



Research report

Syntax in a pianist's hand: ERP signatures of “embodied” syntax processing in music

Daniela Sammler^{a,*}, Giacomo Novembre^{b,1}, Stefan Koelsch^c and Peter E. Keller^b

^aDepartment of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

^bResearch Group “Music Cognition and Action”, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

^cCluster of Excellence “Languages of Emotion”, Free University Berlin, Berlin, Germany

ARTICLE INFO

Article history:

Received 11 November 2011

Reviewed 03 January 2012

Revised 2 April 2012

Accepted 13 June 2012

Action editor Stefano Cappa

Published online 29 June 2012

Keywords:

Music

Syntax

Action

Imitation

Embodiment

ABSTRACT

Syntactic operations in language and music are well established and known to be linked in cognitive and neuroanatomical terms. What remains a matter of debate is whether the notion of syntax also applies to human actions and how those may be linked to syntax in language and music. The present electroencephalography (EEG) study explored syntactic processes during the observation, motor programming, and execution of musical actions. Therefore, expert pianists watched and imitated silent videos of a hand playing 5-chord sequences in which the last chord was syntactically congruent or incongruent with the preceding harmonic context. 2-chord sequences that diluted the syntactic predictability of the last chord (by reducing the harmonic context) served as a control condition. We assumed that behavioural and event-related potential (ERP) effects (i.e., differences between congruent and incongruent trials) that were significantly stronger in the 5-chord compared to the 2-chord sequences are related to syntactic processing. According to this criterion, the present results show an influence of syntactic context on ERPs related to (i) action observation and (ii) the motor programming for action imitation, as well as (iii) participants' execution times and accuracy. In particular, the occurrence of electrophysiological indices of action inhibition and reprogramming when an incongruent chord had to be imitated implies that the pianist's motor system anticipated (and revoked) the congruent chord during action observation. Notably, this well-known anticipatory potential of the motor system seems to be strongly based upon the observer's music-syntactic knowledge, thus suggesting the “embodied” processing of musical syntax. The combined behavioural and electrophysiological data show that the notion of musical syntax not only applies to the auditory modality but transfers – in trained musicians – to a “grammar of musical action”.

© 2012 Elsevier Ltd. All rights reserved.

* Corresponding author. Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany.
E-mail address: sammler@cbs.mpg.de (D. Sammler).

¹ These authors contributed equally to the study.

1. Introduction

From a structural point of view, a linguistic sentence, a musical phrase and a goal-directed action share one key property: all are composed of discrete items (words, tones, motor acts) that are strung together according to specific rules (language-specific grammars, culture dependent tonal systems, motor constraints) to build-up meaning over the course of a sequence (Chomsky, 1957; Lashley, 1951; Lerdahl and Jackendoff, 1983). Modern comparative research further stresses analogies between these domains in terms of hierarchical organization and recursion, concepts that are particularly well established in language (Chomsky, 1957, 1995), and increasingly substantiated in harmonic structure in music (Katz and Pesetsky, 2011; Rohrmeier, 2011) as well as complex actions (Pastra and Aloimonos, 2012). We will refer to this shared property – i.e., the rule-based hierarchical and compositional ordering of discrete elements into sequences – as “syntax”. This term is clearly borrowed from (psycho)linguistics, a field that describes the organization of language, discusses the universals (Chomsky, 1986, 1995; Fitch, 2011; Moro, 2008) and essentials (Hauser et al., 2002) of the language faculty, and – most relevant to us – assumes a neural system that may be dedicated to the processing of syntax in natural languages (Moro et al., 2001; Musso et al., 2003; Pallier et al., 2011).

However, the intriguing parallels of “syntactic structure” in language, music, and perhaps action lead to the question whether this neural system and the cognitive operations necessary to isolate, process, and integrate syntactically organized elements are specific to language or may be shared between domains. The strongest evidence in favour of shared syntactic resources comes from investigations on speech and music perception (Koelsch, 2011; Patel, 2003). Here it has been shown that syntactic violations in the two domains elicit comparable electric brain potentials (Koelsch, 2005; Patel et al., 1998) and activate overlapping brain regions (Abrams et al., 2011; Sammler et al., 2009) including Broca’s area and its right hemisphere homotope (Maess et al., 2001), i.e., areas that have been typically associated with syntactic processing in language (Grodzinsky and Santi, 2008). Moreover, simultaneous presentations of syntactic errors in music and language evoke interference effects (Fedorenko et al., 2009; Koelsch et al., 2005; Slevc et al., 2009; Steinbeis and Koelsch, 2008), neurological patients show parallel syntactic deficits in both domains (Grodzinsky, 2000; Patel et al., 2008; Sammler et al., 2011), and syntactic capabilities in one domain are enhanced after training in the other domain (Jentschke and Koelsch, 2009; Jentschke et al., 2005; Marin, 2009). In other words, these combined findings gave rise to the idea that the brain’s dedicated syntax network (Friederici, 2011; Kaan and Swaab, 2002; Moro et al., 2001; Pallier et al., 2011) may be less language-specific than initially thought.

Since Lashley’s seminal article on the structural principles of goal-related actions (Lashley, 1951), it is a matter of debate whether the notion of syntax also applies to human actions. Most recent work has been dedicated to the formalization of the compositional (Guerra-Filho and Aloimonos, 2012; see also Zacks and Tversky, 2001) and generative organization of actions (Pastra and Aloimonos, 2012) in comparison to the

syntactic organization of language. Moreover, several studies have aimed at clarifying whether the cognitive processes (Allen et al., 2010; Greenfield, 1991) and underlying neural correlates (Farag et al., 2010; van Schie et al., 2006) that operate on compositional action structures are the same as the ones found in language and music. Similar parallels have also been discussed for visuo-spatial sequencing (Bahlmann et al., 2009; Tettamanti et al., 2009), logic (Monti et al., 2009) or arithmetic (Dehaene, 1997; Friedrich and Friederici, 2009; Nuñez-Peña and Honrubia-Serrano, 2004; Scheepers et al., 2011; although the rules of “syntactic” combination have to be explicitly taught in arithmetic, while they are implicitly acquired in language, music and simple actions, such as goal-related grasping).

Most authors adopt the view of a domain-general hierarchical syntax processor in the inferior frontal lobe (Fadiga et al., 2009; Fiebach and Schubotz, 2006; Gelfand and Bookheimer, 2003; Koelsch, 2011; Patel, 2003; Tettamanti and Weniger, 2006), although this is not yet unequivocally proven (Rogalsky et al., 2011). An alternative approach pertains to a polymodal sensorimotor theory of syntax, i.e., the involvement of action-perception circuits to mediate grammar processing in language (Pulvermüller and Fadiga, 2010; van Schie et al., 2006), music (Fadiga et al., 2009), and action (Clerget et al., 2009; Fazio et al., 2009). Although it is not clear how rule-based structures might be processed in sensorimotor areas alone (i.e., by means of a “mirror” mechanism without the recruitment of an extra parser that processes syntactic dependencies, see Tettamanti and Moro, 2012), it is possible that the motor system makes use of syntactic operations during the perception and production of sequences of acts forming goal-directed actions. Some evidence for such a syntax-action link can be inferred from models of incremental planning of serial actions such as speech or music (for a review, see Palmer and Pfordresher, 2003). These models suggest that the ongoing advance construction of motor programs during musical performance is governed by musical structure, e.g., melodic, harmonic or metrical relationships between tones and chords of a musical piece, whose statistical regularities have been acquired over the course of experience (Palmer and van de Sande, 1993, 1995; Restle, 1970).

In a recent behavioural study, Novembre and Keller (2011) explored the impact of syntactic knowledge on musical actions by means of an imitation paradigm. Expert pianists watched and imitated videos displaying one hand performing sequences of chords, including occasional chords that were harmonically, that is syntactically, incongruent with the preceding musical context (i.e., the events that precede the target chord and whose syntactic structure influences how the target chord is perceived). The experiment was run in the absence of sound. Results showed that imitation of chords was faster when they were embedded in a congruent (i.e., syntactically regular) context, suggesting that the harmonic rules implied by the observed actions induced strong expectancies that influenced action execution. Therefore, this study provided evidence in favour of syntactic structures regulating the progression of motor acts associated with producing music. The authors suggested that, as a result of musical training, the rules determining the position of chords within

chord sequences are internalized as a form of “embodied harmony”, i.e., that the motor system of skilled musicians makes use of syntactic rules in the perception and production of musical actions. On a more abstract level, this notion alludes to theories of embodied cognition that ground cognition in the bodily senses and mental simulation (Barsalou, 2008; Gibbs, 2006; Wilson, 2002) instead of segregating body and mind.

The present study set out to test further the hypothesis of “embodied” processing of harmony and zoomed into the neurophysiological correlates of syntactic operations during the observation and imitation of musical actions. Particularly, we aimed to reveal the time course and origin of the syntactic interference effects described by Novembre and Keller (2011), i.e., the influence of syntax on (i) the observation of musical performance, (ii) the translation of observed movements into a motor program, and (iii) the execution of the movements themselves. Therefore, electroencephalogram (EEG) and piano performance were recorded while skilled pianists watched and imitated the videos employed by Novembre and Keller (2011) displaying 5-chord sequences with and without syntactic violations. In order to control for differences between target chords other than syntactic congruity, such as visual appearance or motoric complexity (see Methods), we included an additional set of videos displaying 2-chord sequences. These videos kept the visual and motoric aspects of the target chords invariant, but diluted the music-syntactic predictability of the sequences by reducing harmonic context information (Fig. 1). Note that no sound was presented during the whole experiment, neither in the videos nor on the piano used by the participants in the imitation task.

We predicted that the harmonic rules – as implied by the observed and imitated movements – would induce motor expectations. Thus, the perception and imitation of the last chord should elicit distinct electrophysiological brain responses, imitation time and accuracy, depending on the chord’s

congruency with the preceding harmonic context. From what is known from auditory studies, the observed music-syntactic violations may evoke an early right anterior negativity (ERAN) – i.e., an electrophysiological marker of early musical structure building – and an N500 or P600 both reflecting later stages of syntactic integration (Koelsch, 2009; Koelsch et al., 2000; Patel et al., 1998), although these chord sequences have before never been studied in the context of action. Importantly, based on the observation that the sense of a tonal centre and corresponding music-syntactic expectancies usually gain strength over the course of a musical piece (Bigand and Parncutt, 1999; Koelsch et al., 2000; Leino et al., 2007), we assumed that the 5-chord sequences should induce stronger syntactic expectancies than the 2-chord sequences. Hence, any behavioural or event-related potential (ERP) effect related to the processing of syntax in musical actions should be stronger in the 5-chord than 2-chord sequences, i.e., evidenced by a statistical interaction of Congruency (congruent/incongruent) \times Context (5-chord/2-chord sequences). (The factor Context relates to the number of events that precede the target chords and whose syntactic structure influences how the target chords are perceived and imitated). This criterion, in combination with the excellent temporal resolution of the EEG, should allow us to specify the time course of interaction between perceptuo-motor and syntactic processes. Ultimately, the present paradigm should permit us to specify the neurophysiological signatures and computational underpinnings of putative “embodied” syntactic processes in action.

2. Methods

2.1. Participants

Twenty-seven right-handed pianists (nine males), aged 20–34 years [mean = 24.93, standard deviation (SD) = 3.55], were

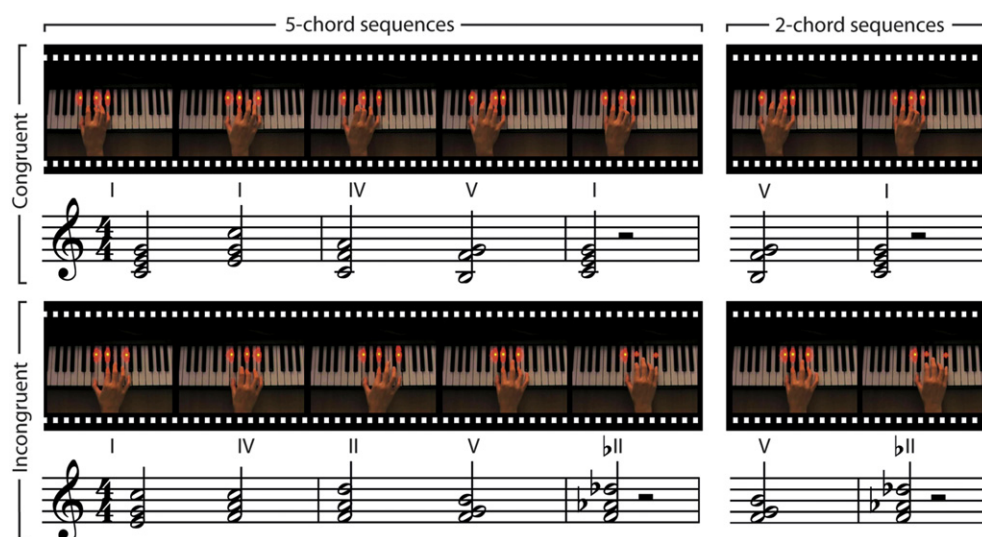


Fig. 1 – Experimental design. Participants watched and imitated silent videos showing a right hand playing chord sequences composed of five or two chords. Two-chord sequences were derived from the 5-chord sequences by deleting the first three chords. The final target chord of each sequence was either syntactically congruent (i.e., Tonic [I]) or incongruent (i.e., the major chord of the lowered second scale degree [bII]). Scores are only shown for illustration and were not presented to the pianists. The experiment was run in the absence of sound.

included in the analysis. Three additional participants were tested but excluded from data analysis because they were not able to perform the task. All pianists had a minimum of 14 years of formal training in classical music, the mean age at which piano studies commenced was 6.31 years ($SD = 1.52$), mean years of piano training was 16.96 years ($SD = 3.45$), and average weekly amount of practice was 7.70 h ($SD = 10.67$). All pianists were naïve with regard to the purpose of the study.

2.2. Stimuli

Stimuli were identical to a portion of those used by [Novembre and Keller \(2011\)](#). They consisted of silent videos showing a female pianist's right hand playing sequences of chords on a muted keyboard (Yamaha EZ200) equipped with red light-emitting diodes (LEDs). These LEDs were illuminated for the duration of each key press and made the identity of the pressed keys clear to the participant (cf. [Novembre and Keller, 2011](#)).

A total of 60 different chord sequences were used for this experiment: 30 were syntactically regular in the sense that they came to a conventional harmonic resolution (congruent condition; upper left panel in [Fig. 1](#)) and 30 were irregular in that they ended in an uncommon and unresolved harmony (incongruent condition; lower left panel in [Fig. 1](#)). For each condition, the chord sequences were in the key of C, D, or F major (10 sequences/key). All chords consisted of three piano keystrokes. The first chord was always the tonic of the given tonal context, and was followed either by a tonic, supertonic, or subdominant at the second position. Chords at the third position were the tonic, subdominant, supertonic or submediant. At the fourth position dominant seventh chords were presented in root position, or in first or third inversion. The chord at the fifth position was different between the two conditions: a tonic chord (congruent condition) or the major chord built on the lowered second scale degree (incongruent condition). Tonic chords were presented in root position, first, and second inversions. Incongruent chords were presented in both first and second inversions. In general, chord sequences had different 'melodic contours' (e.g., starting with the first, third, or fifth degree of the tonic chord) in the top voice.

It is important to note that the videos displayed non-manipulated biological movements (apart from the first chord; see below) – as recorded in natural piano playing – which were intended to maximally activate the observers' motor system ([Buccino et al., 2004](#); [Perani et al., 2001](#); [Stevens et al., 2000](#)). This implied, however, that the spatial trajectory performed by the model hand moving from the penultimate to the incongruent target chords was significantly longer (mean trajectory duration from movement onset to offset = 303 msec, $SD = 86$ msec) than when moving to the congruent targets [mean = 221 msec, $SD = 69$ msec; $t(58) = -4.07$, $p < .001$]. Moreover, other visual aspects such as movement fluency, finger configuration and number of black keys, along with motor task complexity and familiarity could not be kept entirely constant between congruent and incongruent target chords, necessitating an appropriate control condition to isolate syntax-related brain activity. Therefore, we included an additional set of 60 excerpt-videos displaying only the last two chords from the 5-chord sequences described above, i.e., 2-chord sequences (right panel of [Fig. 1](#)). Note that the control videos were truncated

versions of the original videos of the 5-chord sequences. As a result, the target chords of 5- and 2-chord sequences were physically identical (and thus also identical in terms of visual appearance, motoric complexity and familiarity), and merely differed in their syntactic predictability. In other words, the longer (5-chord) music-syntactic context should induce a stronger sense of tonality ([Bigand and Parncutt, 1999](#); [Koelsch et al., 2000](#); [Leino et al., 2007](#)) and thus stronger syntactic expectancies than the shorter (2-chord) sequences. Consequently, any behavioural or electrophysiological effect that is significantly stronger in 5- than 2-chord sequences – i.e., reflected in a statistical interaction of Congruency \times Context – should be clearly attributable to enhanced syntactic processing.

Nevertheless, it should be noted that the 2-chord sequences also contained a certain degree of syntactic information (as they constituted common 2-chord progressions in the Western tonal system). This implies (1) that we manipulated the amount of syntactic information rather than its presence or absence, and (2) that therefore the comparison of 5- and 2-chord sequences may cancel out some syntactic processing aspects.

Each video started with a stationary hand poised to press the three keys associated with the first chord for 3 sec, to give the participant enough time to match the initial position of his or her own hand with the position of the model hand in the video. After that, the model hand executed the chord progression with each chord lasting approximately 2 sec, leading to video durations of 13 sec and 7 sec for 5- and 2-chord sequences, respectively. Because data of interest were the brain responses to the perception and imitation of the last chord in each sequence, the presentation of this chord was time-locked to the video onset. This was done by decomposing each video into its constituent frames (of which there were 30/sec), extending or shortening the first chord, and thus moving the first frame in which the model hand pressed all three keys of the target chord (i.e., all three LEDs were on) to 11 sec (for the 5-chord sequences) or 5 sec (for the 2-chord sequences) after video onset (videos were edited using the software iMovie HD 6.0.3, Apple Computer, Inc.).

2.3. Procedure

Participants were asked to watch and simultaneously imitate the silent videos, which were presented on a computer monitor placed on a musical instrument digital interface (MIDI) piano (Yamaha Clavinova CLP150). They were instructed to imitate both the key presses and the fingerings as fast and correctly as possible with their right hand. Furthermore, they were asked to move as minimally as possible to avoid muscle artefacts in the EEG. Each trial started with a visual fixation cross presented for 500 msec.

Sixty 5-chord and sixty 2-chord sequences were presented separately in two blocks, which were repeated once in order to increase statistical power (resulting in 240 chord sequences in total). The order of the blocks alternated and was counter-balanced across participants (e.g., 5-chords, 2-chords, 5-chords, 2-chords). Trials within each block were randomized individually for each participant. To increase participants' familiarity with the stimuli and accuracy in the task, the experiment started with a short training session consisting of

a short 5-chord and 2-chord block in counterbalanced order, each comprising 20 sequences (10/condition, in the key of G major). To control for individual differences in task strategy, participants were asked to fill in a questionnaire at the end of the experiment. Specifically, they rated (from 1 to 9) to what extent they relied on auditory and/or motor imagery, and their theoretical knowledge of western harmony, in order to perform the task efficiently.

Presentation software (Version 14.2, Neurobehavioral Systems, Inc.) was used to control both stimulus presentation (i.e., videos) and response registration (i.e., keystrokes on the piano). A MIDI interface converted the MIDI key values received from the piano keyboard into a serial signal that was compatible with Presentation software. This permitted us to compute the times at which specific keys were struck in relation to event timing in the video. Additionally, a video camera (Sony, HDR-HC9E) placed above the piano recorded the performed fingering in the participant's hand from an aerial perspective.

2.4. EEG data acquisition

EEG was recorded from 61 Ag/AgCl electrodes mounted in an elastic cap according to the extended international 10–20 system (Sharbrough et al., 1991). The electrode positions were: FPZ, FP1, FP2, AFZ, AF3, AF4, AF7, AF8, FZ, F1, F2, F3, F4, F5, F6, F7, F8, FCZ, FC1, FC2, FC3, FC4, FC5, FC6, FT7, FT8, CZ, C1, C2, C3, C4, C5, C6, T7, T8, CPZ, CP1, CP2, CP3, CP4, CP5, CP6, TP7, TP8, PZ, P1, P2, P3, P4, P5, P6, P7, P8, POZ, PO3, PO4, PO7, PO8, OZ, O1, O2. Left mastoid (M1) served as reference; an additional electrode was placed on the right mastoid bone (M2) and the tip of the nose for off-line re-referencing. The ground electrode was located on the sternum. Horizontal and vertical electrooculograms were bipolarly recorded from electrodes placed on the outer canthus of each eye, as well as above and below the right eye. Impedances were kept below 5 k Ω . Signals were amplified with a 24 bit Brainvision QuickAmp 72 amplifier (Brain Products GmbH, Gilching, Germany) and digitized with a sampling rate of 500 Hz.

2.5. Behavioural data analysis

Errors and response times (RTs) for imitation of the target chord (i.e., the last chord) of each trial were analyzed in accordance with Novembre and Keller (2011). If both the last and the second-last chords had been correctly imitated in terms of the keys pressed and the fingering employed, then a trial was considered to be correct. Chords in which the keystrokes were not synchronous (i.e., when more than 150 msec intervened between the first and the last keystroke) were excluded from analysis (cf. Drost et al., 2005). Errors were counted if the target chord was incorrectly imitated in terms of the keys pressed, the fingering employed, or both. Errors were counted only if the previous chord (i.e., the second-last chord) had been correctly imitated in terms of both keys and fingering.

RTs were measured in correct trials by calculating the time elapsed between the presentation of the target chord (i.e., the frame in which the model hand struck all three target keys) and the participant's execution of the same chord (i.e., mean

of the three keystroke times composing the chord). RTs exceeding 3000 msec were not analyzed (cf. Drost et al., 2005). Statistical analyses were conducted on errors and RT data using separate two-way repeated measures analyses of variance (ANOVAs) with the variables Congruency (congruent/incongruent) and Context (5-chord/2-chord sequences).

2.6. EEG data analysis

EEP 3.2 (ANT-software) was used to re-reference the data to the algebraic mean of both mastoid leads. Further processing steps were done using EEGLAB 6.01 (Delorme and Makeig, 2004) in MATLAB 7.7. Data were filtered using a .3 Hz high-pass filter (fir, 5854 points, Blackman window), and strong muscle artifacts, electrode drifts or technical artifacts were manually rejected before entering the continuous data into an independent component analysis. The resulting component structure was used to reject eye movement and blink artifacts, muscle artifacts and slow drifts. Afterwards, the data were filtered with a 25 Hz lowpass filter (fir, 110 points, Blackman window), and cut into epochs ranging from –800 to 1000 msec relative to the target chord in the videos (i.e., the frame when all three target keys were pressed). Only correct trials (i.e., mean \pm SD long context: 39.02 \pm 10.18, short context: 49.61 \pm 6.61; according to the criteria in the behavioural data) were included in the ERP analysis. Trials were rejected whenever one or more electrodes exhibited voltages of \pm 50 μ V. Altogether, this procedure allowed the complete elimination of movement artifacts caused by the imitation task, e.g., eye movements between screen and keyboard or tension of neck and shoulder muscles during playing. Non-rejected trials were averaged separately for each condition. Averages were aligned to a –800 to –300 msec baseline, i.e., to a time in the video during which the model hand rested on the keys of the penultimate chord, prior to the trajectory onset towards the target chord. An average of 39.17 trials was included for each participant and each condition (mean \pm SD long context: 35.67 \pm 11.16, short context: 42.67 \pm 10.02).

Effects of chord congruency and context length were analyzed time-locked to the target chord in the video, i.e., the point when the model hand struck the keys of the 5th chord in the 5-chord sequences and the 2nd chord in the 2-chord sequences. Statistical analyses were carried out on the mean amplitudes in each condition calculated for specific time windows (see Results) in nine regions of interest (ROIs): (i) left anterior (F3, F5, F7, FC3, FC5, FT7), (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), (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), (viii) right central (C4, C6, T8, CP4, CP6, TP8), and (ix) right posterior (P4, P6, P8, PO4, PO8). Four time windows were defined separately in 5- and 2-chord sequences by visual inspection of the ERPs and topography plots according to the following criteria: assuming that different map topographies and polarities directly indicate different underlying generators, i.e., different cognitive processes (Michel et al., 2004), borders between time windows were set whenever the topography shifted or polarity of the effect flipped (for details, see Results). Note that this approach generated a different border between the first

and second time window in 5-chord (–80 msec) and 2-chord sequences (0 msec). This is most likely due to the better syntactic (and temporal) predictability of the 5- compared to the 2-chord sequences, possibly leading to an acceleration of cognitive processes and their related ERP components.

Statistical evaluation comprised a four-way ANOVA with the repeated measures factors Congruency (congruent/incongruent) \times Context (5-chord/2-chord sequences) \times AntPost (anterior/central/posterior) \times Laterality (left/middle/right). Whenever an interaction involving the factor Congruency was found, follow-up analyses were carried out by splitting up the factorial model.

3. Results

3.1. Behavioural data

Fig. 2A shows mean RTs for correctly produced target chords in each condition. A two-way ANOVA with the repeated measures factors Congruency (congruent/incongruent) and Context (5-chord/2-chord sequences) yielded a significant main effect of Congruency [$F(1,26) = 98.89, p < .001$] and a significant Congruency \times Context interaction [$F(1,26) = 13.98, p < .002$]. This indicates that imitation of congruent chords was overall faster than imitation of incongruent chords, and fastest when a congruent chord was embedded in a 5-chord than 2-chord sequence. Notably, *t*-tests for paired samples showed that the congruent chord in 5-chord sequences was executed significantly faster than in 2-chord sequences [$t(26) = -3.02, p < .007$], whereas no significant difference was found between incongruent chords across long and short contexts [$t(26) = .422, p > .676$]. This suggests that the extended harmonic context facilitated the execution of the congruent chord (rather than interfering with the execution of the incongruent chord). The main effect of Context was not significant [$F(1,26) = 2.09, p > .159$] demonstrating that imitation of the target chords across 5- and 2-chord sequences did not differ in terms of RT.

A similar trend was observed in the mean number of errors, as depicted in Fig. 2B. Less errors were committed during imitation of congruent compared to incongruent chords (main effect of Congruency [$F(1,26) = 23.17, p < .001$]) while errors did not differ between 5- and 2-chord sequences (no main effect of Context [$F(1,26) = 2.59, p > .119$]). Although particularly few errors were produced in the congruent condition in the 5-chord (compared to 2-chord) sequences, the Congruency \times Context interaction fell short of statistical significance [$F(1,26) = 1.56, p > .222$]. Consistently with what was observed for the RTs, *t*-tests for paired samples showed that significantly fewer errors were produced during imitation of congruent target chords in 5- compared to 2-chord sequences [$t(26) = -3.705, p < .002$], whereas no significant difference was found between incongruent chords between long and short contexts [$t(26) = -.251, p > .803$].

3.2. EEG data

In both the 5- and 2-chord sequences a four-phasic ERP pattern was found (Fig. 3), each phase will be described in turn.

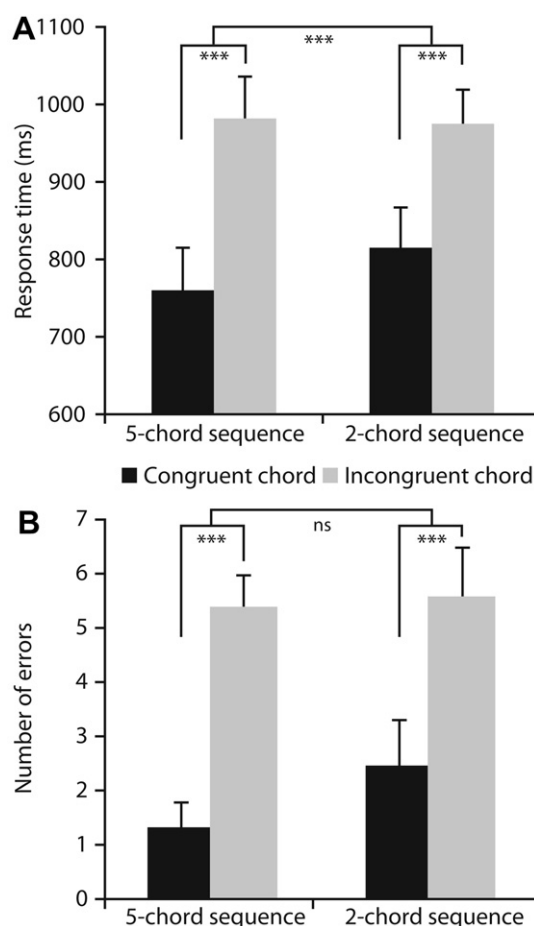


Fig. 2 – Behavioural data. (A) RTs time-locked to the key press in the video. (B) Number of errors. Error bars indicate one standard error of means.

In the 1st phase (shaded in orange in Fig. 3) prior to keystroke in the video, i.e., during the trajectory of the hand towards the target chord in the video, incongruent chords evoked a more positive potential than congruent target chords, in both the 5- and the 2-chord sequences. Yet, the positivity had a shorter duration in the 5- compared to 2-chord sequences: while it gave way to an anterior negativity around –80 msec in the 5-chord sequences, this happened only at 0 msec in the 2-chord sequences. This difference in timing may reflect a speeding-up of the 2nd phase anterior negativity (see below) due to higher predictability of the 5-chord sequences. To account for this difference, time windows for statistical testing were set to –300 to –80 msec in the 5-chord sequences and –300 to 0 msec in the 2-chord sequences. (–300 msec were chosen as onset because the trajectory towards the incongruent chords started on average at –300 msec; see Methods.) An ANOVA with the repeated measures factors Congruency (congruent/incongruent) \times Context (5-chord/2-chord sequences) \times AntPost (anterior/central/posterior) \times Laterality (left/middle/right) revealed a significant main effect of Congruency and an interaction of Congruency \times AntPost \times Laterality indicating a broadly distributed positivity irrespective of sequence length (for statistical values, see Table 1). A significant interaction of Congruency \times Context \times AntPost alluded to the more anterior

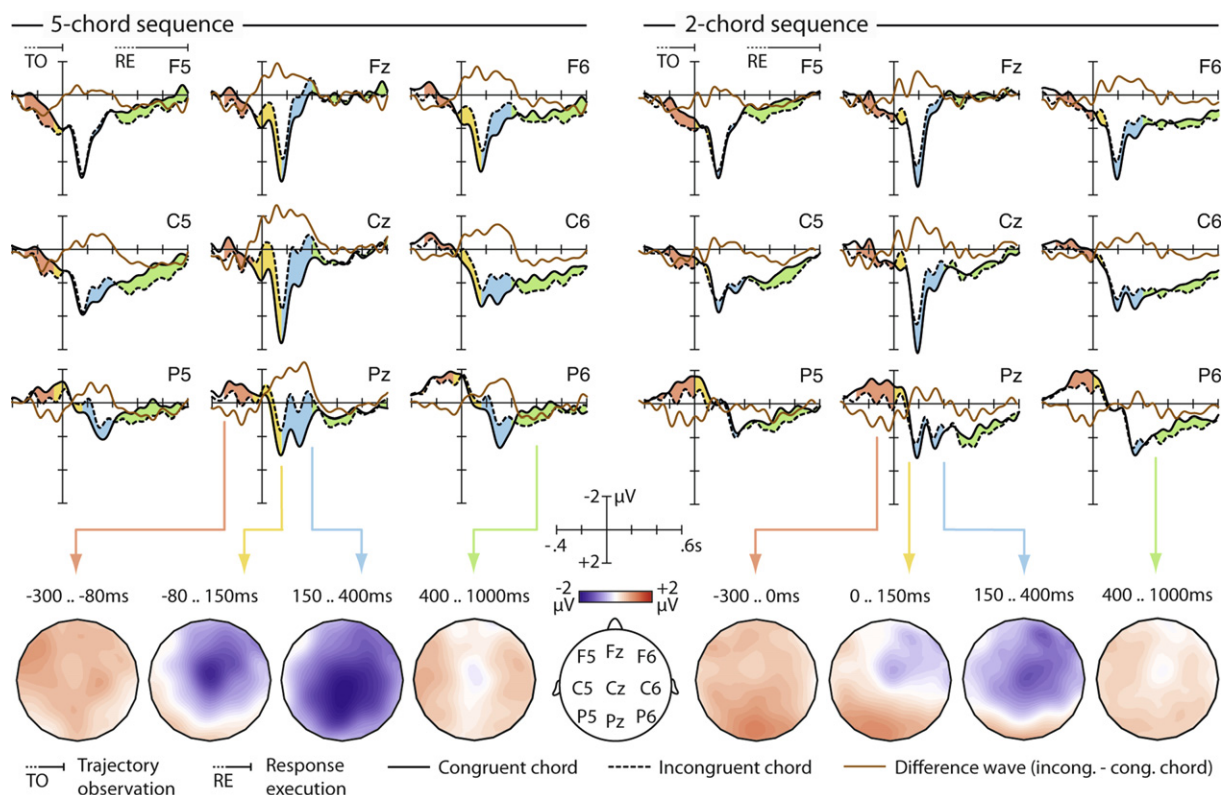


Fig. 3 – ERPs evoked by target chords in 5-chord (left panel) and 2-chord sequences (right panel). Zero demarcates the time when the keys of the target chord were pressed in the video. The legend above electrode F5 indicates the time of the hand trajectory towards the target chord in the video (TO), and the approximate time of response execution by the participant (RE). Time windows of the four neurophysiological phases are shaded in orange (1st), yellow (2nd), blue (3rd) and green (4th) phase. Topography maps in the lower row depict the difference of the potentials of incongruent minus congruent chords within the statistical time window, separately for each phase.

scalp distribution of the positivity in the 5- compared to the 2-chord sequences. Follow-up ANOVAs with the factors Congruency × AntPost × Laterality (computed separately for the long and short contexts) confirmed a left-anterior distribution in the 5-chord sequences [Congruency × AntPost × Laterality: $F(4,104) = 3.34, p < .025, \eta_p^2 = .114$; see topography plots in Fig. 3],

which is in contrast to a trend towards a posterior distribution in the 2-chord sequences [Congruency × AntPost: $F(2,52) = 3.17, p > .073, \eta_p^2 = .109$].

In the 2nd phase (shaded in yellow in Fig. 3), incongruent compared to congruent chords evoked an anterior negativity between -80 and 150 msec in the 5-chord sequences and

Table 1 – Results of the ANOVAs with the factors Congruency × Context × AntPost × Laterality for each time window.

Effect	df	1st phase: -300...-80 versus -300...0 msec ^a			2nd phase: -80...150 versus 0...150 msec ^a			3rd phase: 150...400 msec			4th phase: 400...1000 msec		
		F	p-value	η_p^2	F	p-value	η_p^2	F	p-value	η_p^2	F	p-value	η_p^2
C	1,26	10.77	<.030	.293	<1	>.532	.015	21.41	<.001	.452	3.75	>.063	.126
C × Co	1,26	<1	>.779	.003	6.68	<.016	.204	7.08	<.014	.214	<1	>.863	.001
C × A	2,52	<1	>.760	.006	14.01	<.001	.350	1.96	>.171	.070	<1	>.765	.005
C × A × Co	2,52	4.15	<.049	.138	<1	>.802	.003	5.00	<.032	.161	<1	>.462	.024
C × L	2,52	3.06	>.059	.105	11.40	<.001	.305	11.94	<.001	.315	2.79	>.070	.097
C × L × Co	2,52	<1	>.820	.008	2.80	>.073	.097	1.40	>.256	.051	1.45	>.244	.053
C × A × L	4,104	3.50	<.035	.119	4.01	<.013	.134	9.84	<.001	.275	<1	>.414	.036
C × A × L × Co	4,104	<1	>.485	.031	1.04	>.384	.038	<1	>.454	.033	1.08	>.361	.040

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). C = Congruency, Co = Context, A = AntPost, L = Laterality.

a Note that similar results were found when identical time windows were used for both 5- and 2-chord sequences, i.e., 1st phase -300...-80 msec and 2nd phase 0...150 msec.

between 0 and 150 msec in the 2-chord sequences. (150 msec was chosen as offset because of a remarkable posterior topography shift of the negativity in 5-chord sequences and a return to zero in 2-chord sequences at that time.) The four-way ANOVA showed significant interactions of Congruency \times AntPost, Congruency \times Laterality, and Congruency \times AntPost \times Laterality, demonstrating the middle-to-right frontal maximum of the negativity (Table 1). Follow-up ANOVAs with the factor Congruency computed for each ROI separately confirmed a predominantly middle-to-right anteriorly distributed negativity [middle anterior: $F(1,26) = 10.58$, $p < .004$, $\eta_p^2 = .289$; middle central: $F(1,26) = 4.86$, $p < .037$, $\eta_p^2 = .157$; right anterior: $F(1,26) = 8.36$, $p < .008$, $\eta_p^2 = .243$] accompanied by a left posterior positivity [$F(1,26) = 9.02$, $p < .006$, $\eta_p^2 = .258$; all other p 's $> .158$] that most likely reflects the tail of the 1st phase positivity (see Fig. 3). The negativity was significantly greater in amplitude in the 5- compared to the 2-chord sequences as demonstrated by a significant interaction of Congruency \times Context across all electrodes (Table 1) as well as within single ROIs [middle anterior: $F(1,26) = 4.92$, $p < .036$, $\eta_p^2 = .159$; middle central: $F(1,26) = 7.57$, $p < .011$, $\eta_p^2 = .226$; middle posterior: $F(1,26) = 5.95$, $p < .022$, $\eta_p^2 = .186$; all other p 's $> .061$].

To evaluate in how far the negativity in the 5-chord sequences may have been influenced by (conscious) auditory or motor imagery strategies or the application of music-theoretical knowledge, the mean amplitude of the difference wave (incongruent – congruent) in middle anterior, middle central, and right anterior ROIs was correlated with the ratings obtained in the debriefing. No significant relationships were found for auditory imagery ($r = .189$, $p > .344$, $R^2 = .036$) and motor imagery ($r = .157$, $p > .435$, $R^2 = .025$), whereas the negativity was reduced in amplitude with greater explicit reliance on music-theoretical knowledge ($r = .440$, $p < .023$, $R^2 = .193$), suggesting that the effect was not driven by the conscious identification of the music-syntactic incongruity.

In the 3rd phase (shaded in blue in Fig. 3), incongruent compared to congruent chords elicited a broadly distributed but posteriorly pronounced negativity in 5-chord sequences and an anteriorly distributed negativity in 2-chord sequences, both in the time range from 150 to 400 msec after keystrokes in the video. (400 msec was chosen as offset because the negativities in both contexts gave way to a positivity at that time; see below.) The four-way ANOVA revealed a significant main effect of Congruency, and significant interactions of Congruency \times Laterality and Congruency \times AntPost \times Laterality. Furthermore, interactions of Congruency \times Context and Congruency \times AntPost \times Context were observed (Table 1), indicating a stronger and more posteriorly distributed negativity in 5- compared to 2-chord sequences. Follow-up analyses with the factors Congruency \times Context in each ROI evidenced a significantly stronger negativity in 5- compared to 2-chord sequences at posterior electrodes [interaction of Congruency \times Context; left posterior: $F(1,26) = 11.49$, $p < .003$, $\eta_p^2 = .307$; middle posterior: $F(1,26) = 10.81$, $p < .003$, $\eta_p^2 = .294$; middle central: $F(1,26) = 5.67$, $p < .025$, $\eta_p^2 = .179$; right posterior: $F(1,26) = 9.26$, $p < .006$, $\eta_p^2 = .263$], whereas effects did not differ at anterior and central electrodes (no interaction of Congruency \times Context in the remaining ROIs; all p 's $> .115$).

In the 4th phase (shaded in green in Fig. 3) between 400 and 1000 msec, incongruent chords evoked stronger positive

potentials than congruent chords similarly in both 5- and 2-chord sequences. The four-way ANOVA showed a marginally significant main effect of Congruency and interaction of Congruency \times Laterality (Table 1), suggestive of a stronger positivity in the left and right lateral compared to middle electrodes. No interactions were found between Congruency \times Context, demonstrating that the effects were similar in amplitude and topography in both 5- and 2-chord sequences.

4. Discussion

The present study explored the degree to which musical actions are governed by syntactic processes. Specifically, we aimed to examine the influence of syntax on different aspects of action such as the observation of another person's actions, as well as the programming and execution of one's own actions. To this end, expert piano players simultaneously watched and imitated videos of chord sequences in which the harmonic congruity of the last chord with the preceding syntactic context (congruent or incongruent) and the length of the context (5- or 2-chord sequences) were crossed in a 2×2 factorial design. The experiment was run in the total absence of sound. We defined behavioural and ERP effects that were significantly stronger in 5- compared to 2-chord sequences as related to syntactic processing, because a longer harmonic context establishes more specific syntactic expectancies in the listener (Koelsch et al., 2000; Leino et al., 2007; Tillmann et al., 2003). In other words, the syntactic regulation of motor acts should be reflected in an interaction of Congruency \times Context.

As will be discussed in detail below, the EEG data together with the replication of the behavioural findings reported by Novembre and Keller (2011) show that the observation and imitation of syntactically organized sequences of musical acts evokes motor expectancies that influence skilled pianists' imitation of musical actions. The EEG data extend this finding further by demonstrating that this link between musical syntax and action concerns intermediate processing stages of (i) syntactic analysis of the observed movements and (ii) motor programming for accurate imitation, whereas ERPs related to initial perceptual and late executive stages of the task were not (or only minimally) influenced by the syntactic predictability of the chord sequences. Particularly, electrophysiological indices of action inhibition and reprogramming imply that the observer's motor system anticipates forthcoming actions during imitation based upon his or her long-term music-syntactic knowledge, i.e., suggesting an "embodied" processing of musical harmony. The spatial neighbourhood and dense interconnection (Nieuwenhuys et al., 2008; pp. 841–887) of goal-related action programming in premotor cortex and the syntax-related properties of inferior frontal areas (including Broca's area) might provide a neuroanatomical basis for this interaction. Although it remains to be clarified whether the motor system is informed by an extra syntactic parser (Tettamanti and Moro, 2012) or acts as an independent syntax processor (Pulvermüller and Fadiga, 2010), the combined behavioural and neurophysiological data support the workings of syntax to reach beyond the auditory perception of music to include the action domain.

4.1. Behavioural data

The analysis of the RTs revealed that motor demands differed between congruent and incongruent target chords (main effect of Congruency) but were comparable across 5- and 2-chord sequences as demonstrated by the overall similar RTs in both contexts (no main effect of Context). Most importantly, the imitation of congruent chords was generally faster than imitation of incongruent chords, particularly when the target chord was embedded into a 5-chord sequence, thus replicating the results of [Novembre and Keller \(2011\)](#). This pattern (i.e., an interaction of Congruency \times Context) is entirely in line with our above described criterion for syntax effects in action. Notably, the data suggest that the long syntactic context led to the priming and facilitation of the congruent target chord (i.e., speedup and higher accuracy; possibly reflecting a subliminal modulation of the motor system), instead of processing costs for the incongruent chord (i.e., slowing and lower accuracy) (see also [Tillmann et al., 2003](#)). Overall, this pattern indicates that the harmonic rules implied in the observed action sequences induced strong expectancies in the pianists about forthcoming motor acts and influenced their imitation performance. The EEG data described next, particularly the 2nd and 3rd phases, lead us to argue that this behavioural effect is based on a syntax-driven anticipation of motor programs during action imitation.

4.2. EEG data 1st phase – perceptual processes

In the first phase, i.e., during the presentation of the hand moving towards the target chord in the videos, incongruous chords evoked a more positive potential than congruous chords in both 5- and 2-chord sequences, although with slightly different scalp topography (see below). The early onset of the effect around 300 msec before the hand in the video reached the keys suggests that this ERP component reflects sensory processes related to the perceptually different hand trajectory towards congruent and incongruent targets, i.e., different finger positions, hand shapes and movement onsets (see [Methods](#)). The more pronounced posterior distribution of the effect in 2- compared to 5-chord sequences may reflect the stronger involvement of visual cortical areas due to particular attention of the pianists to these visuo-spatial cues during early stages of musical context build-up (i.e., after the presentation of just one chord when the sense of tonality is still weak) in order to quickly and accurately imitate the observed musical acts. Interestingly, the effect was left-frontally distributed in 5-chord sequences, which raises the possibility of a left inferior frontal source. The left inferior frontal gyrus (IFG) and adjacent ventral premotor cortex (vPMC) have been frequently discussed as a domain-general “grammar processor” ([Fadiga et al., 2009](#); [Fiebach and Schubotz, 2006](#); [Gelfand and Bookheimer, 2003](#); [Koelsch, 2005](#); [Patel, 2003](#); [Tettamanti and Weniger, 2006](#)) involved in the structural sequencing of language ([Friederici, 2011](#); [Grodzinsky and Santi, 2008](#)), music ([Maess et al., 2001](#); [Sammler et al., 2011](#)), and action ([Clerget et al., 2009](#); [Fazio et al., 2009](#)). In this function, and once a clear tonality is established like in the 5-chord sequences, the IFG/vPMC might provide top-down predictions about upcoming chords that include form-based

estimates of the hand trajectory (such as hand shape and finger configurations), i.e., syntactically relevant visuo-motor cues in the movement sequences that are checked against perceptually and motorically salient elements in the video (for similar form-based syntactic estimations in auditory and visual language comprehension, see [Dikker et al., 2009](#); [Herrmann et al., 2009](#)). However, at this stage of research the possibility of top-down syntactic influence on the early perceptual processing of musical actions must remain an interesting hypothesis to test in future studies.

4.3. EEG data 2nd phase – mismatch detection and response conflict

In the second phase, incongruous target chords evoked a right anterior negativity that was significantly stronger and emerged slightly earlier in 5- compared to 2-chord sequences. This interaction of Congruency \times Context is consistent with our criterion indicating syntactic analysis of music performance. Although it remains to be clarified whether this brain response is specifically tied (i) to the detection of the syntactic violation, (ii) to the perception of the incongruous action as a performance error, or (iii) to cognitive control processes related to the participant’s own response, as will be explained in detail below, we will argue that all three views demonstrate the impact of musical grammar on musical actions.

- (i) **Detection of the syntactic violation.** As pointed out earlier (see [Introduction](#)), the auditory presentation of harmonic expectancy violations (such as the ones employed in the current study) evoke an ERAN, an index of (early) music-syntactic processing mediated by the IFG and superior temporal gyrus ([Garza Villarreal et al., 2011](#); [Koelsch, 2009](#); [Sammler et al., 2009](#)). The observed 2nd phase negativity is reminiscent of the ERAN in terms of sensitivity to music-syntactic violations and context length ([Koelsch et al., 2000](#); [Leino et al., 2007](#)), right-anterior scalp topography and polarity inversion at mastoid leads, although the 2nd phase negativity peaked earlier than the ERAN. This acceleration of the effect is most likely due to the ability of skilled pianists to anticipate the congruous or incongruous action outcome in the videos based on the hand trajectory towards the target chord. Pianists may actually use subtle cues in finger configuration – i.e., similar to “coarticulatory” information in speech – to recognize the (in)congruity of the forthcoming chord prior to the actual keystrokes, accounting for the pre-zero onset of the 2nd phase negativity, i.e., an earlier peak than the ERAN in the auditory modality (for a similar action anticipation ability in high-performing athletes, see [Aglioti et al., 2008](#)). Altogether, the above mentioned parallels (despite different timing) may cast the 2nd phase negativity as an equivalent of the ERAN in the visuo-motor modality, and thus provide indirect evidence for modality-independent processing of syntactic irregularities in rule-based harmonic sequences. The idea of such an abstract processing mechanism is supported by experiments showing that reading of unexpected notes in musical scores ([Gunter et al., 2003](#); [Schön and Besson, 2002](#)) evokes early negativities similar to those elicited

when hearing such violations (James et al., 2008; Koelsch, 2005; Patel et al., 1998). Note that the absence of a significant correlation between the negativity's amplitude and the auditory imagery score obtained in the debriefing suggests that the effect is not driven by participants' strategic use of auditory images (Hasegawa et al., 2004; Haslinger et al., 2005; Hubbard, 2010) related to the visually presented stimuli. It more likely reflects the work of a polymodal musical syntax processor that operates on different expressions (i.e., auditory, visual or sensorimotor) of the same syntactic structure. Nevertheless, the possible co-occurrence of auditory images in the context of our motor task is an issue that deserves consideration and is more extensively discussed below.

- (ii) **Perception of a performance error.** In addition to modality-unspecific syntactic processes, the 2nd phase negativity might also reflect an error-related negativity (ERN), or error negativity (N_e), evoked if the incongruous actions in the videos were perceived as erroneous actions (although they were not erroneous *per se*, just unexpected). The ERN is evoked after self-generated errors (Falkenstein et al., 1990; Gehring et al., 1993; Herrojo Ruiz et al., 2009; Mайдhof et al., 2010) as well as errors observed in another person (Miltner et al., 2004; van Schie et al., 2004) suggesting that the observer's own action control system internally simulates the required and perceived action (Iacoboni, 2005; Rizzolatti and Sinigaglia, 2010). The ERN is largest at fronto-central recording sites and is interpreted as the mismatch detection between the actual (i.e., incorrect) action compared to the required (i.e., correct) action (Falkenstein et al., 1990). Notably, the ERN amplitude depends on how well the representation of the required action is established (Falkenstein, 2004), and how strongly the dissimilarity between appropriate and actual response is perceived (Arbel and Donchin, 2011; Bernstein et al., 1995). This property of the ERN could account for its higher amplitude in our 5-chord sequences, which led to a stronger representation of the (required) congruous chord and a greater salience of the incongruous chord, than in 2-chord sequences (Bigand and Parncutt, 1999; Koelsch et al., 2000). Note that such an interpretation puts music-syntactic processes at the origin of a brain response evoked by the observation of an unexpected act. In other words, this finding would demonstrate that syntactic knowledge influences the way in which we perceive another person's action, possibly via simulation of this action in our own motor (syntactic) system (van Schie et al., 2004; Wilson and Knoblich, 2005).
- (iii) **Action control processes.** Beyond these relationships of the 2nd phase negativity to the observation of the incongruous chords in the videos (i.e., ERAN and observer ERN), this brain potential might also be related to the participants' own response. The imitation of the incongruent action sequences may have triggered cognitive control processes such as detection of response conflict and response inhibition to override the prepotent, syntax-driven impulse to produce a congruent sequence ending. In fact, a fronto-centrally distributed N2c or no-go N2 is usually elicited in response priming tasks whenever advance information is invalid (Kopp and Wessel, 2010;

Leuthold, 2004) and a planned response needs to be withheld (Bruin and Wijers, 2002; Falkenstein et al., 1999; Pfefferbaum et al., 1985). It has been suggested that these negativities reflect a control signal that is issued whenever response conflict is detected and is used to temporarily suppress the input to the motor execution system (Stürmer et al., 2002) to adjust or remedy ongoing but inappropriate actions (Kopp et al., 1996). Notably, the N2c amplitude (along with RT costs) has been shown to increase with stronger degree of processing conflict (Botvinick et al., 2001), a condition that is fulfilled especially in the 5-chord sequences. Most importantly, this interpretation does not only imply that the observed syntactically structured sequence of acts triggers an internal representation of the analogous motoric sequence (Rizzolatti and Sinigaglia, 2010). It also alludes to the future-oriented processing of action sequences proposed by incremental models of response preparation (Palmer and Pfordresher, 2003) and, most intriguingly, suggests the automatic advance programming of forthcoming actions (i.e., the congruent target chord) once they can be predicted from the syntactic context (Borroni et al., 2005; Kilner et al., 2004). In other words, the present data argue for an anticipated resonant response in the observer's motor system that does not immediately depend on the realization of the movement in the videos but on context-dependent predictions based on the long-term syntactic knowledge of the pianists.

Taken together, the 2nd phase negativity may be interpreted as an ERAN, an observer ERN, or an N2c/no-go N2 (or a superposition of them; for an overview, see Folstein and Van Petten, 2008), clearly calling for further studies (e.g., with passive observation instead of imitation). Note, importantly, that all three views, irrespective of functional interpretation, demonstrate the operation of musical grammar in the domain of action (observation or programming). On a more abstract level, this triad of processes potentially represents interrelated, syntax-based mechanisms that may play a role during joint musical performance, such as the syntactically guided and modality-unspecific moment-to-moment evaluation and anticipation of other players' musical actions, as well as the syntax-driven programming and flexible revocation of one's own motor acts in concert with other musicians' performances.

4.4. EEG data 3rd phase – response (re-)programming

In the third phase, incongruous (compared to congruous) chords in 5-chord sequences evoked a slightly right-lateralized posterior negativity that was not observed in 2-chord sequences and is therefore – in line with our definition of a Congruency \times Context interaction – most likely related to the syntactic regulation of the musical performance. More precisely, this effect may reflect mechanisms of movement reprogramming following the cancellation of the syntactically prepotent response, i.e., the programming of the incongruent chord in the face of the more dominant congruent chord (Mars et al., 2007). Response priming paradigms comparing the execution of an action after neutral,

valid, and invalid primes have consistently observed large RT costs along with a slightly right-dominant centroparietal negativity in invalid compared to valid and neutral precue conditions (Leuthold and Jentsch, 2002; Vidal et al., 1995). Considering our 2-chord, congruent 5-chord, and incongruent 5-chord sequences as equivalents of the neutral, valid, and invalid conditions, respectively, this finding closely mirrors our results and thus supports the interpretation as response reprogramming. Importantly, the reprogramming of an action provides indirect support for the prior activation of an alternative motor program, i.e., the syntax-driven motoric anticipation of the congruent chord during action observation.

It should be noted that this interpretation does not necessarily exclude the simultaneous auditory anticipation of the congruent chord. In fact, a number of functional magnetic resonance imaging (fMRI) studies showed auditory activations during the observation (Hasegawa et al., 2004; Haslinger et al., 2005) or execution of mute piano performance (Bangert et al., 2006; Lahav et al., 2007; Zatorre et al., 2007), making simultaneous processing in auditory and motor areas likely in the context of the present task. However, we argue that the posterior scalp distribution of the effect is atypical for purely auditory-based potentials and rather compatible with a stronger involvement of the centroparietal motor (as opposed to auditory) system.

No significant centroparietal potential differences were found between congruent and incongruent endings in the 2-chord sequences, suggesting that their programming was less predictive, but rather reactive (i.e., without prior activation of the congruent motor program and thus no need of reprogramming in case of an incongruent chord). Instead, incongruent (compared to congruent) chords evoked a fronto-central negativity with slight right-hemispheric predominance. Note that a similar effect may have been also evoked by 5-chord sequences, and superposed by the prominent posterior negativity. As mentioned earlier (see *Methods*), 2-chord sequences – although syntactically less predictable than 5-chord sequences – were not syntax-free, as they contained very common chord pairs in Western tonal music. Therefore, this effect may reflect some delayed (in 2-chord sequences) or extended (in 5-chord sequences) music-syntactic processing of the visually perceived chord, akin to the negativities observed in the auditory (Koelsch, 2009; Koelsch et al., 2000) and visual modality (Gunter et al., 2003; Schön and Besson, 2002). Alternatively, given that incongruent target chords were equally higher in motoric (i.e., kinematic) complexity than congruent chords in both 5- and 2-chord sequences (see *Methods*), this ERP may be interpreted as a correlate of the greater effort during movement programming, possibly as one aspect of the contingent negative variation (CNV) of externally cued movements (Walter et al., 1964). Consistent with our results, the CNV is fronto-centrally distributed, is assumed to reflect (amongst others) preparatory motor activity, and exhibits larger amplitudes for motorically more complex and less familiar movements (Cui et al., 2000; Kranczioch et al., 2010) (for an overview, see Brunia, 2003).

4.5. EEG data 4th phase – response execution

In the 4th phase, incongruent chords evoked slightly stronger positive potentials than congruent chords at left and right

lateral electrodes in both 5- and 2-chord sequences. This effect may reflect syntactic reanalysis and repair processes (because both 5- and 2-chord sequences contained syntactic information; see above) that are commonly associated with a P300 (Janata, 1995), P600 (Patel et al., 1998), or late positive component (LPC; Besson and Faïta, 1995), although those are usually centro-posteriorly distributed. Alternatively, the latency of the potential between 400 and 1000 msec corresponds to the latency of the participants' motor responses (Fig. 2), suggesting a link of the ERP with the execution of the final chords. Congruent chords required less cognitive and motor programming effort to be executed (e.g., due to a lower number of black keys; see *Methods*) and had significantly shorter RTs than incongruent chords (Fig. 2), possibly leading to an earlier onset of movement-related cortical potentials. These are negative shifts preceding movement onset, such as the slowly rising readiness potential (Bereitschaftspotential) in single self-paced movements (Kornhuber and Deecke, 1965; Shibasaki and Hallett, 2006) or a steady-state movement-related cortical negativity in repetitive movements (Gerloff et al., 1997). Notably, an earlier onset of the negative shift in the congruent conditions would show up as a relative positivity, as the one observed in the present experiment. Furthermore, complex movements (such as in piano playing) are known to recruit not only contra- but also ipsilateral motor areas (Haaland et al., 2004; Nishitani and Hari, 2000; Verstynen et al., 2005), consistent with the bilateral topography of the observed effect.

4.6. Is music-syntactic processing “embodied”?

The present data suggest an influence of music-syntactic knowledge on piano performance, as evidenced by longer RTs, higher error rates, and ERPs reflecting the cancellation (2nd phase) and reprogramming (3rd phase) of a prepotent motor response during imitation of syntactically incongruent chord sequences. In other words, the observed and imitated musical acts evoked an anticipatory response in the pianist's motor system that may be taken as an “embodied” correlate of music-syntactic processing.

Theories of embodiment ground cognitive processes, such as action perception (Wilson and Knoblich, 2005), language processing (Glenberg and Gallese, 2012; Glenberg and Kaschak, 2002; Pulvermüller and Fadiga, 2010), or social cognition (Gallese, 2007; Gallese and Sinigaglia, 2011; Goldman and de Vignemont, 2009), in a bodily format (for reviews, see Barsalou, 2008; Wilson, 2002). Recent attempts in the language domain to unravel bodily representations of syntactic processes in the sensorimotor (“mirror”) system (Pulvermüller and Fadiga, 2010) directly relate to our results, due to the structural and neurofunctional similarities of language and music (see *Introduction*). It should be noted though that our study – although it strongly suggests a role of the sensorimotor system in establishing a bodily image of forthcoming motor acts based on syntactic information – did not explicitly manipulate recursion and thus cannot clarify whether this predictive process involves merely linear chord transitions and overlearned syntactic patterns, or full-fledged recursive dependencies and the hierarchical depth structure of harmony (Katz and Pesetsky, 2011; Rohrmeier, 2011). Overall,

the cognitive instantiation of recursion in both music and action remains an issue that deserves further attention to build a clear theoretical framework on aspects of syntax that are specific to language or shared between domains.

It is conceivable that recursive hierarchy in action cannot be processed by the sensorimotor system alone (i.e., via a “mirror” mechanism) but requires reciprocal interactions with an external syntactic processor. This possibility would be in line with associative (as opposed to non-associative, “mirror”) kinds of embodied cognition theories in a way that a polymodal syntax processor informs and is informed by the sensorimotor system in two-way feedback loops, (back)translating syntactic processes from/into a bodily format (Tettamanti and Moro, 2012). The spatial neighbourhood and dense interconnection (Nieuwenhuys et al., 2008; pp. 841–887) of premotor cortex and putative syntactic properties of inferior frontal regions (including Broca’s area) make such an interaction anatomically plausible. Furthermore, deficits in processing or learning compositional action sequences after damage to left inferior frontal regions (Fazio et al., 2009) or transcranial magnetic stimulation over Broca’s area (Clerget et al., 2011, 2009) is compatible with this notion. The question which level of syntactic complexity the bodily senses themselves are able to parse remains an interesting topic for future research.

5. Conclusion

In sum, the present study shows that music-syntactic knowledge triggers the internal build-up of action expectancies in trained musicians when they imitate other players’ musical actions. In particular, the data suggest that the prediction of musical acts involves a resonant response in the observer’s motor system, i.e., an “embodied” application of the rules determining the position of chords within harmonic sequences (that in turn influences music performance). It is well-known that the motor system is endowed with the ability to predict abstract events (Schubotz, 2007) and to constantly set up anticipatory models of other persons’ actions (Borroni et al., 2005; Iacoboni et al., 2005; Kilner et al., 2004; van Schie et al., 2004) in order to predict other’s intentions ahead of their behavioural realization (for related monkey work, see Umiltà et al., 2001). The novel finding of the present study is that – in a musical context and despite the absence of auditory information – the motor system’s anticipatory potential is strongly based upon the observer’s long-term music-syntactic knowledge. The present study thus shows that the notion of syntax not only applies to music in the auditory modality but transfers – in trained musicians – to an embodied processing of a “grammar of musical action”.

Acknowledgements

First and foremost we thank all pianists who participated in the study. We further thank Katrin Cunitz, Maria Bader, Irene Nenoff, and Katharina Fath for their great help with data acquisition, Kerstin Träger for inviting the pianists, Jan Bergmann and Sven Gutekunst for technical support, Makiko

Sadakata for creating the video stimuli, and three anonymous reviewers for helpful comments.

REFERENCES

- Abrams DA, Bhatara A, Ryali S, Balaban E, Levitin DJ, and Menon V. Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cerebral Cortex*, 21(7): 1507–1518, 2011.
- Aglioti SM, Cesari P, Romani M, and Urgesi C. Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, 11(9): 1109–1116, 2008.
- Allen K, Ibara S, Seymour A, Cordova N, and Botvinick M. Abstract structural representations of goal-directed behavior. *Psychological Science*, 21(10): 1518–1524, 2010.
- Arbel Y and Donchin E. How large the sin? A study of the event related potentials elicited by errors of varying magnitude. *Psychophysiology*, 48(12): 1611–1620, 2011.
- Bahlmann J, Schubotz RI, Mueller JL, Koester D, and Friederici AD. Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Research*, 1298: 161–170, 2009.
- Bangert M, Peschel T, Schlaug G, Rotte M, Drescher D, Hinrichs H, et al. Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. *NeuroImage*, 30(3): 917–926, 2006.
- Barsalou LW. Grounded cognition. *Annual Review of Psychology*, 59: 617–645, 2008.
- Bernstein PS, Scheffers MK, and Coles MG. “Where did I go wrong?” A psychophysiological analysis of error detection. *Journal of Experimental Psychology: Human Perception and Performance*, 21(6): 1312–1322, 1995.
- Besson M and Faïta F. An event-related (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, 21(6): 1278–1296, 1995.
- Bigand E and Parncutt R. Perceiving musical tension in long chord sequences. *Psychological Research*, 62(4): 237–254, 1999.
- Borroni P, Montagna M, Cerri G, and Baldissera F. Cyclic time course of motor excitability modulation during the observation of a cyclic hand movement. *Brain Research*, 1065(1–2): 115–124, 2005.
- Bortz JJ and Döring N. *Forschungsmethoden und Evaluation: Für Human- und Sozialwissenschaftler*. Berlin: Springer, 2003.
- Botvinick MM, Braver TS, Barch DM, Carter CS, and Cohen JD. Conflict monitoring and cognitive control. *Psychological Review*, 108(3): 624–652, 2001.
- Bruin KJ and Wijers AA. Inhibition, response mode, and stimulus probability: A comparative event-related potential study. *Clinical Neurophysiology*, 113(7): 1172–1182, 2002.
- Brunia CHM. CNV and SNP: Indices of anticipatory behavior. In Jahanshahi M and Hallett M (Eds), *The Bereitschaftspotential – Movement-related Cortical Potentials*. New York: Kluwer Academic/Plenum Publishers, 2003: 207–227.
- Buccino G, Lui F, Canessa N, Patteri I, Lagravinese G, Benuzzi F, et al. Neural circuits involved in the recognition of actions performed by nonconpecifics: An fMRI study. *Journal of Cognitive Neuroscience*, 16(1): 114–126, 2004.
- Chomsky AN. *Syntactic Structures*. Den Haag: Mouton, 1957.
- Chomsky AN. *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger, 1986.
- Chomsky AN. *The Minimalist Program*. Cambridge, MA: MIT Press, 1995.
- Clerget E, Poncin W, Fadiga L, and Olivier E. Role of Broca’s area in implicit motor skill learning: Evidence from continuous theta-burst magnetic stimulation. *Journal of Cognitive Neuroscience*, 24(1): 80–92, 2011.

- Clerget E, Winderickx A, Fadiga L, and Olivier E. Role of Broca's area in encoding sequential human actions: A virtual lesion study. *NeuroReport*, 20(16): 1496–1499, 2009.
- Cui RQ, Egkher A, Huter D, Lang W, Lindinger G, and Deecke L. High resolution spatiotemporal analysis of the contingent negative variation in simple or complex motor tasks and a non-motor task. *Clinical Neurophysiology*, 111(10): 1847–1859, 2000.
- Dehaene S. *The Number Sense – How the Mind Creates Mathematics*. New York: Oxford University Press, 1997.
- Delorme A and Makeig S. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1): 9–21, 2004.
- Dikker S, Rabagliati H, and Pyllkkänen L. Sensitivity to syntax in visual cortex. *Cognition*, 110(3): 293–321, 2009.
- Drost UC, Rieger M, Brass M, Gunter TC, and Prinz W. When hearing turns into playing: Movement induction by auditory stimuli in pianists. *The Quarterly Journal of Experimental Psychology Section A*, 58(8): 1376–1389, 2005.
- Fadiga L, Craighero L, and D'Ausilio A. Broca's area in language, action, and music. *Annals of the New York Academy of Sciences*, 1169: 448–458, 2009.
- Falkenstein M. ERP correlates of erroneous performance. In Ullsperger M and Falkenstein M (Eds), *Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring*. Leipzig: Max Planck Institute for Human Cognitive and Brain Sciences, 2004: 5–14.
- Falkenstein M, Hohnsbein J, Hoormann J, and Blanke L. Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In Brunia C, Gaillard A, and Kok A (Eds), *Psychophysiological Brain Research*. Tilburg: Tilburg University Press, 1990.
- Falkenstein M, Hoormann J, and Hohnsbein J. ERP components in Go/Nogo tasks and their relation to inhibition. *Acta Psychologica*, 101(2–3): 267–291, 1999.
- Farag C, Troiani V, Bonner M, Powers C, Avants B, Gee J, et al. Hierarchical organization of scripts: Converging evidence from fMRI and frontotemporal degeneration. *Cerebral Cortex*, 20(10): 2453–2463, 2010.
- Fazio P, Cantagallo A, Craighero L, D'Ausilio A, Roy AC, Pozzo T, et al. Encoding of human action in Broca's area. *Brain*, 132: 1980–1988, 2009.
- Fedorenko E, Patel A, Casasanto D, Winawer J, and Gibson E. Structural integration in language and music: Evidence for a shared system. *Memory & Cognition*, 37(1): 1–9, 2009.
- Fiebach CJ and Schubotz RI. Dynamic anticipatory processing of hierarchical sequential events: A common role for Broca's area and ventral premotor cortex across domains? *Cortex*, 42(4): 499–502, 2006.
- Fitch WT. Unity and diversity in human language. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1563): 376–388, 2011.
- Folstein JR and Van Petten C. Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology*, 45(1): 152–170, 2008.
- Friederici AD. The brain basis of language processing: From structure to function. *Physiological Reviews*, 91(4): 1357–1392, 2011.
- Friedrich R and Friederici AD. Mathematical logic in the human brain: Syntax. *Plos One*, 4(5): e5599, 2009.
- Gallese V. Before and below theory of mind: Embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480): 659–669, 2007.
- Gallese V and Sinigaglia C. What is so special about embodied simulation? *Trends in Cognitive Sciences*, 15(11): 512–518, 2011.
- Garza Villarreal EA, Brattico E, Leino S, Ostergaard L, and Vuust P. Distinct neural responses to chord violations: A multiple source analysis study. *Brain Research*, 1389: 103–114, 2011.
- Gehring WJ, Goss B, Coles MGH, Meyer DE, and Donchin E. A neural system for error-detection and compensation. *Psychological Science*, 4(6): 385–390, 1993.
- Gelfand JR and Bookheimer SY. Dissociating neural mechanisms of temporal sequencing and processing phonemes. *Neuron*, 38(5): 831–842, 2003.
- Gerloff C, Toro C, Uenishi N, Cohen LG, Leocani L, and Hallett M. Steady-state movement-related cortical potentials: A new approach to assessing cortical activity associated with fast repetitive finger movements. *Electroencephalography and Clinical Neurophysiology*, 102(2): 106–113, 1997.
- Gibbs RWJ. *Embodiment and Cognitive Science*. Cambridge: Cambridge University Press, 2006.
- Glenberg AM and Gallese V. Action-based language: A theory of language acquisition, comprehension, and production. *Cortex*, 48(7): 905–922, 2012.
- Glenberg AM and Kaschak MP. Grounding language in action. *Psychonomic Bulletin & Review*, 9(3): 558–565, 2002.
- Goldman A and de Vignemont F. Is social cognition embodied? *Trends in Cognitive Sciences*, 13(4): 154–159, 2009.
- Greenfield PM. Language, tools, and brain – The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences*, 14(4): 531–550, 1991.
- Grodzinsky Y. The neurology of syntax: Language use without Broca's area. *Behavioral and Brain Sciences*, 23(1): 1–71, 2000.
- Grodzinsky Y and Santi A. The battle for Broca's region. *Trends in Cognitive Sciences*, 12(12): 474–480, 2008.
- Guerra-Filho G and Aloimonos Y. The syntax of human actions and interactions. *Journal of Neurolinguistics*, 25(5): 500–514, 2012.
- Gunter TC, Schmidt BH, and Besson M. Let's face the music: A behavioral and electrophysiological exploration of score reading. *Psychophysiology*, 40(5): 742–751, 2003.
- Haaland KY, Elsinger CL, Mayer AR, Durgerian S, and Rao SM. Motor sequence complexity and performing hand produce differential patterns of hemispheric lateralization. *Journal of Cognitive Neuroscience*, 16(4): 621–636, 2004.
- Hasegawa T, Matsuki K, Ueno T, Maeda Y, Matsue Y, Konishi Y, et al. Learned audio–visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study. *Cognitive Brain Research*, 20(3): 510–518, 2004.
- Haslinger B, Erhard P, Altenmüller E, Schroeder U, Boecker H, and Ceballos-Baumann AO. Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience*, 17(2): 282–293, 2005.
- Hauser MD, Chomsky N, and Fitch WT. The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598): 1569–1579, 2002.
- Herrmann B, Maess B, Hasting AS, and Friederici AD. Localization of the syntactic mismatch negativity in the temporal cortex: An MEG study. *NeuroImage*, 48(3): 590–600, 2009.
- Herrojo Ruiz M, Jabusch HC, and Altenmüller E. Detecting wrong notes in advance: Neuronal correlates of error monitoring in pianists. *Cerebral Cortex*, 19(11): 2625–2639, 2009.
- Hubbard TL. Auditory imagery: Empirical findings. *Psychological Bulletin*, 136(2): 302–329, 2010.
- Iacoboni M. Neural mechanisms of imitation. *Current Opinion in Neurobiology*, 15(6): 632–637, 2005.
- Iacoboni M, Molnar-Szakacs I, Gallese V, Buccino G, Mazziotta JC, and Rizzolatti G. Grasping the intentions of others with one's own mirror neuron system. *Plos Biology*, 3(3): e79, 2005.
- James CE, Britz J, Vuilleumier P, Hauert CA, and Michel CM. Early neuronal responses in right limbic structures mediate harmony incongruity processing in musical experts. *NeuroImage*, 42(4): 1597–1608, 2008.
- Janata P. ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*, 7(2): 153–164, 1995.

- Jentschke S and Koelsch S. Musical training modulates the development of syntax processing in children. *NeuroImage*, 47(2): 735–744, 2009.
- Jentschke S, Koelsch S, and Friederici AD. Investigating the relationship of music and language in children: Influences of musical training and language impairment. *Annals of the New York Academy of Sciences*, 1060: 231–242, 2005.
- Kaan E and Swaab TY. The brain circuitry of syntactic comprehension. *Trends in Cognitive Sciences*, 6(8): 350–356, 2002.
- Katz J and Pesetsky D. The Identity Thesis of Language and Music. *lingBuzz/000959*, 2011.
- Kilner JM, Vargas C, Duval S, Blakemore SJ, and Sirigu A. Motor activation prior to observation of a predicted movement. *Nature Neuroscience*, 7(12): 1299–1301, 2004.
- Koelsch S. Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, 15(2): 207–212, 2005.
- Koelsch S. Music-syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. *Psychophysiology*, 46(1): 179–190, 2009.
- Koelsch S. Toward a neural basis of music perception – A review and updated model. *Frontiers in Psychology*, 2: 1–20, 2011.
- Koelsch S, Gunter T, Friederici AD, and Schröger E. Brain indices of music processing: “Nonmusicians” are musical. *Journal of Cognitive Neuroscience*, 12(3): 520–541, 2000.
- Koelsch S, Gunter TC, Wittfoth M, and Sammler D. Interaction between syntax processing in language and in music: An ERP study. *Journal of Cognitive Neuroscience*, 17(10): 1565–1577, 2005.
- Kopp B, Rist F, and Mattler U. N200 in the flanker task as a neurobehavioral tool for investigating executive control. *Psychophysiology*, 33(3): 282–294, 1996.
- Kopp B and Wessel K. Event-related brain potentials and cognitive processes related to perceptual-motor information transmission. *Cognitive, Affective, Behavioral Neuroscience*, 10(2): 316–327, 2010.
- Kornhuber HH and Deecke L. Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, 284(1): 1–17, 1965.
- Kranczioch C, Mathews S, Dean P, and Sterr A. Task complexity differentially affects executed and imagined movement preparation: Evidence from movement-related potentials. *Plos One*, 5(2) 2010.
- Lahav A, Saltzman E, and Schlaug G. Auditory operations of the mirror neuron system: Action recognition network while listening to newly-acquired actions. *Neurology*, 68(12): A106, 2007.
- Lashley KS. The problem of serial order in behavior. In Jeffress LA (Ed), *Cerebral Mechanisms in Behavior*. New York: Wiley, 1951: 112–131.
- Leino S, Brattico E, Tervaniemi M, and Vuust P. Representation of harmony rules in the human brain: Further evidence from event-related potentials. *Brain Research*, 1142: 169–177, 2007.
- Lerdahl F and Jackendoff R. *A Generative Theory of Tonal Music*. Cambridge: MIT Press, 1983.
- Leuthold H. Action monitoring and control: In search of common principles. In Ullsperger M and Falkenstein M (Eds), *Errors, Conflicts, and the Brain. Current Opinions on Performance Monitoring*. Leipzig: Max Planck Institute for Human Cognitive and Brain Sciences, 2004: 70–76.
- Leuthold H and Jentsch I. Spatiotemporal source localisation reveals involvement of medial premotor areas in movement reprogramming. *Experimental Brain Research*, 144(2): 178–188, 2002.
- Maess B, Koelsch S, Gunter TC, and Friederici AD. Musical syntax is processed in Broca's area: An MEG study. *Nature Neuroscience*, 4(5): 540–545, 2001.
- Maidhof C, Vavatzanidis N, Prinz W, Rieger M, and Koelsch S. Processing expectancy violations during music performance and perception: An ERP study. *Journal of Cognitive Neuroscience*, 22(10): 2401–2413, 2010.
- Marin MM. Effects of early musical training on musical and linguistic syntactic abilities. *Annals of the New York Academy of Sciences*, 1169: 187–190, 2009.
- Mars RB, Piekema C, Coles MG, Hulstijn W, and Toni I. On the programming and reprogramming of actions. *Cerebral Cortex*, 17(12): 2972–2979, 2007.
- Michel CM, Murray MM, Lantz G, Gonzalez S, Spinelli L, and Grave de Peralta R. EEG source imaging. *Clinical Neurophysiology*, 115(10): 2195–2222, 2004.
- Miltner WHR, Brauer J, Hecht H, Trippe R, and Coles MGH. Parallel brain activity for self-generated and observed errors. In Ullsperger M and Falkenstein M (Eds), *Errors, Conflict, and the Brain. Current Opinions on Performance Monitoring*. Leipzig: MPI of Cognitive Neuroscience, 2004: 124–129.
- Monti MM, Parsons L, and Osherson DN. The boundaries of language and thought in deductive inference. *Proceedings of the National Academy of Sciences of the United States of America*, 106(30): 12554–12559, 2009.
- Moro A. *The Boundaries of Babel – The Brain and the Enigma of Impossible Languages*. Cambridge, MA: MIT Press, 2008.
- Moro A, Tettamanti M, Perani D, Donati C, Cappa SF, and Fazio F. Syntax and the brain: Disentangling grammar by selective anomalies. *NeuroImage*, 13(1): 110–118, 2001.
- Musso M, Moro A, Glauche V, Rijntjes M, Reichenbach J, Büchel C, et al. Broca's area and the language instinct. *Nature Neuroscience*, 6(7): 774–781, 2003.
- Nieuwenhuys R, Voogd J, and van Huijzen C. *The Human Central Nervous System*. Berlin: Springer, 2008.
- Nishitani N and Hari R. Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2): 913–918, 2000.
- Novembre G and Keller PE. A grammar of action generates predictions in skilled musicians. *Consciousness and Cognition*, 20(4): 1232–1243, 2011.
- Núñez-Peña MI and Honrubia-Serrano ML. P600 related to rule violation in an arithmetic task. *Cognitive Brain Research*, 18(2): 130–141, 2004.
- Pallier C, Devauchelle AD, and Dehaene S. Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences of the United States of America*, 108(6): 2522–2527, 2011.
- Palmer C and Pfordresher PQ. Incremental planning in sequence production. *Psychological Review*, 110(4): 683–712, 2003.
- Palmer C and van de Sande C. Units of knowledge in music performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(2): 457–470, 1993.
- Palmer C and van de Sande C. Range of planning in music performance. *Journal of Experimental Psychology: Human Perception and Performance*, 21(5): 947–962, 1995.
- Pastra K and Aloimonos Y. The minimalist grammar of action. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1585): 103–117, 2012.
- Patel AD. Language, music, syntax and the brain. *Nature Neuroscience*, 6(7): 674–681, 2003.
- Patel AD, Gibson E, Ratner J, Besson M, and Holcomb PJ. Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience*, 10(6): 717–733, 1998.
- Patel AD, Iversen JR, Wassenaar M, and Hagoort P. Musical syntax processing in agrammatic Broca's aphasia. *Aphasiology*, 22(7–8): 776–789, 2008.
- Perani D, Fazio F, Borghese NA, Tettamanti M, Ferrari S, Decety J, et al. Different brain correlates for watching real and virtual hand actions. *NeuroImage*, 14(3): 749–758, 2001.

- Pfefferbaum A, Ford JM, Weller BJ, and Kopell BS. ERPs to response production and inhibition. *Electroencephalography and Clinical Neurophysiology*, 60(5): 423–434, 1985.
- Pulvermüller F and Fadiga L. Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews Neuroscience*, 11(5): 351–360, 2010.
- Restle F. Theory of serial pattern learning: Structural trees. *Psychological Review*, 77(6): 481–495, 1970.
- Rizzolatti G and Sinigaglia C. The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience*, 11(4): 264–274, 2010.
- Rogalsky C, Rong F, Saberi K, and Hickok G. Functional anatomy of language and music perception: Temporal and structural factors investigated using functional magnetic resonance imaging. *Journal of Neuroscience*, 31(10): 3843–3852, 2011.
- Rohrmeier M. Towards a generative syntax of tonal harmony. *Journal of Mathematics and Music*, 5(1): 35–53, 2011.
- Sammler D, Koelsch S, Ball T, Brandt A, Elger CE, Friederici AD, et al. Overlap of musical and linguistic syntax processing: Intracranial ERP evidence. *Annals of the New York Academy of Sciences*, 1169: 494–498, 2009.
- Sammler D, Koelsch S, and Friederici AD. Are left fronto-temporal brain areas a prerequisite for normal music-syntactic processing? *Cortex*, 47(6): 659–673, 2011.
- Scheepers C, Sturt P, Martin CJ, Myachykov A, Teevan K, and Viskupova I. Structural priming across cognitive domains: From simple arithmetics to relative-clause attachment. *Psychological Science*, 22(10): 1319–1326, 2011.
- Schön D and Besson M. Processing pitch and duration in music reading: A RT-ERP study. *Neuropsychologia*, 40(7): 868–878, 2002.
- Schubotz RI. Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences*, 11(5): 211–218, 2007.
- Sharbrough F, Chatrian G-E, Lesser RP, Lüders H, Nuwer M, and Picton TW. American-electroencephalographic-society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, 8(2): 200–202, 1991.
- Shibasaki H and Hallett M. What is the Bereitschaftspotential? *Clinical Neurophysiology*, 117(11): 2341–2356, 2006.
- Slevc LR, Rosenberg JC, and Patel AD. Making psycholinguistics musical: Self-paced reading time evidence for shared processing of linguistic and musical syntax. *Psychonomic Bulletin & Review*, 16(2): 374–381, 2009.
- Steinbeis N and Koelsch S. Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cerebral Cortex*, 18(5): 1169–1178, 2008.
- Stevens JA, Fonlupt P, Shiffrrar M, and Decety J. New aspects of motion perception: Selective neural encoding of apparent human movements. *NeuroReport*, 11(1): 109–115, 2000.
- Stürmer B, Leuthold H, Soetens E, Schröter H, and Sommer W. Control over location-based response activation in the Simon task: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, 28(6): 1345–1363, 2002.
- Tettamanti M and Moro A. Can syntax appear in a mirror (system)? *Cortex*, 48(7): 923–935, 2012.
- Tettamanti M, Rotondi I, Perani D, Scotti G, Fazio F, Cappa SF, et al. Syntax without language: Neurobiological evidence for cross-domain syntactic computations. *Cortex*, 45(7): 825–838, 2009.
- Tettamanti M and Weniger D. Broca's area: A supramodal hierarchical processor? *Cortex*, 42(4): 491–494, 2006.
- Tillmann B, Janata P, Birk J, and Bharucha JJ. The costs and benefits of tonal centers for chord processing. *Journal of Experimental Psychology: Human Perception and Performance*, 29(2): 470–482, 2003.
- Umiltà MA, Kohler E, Gallese V, Fogassi L, Fadiga L, Keysers C, et al. I know what you are doing: A neurophysiological study. *Neuron*, 31(1): 155–165, 2001.
- van Schie HT, Mars RB, Coles MGH, and Bekkering H. Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, 7(5): 549–554, 2004.
- van Schie HT, Toni I, and Bekkering H. Comparable mechanisms for action and language: Neural systems behind intentions, goals, and means. *Cortex*, 42(4): 495–498, 2006.
- Verstynen T, Diedrichsen J, Albert N, Aparicio P, and Ivry RB. Ipsilateral motor cortex activity during unimanual hand movements relates to task complexity. *Journal of Neurophysiology*, 93(3): 1209–1222, 2005.
- Vidal F, Bonnet M, and Macar F. Programming the duration of a motor sequence: Role of the primary and supplementary motor areas in man. *Experimental Brain Research*, 106(2): 339–350, 1995.
- Walter WG, Winter AL, Cooper R, Mccallum WC, and Aldridge VJ. Contingent negative variation – Electric sign of sensorimotor association and expectancy in human brain. *Nature*, 203(494): 380–384, 1964.
- Wilson M. Six views of embodied cognition. *Psychonomic Bulletin and Review*, 9(4): 625–636, 2002.
- Wilson M and Knoblich G. The case of motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131(3): 460–473, 2005.
- Zacks JM and Tversky B. Event structure in perception and conception. *Psychological Bulletin*, 127(1): 3–21, 2001.
- Zatorre RJ, Chen JL, and Penhune VB. When the brain plays music: Auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8(7): 547–558, 2007.