may provide the basis on which the motor system contributes to visual perception and prediction of human behaviour (Novembre and Keller, 2014).

4.2. Music perception

In line with previous findings (Koelsch et al., 2005), listening to harmonically incongruent compared to congruent chords elicited activations in fronto-temporal areas: right inferior frontal gyrus (IFG: BA44, BA45) extending into the insular cortex, and right posterior superior temporal gyrus and sulcus (pSTG/STS: BA22).

Please cite this article as: Bianco, R., et al., Neural networks for harmonic structure in music perception and action, NeuroImage (2016), http://dx.doi.org/10.1016/j.neuroimage.2016.08.025

The IFG and the posterior STG have been associated with structural analysis of auditory musical sequences based on internalised knowledge of harmonic regularities (Koelsch et al., 2005; Maess et al., 2001; Sammler et al., 2013a; Tillmann et al., 2006). The IFG has been proposed to support integration of discrete items into higher-order structures, based on which top-down predictions on forthcoming items can be generated. Greater BOLD responses in IFG may reflect the higher computational demand to integrate incongruent chords that are weakly related to the harmonic context and do not fulfil the prediction. Compared to these higher-order computations in IFG, pSTG/STS has been proposed to support lower-level matching processes between the actually perceived and the predicted sensory information (Sammler et al., 2013a). Indeed, posterior superior temporal areas have been associated with physical feature analysis and short-term representation of sounds (Seger et al., 2013), as well as with the identification of the harmonic functions of chords within musical sequences (Musso et al., 2015).

Our connectivity analysis showed a functional coupling between IFG (BA44/BA45) and pSTG/STS, making it plausible to assume bidirectional dynamic fronto-temporal interactions during structural integration processes (Friston, 2010). On the one hand, early sensory analysis of chord functions may be forwarded from temporal to frontal regions where information is structurally integrated and harmonic predictions are established. On the other hand, these predictions may in turn inform the identification process in pSTG/STS where perceived and predicted items are matched to validate or revise the frontal prediction.

Overall, these data emphasise the crucial role of not just one area, but of a dynamic exchange of information between fronto-temporal areas in providing resources for the parsing of complex harmonically organised sounds (Hyde et al., 2011). Neuroanatomically, the fronto-temporal information exchange may be implemented along dorsal or ventral auditory pathways (see further below) (Loui et al., 2011; Musso et al., 2015; Rauschecker, 2011). The anatomical specification of these pathways, their functional relevance and dependency on musical training are interesting topics for future research.

4.3. Dorsal and ventral streams for musical action and perception

As discussed above, harmonic processing in musical actions and auditory perception relied on dissociable fronto-parietal and fronto-temporal neural networks, respectively. Representations of harmonic regularities in either visual-motor or auditory format are likely to account for the differential involvement of parietal and temporal areas, respectively, that both dynamically interact with computational processes in IFG. Interestingly, these interactions involved distinct posterior-dorsal and anterior-ventral rIFG sub-regions, i.e., BA6/44 in the action imitation task vs. BA44/45 in the audio task. This dissociation may either reflect (i) a task-unspecific sensitivity of IFG to structural processing demands in line with models of general prefrontal cortex specialization, or (ii) a task-specific involvement of dorsal and ventral IFG sub-regions as endpoints of different processing streams.

(i) Investigating harmonic structure processing in perception and action necessarily entails differences in experimental setup that alone suffice to induce different processing demands and shift activation peaks within IFG – even if both peaks may reflect similar structural computations. For example, recent theories propose anterior-posterior (Badre and D'Esposito, 2009; Fuster, 2001; Koechlin and Summerfield, 2007) and/or rostral-caudal (Friederici, 2011) gradients of prefrontal cortex organization along which *similar* functions, e.g., the “integration” of discrete items over time, operate at *different* levels of abstraction (Makuuchi et al., 2012). Along these lines, the more demanding imitation task might have triggered integration over shorter segments in the action sequences (i.e., integration at a lower level of complexity), limiting the activation to dorsal BA44 in the action contrast.

(ii) Alternatively, the divergence of dorsal and ventral rIFG peaks and connectivity profiles may arise from the intrinsically different nature of the tasks – silent musical action imitation vs. listening – in line with dual stream models of the visuo-spatial (Goodale and Milner, 1992; Kravitz et al., 2011) and auditory system (Rauschecker and Scott, 2009; Rauschecker, 2011). According to these models, dorsal portions of IFG are interconnected with the parietal and temporal lobe within dorsal processing streams for time-dependent mechanisms that afford transformation between sensory input (visuo-spatial or sound) and motor representations, thereby supporting action. Dorsal stream involvement has been shown previously for goal-related actions (Kravitz et al., 2011), speech production (Hickok and Poeppel, 2007) and singing (Loui, 2015; Zarate, 2013) and is compatible with our fronto-parietal network observed in pianists during musical action imitation. Moreover, since the audio contrast comprised frontal activation extending to dorsal IFG, it is plausible that also during listening (although without imitation) a dorsal stream of auditory information might have been involved for mapping sound to action simulated by pianists (Zatorre et al., 2007). Notably, the present study adds two new insights: first, we demonstrate that frontal and parietal areas along the dorsal stream provide the neural resources for sequential structure processing during production of musical sequences; second, unlike in singing or speech production, our action imitation task eliminated auditory feedback during self-produced actions, hence, leading us to conclude that music-structural predictions can be grounded in the visual-motor control system.

Ventral IFG, in turn, is known as endpoint of the auditory ventral stream that, in concert with posterior temporal areas, is classically thought to process pitch information during singing (Berkowska and Dalla Bella, 2009; Zarate, 2013) and to map sound to meaning (Hickok and Poeppel, 2007; Rauschecker and Scott, 2009), compatible with our fronto-temporal network observed during listening. Although musical harmony does not have referential meaning as language, harmonic incongruities do have musical significance to listeners – i.e., intra-musical meaning as framed by Koelsch (2011) – in that the harmonic context leads towards a target chord that can be classified as more or less appropriate for musical closure.

Although the current findings do not speak to the causal role of the nodes or streams, they altogether highlight the relevance of considering structural integration in music production and perception as a network capacity by taking into account the connectivity between frontal computational and posterior modality-specific regions. Flexible and proficient music performance is likely to benefit from the dynamic weighting of these dissociable visual-motor and auditory circuits for prediction and motor planning based on internalised knowledge of harmony.

5. Conclusion

The present data provide first neuroimaging evidence that expert pianists predict forthcoming musical chords not only in auditory perception, but also in the processing of actions independently of auditory information. Remarkably, this suggests that, after intensive training, knowledge of structural regularities influences experts' action planning via implicit mechanisms of motor prediction/control, and might in turn increase proficiency of performance on top of fine movement optimization.

Our paradigm, in which pianists acted without listening to sound and listened without acting, dissociated a dorsal action and a ventral audio network for harmonic prediction, potentially acting in concert during real production (i.e., playing with sound). The dorsal and ventral networks both involve frontal computational sub-regions in rIFG, interconnected with parietal and temporal posterior systems of knowledge, respectively. These networks are likely to provide the infrastructure that allows frontal areas to keep track of abstract dependencies in sequential information via dynamic exchange with progressively lower-level modality-specific systems of knowledge. Predictive coding is proposed as an explanatory framework that

Please cite this article as: Bianco, R., et al., Neural networks for harmonic structure in music perception and action, NeuroImage (2016), http://dx.doi.org/10.1016/j.neuroimage.2016.08.025

unifies both networks' functional roles: to optimise predictions in action and perception based on previous exposure and knowledge of harmony.

Acknowledgments

We are grateful to the pianists who participated in this experiment and to Bruno Pace who lent a hand in stimulus preparation. We also thank Mauricio Martins, Tomás Goucha, Gabriele Chierchia and Emiliano Zaccarella for inspiring discussions and Sven Gutekunst and Jöran Lepsien for technical support.

References

Aflalo, T., Kellis, S., Klaes, C., Lee, B., Shi, Y., Pejsa, K., Shanfield, K., Hayes-jackson, S., Aisen, M., Heck, C., Liu, C., Andersen, R.A., 2015. Decoding motor imagery from the posterior parietal cortex of a tetraplegic human. Science 348 (80), 906–910.

Andersen, R.A., Buneo, C.A., 2002. Intentional maps in posterior parietal cortex. Annu. Rev. Neurosci. 25, 189–220.

Badre, D., D'Esposito, M., 2009. Is the rostro-caudal axis of the frontal lobe hierarchical? Nat. Rev. Neurosci. 10, 659–669.

Barany, D.a., Della-Maggiore, V., Viswanathan, S., Cieslak, M., Grafton, S.T., 2014. Feature interactions enable decoding of sensorimotor transformations for goal-directed movement. J. Neurosci. 34, 6860–6873.

Behzadi, Y., Restom, K., Liau, J., Liu, T.T., 2007. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. NeuroImage 37, 90–101.

Berkowska, M., Dalla Bella, S., 2009. Acquired and congenital disorders of sung performance: a review. Adv. Cogn. Psychol. 5, 69–83.

Bharucha, J., Krumhansl, C.L., 1983. The representation of harmonic structure in music: hierarchies of stability as a function of context. Cognition 13, 63–102.

Bianco, R., Novembre, G., Keller, P.E., Scharf, F., Friederici, A.D., Villringer, A., Sammler, D., 2016. Syntax in action has priority over movement selection in piano playing: an ERP study. J. Cogn. Neurosci. 28, 41–54.

Bressler, S.L., Menon, V., 2010. Large-scale brain networks in cognition: emerging methods and principles. Trends Cogn. Sci. 14, 277–290.

Cisek, P., 2006. Integrated neural processes for defining potential actions and deciding between them: a computational model. J. Neurosci. 26, 9761–9770.

Eerola, T., Himberg, T., Toiviainen, P., Louhivuori, J., 2006. Perceived complexity of Western and African folk melodies by Western and African listeners. Psychol. Music 34, 337–371.

Farag, C., Troiani, V., Bonner, M., Powers, C., Avants, B., Gee, J., Grossman, M., 2010. Hierarchical organization of scripts: converging evidence from fmri and frontotemporal degeneration. Cereb. Cortex 20, 2453–2463.

Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A.C., Pozzo, T., Calzolari, F., Granieri, E., Fadiga, L., 2009. Encoding of human action in Broca's area. Brain 132, 1980–1988.

Fiebach, C.J., Schubotz, R.I., 2006. Dynamic anticipatory processing of hierarchical sequential events: a common role for Broca's area and ventral premotor cortex across domains? Cortex 42, 499–502.

Fitch, W.T., Martins, M.D., 2014. Hierarchical processing in music, language, and action: Lashley revisited. Ann. N. Y. Acad. Sci. 1316, 87–104.

Friederici, A.D., 2011. The brain basis of language processing: from structure to function. Physiol. Rev. 91, 1357–1392.

Friston, K., 2010. The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11, 127–138.

Fuster, J.M., 2001. The prefrontal cortex — an update: time is of the essence. Neuron 30, 319–333.

Gallivan, J.P., Culham, J.C., 2015. Neural coding within human brain areas involved in actions. Curr. Opin. Neurobiol. 33, 141–149.

Goodale, M.a., Milner, A.D., 1992. Separate visual pathways for perception and action. Trends Neurosci. 15, 20–25.

Grafton, S.T., Hamilton, A.F.D.C., 2007. Evidence for a distributed hierarchy of action representation in the brain. Hum. Mov. Sci. 26, 590–616.

Hickok, G., Poeppel, D., 2007. The cortical organization of speech processing. Nat. Rev. Neurosci. 8, 393–402.

Hyde, K.L., Zatorre, R.J., Peretz, I., 2011. Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. Cereb. Cortex 21, 292–299.

Kilner, J.M., Friston, K.J., Frith, C.D., 2007. Predictive coding: an account of the mirror neuron system. Cogn. Process. 8, 159–166.

Kim, S.-G., Kim, J.S., Chung, C.K., 2011. The effect of conditional probability of chord progression on brain response: an MEG study. PLoS One 6, e17337.

Koechlin, E., Jubault, T., 2006. Broca's area and the hierarchical organization of human behavior. Neuron 50, 963–974.

Koechlin, E., Summerfield, C., 2007. An information theoretical approach to prefrontal executive function. Trends Cogn. Sci. 11, 229–235.

Koelsch, S., 2005. Neural substrates of processing syntax and semantics in music. Curr. Opin. Neurobiol. 15, 207–212.

Koelsch, S., 2011. Toward a neural basis of music perception — a review and updated model. Front. Psychol. 2, 110.

Koelsch, S., Siebel, W.a., 2005. Towards a neural basis of music perception. Trends Cogn. Sci. 9, 578–584.

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., Schlaug, G., 2005. Adults and children processing music: an fMRI study. NeuroImage 25, 1068–1076.

Kravitz, D.J., Saleem, K.S., Baker, C.I., Mishkin, M., 2011. A new neural framework for visuospatial processing. Nat. Rev. Neurosci. 12, 217–230.

Krumhansl, C.L., 1983. Perceptual structures for tonal music. Music. Percept. 1, 28–62.

Large, E.W., Palmer, C., 2002. Perceiving temporal regularity in music. Cogn. Sci. 26, 1–37.

Lartillot, O., Ayari, M., 2011. Cultural impact in listeners' structural understanding of a Tunisian traditional modal improvisation, studied with the help of computational models. J. Interdiscip. Music Stud. 5, 85–100.

Lerdahl, F., Jackendoff, R., 1983. A Generative Theory of Tonal Music. MIT Press, MA (Cambridge. MA: MIT Press).

Lingnau, A., Downing, P.E., 2015. The lateral occipitotemporal cortex in action. Trends Cogn. Sci. 19, 268–277.

Loui, P., 2015. A dual-stream neuroanatomy of singing. Music. Percept. 32, 232–241.

Loui, P., Wu, E.H., Wessel, D.L., Knight, R.T., 2009. A generalized mechanism for perception of pitch patterns. J. Neurosci. 29, 454–459.

Loui, P., Li, H.C., Schlaug, G., 2011. White matter integrity in right hemisphere predicts pitch-related grammar learning. NeuroImage 55, 500–507.

Maess, B., Koelsch, S., Gunter, T.C., Friederici, A.D., 2001. Musical syntax is processed in Broca's area : an MEG study. Nat. Neurosci. 4, 540–545.

Maidhof, C., Vavatzanidis, N., Prinz, W., Rieger, M., Koelsch, S., 2009. Processing expectancy violations during music performance and perception : an ERP study. J. Cogn. Neurosci. 22, 2401–2413.

Makuuchi, M., Bahlmann, J., Friederici, A.D., 2012. An approach to separating the levels of hierarchical structure building in language and mathematics. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 367, 2033–2045.

McKiernan, K.a., Kaufman, J.N., Kucera-Thompson, J., Binder, J.R., 2003. A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. J. Cogn. Neurosci. 15, 394–408.

Musso, M., Weiller, C., Horn, A., Glauche, V., Umarova, R., Hennig, J., Schneider, A., Rijntjes, M., 2015. A single dual-stream framework for syntactic computations in music and language. NeuroImage 117, 267–283.

Novembre, G., Keller, P.E., 2011. A grammar of action generates predictions in skilled musicians. Conscious. Cogn. 20, 1232–1243.

Novembre, G., Keller, P.E., 2014. A conceptual review on action-perception coupling in the musicians' brain: what is it good for? Front. Hum. Neurosci. 8, 603.

O'Reilly, J.X., Schüffelgen, U., Cuell, S.F., Behrens, T.E.J., Mars, R.B., Rushworth, M.F.S., 2013. Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. Proc. Natl. Acad. Sci. U. S. A. 110, E3660–E3669.

Palmer, C., van de Sande, C., 1995. Range of planning in music performance. J. Exp. Psychol. Hum. Percept. Perform. 21, 947–962.

Patel, A.D., 2003. Language, music, syntax and the brain. Nat. Neurosci. 6, 674–681.

Pearce, Marcus T., et al., 2010. Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. NeuroImage 50 (1), 302–313.

Pfordresher, P.Q., 2012. Musical training and the role of auditory feedback during performance. Ann. N. Y. Acad. Sci. 1252, 171–178.

Power, J.D., Schlaggar, B.L., Petersen, S.E., Power, J.D., Schlaggar, B.L., Petersen, S.E., 2015. Recent progress and outstanding issues in motion correction in resting state fMRI. NeuroImage 105, 536–551.

Rauschecker, J.P., 2011. An expanded role for the dorsal auditory pathway in sensorimotor control and integration. Hear. Res. 271, 16–25.

Rauschecker, J.P., Scott, S.K., 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat. Neurosci. 12, 718–724.

Rohrmeier, M.a., Koelsch, S., 2012. Predictive information processing in music cognition. A critical review. Int. J. Psychophysiol. 83, 164–175.

Rohrmeier, M., Rebuschat, P., 2012. Implicit learning and acquisition of music. Top. Cogn. Sci. 4, 525–553.

Ruiz, M.H., Jabusch, H.-C., Altenmüller, E., 2009. Detecting wrong notes in advance: neuronal correlates of error monitoring in pianists. Cereb. Cortex 19, 2625–2639.

Sammler, D., Koelsch, S., Friederici, A.D., 2011. Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? Cortex 47, 659–673.

Sammler, D., Koelsch, S., Ball, T., Brandt, A., Grigutsch, M., Huppertz, H.-J., Knösche, T.R., Wellmer, J., Widman, G., Elger, C.E., Friederici, A.D., Schulze-Bonhage, A., 2013a. Co-localizing linguistic and musical syntax with intracranial EEG. NeuroImage 64, 134–146.

Sammler, D., Novembre, G., Koelsch, S., Keller, P.E., 2013b. Syntax in a pianist's hand: ERP signatures of 'embodied' syntax processing in music. Cortex 49, 1325–1339.

Schon, D., Anton, J.L., Roth, M., Besson, M., 2002. An fMRI study of music sight-reading. Neuroreport 13, 2285–2289.

Seger, C.A., Spiering, B.J., Quraini, A.G.S.S.I., Alpeter, C., David, J., Thaut, M.H., 2013. Corticostriatal contributions to musical expectancy perception. J. Cogn. Neurosci. 25 (7), 1062–1077.

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Correspondence of the brain's functional architecture during activation and rest. PNAS 106, 13040–13045.

Swain, J.P., 1995. The concept of musical syntax. Music. Q. 79, 281–308.

Tenenbaum, J.B., Kemp, C., Griffiths, T.L., Goodman, N.D., 2011. How to grow a mind: statistics, structure, and abstraction. Science 331, 1279–1285.

Tillmann, B., 2012. Music and language perception: expectations, structural integration, and cognitive sequencing. Top. Cogn. Sci. 4, 568–584.

Tillmann, B., Bharucha, J.J., Bigand, E., 2000. Implicit learning of tonality: A self-organizing approach. Psychol. Rev. 107, 885–913.

Please cite this article as: Bianco, R., et al., Neural networks for harmonic structure in music perception and action, NeuroImage (2016), http://dx.doi.org/10.1016/j.neuroimage.2016.08.025

Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, a.D., von Cramon, D.Y., 2006. Cognitive priming in sung and instrumental music: activation of inferior frontal cortex. NeuroImage 31, 1771–1782.

Tymoczko, D., 2003. Local Harmonic Grammar in Western Classical Music Available at http://www.dmitri.mycpanel.princeton.edu/mozart.pdf.

Uddin, L.Q., Kelly, A.M.C., Biswal, B.B., Castellanos, F.X., Milham, M.P., 2009. Functional connectivity of default mode network components: correlation, anticorrelation, and causality. Hum. Brain Mapp. 30, 625–637.

van der Steen, M.C., Molendijk, E.B.D., Altenmueller, E., Furuya, S., 2014. Expert pianists do not listen: the expertise-dependent influence of temporal perturbation on the production of sequential movements. Neuroscience 269, 290–298.

Vuust, P., Witek, M.a.G., 2014. Rhythmic complexity and predictive coding: a novel approach to modeling rhythm and meter perception in music. Front. Psychol. 5, 1–14.

Wolpert, D.M., Flanagan, J.R., 2001. Motor prediction. Curr. Biol. 11, R729–R732.

Zarate, J.M., 2013. The neural control of singing. Front. Hum. Neurosci. 7, 237.

Zatorre, R.J., Chen, J.L., Penhune, V.B., 2007. When the brain plays music: auditory-motor interactions in music perception and production. Nat. Rev. Neurosci. 8, 547–558.

Please cite this article as: Bianco, R., et al., Neural networks for harmonic structure in music perception and action, NeuroImage (2016), http://dx.doi.org/10.1016/j.neuroimage.2016.08.025

Chapter 3

Hierarchy in action: structure-based planning and motor parameters

Syntax in Action Has Priority over Movement Selection in Piano Playing: An ERP Study

Roberta Bianco[1], Giacomo Novembre[2], Peter E. Keller[2], Florian Scharf[1], Angela D. Friederici[1], Arno Villringer[1], and Daniela Sammler[1]

Abstract

■ Complex human behavior is hierarchically organized. Whether or not syntax plays a role in this organization is currently under debate. The present ERP study uses piano performance to isolate syntactic operations in action planning and to demonstrate their priority over nonsyntactic levels of movement selection. Expert pianists were asked to execute chord progressions on a mute keyboard by copying the posture of a performing model hand shown in sequences of photos. We manipulated the final chord of each sequence in terms of Syntax (congruent/incongruent keys) and Manner (conventional/unconventional fingering), as well as the strength of its predictability by varying the length of the Context (five-chord/two-chord progressions). The production of syntactically incongruent compared to congruent chords showed a response delay that was larger in the long compared to the short context. This behavioral effect was accompanied by a centroparietal negativity in the long but not in the short context, suggesting that a syntax-based motor plan was prepared ahead. Conversely, the execution of the unconventional manner was not delayed as a function of Context and elicited an opposite electrophysiological pattern (a posterior positivity). The current data support the hypothesis that motor plans operate at the level of musical syntax and are incrementally translated to lower levels of movement selection. ■

INTRODUCTION

To facilitate everyday interactions and communication, the brain constantly screens the environment for regularities, forms predictions about upcoming events, and accordingly "pre-engages" potentially relevant neural or cognitive processes (Tenenbaum, Kemp, Griffiths, & Goodman, 2011; Bubic, von Cramon, & Schubotz, 2010; Wilson & Knoblich, 2005). This ability, which does not require deliberate effort or awareness, might be a general function shared by different cognitive domains and pivotal for survival (Perruchet & Pacton, 2006).

Understanding how simple elements are planned and perceived in temporally ordered and coherently structured sequences constitutes a central question in comparative studies across music, language, and action domains (Fitch & Martins, 2014; Tillmann, 2012). The specifically human ability of the brain to variably combine discrete meaningful units into rule-based hierarchical sequences is what is referred to as "syntactic processing" and has been defined as core aspect of language and communication (Friederici, 2011; Hauser, Chomsky, & Fitch, 2002; Lashley, 1951). Over the past years, similarities in the syntactic organization of language and Western music have been increasingly demonstrated (Rohrmeier & Koelsch, 2012; Katz & Pesetsky, 2011; Koelsch, 2005; Patel, 2003). Experimental studies have shown similar neural correlates for syntactic operations in language and music perception (Patel, Gibson, Ratner, Besson, & Holcomb, 1998; Sammler, Koelsch, & Friederici, 2011; Fedorenko, Patel, Casasanto, Winawer, & Gibson, 2009; Koelsch, Gunter, Wittfoth, & Sammler, 2005; Maess, Koelsch, Gunter, & Friederici, 2001) and also in the processing of complex action (Clerget, Winderickx, Fadiga, & Olivier, 2009; Fazio et al., 2009) inviting the hypothesis that syntactic processing might be a general "supramodal" key capability of the human brain (Fadiga, Craighero, & D'Ausilio, 2009; Slevc, Rosenberg, & Patel, 2009; Patel, 2003). Although analogies with the domain of action, in terms of hierarchical and combinatorial organization (Pulvermüller, 2014; Guerra-Filho & Aloimonos, 2012; Pastra & Aloimonos, 2012; Pulvermüller & Fadiga, 2010) remain conceptually controversial, they might be empirically tenable if shifted from actions to action planning (Moro, 2014a, 2014b). Therefore, in this study, we aimed to explore syntax-related mechanisms operating during action motor planning.

Piano performance in the Western classical music tradition provides an ideal test bed for exploring syntax in the action domain. First, playing chord progressions from this tradition is the direct motoric translation of musical syntax, a theoretically established hierarchical system of rules governing music structure (Rohrmeier, 2011). Second, it affords the possibility to investigate different hierarchical

---

[1]Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [2]Western Sydney University, Sydney, Australia

Journal of Cognitive Neuroscience 28:1, pp. 41–54
doi:10.1162/jocn_a_00873

stages in action planning (Keller, 2012; Uithol, van Rooij, Bekkering, & Haselager, 2012; Haggard, 2008; Shaffer, 1981) from lower nonsyntactic levels of movement selection to higher levels of syntax-based action plans.

Sammler, Novembre, et al. (2013) and Novembre and Keller (2011) showed that expert pianists—due to intense practice—have motorically learned syntactic regularities governing musical sequences and therefore generate motor predictions based on their acquired long-term syntactic knowledge. In a priming paradigm, expert pianists were asked to imitate silent videos of a hand-playing chord sequences. The last chord was either syntactically congruent or incongruent with the preceding musical context. Despite the absence of musical sounds, both studies revealed slower imitation times for syntactically incongruent chords as well as motor facilitation (i.e., faster responses) for the syntactically congruent chords. In terms of ERPs (Sammler, Novembre, Koelsch, & Keller, 2013), the imitation of the incongruent chords elicited an early negativity, which was associated with the perception of the syntactic violation (Koelsch, 2009), and a later posterior negativity, indexing the reprogramming (Leuthold & Jentzsch, 2002) of an anticipated motor act (i.e., the congruent chord) primed by the syntactic structure of the musical sequence. In line with models of incremental planning of serial actions (Palmer & Pfordresher, 2003), the authors argued that, during imitation of musical sequences, motor predictions of trained musicians are strongly based on long-term music-syntactic knowledge, as abstract structuring principles translate into a “grammar of musical action.”

However, piano performance not only requires the planning of which chord to play according to the preceding music-syntactic context but also the selection of a specific fingering for an optimal and smooth execution of the musical sequence. Notably, through intensive musical training, frequently occurring musical patterns (i.e., scales, chord progressions) are associated with conventional fingering configurations that are automatically activated during execution of these patterns (Gellrich & Parncutt, 1998; Sloboda, Clarke, Parncutt, & Raekallio, 1998; Clarke, Parncutt, Raekallio, & Sloboda, 1997). From this perspective, it may be suggested that motor pattern familiarity, beyond syntactic knowledge, has a role in motor predictions when playing common chord progressions. This assumption finds support in the facilitated imitation of overlearned (Koeneke, Lutz, Herwig, Ziemann, & Jäncke, 2006; Hund-Georgiadis & von Cramon, 1999) and complex actions that belong to one’s motor repertoire (Aglioti, Cesari, Romani, & Urgesi, 2008; Calvo-Merino, Glaser, Grèzes, Passingham, & Haggard, 2005). To what extent action planning operates at the level of musical syntax or at the level of common transitions of fingering configurations (i.e., the manner) is addressed here.

In the present ERP study, we aimed at untangling two hierarchical stages of musical action planning related to (i) selecting a syntax-based motor program (relative to the musical goal) versus (ii) setting the parameters of this program (the specific movement selection defining the manner of execution). Evidence for the dissociation between program selection and parameter setting has been gleaned from theoretical and empirical work. For example, the framework of “generalized motor programs” (Rosenbaum, Kenny, & Derr, 1983; Keele & Summers, 1976; Schmidt, 1975; Lashley, 1951) posits that action plans consist of core motor programs whose specific movement parameters are only chosen at the time of their use. Furthermore, it has been shown that the performance of advanced pianists is based on abstract conceptual plans and is independent of the specific movement requirements (Palmer & Meyer, 2000). Similarly, a dissociation between a more general and higher versus a more specific and lower level of action processing finds support in the “hierarchical organization of goal-directed actions” theory (see Grafton, 2009). Along these lines, the priority of the goal of an action over the means used to achieve it has been extensively demonstrated in behavioral imitation (Wohlschläger, Gattis, & Bekkering, 2003; Bekkering, Wohlschläger, & Gattis, 2000), neuroimaging (Hamilton & Grafton, 2006; Chaminade, Meltzoff, & Decety, 2002; Koski et al., 2002), and brain stimulation studies (Lago & Fernandez-del-Olmo, 2011; Cattaneo, Caruana, Jezzini, & Rizzolatti, 2009). Therefore, we transferred this hierarchical concept of action planning to music by focusing on predictions at the two levels of the motor hierarchy (goal and manner; see also Novembre & Keller, 2011). We reasoned that motor predictions of expert pianists should concern the musical goal (Syntax) rather than the finger movement selection (Manner), which should be specified only at the time of execution.

We asked expert pianists to watch and execute as fast and accurately as possible chord sequences played by a performing pianist’s hand presented in a series of photos on a computer screen. Moreover, to negate exogenously driven auditory predictive processes, no sound was used. Piano performance (RTs and errors) and an EEG were recorded. In a 2 × 2 factorial design, we manipulated the last chord of the sequences in terms of the identity of the target keys (Syntax congruent/incongruent) to address the syntactic level of action planning and in terms of fingering (Manner correct/incorrect) to address the level of movement selection. To induce different strengths of syntax/manner-based predictions, pianists were presented with five-chord or two-chord sequences (long/short Context). The execution of the long compared to the short context was expected to provide more information, hence lead to a stronger prediction of the last chord to be executed. Crucially, the manipulation of the manner, while keeping the syntax congruent and vice versa, allowed us to dissociate behavioral and neural patterns elicited by the execution of the syntactic violation (Syntax) from those triggered by a general violation of movement patterns (Manner). Additionally, the 2 × 2 factorial design permitted us to investigate syntax-related mechanisms on top of the concurrent manner violation to test whether, in musical action planning, high levels of syntactic operations are prioritized over movement parameter specification.

First, we hypothesized that motor predictions in expert pianists are driven by music-syntactic knowledge more than by motor pattern familiarity. This should be reflected in a stronger priming effect of the long context on the musical goal (Syntax) than and irrespective of the specific movement selection (Manner). Specifically, we expected the execution of the syntactically congruent/incongruent chords to be facilitated/impeded more strongly in the long than in the short contexts, whereas no such effect should occur during the execution of the manner correct/incorrect chords. Second, in terms of neural correlates, we predicted specific response-related patterns evoked by the syntax violation, different from those associated with the processing of the manner violation. To this end, we specifically focused on the syntax-related early and late negativity described by Sammler, Novembre, et al. (2013) and manner-related effects in the same time windows.

METHODS

Participants

Twenty-six pianists (16 women) aged 20–33 years (mean = 25.15; *SD* = 3.55) were included in the analysis. Eight more pianists were tested but excluded because of an insufficient number of valid trials (cutoff = 50% of valid trials). The included pianists possessed between 12 and 27 years of classical music training (mean years of training = 18.21; *SD* = 3.92) and had started to play the piano on average at 6.04 years (*SD* = 2.73). All participants were naive with regard to the purpose of the study. They gave written informed consent to take part in this experiment and received monetary compensation for participation. The study was approved by the local ethics committee.

Stimuli

Stimuli were photos showing a male pianist's right hand playing sequences of chords on the piano (Yamaha Clavinova CLP150; Yamaha Music Europe GmbH, Rellingen, Germany). To maximally address action planning processes, we used photos rather than videos (cf. Sammler, Novembre, et al., 2013; Novembre & Keller, 2011) obtaining more precise onset times of target chord presentation. In two sessions, we presented a total of 72 sequences that were all different from each other in terms of melodic contour. All sequences were composed according to the rules of classical harmony. The first chord always represented the tonic. The second chord could be tonic, mediant, or subdominant. Chords at the third position were subdominant, dominant, or dominant six–four chords, and chords at the fourth position were dominant seventh chords. At the last position, the target chord of each sequence was manipulated in terms of Syntax (syn) and Manner (man) in a 2 × 2 factorial design. Thirty-six sequences were conventional in terms of both Syntax and Manner (syn congruent/man correct), 12 were violated in terms of Syntax (syn incongruent/man correct), 12 in terms of Manner (syn congruent/man incorrect), and 12 in terms of both factors (syn incongruent/man incorrect). More precisely, syn congruent/man correct (S_cM_c) sequences ended with a tonic (a chord typically used to resolve a musical sequence) played with conventional fingering; syn incongruent/man correct (S_iM_c) sequences ended with a Neapolitan chord (a minor subdominant with a diminished sixth instead of a fifth, rarely used in classical harmony to resolve a musical sequence) played with conventional fingering; the syn congruent/man incorrect (S_cM_i) sequences ended on a syntactically congruent tonic chord but played with an unconventional fingering; and finally the double violation (syn incongruent/man incorrect, S_iM_i) was constituted by a Neapolitan chord played with an unconventional fingering. The fingering adopted by the model hand was chosen by a piano teacher with 24 years of experience conforming to the fingering taught in classical piano lessons. This was aimed to achieve smoothness and movement economy between chord transitions (i.e., 124, 125, 135, where 1 represents the thumb; 2 represents the index; and 3, 4, and 5 indicate the middle, the ring, and the little finger, respectively). Conversely, the manner manipulation of the target chord consisted of a fingering that was anatomically awkward and highly unlikely to be used (i.e., 123, 235, 245). The fingering of the last chord was rated by nine pianists on a scale from 1 (*very conventional*) to 9 (*very unconventional*). An ANOVA with the factors Syntax (congruent/incongruent) and Manner (correct/incorrect) on the mean ratings yielded a main effect of Manner [$F(1, 8) = 932.3, p < .001, \eta_p^2 = .991$], but neither main effect of Syntax [$F(1, 8) < 1, p = .721, \eta_p^2 = .017$] nor interaction of the two factors [$F(1, 8) = .206, p = .662; \eta_p^2 = .025$], confirming the motor unconventionality of the chosen fingering, and the independence of the syntax from the manner manipulation. All chords consisted of three keystrokes. Four tonalities with either two or four sharps or flats in the key signatures, that is, D, E, B♭, and A♭ major, were used with equal probability for each condition to balance the average amount of black and white keys in syntactically congruent and incongruent chords and to thus equate their visual surface structure and difficulty of execution. Red circles were superimposed on top of each pressed key for the whole duration of the photo to facilitate the recognition of the pressed keys. Sequences of two different lengths were created: five-chord sequences (long context) and two-chord sequences (short context). The two-chord sequences were identical to the last two chords of the five-chord sequences; thus, the long and short sequences differed only in the strength of the predictability of the last chord. In addition, comparing the four conditions across the two contexts allowed us to control for motor differences between conditions, that is, naturally longer hand trajectories from the penultimate to the syntactically incongruent (S_iM_c; due to different keys) and the manner incorrect (S_cM_i, S_iM_i; due to hand rotation) chords than to the not manipulated chords (S_cM_c). Finally, it should be noted that the sequence-final tonic chords naturally share a number of keys with the harmonically related context. To control for the possibility that final tonics may be merely motorically primed by the repeated use of these keys, we also allowed the final correct manner to be partly motorically primed by presenting the respective fingering on average 1.1 times in the context. The balanced repetition of tonic key configuration (1.3 times) and final chord fingering, that is, the similar likelihood of motor priming, discloses any effect that occurs in the syntax but not manner conditions as related to syntactic (not motor) priming. Each sequence started with 2-sec presentation of a preparatory photo showing a stationary hand poised to press the three keys associated with the first chord. Then the following photos were presented at a rate of 2 sec per photo (total duration: 12-sec long sequences, 6-sec short sequences).

Procedure

Participants were seated in front of a MIDI (musical instrument digital interface) piano (Yamaha Clavinova CLP150) and watched the photo sequences on a computer monitor (100-Hz refresh rate). Simultaneously, they were required to execute the chords they saw, one by one, with their right hand on the piano, as quickly and accurately as possible, both in terms of the keys (syntax) and in terms of fingering (manner). Note that the piano was muted and no sound was presented with the photos, that is, the experiment took place in total absence of musical sounds. Each trial started with a visual fixation cross of 0.5-sec duration and ended with a black screen for 1.5 sec after the final photo of the stimulus sequence (Figure 1).

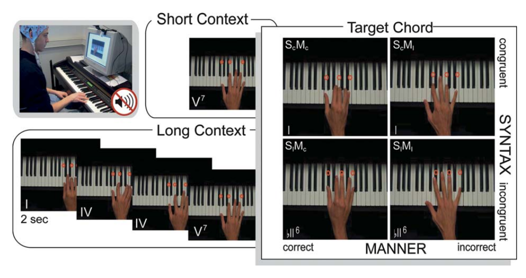


Figure 1. Experimental design. In the total absence of musical sound, pianists executed chord progressions with their right hand by copying (as fast and accurately as possible) the posture of a performing model hand shown in sequences of photos. The target chord of each progression was manipulated in terms of keys (congruent/incongruent Syntax) and fingering (conventional/unconventional Manner) in a 2 × 2 factorial design and was presented at the end of a five- or two-chord sequence (long/short Context) to induce different strengths of predictability.

Participants were invited for two sessions in which they were presented with the same stimuli. Each session consisted of six experimental blocks, one for each of the three violation conditions (S_iM_c, S_cM_i, S_iM_i) and separately for long and short sequences. Each block contained a total of 48 trials: 24 nonviolated trials (S_cM_c) intermixed with 24 trials of the respective violation condition. Block order was counterbalanced across participants and alternated between blocks with long and short sequences. To acquaint participants with the unusual and challenging task and increase accuracy, in the first session, they all received a training of 24 trials (50% nonviolated) for all the violation conditions in both long and short context. Different tonalities were used for the training (G, F, D♭, and B major) than in the main experiment.

At the end of the experiment, participants filled out a questionnaire to assess how much they relied on auditory imagery, motor imagery, and/or theoretical knowledge of western harmony to do the task. Their piano expertise was estimated as the sum of training hours per day across all years of piano lessons.

Stimulus presentation and response registration were controlled by Presentation software (Version 14.9, Neurobehavioral System, Inc., Berkeley, CA). Through a (custom-built) MIDI interface, the MIDI piano key values were converted into a serial signal compatible with Presentation software. This allowed us to compute the RTs of the keystrokes in relation to the onset of the target chord photo. Moreover, a video camera placed above the keyboard recorded the pianist's hand from an aerial view, allowing us to detect (offline) trials in which the pianists used a different fingering from that shown in the photos.

EEG Data Acquisition

The EEG recordings were acquired from 61 Ag/AgCl electrodes (Fpz, Fp1, Fp2, AFz, AF3, AF4, AF7, AF8, Fz, F3, F4, F5, F6, F7, F8, F1, FC2, FCz, FC3, FC4, FC5, FC6, FT7, FT8, FC1, F2, Cz, C1, C2, C3, C4, C5, C6, T7, T8, CPz, CP3, CP4, CP5, CP6, TP7, TP8, P1, P2, Pz, P3, P4, P5, P6, P7, P8, CP1, CP2, POz, PO3, PO4, PO7, PO8, O1, O2, Oz) according to the international 10–20 system (Sharbrough et al., 1991). The left mastoid (M1) served as reference. Three additional electrodes were placed on the sternum as common ground, on the right mastoid bone (M2), and on the tip of the nose for offline re-referencing. The EOG was recorded by two bipolar montages, one with electrodes located above and below the left eye and the other with two electrodes placed on the outer canthus of each eye. Signals were amplified using a 24-bit Brainvision QuickAmp 72 amplifier (Brain Products GmbH, Gilching, Germany) with input impedance below 5 kΩ and digitized at a 500-Hz sampling rate.

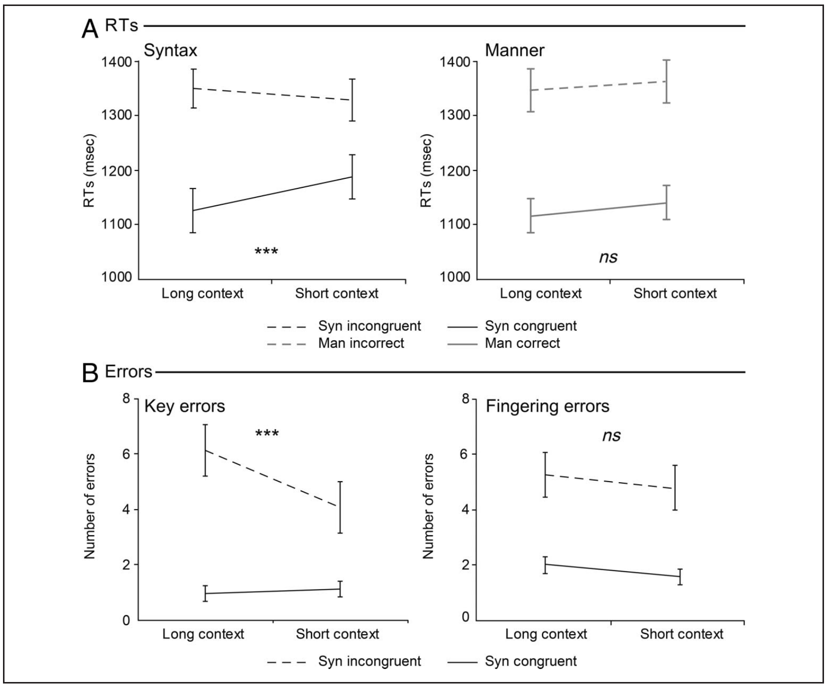


Figure 2. (A) Mean RTs during imitation of syntactically incongruent (dashed line) and congruent chords (solid line; left) and during imitation of manner incorrect (dashed line) and correct chords (solid line; right) in the long and short context. (B) Number of key errors (left) and fingering errors (right) during imitation of syntactically incongruent and congruent chords in the long and short context. Error bars indicate 1 *SEM*. $^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$.

Behavioral Data Analysis

RTs and execution errors of the last chord of each trial were analyzed in accordance with Novembre and Keller (2011). The minimum requirement for including a participant's data into the analysis was correct responses on 50% of trials. Trials were considered valid when three conditions were satisfied: (1) both the last and the second last chord had to be correctly imitated in terms of keys and fingering (for the error analysis we included the trials correctly imitated in the second last but incorrectly in the last chord), (2) the keystrokes within a chord had to be synchronous (i.e., no more than 150 msec should intervene between the first and the last of the three keystrokes), and (3) mean RTs of the three keystrokes after the onset of the target chord photo had to stay within 3000 msec (cf. Drost, Rieger, Brass, Gunter, & Prinz, 2005). RTs were calculated by subtracting the time of execution of the target chord (i.e., mean of the times of three keystrokes composing the chord) from the onset time of the last photo showing the target chord. The fingering performed by each pianist was analyzed through offline inspection of the video recordings in which the fingers employed by the participants were compared with the fingers presented in the stimulus photos. Statistical evaluation of the RT data was done using three-way repeated-measures ANOVA with the factors Syntax (congruent/incongruent), Manner (correct/incorrect), and Context (long/short). Errors were analyzed with an analogous ANOVA, but with the additional within-subject factor Error type (key/fingering error).

EEG Data Analysis

Data analysis was carried out using EEGLAB toolbox 9.01 (Delorme & Makeig, 2004) implemented in MATLAB 7.7. The EEG data were offline re-referenced to the algebraic mean of the mastoids and were 0.3-Hz high-pass filtered (fir, 5854 points, Blackman window). Strong muscle artifacts, electrode drifts, or technical artifacts were manually rejected. Independent component analysis was used for linear decomposition of the continuous data to remove the contributions of artifact sources (slow drifts, eye blink/movement, and muscle artifacts) on the scalp sensors. After 25-Hz low-pass filtering (fir, 110 points, Blackman window), epochs of −200 to 1500 msec, time-locked to the onset of the photo of the last chord, were extracted from the data. Epochs were rejected whenever signal voltages exceeded ±80 μV in one or more electrodes. Nonrejected trials were averaged separately for each condition and baseline-corrected (−200 msec before the onset of the target photo). Only correct trials according to the behavioral analysis were included in the ERP statistical analysis (i.e., mean number of trials ± *SD* for S_cM_c, S_iM_c, S_cM_i, S_iM_i in the long context: 122.9 ± 10, 38.1 ± 5.4, 37.7 ± 4.8, 35.8 ± 5.5; in the short context: 131.3 ± 6, 41.3 ± 4.4, 40.4 ± 4.7, 38.3 ± 6.2).

Effects of Syntax, Manner, and Context were analyzed time-locked to the onset of the last (target) photo of the sequence. Mean amplitudes were computed separately for each condition over nine ROIs and for three specific time windows. The ROIs comprised (i) left anterior (F3, F5, F7, FC3, FC5, FT7, AF3), (ii) left central (C3, C5, T7, CP3, CP5, TP7), (iii) left posterior (P3, P5, P7, PO3, PO7), (iv) middle anterior (F1, FZ, F2, FC1, FCZ, FC2, AFZ), (v) middle central (C1, CZ, C2, CP1, CPZ, CP2), (vi) middle posterior (P1, PZ, P2, POZ), (vii) right anterior (F4, F6, F8, FC4, FC6, FT8, AF4), (viii) right central (C4, C6, T8, CP4, CP6, TP8), (ix) right posterior (P4, P6, P8, PO4, PO8). Three time windows (i) from 210 to 520 msec, (ii) from 520 to 800 msec, and (iii) from 800 to 1200 msec were selected by visual inspection of the ERPs and topography plots shown in Figures 2 and 3. As objective and external criteria independent from the data, borders of the time windows were set to time points at which either a change in polarity or in topography was found, assuming that different map topographies and polarities directly indicate different underlying generators, that is, different cognitive processes (Michel et al., 2004). The same time windows were used in the analysis of the two conditions. Statistical analysis of mean amplitude values was carried out by means of five-way ANOVAs with the repeated-measures factors Syntax (congruent/incongruent) × Manner (correct/incorrect) × Context (long/short) × Laterality (left/middle/right) × AntPost (anterior/central/posterior), separately for each time window.

RESULTS

Behavioral Data

RTs

Statistical values of the 2 × 2 × 2 repeated-measures ANOVA with the factors Syntax (congruent/incongruent), Manner (correct/incorrect), and Context (long/short) are reported in Table 1. These results revealed main effects of Syntax and Manner, indicating that imitation of the syntactically incongruent as well as manner violated chords was generally slower compared to chords that contained no such violations. No main effect of Context was found, showing that RTs for the imitation of target chords was comparable between long and short sequences. Notably, a highly significant Syntax × Context interaction showed that more in the long than in the short context the execution of syntactically congruent chords was faster compared to incongruent chords. Conversely, no interaction between Manner × Context was found (Figure 2A). This finding suggests that the harmonic structure of the musical context rather than the motor pattern familiarity drove the motoric prediction of the target chord and that the

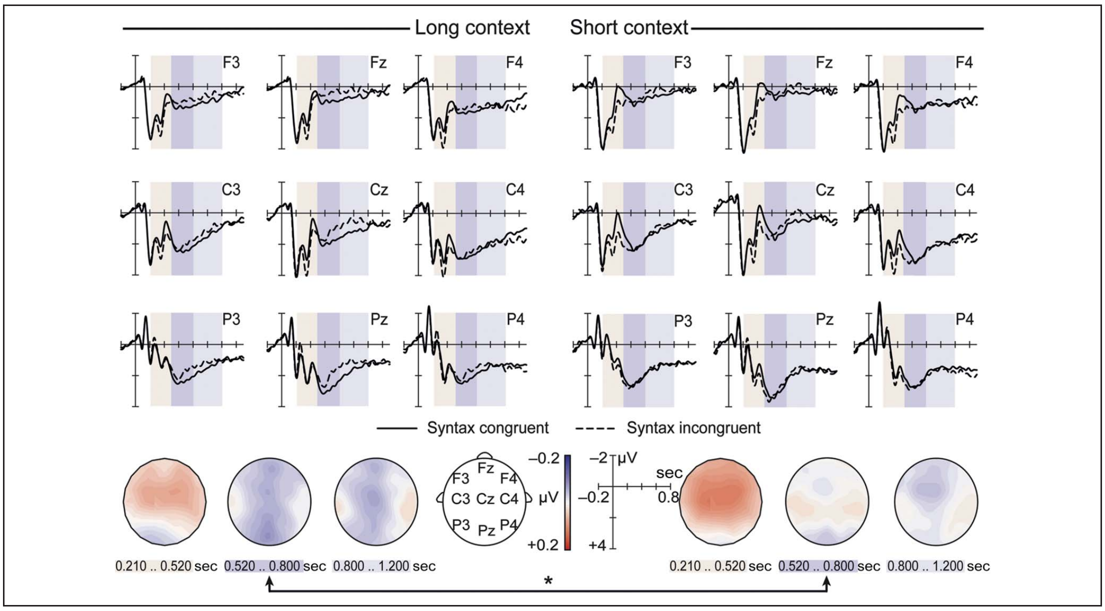


Figure 3. Effect of Syntax. ERPs evoked by syntactically incongruent (dotted line) compared to congruent (solid line) chords in the long (left) and short (right) context across all trials. The three time windows are shaded according to their polarity (red for positivity, blue for negativity). Topography maps for each statistical time window (lower row) depict the difference potentials of syntactically incongruent minus congruent chords (arrows below indicate the interaction between Syntax and Context). ***$p < .001$, **$p < .01$, *$p < .05$.

Table 1. Results of the ANOVA on RTs with the Factors Syntax × Manner × Context

Effect	*df*	*F*	*p*	η_p^2
S	1, 25	**80.16**	**.000**	**.762**
M	1, 25	**133.65**	**.000**	**.842**
C	1, 25	1.522	.229	.057
S × C	1, 25	**52.56**	**.000**	**.678**
M × C	1, 25	.604	.445	.024
S × M × C	1, 25	**4.78**	**.038**	**.160**

Bold values indicate significant results ($p < .05$). Partial eta squared: $\eta_p^2 > .5$, large effect size; $\eta_p^2 > .3$, medium effect size; $\eta_p^2 \leq .1$, small effect size (Bortz & Döring, 2003). S = Syntax; M = Manner; C = Context.

prediction concerned the musical goal (Syntax) rather than the movement selection (Manner).

A three-way interaction of Syntax × Manner × Context suggested a reciprocal influence of syntax and manner processing in relation to the Context. We calculated separate ANOVAs for the manner correct and incorrect trials with the factors Syntax and Context and for the syntax congruent and incongruent trials with the factors Manner and Context. This analysis yielded a Syntax × Context interaction in both manner correct [$F(1, 25) = 71.99, p < .001, \eta_p^2 = .742$] and, although weaker, in the manner incorrect trials [$F(1, 25) = 5.98, p = .022, \eta_p^2 = .193$], whereas a Manner × Context interaction was found only in the syntax congruent trials [$F(1, 25) = 4.505, p = .044, \eta_p^2 = .153$] and not in the syntax incongruent trials [$F(1, 25) = 1.649, p = .211, \eta_p^2 = .062$]. This indicates that while the syntactic prediction effect (Syntax × Context interaction), although weaker, held in presence of the concurrent manner violation, the manner was facilitated (Manner × Context) only when the pianists' syntactic prediction was fulfilled (syntax congruent trials). These data suggest that the syntax of the context primes primarily the motor program of the musical goal, which in turn may trigger information about optimal movement parameters for its execution. In other words, movement selection is facilitated only when the higher plan on the musical goal is confirmed.

Error Analysis

Errors in terms of pressed keys and fingering were counted separately. Trials with both error types were excluded from the analysis (as in Novembre & Keller, 2011). Key and fingering errors are assumed to reflect distinct cognitive processes associated with the musical goal (Syntax) and the specific movement used in the execution (Manner), respectively.

A 2 × 2 × 2 × 2 repeated-measures ANOVA with the factors Syntax, Manner, Context, and Error type (key/fingering errors; for statistical values, see Table 2) revealed that—overall—less errors were committed during execution of syntactically congruent compared to incongruent chords (main effect of Syntax) as well as during the imitation of manner correct compared to incorrect chords (main effect of Manner). Also, less errors were committed in the short compared to the long context (main effect of Context), whereas there was no significant difference between number of key and fingering mistakes (no significant main effect of Error type). The interactions of Syntax × Error type and Manner × Error type revealed that Syntax and Manner conditions were associated with greater amount of key and fingering errors, respectively. Importantly, a Syntax × Context × Error type interaction indicated that key errors, but not fingering errors, were more prevalent in the long than in the short context during the execution of syntactically incongruent chords, irrespective of the manner. Indeed, follow-up ANOVAs with the factors Syntax and Context, calculated separately for the key and the fingering errors (Figure 2B), yielded a significant Syntax × Context interaction for the key errors [$F(1, 25) = 7.164, p = .013, \eta_p^2 = .223$], but not for the fingering errors [$F(1, 25) = 2.599, p = .122, \eta_p^2 = .093$]. These data indicate that the harmonic structure of the context strongly affected the motor program of the musical goal, irrespective of the specific movement selection. Additionally, similar to what was observed in the RTs, we found that during the execution of the manner incorrect chords more fingering errors were committed in the long than in the short context but only when the syntax was congruent [Manner × Context interaction on the fingering errors across syntactically congruent trials: $F(1, 25) = 9.120, p = .006, \eta_p^2 = .267$; across syntactically incongruent trials: $F(1, 25) = 1.161, p = .292, \eta_p^2 = .044$]. This finding confirms that the selection of which fingers to use was facilitated in the long context only when the musical goal matched the (syntactic) predictions.

Table 2. Results of the ANOVA on Number of Errors with the Factors Syntax × Manner × Context × Error Type

Effect	*df*	*F*	*p*	η_p^2
S	1, 25	**39.896**	**.000**	**.615**
M	1, 25	**20.907**	**.000**	**.455**
C	1, 25	**5.779**	**.024**	**.188**
Et	1, 25	3.276	.082	.116
S × Et	1, 25	**9.028**	**.006**	**.265**
M × Et	1, 25	**32.146**	**.000**	**.563**
S × C × Et	1, 25	**8.868**	**.006**	**.262**

Bold values indicate significant results ($p < .05$). Partial eta squared: $\eta_p^2 > .5$, large effect size; $\eta_p^2 > .3$, medium effect size; $\eta_p^2 \leq .1$, small effect size (Bortz & Döring, 2003). S = Syntax; M = Manner; C = Context; Et = Error type.

Table 3. Results of the ANOVAs with the Factors Syntax × Manner × Context × Laterality × AntPost for Each Time Window

		1st tw: 210…520 msec			*2nd tw: 520…800 msec*			*3rd tw: 800…1200 msec*		
Effect	*df*	*F*	*p*	η_p^2	*F*	*p*	η_p^2	*F*	*p*	η_p^2
Effect of Syntax										
S	1, 25	**5.733**	**.024**	**.187**	1.019	.322	.039	<1	.352	.035
S × C	1, 25	1.578	.221	.059	1.673	.208	.063	<1	.650	.008
S × L	2, 50	<1	.417	.033	**3.965**	**.029**	**.137**	**10.894**	**.000**	.304
S × L × C	2, 50	1.481	.238	.056	1.082	.159	.073	<1	.475	.029
S × A	2, 50	**8.225**	**.007**	**.248**	1.543	.227	.058	1.017	.369	.039
S × A × C	2, 50	<1	.912	.001	<1	.835	.003	<1	.557	.023
S × A × L	4, 100	2.015	.120	.075	<1	.660	.021	2.977	.051	.103
S × A × L × C	4, 100	1.083	.360	.042	**2.886**	**.035**	**.103**	1.685	.179	.063
Effect of Manner										
M	1, 25	1.630	.213	.061	**29.014**	**.000**	**.537**	<1	.402	.028
M × C	1, 25	<1	.487	.008	1.715	.202	.064	4.002	.056	.138
M × L	2, 50	**24.202**	**.000**	**.492**	**11.486**	**.000**	**.315**	**9.401**	**.001**	**.273**
M × L × C	2, 50	1.064	.347	.041	<1	.520	.025	1.401	.256	.053
M × A	2, 50	**9.592**	**.003**	**.277**	2.004	.168	.074	<1	.425	.028
M × A × C	2, 50	**10.279**	**.003**	**.291**	**7.833**	**.005**	**.239**	1.230	.283	.047
M × A × L	4, 100	**9.387**	**.000**	**.273**	**8.595**	**.000**	**.256**	**3.601**	**.024**	**.126**
M × A × L × C	4, 100	1.165	.328	.045	1.552	.205	.058	1.558	.210	.059
Syntax and Manner Interaction										
S × M	1, 25	1.164	.291	.044	<1	.869	.001	<1	.964	.000
S × M × C	1, 25	<1	.584	.012	1.586	.219	.060	3.780	.063	.131
S × M × C × L	2, 50	<1	.926	.002	<1	.419	.033	3.039	.068	.108
S × M × A	2, 50	1.962	.172	.073	**6.614**	**.012**	**.209**	<1	.358	.035
S × M × C × A	2, 50	<1	.714	.006	<1	.594	.012	1.128	.304	.043
S × M × L × A	4, 100	<1	.672	.021	1.124	.347	.043	**2.778**	**.039**	**.100**
S × M × L × A × C	4, 100	1.137	.340	.044	1.633	.187	.061	1.249	.297	.048

Bold values indicate the effects due to the difference in strength of potentials. Partial eta squared: $\eta_p^2 > .5$, large effect size; $\eta_p^2 > .3$, medium effect size; $\eta_p^2 \leq .1$, small effect size (Bortz & Döring, 2003). S = Syntax; M = Manner; C = Context; L = Laterality; A = AntPost.

EEG Data

We were interested in distinguishing neurophysiological correlates of (i) the higher level of syntax-based motor programming (prediction of the musical goal, reflected by a Syntax × Context interaction) and (ii) the lower level of specific movement selection (prediction of a conventional optimal movement, reflected by a Manner × Context interaction). To this end, we analyzed (Table 3) the effects of (i) Syntax (Figure 3), (ii) Manner (Figure 4), and (iii) their interaction separately in three time windows (see EEG Data Analysis). Nonparametric cluster-based permutation tests with standard parameters in Fieldtrip (www.fieldtriptoolbox.org) revealed qualitatively similar results to those described below.

The analysis of the effects of Syntax and Manner revealed different neural signatures between 520 and 800 msec, in line with our hypothesis that the planning of a musical goal (Syntax) and the specific movements (Manner) rely on different mechanisms.

Effect of Syntax

In the first time window (210–520 msec), syntactically incongruent compared to congruent chords elicited a positivity with a central scalp distribution as revealed by a significant main effect of Syntax and an interaction of Syntax × AntPost. This early effect did not differ between long and short context (i.e., no interaction of Syntax × Context), suggesting that it was not related to prediction. Most relevant, in the second time window (520–800 msec), syntactically incongruent compared to congruent chords evoked a centroparietal negativity that was present in the long and not in the short context. The four-way ANOVA showed a significant interaction of Syntax × Context × AntPost × Laterality demonstrating that the negativity was particularly strong in the long context and more enhanced in the midline central regions. In the third time window (800–1200 msec), the four-way ANOVA yielded an interaction of Syntax × Laterality, as well as a marginally significant interaction of Syntax × AntPost × Laterality, indicating a predominantly middle-central negativity that did not differ in amplitude between long and short context.

To evaluate whether the negativity between 520 and 800 msec in the long context was influenced by auditory imagery, the difference wave in the middle central ROI (mean = −0.516 ± 1.384 μV) was correlated with the subjective ratings of the extent to which participants actively imagined the sound of the up-coming chord during performance. No significant correlation was found [$r(25) = .261, p = .301, R^2 = .046$].

Effect of Manner

In the first time window (210–520 msec), a significant interaction of Manner × AntPost × Laterality revealed a left middle anterior positivity elicited by the manner incorrect compared with manner correct chords across all trials. This positivity was stronger in the long compared to the short context, as shown by a significant interaction of Manner × Context × AntPost. Follow-up ANOVAs with the factors Manner × Context calculated for each ROI confirmed a left middle anterior distribution of this effect [Manner × Context interaction, middle anterior: $F(1, 25) = 7.920, p = .009$; left anterior: $F(1, 25) = 1.188, p = .027$; right anterior: $F(1, 25) = 3.793, p = .063$; all $ps > .116$ in the other ROIs]. In the second time window (520–800 msec), a main effect of Manner indicated that manner incorrect chords elicited more positive potentials than manner correct chords with a predominately middle to left centroparietal distribution, as confirmed by a Manner × AntPost × Laterality interaction. A Manner × Context × AntPost interaction showed that the positivity in the short context did not extend as far anteriorly as in the long context. This difference in scalp distribution was confirmed by significant Manner × Context interactions in the anterior regions as revealed by follow-up ANOVAs with the factors

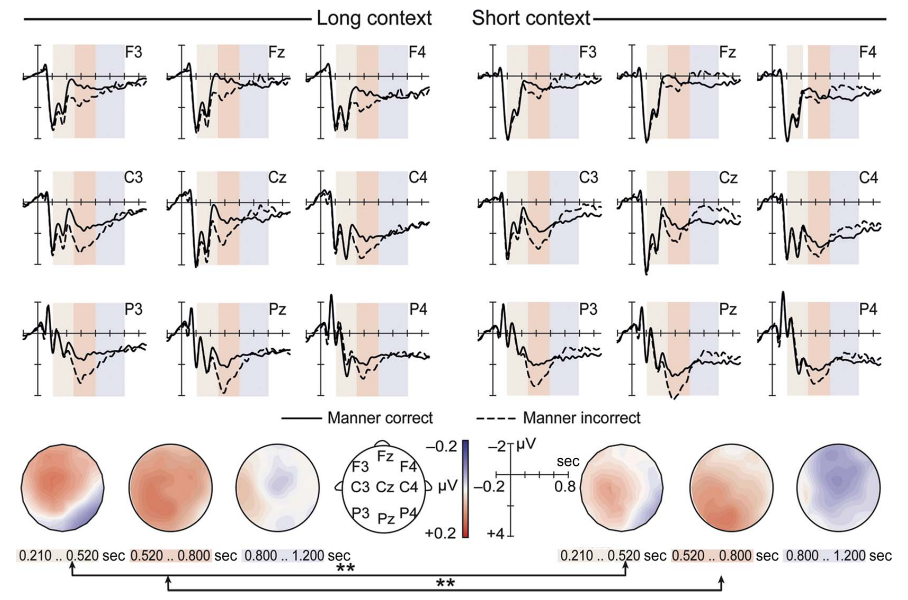


Figure 4. Effect of Manner. ERPs evoked by target chords played with incorrect (dotted line) compared to correct (solid line) manner in the long (left) and short (right) context across all trials. Time windows of the three time windows are shaded according to their polarity (red for positivity, blue for negativity). Topography maps for each statistical time window (lower row) depict the difference potentials of manner incorrect minus correct chords (arrows below indicate the interaction between Manner and Context). ***$p < .001$, **$p < .01$, *$p < .05$.

Manner × Context calculated for each ROI [middle anterior: $F(1, 25) = 7.813, p = .010$; right anterior: $F(1, 25) = 10.884, p = .003$; left anterior: $F(1, 25) = 3.844, p = .061$; all $ps > .161$ in the other ROIs]. In the third time window (800–1200 msec), an interaction of Manner × AntPost × Laterality indicated a middle centrally distributed negativity elicited by the manner incorrect compared to the manner correct chords that tended to be larger in the short compared to the long context [Manner × Context: $F(1, 25) = 4.002, p = .056$].

Interaction Effects

Finally, we analyzed how far the effects of Syntax and Manner described above reciprocally interact. To this end, we tested for interactions that involved the factors Syntax × Manner × Context (and any topographical factor). No such interactions were found in any of the three time windows (see Table 3). Consequently, no further split of the general linear model was performed.

DISCUSSION

Action plans are hierarchically organized, with higher levels representing the general goal of an action and lower levels concerning the specific movements required to realize the goal (Uithol et al., 2012). This study aimed to differentiate action planning based on higher-order syntactic structures (Syntax) from lower nonsyntactic processes of movement selection (Manner) in expert pianists. Therefore, behavioral and neural indices of motor prediction were examined during the execution of chords that contained either a syntax or a manner violation and that were primed by long or short musical contexts.

We found (i) a strong context-dependent priming effect on the execution of syntactic violations (RTs and errors), indicating that plans of musical goals are made ahead according to the musical context. Crucially, (ii) no contextual priming was observed during the execution of manner violations, unless the syntax was congruent. In line with models of action hierarchy (Grafton & Hamilton, 2007), this suggests a priority of planning the goal of the musical action (Syntax) that in turn can prime the selection of the optimal movement parameters (Manner). Finally, (iii) different electrophysiological signals were elicited by the syntactically incongruent chords (centroparietal negativity) and manner incorrect chords (posterior positivity). These signatures may represent the different levels of action planning, pertaining to higher levels of syntax-based motor plans versus lower levels of movement parameter setting, respectively.

Behavior

In line with previous findings (Sammler, Novembre, et al., 2013; Novembre & Keller, 2011), syntactically incongruent chords were executed more slowly and evoked more key mistakes than congruent chords, particularly when primed by a long musical context. Conversely, execution times and number of fingering mistakes were commensurately higher in manner incorrect than correct chords, irrespective of context length. This pattern of results not only excludes an interpretation in terms of mere motor priming (which should have led to similar context effects in syntax and manner; see Methods) but indicates that particularly the syntactic structure of the musical context narrows down the probabilities of chord transitions, thus leading the pianists to (motorically) anticipate the execution of the most likely harmonically coherent chord (Syntax). By contrast, the specific movement parameters seem to be far less strongly determined by the preceding context, despite high familiarity with the types of chord progressions employed in the paradigm. Notably, the context dependency of syntactically incongruent chords was observed irrespective of whether the concurrent manner was correct or incorrect, although it was stronger in the former due to movement familiarity. This shows that, in experts, distal goals are the main drivers of motor predictions regardless of how the goal is realized. At the same time, our data further suggest that the preplanned goal tends to prime the selection of optimal movement parameters required to achieve the goal. We ground this assumption on the observation that manner incorrect chords showed context sensitivity both in terms of execution time and number of fingering errors, but exclusively in syntactically congruent trials (i.e., when the preplanned goal was valid). This suggests that the specific movement selection constitutes a late stage of motor preparation dependent on the action plan concerning the more distal musical goal.

In conclusion, motor predictions concerning the musical goal, prior to the manner, are consistent with the framework of "generalized motor programs" (for a review, see Summers & Anson, 2009), as knowledge structures allow a given class of movements to be executed in different ways, depending on underlying parameter settings. Furthermore, it is reminiscent of imitation studies showing a hierarchical organization of action in which the action goal is prioritized over the short-term selected movements (Wohlschläger et al., 2003; Bekkering et al., 2000).

ERPs

The execution of the syntax and manner violations elicited different electrophysiological patterns between 520 and 800 msec: We found that the syntax violations evoked a centroparietal negativity in the long and not in the short context (similar to Sammler, Novembre, et al., 2013) whereas the execution of the manner violations elicited a positivity with left posterior scalp distribution. In line with integrated models of hierarchical organized motor plans (Grafton & Hamilton, 2007), we claim that this distinction speaks in favor of a motor program level coding for the goal structure of an action (Syntax) and a lower motor level for computing the coordinated movement to a goal (Manner).

More specifically, the centroparietal negativity elicited by the syntax violation was modulated by the length of the musical context and thus matched the context-dependent effects in RTs and errors. This suggests a response-related nature of the negativity, which may be interpreted as a signal of high-level movement reprogramming following the cancellation of the prepotent response in face of the incongruity to be executed (Sammler, Novembre, et al., 2013; Leuthold & Jentzsch, 2002). Importantly, this interpretation implies that the motor program for a structurally coherent musical goal was present at the moment of the target chord presentation, as it had been preplanned based on the syntactic context. Obviously, longer contexts lead to stronger syntax-based predictions of the musical goals that in turn require more effort in their revision (larger amplitude of the negativity, longer RTs, and a higher number of key errors in the long than in the short context). In a later time window (800–1200 msec), a late centrally distributed negativity, which resembles the contingent negative variation (Walter, Cooper, Aldridge, MacCallum, & Winter, 1964), was elicited in all conditions similarly in the two contexts. This time window immediately preceded and partly overlapped with the execution of the final chord and might reflect the computation of muscle-specific commands common to all conditions (Cunnington, Windischberger, Deecke, & Moser, 2003; Rektor, 2000). Indeed, the contingent negative variation is typically elicited before motor responses and specifically in the final stage of response preparation of externally cued movements (Smith, Johnstone, & Barry, 2006).

Crucially, manner violations did not evoke a centroparietal negativity between 520 and 800 msec, as opposed to syntax violations. This discloses the negativity (in line with the behavioral results) as related to syntactic processes not motor priming (see Methods). Furthermore, if one accepts the idea that this negativity reflects the reprogramming of a preplanned motor response, its absence in the manner violations implies that the specific movement for execution had not been programmed at the time of the target chord presentation. This interpretation would be in line with the assumption that the musical goal is planned before movement selection.

Instead, the manner violations elicited a left posterior positivity in the time window between 520 and 800 msec in both contexts, speaking in favor of a different nature of the syntax- and manner-related cognitive processes. This effect was preceded by an earlier positivity (210–520 msec) with anterior scalp distribution. Together, these potentials resemble the P300 complex composed of P3a and P3b, typically elicited by infrequent behaviorally relevant stimuli (Gómez, Flores, Digiacomo, Ledesma, & González-Rosa, 2008) and modulated in amplitude by the probability of the deviant target (Duncan-Johnson & Donchin, 1977, 1982). Interestingly, both the early and late positivities were stronger in the long than in the short context, revealing a context-dependent effect that obviously mismatches with the behavioral data. One explanation might be that, unlike the response-related ERPs in the syntax condition, these positivities rather reflect stimulus-related processes that are contingent on the different sequential probabilities of the manner violation in long and short stimulus sequences. More precisely, given that the manner violation can be recognized as odd even in single photos (see Methods), its occurrence probability amounts to 10% in task blocks with long sequences and 25% in task blocks with short sequences. It should be noted that the same reasoning does not apply to the syntax violations that were only recognizable as part of the sequence. This amounts to an equal occurrence probability of 50% in both long and short sequences and should not lead to amplitude differences. In line with this, the detection of the syntax violations indeed evoked an early positivity (a P3a) that did not differ between long/short contexts (for similar results, see also Sammler, Novembre, et al., 2013). Altogether, the perceptual detection of both syntax- and manner-related violations elicited a P3a; however, rather than a motor reprogramming phase as observed in the syntax violation (centroparietal negativity between 520 and 800 msec), the salient fingering manipulation evoked a following P3b that might reflect memory updating processes dependent on the behaviorally relevant stimulus (Polich, 2007).

As a final remark, we did not find a syntax-related early anterior negativity as is usually evoked by music-syntactic violations in the auditory domain (i.e., an ERAN; Koelsch, 2009; Koelsch, Gunter, Friederici, & Schröger, 2000) and as was found in our previous study, in which chord progressions were presented as videos (Sammler, Novembre, et al., 2013). This suggests that the early anterior negativities might be specifically tied (i) to the auditory detection of music-syntactic irregularities (but see Gunter, Schmidt, & Besson, 2003) and/or (ii) to the perceptual continuity of the musical input as present in dynamic auditory and video streams but less so in discrete photo series. The comparison of music-syntactic processing in perception and production and the potential beneficial effect of real motion on harmonic priming are interesting topics for future research.

Conclusion

In line with the notion of action hierarchy, we distinguished syntax-related motor programs operating at high levels of action planning from lower levels of specific movement selection. Using a priming paradigm involving the execution of chord progressions, we showed that expert pianists make motor predictions concerning the musical goal (Syntax) rather than the manner of execution (Manner). Building on previous findings (Sammler, Novembre, et al., 2013; Novembre & Keller, 2011), our results provide further evidence for motor planning based on long-term music-syntactic knowledge (i.e., a grammar of action based on musical harmony) and for the priority of the motor plan related to the distal goal over the specific ways to achieve it. Although our EEG data do not give specific information on the generators underlying syntax and manner processing, their different electrophysiological and behavioral patterns may indicate different mechanisms in the planning of the musical goal and the movement used to achieve it. We suggest that, during production of musical sequences, motor predictions of the musical goal are driven by the harmonic structure of the musical context recognized through internalized syntactic knowledge of pianists. Critically, given a certain predictable context, the motor program of the distal musical goal might operate at high levels of the action control hierarchy and be incrementally translated to lower levels of movement kinematics at the very late stage of motor preparation. This weighing of action features (i.e., a weak, thus flexible, preselection of the optimal movement associated to the goal) would constitute an advantage in terms of more efficient performance and interactions with unexpected external changes. Finally, the notion that, through years of intensive motor practice, syntactic rules are motorically acquired, that is, a translation of musical syntax into a "grammar of action," might speak for a training-dependent motor plasticity toward an emergent syntax-based motor control. Whether this phenomenon occurs in other human actions associated with syntactic structures, such as speech, is an intriguing prospect for future investigations.

Acknowledgments

We are grateful to the pianists who participated in this experiment and to B. Pace who lent his hand for the stimulus preparation. We also thank S. Gutekunst for technical support. This research was funded by Max Planck Society.

Reprint requests should be sent to Roberta Bianco, Max Planck Institute for Human Cognitive and Brain Sciences, Stephanstraße 1a, 04103 Leipzig, Germany, or via e-mail: bianco@cbs.mpg.de.

REFERENCES

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience, 11,* 1109–1116.

Bekkering, H., Wohlschläger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology, 53,* 153–164.

Bortz, J. J., & Doering, N. (2003). Forschungsmethoden und Evaluation: Fuer Human- und Sozialwissenschaftler. Berlin: Springer.

Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience, 4,* 25.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15,* 1243–1249.

Cattaneo, L., Caruana, F., Jezzini, A., & Rizzolatti, G. (2009). Representation of goal and movements without overt motor behavior in the human motor cortex: A transcranial magnetic stimulation study. *Journal of Neuroscience, 29,* 11134–11138.

Chaminade, T., Meltzoff, A. N., & Decety, J. (2002). Does the end justify the means? A PET exploration of the mechanisms involved in human imitation. *Neuroimage, 15,* 318–328.

Clarke, E., Parncutt, R., Raekallio, M., & Sloboda, J. (1997). Talking fingers: An interview study of pianists' views on fingering. *Musicae Scientiae, 1,* 87–107.

Clerget, E., Winderickx, A., Fadiga, L., & Olivier, E. (2009). Role of Broca's area in encoding sequential human actions: A virtual lesion study. *NeuroReport, 20,* 1496–1499.

Cunnington, R., Windischberger, C., Deecke, L., & Moser, E. (2003). The preparation and readiness for voluntary movement: A high-field event-related fMRI study of the Bereitschafts-BOLD response. *Neuroimage, 20,* 404–412.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134,* 9–21.

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005). When hearing turns into playing: Movement induction by auditory stimuli in pianists. *Quarterly Journal of Experimental Psychology Section A, 58,* 1376–1389.

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology, 14,* 456–467.

Duncan-Johnson, C. C., & Donchin, E. (1982). The P300 component of the event-related brain potential as an index of information processing. *Biological Psychology, 14,* 1–52.

Fadiga, L., Craighero, L., & D'Ausilio, A. (2009). Broca's area in language, action, and music. *Annals of the New York Academy of Sciences, 1169,* 448–458.

Fazio, P., Cantagallo, A., Craighero, L., D'Ausilio, A., Roy, A. C., Pozzo, T., et al. (2009). Encoding of human action in Broca's area. *Brain, 132,* 1980–1988.

Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: Evidence for a shared system. *Memory & Cognition, 37,* 1–9.

Fitch, W. T., & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences, 1316,* 87–104.

Friederici, A. D. (2011). The brain basis of language processing: From structure to function. *Physiological Reviews, 91,* 1357–1392.

Gellrich, M., & Parncutt, R. (1998). Piano technique and fingering in the eighteenth and nineteenth centuries: Bringing a forgotten method back to life. *British Journal of Music Education, 15,* 5–23.

Gómez, C. M., Flores, A., Digiacomo, M. R., Ledesma, A., & González-Rosa, J. (2008). P3a and P3b components associated to the neurocognitive evaluation of invalidly cued targets. *Neuroscience Letters, 430,* 181–185.

Grafton, S. T. (2009). Embodied cognition and the simulation of action to understand others. *Annals of the New York Academy of Sciences, 1156,* 97–117.

Grafton, S. T., & Hamilton, A. F. D. C. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science, 26,* 590–616.

Guerra-Filho, G., & Aloimonos, Y. (2012). The syntax of human actions and interactions. *Journal of Neurolinguistics, 25,* 500–514.

Gunter, T. C., Schmidt, B. H., & Besson, M. (2003). Let's face the music: A behavioral and electrophysiological exploration of score reading. *Psychophysiology, 40,* 742–751.

Haggard, P. (2008). Human volition: Towards a neuroscience of will. *Nature Reviews Neuroscience, 9,* 934–946.

Hamilton, A. F. D. C., & Grafton, S. T. (2006). Goal representation in human anterior intraparietal sulcus. *Journal of Neuroscience, 26,* 1133–1137.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science, 298,* 1569–1579.

Hund-Georgiadis, M., & von Cramon, D. Y. (1999). Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Experimental Brain Research, 125,* 417–425.

Katz, J., & Pesetsky, D. (2011). The identity thesis of language and music. ling.auf.net/lingBuzz/000959.

Keele, S. W., & Summers, J. J. (1976). The structure of motor programs. In G. E. Stelmach (Ed.), *Motor control: Issues and trends* (pp. 109–142). New York: Academic Press.

Keller, P. E. (2012). What movement force reveals about cognitive processes in music performance. In A. Mornell (Ed.), *Art in motion II* (pp. 115–153). Frankfurt, Germany: Peter Lang.

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology, 15,* 207–212.

Koelsch, S. (2009). Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. *Psychophysiology, 46,* 179–190.

Koelsch, S., Gunter, T., Friederici, A. D., & Schröger, E. (2000). Brain indices of music processing: "Nonmusicians" are musical. *Journal of Cognitive Neuroscience, 12,* 520–541.

Koelsch, S., Gunther, T. C., Wittfoth, M., & Sammler, D. (2005). Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience, 17,* 1565–1577.

Koeneke, S., Lutz, K., Herwig, U., Ziemann, U., & Jäncke, L. (2006). Extensive training of elementary finger tapping movements changes the pattern of motor cortex excitability. *Experimental Brain Research, 174,* 199–209.

Koski, L., Wohlschläger, A., Bekkering, H., Woods, R. P., Dubeau, M.-C., Mazziotta, J. C., et al. (2002). Modulation of motor and premotor activity during imitation of target-directed actions. *Cerebral Cortex, 12,* 847–855.

Lago, A., & Fernandez-del-Olmo, M. (2011). Movement observation specifies motor programs activated by the action observed objective. *Neuroscience Letters, 493,* 102–106.

Lashley, K. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–131). New York: Wiley.

Leuthold, H., & Jentzsch, I. (2002). Spatiotemporal source localisation reveals involvement of medial premotor areas in movement reprogramming. *Experimental Brain Research, 144,* 178–188.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience, 4,* 540–545.

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., & Grave de Peralta, R. (2004). EEG source imaging. *Clinical Neurophysiology, 115,* 2195–2222.

Moro, A. (2014a). On the similarity between syntax and actions. *Trends in Cognitive Sciences, 18,* 109–110.

Moro, A. (2014b). Response to Pulvermuller: The syntax of actions and other metaphors. *Trends in Cognitive Sciences, 18,* 221.

Novembre, G., & Keller, P. E. (2011). A grammar of action generates predictions in skilled musicians. *Consciousness and Cognition, 20,* 1232–1243.

Palmer, C., & Meyer, R. K. (2000). Conceptual and motor learning in music performance. *Psychological Science, 11,* 63–68.

Palmer, C., & Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review, 110,* 683–712.

Pastra, K., & Aloimonos, Y. (2012). The minimalist grammar of action. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 367,* 103–117.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6,* 674–681.

Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10,* 717–733.

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10,* 233–238.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118,* 2128–2148.

Pulvermüller, F. (2014). The syntax of action. *Trends in Cognitive Sciences, 18,* 219–220.

Pulvermüller, F., & Fadiga, L. (2010). Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience, 11,* 351–360.

Rektor, I (2000). Parallel information processing in motor systems: Intracerebral recordings of readiness potential and CNV in human subjects. *Neural Plasticity, 7,* 65–72.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music, 5,* 35–53.

Rohrmeier, M. A., & Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology, 83,* 164–175.

Rosenbaum, D. A., Kenny, S. B., & Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 86–102.

Sammler, D., Koelsch, S., Ball, T., Brandt, A., Grigutsch, M., Huppertz, H. J., et al. (2013). Co-localizing linguistic and musical syntax with intracranial EEG. *Neuroimage, 64,* 134–146.

Sammler, D., Koelsch, S., & Friederici, A. D. (2011). Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? *Cortex, 47,* 659–673.

Sammler, D., Novembre, G., Koelsch, S., & Keller, P. E. (2013). Syntax in a pianist's hand: ERP signatures of "embodied" syntax processing in music. *Cortex, 49,* 1325–1339.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82,* 225–260.

Shaffer, L. H. (1981). Performances of Chopin, Bach, and Bartok: Studies in motor programming. *Cognitive Psychology, 13,* 326–376.

Sharbrough, F., Chatrian, G.-E., Lesser, R. P., Lueders, H., Nuwer, M., & Picton, T. W. (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology, 8,* 200–202.

Slevc, L. R., Rosenberg, J. C., & Patel, A. D. (2009). Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin & Review, 16,* 374–381.

Sloboda, J. A., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of finger choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 185–203.

Smith, J. L., Johnstone, S. J., & Barry, R. J. (2006). Effects of pre-stimulus processing on subsequent events in a warned go/nogo paradigm: Response preparation, execution and inhibition. *International Journal of Psychophysiology, 61,* 121–133.

Summers, J. J., & Anson, J. G. (2009). Current status of the motor program: Revisited. *Human Movement Science, 28,* 566–577.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science, 331,* 1279–1285.

Tillmann, B. (2012). Music and language perception: Expectations, structural integration, and cognitive sequencing. *Topics in Cognitive Science, 4,* 568–584.

Uithol, S., van Rooij, I., Bekkering, H., & Haselager, P. (2012). Hierarchies in action and motor control. *Journal of Cognitive Neuroscience, 24,* 1077–1086.

Walter, W. G., Cooper, R., Aldridge, V. J., MacCallum, W. C., & Winter, A. L. (1964). Contingent negative variation: An electric sign of sensori-motor association and expectancy in the human brain. *Nature, 203,* 380–384.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin, 131,* 460–473.

Wohlschläger, A., Gattis, M., & Bekkering, H. (2003). Action generation and action perception in imitation: An instance of the ideomotor principle. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 358,* 501–515.

Chapter 4

Differences of jazz and classical genre on action hierarchy

Genre-specific EEG signatures of musical action planning in classical and jazz pianists

Bianco R.[1], Novembre G.[2], Keller, P. E.[2], Villringer A.[1], Sammler D.[1]

[1] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[2] The MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Sydney, Australia

Abstract

Common sequential actions, from talking to playing music, require flexible action planning at multiple levels. Relatively high-level plans specify the appropriate ordering of discrete acts (such as words or chords) into coherent sequences while lower-level plans specify optimal movement parameters for execution of each single act. In this EEG study we tested whether and how musicians adaptively tune these levels of action planning to master genre-specific demands. Classical and jazz pianists executed 5- and 2-chord progressions (without sound) in a real-time imitation task that assessed (i) structure-based planning of the action sequence by manipulating its conformity with Western tonal harmony (congruent/incongruent), and (ii) parameter specification of single acts by manipulating the fingering used to play (conventional/unconventional). Beyond similar hierarchical core principles of action planning in both groups, behavioural, event-related potential and spectral power measures show (i) that jazz pianists were more flexible in revising high-level structure-based plans in face of harmonic incongruities, while (ii) classical pianists were more accurate to set fingering parameters to structural features. Overall, we show that core processes of action planning can be shaped by the specific musical genre that pianists master, which paves the way to understanding the exceptionality of individual performance.

Keywords: music production, motor system, musical expertise, event-related potentials, oscillations

Introduction

When considering common actions such as talking or drawing, we all perform similar movements of high complexity that nonetheless maintain our particular individual style. Traditionally, motor control research has addressed the complexity of human movements by decomposing a task into its fundamental elements that can then be studied in isolation as *simple acts* (e.g., reaching) (Wolpert, Diedrichsen, & Flanagan, 2011). However, how far can this approach advance our understanding of *complex sequential actions* such as speaking or playing the piano, culminating in outstanding performances like Martha Argerich's magical interpretation of "Gaspard de la Nuit", or Keith Jarrett's tour de force live improvisation of "The Köln Concert"? Moreover, to what extent can current models of action control explain the differential abilities required for instance to play a concerto or to spontaneously improvise, to recite a poem or to freely chat on the phone? Complex sequential behaviours involve multilayer action control processes. At the heart of these processes (i) discrete acts must be appropriately ordered into *sequences*, and (ii) optimal movement parameters should be specified for each *single act* within the sequence. Despite coherent programs of research on these two levels of action planning, they are often investigated independently rather than in integrated scenarios of complex sequential actions, and they are usually regarded as fixed. Whether and how action control may be tuned by habitual action focus adopted to meet the demands of specific performance conditions remains currently unknown.

The present study takes music production as an experimental framework to investigate the multilayer organisation of complex sequential action planning (Uithol, van Rooij, Bekkering, & Haselager, 2012) based on structural knowledge of music. The study particularly investigates how different constituent layers of the action hierarchy are modulated in performers with different action planning demands in their daily practice as illustrated in Figure 1. During performance, musicians rely on their long-term knowledge of musical (e.g., harmonic) structure and the unfolding musical context to (i) build structure-based plans of the *action sequence* they are about to produce (Clarke, 2001). These high-level plans contain information about the structural relationship and ordering of elements in a sequence, based on which forthcoming musical acts are incrementally determined, e.g., the next chord, the "what". At the same time, each *single act* of the sequence needs to be motorically implemented by specifying optimal movement parameters (Verwey, Shea, & Wright, 2015), e.g., appropriate fingering, the "how". We propose that the calibration of these two processes can vary depending on the demands that different types of practice make on the performer, e.g., structure-generative vs. structure-interpretative focus of action as adopted in jazz and classical performance, respectively. Before presenting empirical evidence for this hypothesis, we will first present an overview of prevailing theories of hierarchical action control that focus either on the level of appropriate action sequencing or the implementation of simple acts. Then we will explain how these two levels may be modulated by genre-specific demands and how their experience-dependent plasticity can be comprehensively tested in music production.

Theories of action control agree upon the notion of hierarchical action organisation, yet typically focus on different levels of action planning. One set of theories addresses the planning of *complex sequential actions* with emphasis on how their internal hierarchical structure determines the appropriate ordering of the constituent acts (Lashley, 1951; Schmidt, 1975; Koechlin & Summerfield, 2007; Uithol et al., 2012). The cognitive reality of such structure-based planning has been evidenced in speech (Dell, 1986) and music performance (Palmer & Pfordresher, 2003; Palmer & van de Sande, 1995) in that production errors are not arbitrary: for example, substitution errors are more likely to occur between two non-adjacent but structurally related elements (e.g., nouns) than between two adjacent but structurally unrelated elements (e.g., "writing a mother to my letter" instead of "writing a letter to my mother"). Another set of theories addresses the hierarchical implementation of *single acts* by discriminating between the planning of the goal (the "what") and the subsequent specification of movement parameters (the "how") (Schmidt, 1975; Wohlschläger, Gattis, & Bekkering, 2003; Grafton & Hamilton, 2007; Rosenbaum, Cohen, Jax, Weiss, & van der Wel, 2007; Verwey et al., 2015). Evidence for this distinction comes from behavioural (Bekkering, Wohlschläger, & Gattis, 2000; Rosenbaum et al., 2007) and electrophysiological studies (Fogassi et al., 2005; Cattaneo et al., 2007) showing that action goals (e.g., to drink from a cup or to place it away) determine the choice of optimal movement parameters (e.g., the hand position to grasp the cup by the handle or the rim, respectively). Musical actions can be conceptualised as encompassing both these frameworks: high-level structural plans, e.g., based on harmonic structure, determine the order and identity of single acts, e.g., chords to play ("what"), whose movement parameters are specified at lower levels, e.g., in terms of fingers to use for execution ("how"; see Figure 1).

Evidence for at least two levels of action planning in music production comes from recent behavioural and electrophysiological studies that tested pianists while playing unrehearsed chord sequences without auditory feedback (Novembre & Keller, 2011; Sammler, Novembre, Koelsch, & Keller, 2013; Bianco et al., 2016). Concerning the higher-order sequence level, these studies focused on how action planning in expert classical pianists is influenced by their long-term knowledge of harmony — defining the arrangement of chords into well-structured musical sequences (Swain, 1995). Results showed that, throughout the continuously unfolding musical sequence, pianists construct and maintain representations of harmonic structures that allow them to anticipate structurally pertinent chords, even when playing without auditory feedback. Pianists were found to be faster and more accurate when playing harmonically predictable chords, while they showed behavioural costs and a late negativity in event-related potentials (ERPs) during execution of chords that mismatched the anticipated harmonic structure of the sequence. Notably, these effects were stronger when target chords were preceded by long rather than short musical contexts, demonstrating that — with sufficient harmonic context — knowledge of harmony allows musicians to build *structure-based plans* for the appropriate sequencing of chords. Concerning lower single act level, these studies demonstrated distinctive behavioural (Novembre & Keller, 2011) and EEG signatures (Bianco et al., 2016) associated with goal-related planning (the "what"; context-dependent response time costs and a late negativity) and *specification of movement parameters* of single acts (the "how"; context-independent response time costs and a late positivity) — here defined as the fingering employed during music performance (Shaffer, 1980; Repp, 2000; Clarke, 2001). Altogether, these results provide first neurophysiological evidence for multilayer action control processes in music production by dissociating structure-based planning of musical sequences (i.e., chords within a larger harmonic context) and subsequent parameter specification of single acts (i.e., fingering applied to perform the chord).

An interesting question that has remained unanswered so far concerns how the dynamics between the levels of the action control hierarchy can be modulated by other factors, such as action tendencies developed through specific training styles. For example, playing a classical concerto or improvising jazz makes different demands on the performer and requires focus on different levels of music production, even if both classical and jazz music traditions are similarly rooted in Western tonal harmony (Johnson-Laird, 2002). In fact, jazz musicians (more than classical musicians) adopt a *structure-generative focus* when playing: they particularly focus on building musical sequences in real-time and turning "mistakes" (unexpected sounds) into viable music (Pressing, 1984; Beaty, 2015). These skills not only require perfect mastery of musical conventions and (harmonic) rules in order to be able to creatively deviate from them (Johnson-Laird, 2002); they also imply constant awareness of possible structural alternatives that the musician embraces or revokes more or less on the fly, as explicated by models of improvisation (Pressing, 1987). Structure-generative focus may particularly train the ability to build flexible long-range structural connections between musical elements and foster instantaneous reassessment of structure-based plans in case harmonic novelties must be integrated into on-going performance. In contrast, classical musicians (more than jazz musicians) adopt a *structure-interpretative focus* when playing: they specifically focus on the range of possible expressive features to be applied to the musical structure, which is usually fixed by the composer (Shaffer, 1984). These features include, among others (see Keller, 2012), the choice of particular fingerings (e.g., the thumb being stronger and better suited for accentuation; Parncutt, 2014), which is usually a crucial component in the preparation of classical performance and requires intensive practice of fingering technique (Gellrich & Parncutt, 1998). Through such experience, classical musicians might develop the ability to select the optimal fingering on the spot, even when the music is unrehearsed as in sight-reading. This implies a rapid inference of the most likely forthcoming tone or chord from structural cues, to immediately proceed to expressive stages of action planning (Clarke, 2001; Chaffin, Lemieux, & Colleen, 2007). In other words, structure-interpretative focus may lead musicians to plan ahead along the most probable structural-harmonic line, thus freeing resources for specifying movement parameters, i.e., the manner of appropriately playing a certain element based on its structural properties.

The present EEG study seeks neural evidence for the specialised timing, weighting and dynamics within the motor control hierarchy, which might reflect differences in planning strategies depending on long-term tendencies developed through practice. To do so, we compared EEG and behavioural data from pianists who typically practice with a *structure-generative* (jazz) vs. *structure-interpretative* focus (classical). Both groups of pianists were tested with the same real-time music production task (see Figure 2) addressing two hierarchical stages of action planning, i.e., structure-based planning at sequence level and parameter specification at single act level (Bianco et al., 2016). In complete absence of sound, pianists were required to

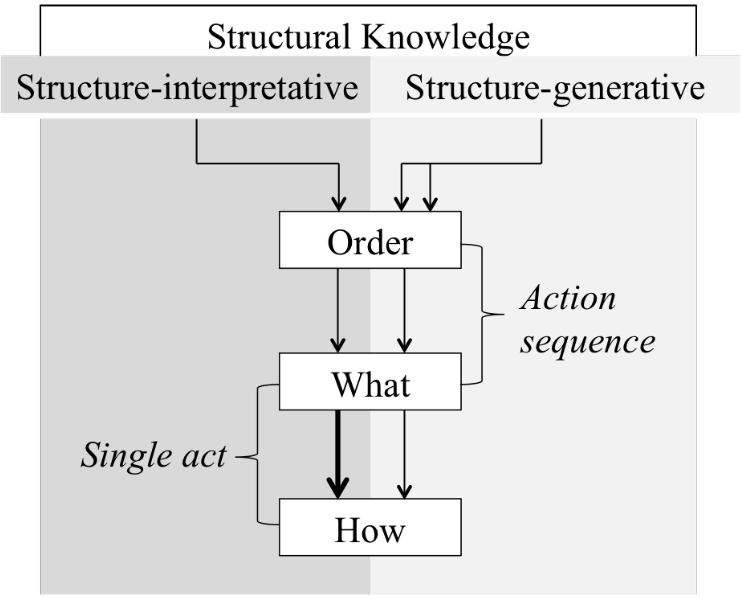


Figure 1: Multiple layers of action planning and influence of action focus. High-level structure-based plans of action sequences are formed based on structural (e.g., harmonic) knowledge and the unfolding musical context. These plans determine the appropriate ordering ("Order") and consequently the identity ("What") of single acts in a sequence. Lower-level parameters (e.g., fingering) are specified at later stages of action planning. These parameters determine the optimal motoric implementation ("How") of the single act. This multilayer organisation of actions may be tuned by (i) structure-generative (jazz) or (ii) structure-interpretative (classical) action focus: (i) requires constant awareness of structural alternatives at the level of the action sequence (as indicated by double arrows), while (ii) requires straightforward planning of the most probable order of the action sequence (single arrow) for rapid assignment of relevant motor parameters of the single act (bold arrow).

.

watch and execute chord sequences played by a performing pianist's hand presented in a series of photos on a computer screen (Figure 2). Absence of sound was chosen (as in previous studies) to focus on the cognitive-motor aspects of the task and to eliminate exogenously driven auditory predictive processes. To address the level of *ordering of the action sequence*, we (i) manipulated the harmonic structure of the chord sequences by rendering the final chord harmonically (in)congruent with the preceding musical context (congruent/incongruent Harmony). We further manipulated the predictability of the final chord by placing it at the end of 5- or 2-chord sequences (long/short Context). Given that long compared to short sequences provide more information for harmonic structure building, the combined manipulation of Harmony and Context allowed us to indirectly probe the build-up of structure-based plans, and to directly examine pianists' reassessment of these plans in face of a harmonic violation. To address the level of *single act implementation*, (ii) the final chord was manipulated not only in terms of chord to play (Harmony; i.e. what to play) but also in terms of fingering used for execution (conventional/unconventional Manner; i.e. how to play). Differences between the signals evoked by manipulations of Harmony and Manner should be indicative of two dissociated levels in the action plan of the single act. Because movement parameters are related to single acts and are specified at late stages of action planning, Manner related effects should occur similarly in both long and short sequences, in line with previous results (Bianco et al., 2016).

In sum, we hypothesised that core hierarchical processes at (i) action sequence and (ii) single act levels would be reflected (i) in context-*dependent* behavioural (response time and key errors) and ERP effects (late negativity) for the processing of harmony as relevant indices of structure-based plan building/reassessment, and (ii) context-*independent* behavioural (response times and fingering errors) and ERP effects (positivity) for the processing of manner as relevant indices of low-level movement parameter setting of single acts. Similar context-dependent and -independent effects were expected in spectral power in task-relevant neural frequency bands for the processing of harmony or manner violations. Regarding the group comparison, we expected that (i) classical and jazz pianists should both show similar indices of a hierarchical core structure of action planning — at the levels of action sequence and single act. However, (ii) if *structure-generative focus* increases awareness of structural alternatives at the level of the action sequence, then jazz compared to classical pianists should show less conflict and greater flexibility in the processing and revision of their motor plans when responding to the harmonically unexpected chord, particularly in the long context; (iii) if *structure-interpretative focus* leads pianists to rely on the most likely structure-plan to rapidly proceed to the setting of manner of execution, then classical compared to jazz pianists should show greater conflict in response to harmonic novelties, yet greater sensitivity to setting appropriate fingering parameters.

Methods

Participants.

Fifteen classical pianists (classical group, CG, 11 females) and 15 jazz pianists (jazz group, JG, 1 female) gave informed consent to participate in the study. Classical pianists were a subset of the participants in (Bianco et al., 2016). All pianists had received formal training at music academies with focus on the classical or jazz genre, e.g., the Hochschule für Musik und Theater "Felix Mendelssohn-Bartholdy" in Leipzig. Data from classical pianists were reanalysed from Bianco et al. (2016) and were selected to match the jazz pianists in the following criteria: The two groups were comparable in age (mean age ± SEM of CG: 25.5 ± 1 years; JG: 25.7 ± 1.3 years; $t(28) = 0.123$, $p > .903$), total accumulated hours of piano training across their life (CG: 11886 ± 1621 hours; JG: 11485 ± 1387 hours; $t(28)$ = -0.188, p > .852), and onset of piano playing (age of onset, CG: 6.7 ± 0.7 years; JG: 9.0 ± 1.0 years; $t(28)$ = 1.88, p > .071). All pianists had at least 6 years of musical training. When they were tested, all participants were similarly active as performers, as revealed by the comparison of average practice hours per week over the past year (CG: 8.9 hours; JG: 13.3 hours; $t(28)$ = 1.19, p > .245). The criteria to qualify as a classical pianist was to have no jazz or improvisation experience, while a minimum of 2 years of jazz piano training was required to be assigned to the jazz group. The JG had in fact more accumulated training hours in jazz than in classical piano (in JG: 7202 ± 954 jazz training hours and 4684 ± 621 classical training hours; $t(28)$ = 2.78, p < .014). On a scale from 1 to 9, the JG reported to practice music more often without reading from scores than the CG (CG: 4.8 ± 0.6; JG: 6.5 ± 0.5; $t(28)$ = 2.22, p < .035), although both groups reported comparable abilities to read scores (CG: 5.73 ± 0.7; JG: 4.26 ± 0.6; $t(28)$ = -1.72, p > .100). All participants were naïve with regard to the purpose of the study and received monetary compensation for participation. The local ethics committee of the University of Leipzig approved the study.

Stimuli.

Stimuli were the same as those used in (Bianco et al., 2016), i.e., photos showing a male pianist's right hand playing sequences of chords on the piano (Yamaha Clavinova CLP150). All participants were presented with 72 sequences that were composed according to the rules of classical harmony in four tonalities with either two or four sharps or flats, i.e., D, E, B♭, and A♭ major. Sequences consisted of chords of three keystrokes each and differed in melodic contour. The last chord of each sequence was manipulated in terms of Harmony (har: a congruent Tonic chord vs. an incongruent Neapolitan chord) and/or Manner (man: conventional fingering - i.e., 124, 125, 135, vs. unconventional fingering - i.e., 123, 235, 245, where 1 represents the thumb, 2 the index and 3, 4 and 5 the middle, the ring and the little finger, respectively) according to a 2 × 2 factorial design (for a more detailed description of the stimuli, see Bianco et al., 2016). 36 sequences were conventional in terms of both Harmony and Manner (har congruent/man correct: H_cM_c), 12 contained violations in terms of Harmony (har incongruent/man correct: H_iM_c), 12 in terms of Manner but not Harmony (har congruent/man incorrect: H_cM_i), and 12 in terms of both factors (har incongruent/man incorrect: H_iM_i). The predictability of the last chord was manipulated by placing it either at the end of five-chord sequences (long context for high predictability), or two-chord sequences (short context for low predictability). The two-chord sequences were identical with the last two chords of the five-chord sequences. Hence, comparing the 4 conditions across the two contexts allowed us both to measure context-dependent predictions of the last chord and to control for motoric differences, i.e., longer movement trajectories from penultimate to incongruent / unconventional chords than in the congruent / conventional conditions. The total duration of the long and short sequences was 12 and 6 sec, respectively, including a 2 sec preparatory photo showing a hand about to press the first chord, followed by photos of the 5- or 2-chord sequences presented at a rate of 2 sec per photo. Each trial started with a visual fixation cross of 0.5 sec and ended with a black screen of 1.5 sec after the final photo of the stimulus sequence. To facilitate recognition of the relevant keys and fingers, red circles were superimposed on top of each pressed key for the whole duration of the photo.

Procedure.

We adopted the same experimental procedure as (Bianco et al., 2016). Pianists were asked to watch the photo series of the model hand on a computer monitor (100 Hz refresh rate) and to simultaneously execute the chords with their right hand on a MIDI (musical instrument digital interface) piano (Yamaha Clavinova CLP150, Yamaha Music Europe GmbH, Rellingen, Germany) (see Figure 2).

Note that the piano was muted and no sound was presented with the photos. Pianists were instructed to execute the chords one by one, as quickly and accurately as possible, both in terms of the keys pressed (Harmony) and in terms of fingering (Manner). The experiment consisted of 6 experimental blocks. Each block contained 24 non-violated trials (H_cM_c) intermixed with 24 trials of one of the violation conditions (H_iM_c, H_cM_i, or H_iM_i), separately for long and short sequences. Block order was counterbalanced across participants and alternated between blocks with long and short sequences. To increase the number of trials, each pianist participated in two sessions with the same stimuli and the same block order. In order to acquaint parti-

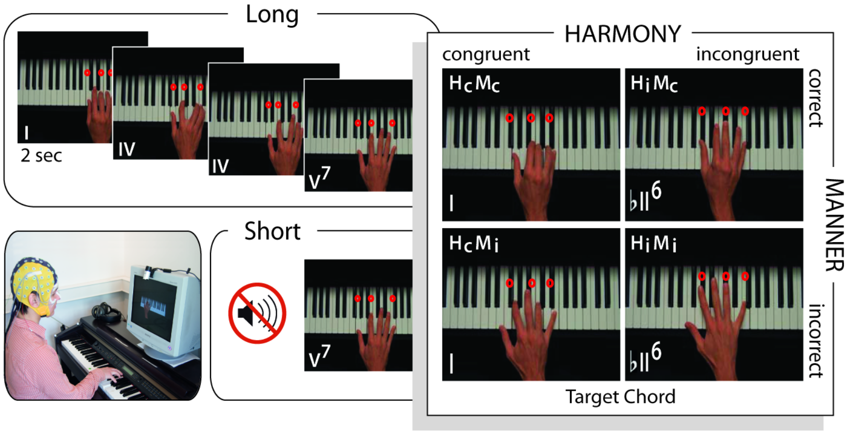


Figure 2: Experimental design. In absence of sound, pianists executed chord progressions by imitating a performing hand presented in series of photos. The sequences were manipulated in their last chord in terms of Harmony (congruent/incongruent) and Manner (correct/incorrect). Furthermore, the length of the Context (long/short) manipulated the overall predictability of the last chord.

.

cipants with the task, the first session started with 6 short blocks of 24 practice trials each (12 H_cM_c and 12 H_iM_c, H_cM_i, or H_iM_i in long or short context) in tonalities that were not used in the main experiment (G, F, D♭, and B major). Stimulus presentation and response registration were controlled by Presentation software (Version 14.9, Neurobehavioural System, Inc.). MIDI piano key values were converted into a serial signal compatible with Presentation software through a custom-built MIDI interface. Participants' right hands were filmed with a video camera placed above the keyboard for (off-line) analysis of their fingering accuracy.

EEG data acquisition.

EEG recordings were acquired from 61 Ag/AgCl electrodes (Fpz, Fp1, Fp2, AFz, AF3, AF4, AF7, AF8, Fz, F3, F4, F5, F6, F7, F8, F1, FC2, FCz, FC3, FC4, FC5, FC6, FT7, FT8, FC1, F2, Cz, C1, C2, C3, C4, C5, C6, T7, T8, CPz, CP3, CP4, CP5, CP6, TP7, TP8, P1, P2, Pz, P3, P4, P5, P6, P7, P8, CP1, CP2, POz, PO3, PO4, PO7, PO8, O1, O2, Oz) placed according to the international 10-20 system (Sharbrough et al., 1991). The left mastoid (M1) served as reference. Three additional electrodes were placed on the sternum as common ground, on the right mastoid bone (M2), and on the tip of the nose for off-line re-referencing. Vertical and horizontal EOG was recorded from two bipolar montages, one with electrodes located above and below the left eye, the other with two electrodes placed on the outer canthus of each eye. Signals were amplified using a 24-bit Brainvision QuickAmp 72 amplifier (Brain Products GmbH, Gilching, Germany) with input impedances below 5 kΩ and digitised at a 500 Hz sampling rate.

Behavioural data analysis.
Response times (RTs) and execution errors of the last chord of valid trials were analysed as in previous studies (Novembre & Keller, 2011; Sammler et al., 2013; Bianco et al., 2016). Trials were considered valid when 1) no key and fingering mistakes occurred in the last and the second last chords, 2) the keystrokes within a chord were synchronous (i.e., no more than 150 msec elapsed between the first and the last of the 3 keystrokes) and 3) RTs were below 3000 msec (cf. Drost, Rieger, Brass, Gunter, & Prinz, 2005). RTs (i.e., average RT of the three keystrokes) were time-locked to the onset of the photo showing the last chord. Fingering of the participants was analysed through off-line inspection of the video recordings of their hands. Statistical evaluation of the RT data was done using four-way analyses of variance (ANOVAs) with the repeated-measures factors Harmony (congruent/incongruent), Context (long/short), Manner (correct/incorrect), and the between-subjects factor Group (classical/jazz). Errors were analysed with two analogous ANOVAs, separately for key and fingering errors.

EEG data analysis.

Pre-processing. EEG data were pre-processed using EEGLAB 9 (Delorme & Makeig, 2004) implemented in MATLAB 7.4. Data were first re-referenced to the algebraic mean of the mastoids and filtered with a 0.3-Hz high-pass filter (fir, 5854 points, Blackman window). Then, electrode drifts, strong muscle and technical artefacts were manually rejected and data were entered into an Independent Component Analysis (ICA) for parcelling out the contribution of artefacts such as eye movements, continuous tension of muscles, slow drifts and technical noise. After 45-Hz low pass filtering (fir, 810 points, Blackman window), epochs were extracted from behaviourally correct trials between -900 and 2200 ms relative to the onset of the last (target) chord. Only trials with signal voltages within ± 60 μV at all electrodes were included in further analyses and averaged separately for each condition. 21% and 38% of the total number of trials were discarded in the classical (CG) and jazz group (JG), respectively. To ensure that group differences cannot be attributed to trial count or signal-to-noise ratio, we randomly eliminated further trials in CG to equate them with the trial numbers of the JG in each condition.

Event-related potentials. For each participant and in each condition, event-related potentials (ERPs) were computed from -200 to 1200 ms relative to the onset of the target photo and baseline correction was performed (baseline: -200 to 0 ms). For statistical analysis, mean amplitudes were extracted separately for each condition from 9 regions of interest (ROIs) in specific time-windows (see below). The ROIs comprised: (i) left anterior (F3, F5, F7, FC3, FC5, FT7, AF3), (ii) left central (C3, C5, T7, CP3, CP5, TP7), (iii) left posterior (P3, P5, P7, PO3, PO7), (iv) middle anterior (F1, FZ, F2, FC1, FCZ, FC2, AFZ), (v) middle central (C1, CZ, C2, CP1, CPZ, CP2), (vi) middle posterior (P1, PZ, P2, POZ), (vii) right anterior (F4, F6, F8, FC4, FC6, FT8, AF4), (viii) right central (C4, C6, T8, CP4, CP6, TP8), (ix) right posterior (P4, P6, P8, PO4, PO8). Three time windows were defined for analyses of the Harmony and the Manner effects according to the following objective criteria: assuming that different map topographies and polarities directly indicate different underlying generators (Michel et al., 2004), borders of the time windows were placed at the average time point (across electrodes) at which changes in polarity and/or topography were found in one or the other group. This procedure yielded time windows (i) from 180 to 370 ms, (ii) from 370 to 550 ms, and (iii) from 550 to 1200 ms for Harmony effects; (i) from 180 to 420 ms, (ii) from 420 to 800 ms, and (iii) from 800 to 1200 ms for Manner effects. Non-parametric cluster-based permutation tests applied to the data with standard parameters in Fieldtrip (http://www.fieldtriptoolbox.org) led to qualitatively similar results as those obtained with the above defined time windows, hence validating the choice of the borders. Statistical analyses of mean amplitude values was carried out by means of six-way ANOVAs with the repeated measures factors Harmony (congruent/incongruent) × Context (long/short) × Manner (correct/incorrect) × Laterality (left/middle/right) × AntPost (anterior/central/posterior) and the between-subjects factor Group (classical/jazz), separately for each time window. Greenhouse-Geisser correction for non-sphericity was applied where appropriate (Keselman & Rogan, 1980).

Time-frequency analysis. Time-frequency (TF) analysis was carried out in FieldTrip (downloaded on 2012-12-05) (Oostenveld, Fries, Maris, & Schoffelen, 2011). To define frequency bands that were generally associated with visual-motor processing of chords regardless of experimental manipulation, epochs were cut from -11000 to 3000 ms for the long sequences and from -5000 to 3000 ms for the short sequences. Then, trials with voltages exceeding a ± 80 μV rejection criterion at one or more electrodes were rejected, TF information was extracted in 1-Hz bins within a 1-45 Hz frequency range using a Hanning-tapered window with 5 cycles and steps of 20 ms (using the '`ft_freqanalysis`' function with '`mtmconvol`' method as implemented in FieldTrip), and collapsed across all conditions (Figure 3). Task-relevant neural frequency bands were identified by visual inspection, including theta (4-8 Hz), alpha (9-13 Hz), and beta (20-30 Hz).

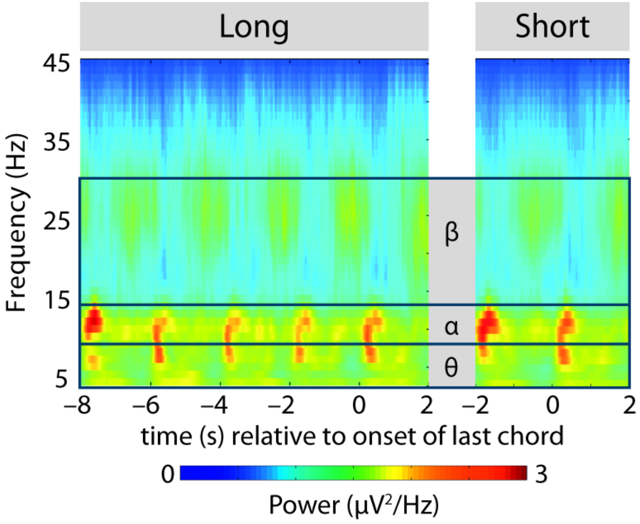


Figure 3: Raw spectral power change during the execution of long (5-chords) and short sequences (2-chords) for frequencies between 4 and 45 Hz across all trials. Frequency bands defined as theta, alpha and beta are indicated. Photos were presented at a rate of 2 seconds. Zero corresponds to the onset of the target chord.

.

In a second step, TF was applied to the same epochs as those that entered the ERP analysis, i.e. spanning -900 to 2200 ms relative to the onset of the final chord. For each group and for each of the eight conditions (H_cM_c, H_iM_c, H_cM_i, H_iM_i in the long and short context), TF information was extracted as described above and analysed in the theta, alpha and beta frequency range identified in the first analysis. To reduce inter-individual variance of absolute power values, individual band power was normalised similarly as in previous EEG (Meyer, Hunnius, Van Elk, Van Ede, & Bekkering, 2011; Novembre, Sammler, & Keller, 2016) and MEG studies (van Ede, de Lange, Jensen, & Maris, 2011). Specifically, for each participant, the average band power across N = 8 conditions (C) was subtracted at each electrode (e), in each frequency (f) and time bin (t) from the individual (condition specific) band power. This difference value was further divided by the average band power to centre values on zero.

$$\hat{P}_{\mathrm{C(e,f,t)}} = \frac{P_{\mathrm{C(e,f,t)}} - \frac{1}{N} \times \sum_{C=1}^{N} P_{\mathrm{C(e,f,t)}}}{1/N \times \sum_{C=1}^{N} P_{\mathrm{C(e,f,t)}}}$$

For statistical analyses, mean power values were extracted for each frequency bin (theta, alpha, beta) in each condition from 9 ROIs in the time windows defined in the previous ERP analysis. Six-way ANOVAs with the repeated measures factors Harmony (congruent/incongruent) x Context (long/short) x Manner (correct/incorrect) x Laterality (left/middle/right) x AntPost (anterior/central/posterior) and the between-subjects factor Group (classical/jazz) were calculated separately for each frequency band. Greenhouse-Geisser correction for non-sphericity was applied where appropriate. Non-parametric cluster-based permutation tests (cf. Maris & Oostenveld, 2007) with specific contrasts of interest (effects of Harmony and Manner for each context and each group) corroborated the main effects yielded by the ANOVAs.

Results

The present study compared classical (CG) and jazz pianists (JG) to assess training-style specific effects on two hierarchical stages of action control, i.e., structure-based planning of complex action sequences and parameter specification of single acts. According to theories of hierarchical action planning, we elucidated behavioural and electrophysiological patterns (ERPs and spectral power) associated with (i) the imitation of harmonic incongruities that conflict with pianists' structure-based action plans, particularly in long sequences (interaction of Harmony × Context), and (ii) the imitation of fingerings that make high demands on the specification of movement parameters in single acts (main effect of Manner). Crucially, (iii) we investigated between-group differences in musical action planning that may emerge from a different use of these action control processes during the execution of the same real-time task. Effects of musical training style on structure-based planning processes should surface as Group × Harmony × Context interactions. Conversely, pianists' differential sensitivity to movement parameter specification of single acts should be reflected by interactions of Group × Manner.

Behavioural data.

Response times. Statistical values of the four-way mixed-measures ANOVA with factors Harmony (congruent/incongruent), Context (long/short), Manner (correct/incorrect), and Group (classical/jazz) are reported in Table 1 and Figure 4A-B. Overall, both groups performed similarly fast (no main effect of Group; $p >$ .594), and imitated harmonically congruent chords faster than incongruent chords (main effect of Harmony; $p <$ .001; no interaction of Group x Harmony; $p >$ .660). However, differences were revealed in the way classical and jazz pianists reacted to harmonic manipulations in the two contexts: A significant interaction of Group × Harmony × Context ($p <$ .046) indicated a stronger effect of context on the execution of harmonically congruent/incongruent chords in the classical (CG) compared to the jazz group (JG). Follow-up ANOVAs with factors Harmony and Context testing CG and JG separately confirmed an interaction of Harmony × Context only in the CG [F(1,14) = 17.51, $p <$ .001, η_p^2 = .56], not in the JG [F(1,14) = 1.37, p > .261, η_p^2 = .09]. This suggests differences in structure-based planning between the two groups, with CG being potentially inclined to build stronger plans than JG, and/or JG being potentially better able than CG to flexibly adapt their plans to unexpected external musical events, as practiced during structure-generative training. Notably, although both groups were overall faster in executing conventional compared to unconventional fingerings, as expected (main effect of Manner; $p <$ .001; no interaction of Group × Manner; $p >$ .146), neither CG nor JG showed significant Manner × Context interactions (no Manner × Context interaction; p > .149; no Group × Manner × Context interaction; $p >$ .745). This suggests that both CG and JG specified movement parameters at the level of single acts (not the level of the action sequence), in line with Bianco et al. (2016) who argued that movement selection (as opposed to harmonic planning) constitutes a low level stage of action planning.

Error analysis. Key and fingering errors were analysed separately (for statistical details, see Table 1) because they are assumed to reflect distinct cognitive processes associated with imitation of the musical goal (i.e., Harmony) or imitation of the specific movements used during execution, respectively ((i.e., Manner; see Bianco et al., 2016; Novembre & Keller, 2011). Overall, pianists committed very few errors (mean ± SEM of key errors: 2.0 ± 0.2%; fingering errors: 9.0 ± 0.8%). With regard to key errors, classical and jazz pianists performed similarly well (CG: 3.0

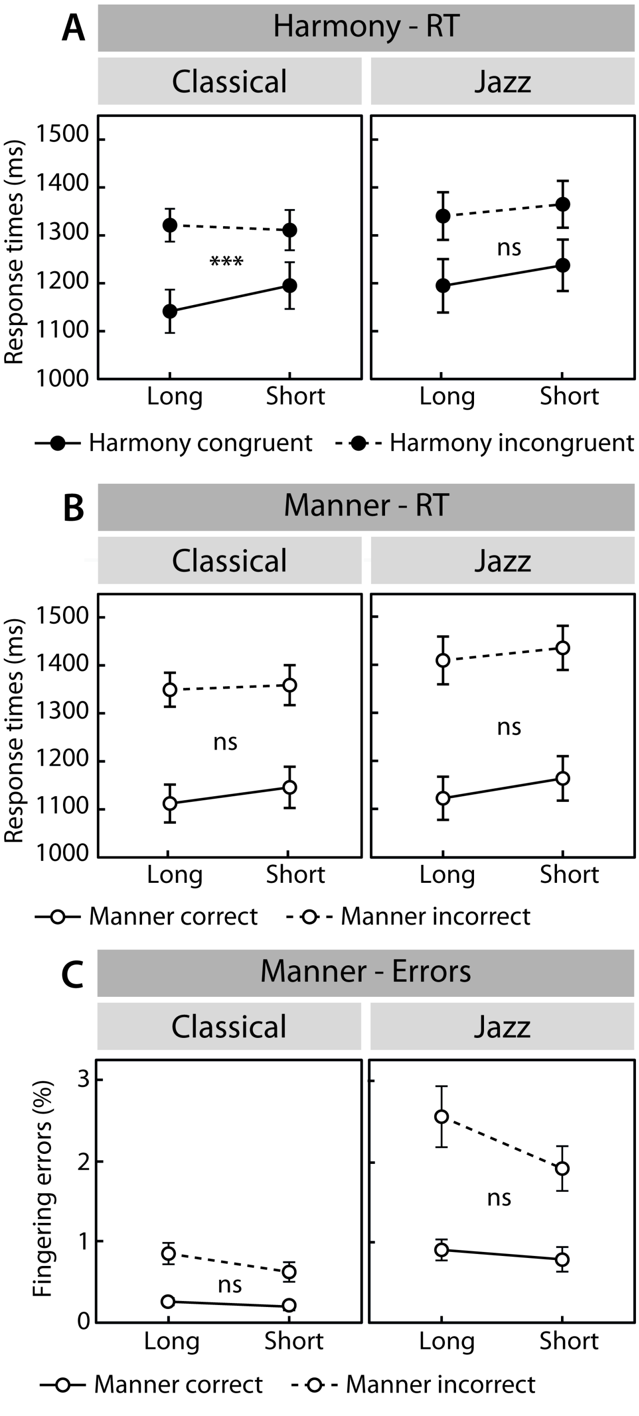


*Figure 4: Mean RTs and Fingering errors for classical (left panels) and jazz pianists (right panels). (A) RTs during imitation of harmonically incongruent (dashed line) and congruent chords (solid line), (B) RTs during imitation of manner incorrect (dashed line) and correct chords (solid line), and (C) fingering errors during imitation of manner incorrect (dashed line) and correct chords (solid line) in the long and short context. Error bars indicate ± 1 SEM. Asterisks indicate significance of the Harmony × Context and Manner × Context interactions: ***p < .001, ns:p> .05.*

.

± 1.0%; JG: 2.0 ± 1.0%; no main effect of Group; p > .166). Both groups produced less key errors during execution of harmonically congruent than incongruent chords (main effect of Harmony; p < .001; no interaction of Group × Harmony; p > .116), especially in the long context (interaction of Harmony × Context; p < .017; no interaction of Group × Harmony × Context; p > .980). Key errors did not differ significantly between manner correct and incorrect trials (no main effect of Manner or interactions with Manner; p's > .114), confirming that key errors are linked to harmony-related structure-based planning of pianists' actions.

With regard to fingering, classical pianists were overall more accurate than jazz pianists (CG: 4.0 ± 1.0 %; JG: 14.0 ± 4.0%; main effect of Group; p < .001), possibly due to generally stronger focus on hand posture in classical than jazz education (see also below). Both groups committed more errors when imitating unconventional compared to conventional fingerings (main effect of Manner; p < .001), however, particularly the JG (interaction of Group × Manner, p < .011) (see Figure 4C). Interestingly, more fingering errors were committed in the harmonically incongruent than congruent chords (main effect of Harmony; p < .049) an effect that tended to be driven mainly by the JG than the CG (marginally significant interaction of Group × Harmony; p > .079). This suggests that the JG compared to CG allocated more resources to the keys when harmonic violations occurred, to the detriment of the manner of execution. Finally, an interaction of Manner × Context (p < .016) suggested that pianists committed more fingering errors when the manner was violated in the long than in the short context. An interaction of Group × Manner × Context × Harmony (F(1, 28) = 7.056, p < .012, η_p^2 = 0.201) clarified that this effect was mainly driven by the CG (Manner × Context × Harmony in CG: F(1, 28) = 7.94, p < .014, η_p^2 = .36; in JG: F(1, 28) = 2.35, p > .148, η_p^2 = .14). In fact, classical pianists committed more fingering errors when the manner was violated on top of harmonically congruent chords in the long context (Manner × Context interaction in harmonically congruent chords: F(1,14) = 10.34, p < .006, η_p^2 = .42; in harmonically incongruent chords F(1,14) = 0.50, p > .490, η_p^2 = .03). This indicates that the CG, more than the JG, associated conventional fingering with congruent harmony, replicating previous results (Bianco et al., 2016). It further suggests that structure-interpretative focus may strengthen the link between structure-based plan and optimal parameter specification.

Post-experiment ratings. In the debriefing questionnaires after the experiment, both groups reported similar active prediction of the next chord during execution of the chord progressions (mean ± SEM on a scale from 1 to 9; CG: 5.64 ± 0.37; JG: 4.26 ± 0.67; $t(28) = -1.79$, $p > .080$). However, prediction strategies differed significantly between groups, with JG preferring to internally name the harmonic chord functions (CG: 3.78 ± 0.62; JG: 6.40 ± 0.43; $t(28) = 3.46$, $p < .002$), and CG paying more attention to the hand as a whole than JG (CG: 5.57 ± 0.32; JG: 3.67 ± 0.57; $t(28) = -2.92$, $p < .007$). These preferences may have tuned JG's structure-based planning and CG's sensitivity to movement parameters, and are compatible with the key and fingering error results described above.

ERP data

Harmony ERPs. In both groups, harmonically incongruent (compared to congruent) chords evoked an early central positivity between 180 and 370 ms, followed by a late posterior negativity between 550 and 1200 ms (see Figure 5A). Only the negativity was stronger in the long than the short context, disclosing it as a marker associated with structure-based planning (see also Sammler et al., 2013; Bianco et al., 2016), and showing that both CG and JG processed the harmonic incongruity. Most importantly, this negativity started significantly earlier in jazz pianists (already between 370 to 550 ms) than in classical pianists (only after 550 ms). This earlier peak might reflect prompt revision of the pre-planned motor act in JG, enabling them to compensate potential behavioural costs during execution. Effects are displayed in Figure 5A and will be statistically assessed below (see also Table 2 for statistical details).

In the *1st time window* (180 to 370 ms), harmonically incongruent chords evoked a positivity with left-central scalp distribution (main effect of Harmony; $p < .015$; interaction of Harmony × AntPost × Laterality; $p < .002$) that was similarly strong in both long and short contexts (no interactions involving Harmony and Context; p's > .256), and in both classical as well as jazz pianists (no interactions involving Harmony and Group; p's > .111). This positivity has been described in earlier studies (Sammler et al., 2013; Bianco et al., 2016) as related to the perceptually different hand postures of congruent and incongruent target chords. In the *2nd time window* (370 to 550 ms), the positivity gave way to a posterior negativity (interaction of Harmony × AntPost; $p < .004$), however, only in the JG as indicated by the significant interaction of Group × Harmony × Context × AntPost × Laterality ($p < .029$). Follow-up ANOVAs with the factors Harmony and Context conducted separately for each group and each ROI confirmed a stronger negativity evoked by harmonically incongruent (than congruent) chords in the long rather than short context in the JG in middle-central, middle- and right-posterior ROIs (interactions of Harmony × Context; p's < .039); an analogous Harmony × Context interaction in the middle-central, left-anterior and left-central ROIs (p's < .012) of the CG was driven by the opposite pattern, i.e., a positivity that was stronger in the short than in the long context. In the *3rd time window* (550 to 1200 ms), both groups displayed a similar late negativity that was stronger in the long than in the short context (interactions of Harmony × Context × Laterality: $p < .023$; Harmony × Context × AntPost: $p < .048$; no interactions involving Group × Harmony × Context; p's > .143). Follow-up ANOVAs calculated for each ROI replicated the centro-parietal topography of the negativity (Harmony × Context interaction in left- and middle-central, left- and middle-posterior ROIs: p's < .045; no interactions with Group: p's > .354).

Manner ERPs. In both CG and JG, manner incorrect compared to correct chords evoked similar neural signatures comprising an early left anterior positivity (180 to 420 ms), followed by a late posterior positivity (420 to 800 ms), and a right central negativity (800 to 1200 ms). None of these potentials was stronger in the long than the short context, similarly in both groups, indicating that both CG and JG planned low-level movement parameters (i.e., fingering) at the level of the single act (not the musical sequence), in line with the behavioural data (no Group × Manner × Context interaction in the RTs) and previous research (Bianco et al., 2016). Effects are displayed in Figure 5B and will be statistically assessed below (see also Table 3 for statistical details).

In the *1st time window* (180 to 420 ms), a left-anterior positivity was elicited by the manner incorrect compared to manner correct chords (interaction of Manner × AntPost, Manner × Laterality, and Manner × AntPost × Laterality; p's < .028), with CG showing a more broadly distributed positivity than JG (interaction of Group × Manner × AntPost × Laterality; $p < .008$). In the *2nd time window* (420 to 800 ms), the positivity attained a predominantly central-posterior distribution in both groups as confirmed by a Manner × AntPost × Laterality interaction ($p < .026$), and no interaction involving the factor Group (p's > .054). Finally, in the *3rd time window* (800 to 1200 ms), both groups showed a right-anteriorly distributed negativity associated with manner incorrect chords (interaction of Manner × AntPost × Laterality; $p < .001$). The negativity was more broadly distributed in the short

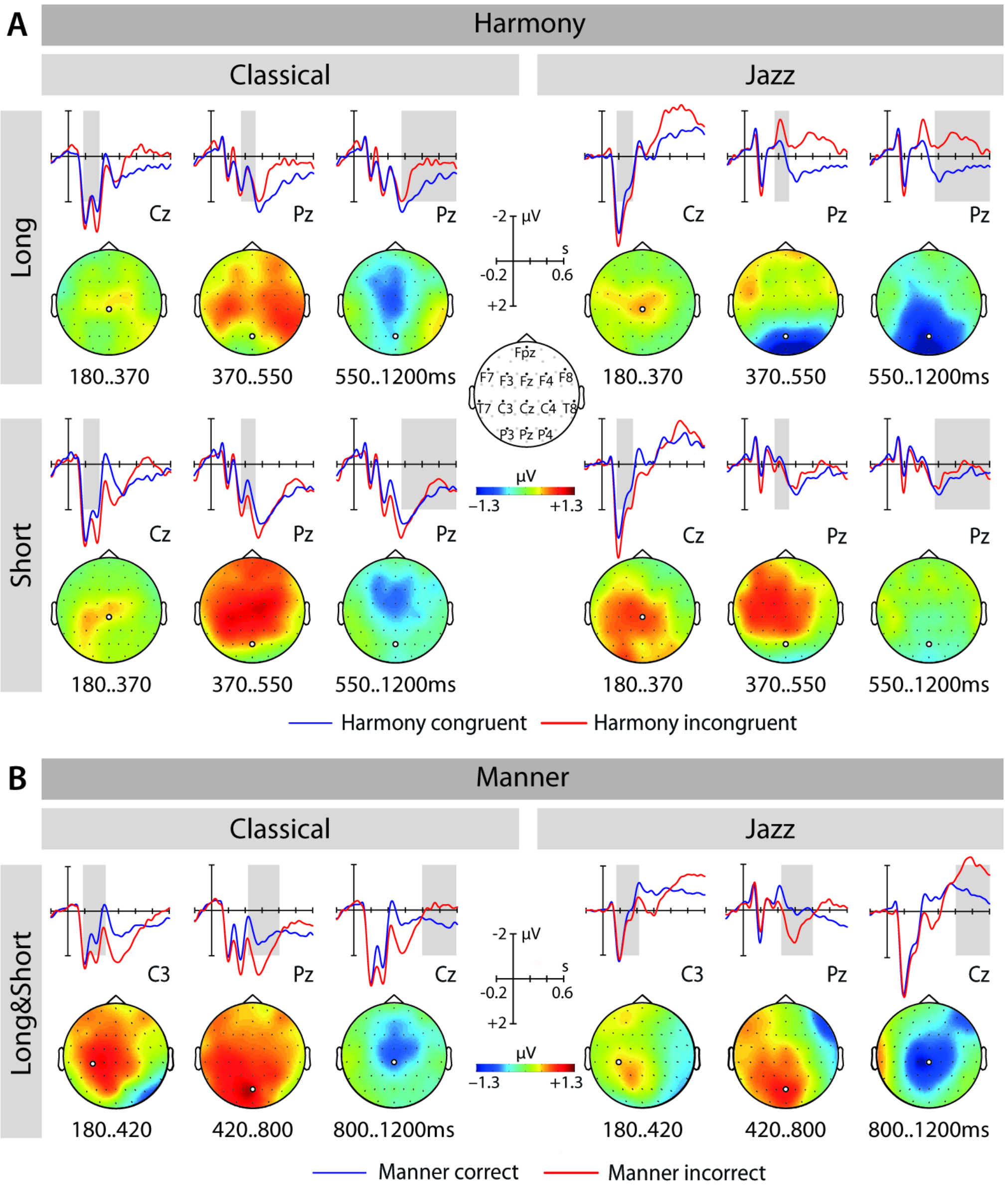


Figure 5: ERP effects of Harmony and Manner. A) ERPs evoked by harmonically incongruent (red line) compared to congruent chords (blue line) in the long (top) and short context (bottom) in CG (left panel) and JG (right panel). B) ERPs evoked by manner incorrect (red line) compared to correct chords (blue line) averaged across long and short contexts in CG (left panel) and JG (right panel). Time windows are shaded in grey. Topography maps for each statistical time window depict the difference potentials of harmonically incongruent / manner incorrect minus congruent / correct chords. Positions of the respective electrodes are indicated as "o" in the head-plots.

than the long context, as indicated by a Manner × Context × AntPost × Laterality interaction ($p < .004$).

Time frequency data.
Time frequency analysis was performed in three frequency bands (theta: 4-8 Hz, alpha: 9-13 Hz, beta: 20-30 Hz, see *Methods* section), in the same time windows as those used for the ERP data analysis. Between-group differences in structure-based planning were found in the *1st time window* in the theta band, and in the *3rd time window* in the alpha band, both displaying differences between harmonically incongruent and congruent conditions that were significantly stronger in the long than in the short context. At the level of movement parameter setting of single acts, we observed group differences in the theta and alpha band in the *2nd time window*, both exhibiting differences between manner incorrect and correct conditions that were significant in both contexts. To ensure the reliability of our results, particular attention was paid to effects that were confirmed by non-parametric cluster-based permutation tests (calculated in Fieldtrip; see *Methods*). Effects are displayed in Figure 6A-B and will be statistically assessed below.

Harmony theta. Only CG, not JG, showed a strong increase of early prefrontal theta power during harmonically incongruent compared to congruent chords, particularly in the long sequences, disclosing this effect as associated with structure-based planning (see Table 4).
In the *1st time window* (180 to 370 ms), a Group × Harmony × Context × AntPost × Laterality interaction ($p < .012$) indicated a right-frontally distributed increase in theta power only in CG and only in long sequences as confirmed by follow-up ANOVAs testing each ROI separately: only the right-anterior and right-central ROIs displayed a Harmony × Context interaction in the CG (p's < .028), but not in the JG (all ROIs: p's > .254). In the *2nd time window* (370 to 550 ms), this effect continued in the CG (Harmony × Context interaction in CG in right-anterior and right-central ROIs: p's < .012), still not in the JG (all ROIs: p's > .139), although group differences fell short of the level of significance in the 5-way ANOVA (no interaction involving Group × Harmony × Context; p's > .206). No relevant effects were found in the *3rd time window*.

Harmony alpha. The 5-way ANOVA indicated that there was no effect related to structure-based planning that differed between groups (see Table 5). No significant effects involving Group × Harmony × Context (p's > .437) were found in the *1st time window*, and the interaction of Group × Harmony × Context the *2nd time window* (p < .035) was not supported by further splits of the general liner model (no interaction involving Harmony × Context in JG: p's > .172; in CG: p's > .474). In the *3rd time window* (550 to 1200 ms), an interaction of Harmony × Context (p < .010) indicated that alpha power decreased more strongly in the long than in the short context. A Group $times$ Harmony × Context interaction fell short of statistical significance (p > .224), although this effect appeared to be mainly driven by the CG, as can be seen in Figure 6A (lower panel). Exploratory ANOVAs for each group separately lend some some support to that interpretation (interaction of Harmony × Context in the CG: p < .038; in the JG: p < .140).

Harmony beta. There were no effects related to structure-based motor planning that differed between groups, therefore, harmony-related effects in beta will not be discussed.

Manner theta. Theta power increased in manner-incorrect compared to correct conditions in both contexts in CG, but not in JG, in line with a stronger focus on movement parameters during classical, structure-interpretative training (see Table 6). In the *2nd time window* (420 to 800 ms), a Group × Manner × Context × AntPost × Laterality interaction (p < .021) indicated a broadly distributed increase in theta power only in the CG, while the interaction with Context was not supported by further splits of the model. This was supported by follow-up ANOVAs in the CG across all ROIs (main effect of Manner: p < .003; no interaction of Manner × Context: p > .905), and in single ROIs (main effect of Manner in all left, all posterior and middle-anterior ROIs: p's < .048; other ROIs: p's > .051; no interaction of Manner × Context: p > .288). No main effect of Manner was observed in the JG in any ROI (all ROIs: p's > .165). An interaction of Manner × Context × AntPost × Laterality in the *1st time window* (p < .042, see Table 6) was not supported by further splits of the general linear model (Manner × Context interaction in all ROIs: p's > .060). No significant effects were found in the *3rd time window*.

Manner alpha. Alpha power increased in manner-incorrect compared to correct conditions in both contexts in CG, but not in JG, providing further evidence for stronger sensitivity to the manner of single act execution in pianists who typically adopt a structure-interpretative focus (see Table 7). In the *2nd time window* (420 to

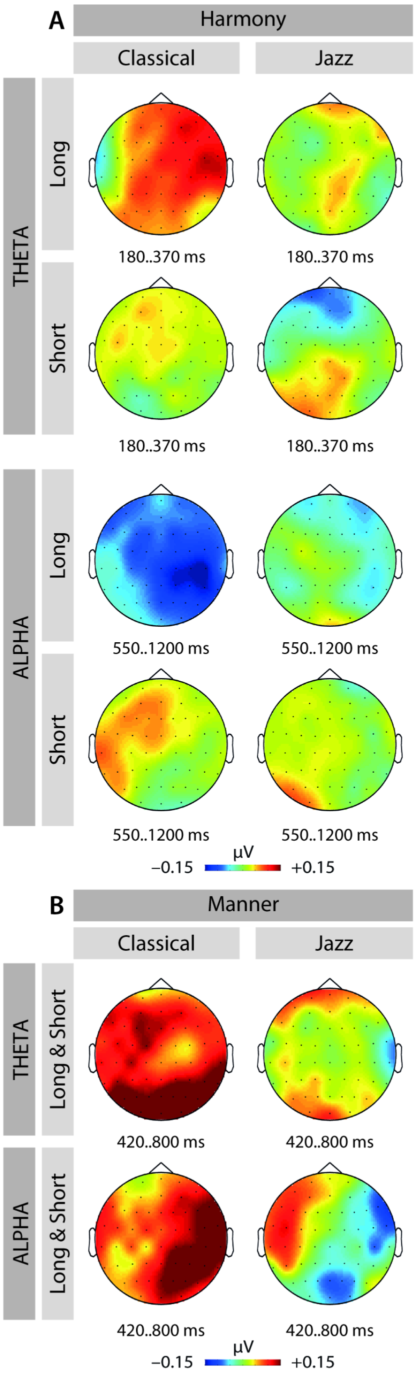


Figure 6: Spectral power effects of Harmony and Manner. A) Topography of normalised power change in the harmonically incongruent minus congruent condition for CG (left panel) and for JG (right panel): in theta band (1^{st} time window), in alpha band (3^{rd} time window) separately for the long (upper row) and the short context (lower row). B) Topography of normalised power change in the manner incorrect minus correct condition for CG (left panel) and for JG (right panel): in theta band (2^{nd} time window), in alpha band (2^{nd} time window) averaged across long and short contexts.

.

800 ms), this stronger manner-related effect in CG compared with JG was indicated by significant interactions of Group × Manner × Laterality ($p < .026$) and confirmed by follow-up ANOVAs split by Laterality (interaction of Manner × Group in right ROI: $p < .042$; in middle and left ROIs: p's > .086). The concurrent 5-way interaction ($p < .011$) was not supported by follow-up ANOVAs indicating that the alpha effect did not differ between long and short contexts (Manner × Context interaction in all ROIs in CG: p's > .340; in JG: p's > .099). No significant group differences were found in the other time windows.

Manner beta. No significant group differences were found, therefore manner-related effects in beta will not be discussed.

Discussion

The present study investigated how the multilayer organisation of actions based on harmonic knowledge of music can be tuned in performers that experience different action planning demands in their daily practice (see Figure 1). Musicians not only have deep abstract musical understanding, but they use this knowledge to inform and flexibly adapt their performance depending on current or long-term demands. Jazz and classical pianists — who are respectively required to focus more on structure-generative or structure-interpretative abilities — were tested with the same production task that addressed two levels of the action control hierarchy: structure-based planning of the action sequence and parameter specification of single acts. By combining behavioural, ERP and power change measures, results showed indices of hierarchical action control at the levels of the sequence and the single act in both groups, but also crucial between-groups differences. Specifically, at the sequence level, jazz pianists were more flexible at revising their structure-plans in face of harmonic incongruities, while at the single act level, classical pianists showed greater accuracy to set movement parameters based on structural features.

Hierarchical action control in both classical and jazz pianists. Both classical and jazz pianists showed indices of hierarchical action control, at the level of the sequence as well as the level of single acts.

At *sequence-level*, results converge on robust structure-based planning in both groups: Classical as well as jazz pianists committed more key errors (Figure 4C) and showed a stronger late negativity (Figure 5A) when imitating incongruent compared to congruent chords, particularly when they were embedded in a long harmonic context (Harmony × Context interaction). This replicates previous findings (Novembre & Keller, 2011; Bianco et al., 2016; Sammler et al., 2013) and indicates that pianists incrementally anticipate and prepare forthcoming musical acts with reference to structure-based plans they construct from their knowledge of harmony coupled with the continuously unfolding musical context (Palmer & Pfordresher, 2003; Clarke, 2001). Notably, the context-dependency of both behavioural and electrophysiological effects (reflected in Harmony × Context interactions) suggests that a longer harmonic context allows performers to identify larger-scale relationships between constituent acts, i.e., to build a harmonic representation of greater hierarchical depth that strengthens the structure-based plan and narrows down the possibilities of likely forthcoming events in the sequence ("what" to play). Imitation of a chord that mismatches the anticipated event requires the "reprogramming" of the prepotent response, plausibly more strongly in the long than short context. This process is typically associated with behavioural costs (a slowing of response times which was observed only in classical pianists as will be explained below, and an increase in key mistakes observed in both groups), and a late negativity (Bianco et al., 2016; Leuthold & Jentzsch, 2002; Sammler et al., 2013) such as the one found in the *3rd time window* (550 to 1200 ms) in both groups. Altogether, the present data argue for high-level structure-based plans that govern the appropriate ordering of constituent acts during both classical and jazz performance (Clarke, 2001).

At *single act level*, both groups showed distinct signatures of goal-related planning that determines "what" to play (e.g., which chord; Harmony) and subsequent movement parameter specification that determines "how" to play (e.g., which fingering; Manner): while imitation of harmonically incongruent chords (goal) induced context-dependent late negativity and an increase of key errors (see above), imitation of unconventional fingering (parameters) was associated with context-independent late positivity (Figure 5B) and response time slowing (Figure 4B), in line with previous results (Bianco et al., 2016). Notably, the context-independence of the manner effects (i.e., no Manner × Context interactions) demonstrates that performers specify the manner of execution at single act rather than sequence level, i.e., at late stages of action planning. This assumption resonates with the occurrence of a positivity (rather than a negativity) in the *2nd time window* (420 to 800 ms) that has been previously related to visual-spatial processes that precede action programming, not to the (re)programming of a motor act itself that is typically reflected in a late negativity (Leuthold & Jentzsch, 2002). In other words, the specification of the fingering is subordinate to the planning of the chord in line with hierarchical action theories (see Figure 1, Rosenbaum et al., 2007; Vallacher & Wegner, 1987; Wohlschläger et al., 2003). Altogether, the present data argue for multilayer planning processes at single act level with musical goal taking priority over manner, in both classical and jazz performance.

Greater structural flexibility in jazz pianists. Despite the similar hierarchical core structure of action planning in classical and jazz pianists, the present data highlight nuanced group differences that may indicate a variable tuning of the action control hierarchy, possibly shaped by action tendencies developed through training.

At the *sequence level*, jazz compared to classical pianists showed indices for a greater readiness and flexibility to revise their structure-based plans in case of harmonically unexpected chords: While classical pianists displayed a context-dependent slowing of response times, increase of right-frontal theta and decrease of late alpha power when asked to imitate harmonic violations (Harmony × Context interaction), jazz pianists showed none of these effects (Harmony × Context × Group interaction; Figures 4A and 6A). Instead, they reacted to harmonic violations with an increase of fingering mistakes that was not seen in classical pianists and an earlier onset of the context-dependent "reprogramming" negativity, already in the *2nd* (370 to 550 ms), not only in the *3rd time window* (550 to 1200 ms).
Our interpretation of these findings is twofold: first, jazz pianists seem to be able to reassess their structure-based action plans more quickly than classical pianists (earlier "reprogramming" negativity). This prevents response time costs (no context-dependent slowing) at the expense of the manner (more fingering errors). Second, classical pianists seem to display more cognitive conflict and effort when a structural revision is required. In terms of conflict, the increase of early right-frontal theta power (180 to 370 ms) can be interpreted as a mismatch signal that calls for the initiation of other alternatives (Cavanagh & Frank, 2014; Cohen, 2014). Its right-anterior scalp distribution is in line with local frontal control functions (Miller & Cohen, 2001; Botvinick, Cohen, & Carter, 2004; Cavanagh, Cohen, & Allen, 2009) that inform action selection based on goal-relevant information (Womelsdorf, Johnston, Vinck, & Everling, 2010), including the inhibition of prepotent and initiation of goal-relevant responses (Miller & Cohen, 2001; Munakata et al., 2011;

Aron, Robbins, & Poldrack, 2014). In terms of effort, the decrease of late alpha power (550 to 1200 ms) may indicate cognitive revision processes (Palva & Palva, 2007) such as context-updating and rule-use/switching (Mansfield, Karayanidis, & Cohen, 2012; Cooper, Darriba, Karayanidis, & Barceló, 2016). Note that the right-lateralisation of the effect excludes a purely sensorimotor interpretation (Pfurtscheller & Lopes, 1999), because pianists were performing with the ipsilateral right hand.

Altogether, these combined behavioural and electrophysiological findings suggest qualitative differences in structure-based planning that may emerge from action planning requirements in jazz and classical performance: The spontaneous generation and intentional violation of musical structures in jazz may train pianists to dynamically arrange and re-arrange structure-based plans allowing them to flexibly adapt to harmonic novelties. This structure-generative tendency and readiness for change may be grounded in the simultaneous pre-activation of several possible harmonic options, the most likely of which has right of way but can be rapidly cancelled and turned into one of the other alternatives (Cisek, 2006). The greater number of competing harmonic alternatives available to jazz pianists may imply that more resources are allocated to high-level planning, coming at the expense of low-level fingering specification, especially under time pressure. Indeed, jazz compared to classical pianists had higher rates of fingering errors that might reflect a speed-accuracy trade-off (Berlyne, 1957) observed across two levels of action planning. Classical pianists' structure-interpretative focus in turn may lead to greater narrowing of harmonic possibilities and shallower processing of long-term dependencies in order to rapidly proceed to expressive stages of action planning (the "how", see below), hence, evoking greater cognitive effort to initiate other alternatives in case of structural-harmonic deception. These strategic differences fit with the self-reported higher understanding of harmonic relationships between chords in JG than CG. Furthermore, cognitive psychological models of improvisation (Pressing, 1987; Clarke, 2001) emphasise the importance of combinatorial strategies to embrace the wealth of possible structural relationships between musical elements, in order to either intentionally maintain rigorous musicality or subvert the audience's expectations. In turn, models of expressive musical interpretation (Shaffer, 1984; Clarke, 2001) highlight the importance of fast encoding of given musical structures to rapidly free cognitive resources for generation of expressive intentions.

Greater sensitivity to movement parameters in classical pianists. At *single act level*, classical compared to jazz pianists revealed a higher propensity to encode and accurately set movement parameters (as reflected by Group × Manner interactions): Classical pianists showed a broader early positivity upon detection of the fingering violation (180 to 420 ms; Figure 5B), an increase of late posterior theta and alpha power (420 to 800 ms, that was not seen in JG) upon programming of the uncommon fingering (Figure 6B), and an overall more accurate imitation of the fingering (Figure 4C) that was partly determined by pianists' higher-level structural-harmonic predictions (Manner × Harmony × Context interaction in CG, not JG).

These combined data seem to indicate that classical pianists are more sensitive than jazz pianists to the way musical elements are motorically rendered: This possibly reflects their trained structure-interpretative abilities to rapidly plan expressive features on top of the musical structure by relying on strong associations between fingering and frequent musical patterns (Gellrich & Parncutt, 1998). Classical pianists' expressive tendency may have enhanced their sensorimotor preparedness to rapidly react to observed erroneous finger movements (Candidi, Maria Sacheli, Mega, & Aglioti, 2014) and may have even boosted the processing of stimulus-features in the model hand that were relevant for the fine-tuning of the pianists' own actions (see also Fagioli, Ferlazzo, & Hommel, 2007; Hommel, 2010). The latter point receives support from the stronger self-reported focus of classical pianists on hand posture and the more broadly distributed early positivity in CG than JG, a component that has recently been associated with the detection of action-relevant stimulus-properties (Polich, 2007; Bianco et al., 2016). The subsequent posterior theta and alpha power increase is more likely tied to the correct implementation and consolidation of the uncommon fingering, e.g., via appropriate visuo-motor translation of the unconventional stimulus into the required response (Culham, Cavina-Pratesi, & Singhal, 2006; Nigbur, Cohen, Ridderinkhof, & Stürmer, 2012). Notably, classical pianists seemed to be committed to the higher-level structure-based plan in their lower-level choice of fingering: The fact that fingering mistakes occurred particularly frequently during imitation of unconventional manner on top of structurally congruent chords suggests that pianists had prepared the fingering that matched the structure-based plan in advance (Manner × Harmony × Context interaction only in CG), in line with previous results (Bianco et al., 2016). In other words, structure-interpretative focus may lead pianists to rapidly project optimal movement parameters as soon as the structure-based plan has emerged from the context.

Altogether, these findings indicate that classical and jazz pianists give different weights to the hierarchical levels of action planning, intuitively and despite identical instruction and material: While habitual action focus on structure generation in jazz may inadvertently bind resources to high levels of planning, even in a task that did not require creative improvisation, the building of solid associations between finger configurations and frequent structural figures (e.g., scales, intervals, cadences) in classical training (Clarke, Parncutt, Raekallio, & Sloboda, 1997; Gellrich & Parncutt, 1998; Parncutt, 2014) may spread focus from high to low levels to ease the structure-based selection of fine-grained expressive features that shade interpretative performance (Clarke, 2001).

Conclusion

Complex actions are not rigid entities but their planning and execution can be variably shaped by the demands of our daily life. The present study used music production as experimental framework to investigate hierarchical principles of complex sequential action planning and the influence of habitual action focus. We demonstrate a general dissociation between (i) structure-based planning at the level of the action sequence and (ii) movement parameter specification at the level of single acts. Furthermore, we show that these action control processes are not fixed, but are tuned differently — despite equal instruction and task — depending on action tendencies developed by classical and jazz musicians during daily practice. The observed structure-generative flexibility in jazz pianists and structure-interpretative speed in classical pianists may reflect an adaptive tuning of the action control system that supports musicians to either immerse in creative improvisation or enthralling interpretation. In other words, playing jazz or classical music may not only be a matter of different musical styles, but an "inner approach to music" (Keith Jarrett in Rosenthal, 1997) that drives the same cognitive machinery differently just as we maintain our particular individual style during talking or drawing.

Acknowledgments

We are grateful to the classical and jazz pianists who participated to this experiment and to Prof. Michael Wollny. We also thank Corinna Eidner and Elias Matin for helping with data acquisition, Sven Gutekunst for technical support, and Gabriele Chierchia for valuable discussions.

References

Aron, A. R., Robbins, T. W., and Poldrack, R. (2014). Inhibition and the right inferior frontal cortex: one decade on. *Trends in Cognitive Sciences*, 18(4):177–185.

Beaty, R. E. (2015). The neuroscience of musical improvisation. *Neuroscience & Biobehavioral Reviews*, 51:108–117.

Bekkering, H., Wohlschläger, A., and Gattis, M. (2000). Imitation of gestures in children is goal-directed. *The Quarterly Journal of Experimental Psychology*, 53(1):153–64.

Berlyne, D. E. (1957). Uncertainty and conflict: a point of contact between information-theory and behavior-theory concepts. *Psychological Review*, 64(6):329–339.

Bianco, R., Novembre, G., Keller, P. E., Scharf, F., Friederici, A. D., Villringer, A., and Sammler, D. (2016). Syntax in action has priority over movement selection in piano playing: an ERP study. *Journal of Cognitive Neuroscience*, 28(1):41–54.

Botvinick, M. M., Cohen, J. D., and Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, 8(12):539–546.

Candidi, M., Maria Sacheli, L., Mega, I., and Aglioti, S. M. (2014). Somatotopic mapping of piano fingering errors in sensorimotor experts: TMS studies in pianists and visually trained musically naives. *Cerebral Cortex*, 24(2):435–443.

Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G., and Rizzolatti, G. (2007). Impairment of actions chains in autism and its possible role in intention understanding. *Proceedings of the National Academy of Sciences of the United States of America*, 104(45):17825–30.

Cavanagh, J. F., Cohen, M. X., and Allen, J. J. B. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *The Journal of Neuroscience*, 29(1):98–105.

Cavanagh, J. F. and Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, 18(8):414–421.

Chaffin, R., Lemieux, A. F., and Colleen, C. (2007). It is different each time I play: variability in highly prepared musical performance. *Music Perception*, 24(5):455–472.

Cisek, P. (2006). Integrated neural processes for defining potential actions and deciding between them: a computational model. *The Journal of Neuroscience*, 26(38):9761–70.

Clarke, E. (2001). Generative principles in music performance. In *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition*, pages 1–26. Oxford University Press, New York.

Clarke, E., Parncutt, R., Raekallio, M., and Sloboda, J. (1997). Talking fingers: an interview study of pianists' views on fingering. *Musicae Scientiae*, 1(1):87–107.

Cohen, M. X. (2014). A neural microcircuit for cognitive conflict detection and signaling. *Trends in Neurosciences*, 37(9):480–490.

Cooper, P. S., Darriba, Á., Karayanidis, F., and Barceló, F. (2016). Contextually sensitive power changes across multiple frequency bands underpin cognitive control. *NeuroImage*, 132:499–511.

Culham, J. C., Cavina-Pratesi, C., and Singhal, A. (2006). The role of parietal cortex in visuomotor control: what have we learned from neuroimaging? *Neuropsychologia*, 44:2668–2684.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Research*, 93(3):283–321.

Delorme, A. and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1):9–21.

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., and Prinz, W. (2005). Action-effect coupling in pianists. *Psychological Research*, 69(4):233–41.

Fagioli, S., Ferlazzo, F., and Hommel, B. (2007). Controlling attention through action: Observing actions primes action-related stimulus dimensions. *Neuropsychologia*, 45(14):3351–3355.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, 308(5722):662–667.

Gellrich, M. and Parncutt, R. (1998). Piano Technique and Fingering in the Eighteenth and Nineteenth Centuries: Bringing a forgotten method back to life. *British Journal of Music Education*, 15(01):5–23.

Grafton, S. T. and Hamilton, A. F. D. C. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, 26(4):590–616.

Hommel, B. (2010). Grounding attention in action control: The intentional control of selection. In B. J. Bruya, editor, *Effortless attention: A new perspective in the cognitive science of attention and action*, pages 121–140. MA: MIT Press, Cambridge.

Johnson-Laird, P. N. (2002). How Jazz Musicians Improvise. *Music Perception: An Interdisciplinary Journal*, 19(3):415–442.

Keller, P. E. (2012). What movement force reveals about cognitive processes in music performance. In *Art in motion II*, pages 115–153. Peter Lang, Frankfurt.

Keselman, H. J. and Rogan, J. C. (1980). Repeated measures F tests and psychophysiological research: controlling the number of false positives. *Psychophysiology*, 17:499–503.

Koechlin, E. and Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, 11(6):229–235.

Lashley, K. (1951). The problem of serial order in behavior. In Jeffress, L., editor, *Cerebral Mechanisms in Behavior*, pages 112–131. New York: Wiley.

Leuthold, H. and Jentzsch, I. (2002). Spatiotemporal source localisation reveals involvement of medial premotor areas in movement reprogramming. *Experimental Brain Research*, 144(2):178–88.

Mansfield, E. L., Karayanidis, F., and Cohen, M. X. (2012). Switch-related and general preparation processes in task-switching: evidence from multivariate pattern classification of EEG data. *Journal of Neuroscience*, 32(50):18253–8.

Maris, E. and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164:177–190.

Meyer, M., Hunnius, S., Van Elk, M., Van Ede, F., and Bekkering, H. (2011). Joint action modulates motor system involvement during action observation in 3-year-olds. *Experimental Brain Research*, 211(3-4):581–592.

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., and Grave de Peralta, R. (2004). EEG source imaging. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, 115(10):2195–222.

Miller, E. K. and Cohen, J. D. (2001). An integrative theory of preforntal cortex function. *Annual Review of Neuroscience*, 24:167–202.

Munakata, Y., Herd, S. a., Chatham, C. H., Depue, B. E., Banich, M. T., and O'Reilly, R. C. (2011). A unified framework for inhibitory control. *Trends in Cognitive Sciences*, 15(10):453–9.

Nigbur, R., Cohen, M. X., Ridderinkhof, K. R., and Stürmer, B. (2012). Theta dynamics reveal domain-specific control over stimulus and response conflict. *Journal of Cognitive Neuroscience*, 24(5):1264–1274.

Novembre, G. and Keller, P. E. (2011). A grammar of action generates predictions in skilled musicians. *Consciousness and Cognition*, 20(4):1232–43.

Novembre, G., Sammler, D., and Keller, P. E. (2016). Neural alpha oscillations index the balance between self-other integration and segregation in real-time joint action. *Neuropsychologia*, 89:414–425.

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, 2011.

Palmer, C. and Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review*, 110(4):683–712.

Palmer, C. and van de Sande, C. (1995). Range of planning in music performance. *Journal of Experimental Psychology. Human perception and Performance*, 21(5):947–62.

Palva, S. and Palva, J. M. (2007). New vistas for α-frequency band oscillations. *Trends in Neurosciences*, 30(4):150–158.

Parncutt, R. (2014). Fingering. In *Music in the social and behavioral sciences: An encyclopedia*, pages 481–484. Los Angeles: Sage, w. f. thom edition.

Pfurtscheller, G. and Lopes, F. H. (1999). Event-related EEG / MEG synchronization and desynchronization : basic principles. *Clinical Neurophysiology*, 110:1842–1857.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, 118(10):2128–48.

Pressing, J. (1984). The history of classical improvisation. Baroque music and beyond. *Keyboard*.

Pressing, J. (1987). Improvisations: methods and models. In *Generative Processes in Music*, pages 129–178. Oxford University Press.

Repp, B. H. (2000). Pattern typicality and dimensional interactions in pianists' imitation of expressive timing and dynamics. *Music Perception*, 18(2):173–211.

Rosenbaum, D. a., Cohen, R. G., Jax, S. a., Weiss, D. J., and van der Wel, R. (2007). The problem of serial order in behavior: Lashley's legacy. *Human Movement Science*, 26(4):525–54.

Rosenthal, T. (1997). The "insanity" of doing more than one (musical) thing. *Piano and Keyboard*, pages 1–16.

Sammler, D., Novembre, G., Koelsch, S., and Keller, P. E. (2013). Syntax in a pianist's hand: ERP signatures of "embodied" syntax processing in music. *Cortex*, 49(5):1325–1339.

Schmidt, R. a. R. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, 82(4):225–260.

Shaffer, L. H. (1980). 26 Analysing Piano Performance: A Study of Concert Pianists. *Advances in Psychology*, 1:443–455.

Shaffer, L. H. (1984). Timing in solo and duet piano performances. *The Quarterly Journal of Experimental Psychology*, 36(4):577–595.

Sharbrough, F., Chatrian, G., Lesser, R., Luders, H., Nuwer, M., and Picton, T. (1991). American Electroencephalographic Society guidelines for standard electrode position nomenclature. *Journal Clinical Neurophysiology*, (8):200–202.

Swain, J. P. (1995). The concept of musical syntax. *Musical Quarterly*, 79(2):281–308.

Uithol, S., van Rooij, I., Bekkering, H., and Haselager, P. (2012). Hierarchies in action and motor control. *Journal of Cognitive Neuroscience*, 24(5):1077–86.

Vallacher, R. R. and Wegner, D. M. (1987). Action identification and human behavior. *Psychological Review*, 94(1):3–15.

van Ede, F., de Lange, F., Jensen, O., and Maris, E. (2011). Orienting attention to an upcoming tactile event involves a spatially and temporally specific modulation of sensorimotor alpha- and beta-band oscillations. *The Journal of neuroscience*, 31(6):2016–2024.

Verwey, W. B., Shea, C. H., and Wright, D. L. (2015). A cognitive framework for explaining serial processing and sequence execution strategies. *Psychological Bulletin*, 22:54–77.

Wohlschläger, A., Gattis, M., and Bekkering, H. (2003). Action generation and action perception in imitation: an instance of the ideomotor principle. *Philosophical Transactions of the Royal Society of London. Series B, Biological* *sciences*, 358(1431):501–15.

Wolpert, D. M., Diedrichsen, J., and Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews. Neuroscience*, 12(12):739–51.

Womelsdorf, T., Johnston, K., Vinck, M., and Everling, S. (2010). Theta-activity in anterior cingulate cortex predicts task rules and their adjustments following errors. *Proceedings of the National Academy of Sciences of the United States of America*, 107(11):5248–5253.

Table 1 – ANOVAs for response times and errors with the factors Harmony x Context x Manner x Group on the behavioural data.

Effect	F(1,28)	p	η_p^2
Response times			
G	<1	.594	.001
H	124.85	< .001	.82
M	190.65	< .001	.87
H x C	14.13	< .001	.34
M x C	2.20	.149	.07
G x H	0.20	.660	.01
G x M	2.24	.146	.07
G x H x C	4.34	.046	.13
G x M x C	<1	.745	< .01
Key errors			
G	2.02	.166	.066
H	33.59	< .001	.545
M	2.66	.114	.09
H x C	6.40	.017	.186
M x C	<1	.799	< .01
G x H	2.64	.116	.09
G x M	<1	.964	< .01
G x H x C	<1	.980	< .01
G x M x C	1.10	.302	.04
Fingering errors			
G	17.36	< .001	.383
H	4.23	.049	.131
M	34.70	< .001	.553
H x C	<1	.790	< .01
M x C	6.51	.016	.188
G x H	3.33	.079	.11
G x M	7.36	.011	.208
G x H x C	2.40	.133	.08
G x M x C	1.75	.197	.06

Bold values indicate significant results ($p < .05$). Partial eta squared $\eta_p^2 > .5$ = large effect size, $\eta_p^2 > .3$ = medium effect size, $\eta_p^2 < .1$ = small effect size (Bortz and Döring, 2003). H = Harmony, M = Manner, C = Context, G = Group.

Table 2 – ANOVA on amplitudes values with the factors Group x Context x Harmony x Manner x AntPost x Laterality in each time window. N= 15 vs. 15. (HARMONY effects).

		1st tw: 180…370 ms			*2nd tw: 370…550 ms*			*3rd tw: 550…1200 ms*		
Effect	df	F	p	η^2_P	F	p	η^2_P	F	p	η^2_P
G	1,28	2.24	.146	0.07	3.92	.058	0.12	3.00	.094	0.1
H	1,28	6.75	.015	0.19	4.97	.034	0.15	8.32	.007	0.23
H x AP	2,56	0.35	.589	0.01	9.09	.004	0.25	1.93	.173	0.06
H x L	2,56	1.44	.245	0.05	1.51	.230	0.05	5.56	.013	0.17
H x AP x L	4,112	5.13	.002	0.15	1.00	.398	0.03	2.83	.043	0.09
H x C	1,28	1.35	.256	0.05	8.34	.007	0.23	3.88	.059	0.12
H x C x AP	2,56	1.05	.321	0.04	2.05	.160	0.07	4.00	.048	0.13
H x C x L	2,56	1.15	.317	0.04	3.98	.024	0.12	4.46	.023	0.14
H x C x AP x L	4,111	0.49	.683	0.02	1.03	.377	0.04	1.43	.241	0.05
G x H	1,28	0.27	.609	0.01	3.77	.062	0.12	0.28	.601	0.01
G x H x AP	2,56	0.00	.978	0.01	1.79	.191	0.06	3.92	.051	0.12
G x H x L	2,56	0.47	.627	0.02	0.30	.742	0.01	0.07	.878	0.01
G x H x AP x L	4,112	0.56	.655	0.02	0.37	.778	0.01	1.17	.326	0.04
G x H x C	1,28	0.19	.663	0.01	0.00	.984	0.01	0.34	.562	0.01
G x H x C x AP	2,56	0.19	.690	0.01	1.54	.227	0.05	0.09	.808	0.01
G x H x C x L	2,56	0.28	.714	0.01	0.57	.567	0.02	0.44	.607	0.02
G x H x C x AP x L	4,112	2.09	.111	0.07	3.38	.029	0.11	1.77	.163	0.06

Bold values indicate significant results ($p < .05$). Partial eta squared $\eta^2_P > 0.5$ = large effect size, $\eta^2_P > 0.3$ = medium effect size, $\eta^2_P > 0.1$ = small effect size (Bortz & Döring, 2003). G = group, H = Harmony, C = Context, AP = AntPost, L = Laterality. Degrees of freedom were corrected after Huynh-Feldt if necessary.

Table 3 – ANOVA on amplitudes values with the factors Group x Context x Harmony x Manner x AntPost x Laterality in each time window. N= 15 vs. 15. (MANNER effects).

		1st tw: 180…420 ms			*2nd tw: 420…800 ms*			*3rd tw: 800…1200 ms*		
Effect	df	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
G	1,28	2.92,	.099,	0.09	3.51,	.071,	0.11	2.52,	.124,	0.08
M	1,28	0.35,	.557,	0.01	10.12,	.004,	0.27	5.10,	.032,	0.15
M x AP	2,56	4.99,	.028,	0.15	4.81,	.030,	0.15	3.17,	.078,	0.1
M x L	2,56	9.78,	.001,	0.26	8.35,	.001,	0.23	13.01,	.001,	0.32
M x AP x L	4,112	4.95,	.004,	0.15	3.47,	.026,	0.11	5.99,	.002,	0.18
C x M	1,28	0.22,	.640,	0.01	0.31,	.579,	0.01	1.83,	.187,	0.06
C x M x AP	2,56	2.65,	.112,	0.09	2.16,	.148,	0.07	0.66,	.436,	0.02
C x M x L	2,56	1.48,	.239,	0.05	1.31,	.277,	0.04	1.59,	.212,	0.05
C x M x AP x L	4,111	2.49,	.047,	0.08	2.03,	.117,	0.07	4.03,	.004,	0.13
G x M	1,28	1.06,	.311,	0.04	1.72,	.200,	0.06	0.54,	.469,	0.02
G x M x AP	2,56	0.07,	.827,	0.01	0.04,	.875,	0.01	0.67,	.444,	0.02
G x M x L	2,56	0.09,	.913,	0.01	0.63,	.538,	0.02	1.32,	.275,	0.05
G x M x AP x L	4,112	4.42,	.008,	0.14	2.79,	.054,	0.09	1.98,	.130,	0.07
G x C x M	1,28	0.00,	.981,	0.01	0.01,	.940,	0.01	0.15,	.698,	0.01
G x C x M x AP	2,56	0.87,	.366,	0.03	0.03,	.899,	0.01	0.14,	.733,	0.01
G x C x M x L	2,56	0.30,	.699,	0.01	2.77,	.071,	0.09	3.63,	.033,	0.11
G x C x M x AP x L	4,112	1.10,	.358,	0.04	1.98,	.123,	0.07	1.40,	.239,	0.05

Bold values indicate significant results ($p < .05$). Partial eta squared $\eta_p^2 > 0.5$ = large effect size, $\eta_p^2 > 0.3$ = medium effect size, $\eta_p^2 > 0.1$ = small effect size (Bortz & Döring, 2003). M = Manner, C = Context, G = group, AP = AntPost, L = Laterality. Degrees of freedom were corrected after Huynh-Feldt if necessary.

Table 4 – ANOVA on THETA power values with the factors Group x Context x Harmony x Manner x AntPost x Laterality in each time window. N= 15 vs. 15. (HARMONY effects).

		1^{st} tw: 180…370 ms			*2^{nd} tw: 370…550 ms*			*3^{rd} tw: 550…1200 ms*		
Effect	df	F	p	η_p^2	F	p	η_p^2	F	p	η_p^2
G	1,28	5.78	.023	0.17	4.12	.052	0.13	6.15	.019	0.18
H	1,28	0.83	.370	0.03	0.21	.650	0.01	0.04	.843	0.01
H x AP	1,28	0.24	.677	0.01	0.28	.644	0.01	0.21	.690	0.01
H x L	1,28	0.93	.402	0.03	0.45	.637	0.02	2.53	.089	0.08
H x AP x L	2,56	1.77	.167	0.06	0.88	.480	0.03	0.34	.744	0.01
H x C	2,56	1.11	.302	0.04	4.11	.052	0.13	0.06	.813	0.01
H x C x AP	1,28	0.29	.697	0.01	0.09	.852	0.01	0.10	.842	0.01
H x C x L	2,56	1.64	.208	0.06	10.55	.001	0.27	3.12	.063	0.1
H x C x AP x L	2,56	0.91	.437	0.03	0.37	.710	0.01	0.65	.556	0.02
G x H	2,56	0.04	.849	0.01	0.11	.738	0.01	1.72	.200	0.06
G x H x AP	2,56	1.18	.296	0.04	1.46	.240	0.05	0.95	.352	0.03
G x H x L	4,112	1.83	.169	0.06	1.55	.222	0.05	0.85	.432	0.03
G x H x AP x L	2,56	1.02	.381	0.04	1.95	.107	0.07	1.74	.179	0.06
G x H x C	2,56	0.87	.359	0.03	0.74	.397	0.03	0.85	.364	0.03
G x H x C x AP	4,112	1.47	.240	0.05	0.95	.369	0.03	0.22	.728	0.01
G x H x C x L	4,112	0.17	.79	0.01	0.0	.911	0.01	0.86	.410	0.03
G x H x C x AP x L	4,112	3.94	.012	0.12	1.61	.206	0.05	1.91	.145	0.06

Bold values indicate significant results ($p < .05$). Partial eta squared $\eta_p^2 > 0.5$ = large effect size, $\eta_p^2 > 0.3$ = medium effect size, $\eta_p^2 > 0.1$ = small effect size (Bortz & Döring, 2003). H = Harmony, C = Context, G = group, AP = AntPost, L = Laterality. Degrees of freedom were corrected after Huynh-Feldt if necessary.

Table 5 – ANOVA on ALPHA power values with the factors Group x Context x Harmony x Manner x AntPost x Laterality in each time window. N= 15 vs. 15. (HARMONY effects).

		1st tw: 180…370 ms			*2nd tw: 370…550 ms*			*3rd tw: 550…1200 ms*		
Effect	df	F	p	η_P^2	F	p	η_P^2	F	p	η_P^2
G	1,28	12.25	.002	0.3	2.89	.100	0.09	0.01	.934	0.01
H	1,28	8.03	.008	0.22	0.09	.770	0.01	0.71	.405	0.02
H x AP	1,28	0.63	.476	0.02	0.69	.454	0.02	0.44	.644	0.02
H x L	1,28	1.19	.313	0.04	1.85	.167	0.06	7.69	.001	0.22
H x AP x L	2,56	0.74	.526	0.03	1.24	.300	0.04	0.97	.398	0.03
H x C	2,56	0.01	.908	0.01	0.99	.328	0.03	7.61	.010	0.21
H x C x AP	1,28	0.57	.532	0.02	0.06	.862	0.01	1.76	.192	0.06
H x C x L	2,56	3.03	.070	0.1	0.65	.496	0.02	0.11	.862	0.01
H x C x AP x L	2,56	1.20	.312	0.04	0.25	.734	0.01	0.77	.489	0.03
G x H	2,56	0.04	.838	0.01	0.21	.653	0.01	1.37	.252	0.05
G x H x AP	2,56	2.67	.101	0.09	2.79	.092	0.09	0.89	.415	0.03
G x H x L	4,112	1.09	.345	0.04	0.39	.677	0.01	0.69	.505	0.02
G x H x AP x L	2,56	1.10	.354	0.04	0.27	.842	0.01	1.23	.304	0.04
G x H x C	2,56	0.00	.961	0.01	0.74	.398	0.03	1.55	.224	0.05
G x H x C x AP	4,112	0.78	.437	0.03	4.44	.035	0.14	0.49	.550	0.02
G x H x C x L	4,112	0.29	.693	0.01	0.07	.900	0.01	0.33	.682	0.01
G x H x C x AP x L	4,112	0.57	.584	0.02	1.93	.164	0.06	1.52	.223	0.05

Bold values indicate significant results ($p < .05$). Partial eta squared $\eta_P^2 > 0.5$ = large effect size, $\eta_P^2 > 0.3$ = medium effect size, $\eta_P^2 > 0.1$ = small effect size (Bortz & Döring, 2003). H = Harmony, C = Context, G = group, AP = AntPost, L = Laterality. Degrees of freedom were corrected after Huynh-Feldt if necessary.

Table 6 – ANOVA on THETA power values with the factors Group x Context x Harmony x Manner x AntPost x Laterality in each time window. N= 15 vs. 15. (MANNER effects).

		1st tw: 180…420 ms			*2nd tw: 420…800 ms*			*3rd tw: 800…1200 ms*		
Effect	df	F	p	η_P^2	F	p	η_P^2	F	p	η_P^2
G	1,28	6.01	.021	0.18	5.74	.024	0.17	4.18	.050	0.13
M	1,28	2.52	.123	0.08	8.24	.008	0.23	3.88	.059	0.12
M x AP	1,28	0.44	.581	0.02	1.86	.177	0.06	3.38	.059	0.11
M x L	1,28	0.09	.880	0.01	0.87	.426	0.03	0.65	.528	0.02
M x AP x L	2,56	1.13	.342	0.04	0.15	.910	0.01	0.11	.933	0.01
M x C	2,56	0.11	.747	0.01	0.13	.718	0.01	0.01	.917	0.01
M x C x AP	1,28	0.12	.887	0.01	1.02	.355	0.04	1.48	.239	0.05
M x C x L	2,56	1.17	.317	0.04	1.30	.279	0.04	0.15	.859	0.01
M x C x AP x L	2,56	3.14	.042	0.10	2.53	.075	0.08	0.56	.588	0.02
G x M	2,56	1.07	.309	0.04	3.42	.075	0.11	1.45	.238	0.05
G x M x AP	2,56	0.18	.767	0.01	0.17	.775	0.01	2.49	.111	0.08
G x M x L	4,112	0.18	.792	0.01	0.08	.928	0.01	0.10	.908	0.01
G x M x AP x L	2,56	0.19	.894	0.01	0.23	.851	0.01	0.70	.532	0.02
G x M x C	2,56	0.38	.542	0.01	0.03	.861	0.01	0.06	.809	0.01
G x M x C x AP	4,112	1.09	.343	0.04	0.01	.987	0.01	1.93	.168	0.06
G x M x C x L	4,112	1.13	.329	0.04	0.02	.975	0.01	0.05	.953	0.01
G x M x C x AP x L	4,112	1.72	.183	0.06	3.74	.021	0.12	1.02	.372	0.04

Bold values indicate significant results ($p < .05$). Partial eta squared $\eta_P^2 > 0.5$ = large effect size, $\eta_P^2 > 0.3$ = medium effect size, $\eta_P^2 > 0.1$ = small effect size (Bortz & Döring, 2003). M = Manner, C = Context, G = group, AP = AntPost, L = Laterality. Degrees of freedom were corrected after Huynh-Feldt if necessary.

Table 7 – ANOVA on ALPHA power values with the factors Group x Context x Harmony x Manner x AntPost x Laterality in each time window. N= 15 vs. 15. (MANNER effects).

		1st tw: 180…420 ms			*2nd tw: 420…800 ms*			*3rd tw: 800…1200 ms*		
Effect	df	F	p	η_P^2	F	p	η_P^2	F	p	η_P^2
G	1,28	13.18	.001	0.32	1.42	.244	0.05	0.07	.800	0.01
M	1,28	1.87	.182	0.06	1.78	.193	0.06	0.14	.707	0.01
M x AP	1,28	0.25	.651	0.01	0.12,	.767	0.01	0.89	.376	0.03
M x L	1,28	0.11	.859	0.01	0.08	.921	0.01	0.12	.848	0.01
M x AP x L	2,56	2.31	.079	0.08	0.54	.613	0.02	0.85	.494	0.03
M x C	2,56	0.05	.828	0.01	0.00	.988	0.01	0.00	.973	0.01
M x C x AP	1,28	0.98	.352	0.03	0.63	.537	0.02	4.20	.030	0.13
M x C x L	2,56	0.25	.777	0.01	0.44	.647	0.02	0.35	.709	0.01
M x C x AP x L	2,56	0.26	.905	0.01	1.50	.228	0.05	0.74	.489	0.03
G x M	2,56	0.89	.354	0.03	2.70	.111	0.09	1.66	.208	0.06
G x M x AP	2,56	0.22	.676	0.01	0.99	.337	0.03	0.49	.533	0.02
G x M x L	4,112	1.14	.320	0.04	3.88	.026	0.12	0.62	.512	0.02
G x M x AP x L	2,56	0.32	.821	0.01	2.55	.078	0.08	1.78	.138	0.06
G x M x C	2,56	0.05	.831	0.01	0.22	.641	0.01	0.56	.459	0.02
G x M x C x AP	4,112	0.97	.355	0.03	0.54	.587	0.02	0.19	.770	0.01
G x M x C x L	4,112	0.77	.46	0.03	0.16	.849	0.01	1.13	.331	0.04
G x M x C x AP x L	4,112	1.24	.297	0.04	4.46	.011	0.14	0.12	.902	0.01

Bold values indicate significant results ($p < .05$). Partial eta squared $\eta_P^2 > 0.5$ = large effect size, $\eta_P^2 > 0.3$ = medium effect size, $\eta_P^2 > 0.1$ = small effect size (Bortz & Döring, 2003). M = Manner, C = Context, G = group, AP = AntPost, L = Laterality. Degrees of freedom were corrected after Huynh-Feldt if necessary.

Chapter 5

Zusammenfassung der Arbeit

Dissertation zur Erlangung des akademischen Grades Dr. rer. med. Titel
Principles of action planning in music production: evidence from fMRI and EEG studies
eingereicht von
Bianco Roberta
angefertigt in
Max Planck Institute for Human Cognitive and Brain Sciences
betreut von
Arno Villringer and Daniela Sammler

Oktober 2016

> We don't play piano with our
> fingers but with our mind.
>
> *Glenn Gould*

Playing music is not only moving fingers on an instrument. It is a very demanding skill that emerges from the interaction of complex cognitive operations and optimal motor control through intensive practice (Shaffer, 1980). Music production, from jazz improvisation to rehearsed classical concert interpretations, relies on musicians' ability to translate their musical thoughts into bodily action (Keller, 2012). At the heart of this process, musical elements should be appropriately sequenced, and movement parameters of single acts optimally specified. While behavioural and neuroimaging studies have extensively addressed the mechanisms and neural bases

of complex motor behaviours, they often regards fixed sequences of movements (e.g., serial reaction time task) (Hikosaka et al., 2002; Penhune & Steele, 2012), and may fail to account for actions that share the same global structure, but not specific movements. Psychological approaches (Palmer & Pfordresher, 2003; Clarke, 2001) and recent electrophysiological evidence (Sammler et al., 2013) suggest that acquired knowledge of musical structure, e.g., harmony, allows musicians to mentally represent and order musical elements in a sequence based on their structural properties. Therefore, when exposed to unrehearsed harmonic musical contexts, musicians can generate structure-based predictions about the forthcoming (i) sound (Koelsch, 2005), and also (ii) action used to produce that sound, e.g., the next chord, the "what". Then, once they know what to play, each single act of the sequence is motorically implemented by specifying (iii) optimal movement parameters (Verwey et al., 2015), e.g., which fingers to use, the "how". The general question of the present thesis is: How does the musician's brain process all these information and to create meaningful output and to meet specific performance-demands?
Multiple cognitive functions need to interact with multilayer action control processes during performance and the integration of discrete elements into coherent sequences must have a pivotal role. The following studies take expert pianists' long-term knowledge of harmony — defining the arrangement of chords into well-structured musical sequences (Swain, 1995; Rohrmeier & Koelsch, 2012), as the basis to breaking down the complexity of musical sequential actions into simpler components that may be differentially tuned to adapt to specific performance conditions. To do so, our general approach was to test musicians while generating musical actions, but not music (auditory feedback was eliminated from performance). Three hypotheses were investigated. First (Study I), we examined whether high-level action planning based on pianists' knowledge of harmony can be dissociated from auditory information processing, usually coupled in musicians' brains (Zatorre et al., 2007). Second (Study II), we zoomed into fine-grained aspects of action control to find evidence for dissociable high-level structure-based planning of the *action sequence* (Sammler et al., 2013) and low-level parameter specification of *single act*, in a integrated scenario of complex musical actions. Finally (Study III), we explored whether and how musicians adaptively tune these levels of action control to optimally master the specific demands of their genre, i.e., jazz or classical.

Study I (Bianco et al., *in press*) examined whether pianists can motorically represent structural relationships between musical chords of a sequence regardless of auditory feedback and seeks evidence for dissociable contributions of motor and auditory information streams in structure-based predictions. To do so we combined functional connectivity at rest with functional magnetic resonance imaging (fMRI) during an *action* or an *audio* task: classical pianists imitated 5-chord progressions *without sound* by copying a model hand shown in sequences of photos, or they listened to the same progressions *without acting*. Progressions were manipulated in their conformity with Western tonal harmony (harmonically congruent/incongruent), to address structure-based predictions in response to harmonic violations. Greater brain responses elicited by incongruent compared to congruent chords, separately in action and audio task, should reflect more costly computations to integrate elements that mismatch the predicted harmonic structure. Results yielded dissociated dorsal fronto-parietal and ventral fronto-temporal networks for harmonic processing in action and perception, respectively. Unified under the framework of predictive coding (Friston, 2010), these networks are likely to provide the infrastructure that allows frontal areas to keep track of structural relationships in sequential information via dynamic exchange with progressively lower-level modality-specific systems of knowledge, in the parietal or temporal lobe. Importantly, the absence of auditory activation in the action task demonstrates that pianists flexibly decoupled motor from auditory images of the forthcoming chord in the sequence, and that structural knowledge of music can be grounded in the musicians' visual-motor control system.

Study II (Bianco et al., 2016) investigated whether the processes involved in structure-based planning of the *action sequence* (i.e., the order of chords within a larger harmonic context) can be dissociated from how pianists perform *single acts* of the sequence (i.e., fingering applied to each single chord). EEG was recorded from classical pianists while executing on a muted MIDI piano (i) the same harmonically congruent/incongruent chord progressions as in Study I. These were manipulated in two additional factors: Harmonically (in)congruent target chords were placed at the end (ii) of 5- and 2-chord progressions (long/short context), and were further manipulated in terms of (iii) fingering used for execution (conventional/unconventional manner). Note that the long, compared to the short context was expected to provide pianists with greater amount of harmonic information to construct structure-based plans of the *action sequence* such to anticipate forthcoming chords (i.e., the "what"); in turn, signals evoked by execution of harmonic or manner (i.e., the "how") violations should be indicative of two dissociated processes at the level of the *single act*. At the sequence level, results showed that, more in the long than in the short context, pianists were faster and more accurate when playing harmonically congruent chords, while they showed behavioural costs and a late negativity in event-related potentials (ERPs) during execution of chords that harmonically violated the anticipated structure-based plan. At the lower level of single acts, results distinguished these signatures associated with the harmonic plan from those associated with specification of fingering parameters (context-independent response time costs and a late positivity). Therefore, musical actions based on pianists' long-term structural knowledge can be conceptualised as a multilayer action control hierarchy encompassing high-level structural plans, determining the order of single acts in the sequence, i.e., the next chords to play, whose movement parameters are specified at lower level of single acts, i.e., fingering to use.

In **Study III**, we used the same task as in Bianco et al. (2016) that inclusively addressed multilevel action planning during execution of chord sequences. We explored whether the two identified layers of action planning may be modulated in performers with different action focus in their daily practice. Jazz pianists are in fact required to focus more on their structure-generative abilities — to create harmonic sequences in real-time and tune "mistakes" into viable music (Pressing, 1987; Beaty, 2015), whereas classical pianists focus more on their structure-interpretative abilities — to rapidly decode (harmonic) structural cues of the given piece around which expressive strategies (e.g., touch, tempo, choice of fingerings) can then be optimised (Clarke, 2001). Combined behavioural, ERP and time frequency EEG measures showed similar hierarchical core action control processes in both groups. Crucially, they also showed that classical and jazz pianists gave different weights to the two investigated levels, intuitively and despite identical instruction and material. Specifically, at the sequence level, classical pianists displayed indices of cognitive effort to revise the expected structure-plan in response to harmonic violations (context-dependent effects on response time costs, late negativity, and conflict-related power changes in theta band), while jazz pianists were more flexible (no response time costs, earlier negativity, and no conflict-related power change signals). At the single act level, classical compared to jazz pianists showed greater accuracy to set uncommon movement parameters (in terms of fingering errors, power changes in theta and alpha band associated with visuo-motor translation processes), and tended to predict the optimal fingerings to match the high-level harmony-based plan. The observed structure-generative flexibility in jazz pianists and structure-interpretative speed in classical pianists may reflect the adaptive tuning of the action control system that specialises musicians in either creative improvisation or enthralling interpretation. Therefore, the notion of a pliable multilayer organisation of action might help to explain variability in complex behaviours that makes each individual so exceptional.

In conclusion, acquired knowledge of harmony influences pianists' performance at high levels of the action control hierarchy to meaningfully and creatively integrate discrete elements into sequences. Once established, this cognitive-motor interaction does not depend on the auditory system. On top of low-level movement optimisation, the internalisation of high-level/cognitive structures in pianists' motor system may be a key step during training for reaching that extraordinary facility to play, that most other people don't have. In other words, once musicians intuitively know how to play (which fingers to use) and what to play (the musical structure), they can free cognitive resources for more interpretative or improvisational purposes forgetting about the fingers and playing with their mind.

References

Badre, D. and D'Esposito, M. (2009). Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature Reviews. Neuroscience*, 10(9):659–669.

Beaty, R. E. (2015). The neuroscience of musical improvisation. *Neuroscience & Biobehavioral Reviews*, 51:108–117.

Bharucha, J. and Krumhansl, C. L. (1983). The representation of harmonic structure in music: hierarchies of stability as a function of context. *Cognition*, 13(1):63–102.

Bianco, R., Novembre, G., Keller, P. E., Scharf, F., Friederici, A. D., Villringer, A., and Sammler, D. (2016). Syntax in action has priority over movement selection in piano playing: an ERP study. *Journal of Cognitive Neuroscience*, 28(1):41–54.

Brown, R. M., Zatorre, R. J., and Penhune, V. B. (2015). Expert music performance: cognitive , neural , and developmental bases. *Progress in Brain Research*, 217:57–86.

Candidi, M., Maria Sacheli, L., Mega, I., and Aglioti, S. M. (2014). Somatotopic mapping of piano fingering errors in sensorimotor experts: TMS studies in pianists and visually trained musically naives. *Cerebral Cortex*, 24(2):435–443.

Cattaneo, L., Fabbri-Destro, M., Boria, S., Pieraccini, C., Monti, A., Cossu, G., and Rizzolatti, G. (2007). Impairment of actions chains in autism and its possible role in intention understanding. *Proceedings of the National Academy of Sciences of the United States of America*, 104(45):17825–30.

Chaffin, R. and Logan, T. (2006). Practicing perfection: How concert soloists prepare for performance. *Advances in Cognitive Psychology*, 2(2):113–130.

Cisek, P. (2006). Integrated neural processes for defining potential actions and deciding between them: a computational model. *The Journal of Neuroscience*, 26(38):9761–70.

Clarke, E. (2001). Generative principles in music performance. In *Generative Processes in Music: The Psychology of Performance, Improvisation, and Composition*, pages 1–26. Oxford University Press, New York.

Clarke, E., Parncutt, R., Raekallio, M., and Sloboda, J. (1997). Talking fingers: an interview study of pianists' views on fingering. *Musicae Scientiae*, 1(1):87–107.

Cloutman, L. L. (2013). Interaction between dorsal and ventral processing streams: Where, when and how? *Brain and Language*, 127(2):251–263.

Conard, N. J., Malina, M., and Münzel, S. C. (2009). New flutes document the earliest musical tradition in southwestern Germany. *Nature*, 460(7256):737–40.

Dalla Bella, S., Giguère, J. F., and Peretz, I. (2007). Singing proficiency in the general population. *The Journal of the Acoustical Society of America*, 121(2):1182–1189.

D'Ausilio, A., Novembre, G., Fadiga, L., and Keller, P. E. (2015). What can music tell us about social interaction? *Trends in Cognitive Sciences*, 19(3):111–114.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Research*, 93(3):283–321.

Fagioli, S., Ferlazzo, F., and Hommel, B. (2007). Controlling attention through action: Observing actions primes action-related stimulus dimensions. *Neuropsychologia*, 45(14):3351–3355.

Finney, S. a. and Palmer, C. (2003). Auditory feedback and memory for music performance: sound evidence for an encoding effect. *Memory & Cognition*, 31(1):51–64.

Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., and Rizzolatti, G. (2005). Parietal lobe: from action organization to intention understanding. *Science*, 308(5722):662–667.

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological Reviews*, 91(4):1357–92.

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews. Neuroscience*, 11(2):127–138.

Fuster, J. M. (2001). The prefrontal cortex - An update: Time is of the essence. *Neuron*, 30(2):319–333.

Gallivan, J. P. and Culham, J. C. (2015). Neural coding within human brain areas involved in actions. *Current Opinion in Neurobiology*, 33:141–149.

Gellrich, M. and Parncutt, R. (1998). Piano Technique and Fingering in the Eighteenth and Nineteenth Centuries: Bringing a forgotten method back to life. *British Journal of Music Education*, 15(01):5–23.

Goodale, M. a. and Milner, A. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1):20–25.

Grafton, S. T. and Hamilton, A. F. D. C. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science*, 26(4):590–616.

Hickok, G. and Poeppel, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2):67–99.

Hickok, G. and Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews. Neuroscience*, 8(5):393–402.

Hikosaka, O., Nakamura, K., Sakai, K., and Nakahara, H. (2002). Central mechanisms of motor skill learning. *Current Opinion in Neurobiology*, 12(2):217–222.

Hommel, B. (2010). Grounding attention in action control: The intentional control of selection. In B. J. Bruya, editor, *Effortless attention: A new perspective in the cognitive science of attention and action*, pages 121–140. MA: MIT Press, Cambridge.

Hommel, B. (2016). Sequential Action Planning. In Springer, editor, *Human Action Control*, chapter 7, pages 147–166. Switzerland, Springer International Publishing.

Hommel, B., Müsseler, J., Aschersleben, G., and Prinz, W. (2001). The Theory of Event Coding (TEC): a framework for perception and action planning. *The Behavioral and Brain Sciences*, 24(5):849–878.

Johnson-Laird, P. N. (2002). How Jazz Musicians Improvise. *Music Perception: An Interdisciplinary Journal*, 19(3):415–442.

Keller, P. E. (2012). What movement force reveals about cognitive processes in music performance. In *Art in motion II*, pages 115–153. Peter Lang, Frankfurt.

Keller, P. E. and Koch, I. (2008). Action planning in sequential skills: relations to music performance. *Quarterly Journal of Experimental Psychology*, 61(2):275–91.

Koechlin, E. and Summerfield, C. (2007). An information theoretical approach to prefrontal executive function. *Trends in Cognitive Sciences*, 11(6):229–235.

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, 15(2):207–212.

Koelsch, S. (2011). Toward a neural basis of music perception - a review and updated model. *Frontiers in Psychology*, 2:110.

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., and Schlaug, G. (2005). Adults and children processing music: an fMRI study. *Neuroimage*, 25(4):1068–1076.

Koelsch, S. and Siebel, W. a. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, 9(12):578–84.

Kravitz, D. J., Saleem, K. S., Baker, C. I., and Mishkin, M. (2011). A new neural framework for visuospatial processing. *Nature Reviews. Neuroscience*, 12:217–230.

Krumhansl, C. L. (1983). Perceptual structures for tonal music. *Music Perception*, 1(1):28–62.

Large, E. W. and Palmer, C. (2002). Perceiving temporal regularity in music. *Cognitive Science*, 26(1):1–37.

Lashley, K. (1951). The problem of serial order in behavior. In Jeffress, L., editor, *Cerebral Mechanisms in Behavior*, pages 112–131. New York: Wiley.

Lerdahl, F. and Jackendoff, R. (1983). *A generative theory of tonal music.* MIT Press, Cambridge.

Levitin, D. J. (2006). What Makes a Musician? Expertise Dissected. In *This Is Your Brain on Music: The Science of a Human Obsession*, pages 189–216. Dutton, New York.

Loui, P. (2015). A Dual-Stream Neuroanatomy of Singing. *Music perception*, 32(3):232–241.

MacKay, D. G. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review*, 89(5):483–506.

Maess, B., Koelsch, S., Gunter, T. C., and Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, 4(5):540–545.

Maidhof, C., Vavatzanidis, N., Prinz, W., Rieger, M., and Koelsch, S. (2009). Processing expectancy violations during music performance and perception: An ERP Study. *Journal of Cognitive Neuroscience*, 22(10):2401–2413.

Münte, T. F., Altenmüller, E., and Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews. Neuroscience*, 3(6):473–8.

Novembre, G. and Keller, P. E. (2011). A grammar of action generates predictions in skilled musicians. *Consciousness and Cognition*, 20(4):1232–43.

Palmer, C. and Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review*, 110(4):683–712.

Palmer, C. and van de Sande, C. (1995). Range of planning in music performance. *Journal of Experimental Psychology. Human perception and Performance*, 21(5):947–62.

Parncutt, R. (2014). Fingering. In *Music in the social and behavioral sciences: An encyclopedia*, pages 481–484. Los Angeles: Sage, w. f. thom edition.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, 6(7):674–81.

Penhune, V. B. and Steele, C. J. (2012). Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. *Behavioural Brain Research*, 226(2):579–591.

Pfordresher, P. Q. (2006). Coordination of perception and action in music performance. *Advances in Cognitive Psychology*, 2(2):183–198.

Pfordresher, P. Q. (2012). Musical training and the role of auditory feedback during performance. *Annals of the New York Academy of Sciences*, 1252(1):171–178.

Pinho, A. L., de Manzano, O., Fransson, P., Eriksson, H., and Ullén, F. (2014). Connecting to create: expertise in musical omprovisation is associated with increased functional connectivity between premotor and prefrontal Areas. *The Journal of Neuroscience*, 34(18):6156–63.

Pressing, J. (1984). The history of classical improvisation. A thousand years of fluid performance traditions. *Keyboard*, pages 64–68.

Pressing, J. (1987). Improvisations: methods and models. In *Generative Processes in Music*, pages 129–178. Oxford University Press.

Prinz, W. (2002). What re-enactment earns us. *Cortex*, 42:515–517.

Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing Research*, 271(1-2):16–25.

Rauschecker, J. P. and Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature neuroscience*, 12(6):718–24.

Repp, B. H. (2000). Pattern typicality and dimensional interactions in pianists' imitation of expressive timing and dynamics. *Music Perception*, 18(2):173–211.

Rohrmeier, M. (2011). Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, 5(1):35–53.

Rohrmeier, M. and Koelsch, S. (2012). Predictive information processing in music cognition. A critical review. *International Journal of Psychophysiology*, 83(2):164–175.

Rosenbaum, D. a., Cohen, R. G., Jax, S. a., Weiss, D. J., and van der Wel, R. (2007). The problem of serial order in behavior: Lashley's legacy. *Human Movement Science*, 26(4):525–54.

Sammler, D., Grosbras, M.-H., Anwander, A., Bestelmeyer, P. E., and Belin, P. (2015). Dorsal and ventral pathways for prosody. *Current Biology*, 25(23):3079–3085.

Sammler, D., Koelsch, S., and Friederici, A. D. (2011). Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? *Cortex*, 47(6):659–73.

Sammler, D., Novembre, G., Koelsch, S., and Keller, P. E. (2013). Syntax in a pianist's hand: ERP signatures of "embodied" syntax processing in music. *Cortex*, 49(5):1325–1339.

Schmidt, R. a. R. (1975). A schema theory of discrete motor skill learning. *Psychological Review*, 82(4):225–260.

Shaffer, L. H. (1980). 26 Analysing Piano Performance: A Study of Concert Pianists. *Advances in Psychology*, 1:443–455.

Shaffer, L. H. (1984). Timing in solo and duet piano performances. *The Quarterly Journal of Experimental Psychology*, 36(4):577–595.

Sloboda, J. a., Clarke, E. F., Parncutt, R., and Raekallio, M. (1998). Determinants of finger choice in piano sight-reading. *Journal of Experimental Psychology: Human Perception and Performance*, 24(1):185–203.

Swain, J. P. (1995). The concept of musical syntax. *Musical Quarterly*, 79(2):281–308.

Tillmann, B. (2012). Music and language perception: expectations, structural integration, and cognitive sequencing. *Topics in Cognitive Science*, 4:568–584.

Trehub, S. E. (2003). The developmental origins of musicality. *Nature Neuroscience*, 6(7):669–73.

Uithol, S., van Rooij, I., Bekkering, H., and Haselager, P. (2012). Hierarchies in action and motor control. *Journal of Cognitive Neuroscience*, 24(5):1077–86.

Vallacher, R. R. and Wegner, D. M. (1987). Action identification and human behavior. *Psychological Review*, 94(1):3–15.

van der Steen, M. C., Molendijk, E. B. D., Altenmueller, E., and Furuya, S. (2014). Expert pianists do not listen: The expertise-dependent influence of temporal perturbation on the production of sequential movements. *Neuroscience*, 269:290–298.

Verwey, W. B., Shea, C. H., and Wright, D. L. (2015). A cognitive framework for explaining serial processing and sequence execution strategies. *Psychological Bulletin*, 22:54–77.

Wohlschläger, A., Gattis, M., and Bekkering, H. (2003). Action generation and action perception in imitation: an instance of the ideomotor principle. *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences*, 358(1431):501–15.

Wolpert, D. M., Diedrichsen, J., and Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews. Neuroscience*, 12(12):739–51.

Wulf, G., Höß, M., and Prinz, W. (1998). Instructions for motor learning: Differential effects of internal vs. external focus of attention. *Journal of Motor Behavior*, 30:169–179.

Wulf, G., McNevin, N., and Shea, C. H. (2001). The automaticity of complex motor skill learning as a function of attentional focus. *Quarterly Journal of Experimental Psychology*, 54(4):1143–1154.

Wulf, G. and Shea, C. H. (2002). Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychonomic Bulletin & Review*, 9(2):185–211.

Zarate, J. M. (2013). The neural control of singing. *Frontiers in Human Neuroscience*, 7(237):237.

Zarate, J. M., Wood, S., and Zatorre, R. J. (2010). Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers. *Neuropsychologia*, 48(2):607–618.

Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nature Reviews. Neuroscience*, 8(7):547–558.

List of Figures

List of Abbreviations

BA	Brodmann area
CG	Classical group
JG	Jazz group
EEG	Electroencephalography
ERP	Event Related Potential
fMRI	functional Magnetic Resonance Imaging
IFG	Inferior Frontal Gyrus
MEG	Magnetoelectroencephalography
MIDI	Musical Instrument Digital Interface
pSTG	posterior Superior Temporal Gyrus
SPL	Superior Parietal Lobe
TMS	Trancranial Magnetic Stimulation
vPMC	ventral Premotor Cortex

Erklärung über die eigenständige Abfassung der Arbeit

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne unzulässige Hilfe oder Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Ich versichere, dass Dritte von mir weder unmittelbar noch mittelbar eine Vergütung oder geldwerte Leistungen für Arbeiten erhalten haben, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen, und dass die vorgelegte Arbeit weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde zum Zweck einer Promotion oder eines anderen Prüfungsverfahrens vorgelegt wurde. Alles aus anderen Quellen und von anderen Personen übernommene Material, das in der Arbeit verwendet wurde oder auf das direkt Bezug genommen wird, wurde als solches kenntlich gemacht. Insbesondere wurden alle Personen genannt, die direkt an der Entstehung der vorliegenden Arbeit beteiligt waren. Die aktuellen gesetzlichen Vorgaben in Bezug auf die Zulassung der klinischen Studien, die Bestimmungen des Tierschutzgesetzes, die Bestimmungen des Gentechnikgesetzes und die allgemeinen Datenschutzbestimmungen wurden eingehalten. Ich versichere, dass ich die Regelungen der Satzung der Universität Leipzig zur Sicherung guter wissenschaftlicher Praxis kenne und eingehalten habe.

Roberta Bianco
Leipzig, den 27. Oktober 2016