Fachbereich Erziehungswissenschaft und Psychologie der

Freien Universität Berlin

The Impact of Cognitive- and Action- Predictive Processes on Neural Correlates of Syntactic Processing in Music

Dissertation

zur Erlangung des akademischen Grades
Doktorin der Philosophie (Dr. phil.)
Doctor of Philosophy (Ph.D.)
im Fach Psychologie

Vorgelegt von
Shuang Guo (M.Ed., B.Sc.)
Berlin, 2017

Erstgutachter:

Prof. Dr. Stefan Koelsch

Zweitgutachter:

Prof. Dr. Dirk Ostwald

Tag der Disputation: 10. Juli 2017

To my beloved parents

Acknowledgments

My first thank goes to, music, which gives me consolation, understanding, belonging, inspiration, and motivation for those bittersweet moments.

I am extremely grateful to my supervisor Prof. Stefan Koelsch. Without his support at various stages, I could not finish my Ph.D. study. Many thanks to him for offering me the opportunity to study in Germany, leading me to a better understanding of my research field, giving me helpful and patient guidance and invaluable advice, and for being a role model for me to do research in depth. Thank you for opening a door for me to know the scientific world better.

My sincere gratitude goes to Dr. Sebastian Jentschke for not only training me in EEG data analysis, but also giving me generous help and advice in many aspects. I am equally grateful to Dr. Clemens Maidhof who is supportive in many ways to my research. I am greatly benefited from many stimulating and encouraging discussions and comments.

I have to say that how fortunate I am to be a member of the Neuromusic group, where I have the chance to share thoughts and discuss with these very helpful and nice colleagues, especially Maria Sumpf, Moritz Lehne, Liila Taruffi, Corinna Pehrs, Stavros Skouras, Martin Rohrmeier, and Niloofar Keshtiari. Thanks also go to staffs of Freie Universität Berlin, in particular Franca Fabiunke, Christian Kainz, Regula von Schintling, Aleksandra Gulka, and Silke Witte.

I would like to thank Kamila Borowiak, Laura Hahn, Bas Visser, Stephanie Appelhans, Anne Märtin, Vienna Doenni, Yu Fukuda, Daniela Ordóñez, and Anastasia Glushko, for giving me assistance in recruiting participants and doing experiments. In addition, my thanks go to Guangzhan Fang and Kurt Steinmetzger for helping with programming and data analysis.

Thanks particular to anonymous reviewers who offered a host of insightful comments and criticisms, and to my participants who have endured many hours of experiment in the lab.

I shared many precious and memorable moments with my friends Kai Zhang, Hong Zhang, Yan Shen, and Yan Fan. Thank you for being with me, comforting me, and continuously motivating me. Special thanks go to Anne Roosen-Runge, Ms. Li, and Xixi for taking so much care of me. I do cherish in my heart.

Last but not the least, I want to express my gratitude to my beloved parents. I know from the bottom of my heart that they are always there, giving me unconditional support no matter how far away I am from them and how long my Ph.D. study takes. Thank them for understanding my struggle, encouraging my thought, cherishing my dream, respecting my choice, and giving me selfless and eternal love. This dissertation is dedicated to them. I cannot mention other family members here one by one, but they are my inner strength no matter what I am doing.

I was guided by different expectations throughout the course of the Ph.D. study. Until I was approaching the end of this study, did I realize that living in and cherishing the moment is more important than expecting the future.

Just like my research findings, even if I would have perfect knowledge of the future, it would never change my yearning for experiencing all the moments in this charming world.

Table of contents

Short summary

One of the primary consequences of a music-syntactic system is that listeners and performers familiar with this system automatically develop top-down predictions about upcoming notes or chords. It is a curious phenomenon that, even after having heard or played a piece of music numerous times, we still perceive music-syntactically irregular harmonies of this piece as unexpected. It is unknown whether neural correlates of music-syntactic processing would be modulated by auditory- and motor- predictive processes (as they occur during music perception and production) with prior knowledge about an upcoming musical event. The present dissertation therefore aims to address this issue by using event-related potentials.

Study 1 and 2 focused on the impact of cognitive (here referring to auditory) predictive processes. Study 3 provided a step forward in understanding the impact of action (here referring to motor) predictive processes and compared action predictive processes with cognitive predictive processes.

Study 1 investigated whether auditory predictive processes (through the course of short-term learning phase: over the span of two minutes) would modulate music-syntactic processing. For this purpose, participants (non-musicians and amateur musicians) were informed by a cue as to whether the following excerpt contained a regular or less regular chord. Study 2 investigated whether auditory predictive processes (through extensive learning phase: over a period of about half an hour) would influence the processing of music-syntactic regularities. For this purpose, participants (non-musicians and musicians) were either informed or not informed about whether the following sequence would end on a regular or irregular chord. Study 3 investigated whether auditory- and motor- predictive processes would modulate music-syntactic processing. For this purpose, professional pianists either played chord sequences with or without auditory feedback, or listened to these sequences without playing.

The main results are:

(1) Auditory predictive processes (through the course of short-term learning phase) modulated the late controlled, but not early, partially automatic, neural mechanisms of music-syntactic processing (Study 1).

(2) Auditory predictive processes (through extensive learning phase) modulated only the speed, but not the principle mechanisms, of music-syntactic processing (Study 2).

(3) Motor predictive processes did not modulate the neural mechanisms of music-syntactic processing (Study 3).

In summary, this dissertation indicates that top-down predictions of upcoming syntactic errors do not modify the bottom-up perceptual processing of syntactically irregular events.

Kurzzusammenfassung

Eine der wichtigsten Konsequenzen des musikalisch-strukturellen Systems ist, dass in mit diesem System vertraute HörerInnen und SpielerInnen automatisch top-down Erwartungen bezüglich kommender Noten und Akkorden gebildet werden. Es ist ein eigenartiges Phänomen, dass wir – selbst wenn wir ein Stück viele Male gehört oder gespielt haben – musikalisch-syntaktisch irreguläre Akkorde immer noch als unerwartet wahrnehmen. Es ist jedoch unklar, ob auditorisch- und motorisch-prädiktive Prozesse, wie sie bei der Musikwahrnehmung und -produktion auftreten, die neuronalen Korrelate der musikalisch-syntaktischen Verarbeitung modulieren, wenn vorheriges Wissen über ein zukünftiges musikalisches Ereignis vorhanden ist. Die vorliegende Dissertation widmet sich dieser Fragestellung unter Zuhilfenahme der Methode der Ereignis-korrelierten Potentiale.

Studie 1 und 2 fokussieren dabei auf den Einfluss von kognitiv-prädiktiven (hier als auditorische) Prozessen. Studie 3 widmet sich dem Einfluss von handlungsgeleiteten prädiktiven (motorischen) Prozessen und vergleicht dabei diese mit kognitiv-prädiktiven Prozessen.

Die erste Studie untersuchte, ob auditorisch-prädiktive Prozesse nach einer Lernphase von 2 Minuten Länge die musiksyntaktische Verarbeitung beeinflussen. Dazu wurden sowohl Nicht-Musikern als auch Musikern ein Hinweisreiz präsentiert, der anzeigte, ob die folgende musikalische Sequenz einen regulären oder irregulären Akkord enthielt. Die zweite Studie untersuchte, ob auditorisch-prädiktive Prozesse nach einer Lernphase von 30 Minuten Länge die musiksyntaktische Verarbeitung beeinflussen. Dazu wurden sowohl Nicht-Musiker als auch Musiker darüber informiert, ob die folgende musikalische Sequenz mit einem regulären oder irregulären Akkord endet. Die dritte Studie untersuchte, ob kognitiv-prädiktive und motorisch-prädiktive Prozesse die musiksyntaktische Verarbeitung beeinflussen.

Dazu spielten professionelle Pianisten Akkordsequenzen entweder mit auditorischem Feedback, ohne auditorischem Feedback oder hörten die Akkordsequenzen lediglich (ohne diese zu spielen).

Die Ergebnisse lassen sich wie folgt zusammenfassen:

(1) Nach einer kurzen Lernphase beeinflussten auditorisch-prädiktive Prozesse späte und kontrollierte, nicht jedoch frühe und teilweise automatische neuronale Mechanismen der musiksyntaktischen Verarbeitung (Studie 1).

(2) Nach einer langen Lernphase beeinflussten auditorisch-prädiktive Prozesse lediglich die Geschwindigkeit, nicht jedoch die grundlegenden Mechanismen der musiksyntaktischen Verarbeitung (Studie 2).

(3) Motorisch-prädiktive Prozesse beeinflussten die neuronalen Mechanismen der musiksyntaktischen Verarbeitung nicht (Studie 3).

Zusammenfassend liefert diese Dissertation starke Hinweise darauf, dass top-down Vorhersagen über kommende syntaktische Fehler nicht die bottom-up, perzeptuelle Verarbeitung von syntaktisch irregulären Ereignissen verändern.

List of publications

This dissertation is a cumulative work of manuscripts, either published or in preparation.

Study 1

Guo S, Koelsch S. 2015. The effects of supervised learning on event-related potential correlates of music-syntactic processing. Brain Res. 1626:232-246.

Study 2

Guo S, Koelsch S. 2016. Effects of veridical expectations on syntax processing in music: Event-related potential evidence. Sci Rep. 6:1-11.

Study 3

Guo S, Maidhof C, Koelsch S. (in preparation). Knowing that my prediction is wrong: The impact of auditory- and motor-predictions on neural correlates of syntactic processing in music.

List of figures

List of tables

List of abbreviations

AAHC	atomize-agglomerate hierarchical clustering
ANOVA	analysis of variance
BPM	beats per minute
dB	decibels
DD	double dominant
EEG	electroencephalography or electroencephalogram
EMG	electromyography or electromyogram
EOG	electrooculograms
ERAN	early right anterior negativity
ERP	event-related potential
FIR	finite impulse response
fMRI	functional magnetic resonance imaging
GEV	global explain variance
GFP	global field power
ICA	independent component analysis
IFG	inferior frontal gyrus
IOI	inter-onset interval
IR	incongruency response
LAN	left anterior negativity
LPC	late positive component
LTN	late tonic negativity
MEG	magnetoencephalography
MIDI	musical instrument digital interface
MMN	mismatch negativity
MNS	mirror neuron system
MRI	magnetic resonance imaging
MTG	middle temporal gyrus

RATN	right anterior-temporal negativity
ROI	region of interest
RP	repetition positivity
RT	reaction time
SCR	skin conductance response
SD	standard deviation
SEM	standard error of mean
SNR	signal-to-noise ratio
STG	superior temporal gyrus
TMS	transcortical magnetic stimulation

1. General introduction

"We cannot turn off the mind's tendency to anticipate events and
we are usually unaware of the mind's disposition to make predictions."
(Huron, 2006)

"We are hostages to our prior beliefs."
(Hohwy, 2013)

The brain is constantly predicting future events. The ability of prediction has evolutionary advantages for successful adaptation (Widmann et al., 2004). With these predictions, humans prepare fast adequate responses to forthcoming events based on actual incoming information and long-term knowledge, no matter it occurs in the sensory, cognitive, or motor domain (Bubic et al., 2009; Trainor, 2012; Rohrmeier and Koelsch, 2012).

Music is a unique auditory stimulus, encompassing events unfolding over time (Miles et al., 2016). With its properties, music provides us a helpful tool to study both cognitive processes and the neural correlates underlying these processes (Limb, 2006). This dissertation utilizes music as a tool for understanding the functions of the brain.

1.1 Theoretical background

1.1.1 Free-energy principle

Tracing back to 1870, von Helmholtz proposed that our perception of the environment is influenced by predictions (Omigie et al., 2012). The combination of his views and modern-day theories could explain a wide range of neurobiological facts.

Free-energy is the expected energy minus the entropy (or surprise) of predictions. Free-energy quantifies the amount of prediction errors (Clark, 2013), and the prediction error is regarded as free-energy in inference (Friston et al., 2006). A basic notion of free-energy principle is minimizing prediction errors over time (Friston et al., 2006; Friston, 2010; Hohwy, 2013; Winkler and Czigler, 2012; Schröger et al., 2014). The better prediction fits the actual input, the more information-theoretic free-energy is minimized.

1.1.2 Predictive coding theories

Nowadays, in cognitive science and neuroscience, predictive coding theories become the most fundamental and dominant accounts to explain the functions of the brain (Maes, 2016), and are regarded as the most basic learning mechanism in the brain (Trainor, 2012).

In the framework of predictive coding hierarchy, perception is understood as a process of perceptual probabilistic inference, and the perceptual system is regarded as a hierarchically organized model (Friston, 2005; Maes, 2016; Hohwy, 2013). Predictions from higher levels over longer time scales (top-down) are tested on actual sensory input from lower levels over shorter time scales (bottom-up) emerging one level lower with discrepancy (i.e., “prediction error”; Friston, 2005; 2009; 2010). Predictions are generated from internal generative models of the environment where an organism's prior knowledge about the world is encoded in (taken as empirical priors). In this hierarchy, predictions are passed down to explain the error away, and prediction errors are passed upwards to adjust the prediction at higher levels. No level within this hierarchy is of special relevance or carries more or less information than other levels (Winkler and Czigler, 2012; Baldeweg, 2007; Hohwy, 2013). The primary objective of the perceptual system is to minimize the prediction errors that cannot be explained by the current predictive representations, and lead to more accurate and optimal perceptions under noisy and ambiguous conditions (Schröger et al., 2014; Hohwy, 2013). The more prediction errors are minimized, the better predictions are made (Hohwy, 2013).

Prediction error, as a core concept of predictive coding framework, is essentially a feedback signal to higher levels to adjust predictions (Vuust et al., 2009; Hohwy, 2013; Wacongne et al., 2012), which drives both cognitive processing and learning (Feldman and Friston, 2010). It has been proposed that surprise feeling is associated with prediction errors at several hierarchical levels (Summerfield and Egner, 2009). Prediction errors can be minimized by either changing predictions or the things that we predict (Clark, 2013). For example, by optimizing the sensory input through attention processes or action (Feldman and Friston, 2010), by introducing perceptual bias (Geisler and Kersten, 2002), or by adjusting and optimizing higher level representations (Hohwy, 2013).

The predictive coding theories deal with probability distributions (Knill and Pouget, 2004; Winkler and Czigler, 2012; Hohwy, 2013). Relying on current sensory input and prior predictions based on experience, the brain is compelled to infer the probable causes of its sensations and constantly make inferences of sensory events, which is refer to as “Bayesian brain” (Maes, 2016; Knill and Pouget, 2004; Vuust and Witek, 2014), and the perceptual system follows the Bayesian principles unconsciously (Hohwy, 2013). Bayesian principles have been adopted to explain auditory processing in music (Clark, 2013).

Previous empirical studies have provided clear evidence for the predictive coding theories. Predictive coding models of perception are supported by findings from Nazimek et al. (2013). Besides, sensory deviances evoked event-related potentials (ERPs; e.g., the mismatch responses), which are thought to map the processing of prediction errors (see also *Impact of predictive processes on auditory sensory memory operations*) and results in an update of a regularity-based internal model (Winkler and Czigler, 2012). Even when an expected sound was absent (i.e., no auditory input), omission-related ERP responses were observed, indicating that the prediction errors are generated (Bubic et al., 2010; Wacongne et al., 2011). By contrast, prediction errors are reduced with stimulus repetition by increasing predictability, which result in the attenuation of neural activity (e.g., magnetoencephalographic responses; Furl et al., 2011). Repetition positivity (RP) component may reflect this process, which has not been directly tested yet (Haenschel et al., 2005; Costa-Faidella et al., 2011; Schröger et al., 2014; Baldeweg, 2007). When the incoming stimulus matches the prediction, no updating of the current model is required as it is part of the internal generative models. Thus, the prediction error is avoided (Bubic et al., 2010).

In the music domain, the role of prediction has been studied and framed in the context of predictive coding theories (Friston and Kiebel, 2009a; Summerfield and Egner, 2009; Wacongne et al., 2012; Vuust et al., 2009; Schröger et al., 2015a; Agres et al., 2017). For example, the perception of pitch (Furl et al., 2011; Kumar et al., 2011), rhythm and meter (Vuust and Witek, 2014; Vuust et al., 2009), melody (Pearce et al., 2010), harmony (Rohrmeier and Koelsch, 2012), the relationship between perception and action in music performance (Bianco et al., 2016b) (for a review, see Maes, 2016). According to Trainor (2012), predictive coding in music is present in 6-month-old infants by measuring response to many types of deviations (e.g., pitch, duration, timbre, rhythmic and melodic pattern), suggesting that the predictive coding is an early-developing learning mechanism (Trainor, 2012).

1.1.3 Common coding theory

1.1.3.1 Auditory-motor associative learning

James (1890) proposed that the perception and action may share common cognitive roots (James, 1890). Planning or execution of action and perception of the sensory consequences of action are tightly coupled and may prime or interfere with each other (Prinz, 1990; 1997; Maes et al., 2014; for a review, see Hommel et al., 2001). They have the common representational format and functional principles (Schröger et al., 2015a), and are linked by recruiting the same neural resources (i.e. sensorimotor areas) (Koelsch, 2012).

In the music domain, the core mechanism underlying music production is the common coding of perception and action through sensorimotor associative learning mechanisms (during music-listening: Haslinger et al., 2005; Brown and Martinez, 2007; Bangert et al., 2006; during music-production: Baumann et al., 2005; Bangert et al., 2006). The sensorimotor association refers to the reciprocal or bidirectional interactions between successions of auditory outcomes and corresponding motor programs (Elsner and Hommel, 2001; 2004). To be more specific, when playing a musical instrument, motor programs are selected to produce sounds. Based on the sounds, motor output is adjusted accordingly. The auditory-motor associative learning processes can be regarded as a central mechanism underlying the real-time predictions based on the internal models (Maes et al., 2014; Novembre and Keller, 2014), which may develop over time with practice (e.g., one week of piano training) in non-musicians (Lappe et al., 2011; 2008; Lahav et al., 2007), and are shaped through systematically years of musical training (Pfordresher, 2012; Jäncke, 2012). The notion of internal models is compatible with predictive coding theories (Friston et al., 2010; Friston and Kiebel, 2009b; Schröger et al., 2015a; Lange, 2013). There are two kinds of internal models, which are internal forward models and internal inverse models.

1.1.3.2 Internal forward models

Forward models represent the causal relationship between planned action (i.e. motor commands) and their sensory consequences (Wolpert et al., 1995; Kilner et al., 2007; Keller, 2012). Music is perceived in the way that listeners act upon and interact with it (Halpern and Zatorre, 1999). With regard to music production, when an action is planned or executed, the corresponding representations of the “intended” sounds are automatically formed well ahead of their generation (Novembre and Keller, 2014). Forward models are supported by motor dysfunctions which cause perceptual disabilities (Maes et al., 2014). A series of studies have found that when musicians played an instrument without auditory feedback, auditory imagery was evoked (Zatorre et al., 2007) and auditory-related areas were activated (Bangert et al., 2006; Bangert and Altenmüller, 2003; Baumann et al., 2005; 2007). Within the context of common coding theory, forward models automatically predict representations of the expected outcomes of actions and compare them with the actual sensory input (Timm et al., 2016), which are essential to detect deviations and help to process deviations more efficiently (Rohrmeier and Koelsch, 2012; Salimpoor et al., 2015). When the sensory input is ambiguous, perception of these events is modulated by forward models in the predicted direction (Repp and Knoblich, 2007).

There is ample evidence supporting that action can modulate perception (Bäss et al., 2008; Hughes et al., 2013; Jones et al., 2013; Loehr, 2013; SanMiguel et al., 2013; Phillips-Silver and Trainor, 2007; 2005). For example, the rhythmic movement could influence both adult and infant's rhythm perception (Phillips-Silver and Trainor, 2007; 2005). Previous research has indicated that the cerebellum is a critical locus for generating forward predictions (Halpern and Zatorre, 1999; Leaver et al., 2009; Lahav et al., 2005).

1.1.3.3 Internal inverse models

Inverse models represent that the intended or perceived sensory consequences are transferred into corresponding motor commands (Wolpert et al., 1995). It has been previously shown that merely listening to well-rehearsed pieces without performing any movement is likely to elicit conscious attempts at motor imagery (Zatorre et al., 2007) and automatically activate motor-related network responsible for producing these pieces (even when imagining familiar tunes; Halpern and Zatorre, 1999; Leaver et al., 2009) (Lahav et al., 2005; 2007; D'Ausilio et al., 2006; Bangert et al., 2006; Bangert and Altenmüller, 2003; Kristeva et al., 2003; Haueisen and Knösche, 2001; Baumann et al., 2007).

In forward models, action is perceptually guided (i.e., effects of action on perception). In inverse models, perception is action-oriented (i.e., effects of perception on action) (Repp and Knoblich, 2007).

1.1.4 Action-effect principle

Actions are planned and controlled by anticipating the perceivable effects produced by thess actions (i.e. action effects). Action effects are bidirectional, that is, action representations are coded in terms of the associated representation of its effect, and vice versa (Elsner and Hommel, 2001; Drost et al., 2005a). The ability to anticipate the perceptual effect (e.g., auditory consequences) of actions is one of the core mechanisms of action-effect principle (Koelsch, 2012; Bianco et al., 2016b).

In music perception, action effects could be explained by internal inverse models, which refers to embodied cognition (Maes et al., 2014). Action effects have the same representational format as perceived (Drost et al., 2005a; 2005b; Keller and Koch, 2008) and imagined (Stock and Stock, 2004; D'Ausilio et al., 2006; Rizzolatti, 2005; Keller and Koch, 2008) stimuli. During action planning (i.e., selecting which potential responses to execute), action effects are imagined (James, 1890; Hommel et al., 2001; Novembre and Keller, 2014). The role of action effects in action planning, control and execution is theoretically based on ideomotor principle, which is first proposed by Hermann Lotze and recently revisited (Hommel et al., 2001; Stock and Stock, 2004). Ideomotor principle and common coding theory share features of internal (forward and inverse) models. Studies of stimulus-response compatibility and response-effect compatibility have been found to provide support for the ideomotor principle (D'Ausilio et al., 2006; Knuf et al., 2001; Keller and Koch, 2008).

There is growing converging evidence from behavioral (Drost et al., 2005a), electrophysiological (Bangert and Altenmüller, 2003), functional magnetic resonance imaging (fMRI; Haslinger et al., 2005; Baumann et al., 2005; Bangert et al., 2006; Lahav et al., 2007), magnetoencephalography (MEG; Haueisen and Knösche, 2001), and transcortical magnetic stimulation (TMS; D'Ausilio et al., 2006) studies (Rizzolatti, 2005; Jäncke, 2009), consistently supporting that action and their effects must be learned through previous extensive experience (e.g., musical training) before they can be used for goal-directed action (Zatorre et al., 2007; Koelsch, 2012).

1.1.5 Mirror neuron system

In the framework of internal inverse models, the perception of actions leads to activations in the brain regions involved in executing these actions (Rizzolatti, 2005). When professional pianists listened to musical sequences they knew how to play, the MEG activity was shown in sensorimotor regions, which are involved in the control of movements (Haueisen and Knösche, 2001). These neural networks seem to behave in a similar way as the mirror neurons, which are activated by observing goal-directed actions and associated sounds produced in these actions (Kohler et al., 2002). The mirror neuron system (MNS) is regarded as the neural or physiological correlates of common coding of perception and action, and it is facilitated by sensorimotor associative learning (Maes et al., 2014; Koelsch, 2012).

1.2 Predictive processes in perception

1.2.1 Predictive brain

Our brain makes prediction all the time in a dynamic environment. In recent years, cognitive science and neuroscience have been studying the general concept of “predictive brain” in an attempt to understand the principles of brain’s function (Tillmann et al., 2014). The brain has been taken as an anticipatory machine (Bubic et al., 2010; Friston, 2005), and its information processing system has the anticipatory nature (Schröger et al., 2015b).

Prediction can take place at different time scales (Trainor, 2012) and plays a crucial role across multiple domains including perception, learning, memory, decision-making, language comprehension, and action planning (Omigie et al., 2012). Both perception and action rely on prediction (Bubic et al., 2009; Schröger et al., 2015b). A number of studies have shown the impact of prediction on processing of temporally structured events in the auditory (Hoen et al., 2006; Rüsseler and Rösler, 2000) and motor domains (Ashe et al., 2006; Keele et al., 2003).

In this dissertation, the term *prediction* is used as a general term for the overall process of future-directed information processing, and *expectation* as the representation of what is predicted to occur, including a probability distribution, and thus not necessarily representing a single event only (Rohrmeier and Koelsch, 2012).

1.2.2 Predictive processes in auditory perception

1.2.2.1 Auditory sensory memory operations

In the auditory domain, the mechanisms of predictive processes are extensively studied (Bendixen et al., 2012). The MMN (mismatch negativity, deviant- minus standard difference waves) is taken as an electrophysiological marker of predictive processing (SanMiguel et al., 2013). It has been investigated in the context of mismatch paradigms (i.e., auditory oddball paradigm; Näätänen et al., 1978; Schröger, 1998; for reviews, see Näätänen et al., 2011; Kujala et al., 2007), where frequent standard stimuli are randomly interspersed by rare deviant stimuli in the aspects of physical parameters (such as frequency, pitch, intensity, timbre, duration, rhythm, spatial localization; physical MMN) (Sams et al., 1985; Vuust et al., 2009; Schwartze et al., 2012; Tervaniemi and Huotilainen, 2003) or more abstract properties (such as auditory local / global regularities that are extracted online from acoustic environments; abstract MMN) (Paavilainen et al., 2001), or an omission of sound in a pattern (Kujala et al., 2007; Näätänen et al., 2001). Therefore, generations of the MMN are not only based on the established memory for previously presented repetitive sound, but also the regularities extrapolated from the preceding sounds (Widmann et al., 2004). Neither behavioral response nor attention to the stimuli is required in the elicitation of the MMN (Näätänen and Winkler, 1999; Kathmann et al., 1999; Näätänen et al., 1978; Paavilainen et al., 1993).

The peak latency of the MMN is about 100-250 ms after the onset of the deviance and the MMN often occurs together with an inversion of the polarity at the mastoids (when nose reference is used; Widmann et al., 2004; Leino et al., 2007). The amplitude and latency of the MMN depend upon the magnitude of the difference between standard and deviant stimuli (Näätänen et al., 1982; Sams et al., 1985) and perceptual discriminability. Accordingly, when the difference increases, the amplitude of the MMN becomes larger, and the latency of the MMN becomes shorter (Sams et al., 1985; Näätänen and Winkler, 1999). The MMN tends to be stable when the difference is above 10% (Scherg et al., 1989). The distribution of the MMN is characterized by a bilateral superior temporal gyrus (STG) with sources in and around the primary auditory cortex (Alho, 1995), and in the frontal areas as additional source (Escera et al., 2000).

In addition to the mismatch paradigms, the mechanisms of predictive processes are investigated in statistical learning paradigms (Hughes et al., 2001) and omission paradigms (Wacongne et al., 2011; Hughes et al., 2001). More recently, match and self-generation paradigms in which the incoming stimuli match predictions are adopted (Hughes et al., 2013).

Predictive coding theories (Friston, 2005; 2009; 2010) have been applied to explain the MMN in the last couple of years (Winkler and Czigler, 2012; Schröger et al., 2014). In this framework, the MMN reflects the representation of the regularities updating process (Winkler, 2007; Schröger et al., 2014), and is taken as an instance of prediction error (Vuust et al., 2009). The shorter latency of the MMN reflects that the prediction error appears in the first-order (Tavano et al., 2014).

1.2.2.2 Impact of predictive processes on auditory sensory memory operations

There are a number of studies suggesting the impact of higher cognitive systems on the deviance detection system. For example, it has been reported that the MMN is sensitive to attention (Alain and Woods, 1997): The amplitude of the MMN is larger in attended condition than in unattended conditions (Woldorff et al., 1991; 1998). In order to determine whether the deviance detection system underlying the MMN is influenced by knowledge about forthcoming stimuli available to a higher cognitive system, a series of ERP studies were conducted.

In the study by Ritter et al. (1999), participants were informed to respond to rare target tones. In a predictable condition, a rare visual cue signaled that a rare tone was about to occur, and a frequent visual cue signaled that a frequent tone would be presented. In an unpredictable condition, the visual cue was identical between rare and frequent tones. The reaction times (RTs) were longer in the unpredictable condition than in the predictable condition. Besides, the MMN was elicited by the rare tones, even if participants were informed by a visual cue that the next tone would be a deviant. The P3 was elicited by the rare visual cue but not the rare tones in either predictable or unpredictable condition. From these results, there was a dissociation between higher order cognitive system (i.e. expectation) and lower order system (i.e., what the deviance detection system was prepared for) immediately before target tones in the predictable condition: Participants did expect a deviant on the basis of the visual cue, which was indicated by the small or absent P3 elicited by the rare tones and shortened RTs. However, the deviance detection system was prepared for a frequent tone. These results indicate that the deviance detection system operates in a modular manner that does not make use of information available to higher order cognitive system (Ritter et al., 1999). On the basis of this study, Widmann et al. (2004) presented participants with auditory stimuli which were either congruent or occasionally incongruent with the preceding corresponding visual symbols. Results suggest that the deviance detection system generates expectations for sound with the same frequency as previous sounds, but the cognitive processes predict a deviant sound based on the visual symbol (Widmann et al., 2004). In line with these findings, Sussman et al. (2003; see also Wetzel and Schröger, 2007) manipulated visual cues in terms of the predictive value across conditions and showed which stages would be influenced by the visual cues. In the predictable condition, each visual cue corresponded to each pitch to signal which type of pitch would be presented. In the unpredictable condition, the visual cue was not predictive with respect to the type of pitch. The P3a was observed only in the unpredictable condition. However, the MMN reflecting the first-stage processes were not affected by the visual cues (Sussman et al., 2003). In order to specify whether brain utilizes general (i.e. probability) or specific (i.e. pitch) information of the cue, Horváth et al. (2011) then conducted an experiment in which the visual cue either signaled the specific pitch of the forthcoming deviant (fully predictable condition), or signaled the next tone was deviant or standard (predictable sound probability condition), or the visual cue was uninformative of the subsequent sound (unpredictable condition). Results suggest that the early P3a reflects processes of specific stimulus, and the late P3a reflects a more general process (Horváth et al., 2011). Based on the paradigm adopted by Ritter et al. (1999), Pieszek et al. (2013) provided participants with a visual cue which signaled contradictory predictions by occasionally violating visually induced and automatically generated auditory predictions. The MMN was elicited by the deviant even when it was congruently cued by the visual stimulus (i.e., consciously expected by participants) (Pieszek et al., 2013). These results suggest that the auditory deviance detection system underlying the generation of the MMN is not influenced by the information available to a higher cognitive system underlying the generation of the P3.

1.2.3 Predictive processes in music perception

1.2.3.1 Musical expectation

One of the primary consequences of a music-syntactic system is that listeners and performers familiar with this system automatically develop expectations (*musical expectation*; for example, the experience of “waiting for the next tone or chord”) (Tillmann et al., 2014). Expectation is therefore considered as an especially important component of music perception and production. The generation of expectations is conceived as a form of mental belief about ‘what’ and ‘when’ musical events or structures are likely to occur in an unfolding piece, based on a priori internalized musical knowledge, actual memory of musical pieces, or current information (Janata, 1995; Limb, 2006; Schmuckler, 1997). Musical expectations seem to exist in all types of music from all cultures (Janata, 1995), but responses to expectations may vary between cultures. There are two different dimensions of musical expectations, which are strength and specificity. The former refers to stronger or weaker expectation, and the latter refers to specific or unfocused expectation (Schmuckler, 1989).

Studies have repeatedly shown that the ability to form expectations in music does not require musical training, nor any explicit knowledge of listeners (Bigand et al., 2000; Koelsch et al., 2002b). For example, harmonic expectations are generated by automatic activating of implicit knowledge acquired through sufficient exposure to syntactic relationships within musical structures (Bigand et al., 2001; Koelsch et al., 2000).

Musical expectations can be driven either by a sensory or a cognitive process. Sensory process is a result of physical similarity or identity between events (e.g., step inertia, pitch proximity, melodic regression and melodic arches; Tekman and Bharucha, 1998; Justus and Bharucha, 2001; Hutchins and Palmer, 2008; Salimpoor et al., 2015), and cognitive process is a result of activation of listeners' implicit or explicit knowledge of regularities (Bharucha and Stoeckig, 1986; 1987; Schmuckler, 1989; Schmuckler and Boltz, 1994).

Musical expectations play a central role in music theory (Meyer, 1956; 1973) and psychology (Bharucha and Stoeckig, 1987; Schmuckler, 1989). In psychology, the role of musical expectations has been wildly addressed in earlier behavioral studies on music perception (Krumhansl, 1995; Schmuckler, 1989) in terms of melody, rhythm (Jones and Boltz, 1989), as well as harmonic structure (Bharucha and Stoeckig, 1986; 1987), which are fundamental aspects of music. In addition, musical expectations, which are at the origin of mood induction, or aesthetic experiences creation (e.g., surprise, tension, anticipation, suspense, disappointment, resolution, pleasure, relaxation; Meyer, 1956; Huron, 2006; for a review, see Koelsch, 2014), play a fundamental role in appreciation of music (Tillmann, 2005). Tension-relaxation patterns are implicitly expected by Western listeners in a way that tension should be followed by relaxation (i.e. a musically more stable event) (Tillmann et al., 2014), but tension is not always immediately or completely satisfied, and may be temporarily delayed or blocked. These disruptions, violations, and resolutions or satisfactions of musical expectations are often taken by composers through manipulating structural and temporal aspects of music as an attempt to please the audience (Tillmann and Bigand, 2010; Salimpoor et al., 2011). Otherwise, listeners would feel boring with too predictable music (Trainor, 2008).

1.2.3.2 Measuring musical expectation

In previous studies, production tasks are chosen to investigate musical expectations of musicians. Participants are informed to either sing or play continuation on a piano after listening to short excerpts (Schellenberg, 1996; Schmuckler, 1989). For non-musicians, musical expectations are measured by rating how well the final chord fits their expectations, or how strong the feeling of completion is after listening to the excerpts (Cuddy and Badertscher, 1987; Cuddy and Thompson, 1992).

Priming is a useful and powerful tool to investigate the impact of preceding context on the processing of target event (Bharucha and Stoeckig, 1987), which is extensively used in psycholinguistics (Tillmann et al., 2014). The priming paradigm has been employed by Bharucha and Stoeckig (1986) in music perception. In the priming paradigm, participants are presented with two chords, which are either harmonically-related or harmonically-unrelated. The processing of harmonically related targets is facilitated (i.e., greater consonance ratings, faster and more accurate responses) relative to harmonically unrelated targets, referred to as harmonic priming effects (Bharucha and Stoeckig, 1986). A connectionist model proposed by Bharucha and Stoeckig (1987) provides an explanation of harmonic priming effects. In this model, tone units are linked to chord units. Chord units are connected with each other via their parent key units. The spreading activation among musical units is determined according to the strength of the connections between them: more activation is transmitted by strong connection than weak connection (Bharucha and Stoeckig, 1987).

Harmonic priming effects have gained further support from numerous studies nowadays using priming paradigm (Tillmann et al., 2000; 2003b; Tillmann and Lebrun-Guillaud, 2006; Tillmann and Marmel, 2013). The harmonic context could either be a single chord (Tillmann and Bharucha, 2002), or a long chord sequence (Bigand et al., 2003). The harmonic priming paradigm, as an implicit access to sophisticated knowledge of structural regularities (i.e., participants are not required to make direct judgments on the relation between the prime and the target), has been employed in measuring musical expectations of non-musicians (Tillmann, 2005; Marmel et al., 2010), children and patients (Tillmann, 2005).

In addition to top-down expectations based on knowledge, harmonic priming effects could be realized by bottom-up expectations (i.e. relatively low-level perceptual processes) (Bharucha and Stoeckig, 1987; Bigand et al., 1996; Tillmann and Poulin-Charronnat, 2010; Krumhansl et al., 1999). According to Schmuckler (1989), "a chord sharing component tones, or overtones, with a preceding chord will be more highly anticipated than a continuation containing no overlapping frequencies with its predecessor" (Schmuckler, 1989). However, the harmonic priming effects are observed even when unrelated targets share more tones with the prime than related targets (Bigand et al., 2003; Regnault et al., 2001), or the confound of sensory differences between related and unrelated conditions is ruled out (e.g., the prime and the target do not share any component tones; Bharucha and Stoeckig, 1987), revealing that cognitive harmonic priming prevails over sensory repetition priming (Tillmann, 2005).

1.2.3.3 Forms of musical expectation

Top-down predictions play an important role in influencing harmonic expectations (Bigand et al., 1996). Two different kinds of top-down predictions can be differentiated, usually referred to as “veridical expectations” and “schematic expectations” (Krumhansl et al., 1999).

Regarding *schematic expectations*, according to Bharucha (1987), listeners can generalize common musical rules through years of exposure to various musical sounds, and store knowledge of general musical patterns about how single musical events are combined into musical sequences in long-term memory (i.e. *schematic knowledge*; Bharucha, 1987; Justus and Bharucha, 2001; Tillmann and Bigand, 2004; 2010; Huron, 2006). Studies have repeatedly shown that the formation of schematic expectations is passive (without conscious efforts and awareness of learning; Seger, 1994), culturally received (e.g., in Western tonal system) and occurs probably over many years (Bigand et al., 2000; Koelsch et al., 2002b). When a sound experience is coded schematically, it is taken as a style or genre (Huron, 2006). Schematic expectations can either be implicit (e.g., in non-musicians) or explicit (e.g., in musicians) (Krumhansl, 2000). Schematic expectations are often investigated with harmonic priming paradigm (see also *Measuring musical expectation*; List and Justus, 2010).

Schematic expectations can be contrasted with *veridical expectations*, which are generated either by repeated listening (i.e., the activation of memory traces for particular pieces), or by context-specific prior knowledge of the event structures (i.e., the anticipation of a specific upcoming note, chord etc.; *veridical knowledge*) (Bharucha, 1987; Huron, 2006; Miles et al., 2016). That is, the veridical expectations are formed by one's own direct experience. When a sound experience is coded veridically, it is taken as a work (Huron, 2006). Supervised learning (i.e., learning due to specific training signals or cues from the environment; Large and Palmer, 2002) serves as a way to gain veridical knowledge. Veridical expectations are usually explicit (e.g., the memory representation of the song *Happy Birthday*), but can also be implicit (e.g., when playing a piece of music by heart, the motor sequences are often executed automatically, without the player being able to tell explicitly which notes will follow) (Krumhansl, 2000). Veridical expectations are often investigated with repetition priming paradigm (List and Justus, 2010).

According to Huron (2006), musical expectations derive from experience and are thus linked to memory. He associated schematic expectations with long-term (semantic) memory, veridical expectations with episodic memory, *dynamic expectations* (i.e., shaped by immediate experience and updated in real-time through listening) with short-term and intermediate-term memory, and *conscious expectations* (i.e., listeners' conscious reflection on how musical piece will sound) with working memory. The first three expectations may operate concurrently and in parallel especially during exposure to a novel auditory experience (Huron, 2006).

1.2.4 Schematic expectations

1.2.4.1 Musical syntax

Meyer (1956, 1957) tended to investigate the relationships between musical expectations and musical structures (Egermann et al., 2013). Musical structures (e.g., melodic and rhythmic constraints, tonal and harmonic hierarchies; Janata, 1995) create musical expectations about future events (Schmuckler, 1989; Large and Palmer, 2002). Therefore, musical expectation is one reflection of the psychological reality of musical syntax (Koelsch, 2012), and the principles underlying musical expectations reflect the principles of musical structure (Huron, 2006).

It is widely accepted that humans are capable of integrating discrete basic elements (e.g., words in the language) that are arranged in a rule-governed way into higher-order structures (e.g., sentences in the language) according to a set of combinatorial principles (Kunert et al., 2015; 2016; Asano and Boeckx, 2015). *Musical syntax* is regarded as a “grammar of music” (Patel et al., 1998). It refers to a set of specific grammatical rules (e.g., harmonic syntax) governing the hierarchical combination of series of perceptually discrete elements (e.g., chords) into long complex sequences (e.g., tonal and harmonic structure; Bharucha and Krumhansl, 1983) (Asano and Boeckx, 2015; Koelsch and Siebel, 2005). Instead of syntax, Lerdahl (1991) tended to use the term ‘schema’ (i.e. style-specific knowledge) (Lerdahl, 1991), and activation of the schemata is implicitly in the brain (Bigand et al., 2003; Tillmann and Bigand, 2010). Different genres have different schemata, and common patterns of schemata exist between genres (Huron, 2006).

Musical syntax consists of a set of rules related to, for example, melody, metre, rhythm, harmony, and timbral structure (Large and Palmer, 2002; Rohrmeier and Koelsch, 2012; Asano and Boeckx, 2015). Western tonal music is dominated by harmonic syntax (Trainor, 2008). In this dissertation, we consider musical syntax as the knowledge of harmonic regularities in Western tonal music, which has been the subject of a considerable body of theoretical and neuroscientific studies (e.g., Koelsch and Siebel, 2005; Koelsch, 2005).

In this music system, the frequency range and its next higher harmonic (i.e. the octave) is divided into twelve equally spaced steps, which are called semitones (6% pitch difference) (Janata, 1995). In music theory, tone (i.e. a pitch in a musical scale) that occurs with greater frequency (e.g., the tonic, the fifth) typically serves as the most important cognitive reference point of a given key (Tillmann, 2012; Tillmann et al., 2000). From a psychological point of view, every tone is heard in relation to a hierarchically most stable tone (Tillmann, 2012; Loui et al., 2005). Each of the twelve tones can serve as a tonal center and give the name of the key (e.g., the C-major key is named after the tonic C) (Tillmann et al., 2000; Musso et al., 2015). Other pitches around the tone vary in their relative importance in the hierarchy (Jebb and Pfordresher, 2016). The restricted set of twelve different tones builds the chromatic scale (referred to as C, C# / D*b*, D, D# / E*b*, E, F, F# / G*b*, G, G# / A*b*, A, A# / B*b*, B; Tillmann et al., 2000), which is the basic element of Western tonal music (Koelsch, 2012).

A chord is comprised of three or more pitches (i.e. tones) occurring simultaneously, viewed as the vertical organizational dimensions in music (Tramo et al., 2001). The three pitches in a chord are usually called the root (the reference pitch within the chord), the third and the fifth (Loui, 2012). The degree on which the chord is built determines the structural function of this chord (Tillmann et al., 2000). Chords built on the first, the fourth, and the fifth scale degrees are major chords (e.g., a C-major chord is comprised of the pitches C, E, and G), which usually have a central syntactic function. Chord built on the first scale degree is called tonic chord (notated with the Roman numeral ‘I’), which is typically the most frequently occurring chord in a musical piece. The tonic chord tends to be used to resolve a musical sequence, providing a sense of finality (Tillmann et al., 2000; Janata, 1995). Chord built on the fourth scale degree is called subdominant chord (notated with ‘IV’), typically with its root four increments away from the root of the tonic chord. Major chord built on the fifth scale degree is called dominant chord (notated with ‘V’) (Loui et al., 2005). Major chord consists of major third interval (i.e. four semitones) and a perfect fifth interval (i.e., three semitones or a minor third interval away from the major third interval) (Limb, 2006). Chords built on the second, third, and sixth scale degrees are minor chords (Tillmann et al., 2000). A minor subdominant with a diminished sixth instead of a fifth is called Neapolitan sixth chord (N^6) (Koelsch, 2012), which is a prominent stylistic chord of tonal music (Koelsch and Friederici, 2003). Neapolitan sixth chord contains two out-of-key tones in major keys and one out-of-key tone in minor keys (Steinbeis et al., 2006). It is consonant when played in isolation, but sounds abnormal at the end of a standard harmonic progression (i.e., in place of the final tonic), which violates typical musical syntax (Loui et al., 2005). A secondary dominant is called dominant to the dominant or double dominant (DD) (Koelsch, 2012). Chord built on the seventh scale degree is diminished chord, in which the third is minor, and the fifth is diminished (Tillmann et al., 2000; 2003b).

Tones and chords over longer temporal windows form keys, which are the largest building blocks of the Western tonal system (Tillmann et al., 2000). A key consists of two subsets of pitches: subsets of seven tones are called diatonic scales, and subsets of five tones are called nondiatonic set (Janata, 1995). Major and minor keys consist of overlapping sets of seven tones (Morimoto et al., 2016). Western tonal music relies on a formal geometric structure, which is referred to as the Circle of Fifths (see Figure 1) (Bigand and Pineau, 1997; Musso et al., 2015). The Circle of Fifths determines the harmonic distance relationships among two keys by the number of steps between them (whatever the direction of the rotation) (Musso et al., 2015). For example, C- and F- or G- major keys are closely related, in contrast to C- and F#- major keys, which are less-related.

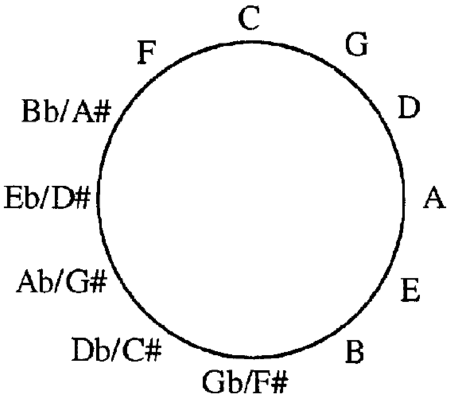

Figure 1. General introduction - The Circle of Fifths (for major keys; Tillmann et al., 2000).

Same chords may belong to the several keys, for example, the C-major chord belongs to the C-, F-, and G- major keys, the A- and E- minor keys. Thus, the same chord (e.g., C-major chord) may function as a stable tonic chord acting the most referential role in one context (e.g., in a C-major key), and function as a less stable dominant or subdominant chord in another context (e.g., in a F- or G- major key), or even function as an out-of-key chord (Bigand and Pineau, 1997; Tillmann et al., 2000; 2003b). Some chords may belong to the same parent keys, for example, the C and B*b* chords share the key of F (Bigand and Pineau, 1997). Chords belonging to the same keys are harmonically related and more likely to co-occur in a given musical piece (Tillmann et al., 2000; 2003b).

1.2.4.2 Music-syntactic processing

Music-syntactic processing is based on the predictive models for the regularities of tonal structure. Musically naive listeners acquire implicit knowledge of music-syntactic regularities incidentally over long-term exposure to the tonal relationships in everyday life (through statistical learning and other learning mechanisms; McMullen and Saffran, 2004; Rohrmeier and Rebuschat, 2012) (Tillmann et al., 2000; Bigand and Poulin-Charronnat, 2006). Numerous research has provided evidence that listeners familiar with Western tonal music are sensitive to the tonal structure and its underlying regularities, which may not necessarily be verbalized by them (Tillmann and Bigand, 2004). Sensitivity to music-syntactic regularities is observed early in life (e.g., with a few years of passive exposure to tonal music) (Koelsch et al., 2003a; Schellenberg et al., 2005; Trainor, 2005). Findings suggest that tonal schemata are even observed in the brain of amusia (Zendel et al., 2015).

1.2.4.3 Neural correlates of music-syntactic processing

Previous neurophysiological studies have investigated ERP responses to music-syntactic processing associated with the violations of harmonic contexts. Patel et al. (1998) found that a right anterior-temporal negativity (RATN), peaking at 300 to 400 ms, was elicited by harmonically unexpected chords in the middle of musical sequences. The RATN is speculated to reflect the application of music-syntactic regularities and working memory for musical events (Patel et al., 1998). Koelsch et al. (2000) investigated harmonic regularity violations by presenting participants music-syntactically irregular / incongruous / unexpected chords in the middle or the end of chord sequences. The irregular chords elicited a negative deflection typically observed at around 150 to 280 ms and peaked at around 180 ms over anterior regions with a tendency of right-hemispheric lateralization. The early right anterior negativity was denoted by Koelsch et al. (2000) as the ERAN (Koelsch et al., 2000), which is the most heavily investigated component taken to reflect the violations of music-syntactic regularities and regarded as a special index of Western tonal system (Koelsch et al., 2000; Maess et al., 2001; Koelsch et al., 2002b; 2002c; Loui et al., 2005; Steinbeis et al., 2006; Leino et al., 2007; Miranda and Ullman, 2007; Carrión and Bly, 2008; Koelsch and Jentschke, 2008; Garza Villarreal et al., 2011; Tervaniemi et al., 2012; Kalda and Minati, 2012; Sammler et al., 2013a; Brattico et al., 2013; Koelsch et al., 2013; Guo and Koelsch, 2015; Ellison et al., 2015; Zendel et al., 2015; Guo and Koelsch, 2016; Jaśkiewicz et al., 2016; for a review, see Koelsch, 2012). In numerous subsequent studies, the ERAN is elicited by, for example, DD, N^6, deviant notes (in melodies), or irregular versions of excerpts from authentic musical stimuli.

The sensitivity of amplitude and latency of the ERAN has been demonstrated to reflect the salience of harmonic irregularities (Garza Villarreal et al., 2011; Koelsch and Sammler, 2008; Kim et al., 2011). The stronger music-syntactic irregularity is, the larger ERAN is elicited (Koelsch et al., 2000; Steinbeis et al., 2006). However, the long-term representations of syntactic regularities (as reflected in the ERAN) are remarkably stable: In a study by Koelsch and Jentschke (2008), two types of short (3.6 s) chord sequences, of which the final chord was syntactically either regular or irregular, were presented to participants for approximately two hours (participants were watching a silent movie with subtitles). It was found that the amplitude of the ERAN declined over the course of the experiment, but was still significant at the end of the experiment, indicating that cognitive representations of music-syntactic regularities cannot easily be modified (Koelsch and Jentschke, 2008). This finding is consistent with results of a study by Carrión and Bly (2008) which reported that the amplitude of the ERAN elicited by irregular chords did not increase when participants underwent an implicit learning session in which they were presented with eighty-four training sequences that ended on syntactically correct chords (Carrión and Bly, 2008).

The lateralization of the ERAN has not reached a consistent conclusion in previous studies. It is generally right-lateralized (Patel et al., 1998; Koelsch et al., 2000; 2002b; 2002c) or bilateral (Loui et al., 2005; Leino et al., 2007; Guo and Koelsch, 2015; 2016). Moreover, the lateralization of the ERAN is modulated by attention (Koelsch et al., 2001; 2002c) and gender in children (Koelsch et al., 2003a).

Attention is a critical factor in the processing of musical syntax. In most of previous studies, the allocation of attentional resources was manipulated through task-relevance. The ERAN was elicited when participants were instructed to focus on chords with a different timbre (i.e., attention was diverted from detecting the harmonical irregularities; Koelsch et al., 2000), perform attention-demanding reading tasks (e.g., read a self-brought book or comprehension passages; Koelsch et al., 2002c; Loui et al., 2005), or when their attention is directed to a concurrent speech stimulus (Maidhof and Koelsch, 2011). The ERAN was appeared even during sedation, but the amplitude of the ERAN decreased with increasing sedation, and was completely abolished during unconsciousness (Heinke et al., 2004). Thus, the elicitation of the ERAN operates in the absence of attention, but the amplitude of the ERAN is modulated by attentional demands (Koelsch, 2012). The degree of the impact of attention on music-syntactic processing was lesser in musicians than in non-musicians (Loui et al., 2005).

Both the MMN and the ERAN are taken as reflection of detecting deviations of regularities (Koelsch, 2009). The ERAN has previously been suggested to be regarded as a "music-syntactic MMN" (Koelsch et al., 2002b; 2002c; Koelsch and Friederici, 2003; Koelsch et al., 2003b). With respect to the differences between the MMN and the ERAN, the MMN reflects the detection of abstract-feature deviants represented in short-term auditory sensory memory, whereas the ERAN is elicited by highly abstract feature format (i.e. music-syntactic regularities) represented in long-term memory (for a review, see Koelsch, 2009).

The ERAN is followed by a late bilateral frontal negativity, namely the N5, which is elicited by harmonically incongruous chords. The N5 is observed at around 380 to 550 ms and peaks at around 500 to 550 ms, which has a bilateral frontal distribution in musicians and a right-lateralization in non-musicians (Koelsch et al., 2000; 2002b; 2002c; Loui et al., 2005; Miranda and Ullman, 2007; Koelsch et al., 2013). It has been proposed to reflect the musical knowledge-based process of harmonic integration of each incoming chord into the ongoing sequence on a cognitive level (Koelsch et al., 2000; 2002c; Loui et al., 2005; Steinbeis et al., 2006; Poulin-Charronnat et al., 2006). The amplitude of the N5 is sensitive to the degree of difficulty (i.e. processing effort) of integration. The N5 is found to be less automatic than the ERAN (Koelsch et al., 2000; Loui et al., 2005) and is heavily modulated by attention (Koelsch et al., 2002c). Both the schematic (i.e. rule-based) and veridical (i.e. memory-based) expectations have impact on this component (Miranda and Ullman, 2007).

Besides the ERAN, musical expectations have been reported to be reflected in a number of other neurophysiological markers, such as the P300 (Regnault et al., 2001; Patel et al., 1998; Janata, 1995), late positive component (LPC; Besson and Faïta, 1995; Regnault et al., 2001; Besson et al., 1994) or P600 (Patel et al., 1998;

Featherstone et al., 2013), which are elicited in chord sequences and melodies. These components are thought to reflect structural re-analysis and task-related decision processes (Koelsch, 2011; 2012).

Among these components, the P300 complex consists of two different components: the P3a and the P3b. The ERAN is followed by a P3a when the incongruous event is highly salient or participants are asked to ignore the auditory stimuli (Koelsch et al., 2000; Regnault et al., 2001; Koelsch et al., 2007; Miranda and Ullman, 2007; Carrión and Bly, 2008; Brattico et al., 2013), which is a fronto-centrally distributed positive deflection between 200 to 400 ms, and peaks at about 300 ms (Polich, 2007; Friedman et al., 2001). The P3a has been taken to reflect an involuntary shift or orienting of attentional focus toward deviant sounds (Münte et al., 1998; Escera et al., 2000), and the generation of the P3a is sensitive to musical training (Regnault et al., 2001). In the framework of predictive coding theories (Friston, 2005; 2009; 2010), the P3a has been proposed as the outcome of the evaluation process between listeners' predictions and representations of incoming sounds (Schröger et al., 2014). Within the P300 complex, the P3a is usually followed by another positive going component P3b (often just 'P300'; Koelsch et al., 2000; Carrión and Bly, 2008; Ellison et al., 2015), which exhibits a more posterior scalp distribution within a range of 300 to 800 ms (Comerchero and Polich, 1999), and is commonly elicited by infrequently occurring target events that are attended and task-relevant (Friedman et al., 2001; Katayama and Polich, 1998; Polich, 1988). The P3b is generally interpreted as updating of working memory representations or decisional processes (Donchin and Coles, 1988; Polich, 2007). The amplitude of the P3b is sensitive to the probability of the target events and the effort participants devote to a task (Luck, 2014).

The P300 evoked by unexpected musical events was referred to as LPC. The label LPC was used due to the prolonged morphology of this P300-like component, and due to the fact that task-relevant music-syntactic irregularities can evoke an LPC. Thus, the LPC resembles the P600 elicited in language experiments, and is taken to reflect controlled (intentional and task-dependent) processes of structural reanalysis and integration during the processing of rule-based sequences when participants attend to the stimulus and consciously detect target chords (or tones; Besson and Schön, 2001; Patel et al., 1998). The LPC / P600 develops between 300 and 800 ms and peaks at around 600 ms with central and posterior distributions (Besson et al., 1994; Besson and Faïta, 1995; Patel et al., 1998; Regnault et al., 2001). The amplitude and latency of the LPC / P600 are sensitive to the familiarity with the stimuli (Besson and Faïta, 1995), the degree of unexpectedness (Besson and Faïta, 1995; Patel et al., 1998), the timing of the incongruous event (Nittono et al., 2000), and musical training (Besson and Faïta, 1995).

1.2.4.4 Neuroanatomical correlates of music-syntactic processing

Previous studies have suggested the neural generators of the ERAN are located in inferior frontal gyrus (IFG: BA 44, the left hemisphere is referred to as part of Broca's area; Maess et al., 2001; Koelsch et al., 2005; Garza Villarreal et al., 2011), extending into the insular cortex, and are co-activated with the anterior / posterior superior temporal gyrus (aSTG / pSTG). The IFG has been suggested to support integrating discrete items into higher-order structures (Koelsch et al., 2002a; 2005; Tillmann et al., 2003a; 2006; Minati et al., 2008; Sammler et al., 2011; Kim et al., 2011; Musso et al., 2015; Bianco et al., 2016b). Musicians show stronger activations of IFG than non-musicians in the processing of harmonically incongruent chords (Janata et al., 2002). It has been previously shown that the white matter connectivity in IFG and STG is associated with the structural analysis of auditory sequences (Loui et al., 2011). Congenital amusia (with deficits in music perception) is suffered from the disruption of STG-IFG pathways (Loui and Schlaug, 2009; Hyde et al., 2011).

1.2.5 Veridical expectations

1.2.5.1 Mere exposure effect

The veridical expectations have been found to facilitate perception (Besson and Faïta, 1995). Familiar musical sequences are frequently used in studies addressing veridical expectations by activation of memory traces for particular pieces. Familiarity with musical excerpts leads to more accurate veridical expectations, which makes the music more predictable (Besson and Faïta, 1995; Huron, 2006). Considerable empirical studies have reported that familiar music is more favored than unfamiliar music (Heyduk, 1975; Hargreaves, 1987; Peretz et al., 1998; Loui, 2012), which refers to (mere) exposure effect. That is, familiar music may induce more positive emotional responses and tends to be experienced as more pleasurable than unfamiliar music (Huron, 2006; van den Bosch et al., 2013; Declos, 2014).

1.2.5.2 Neural correlates of veridical expectations

Up to now, most previous neuropsychological studies concerned with musical expectations are mainly focused on schematic expectations (see also *Neural correlates of music-syntactic processing*). The role of veridical expectations has been investigated using either familiar musical sequences (behavioral studies: Bigand et al., 2005; Dowling et al., 1987; McAuley et al., 2004; ERP studies: Besson and Macar, 1987; Verleger, 1990; Paller et al., 1992; Besson et al., 1994; Besson and Faïta, 1995; Miranda and Ullman, 2007), or implicit learning tasks (e.g., prior exposure to regular chord sequences; Carrión and Bly, 2008).

The P300, N400, LPC / P600 are elicited by wrong tones occurring in familiar melodies (i.e. violations of veridical expectations; Besson and Macar, 1987; Verleger, 1990; Paller et al., 1992; Besson et al., 1994; Besson and Faïta, 1995; Patel et al., 1998). For example, Besson and Faïta (1995) found that the amplitude of the LPC elicited by the incongruous ending of familiar melodies was larger than of unfamiliar melodies (Besson and Faïta, 1995). Veridical expectations are investigated particularly in comparison to schematic expectations. Miranda and Ullman (2007) investigated the dissociation between schematic and veridical expectations and found that out-of-key deviant notes that violated tonal harmony rules in unfamiliar melodies (i.e. schematic expectation violations) elicited the ERAN, and in-key deviant notes in familiar melodies (i.e. veridical expectation violations) elicited an N400 (Miranda and Ullman, 2007).

1.2.6 Relationship between schematic and veridical expectations

Many scholars attempted to distinguish between schematic and veridical expectations. Dowling and Harwood (Dowling and Harwood, 1986), as well as Meyer (Meyer, 1961) proposed that schematic expectations may work on an automatic, or "subconscious" level (Tillmann and Bigand, 2010). Bharucha and Stoeckig (1987) suggested that schematic expectations may conflict with veridical expectations in terms of favoring different chords in a deceptive cadence (e.g., schematic expectations: tonic chord, veridical expectations: subdominant chord) (Bharucha, 1987). From the point of view given by Tillmann and Bigand (2010), schematic and veridical expectations should tap into two different cognitive processes that work independently (Tillmann and Bigand, 2010). In other words, schematic knowledge works independently as a module (i.e., "'informationally encapsulated' from long-term memory of pieces"), which has no access to veridical knowledge (Jackendoff, 1991; Margulis, 2005). According to Huron (2006), "the schematic brain is surprised by the deception while the veridical brain is not" (Huron, 2006). Therefore, it is possible to leave schematic surprise intact while reducing veridical surprise (Huron, 2006). Salimpoor et al. (2011) pointed out that schematic and veridical expectations are not mutually exclusive in the aspect of developing anticipation (Salimpoor et al., 2011).

It has been proposed that schematic and veridical expectations operate in a highly functionally modular fashion in terms of memory system (Huron, 2006), cognitive processes (Justus and Bharucha, 2001; Tillmann and Bigand, 2004; 2010), and neural correlates (Miranda and Ullman, 2007; Huron, 2006). Previous empirical studies have explored the relationship between schematic and veridical expectations, which is always susceptible to change, depending on the listener's auditory experience (Huron, 2006) and familiarity with a given stylistic tradition (Krumhansl et al., 2000).

Peretz (1996) reported that patients with severe musical agnosia showed damaged memories for well-known musical excerpts (due to veridical knowledge), but little-impaired capacity of experiencing (and enjoying) music (due to schematic knowledge) (Peretz, 1996). Krumhansl et al. (1999) found that schematic expectations were stronger for nonexperts than experts, and veridical expectations were stronger for experts than nonexperts (Krumhansl et al., 1999). A few behavioral studies directly investigated the mutual influence of schematic and veridical expectations on the processing of harmonic regularities. Justus and Bharucha (2001) conducted a series of harmonic priming experiments by manipulating the veridical expectations. The veridical knowledge was provided in various situations, which were prime-target previews, local transition probabilities, or valid versus invalid previews. Results showed that the schematic harmonic priming effects (i.e., expected chord pairs were processed faster than unexpected chord pairs) remained unchanged even when unexpected chord pairs occurred more often or had been processed directly before the target pair. This finding suggests that the impact of schematic expectations is stronger than veridical expectations on chord processing. It has been argued that the weak influence of veridical expectations might be due to the use of short musical stimuli (i.e. two chords) because complete forms are more easily stored in memory than fragments in the linguistic and visual domains. In order to generate strong veridical expectations, similar studies were conducted by Tillmann and Bigand (2004; 2010) in which longer musical stimuli were used in the prime context. They investigated how far the schematic expectations could be influenced by veridical expectations induced by repetition priming in the exposition phase prior to the experiment. During the exposition phase, half of the participants were presented with numerous sequences ending on the most related chord (i.e. the tonic chord), and another half of participants were presented with sequences ending on the less-related chord (i.e. the subdominant chord). Results showed that the related chords were processed faster (i.e. a facilitation effect of the tonic chord), even after listening to chord sequences ending on the subdominant chord repeatedly. This finding underlines the strength of schematic expectations (Tillmann and Bigand, 2004; 2010). However, Creel investigated whether listeners activated veridical memory for a particular melody when processing musical harmony and meter (i.e. schematic memory). It was found that for both harmonic and metrical information, previous familiarization shifted listeners' preferences towards a probe they had become familiarized with, indicating the strong effects of veridical memory in addition to listeners' schemata (Creel, 2011).

The above-mentioned studies (Justus and Bharucha, 2001; Tillmann and Bigand, 2004; 2010) are in line with our daily listening experience that we like familiar music even though we have specific veridical expectations for what will come next (see also *Mere exposure effect*). Otherwise, well-known or highly familiar music would lose its attraction.

1.3 Predictive processes in action

1.3.1 Syntax of music production

Music performance is a time-based sequential behavior which demands functional integrations of perceptual (e.g., auditory, visual), motor (e.g., somatomotor), cognitive, and emotive skills. It relies heavily on the rapid and precise mapping between motor activity (i.e. hierarchically organized hand movements) and their resulting auditory effects (i.e. perceived sounds; Zatorre et al., 2007; Launay et al., 2016), in combination with predictions acquired through years of musical training and extensive motor practice (Pfordresher and Palmer, 2006; Maes, 2016), which remind musicians to prepare the motor commands of current and peripheral evetns (Ruiz et al., 2009).

The notion of investigating syntax with regard to goal-directed actions was first introduced by Lashley (1951) (Asano and Boeckx, 2015; Ruiz et al., 2009). Recently, a great deal of attention has been paid to investigate syntax of music and language in terms of action (Jackendoff, 2009; Boeckx and Fujita, 2014; Fitch and Martins, 2014; Pulvermüller, 2014). The hypothesis that processing of syntactic-like hierarchical structure is a prominent feature of music perception, which is considered to be shared with musical actions was raised (Fadiga et al., 2009; Koelsch, 2012). Novembre and Keller (2011) explored the impact of syntactic knowledge on musical actions using imitation paradigm. Results showed that imitation of syntactically incongruent chords was slower and had a higher number of errors than imitation of congruent chords, suggesting that syntactic rules determining the observation of musical actions evoke motor expectations that influence pianists' action execution (Novembre and Keller, 2011). Jebb and Pfordresher (2016) found that the action-perception circuits in musical actions are also based on schematic knowledge (Jebb and Pfordresher, 2016). These findings indicated that music-syntactic information regulates not only music perception but also motor programs in musical action, which has been acquired over years of extensive practice.

1.3.2 Neural correlates of syntactic processing during music production

Recently, a growing number of neurophysiological studies have investigated the processing of syntactic regularities during music production (ERP: Sammler et al., 2013b; Bianco et al., 2016a, fMRI: Bianco et al., 2016b). In the ERP studies with priming paradigms, expert pianists were asked to watch silent videos (Sammler et al., 2013b) or sequences of photos (Bianco et al., 2016a) of a model hand playing five-chord sequences (i.e. visually guided), and to simultaneously imitate these chord sequences with one hand on a muted keyboard. The last chord was either syntactically congruent or incongruent with the preceding harmonic context. To induce different strength of predictability, pianists were presented with two-chord sequences as well. During the imitation of incongruent chord sequences, a right anterior negativity was elicited in the study by Sammler et al. (2013b). This might be interpreted as an ERAN in the visuomotor modality, evoked by expecting the action outcome in the videos based on the hand trajectory towards the final chord (Sammler et al., 2013b). The ERAN was not found in the study by Bianco et al. (2016a), which was interpreted as lacking auditory input (i.e., tied to auditory detection of syntactic irregularities) and / or the way of presenting model hand (i.e. perceptual continuity) (Bianco et al., 2016a). In addition, a right-lateralized posterior negativity (Sammler et al., 2013b) and a late centroparietal negativity (Bianco et al., 2016a) were elicited by incongruent chords, reflecting mechanisms of re-programming of preplanned dominant responses (i.e., playing congruent chords). Taken together, these studies show that motor system of skilled pianists makes use of long-term music-syntactic knowledge for expecting forthcoming actions (i.e. motor predictions) during action observation and imitation (Sammler et al., 2013b). The musical syntactic rules acquired motorically are internalized as a form of "embodied harmony", which refers to as "grammar of musical action" (Sammler et al., 2013b; Bianco et al., 2016a).

1.3.3 Neuroanatomical correlates of syntactic processing during music production

One fMRI study investigated how music-syntactic processing in musical actions (without auditory feedback) related to this processing in music perception. Imitating harmonic sequences using one hand on a glass-board were compared with listening to the same chord progressions, and results showed distinct sub-regions in the right inferior frontal gyrus (rIFG) interconnected with parietal and temporal areas for processing silent action and audio sequences separately (Bianco et al., 2016b).

1.4 Effects of musical training

1.4.1 Structural changes in the brain

Numerous previous empirical studies have investigated the structural and functional changes taken place in the brain between musicians and non-musicians with various experimental techniques, such as behavioral assessment, EEG (electroencephalography), MEG, magnetic resonance imaging (MRI) and fMRI. These studies have gained insights into neuroplasticity of sensory and motor regions (Tervaniemi et al., 2012).

In terms of structural changes in the brain, it has been previously shown that musicians have increased density of right medial orbital gyrus (in processing of tonal sensitivity), left IFG (in processing of syntactic processing), and bilateral perirolandic and striatal areas (Brown and Palmer, 2013). In addition, musicians have enhanced white matter tracts related to dorsal and ventral auditory processing (Halwani et al., 2011), higher tract consistency of the right hemisphere ventral stream pathway (Oechslin et al., 2017), increased volume of gray matter in auditory (Schneider et al., 2002) and motor regions (Gaser and Schlaug, 2003; Gaab and Schlaug, 2003), and reduced volume middle temporal gyrus (MTG) and the pars opercularis (Oechslin et al., 2017).

1.4.2 The impact of musical expertise on music-syntactic processing

Musicians, who have established internalized knowledge of Western tonal music due to long-term musical training, are superior to non-musicians in various stages of auditory processing. With regard to music-syntactic processing, (amateur) musicians are more proficient at detecting violations of musical syntactic regularities as reflected in the shorter RTs, higher correct responses (Bigand et al., 1999), and larger amplitude of the ERAN compared to non-musicians (Koelsch and Siebel, 2005; Koelsch et al., 2002b; Steinbeis et al., 2006; Koelsch and Jentschke, 2008; James et al., 2008). Although such group difference is just above or did not reach statistical significance in some studies (Regnault et al., 2001; Koelsch et al., 2007; Koelsch and Jentschke, 2008; Steinbeis et al., 2006; Koelsch and Sammler, 2008; Guo and Koelsch, 2015; 2016), the amplitude of the ERAN is larger for (amateur) musicians than non-musicians. In the framework of predictive coding theories (Friston, 2005; 2009; 2010), musicians are more sensitive to auditory prediction errors on the basis of more specific representations of music-syntactic regularities (Vuust et al., 2009).

1.5 Research aims

The primary aim of the present dissertation is to investigate the impact of auditory- (Study 1 and 2) and motor- (Study 3) predictive processes on neural correlates of syntactic processing in music.

Investigating the impact of predictive processes on the processing of music-syntactic regularities is important and promising because syntactic processing established through long-term learning should have some flexibility on the one hand, however, on the other hand, it should not easily be modified by each specific experience. These questions are of fundamental importance because they not only allow us to further our understanding of predictive processes in the auditory and motor domains, but also provide us how music production would influence neural correlates taken to reflect music perception (particularly with regard to music-syntactic processing). In addition, it may help us to better understand the particularly curious phenomenon that highly familiar musical pieces (in contrast to, for example, literature, movies) continue to be interesting, even when we are perfectly capable of predicting the irregular events after having heard or played the pieces numerous times.

2. General methodology

In order to assess changes in the brain electrical activity associated with the online processing of a specific event, electroencephalography (EEG) data was recorded. The general methodology provides a brief introduction to the event-related potentials (ERPs) technique and the rationale for adopting spatiotemporal ERP analyses (in Study 3).

2.1 Neurophysiological bases of the EEG

The activity of large numbers of neurons at the same time primarily in the neocortex results in summation of electrical activity (Huron, 2006). Using suitably sensitive electrodes situated at the scalp, electrical activity originating from the brain activity is recorded, which is called the EEG, with the amplitudes varying between 1-100 μV. The term EEG was first introduced by Berger (1929) when measuring human brain activity (Koelsch, 2012).

2.2 Recording of ERPs

After repeatedly presenting the same or similar stimulus in each every trial of an experiment, the signal (but not the noise) in the brain activities systematically correlates with the presentation of the stimuli. In order to obtain a higher signal-to-noise ratio (SNR), activities of all the valid trials are averaged (the stimulus-uncorrelated noise is averaged out) relative to the onset of a critical stimulus, making it possible to extract the stimulus-correlated signal from the noise. An ERP waveform is yielded after this time-locked average (Koelsch, 2012).

ERPs have a number of main advantages. As a noninvasive technique, ERPs offer a high temporal resolution (i.e. in a millisecond range; Koelsch and Friederici, 2003; Regnault et al., 2001). In addition, neither overt (e.g., verbal) responses nor conscious awareness are required in the ERPs recording (Huron, 2006), which is capable of detecting implicit knowledge of participants (Jentschke et al., 2014). Besides, ERPs may determine which neuro-cognitive process is influenced by an experimental manipulation (Luck, 2014).

ERPs consist of a series of peaks and troughs, which are usually characterized by their peak polarity (positive or negative), latency (i.e., the time point where a component reaches its maximum amplitude; Coulson et al., 1998), scalp distribution, and functional significance (Koelsch, 2012). The peak polarity (i.e. ‘P’ or ‘N’) and approximate peak latency (i.e. in milliseconds) are used to label an ERP component (Coulson et al., 1998). Whether and how (in terms of the amplitude and / or the latency) a particular ERP component is modulated by an experimental manipulation can be used to test hypotheses about the underlying processes (Coulson et al., 1998).

2.3 Spatiotemporal ERP analyses

Differences observed in amplitudes can be a consequence of a modulation in the strength of the electric field, of a scalp topographic change of the electric field (revealing distinguishable brain generators), or of latency shifts in brain processes. To differentiate these effects, spatiotemporal ERP analyses (i.e. topographic analyses) is performed to the grand-averaged ERPs (Michel et al., 2004; 1999a; 1999b) using the software Cartool (Functional Brain Mapping Laboratory and CIBM, Geneva, Switzerland; http://brainmapping.unige.ch/Cartool.htm) (Brunet et al., 2011). Topographic analyses aim to objectively identify the successive stable microstate maps (i.e. map topographies) and the time intervals where they rapidly change from one stable configuration into another (i.e. segment borders) (Arzy et al., 2007; Michel et al., 2001; Murray et al., 2008), following the notion that microstate maps are not randomly distributed across time, but remain stable for successive processing states before an abrupt qualitative change in the electric field configuration occurs (Lehmann et al., 1987; Michel et al., 1999a; 1999b). Each map presumably reflects different functional stages of the brain, which are referred to as “functional microstates" (Brandeis and Lehmann, 1986a; Pascual-Marqui et al., 1995).

First, topographic pattern analysis is conducted, based on an Atomize-Agglomerate Hierarchical Clustering (AAHC) method (Britz et al., 2009; Murray et al., 2008), which is used to identify the most dominant map topographies that represent the microstates (Michel et al., 2001; Pascual-Marqui et al., 1995). The optimal number of microstate maps is determined by means of a cross-validation criterion (Pascual-Marqui et al., 1995) and a Krzanowski-Lai criterion (Brunet et al., 2011; Michel et al., 2004; Murray et al., 2008; Pascual-Marqui et al., 1995). Map durations of less than 20 ms are considered physiologically implausible, and a temporal criterion of at least 10 consecutive time frames (≥20 ms) is set accordingly. This approach allows summarizing ERP data into a limited number of electrocortical map configurations and identifying time periods during which different experimental conditions evoke different maps (Michel et al., 2001).

Next, a fitting procedure is applied to individual data to determine whether the microstate maps identified in the grand-averaged ERPs could be identified in individual subject. This procedure consists of calculating for each subject strength-independent spatial correlation coefficients between the microstate maps identified by the topographic pattern analysis and the ERPs of each subject in corresponding time intervals (for a review, see Michel et al., 2001). Statistical analyses are conducted to assess how frequently each map is identified in each subject. A significant statistical difference indicates that one condition was significantly better explained by one given map than another. This map is thus specific to this condition. The global variance (GEV) that each map explained is calculated as well. The GEV equals the sum of the explained variances over the time windows of interest, weighted by the strength of the map at each moment in time. Statistical analyses are also performed on the first time frame of appearance (FOnset) and the time of occurrence of the highest correlation value yielded (TFBCorr) for a given map. These indexes are used to reveal differences in timing, which are topographic onset and latency. Moreover, analyses on the position of the max Global Field Power (TFMaxGFP) over the segment fitted by a given map are conducted, which represent the topographic peak latency (Murray et al., 2008). The above-mentioned parameters (FOnset, TFBCorr and TFMaxGFP) are used to determine the onset and offset of time windows for further ERP waveform analyses.

In contrast to conventional waveform analyses, topographic analyses rely on the distribution of the electric potential at the scalp (i.e. the topographic map), which is recorded from a multichannel electrode array. Moreover, topographic analyses give non-ambiguous and reference-independent information about changes in the underlying generators in time and between conditions (Murray et al., 2008).

3. Empirical studies

3.1 Study 1-3: Research questions and hypotheses

The general research question of this dissertation is whether predictive processes in the auditory and motor domains would modulate the neural correlates of (partially) automatic and involuntary processing of schematic knowledge stored in long-term memory. To be more specific, whether processing of an ‘unexpected’ chord would be influenced by the fact that participants knew the to-be-expected event was an ‘unexpected’ chord.

In the three studies, participants were informed by a cue as to whether the following excerpt or sequence they would listen to or play contained a regular or irregular chord. This method was chosen to serve as a way of generating veridical expectations.

The general research question can be partitioned into three sub-questions:

a) Whether veridical expectations acquired through the course of short-term learning phase (over the span of two minutes) would modulate music-syntactic processing (Study 1).

b) Whether veridical expectations acquired through extensive learning phase (over a period of about half an hour) would modulate music-syntactic processing (Study 2).

c) Whether music-syntactic processing would be modulated by auditory- and motor- predictive processes as they occur during music performance (Study 3).

It is also of interest whether long-term musical training would impact on the processing of expectations based on veridical and schematic knowledge. In order to address this issue, Study 1 and 2 compared (amateur) musicians who have specific representations of musical regularities with non-musicians who do not have explicit knowledge of syntactic regularities. Study 3 provided insights into professional musicians whose brain is an excellent example of auditory-motor coupling where actions and intended sounds become strongly associated after long-term musical training and extensive practice.

Based on the research questions above, hypotheses were put forward:

a) Music-syntactic processing would be modulated by veridical expectations acquired through the course of short-term learning phase (Study 1).

b) Music-syntactic processing would be modulated by veridical expectations acquired through extensive learning phase (Study 2).

c) Music-syntactic processing would be modulated by auditory- and motor-predictive processes (Study 3).

3.2 Study 1: The effects of supervised learning on event-related potential correlates of music-syntactic processing[1]

---

[1] This part was published as

Guo S, Koelsch S. 2015. The effects of supervised learning on event-related potential correlates of music-syntactic processing. Brain Res. 1626:232-246.

Accessible online at http://dx.doi.org/10.1016/j.brainres.2015.01.046

3.3 Study 2: Effects of veridical expectations on syntax processing in music: Event-related potential evidence[1]

---

[1] This part was published on

Guo S, Koelsch S. 2016. Effects of veridical expectations on syntax processing in music: Event-related potential evidence. Sci Rep. 6:1-11.

Accessible online at http://dx.doi.org/10.1038/srep19064

3.4 Study 3: Knowing that my prediction is wrong: The impact of auditory- and motor-predictions on neural correlates of syntactic processing in music[1]

3.4.1 Introduction

Our daily listening experience suggests that after having heard a piece of music numerous times (i.e., with auditory predictions based on knowledge of the piece), music-syntactically irregular harmonies of this piece are still perceived as unexpected. However, when pianists play the harmonically irregular chord in a known piece (i.e., with both auditory- and motor-predictions), it is unknown whether processing of this chord would differ from listening to the same chord in a known piece. The present ERP study aimed at investigating the impact of top-down auditory- and motor-predictions on neural correlates of the bottom-up perceptual processing of music-syntactic information.

One of the primary consequences of a music-syntactic system is that listeners and performers familiar with this system automatically develop top-down predictions about upcoming notes or chords. The generation of predictions is conceived as a form of mental belief about ‘what’ and ‘when’ musical events or structures are likely to occur in an unfolding piece, based on a priori internalized (often implicit) musical knowledge, actual memory for pieces of music, or current information (Rohrmeier and Koelsch, 2012; Janata, 1995; Limb, 2006; Schmuckler, 1997).

Top-down predictions play an important role in processing music (Bigand et al., 1996). Two different kinds of top-down predictions can be differentiated, usually referred to as “veridical expectations” and “schematic expectations” (Krumhansl et al., 1999). These types of predictions have been proposed to operate in a highly functionally modular fashion in terms of memory system (Huron, 2006), cognitive

---

[1] This part is in preparation as
Guo S, Maidhof C, Koelsch S. Knowing that my prediction is wrong: The impact of auditory- and motor- predictions on neural correlates of syntactic processing in music.

processes (Justus and Bharucha, 2001; Tillmann and Bigand, 2004; 2010), and neural correlates (Miranda and Ullman,2007). According to Bharucha (1987), veridical expectations are generated either by repeated listening (i.e., the activation of memory traces for particular pieces), or by context-specific prior knowledge of the event structures (i.e., the anticipation of a specific upcoming note, chord etc.; veridical knowledge) (Bharucha, 1987; Huron, 2006; Miles et al., 2016). Veridical expectations are usually explicit (e.g., the memory representation of the song *Happy Birthday*), but can also be implicit (e.g., when playing a piece of music by heart, the motor sequences are often executed automatically, without the player being able to tell explicitly which notes will follow) (Krumhansl, 2000). Regarding schematic expectations, listeners can generalize common musical rules and store knowledge of general musical patterns determining how single musical events are combined into musical sequences of a cultural system (e.g., Western tonal system) in long-term memory (i.e. schematic knowledge; Bharucha, 1987; Justus and Bharucha, 2001; Tillmann and Bigand, 2004; 2010; Huron, 2006). Schematic knowledge is represented in the form of schemata, e.g. statistical properties of music-syntactic regularities, which shape our perception of music and our predictions (i.e. "schematic expectations") (Krumhansl et al., 1999; Tillmann and Bigand, 2004; Maes, 2016). Schematic expectations can either be implicit (e.g., in non-musicians) or explicit (e.g., in musicians) (Krumhansl, 2000). Studies have repeatedly shown that the formation of schematic expectations is passive and can probably be acquired over many years (Bigand et al., 2000; Koelsch et al., 2002b).

Previous neurophysiological studies have investigated ERP responses to violations of music-syntactic regularities (e.g., Patel, 1998; Koelsch et al., 2000; Regnault et al., 2001; Besson and Faïta, 1995; Paller et al., 1992). Comparing ERPs elicited by harmonically congruous and incongruous chords, the ERAN is elicited, taken as an electrophysiological reflection of processing of music-syntactic violations. It is a negative deflection typically observed at around 150 to 280 ms and peaks at about 180 ms over anterior regions with a slight right-hemispheric lateralization (Koelsch et al., 2000; Garza Villarreal et al., 2011; Kalda and Minati, 2012; Koelsch et al., 2013;

Sammler et al., 2013a; Ellison et al., 2015; Zendel et al., 2015; Jaśkiewicz et al., 2016; for a review, see Koelsch, 2012). When the unexpected event is highly salient, or participants are asked to ignore the auditory stimuli, the ERAN is followed by a fronto-centrally P3a around 300 ms (Koelsch et al., 2000; Koelsch et al., 2007; Miranda and Ullman, 2007; Carrión and Bly, 2008; Brattico et al., 2013). Within the P300 complex, the P3a is usually followed by a more posterior distributed P3b within a range of 300 to 800 ms (often just 'P300'; Koelsch et al., 2000; Carrión and Bly, 2008; Ellison et al., 2015), which is commonly elicited by infrequently occurring target events that are attended and task-relevant (Friedman et al., 2001; Katayama and Polich, 1998; Polich, 1988). The ERAN is usually followed by a late bilateral frontal negativity N5 within a range of 500 to 550 ms (Koelsch et al., 2000; Loui et al., 2005). Several previous studies have provided evidence that rule-based syntactic knowledge stored in long-term memory is remarkably stable and cannot easily be modified (Koelsch and Jentschke, 2008; Carrión and Bly, 2008; Guo and Koelsch, 2015; 2016). Koelsch and Jentschke (2008) found that the amplitude of the ERAN declined over the course of an experimental session of about two hours, but was still significant at the end of the experiment, indicating that cognitive representations of music-syntactic regularities cannot easily be modified (Koelsch and Jentschke, 2008). In our last two studies (Guo and Koelsch, 2015; 2016), participants were informed by a cue as to whether the following sequence ended on a regular or an irregular chord. Despite being able to predict the irregular sequence endings, these events elicited a clear ERAN, indicating that the bottom-up principle mechanisms of music-syntactic processing (as reflected in the ERAN) are surprisingly robust against top-down auditory predictions.

Much research has been traditionally concentrated on understanding the cognitive and neural mechanisms of musical syntactic processing in the auditory domain (e.g., Patel, 1998; Koelsch et al., 2000; Regnault et al., 2001; Brattico et al., 2013; Guo and Koelsch, 2015; Ellison et al., 2015; Zendel et al., 2015; Guo and Koelsch, 2016; Jaśkiewicz et al., 2016). Recently, a growing number of studies have explored the processing of musical regularities in the motor domain (Novembre and Keller, 2011;

Jebb and Pfordresher, 2016; Sammler et al., 2013b; Bianco et al., 2016a, 2016b). In a recent behavioral study, participants played either tonal or atonal short melodies from memory, with auditory feedback varying in tonal class. Results indicate that schematic tonal class knowledge was highly active in determining action-perception relations during music performance (Jebb and Pfordresher, 2016). Novembre and Keller (2011) explored the impact of syntactic knowledge on musical actions using an imitation paradigm. Results showed that imitation of syntactically incongruent chords was slower and had a higher number of errors than imitation of congruent chords, suggesting that syntactic rules determining the observation of musical actions evoke motor expectations that influence pianists' action execution (Novembre and Keller, 2011). In subsequent ERP studies using priming paradigms, expert pianists were asked to watch silent videos (Sammler et al., 2013b) or sequences of photos (Bianco et al., 2016a) of a model hand playing five-chord sequences, and to simultaneously imitate these chord sequences with one hand on a muted keyboard. The last chord was either syntactically congruent or incongruent with the preceding harmonic context. To induce different strengths of predictability, pianists were presented with two-chord sequences as well. Compared to congruent chords, imitation of incongruent chords was slower and had a higher number of errors (Novembre and Keller, 2011; Sammler et al., 2013b; Bianco et al., 2016a). A right anterior negativity was elicited in the study by Sammler et al. (2013b), which might be interpreted as an ERAN in the visuomotor modality, evoked by expecting the action outcome in the videos based on the hand trajectory towards the final chord (Sammler et al., 2013b). The finding of no ERAN in the study by Bianco et al. (2016a) was assumed to be due to lacking auditory input (i.e., detection of syntactic irregularities present in the auditory domain) and / or the way of presenting the model hand (i.e. perceptual continuity) (Bianco et al., 2016a). Thus, it remains an open question whether the elicitation of the ERAN requires auditory (acoustic) input or not. In addition, a right-lateralized posterior negativity (Sammler et al., 2013b) and a late centroparietal negativity (Bianco et al., 2016a) were elicited by incongruent chords, reflecting mechanisms of re-programming of preplanned dominant responses (i.e., playing congruent chords). From these studies, it might be argued that the motor system of skilled pianists makes use of long-term music-syntactic knowledge for expecting forthcoming actions during action observation and imitation ("embodied processing of musical syntax"; Sammler et al., 2013b; Bianco et al., 2016a). Following these two ERP studies, another fMRI study investigated predictions in musical actions (without auditory feedback) and how they relate to music perception. Imitating harmonic sequences using one hand on a glass-board were compared with listening to the same chord progressions, and results showed distinct sub-regions in the right rIFG interconnected with parietal and temporal areas for processing silent action and audio sequences separately (Bianco et al., 2016b).

Note that in the above-mentioned studies (Sammler et al., 2013b; Bianco et al., 2016a; 2016b; Novembre and Keller, 2011), pianists did not know the forthcoming chord until the chord performed by a model hand was shown (i.e. without veridical knowledge). They generated predictions for forthcoming actions based on schematic knowledge (i.e. internalized music-syntactic knowledge). Our last two studies (Guo and Koelsch, 2015; 2016) investigated that with veridical knowledge, the impact of auditory predictions on neural correlates of music-syntactic processing during music-listening. However, it is unknown whether music-syntactic processing (as reflected in the ERAN, and other, motor-related ERPs) would be modulated by, or interact with auditory- and motor-predictions based on veridical knowledge as they occur during music performance. In addition, it is noteworthy that previous studies directly compared the auditory-only and the motor-only representations of music-syntactic regularities except one fMRI study (Bianco et al., 2016b), however without auditory feedback (in order to control the influence of auditorily learned syntactic regularities) and with only one hand during playing (similarly for other studies: Novembre and Keller, 2011; Sammler et al., 2013b; Bianco et al., 2016a). Music performance in the real world requires precise timing between both hands and sensorimotor dynamics between motor activity and its auditory effects (i.e., the associations of action with perception). Musicians are both producers and perceivers of a musical piece (Zatorre et al., 2007). Therefore, our present study attempted to compare the impact of auditory-motor, motor-only, and auditory-only predictions on the processing of music-syntactic regularities. To investigate these issues, two versions of chord sequences in which the final chord was syntactically either regular or irregular were used. Professional pianists either played these sequences with auditory feedback (on a digital piano, “auditory-motor modality”), or played them without auditory feedback (on a muted piano, “motor-only modality”), or listened to them (without playing, “auditory-only modality”). Note that pianists had veridical knowledge (and action goals) because each chord sequence was cued.

We hypothesized that a) there would be processing costs for playing irregular chords in both the auditory-motor and the motor-only modalities, b) the ERAN would be elicited in the auditory-only modality (whether an ERAN would be elicited in the motor-only modality was an open question), and c) the ERAN would be modulated (smaller amplitude and shorter latency), or even absent in the auditory-motor modality due to auditory- and motor-predictions.

3.4.2 Methods

3.4.2.1 Participants

Eighteen professional pianists participated in the experiment (age-range: 21-32 years, *M* = 24.06 years, *SD* = 2.9 years, 10 males and 8 females). The mean age at which piano studies commenced was 5.94 years (range: 3-9 years, *SD* = 1.66 years). All pianists had, on average 18.11 years (range: 13-26 years, *SD* = 3.45 years) of continuous formal piano training (quantified by subtracting the age at which piano studies commenced from each pianist’s age at the time of the current experiment), and the average daily amount of piano practice was 4.53 h (*SD* = 1.28 h; estimated from a musical background questionnaire answered by pianists). Eleven pianists reported to have absolute pitch, the remaining 7 pianists reported to not have absolute pitch. All pianists were current or former students at Universität der Künste (Berlin, Germany) or Hochschule für Musik Hanns Eisler (Berlin, Germany), and none of the participants knew any details about the purpose of the study. They had normal hearing and no neurological or psychiatric disorder (according to self-report). Pianists gave written informed consent prior to participation in this study. The study was approved by the ethics committee of the Psychology department of the Freie Universität Berlin and was performed in accordance with the Declaration of Helsinki.

3.4.2.2 Stimuli and apparatus

The stimuli were four-part piano chord sequences that had been used in previous studies (Jentschke and Koelsch, 2009; Jentschke et al., 2008; Koelsch and Jentschke, 2008; Koelsch et al., 2007). There were two versions of sequences, each version consisting of five chords. Both versions began with a tonic chord, followed by a subdominant, a supertonic and a dominant (see Figure 10A). In the regular version, the final chord was a tonic chord, whereas in the irregular version, the final chord was a supertonic (see Figure 10B). Presentation time of each chord was 600 ms, except for the final chord which lasted 1200 ms. Thus, each chord sequence had a duration of 3.6 s. Sequences were transposed to five major keys (C, D, F, G and Bb), resulting in 10 (2 × 5) different sequences. Note that syntactic irregularity did not co-occur with physical deviance (Koelsch et al., 2007), and syntactically irregular chords (supertonics) were even more congruent with the information stored in the echoic memory than regular chords (tonics). Therefore, any mismatch response evoked by irregular chords (supertonics) compared to regular chords (tonics) could not simply be due to the processing of deviant pitches.

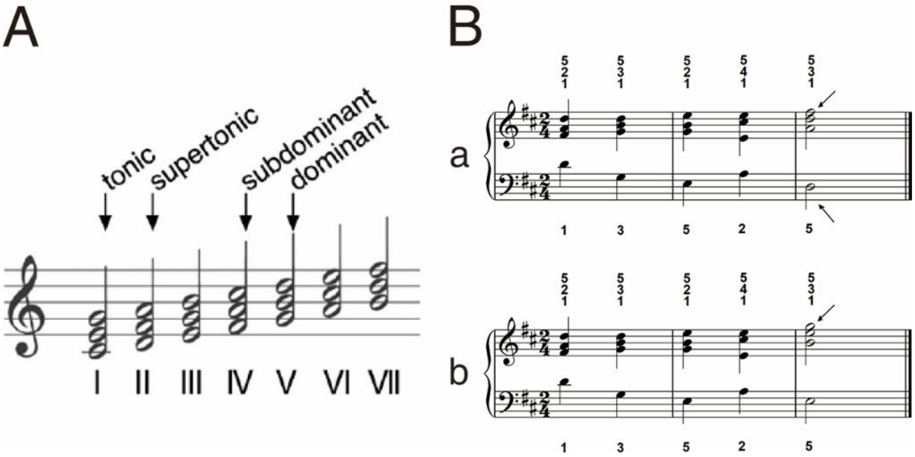

Figure 2. Study 3 - Illustrations of chord functions and examples of experimental stimuli. The left panel (**A**) illustrates chord functions (in C-major). In tonal music, the harmonies built on the steps of a scale are denoted as chord functions. For example, the chord built on the first scale tone is denoted as the tonic, the chord on the second scale tone as the supertonic, the chord on the fourth scale tone as the subdominant, and the chord on the fifth scale tone as the dominant. The right panel (**B**) illustrates the experimental stimuli (in D-major, note that in the experiment sequences from different keys were used). Both sequences began with a tonic chord, followed by a subdominant, a supertonic and a dominant. Sequence (a) ends on a tonic chord (regular version), and sequence (b) on a supertonic chord (irregular version). Pianists were instructed to play chord sequences bimanually: With their right hand they played the chord formed by three tones of the treble stave, and with their left hand they played one tone of the bass stave. The numerals above and below the notes indicated the fingerings (1: thumb, 2: index finger, 3: middle finger, 4: ring finger, 5: little finger). Sequences were used in the auditory-motor (A-M), auditory-only (A-only), and the motor-only (M-only) blocks.

3.4.2.3 Procedure

This study consisted of six phases: pre-experiment, training, test, practice, EEG experiment and post-experiment (see Figure 11A for an illustration).

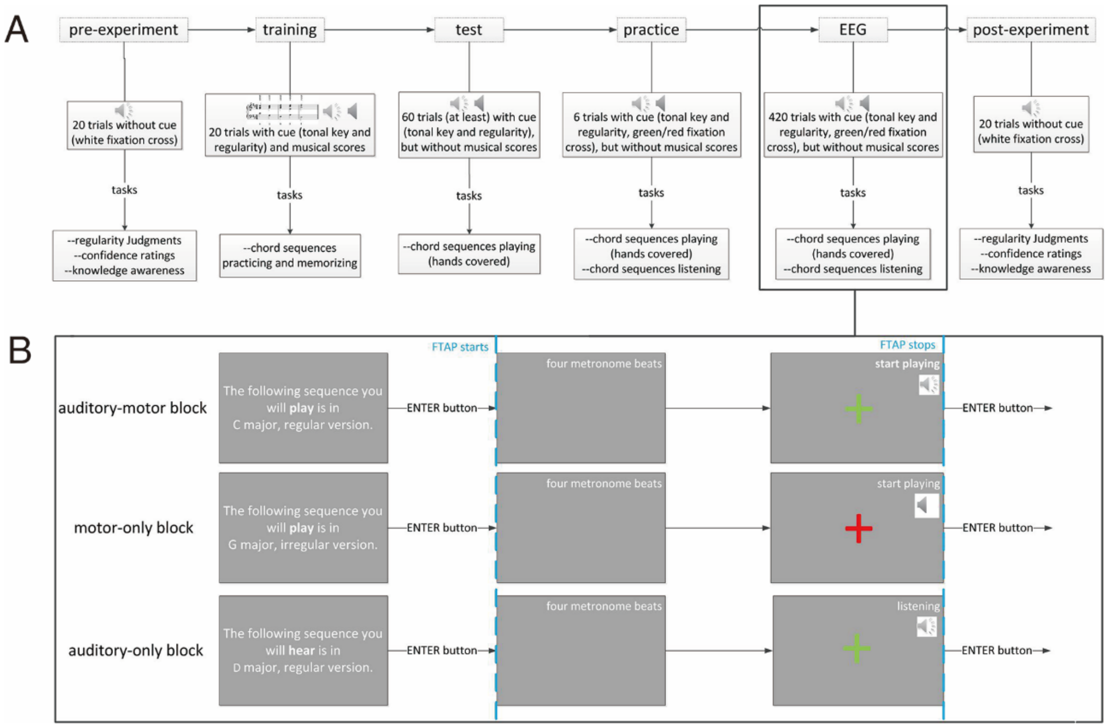

Figure 3. Study 3 - Procedure of the study (A) and the EEG experiment (B). The upper panel (**A**) illustrates the procedure of the study: There were six phases, which consisted of 20 trials of pre-experiment, 20 trials of training, 60 trials of test, 6 trials of practice, 420 trials of EEG experiment, and 20 trials of post-experiment. The lower panel (**B**) illustrates procedure of the EEG experiment: Each block was repeated twice, resulting in 6 blocks. Each block had a duration of about 12 min, and featured the two versions of chord sequences (see Figure 10). In the auditory-motor (A-M) block, pianists were informed of the tonal key and version of the sequence ending (regular or irregular). To remind pianists of the correct tempo, four metronome beats at 100 beats per minute (bpm) were presented with a blank screen. The metronome beats were followed by a green or red fixation cross, which cued pianists to start playing a sequence with a regular (green cross) or irregular (red cross) ending. Pianists played the corresponding sequence while hearing auditory feedback of their performances. In the motor-only (M-only) block, all the procedure was the same as in the A-M block, except that pianists could not hear auditory feedback of their performances. In the auditory-only (A-only) block, the procedure was similar as in the A-M block. Instead of playing, pianists listened to sequences with a regular (green cross) or irregular (red cross) ending. The FTAP software for processing the musical data was started internally when metronome beats were presented, and stopped internally when pianists pressed a button after they finished playing in the A-M or the M-only block, or after the presented sequence ended in the A-only block.

Pianists first completed a musical background questionnaire, after which the pre-experiment with 20 trials was carried out (see Figure 11A for an illustration). The purpose of the pre-experiment was to assess pianists' capability to differentiate between regular and irregular sequence endings and serve as initial baseline. In each trial, a sequence with regular or irregular ending was presented equiprobably, while pianists were instructed to look at a white fixation cross on a gray screen (i.e., pianists did not know whether they were presented with a regular or irregular sequence). After the presentation of each sequence, they were asked to indicate whether the final chord of the sequence sounded regular or irregular by pressing one of two response buttons as fast as possible ('R' button for the regular version, and 'I' button for the irregular version). The reason to stress the RTs was due to our last two studies which suggest that (amateur) musicians had no difficulty in differentiating between regular and irregular versions (Guo and Koelsch, 2015; 2016). There was no feedback on their regularity judgments. In order to determine whether pianists consciously knew that their judgments were correct, and what these judgments were based on (Dienes and Scott, 2005), confidence ratings and pianists' awareness of their knowledge guiding regularity judgments were collected. Using a Likert five-point scale (1 corresponding to "not confident at all" and 5 corresponding to "very confident"), pianists rated how confident they were for each regularity judgment and what their judgments were based on ((i) guessing, (ii) intuition, or (iii) rule knowledge). No time limit was imposed for the confidence ratings and pianists' awareness of their knowledge. Duration of the pre-experiment was around 4 min.

Subsequently, a training phase was delivered to pianists to rehearse and memorize scores of the chord sequences (see Figure 11A for an illustration). To remind pianists of the correct tempo, each trial began with four metronome beats at 100 beats per minute (bpm). Then, a score of one chord sequence appeared in the middle of the screen which was located behind the digital piano. The tonal key (e.g., D major) and the regularity of sequence ending (e.g., irregular, see also Figure 10 and *Stimuli* in Study 3) were shown above the score. To control for general effects of hand movements, all pianists had to produce the sequences with the same fingerings (chosen by two experienced piano teachers), which were indicated by the numerals above and below the notes which were identical for regular and irregular chord sequences. Pianists were instructed to play each chord sequence on a Yamaha digital piano (Clavinova CLP 130) bimanually: The right hand played the chord formed by three tones of the treble stave, and the left hand played one tone of the bass stave (see Figure 10B). They were required to play all chord sequences: (a) using assigned fingerings, (b) in the instructed tempo (guided by the metronome beats), (c) with same intensity, (d) without emotional expression, (e) without looking at their hands (after the first few practice trails), (f) without looking at the musical scores (after the first few practice trials). Moreover (g), it was emphasized that they should play the irregular sequences exactly in the same way as the regular sequences in terms of duration and velocity. They listened to the metronome beats and auditory feedback of their performances via studio headphones (Philips SHP 2500) plugged into a MR-Rack synthesizer (ENSONIQ, USA) at comfortable listening levels (approximately 60 dB when played with medium velocity of the key-presses, as instructed). All recorded chord sequences had the standard MIDI piano timbre generated by the synthesizer. There were 20 trials in the training: 2 (regularity: regular, irregular) × 5 (tonal key: C, D, F, G and B*b*) × 2 (auditory feedback: with, without). Once pianists could play all the sequences following the above-mentioned instructions (see (a)-(g)), the test started.

In the test phase, pianists were first informed of the tonal key and regularity of the sequence ending (e.g., regular sequence F major; see Figure 11A for an illustration). Then they were asked to play the corresponding sequence by heart (i.e., without musical scores) using assigned fingerings in the instructed tempo. Their hands and the piano keyboard were covered with a board to exclude a possible influence of visual perception of different hand shapes, and to decrease the likelihood of eye artifacts caused by visually tracking finger movements in the later EEG experiment. After three consecutive error-free performances of a sequence, the experimenter started the next sequence. That is, each pianist needed to play at least 60 error-free sequences in the test: 2 (regularity) × 5 (tonal key) × 2 (auditory feedback) × 3 (repetitions). All pianists passed the test and were admitted to the EEG experiment.

Afterwards, a practice phase with six trials was delivered to familiarize pianists with the task in the EEG experiment (see Figure 11A for an illustration, see EEG experiment below for details), after which pianists were informed to warm up on the piano keyboard until they felt ready to start the EEG experiment.

The EEG experiment was defined as a within-subjects design by modality (auditory-motor, motor-only, and auditory-only) × regularity (regular, irregular) (see Figure 11A for an illustration). For the sake of better readability, we refer to auditory-motor block as A-M block, motor-only block as M-only block, and auditory-only block as A-only block. In the A-M block (see Figure 11B for an illustration), pianists were informed of the tonal key and regularity of the sequence ending. Then four metronome beats at 100 bpm were presented. The metronome beats were followed by a green or red fixation cross, which cued pianists to start playing a sequence with a regular (green cross) or irregular ending (red cross). Pianists' hands and the piano keyboard were covered with a board. The task for pianists was to play chord sequences from memory following the same instructions (see (a)-(g)) as in the training. In addition, pianists were instructed to (h) play as accurately as possible from beginning to end without stopping to correct any error, (i) keep their eyes open and focus on the fixation cross, (j) relax in order to reduce tension in the neck and the shoulder muscles during playing, (k) refrain from eye-blinking and eye-movements as much as possible, and (l) move as minimally as possible to avoid muscle artefacts. Pianists were asked to press a button after they finished playing. In the M-only block (see Figure 11B for an illustration), both procedure and instructions were the same as in the A-M block, except that pianists could not hear the auditory feedback of their performance. In the A-only block (see Figure 11B for an illustration), procedure and instructions were similar as in the A-M block. Instead of playing, pianists were instructed to listen attentively to the following chord sequence which was cued by a green (regular ending) or red (irregular ending) cross, and not to move their fingers while listening. They were not asked to detect regular or irregular endings, in order to avoid that the ERAN elicited by irregular chords would be overlapped with N2b and P3 potentials. Pianists were asked to press a button after each sequence. The stimuli in the A-only block were earlier pre-recorded chord sequences played by a previous pianist with the same order of keys and regularity of the sequence ending. This was done to avoid confounding ordering effects, or any other auditory effects, that could possibly emerge from pianists always listening to their own (previous) performance. The first pianist listened to chord sequences played by an invited pianist in a pre-experiment, which was used as the stimuli in the A-only block of practice phase for all pianists as well. Note that, thus, the auditory input of A-M and A-only blocks was nearly identical across our sample. Finger movements in all blocks were recorded with a video camera (Praktica, DVC 5.4 HDMI) placed behind pianists (between the covered board and the piano keyboard).

In the entire EEG experiment, each sequence was presented 42 times, amounting to 420 trials in total: 2 (regularity) × 5 (tonal key) × 42 (repetitions), and 140 trials in each block (A-M, M-only, and A-only). Regular and irregular sequences occurred equiprobably and were pseudorandomly intermixed, with the constraints that no more than three sequences of the same version (regular, irregular) followed each other, and that each sequence was presented in a tonal key that differed from the key of the preceding sequence. The order of blocks was counterbalanced across pianists, and they were notified of the beginning of a new block (and thus task). To ensure that pianists attended to the cues (indicating tonal key and regularity), and that they would thus have veridical knowledge of what to play or listen to, all trials were self-paced (i.e., pianists determined when to proceed with the next trial by pressing a button). There were short breaks in the middle and after each block, and pianists could continue the experiment whenever they were ready by pressing a button. Each block had a duration of around 24 min, amounting to a duration of the entire EEG experiment of approximately 90 min (including breaks).

The EEG experiment was followed by the post-experiment with 20 trials, which was identical to the pre-experiment (see Figure 11A for an illustration). After the post-experiment, pianists rated to what extent they actively anticipated the sounds and finger movements in the A-M, M-only, and A-only blocks separately (9-point scale, 1 corresponding to “not at all” and 9 to “very much”).

3.4.2.4 Data acquisition

3.4.2.4.1 Behavioral data

The behavioral data of pre- and post- experiments were recorded with Presentation software (Version 17.2, Neurobehavioral Systems, Inc.).

3.4.2.4.2 Musical data

Musical data were processed in MIDI format with a modified version of FTAP software (Finney, 2001a; 2001b; Maidhof et al., 2014). To synchronize MIDI and EEG data, this program sent synchronization signals concurrently with every sixth key-press to the EEG acquisition computer (see TTL pulses with key-presses in Figure 12). The MIDI information (including pitch event, key-press onset, duration, and velocity) was used for offline performance analyses, so that triggers for all key-presses could be reconstructed offline for the EEG data evaluation. The digital piano was connected to a Linux-computer (see PC2 in Figure 12) via a 2 x 2 USB-MIDI-Interface (E-MU). A program written in Perl controlled the experiment procedure (see box of PC2 in Figure 12).

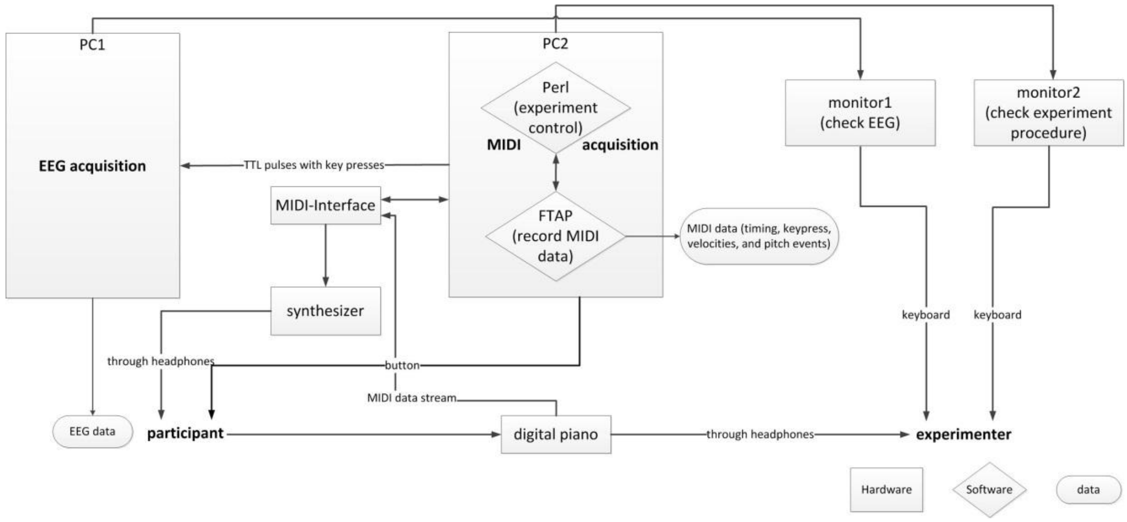

Figure 4. Study 3 - Schematic illustration of the setup combining MIDI and EEG recordings.

3.4.2.4.3 EEG data

The EEG was recorded with a BrainAmp MR plus amplifier system (Brain Products

Inc., Gilching / Germany) from 59 electrodes (Fp1, Fpz, Fp2, AF7, AF3, AF4, AF8, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FT7, FC5, FC3, FC1, FC2, FC4, FC6, FT8, T7, C5, C3, C1, Cz, C2, C4, C6, T8, TP7, CP5, CP3, CP1, CPz, CP2, CP4, CP6, TP8, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO3, POz, PO4, PO8, O1, Oz, O2; placed according to the extended international 10-20 system), referenced to the left mastoid (M1). Four electrodes were used for recording the vertical and the horizontal EOGs. The ground electrode was located on the sternum. The EEG was digitized at a rate of 500 Hz (low and high cut off were D.C. and 100 Hz, respectively) and the impedances were kept below 5 kΩ.

3.4.2.5 Data analysis

3.4.2.5.1 Behavioral and musical data

The following behavioral indices were evaluated: a) time windows of synchronous keypresses within one chord (onset of last keypress minus onset of first keypress), b) IOIs between penultimate (dominant chord preceding the final chord) and final chords, and IOIs between the third and the penultimate (the fourth) chords (note that the mean IOIs between successive chords provided an indicator of how well pianists adjusted to the given tempo, c) duration of final and penultimate (the dominant chord preceding the final chord) chords (onset of last key-release minus onset of first keypress), and d) keypress velocities (i.e. the speed with which pianists pressed a key) of the final and the penultimate chords.

Errors were identified offline by using the MIDI Toolbox for MATLAB (Eerola and Toiviainen, 2004) and its extension for matching a music performance to its corresponding notation (Large, 1993), and erroneous trials were excluded from analysis. Errors were defined in terms of pitch (wrong note, omission note, insertion / extra note, corrected note), synchronization of the key-presses (time window of synchronous key-presses exceeded 70 ms), chord timing (IOI between two consecutive chords smaller than 300 ms or greater than 900 ms), and fingering (not using assigned fingerings).

Used fingerings were analyzed based on the video recordings. In addition, the videos allowed us to ensure that the motor transitional or spatial trajectories performed by both hands moving from the penultimate to the final chords between regular and irregular versions were basically the same.

3.4.2.5.2 EEG data

Data were analyzed offline using EEGLAB V12.0.2 (Delorme and Makeig, 2004). Raw data were filtered with a 49-51 Hz band-stop filter with FIR (filter order: 2750 points) to eliminate line noise, and a 0.25 high-pass filter (FIR, filter order: 13750 points) to remove slow waves (such as electrode saturation or drifts). An ICA was carried out, and components representing artifacts (eye blinks, eye movements, and muscle activities) were removed. Afterwards, data were filtered with a 25 Hz low-pass filter (FIR, filter order: 550 points) to remove remaining high-frequency noise (such as muscle activity that was not removed using the ICA). Subsequently, data were epoched relative to the mean onset time of the final chords (mean onset time refers to the average of the onset times of all four notes of one chord; for example, when onset times of four notes of one chord were 100 ms, 120 ms, 160 ms and 180 ms, the mean onset time of this chord would be 140 ms). To remove further possible artifacts, sampling points were rejected whenever the SD of a 100-ms or 700-ms gliding window exceeded 25 μV at any EEG electrode. Then, data were re-referenced to the arithmetical mean of left and right mastoid leads. Finally, non-rejected epochs were averaged from -100 to 1000 ms relative to the mean onset time of the final chord with a -100 to 0 ms baseline. An average of 57 artifact-free trials was included for each pianist in each condition: (56 ± 11) trials in the A-M modality, (59 ± 7) trials in the M-only modality, and (58 ± 9) trials in the A-only modality (*M* ± *SD*).

Spatiotemporal ERP analyses (microstate segmentation). To differentiate these effects, spatiotemporal ERP analyses (i.e. topographic analyses) were performed on the grand-averaged ERPs (Michel et al., 2004; 1999a; 1999b) using the software Cartool (Functional Brain Mapping Laboratory and CIBM, Geneva, Switzerland; http://brainmapping.unige.ch/Cartool.htm) (Brunet et al., 2011). Topographic analyses aim to objectively identify the successive stable microstate maps (i.e. map topographies) and the time intervals where they rapidly change from one stable configuration into another (i.e. segment borders) (Arzy et al., 2007; Michel et al., 2001; Murray et al., 2008), following the notion that microstate maps are not randomly distributed across time, but remain stable for successive processing states before an abrupt qualitative change in the electric field configuration occurs (Lehmann et al., 1987; Michel et al., 1999a; 1999b). Each map presumably reflects different functional stages of the brain, which are referred to as "functional microstates" (Brandeis and Lehmann, 1986a; Pascual-Marqui et al., 1995).

First, topographic pattern analysis was conducted, based on an AAHC method (Britz et al., 2009; Murray et al., 2008), which was used to identify the most dominant map topographies that represent the microstates (Michel et al., 2001; Pascual-Marqui et al., 1995). The optimal number of microstate maps was determined by means of a cross-validation criterion (Pascual-Marqui et al., 1995) and a Krzanowski-Lai criterion (Brunet et al., 2011; Michel et al., 2004; Murray et al., 2008; Pascual-Marqui et al., 1995). Map durations of less than 20 ms are considered physiologically implausible, and we set a temporal criterion of at least 10 consecutive time frames (≥20 ms) accordingly. This approach allows summarizing ERP data into a limited number of electrocortical map configurations and identifying time periods during which different experimental conditions evoke different maps (Michel et al., 2001).

Next, we determined whether the microstate maps identified in the grand-averaged ERPs could be identified in individual subjects by means of a fitting procedure applied to individual data. This procedure consisted of calculating for each subject strength-independent spatial correlation coefficients between the microstate maps identified by the topographic pattern analysis and the ERPs of each subject in corresponding time intervals (for a review, see Michel et al., 2001). Statistical analyses were conducted to assess how frequently each map was identified in each subject. A significant statistical difference indicates that one condition was significantly better explained by one given map than another. This map was thus specific to this condition. The GEV each map explained was calculated as well. The

GEV equals the sum of the explained variances over the time windows of interest, weighted by the strength of the map at each moment in time. Statistical analyses were also performed on the FOnset and the TFBCorr for a given map. These indexes were used to reveal differences in timing, which were topographic onset and latency. Moreover, we conducted analyses on the TFMaxGFP over the segment fitted by a given map, which represented the topographic peak latency (Murray et al., 2008). We used above-mentioned parameters (FOnset, TFBCorr, and TFMaxGFP) to determine the onset and offset of time windows for further ERP waveform analyses.

In contrast to conventional waveform analyses, spatiotemporal ERP analyses rely on the distribution of the electric potential at the scalp (i.e. the topographic map), which is recorded from a multichannel electrode array. Moreover, spatiotemporal ERP analyses give non-ambiguous and reference-independent information about changes in the underlying generators in time and between conditions (Murray et al., 2008). Differences observed in amplitudes can be a consequence of a modulation in the strength of the electric field, of a scalp topographic change of the electric field (revealing distinguishable brain generators), or of latency shifts in brain processes.

ERP waveform analyses. For statistical analyses, mean amplitude values of the ERAN, P3a, P3b, N5, and LTN (late tonic negativity) were computed for different ROIs: right anterior (F4, F6, FC4, FC6) and right central (C4, C6, CP4, CP6) for the ERAN and LTN, frontal (FPz, AF3, AF4, Fz) for the P3a, posterior (Pz, PO3, PO4, POz) for the P3b, and right anterior (F4, F6, FC4, FC6) for the N5. The electrodes for the ROIs were chosen according to microstate maps of irregular chords from the microstate segmentation. These ROIs were consistent with both isopotential maps of the ERP components analyzed (ERAN, P3a, P3b, N5, and LTN) and with previous studies.

Global ANOVAs with factors modality (A-only, A-M, and M-only), regularity, and ROIs were conducted for the amplitudes of the ERAN, P3a, P3b, and LTN. Paired-sample *t*-test was conducted in the A-only modality for the amplitude of the N5. Greenhouse-Geisser corrections were performed for analyses when necessary. Whenever an interaction involving modality and / or regularity was found, follow-up analyses were carried out by splitting up the factorial model. Only significant results in linear correlation analyses were reported. Cohen's *d* (for *t*-tests) and partial eta squares (partial η^2, for ANOVAs) were used as measures of effect sizes. Cohen's *d* was calculated using the formula provided in Cohen (1988) (d-value of 0.2 corresponds to small effect-size, 0.5 to medium effect-size, and 0.8 to large effect-size). Partial η^2 was calculated in SPSS (partial η^2 value of 0.1 is defined as small effect-size, 0.3 as medium effect-size, and 0.5 as large effect-size; Bakeman, 2005).

Peak latencies and time windows of those effects (the ERAN, P300(P3b), and N5) for which we had hypotheses in the A-only modality were determined by the appropriate parameters (TFMaxGFP, Fonset, TFBCorr, and TFMaxGFP) of the spatiotemporal ERP analyses. The same time window of the ERAN was adopted in the A-M and M-only modalities as in the A-only modality in order to make comparisons between these three modalities. Peak latency and time window of the LTN in the A-M and the M-only modalities were determined using the grand-average difference-waveforms of ERPs (regular subtracted from irregular chords). These time windows were 120 to 240 ms (ERAN), 240 to 310 ms (P3a), 310 to 550 ms (P3b), 550 to 850 ms (N5), and 250 to 850 ms (LTN). For presentation purposes, averaged data were filtered after statistical evaluation (10 Hz low-pass, 41 points, FIR).

3.4.3 Results

3.4.3.1 Behavioral results

3.4.3.1.1 Pre- and post- experiments

The mean percentages of correct responses of regularity judgments in the pre- and the post- experiments were 96.89% (*SD* = 7.9%) and 100% (*SD* = 0%) ($p = 0.11$, paired-sample *t*-test). The regularity judgments were faster in the post-experiment ($M = 1.25$ s, $SD = 0.53$ s) than in the pre-experiment ($M = 1.86$ s, $SD = 1.12$ s) ($t = 3.05$, $p = .007$, Cohen's $d = 0.71$, paired-sample *t*-test).

All pianists were very confident (5 points) in the confidence rating of regularity judgments in both pre- and post- experiments. None of the judgments was based on guessing. Pianists reported that 77.8% (pre-experiment) and 72.2% (post-experiment) of their judgments were based on rule knowledge, and 22.2% (pre-experiment) and 27.8% (post-experiment) of their judgments were based on intuition. These results confirm that their knowledge of music-syntactic regularities was predominantly explicit and consciously available.

3.4.3.1.2 Training and test

The duration of achieving error-free performance with the chord sequences in the training was highly variable across pianists (*M* = 11.33 min, *SD* = 5.5 min, *ω* = 17.48 min). Due to the fact that the auditory-motor (A-M) modality was taken before the motor-only (M-only) modality, the time used in the A-M modality (*M* = 8.42 min, *SD* = 4.85 min, *ω* = 15.71 min) was longer than in the M-only modality (*M* = 2.91 min, *SD* = 1.14 min, *ω* = 3.6 min) to achieve error-free performance. A positive correlation between the age at which piano studies commenced and the time used in the training session to achieve error-free performance was found (ρ(16) = 0.63, p = .005), suggesting that the younger pianists started to learn playing piano, the short time they needed to reach error-free performance in the training session. The mean time to pass (i.e., achieve error-free performance) in the test was 7.48 min (*SD* = 2.03 min, *ω* = 7.51 min). As in the training, the A-M modality (*M* = 4.4 min, *SD* = 1.36 min, *ω* = 5.45 min) was taken before the M-only modality (*M* = 3.08 min, *SD* = 0.9 min, *ω* = 3.43 min).

3.4.3.1.3 EEG experiment

The overall note accuracy (i.e., all notes of a sequence) was 94.54 % (*SD* = 4.21 %). Pianists committed more errors when playing irregular sequences (mean accuracy: 93.41 %, *SD* = 5.18 %) compared to regular sequences (mean accuracy: 95.67 %, *SD* = 3.69 %) (*t* = 3.06, *p* = .007, Cohen's *d* = 0.50), independent of whether they played with or without auditory feedback. In the A-M modality, the strength of anticipating the sounds was significantly correlated with the note accuracy (ρ(16) = 0.59, *p* = .009), indicating that the stronger pianists anticipated the sounds, the higher note accuracy they played chord sequences.

The four notes of each chord were pressed within a time window ranging from 9.4 ms to 24.6 ms across pianists (*M* = 16.1 ms, *SD* = 3.7 ms). An ANOVA with factors modality (A-M and M-only) and regularity (regular and irregular) did not indicate any main effects or interaction.

Duration of the final chord in the A-M modality (*M* = 1.18 s, *SD* = 0.24 s) was longer than the duration of the final chord in the M-only modality (*M* = 1.10 s, *SD* = 0.24 s) (*t* = 2.06, *p* = 0.055, Cohen's *d* = 0.35, paired-sample *t*-test). Notably, in the M-only modality, the duration of the penultimate chord in regular sequence (*M* = 0.38 s, *SD* = 0.13 s) was longer than in irregular sequence (*M* = 0.35 s, *SD* = 0.11 s) (*t* = 2.97, *p* = .009, Cohen's *d* = 0.27, paired-sample *t*-test), as confirmed by an ANOVA with factors modality and regularity revealing an interaction between modality and regularity (*F*(1, 17) = 10.28, *p* = .005, partial η^2 = 0.38).

Regarding the IOIs between the penultimate (the fourth chord in the sequence) and all final chords (the fifth chord in the sequence, both regular and irregular), and between the third and the penultimate (the fourth) chords, pianists showed performance slowing in the A-M modality (*M* = 604.6 ms, *SD* = 24.2 ms) compared to the M-only modality (*M* = 591.5 ms, *SD* = 26.2 ms) (*t* = 2.89, *p* = .01, Cohen's *d* = 0.52), as confirmed by an ANOVA with factors modality, regularity and chord position (final and penultimate) showing a main effect of modality (F(1, 17) = 8.37, *p* = .01, partial η^2 = 0.33). Irregular final chord showed a greater performance slowing (*M* = 621.6 ms, *SD* = 24.5 ms) than regular final chord (*M* = 613.3 ms, *SD* = 24.1 ms) (*t* = -5.82, *p* < .001, Cohen's *d* = 0.34), as confirmed by an ANOVA with factors modality, regularity and chord position showing an interaction between regularity and chord position (F(1, 17) = 22.72, *p* < .001, partial η^2 = 0.57).

Although no interaction between modality, regularity, and chord position was found (*p* = 0.12), further paired-sample *t*-tests were computed to compare IOIs between regular and irregular final chords, conducted separately for A-M and M-only modalities, and between the final chords in A-M and M-only modalities, conducted separately for regular and irregular sequences. Similarly, paired-sample *t*-tests were computed for the penultimate chord as the final chord (see Table 6 for detailed results).

Table 1. Study 3 - Summary of inter-onset intervals (IOIs) and results of paired-sample *t*-tests for final and penultimate chords in regular and irregular chord sequences in the auditory-motor (A-M) and motor-only (M-only) modalities for all pianists (in ms). Bold font indicates significant results.

modalities	all pianists (n = 18)				
	M	*SD*	*t*	*p*	Cohen's *d*
Regular_final_A-M	619.06	25.17	-4.33	**< .001**	0.33
Irregular_final_A-M	627.24	23.84			
Regular_fianl_M-only	607.47	27.54	-5.51	**< .0001**	0.29
Irregular_fianl_M-only	615.88	29.74			
A-M_final_regular	619.16	25.17	2.27	**.036**	0.44
M-only_final_regular	607.47	27.54			
A-M_final_irregular	627.24	23.84	2.15	**.046**	0.42
M-only_final_irregular	615.88	29.74			
Regular_penultimate_A-M	584.84	28.34	-1.91	.073	0.09
Irregular_penultimate_A-M	587.18	25.21			
Regular_penultimate_M-only	571.30	25.60	0.095	.93	0.003
Irregular_penultimate_M-only	571.23	26.09			
A-M_penultimate_regular	584.84	28.34	2.89	**.01**	0.50
M-only_penultimate_regular	571.30	25.60			
A-M_penultimate_irregular	587.18	25.21	3.70	**.002**	0.62
M-only_penultimate_irregular	571.23	26.09			

The overall tempo was 102 bpm (min = 96 bpm, max = 110 bpm, *SD* = 4 bpm), suggesting that pianists could play chord sequences with very low temporal variability. The overall performed tempo was significantly faster than the instructed tempo (i.e., 100 bpm, *t* = 2.25, *p* = .038, one-sample *t*-test). Given the high note accuracy (94.54 %), these results suggest that pianists did not have any problems playing these chord sequences. The tempo in the M-only modality (*M* = 103 bpm, *SD* = 5 bpm) was faster than in the A-M modality (*M* = 101 bmp, *SD* = 4 bpm) (*t* = -2.59, *p* = .019, Cohen's *d* = 0.46, paired-sample *t*-test), and the tempo of irregular chord sequences (*M* = 101 bpm, *SD* = 4 bpm) was slower than regular chord sequences (*M* = 102 bpm, *SD* = 4 bpm) (*t* = 3.05, *p* = .007, Cohen's *d* = 0.14, paired-sample *t*-test).

There were no differences in keypress velocities between regular and irregular sequences, and no differences between modalities (an ANOVA with factors modality and regularity showed no main effects or interaction).

Analyses of the video recordings in the EEG experiment revealed that pianists did not move their fingers in the auditory-only (A-only) modality and that only few fingering errors were committed during playing with auditory feedback (final regular chords: M = 99.44%, SD = 1.71%; final irregular chords: M = 99.13%, SD = 2.09%), and without auditory feedback (final regular chords: M = 99.52%, SD = 1.21%; final irregular chords: M = 99.21%, SD = 2.41%). Fingering accuracies did not differ between regular and irregular sequences and between playing with and without auditory feedback, as confirmed by an ANOVA with factors modality and regularity showing no main effects or interaction.

3.4.3.1.4 Post-EEG debriefing

The strength of actively anticipating finger movements was strongest in the M-only modality and weakest in the A-only modality (see Table 7 for detailed rating points), and the strength of actively anticipating the sounds was stronger than anticipating finger movements in both A-only (t = 5.46, p < .001, Cohen's d = 1.31) and A-M (t = 2.40, p = .028, Cohen's d = 0.71) modalities, which were confirmed by an ANOVA with factors modality (A-only, A-M, and M-only) and anticipation content (sounds and finger movements), showing an interaction between factors modality and anticipation content ($F(2, 34) = 11.60$, $p < .001$, partial $\eta^2 = 0.41$; see Figure 13).

Table 2. Study 3 - Summary of rating points of anticipating the sounds and finger movements in the auditory-only (A-only), auditory-motor (A-M), and motor-only (M-only) modalities for all pianists (1 corresponding to “not anticipating at all” and 9 to “anticipating actively”) (means, with standard deviation in parentheses).

modality	all pianists (n = 18)	
	sounds	finger movements
A-only	7.39 (1.85)	4.06 (3.10)
A-M	7.50 (1.25)	5.94 (2.86)
M-only	7.94 (1.26)	7.61 (1.75)

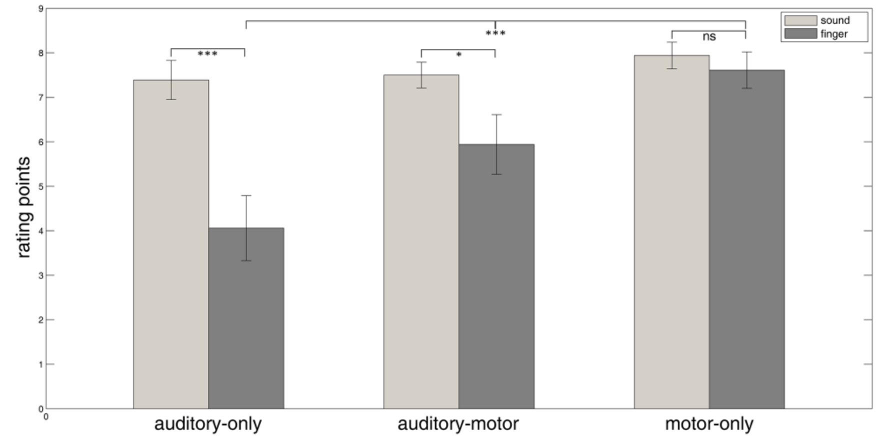

Figure 5. Study 3 - Histograms depict to what extent pianists anticipated the sounds (light gray bars) and finger movements (dark gray bars) in the A-only, A-M, and M-only modalities in the post-EEG debriefing. The data showed that the strength of anticipating finger movements between the auditory-only (A-only), auditory-motor (A-M), and motor-only (M-only) modalities were statistically significant. The strength between anticipating the sounds and finger movements were significant in the A-only and the A-M modalities. Error bars indicate standard error of means (SEMs). Asterisks denote statistical significant differences ($^{*}p \leq .05$, $^{**}p \leq .01$, $^{***}p \leq .001$).

In the A-only modality, the correlation between the age at which piano studies commenced and the strength of anticipating finger movements reached significance ($\rho(16) = -0.55$, $p = .018$), indicating that the earlier pianists started to learn piano, the stronger they anticipated finger movements. The more pianists practiced playing piano a day, the stronger they anticipated the sounds, as confirmed by the significant correlation between hours of piano-practice a day and the strength of anticipating the sounds ($\rho(16) = 0.49$, $p = .038$).

3.4.3.2 ERP results

3.4.4.2.1 Spatiotemporal ERP analyses (microstate segmentation)

Microstate segmentations were computed for auditory-only (i.e., pianists listened to earlier pre-recorded sequences played by a previous pianist), auditory-motor (i.e., pianists played chord sequences with auditory feedback), and motor-only (i.e., pianists played chord sequences without auditory feedback) modalities, to identify ERP components and their time-windows, and select electrodes included in the ROIs objectively.

The spatiotemporal segmentation procedure identified 14 distinct ERP microstate maps in the auditory-only modality, which explained 79% of the variance in regular and irregular conditions during the time period from -100 to 1000 ms. Figure 14A illustrates the global field power (GFP) traces for grand-average ERPs of regular and irregular conditions. This illustration shows that while some of the map segments exhibited varying onset times (i.e., shifts in terms of time of occurrence) and durations in the grand-average ERPs (e.g., maps 7 and 8), some maps only appeared in regular but not irregular condition (e.g., maps 10 to 12). The corresponding scalp voltage configurations labelled 1 to 14 can be seen in Figure 14B, as a function of the order of their appearance in the grand-average ERPs.

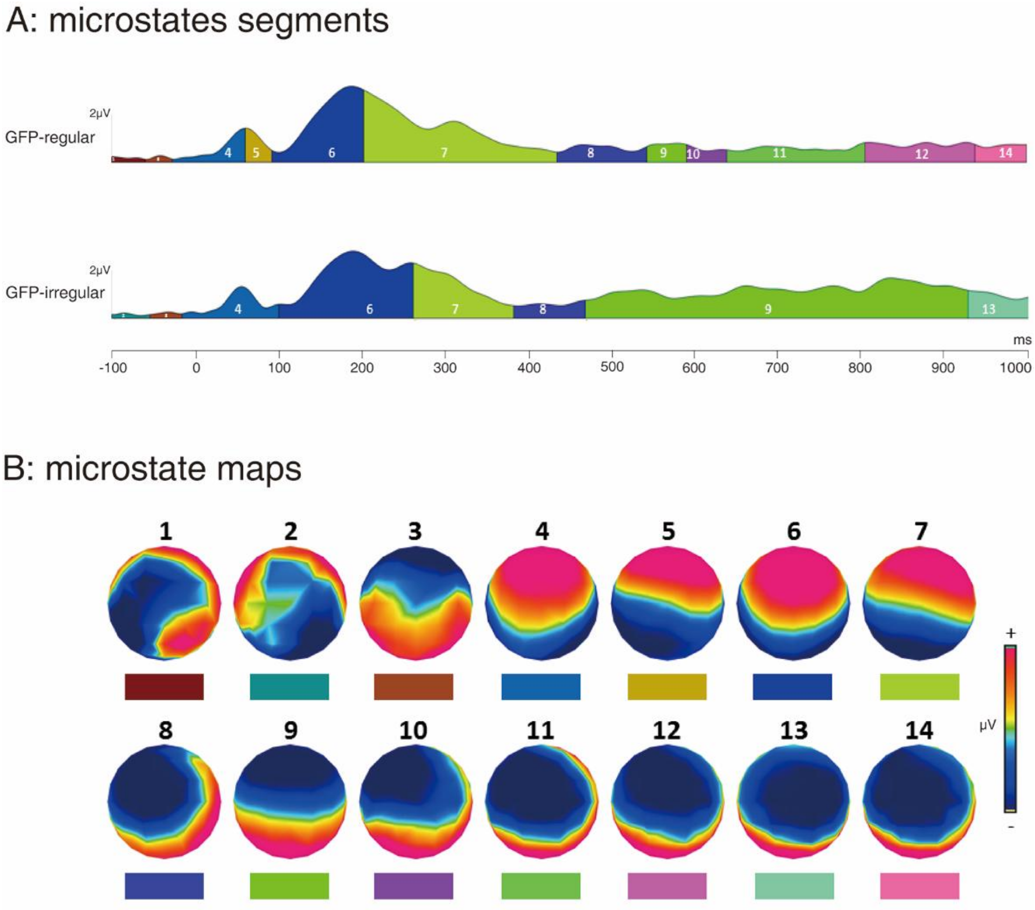

Figure 6. Study 3 - Spatiotemporal grand-average ERP analyses in the auditory-only (A-only) modality. The upper panel (**A**) shows the Global Field Power (GFP) curves of regular (upper plot) and irregular (lower plot) chords, illustrating the time periods during which each of these

microstate maps was most represented in all subjects. Fourteen distinct segments were yielded by the temporal spatial segmentation procedure (AAHC analysis). These segments are indicated by color codes and map numbers. Same color codes are used for corresponding maps in panel (B). High GFP shows 'flat' appearance of a segment, and low GFP shows 'gradient' appearance of a segment. The bottom panel (**B**) shows a series of distinct scalp configurations of the microstate maps. These maps are labelled 1 to 14 (see number above each map) and color-coded (see color bar under each map). Blue values represent negative potentials and red values represent positive potentials. These maps are 2-D projections of the 3-D electrode configuration (view from above, nasion on top).

Although the spatiotemporal segmentation analyses were conducted over the -100-1000 ms time period to identify distinct topographic maps, our statistical comparisons between regular and irregular conditions focused on the period of 100-850 ms. In this period, four maps (maps 6 to 9) arose successively in the same order for regular and irregular conditions. Visual inspection might suggest that the map segments did not differ between these two conditions. In the periods of 100-250 ms, 250-400 ms, and 400-550 ms, map 6, 7 and 8 gave rise to strong amplitudes. In the period of 550-850 ms, different maps occurred in regular (maps 9 to 12) and irregular (map 9) conditions. The total number of time frames of appearance (NumTFs) was calculated for maps 10 to 12, and maps 10 to 12 did not appear in most subjects (all > 12) in either regular or irregular condition. However, map 9 could be identified in each subject for both regular and irregular conditions.

Based on the time of occurrence of the different maps in the grand-average ERPs, we fitted maps 6 to 9 for the time window from 100 to 850 ms (see Table 8 for detailed results). Paired-sample *t*-tests between regular and irregular conditions were computed separately for the first time frame of occurrence (FOnset), the time frame of the highest correlation value (TFBCorr), and the time frame of the max Global Field Power (TFMaxGFP) over maps 6 to 9.

Table 3. Study 3 - Summary of the first time frame of appearance (FOnset), the time frame of the highest correlation value (TFBCorr) yielded for a given map, and the time frame of the max Global Field Power (TFMaxGFP) over the segment fitted by a given map (maps 6 to 9), and results of paired-sample *t*-tests between regular and irregular conditions separately for these parameters.

microstate	condition	FOnset (in ms)	*p*	TFBCorr (in ms)	*p*	TFMaxGFP (in ms)	*p*
map 6	regular	119	0.81	270	0.59	234	0.48
	irregular	127		241		181	
map 7	regular	200	0.45	291	0.24	314	0.86
	irregular	238		325		298	
map 8	regular	315	0.90	480	0.91	434	0.65
	irregular	307		471		379	
map 9	regular	540	0.78	533	0.97	734	0.54
	irregular	470		600		819	

As can be seen in Table 8, no statistically significant differences between regular and irregular conditions in the FOnset (determining topographic onset), the TFBCorr (defining topographic latency), and the TFMaxGFP (determining topographic peak latency) were obtained.

Time windows of the auditory-only modality for statistical ERP analysis were 120 to 240 ms (ERAN), 240 to 310 ms (P3a), 310 to 550 ms (P3b), and 550 to 850 ms (N5). For the ERAN, the same time window (120 to 240 ms) was adopted for the auditory-motor and motor-only modalities in order to make adequate comparisons between these three modalities.

3.4.4.2.2 ERP waveform analyses.

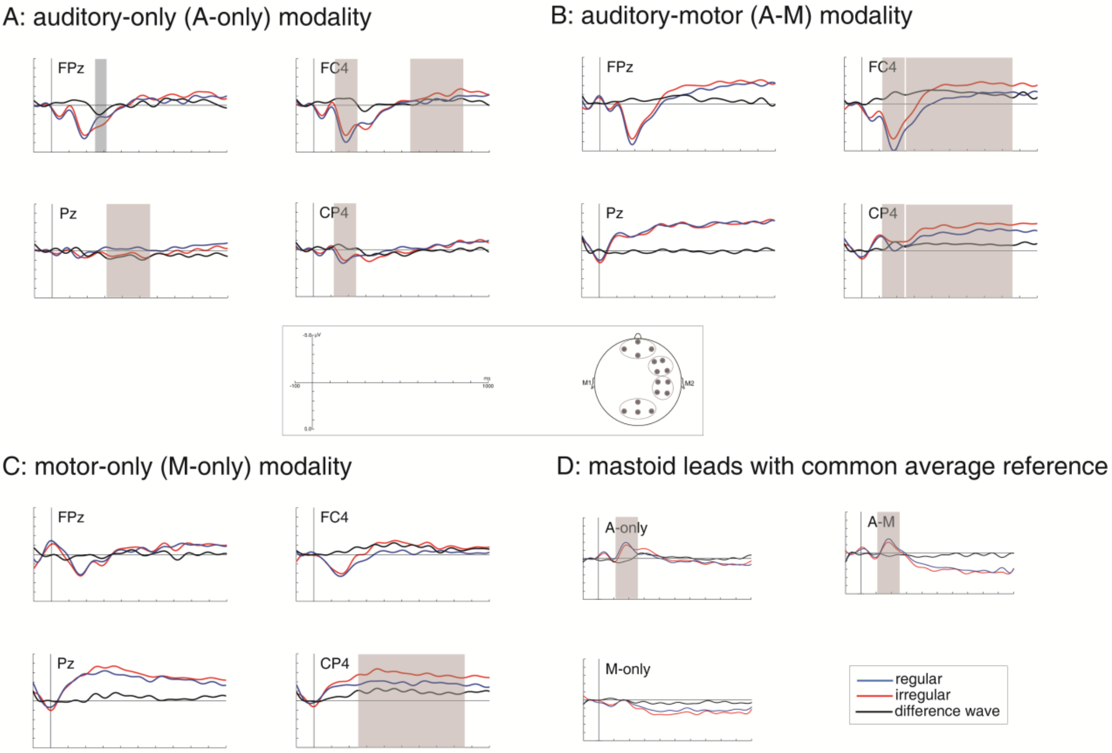

Figure 7. Study 3 - Grand-average of ERPs elicited by the regular (blue) and irregular (red) chords. The black waveform shows difference potentials (regular subtracted from irregular chords). The upper panel (**A**): When pianists listened to sequences without playing (auditory-only), irregular chords, compared to regular chords, elicited an early right anterior negativity (ERAN), a slight P3a, a P3b, and an N5. Gray-shaded areas indicate time windows used for the statistical analysis of these ERP components. The middle panel (**B**): When pianists played chord sequences on a digital piano with auditory feedback (auditory-motor), irregular chords, compared to regular chords, elicited an ERAN and a late tonic negativity (LTN, indicated by gray-shaded areas). The inset shows the ROIs used for statistical analyses (shaded in gray).The bottom panel (**C**): When pianists played chord sequences on a muted piano without auditory feedback (motor-only), irregular chords, compared to regular chords, elicited a late tonic negativity (LTN, indicated by gray-shaded areas). ERPs in (A)-(C) are referenced to the arithmetical mean of left and right mastoid leads. (D) When ERPs are referenced to the common average reference, the ERAN observed in the A-only and A-M modalities inverted polarity at mastoid leads (indicated by gray-shaded areas, shown for the right mastoid lead, M2).

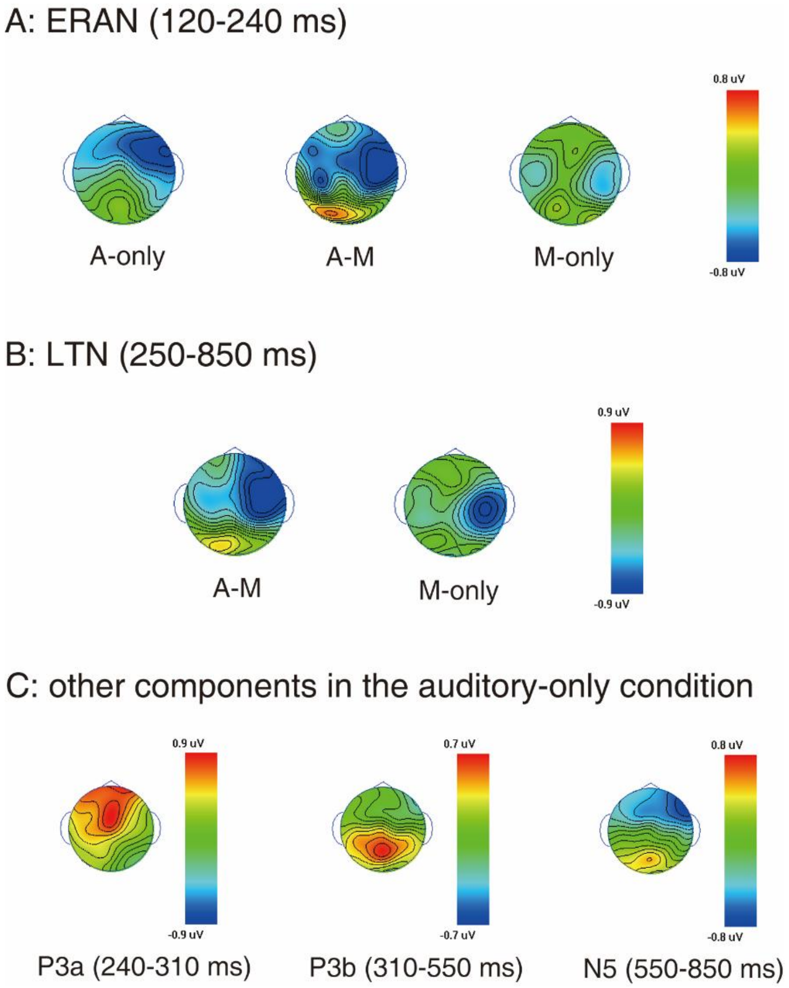

Figure 8. Study 3 - Isopotential maps of the ERAN, LTN, and other components. The upper panel (**A**) shows isopotential maps of the ERAN (difference potentials, regular subtracted from irregular chords) in the auditory-only (A-only), auditory-motor (A-M), and motor-only (M-only) modalities (referenced to the mean of M1 and M2). The middle panel (**B**) shows isopotential maps of the LTN (difference potentials, regular subtracted from irregular chords) in the A-M and M-only modalities. The lower panel (**C**) shows isopotential maps of the P3a, P3b, and N5 in the A-only modality.

ERAN. Figure 15A shows the grand-average ERPs of the A-only modality. Syntactically irregular (compared to regular) chords elicited an ERAN that was broadly distributed over anterior scalp sites (with the largest amplitude at the FPz electrode), and a mean peak latency of 180 ms (*SD* = 35 ms). When referenced to the common average reference, the ERAN inverted polarity at mastoid leads (see Figure 15D for M2). Figure 15B shows the grand-average ERPs of the A-M modality. Again, an early negativity was elicited by irregular chords which resembles the ERAN elicited in the A-only modality in the following ways: Its time course (from 120 to 240 ms), peak latency (*M* = 180 ms, *SD* = 31 ms), negative maximum over right-anterior electrode sites (maximal over FC4), and polarity inversion at mastoid leads when referenced to the common average reference (see Figure 15D for M2).

We investigated the statistical significance of the differences between modalities (A-only, A-M, and M-only) for the time window from 120 to 240 ms and two regions of interest (a right anterior and a right central ROI). Electrodes included in the ROIs were selected based on the microstate segmentation (see *Methods* in Study 3). A global ANOVA with factors modality (A-only, A-M, and M-only), regularity (regular, irregular), and ROIs (right anterior, right central) indicated a marginally significant interaction between modality, regularity and ROIs ($F(2,34) = 3.45$, $p = 0.056$, partial $\eta^2 = 0.17$; see Table 9 for all significant results), reflecting that a clear ERAN was elicited in the A-only and A-M modalities, but not in the M-only modality. Follow-up ANOVAs computed separately for the A-only, A-M, and M-only modalities (with factors regularity and ROI) indicated a main effect of regularity ($F(1,17) = 6.09$, $p = .025$, partial $\eta^2 = 0.26$) in the A-M modality and a marginally significant interaction between regularity and ROIs ($F(1,17) = 4.26$, $p = 0.055$, partial $\eta^2 = 0.20$) in the A-only modality. Paired-sample *t*-tests showed that the amplitude differences between regular and irregular chords were statistically significant for the right anterior ($t = 2.88$, $p = .01$, Cohen's $d = 0.54$) and right central ROIs ($t = 2.53$, $p = .022$, Cohen's $d = 0.46$) in the A-only modality. No main effect of regularity ($p = 0.28$) or interaction between regularity and ROIs ($p = 0.17$) was found in the M-only modality. These results reflect that a significant ERAN was elicited in the A-only and A-M modalities, but not in the M-only modality, and that the scalp distribution of the ERAN differed slightly between the A-only and the A-M modality.

P3a. Visual inspection of the waveforms, as well as the microstate segmentation, of the A-only modality suggest that the ERAN was followed by a positivity that was maximal at frontal scalp sites (FPz) and had a mean peak latency of 275 ms (*SD* = 19 ms), consistent with the characteristics of a P3a (see Figure 15A). However, a global

ANOVA for the time window from 240 to 310 ms with factors modality and regularity (computed for a frontal ROI, see *Methods* in Study 3) did not indicate a main effect of regularity (*p* = 0.59) nor an interaction between modality and regularity (*p* = 0.12).

P3b. A P3b was observed that was maximal at 401 ms (*SD* = 69 ms) and predominant at posterior scalp sites (Pz; see Figure 15A) in the A-only modality. A global ANOVA for the time window from 310 to 550 ms with factors modality and regularity for a posterior ROI (see *Methods* in Study 3) indicated an interaction between modality and regularity (*F*(2,34) = 2.59, *p* = .045, partial η^2 = 0.15; see Table 9 for all significant results). Follow-up paired-sample *t*-tests computed to compare ERPs elicited by regular and irregular chords, separately for the A-only (*p* = .037), A-M (*p* = 0.29), and M-only (*p* = 0.45) modalities, reflecting that the P3b was elicited only in the A-only modality over posterior ROI.

Upon visual inspection of the grand-average ERPs and scalp maps (see Figure 15B, 15C and Figure 16), neither the P3a nor the P3b was shown in the A-M and M-only modalities.

LTN and N5. In both the A-M and the M-only modalities (i.e., in both motor-modalities), a late tonic negativity (LTN) was elicited from around 250 to 850 ms, being maximal over FC4 in the A-M modality (see Figure 15B), and over CP4 in the M-only modality (see Figure 15C). A global ANOVA for the time window from 250 to 850 ms with factors modality, regularity, and ROIs (right anterior, right central; see *Methods* in Study 3) indicated an interaction between modality, regularity and ROIs (*F*(2,34) = 7.63, *p* = .002, partial η^2 = 0.31; see Table 9 for all significant results). Follow-up ANOVAs computed for the A-only, A-M, and M-only modalities indicated a main effect of regularity (*F*(1,17) = 12.79, *p* = .002, partial η^2 = 0.43) in the A-M only modality and an interaction between regularity and ROIs (*F*(1,17) = 9.45, *p* = .007, partial η^2 = 0.36) in the M-only modality. Paired-sample *t*-tests showed that, in the M-only modality, the amplitude differences between regular and irregular chords was statistically significant for the right central ROI (*t* = 2.63, *p* = .018, Cohen's *d* = 0.46). These results reflect that a significant LTN was elicited in the A-M and M-only modalities. No main effect of regularity (*p* = 0.70) or interaction between regularity and ROIs (*p* = 0.32) was found in the A-only modality. However, in the A-only modality, irregular chords (compared to regular chords) elicited an N5 with an onset at around 550 ms, peaking at 722 ms (*SD* = 84 ms) at frontal scalp sites (FPz; see Figure 15A). A paired-sample *t*-test for the time window from 550 to 850 ms at the right anterior ROI (see *Methods* in Study 3) showed that the amplitude difference between regular and irregular chords was marginally significant (*t* = 2.03, *p* = 0.058, Cohen's *d* = 0.42), reflecting that the N5 was significant over right anterior electrode leads.

Table 9. Study 3 - Summary of global ANOVAs for the amplitudes of the ERAN, P3b, and late tonic negativity (LTN) with factors modality (A-only, A-M, and M-only), regularity (regular, irregular), and ROIs. Only significant results (main effects and interactions) with *p* ≤ .05 are listed. Significance of p values is indicated by asterisks ($^{*}p \le .05$, $^{**}p \le .01$, $^{***}p \le .001$).

factors	ERAN		P3b		LTN	
	df	F	*df*	F	*df*	F
modality	2,34	7.00**	2,34	10.41***	2,34	7.31**
regularity	1,17	8.00*			1,17	6.87*
modality × regularity			2,34	2.59*	2,34	4.10*
modality × ROIs	2,34	9.07**			2,34	6.99**
modality × regularity × ROIs	2,34	3.45*			2,34	7.63**

3.4.4 Discussion

The present study aimed at investigating the impact of auditory- and motor-predictions on neural correlates of music-syntactic processing. To address this issue, professional pianists listened to chord sequences (without playing, auditory-only modality), or played sequences with (auditory-motor modality) or without auditory feedback (motor-only modality).

Before discussing the results in detail, one important issue that needs consideration is how specific veridical expectations (prompted by the cue) were. In the auditory-only modality, the task for pianists was attentive listening and they had relatively high scores of auditory imagery obtained in the post-EEG debriefing (*M* =

7.39, 9-point scale), indicating that cues (i.e., informing participants the tonal key and regularity of the sequence ending) were utilized to generate specific veridical expectations. To play the cued sequence in the auditory-motor and the motor-only modalities, pianists should form specific veridical expectations (i.e., they knew exactly what to play), which was supported by relatively high scores of motor imagery obtained in the post-EEG debriefing (in the auditory-motor modality: M = 5.94, in the motor-only modality: M = 7.61, 9-point scale).

3.4.4.1 Behavioral results

In both pre- and post- experiments, the high correct responses of regularity judgments and confidence ratings, in combination with that none of these judgments was based on guessing, provides assurance that pianists had strong and specific representations of Western harmonic principles. The correct responses of regularity judgments did not increase significantly from pre- to post- experiment. This is most probably due to the ceiling effect in the post-experiment (100% correct response). The shortened RT of regularity judgments from pre- to post- experiment indicates that the veridical expectations were well formed by the cue.

Note accuracies, IOIs, and tempo of final chords suggest that irregular chords were executed with lower accuracy and more slowly than regular chords in both the auditory-motor and the motor-only modalities. These findings are consistent with previous studies of imitating chord sequences on a muted piano (see also *Introduction* in Study 3) (Novembre and Keller, 2011; Sammler et al., 2013b; Bianco et al., 2016a). The processing costs for playing irregular chords indicate that schematic expectations (for playing regular chord) are resistant to veridical expectations (for playing irregular chord) on a behavioral level. Thus, more cognitive and motor programming effort is required to play irregular chords (due to inhibiting preplanned dominant responses of playing regular chords), even when irregular chords had lower motoric complexity (due to lower number of black keys).

It is noteworthy that the production accuracy was not modulated by the auditory feedback (i.e., there was no difference between the auditory-motor and the motor-only modalities), similar to previous studies indicating negligible effects of auditory feedback on performance accuracy and error-monitoring in trained musicians (Banton, 1995; Finney, 1997; Repp, 1999; Finney and Palmer, 2003; Pfordresher, 2005; for a review, see Maidhof, 2013). However, the duration of final chords (i.e., longer duration of final chords in the auditory-motor than the motor-only modality), the IOIs of final chords (i.e., longer IOIs of final chords in the auditory-motor than the motor-only modality), and the overall tempo (i.e., slower tempo of auditory-motor than motor-only modality) were influenced by the auditory feedback. This might suggest that during the production of multivoiced music, auditory feedback is important for accurate timing; however, this needs to be specified in future studies.

According to Zatorre and Halpern (2005), motor imagery is likely to be elicited when listening to well-rehearsed piece, and auditory imagery was induced when executing finger movements in trained musicians (Zatorre and Halpern, 2005). In the current study, the strength of anticipating finger movements (motor imagery) was strongest in the motor-only modality and weakest in the auditory-only modality, and the strength of anticipating the sounds (auditory imagery) was stronger than motor imagery in both the auditory-only and the auditory-motor modalities, suggesting that pianists used auditory imagery more than motor imagery when they listened to and played chord sequences with auditory feedback, and used both auditory and motor imagery when they played chord sequences without auditory feedback. In the auditory-motor modality, correlation analysis between the strength of auditory imagery and note accuracy indicates that pianists who used stronger auditory imagery could perform better. In the auditory-only modality, negative correlation between the age at which piano studies commenced and the strength of motor imagery, and positive correlation between hours of piano-practice a day and the strength of auditory imagery were found, suggesting that long-term extensive musical training (as reflected in early commencement of musical training and more hours of piano-practice a day) would strengthen the audio-motor coupling of pianists, and improved their auditory and motor imagery.

3.4.4.2 ERP results

Prediction research is more often discussed in the framework of predictive coding theories nowadays (Schröger et al., 2015b). In this framework, perception is a process of hierarchical inference (Vassena et al., 2016), and the perceptual system is regarded as a hierarchically organized model, viewed as the “compromise” between sensations of actual sensory input from lower levels (bottom-up) and predictions generated from internal generative models of the environment from higher levels (top-down), with the primary objective to minimize the discrepancy (termed ”prediction error”) between these levels (Friston, 2005; 2009; 2010; Winkler and Czigler, 2012). There is growing empirical evidence in the music domain supporting the predictive coding theories, for example, the perception of pitch (Furl et al., 2011; Kumar et al., 2011), rhythm and meter (Vuust and Witek, 2014; Vuust et al., 2009), and the relationship between perception and action in music performance (Bianco et al., 2016b) (for a review, see Maes, 2016).

Within the predictive coding framework, it is assumed that top-down predictions influence bottom-up processing. Curiously, we have found in our previous studies (Guo and Koelsch, 2015; 2016) that bottom-up syntactic processing is surprisingly robust against top-down predictions. However, predictions in our previous studies were only auditory. The question of the current study is whether bottom-up music-syntactic processing would be less robust against top-down motor predictions (as one would expect based on the framework of predictive coding theories).

3.4.4.2.1 ERAN

Our experiment showed that although pianists had veridical knowledge of an irregular ending, irregular chords elicited a clear ERAN when pianists listened to chord sequences without playing (auditory-only modality), which replicates our previous findings (Guo and Koelsch, 2015; 2016). The peak latency of the ERAN (180 ± 35 ms) was consistent with the peak latency (183 ± 9 ms) of our last study (Guo and Koelsch, 2016) when musicians were informed as to whether the final chord would be regular or irregular. During the perception of an irregular chord, we assume that auditory predictions generated by schematic syntactic knowledge stored in long-term memory consistently suggested a regular ending. In contrast, auditory predictions generated by veridical knowledge on the basis of the preceding cue suggested an irregular ending. Converging evidence has been provided by behavioral studies, indicating that the influence of schematic expectations remains stronger than veridical expectations on chord processing (Justus and Bharucha, 2001; Tillmann and Bigand, 2004; 2010). These findings gain further support from our last two ERP studies (Guo and Koelsch, 2015; 2016) and the processing costs for playing irregular chords of the current study which underline the strength of schematic expectations. Therefore, top-down prediction outcomes generated by veridical and schematic knowledge were in part "wrong", and the actually perceived sound of irregular chord (bottom-up) conflicted with predicted regular chord (top-down). In terms of the predictive coding theories (Friston, 2005; 2009; 2010), neural networks form predictions and generate prediction errors when the actual input does not match the predictions (Nazimek et al., 2013). The prediction error here was assumed to be reflected in the ERAN, representing the detection of music-syntactic violations (i.e. syntactic structure building). In contrast, when processing a syntactically regular chord, bottom-up perceived sound of regular chord matched top-down predictions of regular chord, thus no prediction error appeared. The ERAN was frontal- and right-lateralized in the present study, which gains support from a recent finding showing activation of right IFG when pianists listened to chord sequences they played in this experiment (Bianco et al., 2016b).

Playing irregular chord sequences with auditory feedback (auditory-motor modality) evoked the ERAN in pianists. During the production of a syntactically irregular chord with auditory feedback, top-down predictions are generated by veridical and schematic knowledge.

a) Veridical knowledge. Action goals (serving as the mental reference point; Asano and Boeckx, 2015) were established based on veridical knowledge (prompted by the cue). Prior to the actual execution of movement, pianists first programmed a series of actions (i.e., key-presses on the piano with assigned fingerings). The formation of action programs led to copies of the motor commands (e.g., “efference copies” or “corollary discharge”; Maes et al., 2014) to expect the corresponding action effects (here referring to internal representation in which the motor commands and the sensory consequences were linked; Koelsch, 2012; Hommel et al., 2001; Koch et al., 2004; Dick et al., 2011; Wolpert and Ghahramani, 2000; for reviews, see Hoffmann et al., 2004). Activated by the sensory consequences in an anticipatory fashion, motor commands were sent, which drove and controlled movement to produce the required sounds (Elsner and Hommel, 2001; Thompson et al., 2006) ("correct" auditory- and motor-predictions).

b) Schematic knowledge. Based on internalized schematic knowledge acquired through years of musical training, pianists predicted the action effects (i.e. auditory predictions) of the most likely harmonically regular chord. Previous studies indicate that the schematic knowledge (i.e. syntactic regularities) is acquired motorically in expert pianists due to extensive practice, thus pianist predicted the execution (i.e. combinations of movement) of regular chord as well (motor predictions; Sammler et al., 2013b; Bianco et al., 2016a; 2016b; Novembre and Keller, 2011) ("wrong" auditory- and motor-predictions).

Several studies have provided evidence that schematic expectations are resistant to veridical expectations (Justus and Bharucha, 2001; Tillmann and Bigand, 2004; 2010; Guo and Koelsch, 2015; 2016). We assume that top-down prediction outcomes based on veridical and schematic knowledge (i.e. regular chord) were compared against bottom-up actual consequences of motor commands (i.e. sensorimotor feedback of irregular chord), which might generate the prediction error (as reflected in the ERAN). The finding that the ERAN showed a right-lateralization was in line with studies by Sammler et al. (2013b) and Bianco et al. (2016b).

The ERAN was not shown when pianists played chord sequences without auditory feedback (motor-only modality), which is consistent with the finding reported by Bianco et al. (2016a). Two possible explanations can be put forward for the absence of the ERAN: a) the elicitation of the ERAN requires auditory (acoustic) input, b) the ERAN is abolished by motor predictions. It must be noted that the ERAN was elicited in the auditory-motor modality in which the motor predictions were generated, thus the second explanation was excluded.

Within the context of predictive coding theories, action affords the opportunity to bring actual input to fit predictions (Friston et al., 2010; Feldman and Friston, 2010; Hohwy, 2013). Comparing to music perception, music production should be more efficient in minimizing prediction errors. In accordance with this theory, the ERAN (taken to reflect a prediction error) would differ between the auditory-only and the auditory-motor modalities. However, our results suggest that music-syntactic processing (with auditory input) was not modulated by motor predictions (see also *Results* in Study 3).

3.4.4.2.2 P300(P3b)-N5

The ERAN was followed by a trend of frontal P3a in the auditory-only modality. The P3a has been taken to reflect an involuntary shift or orienting of attentional focus toward deviant sounds (Münte et al., 1998; Escera et al., 2000). However, the P3a was statistically not significant, similar to the finding of our last study showing the absence of the frontal P300 (the P3a) when musicians listened to chord sequences with the knowledge about the upcoming event (Guo and Koelsch, 2016). Although the P3a was not statistically observed here, visual inspection of the waveforms and the isopotential maps (see Figure 15A and Figure 16) suggests a trend of this effect. Following the trend of the P3a, a posterior distributed P3b was elicited in the auditory-only modality, which is generally interpreted as updating of working memory representations or decision-related processes of structural re-analysis (Donchin and Coles, 1988; Polich, 2007). This finding is consistent with studies reporting a P3b after harmonic violations of learned familiar sequences (Guo and Koelsch, 2016; Carrión and Bly, 2008). The P3b was followed by an anterior distributed N5 in the auditory-only modality, which reflects processing of syntactic integration of each incoming chord into the ongoing sequence (Koelsch et al., 2000; 2002c).

Taken together, the pattern (ERAN-P300(P3b)-N5) observed in the auditory-only modality is reminiscent of ERPs (ERAN-P300) elicited by irregular chords when musicians were informed about whether the sequence they listened to would end on regular or irregular chord (Guo and Koelsch, 2016).

3.4.4.2.3 LTN

Instead of the P300 and N5, the ERAN was followed by an LTN in the auditory-motor modality. According to the predictive coding theories (Friston, 2005; 2009; 2010), during the production of a syntactically irregular chord with auditory feedback, top-down auditory- and motor- prediction outcomes were in part "wrong" (see also top-down predictions in the elicitation of the ERAN in the auditory-motor modality), which conflicted with bottom-up sensorimotor feedback. The prediction error here was assumed to be reflected in the LTN. Note that the LTN was also evoked in the motor-only modality, suggesting that even without auditory input, sensorimotor feedback (i.e. proprioceptive and tactile information) alone was sufficient to generate this error signal. The LTN elicited during music production may be interpreted as a signal of differentiation and correction (maybe integration as well), which parallel processes of syntactic structure building and reanalysis (maybe syntactic integration as well) in music perception (as reflected in the ERAN, P300, and probably N5 elicited in the auditory-only modality; Koelsch, 2012). It may also explain the long-lasting feature of the LTN (i.e. 250-850 ms) which is assumed to encompass at least two distinct processes (i.e. discrepancy differentiation and correction). However, this interpretation remains to be specified in future studies. It is worth noting that both the ERAN and LTN were elicited in the auditory-motor modality, indicating that the ERAN may be associated with auditory-related differentiation and the early stage of the LTN appears to be associated with motor-related differentiation. If this notion is correct, it provides an explanation of the missing LTN in the auditory-only modality (due to absence of finger movements) and the missing ERAN in the motor-only modality (due to absence of auditory input). An explanation of the anterior distribution of the LTN may reside in the fact that action-specific processes are located in anterior brain regions (Maffongelli et al., 2015). The right-hemispheric dominance of the LTN is probably due to that the right-hemispheric lateralization of perceptual music-syntactic processing (i.e. the

ERAN) transfers to motor-related processes.

It is noteworthy that the amplitude of the LTN did not differ between the auditory-motor and the motor-only modalities (see also *Results* in Study 3). A series of studies have found that auditory-related areas were activated when musicians played an instrument without auditory feedback (Bangert et al., 2006; Bangert and Altenmüller, 2003; Baumann et al., 2005). Previous studies have showed that musical imagery could compensate the absence of auditory feedback (Bishop et al., 2013; Keller et al., 2010; Repp, 1999). Imagining and perceiving auditory stimuli share common psychological processes and engage similar brain regions (Halpern and Zatorre, 1999; Leaver et al., 2009; Ohnishi et al., 2001; Zatorre et al., 1998; Schürmann et al., 2002; Martin et al., 2017). Although bottom-up auditory input is missing in the motor-only modality, it is likely that auditory representations were activated, which is supported by the fact that the high auditory imagery scores in the motor-only (M = 7.94, in a 9-point scale) and auditory-motor (M = 7.50, in a 9-point scale) modalities are statistically the same. Thus, the LTN, interpreted as a signal of discrepancy differentiation and correction, was not modulated by the auditory feedback, consistent with our behavioral finding that production accuracy was not modulated by the auditory feedback (see also *Results* in Study 3).

In auditory prediction studies, effects of prediction on the auditory ERP can be strongly modulated by others factors like attention (for reviews, see Lange, 2013; Schröger et al., 2015b). In our study, the explicit cue may allow pianists to direct their attention toward the upcoming sequence ending, and action may probably enhance orienting of attention. However, we did no assess how much attention pianists paid in different modalities. Therefore, future studies could control attention by giving a specific task to participants to ensure that they pay the same amount of attention to each condition.

3.4.5 Conclusion

In this study, we found that when pianists listened to and played chord sequences, irregular chords elicit an ERAN. Thus, it appears that the detection of music-syntactic irregularities is not modulated by motor predictions (when auditory input is present). Additionally, when pianists listened to and played chord sequences without auditory feedback, irregular chords elicited an LTN. Thus, the LTN is not modulated by auditory feedback. Our findings indicate that top-down motor-predictions of upcoming syntactic errors (even if the error is produced deliberately by an individual) do not modify the bottom-up perceptual processing of syntactically irregular events. Thus, musicians will make wrong auditory- and motor-predictions when playing a deceptive cadence in a known piece of music, and the expected deceptive cadence will still be perceived as "deceptive".

3.5 Summary

Study 1 investigated whether the acquisition of veridical knowledge of music-syntactic regularities (acquired through supervised learning) modulates early, partially automatic, music-syntactic analysis (as reflected in the ERAN), and / or late controlled processes of syntactic re-analysis (as reflected in the LPC). Excerpts of piano sonatas with syntactically regular and less regular chords were presented repeatedly (ten times) to non-musicians and amateur musicians. Participants were informed by a cue as to whether the following excerpt contained a regular or less regular chord. Results showed that the repeated exposure to several presentations of regular and less regular excerpts did not influence the ERAN elicited by less regular chords. By contrast, amplitudes of the LPC (as well as of the P3a evoked by less regular chords) decreased systematically across the learning trials. These results reveal that late controlled, but not early, partially automatic, neural mechanisms of music-syntactic processing are modulated by veridical expectations.

Study 2 investigated how prior knowledge about an upcoming syntactically irregular event modulates neural correlates of music-syntactic processing. Two versions of a short chord sequence were presented repeatedly to non-musicians and musicians. One sequence version ended on a syntactically regular chord, and the other one ended on a syntactically irregular chord. Participants were either informed (cued condition), or not informed (non-cued condition) about whether the sequence would end on the regular or the irregular chord. Results indicate that in the cued condition (compared to the non-cued condition) the peak latency of an ERAN, elicited by irregular chords, was earlier in both non-musicians and musicians. However, the expectations due to the knowledge about the upcoming event (veridical expectations) did not influence the amplitude of the ERAN. These results suggest that veridical expectations modulate only the speed, but not the principle mechanisms, of music-syntactic processing.

Study 3 investigated whether music-syntactic processing is influenced by top-down motor predictions. Two different versions of chord sequences differing in the final chord (syntactically regular or irregular) were used as experimental stimuli. Professional pianists either played these two versions with auditory feedback (on a digital piano, “auditory-motor modality”), or played them without auditory feedback (on a muted piano, “motor-only modality”), or listened to them (without playing, “auditory-only modality”). On a behavioral level, results showed processing costs for playing irregular chords (i.e., lower accuracy and longer inter-onset intervals between the penultimate and final chords) in both the auditory-motor and the motor-only modalities. On a neurophysiological level, an early right anterior negativity (ERAN) was elicited in the auditory-only and the auditory-motor modalities, reflecting the detection of music-syntactic irregularities (with auditory input). However, the ERAN was not modulated by motor predictions. The ERAN was followed by a P300(P3b)-N5 pattern in the auditory-only modality. A late tonic negativity (LTN) was evoked in the auditory-motor and the motor-only modalities. The LTN was not modulated by the auditory feedback. These results indicate that top-down motor-predictions of upcoming syntactic errors (even if the error is produced deliberately by an individual) do not modify the bottom-up perceptual processing of syntactically irregular events.

4. General discussion

4.1 In the framework of predictive coding theories

Results of Study 1 and 2 showed that cognitive predictive processes modulated the speed of early and partially automatic music-syntactic processing, and the late controlled music-syntactic processing. In other words, veridical expectations about what are coming next, especially those “unexpected” events, do not alter the schematic expectations from higher order cognitive system associated with music-syntactic processing. Findings of Study 3 indicated that action predictive processes did not modulate the processing of music-syntactic regularities. Results of Study 1-3 suggest that sometimes we know that we will make wrong predictions when listening to or playing a deceptive cadence in a known piece of music.

Within the context of predictive coding theories and free-energy minimization (Friston, 2005; 2009; 2010; Clark, 2013), prediction errors can be reduced by either changing predictions or the things that we predict (Clark, 2013). The prediction error in music-syntactic processing is assumed to be reflected in the ERAN (see also *Discussion* in Study 3). In Study 1 and 2, veridical knowledge afforded the opportunity of resolving uncertainty in the future. As a method of adjusting or optimizing higher level representations, it failed to reduce the prediction error. In Study 3, action affords the opportunity to bring actual input to fit predictions (Friston et al., 2010; Feldman and Friston, 2010; Hohwy, 2013). However, it failed to minimize the prediction error as well. Although it is theoretically possible to modulate the long-term memory representations of music-syntactic processing through veridical knowledge and motor predictions, our findings suggest that cognitive representations of music-syntactic regularities cannot easily be modified (probably due to the culturally received and ‘generic’ learned musical schema).

4.2 In the framework of common coding theory

Comparing the auditory predictions generated in Study 1 and 2 with auditory-only condition in Study 3, one notable difference lies in whether sequences were played by participants before or not (motorically known or not). Excerpts or chord sequences were only perceived without playing in Study 1 and 2, but sequences were played (with and without auditory feedback) in Study 3.

Based on the common coding theory and action-effect principle (see also *Common coding theory*), once action (e.g., pressing a key) and its consequence (e.g., a tone) is learned, internal forward models predict the sensory consequences of movements, similarly to the action-effect principle that actions can prime the corresponding action-effects. According to this notion, pianists in Study 3 generated auditory predictions in the auditory-motor and the motor-only conditions. However, when comparing auditory-motor with auditory-only condition, it is difficult to control how much auditory predictions were generated. In terms of internal inverse models, corresponding motor commands are selected to achieve the intended or perceived sensory consequences. Similarly to the action-effect principle that anticipating action-effects guides the selection of corresponding actions, pianists in Study 3 generated motor predictions in the auditory-only condition as well. Again, it is difficult to control how much motor predictions were generated when comparing auditory-only with auditory-motor condition.

4.3 Tentative theoretical model and research summary

A tentative theoretical model which integrates predictive coding theories with common coding theory is proposed (see Figure 17).

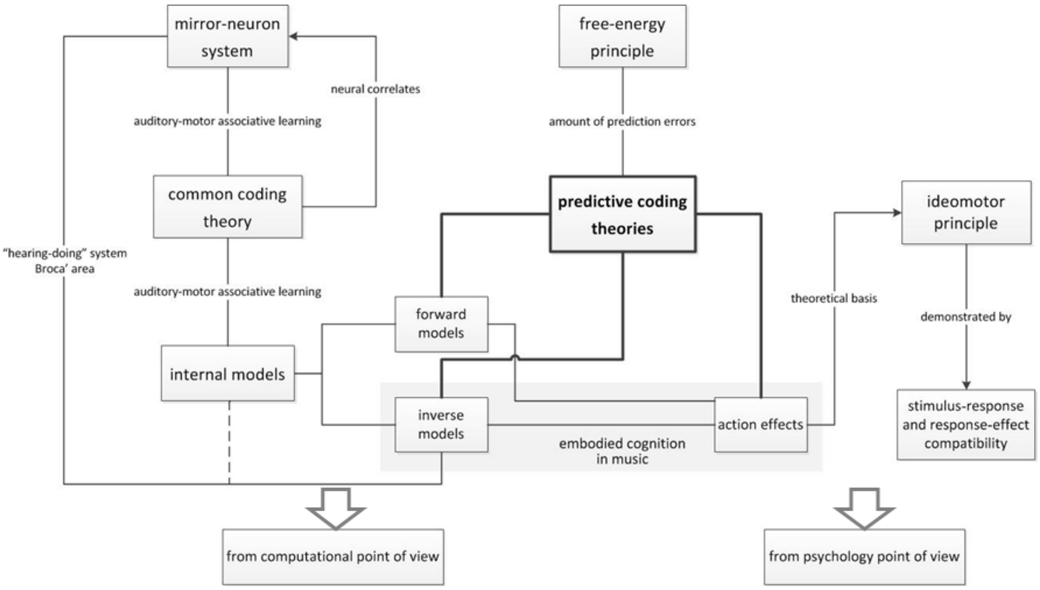

Figure 9. General discussion - A tentative theoretical model. This model integrates predictive coding theories with common coding theory (see also *Theoretical background*). Theories and models used to explain results of studies are highlighted in bold. Note that the distance between theories in this model (e.g., common coding theory and ideomotor principle) does not represent the actual correlation (e.g., common coding theory and ideomotor principle are tightly linked). For details, please refer to the main text below.

James (1890) proposed that the perception and action may share common cognitive roots (James, 1890). Internal models suggest that perception and action are tightly linked through auditory-motor associative learning (Maes et al., 2014; Novembre and Keller, 2014). From a computational point of view, there are two kinds of internal models, which are forward and inverse models. The notion of internal models is compatible with predictive coding theories (Friston et al., 2010; Schröger et al., 2015a; Lange, 2013). Prediction error is a core concept of predictive coding framework (Vuust et al., 2009; Hohwy, 2013), and free-energy quantifies the amount of prediction errors (Clark, 2013). From a psychology point of view, actions are planned and controlled by anticipating the perceivable effects produced by these actions (i.e. action effects). Action effects are bidirectional, which is similar to the internal forward and inverse models. Inverse models used to explain action effects on music perception refer to embodied cognition in music (Maes et al., 2014). The role of action effects in action planning, control and execution is theoretically based on ideomotor principle (Hommel et al., 2001; Stock and Stock, 2004). Studies of stimulus-response compatibility and response-effect compatibility have been found to provide support for the ideomotor principle (D'Ausilio et al., 2006; Keller and Koch, 2008). Internal inverse models are supported by activations of motor-related network when listening to well-rehearsed pieces without performing any movement (Lahav et al., 2007; D'Ausilio et al., 2006; Bangert et al., 2006; Bangert and Altenmüller, 2003; Haueisen and Knösche, 2001). These findings support the hypothesis of a “hearing-doing” system, which consists of Broca’s area (Lahav et al., 2007). It has been proposed that Broca’s area is a crucial node of mirror neuron system (Fazio et al., 2009). The mirror neuron system is regarded as the neural correlates of common coding of perception and action, and it is facilitated by auditory-motor associative learning (Maes et al., 2014; Koelsch, 2012).

Study 1-3 are summarized according to the framework of predictive coding theories (Friston, 2005; 2009; 2010) (see Figure 18). In Study 1 and 2, auditory prediction outcomes (based on the schematic knowledge and modulated by the veridical knowledge) from higher levels (top-down) are compared with actually perceived sound from lower levels (bottom-up), which result either in “match” or “mismatch”. The mismatch signals are regarded as prediction errors occur at a series of hierarchical levels, assumed to be reflected in the ERAN and P300 (in Study 1-3), LPC (in Study 1) and N5 (in Study 3). Prediction errors are passed upwards to adjust the prediction at higher levels. Results showed that the ERAN is not modulated by auditory predictions generated by veridical knowledge. However, the prediction errors of the P300 and LPC (in Study 1) were minimized (as reflected by the systematical decrease of the P300 and LPC during the auditory predictive processes). In Study 3, based on the establishment of efference copies of motor commands (modulated by the veridical knowledge), auditory- and motor- prediction outcomes (based on the schematic knowledge) are compared with actual consequences of motor commands (i.e. sensorimotor feedback), which result either in “match” or “mismatch”. The mismatch signals are regarded as prediction errors, assumed to be reflected in the ERAN and LTN (in the auditory-motor condition) and LTN (in the motor-only condition). Results showed that the ERAN is not modulated by motor predictions.

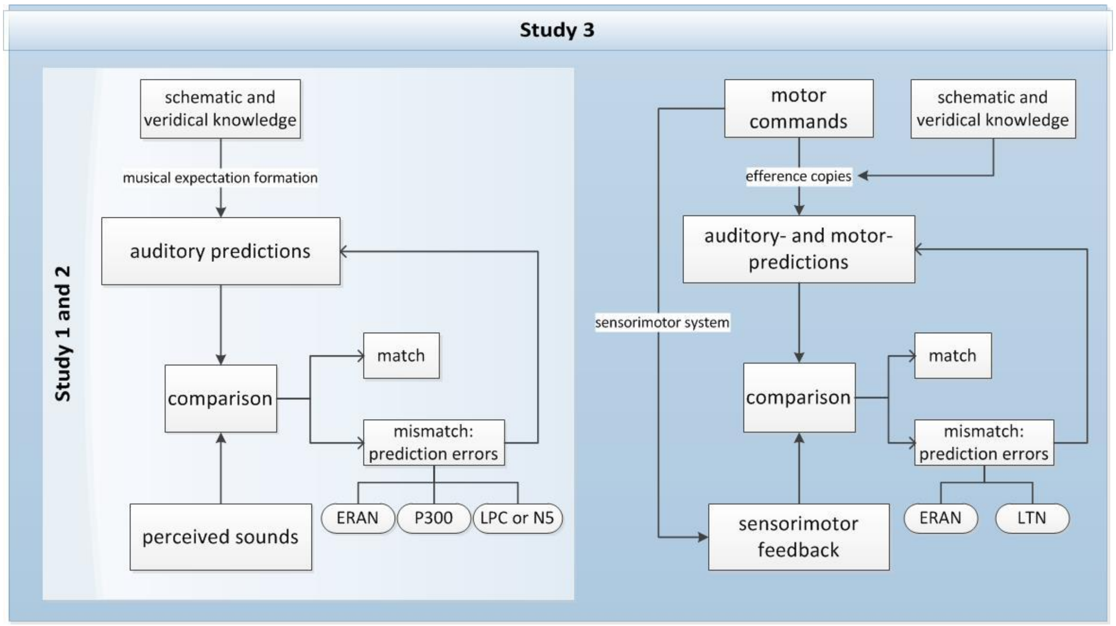

Figure 10. General discussion - Summary of Study 1-3. Study 1 and 2 investigated whether auditory predictions modulated music-syntactic processing (as reflected in the ERAN). Study 3 investigated the impact of auditory- and motor-predictions on neural correlates of music-syntactic processing. For details, please refer to the main text above.

In summary, Study 1-3 provides strong evidence that music-syntactic processing occurs relatively independent of auditory- and motor- predictive processes on a neurophysiological level.

4.4 Future directions

a) Investigating whether the fulfillment and violation of schematic expectations would be modulated by, or interact with the validity of the veridical expectations generated by competing information of what to come (e.g., providing participants with wrong cues of the sequences using a cue-validity method).

b) Measuring participants repeatedly across the same EEG-method taking place on several days (i.e. over longer periods of exposure).

c) Schematic expectations in this dissertation were developed on the basis of schematic syntactic knowledge of Western tonal music. Future studies could explore whether the influence of schematic expectations would remain stronger than veridical expectations in other musical styles.

References

Agres K, Abdallah S, Pearce M. 2017. Information-theoretic properties of auditory sequences dynamically influence expectation and memory. Cogn Sci. in press.

Alain C, Woods DL. 1997. Attention modulates auditory pattern memory as indexed by event-related brain potentials. Psychophysiology. 34:534-546.

Alho K. 1995. Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. Ear Hear. 16:38-51.

Arzy S, Mohr C, Michel CM, Blanke O. 2007. Duration and not strength of activation in temporo-parietal cortexpositively correlates with schizotypy. Neuroimage. 35:326-333.

Asano R, Boeckx C. 2015. Syntax in language and music: What is the right level of comparison? Front Psychol. 6:1-16.

Ashe J, Lungu OV, Basford AT, Lu X. 2006. Cortical control of motor sequences. Curr Opin Neurobiol. 16:213-221.

Bäss P, Jacobsen T, Schröger E. 2008. Suppression of the auditory N1 event-related potential component with unpredictable self-initiated tones: Evidence for internal forward models with dynamic stimulation. Int J Psychophysiol. 70:137-143.

Bakeman R. 2005. Recommended effect size statistics for repeated measures designs. Behav Res Methods. 37:379-384.

Baldeweg T. 2007. ERP repetition effects and mismatch negativity generation: A predictive coding perspective. J Psychophysiol. 21:204-213.

Bangert M, Altenmüller EO. 2003. Mapping perception to action in piano practice: A longitudinal DC-EEG study. BMC Neurosci. 4:1-14.

Bangert M, Peschel T, Schlaug G, Rotte M, Drescher D, Hinrichs H, Heinze HJ, Altenmüller E. 2006. Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction. Neuroimage. 30:917-926.

Banton L. 1995. The role of visual and auditory feedback during the sight-reading of music. Psychol Music. 23:3-16.

Baumann S, Koeneke S, Meyer M, Lutz K, Jäncke L. 2005. A network for sensory-motor integration: What happens in the auditory cortex during piano playing without acoustic feedback? Ann N Y Acad Sci. 1060:186-188.

Baumann S, Koeneke S, Schmidt CF, Meyer M, Lutz K, Jancke L. 2007. A network for audio-motor coordination in skilled pianists and non-musicians. Brain Res. 1161:65-78.

Bechara A, Damasio AR. 2005. The somatic marker hypothesis: A neural theory of economic decision. Games Econ Behav. 52:336-372.

Bendixen A, SanMiguel I, Schröger E. 2012. Early electrophysiological indicators for predictive processing in audition: A review. Int J Psychophysiol. 83:120-131.

Bendixen A, Schröger E, Winkler I. 2009. I heard that coming: Event-related potential evidence for stimulus-driven prediction in the auditory system. J Neurosci. 29:8447-8451.

Besson M, Faïta F, Requin J. 1994. Brain waves associated with musical incongruities differ for musicians and non-musicians. Neurosci Lett. 168:101-105.

Besson M, Faïta F. 1995. An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. J Exp Psychol Hum Percept Perform. 21:1278-1296.

Besson M, Macar F. 1987. An event-related potential analysis of incongruity in music and other non-linguistic contexts. Psychophysiology. 24:14-25.

Besson M, Schön D. 2001. Comparison between language and music. Ann N Y Acad Sci. 930:232-258.

Bharucha JJ, Krumhansl CL. 1983. The representation of harmonic structure in music: Hierarchies of stability as a function of context. Cognition. 13:63-102.

Bharucha JJ. 1987. Music cognition and perceptual facilitation: A connectionist framework. Music Percept. 5:1-30.

Bharucha JJ. 1994. Tonality and expectation. Oxford University Press.

Bharucha JJ, Stoeckig K. 1987. Priming of chords: Spreading activation or overlapping frequency spectra? Percept Psychophys. 41:519-524.

Bharucha JJ, Stoeckig K. 1986. Reaction time and musical expectancy: Priming of chords. J Exp Psychol Hum Percept Perform. 12:403-10.

Bianco R, Novembre G, Keller PE, Scharf F, Friederici AD, Villringer A, Sammler D. 2016a. Syntax in action has priority over movement selection in piano playing: An ERP study. J Cogn Neurosci. 28:41-54.

Bianco R, Novembre G, Keller PE, Seung-Goo K, Scharf F, Friederici AD, Villringer A, Sammler D. 2016b. Neural networks for harmonic structure in music perception and action. Neuroimage. 142:454-464.

Bigand E, Madurell F, Tillmann B, Pineau M. 1999. Effect of global structure and temporal organization on chord processing. J Exp Psychol Hum Percept Perform. 25:184-197.

Bigand E, McAdams S, Foret S. 2000. Divided attention in music. Int J Psychol. 35:270-278.

Bigand E, Parncutt R, Lerdahl F. 1996. Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training. Percept Psychophys. 58:124-141.

Bigand E, Pineau M. 1997. Global context effects on musical expectancy. Percept Psychophys. 59:1098-1107.

Bigand E, Poulin B, Tillmann B, Madurell F, D'Adamo DA. 2003. Sensory versus cognitive components in harmonic priming. J Exp Psychol Hum Percept Perform. 29:159-171.

Bigand E, Poulin-Charronnat B. 2006. Are we “experienced listeners”? A review of the musical capacities that do not depend on formal musical training. Cognition. 100:100-130.

Bigand E, Tillmann B, Poulin B, D'Adamo DA. 2001. The effect of harmonic context on phoneme monitoring in vocal music. Cognition. 81:B11-B20.

Bigand E, Tillmann B, Poulin-Charronnat B, Manderlier D. 2005. Repetition priming: Is music special? Q J Exp Psychol A. 58:1347-1375.

Bishop L, Bailes F, Dean R. 2013. Musical imagery and the planning of dynamics and articulation during performance. Music Percept. 31:97-117.

Boeckx CA, Fujita K. 2014. Syntax, action, comparative cognitive science, and Darwinian thinking. Front Psychol. 5:627.

Brandeis D, Lehmann D. 1986. Event-related potentials of the brain and cognitive processes: Approaches and applications. Neuropsychologia. 24:151-168.

Brattico E, Tervaniemi M, Näätänen R, Peretz I. 2006. Musical scale properties are automatically processed in the human auditory cortex. Brain Res. 1117:162-174.

Brattico E, Tupala T, Glerean E, Tervaniemi M. 2013. Modulated neural processing of Western harmony in folk musicians. Psychophysiology. 50:653-663.

Britz J, Landis T, Michel CM. 2009. Right parietal brain activity precedes perceptual alternation of bistable stimuli. Cereb Cortex. 19:55-65.

Brown RM, Palmer C. 2013. Auditory and motor imagery modulate learning in music performance. Front Hum Neurosci. 7:1-13.

Brown S, Martinez MJ. 2007. Activation of premotor vocal areas during musical discrimination. Brain Cogn. 63:59-69.

Brunet D, Murray MM, Michel CM. 2011. Spatiotemporal analysis of multichannel EEG: CARTOOL. Comput Intell Neurosci. 2011:1-15.

Bubic A, Von Cramon DY, Jacobsen T, Schröger E, Schubotz RI. 2009. Violation of expectation: Neural correlates reflect bases of prediction. J Cogn Neurosci. 21:155-168.

Bubic A, Von Cramon DY, Schubotz RI. 2010. Prediction, cognition and the brain. Front Hum Neurosci. 4:1-15.

Bunce SC, Bernat E, Wong PS, Shevrin H. 1999. Further evidence for unconscious learning: Preliminary support for the conditioning of facial EMG to subliminal stimuli. J Psychiatr Res. 33:341-347.

Carrión RE, Bly BM. 2008. The effects of learning on event-related potential correlates of musical expectancy. Psychophysiology. 45:759-775.

Clark A. 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci. 36:181-204.

Cohen J. 1988. Statistical power analysis for the behavioral sciences. Lawrence

Earlbaum Associates.

Comerchero MD, Polich J. 1999. P3a and P3b from typical auditory and visual stimuli. Clin Neurophysiol. 110:24-30.

Costa-Faidella J, Grimm S, Slabu L, Díaz-Santaella F, Escera C. 2011. Multiple time scales of adaptation in the auditory system as revealed by human evoked potentials. Psychophysiology. 48:774-783.

Coulson S, King JW, Kutas M. 1998. Expect the unexpected: Event-related brain response to morphosyntactic violations. Lang Cogn Process. 13:21-58.

Creel SC. 2011. Specific previous experience affects perception of harmony and meter. J Exp Psychol Hum Percept Perform. 37:1512-1526.

Cuddy LL, Badertscher B. 1987. Recovery of the tonal hierarchy: Some comparisons across age and levels of musical experience. Percept Psychophys. 41:609-620.

Cuddy LL, Thompson WF. 1992. Asymmetry of perceived key movement in chorale sequences: Converging evidence from a probe-tone analysis. Psychol Res. 54:51-59.

Daikoku T, Yatomi Y, Yumoto M. 2014. Implicit and explicit statistical learning of tone sequences across spectral shifts. Neuropsychologia. 63:194-204.

D'Ausilio A, Altenmüller E, Olivetti Belardinelli M, Lotze M. 2006. Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. Eur J Neurosci. 24:955-958.

Declos A. 2014. The aesthetic and cognitive value of surprise. Proceedings of the European Society for Aesthetics. 6:52-69.

Delorme A, Makeig S. 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods. 134:9-21.

Dienes Z, Broadbent DE, Berry DC. 1991. Implicit and explicit knowledge bases in artificial grammar learning. J Exp Psychol Learn Mem Cogn. 17:875-887.

Dienes Z, Longuet-Higgins C. 2004. Can musical transformations be implicitly learned? Cogn Sci. 28:531-558.

Dienes Z, Scott R. 2005. Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. Psychol Res. 69:338-351.

Donchin E, Coles M. 1988. Is the P300 component a manifestation of context updating? Behav Brain Sci. 11:355-425.

Dowling WJ, Harwood DL. 1986. Music cognition. Academic Press.

Dowling WJ, Lung KM, Herrbold S. 1987. Aiming attention in pitch and time in the perception of interleaved melodies. Percept Psychophys. 41:642-656.

Drost UC, Rieger M, Brass M, Gunter TC, Prinz W. 2005a. Action-effect coupling in pianists. Psychol Res. 69:233-241.

Drost UC, Rieger M, Brass M, Gunter TC, Prinz W. 2005b. When hearing turns into playing: Movement induction by auditory stimuli in pianists. Q J Exp Psychol A. 58:1376-1389.

Eerola T, Toiviainen P. 2004. MIDI Toolbox: MATLAB Tools for Music Research. University of Jyväskylä: Kopijyvä, Jyväskylä, Finland.

Egermann H, Pearce MT, Wiggins GA, McAdams S. 2013. Probabilistic models of expectation violation predict psychophysiological emotional responses to live concert music. Cogn Affect Behav Neurosci. 13:533-553.

Ellison D, Moisseinen N, Fachner J. 2015. Affective versus cognitive responses to musical chords: An ERP and behavioral study. Psychomusicology: Music, Mind, and Brain. 25:423-434.

Elsner B, Hommel B. 2004. Contiguity and contingency in action-effect learning. Psychol Res. 68:138-154.

Elsner B, Hommel B. 2001. Effect anticipation and action control. J Exp Psychol Hum Percept Perform. 27:229-240.

Escera C, Alho K, Schröger E, Winkler I. 2000. Involuntary attention and distractibility as evaluated with event-related brain potentials. Audiol Neurootol. 5:151-166.

Fadiga L, Craighero L, D'Ausilio A. 2009. Broca's area in language, action, and music. Ann N Y Acad Sci. 1169:448-458.

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

Featherstone CR, Morrison CM, Waterman MG, MacGregor LJ. 2013. Semantics, syntax or neither? A case for resolution in the interpretation of N500 and P600 responses to harmonic incongruities. PloS ONE, 8:1-13.

Feldman H, Friston KJ. 2010. Attention, uncertainty, and free-energy. Front Hum Neurosci. 4:1-23.

Finney SA. 1997. Auditory feedback and musical keyboard performance. Music Percept. 15:153-174.

Finney SA. 2001a. FTAP: A Linux-based program for tapping and music experiments. Behav Res Methods Instrum Comput. 33:65-72.

Finney SA. 2001b. Real-time data collection in Linux: A case study. Behav Res Methods Instrum Comput. 33:167-173.

Finney SA, Palmer C. 2003. Auditory feedback and memory for music performance: Sound evidence for an encoding effect. Mem Cognit. 31:51-64.

Fitch WT, Martins MD. 2014. Hierarchical processing in music, language, and action: Lashley revisited. Ann N Y Acad Sci. 1316:87-104.

Fitzroy AB, Sanders LD. 2013. Musical expertise modulates early processing of syntactic violations in language. Front Psychol. 3:1-15.

Friedman D, Cycowicz YM, Gaeta H. 2001. The novelty P3: An event-related brain potential (ERP) sign of the brain's evaluation of novelty. Neurosci Biobehav Rev. 25:355-373.

Friedman D, Kazmerski VA, Cycowicz YM. 1998. Effects of aging on the novelty P3 during attend and ignore oddball tasks. Psychophysiology. 35:508-520.

Friston KJ. 2005. A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci. 360:815-836.

Friston KJ. 2009. The free-energy principle: A rough guide to the brain? Trends Cogn Sci. 13:293-301.

Friston KJ. 2010. The free-energy principle: A unified brain theory? Nat Rev Neurosci. 11:127-138.

Friston KJ, Kiebel S. 2009a. Cortical circuits for perceptual inference. Neural Netw. 22:1093-1104.

Friston KJ, Kiebel S. 2009b. Predictive coding under the free-energy principle. Philos.

Trans. R. Soc. Lond., B, Biol. Sci. 364:1211-1221.

Friston KJ, Kilner J, Harrison L. 2006. A free energy principle for the brain. J. Physiol. Paris. 100:70-87.

Friston KJ, Daunizeau J, Kilner J, Kiebel SJ. 2010. Action and behavior: A free-energy formulation. Biol Cybern. 102:227-260.

Furl N, Kumar S, Alter K, Durrant S, Shawe-Taylor J, Griffiths TD. 2011. Neural prediction of higher-order auditory sequence statistics. Neuroimage. 54:2267-2277.

Gaab N, Schlaug G. 2003. The effect of musicianship on pitch memory in performance matched groups. Neuroreport. 14:2291-2295.

Garza Villarreal EA, Brattico E, Leino S, Ostergaard L, Vuust P. 2011. Distinct neural responses to chord violations: A multiple source analysis study. Brain Res. 1389:103-114.

Gaser C, Schlaug G. 2003. Brain structures differ between musicians and non-musicians. J Neurosci. 23:9240-9245.

Geisler WS, Kersten D. 2002. Illusions, perception and Bayes. Nature neuroscience. 5:508-510.

Gunter TC, Stowe LA, Mulder G. 1997. When syntax meets semantics. Psychophysiology. 34:660-676.

Guo S, Koelsch S. 2016. Effects of veridical expectations on syntax processing in music: Event-related potential evidence. Sci Rep. 6:1-11.

Guo S, Koelsch S. 2015. The effects of supervised learning on event-related potential correlates of music-syntactic processing. Brain Res. 1626:232-246.

Haenschel C, Vernon DJ, Dwivedi P, Gruzelier JH, Baldeweg T. 2005. Event-related brain potential correlates of human auditory sensory memory-trace formation. J Neurosci. 25:10494-10501.

Halpern AR, Zatorre RJ. 1999. When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. Cereb Cortex. 9:697-704.

Halwani GF, Loui P, Rüber, T, Schlaug G. 2011. Effects of practice and experience on the arcuate fasciculus: Comparing singers, instrumentalists, and non-musicians.

Front Psychol. 2:1-9.

Hargreaves DJ. 1987. Verbal and behavioral responses to familiar and unfamiliar music. Curr Psychol Res Rev. 6:323-330.

Haslinger B, Erhard P, Altenmüller E, Schroeder U, Boecker H, Ceballos-Baumann AO. 2005. Transmodal sensorimotor networks during action observation in professional pianists. J Cogn Neurosci. 17:282-293.

Haueisen J, Knösche TR. 2001. Involuntary motor activity in pianists evoked by music perception. J Cogn Neurosci. 13:786-792.

Heinke W, Kenntner R, Gunter TC, Sammler D, Olthoff D, Koelsch S. 2004. Sequential effects of increasing propofol sedation on frontal and temporal cortices as indexed by auditory event-related potentials. Anesthesiology. 100:617-625.

Heyduk RG. 1975. Rated preference for musical compositions as it relates to complexity and exposure frequency. Percept Psychophys. 17:84-90.

Hoen M, Pachot-Clouard M, Segebarth C, Dominey PF. 2006. When Broca experiences the Janus syndrome: An ER-fMRI study comparing sentence comprehension and cognitive sequence processing. Cortex. 42:605-623.

Hohwy J. 2013. The predictive mind. Oxford University Press.

Hommel B, Müsseler J, Aschersleben G, Prinz W. 2001. The Theory of Event Coding (TEC): A framework for perception and action planning. Behav Brain Sci. 24:849-878.

Horváth J, Bendixen A. 2012. Preventing distraction by probabilistic cueing. Int J Psychophysiol. 83:342-347.

Horváth J, Sussman E, Winkler I, Schröger E. 2011. Preventing distraction: Assessing stimulus-specific and general effects of the predictive cueing of deviant auditory events. Biol Psychol. 87:35-48.

Hughes G, Desantis A, Waszak F. 2013. Attenuation of auditory N1 results from identity-specific action-effect prediction. Eur J Neurosci. 37:1152-1158.

Hughes HC, Darcey TM, Barkan HI, Williamson PD, Roberts DW, Aslin CH. 2001. Responses of human auditory association cortex to the omission of an expected acoustic event. Neuroimage. 13:1073-1089.

Huron D. 2006. Sweet anticipation: Music and the psychology of expectation. MIT press.

Hutchins S, Palmer C. 2008. Repetition priming in music. J Exp Psychol Hum Percept Perform. 34:693-707.

Hyde KL, Zatorre RJ, Peretz I. 2011. Functional MRI evidence of an abnormal neural network for pitch processing in congenital amusia. Cereb Cortex. 21:292-299.

Jackendoff R. 1991. Musical parsing and musical affect. Music Percept. 9:199-229.

Jackendoff R. 2009. Parallels and nonparallels between language and music. Music Percept. 26:195-204.

Jäncke L. 2009. Music drives brain plasticity. F1000 Biol. Rep:1-6.

Jäncke L. 2012. The dynamic audio-motor system in pianists. Ann N Y Acad Sci. 1252:246-252.

James CE, Britz J, Vuilleumier P, Hauert CA, Michel CM. 2008. Early neuronal responses in right limbic structures mediate harmony incongruity processing in musical experts. Neuroimage. 42:1597-1608.

James W. 1890. The principles of psychology. Holt and company.

Janata P. 1995. ERP measures assay the degree of expectancy violation of harmonic contexts in music. J Cogn Neurosci. 7:153-164.

Janata P, Birk JL, Van Horn JD, Leman M, Tillmann B, Bharucha JJ. 2002. The cortical topography of tonal structures underlying Western music. Science. 298:2167-2170.

Jaśkiewicz M, Francuz P, Zabielska-Mendyk E, Zapała D, Augustynowicz P. 2016. Effects of harmonics on aesthetic judgments of music: An ERP study involving laypersons and experts. Acta Neurobiol Exp (Wars). 76:142-151.

Jebb AT, Pfordresher PQ. 2016. Exploring perception-action relations in music production: The asymmetric effect of tonal class. J Exp Psychol Hum Percept Perform. 42:1-13.

Jentschke S, Friederici AD, Koelsch S. 2014. Neural correlates of music-syntactic processing in two-year old children. Dev Cogn Neurosci. 9:200-208.

Jentschke S, Koelsch S. 2009. Musical training modulates the development of syntax processing in children. Neuroimage. 47:735-744.

Jentschke S, Koelsch S, Sallat S, Friederici AD. 2008. Children with specific language impairment also show impairment of music-syntactic processing. J Cogn Neurosci. 20:1940-1951.

Johnson R. 1988. The amplitude of the P300 component of the event-related potential: Review and synthesis. Adv. Psychophys. 3:69-137.

Jones A, Hughes G, Waszak F. 2013. The interaction between attention and motor prediction. An ERP study. Neuroimage. 83:533-541.

Jones MR, Boltz M. 1989. Dynamic attending and responses to time. Psychol Rev. 96:459-491.

Justus TC, Bharucha JJ. 2001. Modularity in musical processing: The automaticity of harmonic priming. J Exp Psychol Hum Percept Perform. 27:1000-1011.

Kalda T, Minati L. 2012. Detecting scale violations in absence of mismatch requires music-syntactic analysis: A further look at the early right anterior negativity (ERAN). Brain Topogr. 25:285-292.

Katayama J, Polich J. 1998. Stimulus context determines P3a and P3b. Psychophysiology. 35:23-33.

Kathmann N, Frodl-Bauch T, Hegerl U. 1999. Stability of the mismatch negativity under different stimulus and attention conditions. Clin Neurophysiol. 110:317-323.

Keele SW, Ivry R, Mayr U, Hazeltine E, Heuer H. 2003. The cognitive and neural architecture of sequence representation. Psychol Rev. 110:316-339.

Keller PE. 2012. Mental imagery in music performance: Underlying mechanisms and potential benefits. Ann N Y Acad Sci. 1252:206-213.

Keller PE, Dalla Bella S, Koch I. 2010. Auditory imagery shapes movement timing and kinematics: Evidence from a musical task. J Exp Psychol Hum Percept Perform. 36:508-513.

Keller PE, Koch I. 2008. Action planning in sequential skills: Relations to music performance. Q J Exp Psychol (Hove). 61:275-291.

Kilner JM, Friston KJ, Frith CD. 2007. Predictive coding: An account of the mirror neuron system. Cogn Process. 8:159-166.

Kim SG, Kim JS, Chung CK. 2011. The effect of conditional probability of chord progression on brain response: An MEG study. PLoS ONE. 6:1-9.

Knill DC, Pouget A. 2004. The Bayesian brain: The role of uncertainty in neural coding and computation. Trends Neurosci. 27:712-719.

Knuf L, Aschersleben G, Prinz W. 2001. An analysis of ideomotor action. J Exp Psychol Gen. 130:779-798.

Koch I, Keller P, Prinz W. 2004. The ideomotor approach to action control: Implications for skilled performance. Int J Sports Psychol. 2:362-375.

Koelsch S. 2012. Brain and music. Wiley-Blackwell.

Koelsch S. 2014. Brain correlates of music-evoked emotions. Nat Rev Neurosci. 15:170-180.

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

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, Friederici AD. 2003. Toward the neural basis of processing structure in music. Comparative results of different neurophysiological investigation methods. Ann N Y Acad Sci. 999:15-28.

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

Koelsch S, Grossmann T, Gunter TC, Hahne A, Schröger E, Friederici AD. 2003a. Children processing music: Electric brain responses reveal musical competence and gender differences. J Cogn Neurosci. 15:683-693.

Koelsch S, Gunter T, Friederici AD, Schröger E. 2000. Brain indices of music processing: "nonmusicians" are musical. J Cogn Neurosci. 12:520-541.

Koelsch S, Gunter TC, Schröger E, Tervaniemi M, Sammler D, Friederici AD. 2001. Differentiating ERAN and MMN: An ERP study. Neuroreport. 12:1385-1389.

Koelsch S, Gunter TC, Von Cramon DY, Zysset S, Lohmann G, Friederici AD. 2002a. Bach speaks: A cortical "language-network" serves the processing of music. Neuroimage. 17:956-966.

Koelsch S, Heinke W, Sammler D, Olthoff D. 2006. Auditory processing during deep propofol sedation and recovery from unconsciousness. Clin Neurophysiol. 117:1746-1759.

Koelsch S, Jentschke S. 2008. Short-term effects of processing musical syntax: An ERP study. Brain Res. 1212:55-62.

Koelsch S, Jentschke S, Sammler D, Mietchen D. 2007. Untangling syntactic and sensory processing: An ERP study of music perception. Psychophysiology. 44:476-490.

Koelsch S, Kilches S, Steinbeis N, Schelinski S. 2008. Effects of unexpected chords and of performer's expression on brain responses and electrodermal activity. PLoS ONE. 3:1-10.

Koelsch S, Maess B, Grossmann T, Friederici AD. 2003b. Electric brain responses reveal gender differences in music processing. NeuroReport. 14:709-713.

Koelsch S, Mulder J. 2002. Electric brain responses to inappropriate harmonies during listening to expressive music. Clin Neurophysiol. 113:862-869.

Koelsch S, Rohrmeier M, Torrecuso R, Jentschke S. 2013. Processing of hierarchical syntactic structure in music. Proc Natl Acad Sci. U.S.A. 110:15443-15448.

Koelsch S, Sammler D. 2008. Cognitive components of regularity processing in the auditory domain. PLoS ONE. 3:1-7.

Koelsch S, Schmidt BH, Kansok J. 2002b. Effects of musical expertise on the early right anterior negativity: An event-related brain potential study. Psychophysiology. 39:657-663.

Koelsch S, Schröger E, Gunter TC. 2002c. Music matters: Preattentive musicality of the human brain. Psychophysiology. 39:38-48.

Koelsch S, Siebel WA. 2005. Towards a neural basis of music perception. Trends Cogn Sci. 9:578-584.

Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V, Rizzolatti G. 2002. Hearing sounds, understanding actions: Action representation in mirror neurons. Science. 297:846-848.

Kristeva R, Chakarov V, Schulte-Mönting J, Spreer J. 2003. Activation of cortical areas in music execution and imagining: A high-resolution EEG study. Neuroimage. 20:1872-1883.

Krumhansl CL. 1995. Music psychology and music theory: Problems and prospects. Music Theory Spectrum. 17:53-90.

Krumhansl CL. 2000. Rhythm and pitch in music cognition. Psychol Bull. 126:159-179.

Krumhansl CL, Louhivuori J, Toiviainen P. 1999. Melodic expectation in Finnish spiritual folk hymns: Convergence of statistical, behavioral, and computational approaches. Music Percept. 17:151-195.

Krumhansl CL, Toivanen P, Eerola T, Toiviainen P, Järvinen T, Louhivuori J. 2000. Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. Cognition. 76:13-58.

Kuhn G, Dienes Z. 2005. Implicit learning of nonlocal musical rules: Implicitly learning more than chunks. J Exp Psychol Learn Mem Cogn. 31:1417-1432.

Kujala T, Tervaniemi M, Schröger E. 2007. The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. Biol Psychol. 74:1-19.

Kumar S, Sedley W, Nourski KV, Kawasaki H, Oya H, Patterson RD, Howard III MA, Friston KJ, Griffiths TD. 2011. Predictive coding and pitch processing in the auditory cortex. J Cogn Neurosci. 23:3084-3094.

Kunert R, Willems RM, Casasanto D, Patel AD, Hagoort P. 2015. Music and language syntax interact in Broca's area: An fMRI study. PLoS ONE. 10:1-16.

Kunert R, Willems RM, Hagoort P. 2016. Language influences music harmony perception: Effects of shared syntactic integration resources beyond attention. R Soc Open Sci. 3:1-29.

Lahav A, Boulanger A, Schlaug G, Saltzman E. 2005. The power of listening: Auditory-motor interactions in musical training. Ann N Y Acad Sci. 1060:189-194.

Lahav A, Saltzman E, Schlaug G. 2007. Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. J Neurosci. 27:308-314.

Lange K. 2013. The ups and downs of temporal orienting: A review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory N1 with opposite effects of attention and prediction. Front Hum Neurosci. 7:1-14.

Lappe C, Herholz SC, Trainor LJ, Pantev C. 2008. Cortical plasticity induced by short-term unimodal and multimodal musical training. J Neurosci. 28:9632-9639.

Lappe C, Trainor LJ, Herholz SC, Pantev C. 2011. Cortical plasticity induced by short-term multimodal musical rhythm training. PLoS ONE. 6:1-8.

Large EW. 1993. Dynamic programming for the analysis of serial behaviors. Behav Res Methods Instrum Comput. 25:238-241.

Large EW, Palmer C. 2002. Perceiving temporal regularity in music. Cogn Sci. 26:1-37.

Launay J, Dean RT, Bailes F. 2016. Rapid learning of associations between sound and action through observed movement. A TMS study. Psychomusicology. 26:35-42.

Leaver AM, Van Lare J, Zielinski B, Halpern AR, Rauschecker JP. 2009. Brain activation during anticipation of sound sequences. J Neurosci. 29:2477-2485.

Lehmann D, Ozaki H, Pal I. 1987. EEG alpha map series: Brain micro-states by space-oriented adaptive segmentation. Electroencephalogr Clin Neurophysiol. 67:271-288.

Leino S, Brattico E, Tervaniemi M, Vuust P. 2007. Representation of harmony rules in the human brain: Further evidence from event-related potentials. Brain Res. 1142:169-177.

Lerdahl F. 1991. Underlying musical schemata. Representing musical structure. 5:273.

Limb CJ. 2006. Structural and functional neural correlates of music perception. Anat Rec A Discov Mol Cell Evol Biol. 288:435-446.

List A, Justus T. 2010. Relative priming of temporal local-global levels in auditory hierarchical stimuli. Atten Percept Psychophys. 72:193-208.

Loehr JD. 2013. Sensory attenuation for jointly produced action effects. Front Psychol. 4:1-9.

Loui P. 2012. Learning and liking of melody and harmony: Further studies in artificial grammar learning. Top Cogn Sci. 4:554-567.

Loui P, Grent-'t-Jong T, Torpey D, Woldorff M. 2005. Effects of attention on the neural processing of harmonic syntax in Western music. Cogn Brain Res. 25:678-687.

Loui P, Li HC, Schlaug G. 2011. White matter integrity in right hemisphere predicts pitch-related grammar learning. Neuroimage. 55:500-507.

Loui P, Schlaug G. 2009. Investigating musical disorders with diffusion tensor imaging: A comparison of imaging parameters. Ann N Y Acad Sci. 1169:121-125.

Luck SJ. 2014. An introduction to the event-related potential technique. MIT press.

Maes PJ. 2016. Sensorimotor grounding of musical embodiment and the role of prediction: A review. Front Psychol. 7:1-10.

Maes PJ, Leman M, Palmer C, Wanderley MM. 2014. Action-based effects on music perception. Front Psychol. 4:1-14.

Maess B, Koelsch S, Gunter TC, Friederici AD. 2001. Musical syntax is processed in Broca's area: An MEG study. Nat Neurosci. 4:540-545.

Maffongelli L, Bartoli E, Sammler D, Koelsch S, Campus C, Olivier E, Fadiga L, D'Ausilio A. 2015. Distinct brain signatures of content and structure violation during action observation. Neuropsychologia, 75:30-39.

Maidhof C. 2013. Error monitoring in musicians. Front Hum Neurosci. 7:1-8.

Maidhof C, Kästner T, Makkonen T. 2014. Combining EEG, MIDI, and motion capture techniques for investigating musical performance. Behav Res Methods. 46:185-195.

Maidhof C, Koelsch S. 2011. Effects of selective attention on syntax processing in music and language. J Cogn Neurosci. 23:2252-2267.

Margulis EH. 2005. A model of melodic expectation. Music Percept. 22:663-714.

Marmel F, Perrin F, Tillmann B. 2011. Tonal expectations influence early pitch processing. J Cogn Neurosci. 23:3095-3104.

Marmel F, Tillmann B, Delbé C. 2010. Priming in melody perception: Tracking down the strength of cognitive expectations. J Exp Psychol Hum Percept Perform. 36:1016-1028.

Martin S, Mikutta C, Leonard MK, Hungate D, Koelsch S, Chang EF, Millan JdR, Knight RT, Pasley BN. 2017. Neural encoding of auditory features during music perception and imagery: Insight into the brain of a piano player. bioRxiv, 106617.

McAuley JD, Stevens C, Humphreys MS. 2004. Play it again: Did this melody occur more frequently or was it heard more recently? The role of stimulus familiarity in episodic recognition of music. Acta Psychol (Amst). 116:93-108.

McMullen E, Saffran JR. 2004. Music and language: A developmental comparison. Music Percept. 21:289-311.

Meyer LB. 1956. Emotion and meaning in music. Universityof Chicago Press.

Meyer LB. 1973. Explaining music: Essays and explorations. University of Chicago Press.

Meyer LB. 1961. On rehearing music. J A M S. 14:257-267.

Michel CM, Grave de Peralta R, Lantz G, Gonzalez Andino S, Spinelli L, Blanke O, Landis T, Seeck M. 1999a. Spatiotemporal EEG analysis and distributed source estimation in presurgical epilepsy evaluation. J Clin Neurophysiol. 16:239-266.

Michel CM, Murray MM, Lantz G, Gonzalez S, Spinelli L, Grave de Peralta R. 2004. EEG source imaging. Clin Neurophysiol. 115:2195-2222.

Michel CM, Seeck M, Landis T. 1999b. Spatiotemporal dynamics of human cognition. News Physiol Sci. 14:206-214.

Michel CM, Thut G, Morand S, Khateb A, Pegna AJ, Grave de Peralta R, Gonzalez S, Seeck M, Landis T. 2001. Electric source imaging of human brain functions. Brain Res Rev. 36:108-118.

Miles SA, Miranda RA, Ullman MT. 2016. Sex differences in music: A female advantage at recognizing familiar melodies. Front Psychol. 7:1-12.

Minati L, Rosazza C, D'Incerti L, Pietrocini E, Valentini L, Scaioli V, Loveday C, Bruzzone MG. 2008. FMRI/ERP of musical syntax: Comparison of melodies and unstructured note sequences. Neuroreport. 19:1381-1385.

Miranda RA, Ullman MT. 2007. Double dissociation between rules and memory in music: An event-related potential study. Neuroimage. 38:331-345.

Molinaro N, Carreiras M. 2010. Electrophysiological evidence of interaction between contextual expectation and semantic integration during the processing of collocations. Biol Psychol 83:176-190.

Morimoto S, Remijn GB, Nakajima Y. 2016. Computational-model-based analysis of context effects on harmonic expectancy. PLoS ONE. 11:1-17.

Muller-Gass A, Macdonald M, Schröger E, Sculthorpe L, Campbell K. 2007. Evidence for the auditory P3a reflecting an automatic process: Elicitation during highly-focused continuous visual attention. Brain Res. 1170:71-78.

Murray MM, Brunet D, Michel CM. 2008. Topographic ERP analyses: A step-by-step tutorial review. Brain Topogr. 20:249-264.

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.

Müller M, Höfel L, Brattico E, Jacobsen T. 2010. Aesthetic judgments of music in experts and laypersons -An ERP study. Int J Psychophysiol. 76:40-51.

Münte TF, Schuppert M, Johannes S, Wieringa BM, Kohlmetz C, Altenmüller E. 1998. Brain potentials in patients with music perception deficits: Evidence for an early locus. Neurosci Lett. 256:85-88.

Näätänen R, Gaillard AW, Mäntysalo S. 1978. Early selective-attention effect on evoked potential reinterpreted. Acta Psychol (Amst). 42:313-329.

Näätänen R, Kujala T, Winkler I. 2011. Auditory processing that leads to conscious perception: A unique window to central auditory processing opened by the mismatch negativity and related responses. Psychophysiology. 48:4-22.

Näätänen R, Simpson M, Loveless NE. 1982. Stimulus deviance and evoked potentials. Biol Psychol. 14:53-98.

Näätänen R, Tervaniemi M, Sussman E, Paavilainen P, Winkler I. 2001. 'Primitive intelligence' in the auditory cortex. Trends Neurosci. 24:283-288.

Näätänen R, Winkler I. 1999. The concept of auditory stimulus representation in cognitive neuroscience. Psychol Bull. 125:826-859.

Nazimek JM, Hunter MD, Hoskin R, Wilkinson I, Woodruff PW. 2013. Neural basis of auditory expectation within temporal cortex. Neuropsychologia. 51:2245-2250.

Nittono H, Bito T, Hayashi M, Sakata S, Hori T. 2000. Event-related potentials elicited by wrong terminal notes: Effects of temporal disruption. Biol Psychol. 52:1-16.

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

Novembre G, Keller PE. 2011. A grammar of action generates predictions in skilled musicians. Conscious Cogn. 20:1232-1243.

Oechslin MS, Gschwind M, James CE. 2017. Tracking training-related plasticity by combining fMRI and DTI: The right hemisphere ventral stream mediates musical syntax processing. Cereb Cortex. 1:1-10.

Ohnishi T, Matsuda H, Asada T, Aruga M, Hirakata M, Nishikawa M, Katoh A, Imabayashi E. 2001. Functional anatomy of musical perception in musicians. Cereb Cortex. 11:754-760.

Omigie D, Pearce MT, Stewart L. 2012. Tracking of pitch probabilities in congenital amusia. Neuropsychologia. 50:1483-1493.

Omigie D, Pearce MT, Williamson VJ, Stewart L. 2013. Electrophysiological correlates of melodic processing in congenital amusia. Neuropsychologia. 51:1749-1762.

Paavilainen P. 2013. The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: A review. Int J Psychophysiol. 88:109-123.

Paavilainen P, Simola J, Jaramillo M, Näätänen R, Winkler I. 2001. Preattentive extraction of abstract feature conjunctions from auditory stimulation as reflected by the mismatch negativity (MMN). Psychophysiology. 38:359-365.

Paavilainen P, Tiitinen H, Alho K, Näätänen R. 1993. Mismatch negativity to slight pitch changes outside strong attentional focus. Biol Psychol. 37:23-41.

Paller KA, McCarthy G, Wood CC. 1992. Event-related potentials elicited by deviant endings to melodies. Psychophysiology. 29:202-206.

Pascual-Marqui RD, Michel CM, Lehmann D. 1995. Segmentation of brain electrical activity into microstates: Model estimation and validation. IEEE Trans Biomed Eng. 42:658-665.

Patel AD, Gibson E, Ratner J, Besson M, Holcomb PJ. 1998. Processing syntactic relations in language and music: An event-related potential study. J Cogn Neurosci. 10:717-733.

Pearce MT, Ruiz MH, Kapasi S, Wiggins GA, Bhattacharya J. 2010. Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. Neuroimage. 50:302-313.

Peretz I. 1996. Can we lose memory for music? A case of music agnosia in a nonmusician. J Cogn Neurosci. 8:481-496.

Peretz I, Gaudreau D, Bonnel AM. 1998. Exposure effects on music preference and recognition. Mem Cognit. 26:884-902.

Pfordresher PQ. 2005. Auditory feedback in music performance: The role of melodic structure and musical skill. J Exp Psychol Hum Percept Perform. 31:1331-1345.

Pfordresher PQ. 2012. Musical training and the role of auditory feedback during performance. Ann N Y Acad Sci. 1252:171-178.

Pfordresher PQ, Palmer C. 2006. Effects of hearing the past, present, or future during music performance. Percept Psychophys. 68:362-376.

Phillips-Silver J, Trainor LJ. 2005. Feeling the beat: Movement influences infant rhythm perception. Science. 308:1430.

Phillips-Silver J, Trainor LJ. 2007. Hearing what the body feels: Auditory encoding of rhythmic movement. Cognition. 105:533-546.

Pieszek M, Widmann A, Gruber T, Schröger E. 2013. The human brain maintains contradictory and redundant auditory sensory predictions. PLoS ONE. 8:1-13.

Polich J. 1988. Bifurcated P300 peaks: P3a and P3b revisited? J Clin Neurophysiol. 5:287-294.

Polich J. 2007. Updating P300: An integrative theory of P3a and P3b. Clin Neurophysiol. 118:2128-2148.

Poulin-Charronnat B, Bigand E, Koelsch S. 2006. Processing of musical syntax tonic versus subdominant: An event-related potential study. J Cogn Neurosci.

18:1545-1554.

Prinz W. 1990. A common coding approach to perception and action. Relationships between perception and action. Springer:167-201.

Prinz W. 1997. Perception and action planning. Eur J Cogn Psychol. 7:129-154.

Pulvermüller F. 2014. The syntax of action. Trends Cogn Sci. 18:219-220.

Reber AS. 1967. Implicit learning of artificial grammar. J Verb Learn Verb Beha. 6:855-863.

Regnault P, Bigand E, Besson M. 2001. Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related brain potentials. J Cogn Neurosci. 13:241-255.

Repp BH. 1999. Effects of auditory feedback deprivationon on expressive piano performance. Music Percept. 16:409-438.

Repp BH, Knoblich G. 2007. Action can affect auditory perception. Psychol Sci. 18:6-7.

Rinne T, Antila S, Winkler I. 2001. Mismatch negativity is unaffected by top-down predictive information. Neuroreport. 12:2209-2213.

Ritter W, Sussman E, Deacon D, Cowan N, Vaughan HG. 1999. Two cognitive systems simultaneously prepared for opposite events. Psychophysiology. 36:835-838.

Rizzolatti G. 2005. The mirror neuron system and its function in humans. Anat Embryol. 210:419-421.

Roehm D, Bornkessel-Schlesewsky I, Rösler F, Schlesewsky M. 2007. To predict or not to predict: Influences of task and strategy on the processing of semantic relations. J Cogn Neurosci. 19:1259-1274.

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

Rohrmeier M, Rebuschat P, Cross I. 2011. Incidental and online learning of melodic structure. Conscious Cogn. 20:214-222.

Rohrmeier MA, Koelsch S. 2012. Predictive information processing in music cognition. A critical review. Int J Psychophysiol. 83:164-175.

Rugg MD, Curran T. 2007. Event-related potentials and recognition memory. Trends Cogn Sci. 11:251-257.

Ruiz MH, Jabusch HC, Altenmüller E. 2009. Detecting wrong notes in advance: Neuronal correlates of error monitoring in pianists. Cereb Cortex. 19:2625-2639.

Rüsseler J, Rösler F. 2000. Implicit and explicit learning of event sequences: Evidence for distinct coding of perceptual and motor representations. Acta Psychol (Amst). 104:45-67.

Salimpoor VN, Benovoy M, Larcher K, Dagher A, Zatorre RJ. 2011. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. Nat Neurosci. 14:257-262.

Salimpoor VN, Zald DH, Zatorre RJ, Dagher A. 2015. Predictions and the brain: How musical sounds become rewarding. Trends Cogn Sci. 19:86-91.

Sammler D, Koelsch S, Ball T, Brandt A, Grigutsch M, Huppertz HJ, Knösche TR, Wellmer J, Widman G, Elger CE, et al. 2013a. Co-localizing linguistic and musical syntax with intracranial EEG. Neuroimage. 64:134-146.

Sammler D, Koelsch S, Friederici AD. 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 PE. 2013b. Syntax in a pianist's hand: ERP signatures of "embodied" syntax processing in music. Cortex. 49:1325-1339.

Sams M, Paavilainen P, Alho K, Näätänen R. 1985. Auditory frequency discrimination and event-related potentials. Electroencephalogr Clin Neurophysiol. 62:437-448.

SanMiguel I, Todd J, Schröger E. 2013. Sensory suppression effects to self-initiated sounds reflect the attenuation of the unspecific N1 component of the auditory ERP. Psychophysiology. 50:334-343.

Schellenberg EG. 1996. Expectancy in melody: Tests of the implication-realization model. Cognition. 58:75-125.

Schellenberg EG, Bigand E, Poulin-Charronnat B, Garnier C, Stevens C. 2005. Children's implicit knowledge of harmony in Western music. Dev Sci. 8:551-566.

Scherg M, Vajsar J, Picton TW. 1989. A source analysis of the late human auditory evoked potentials. J Cogn Neurosci. 1:336-355.

Schmuckler M. 1997. Expectancy effects in memory for melodies. Can J Exp Psychol. 51:292-305.

Schmuckler MA. 1989. Expectation in music: Investigation of melodic and harmonic processes. Music Percept. 7:109-150.

Schmuckler MA, Boltz MG. 1994. Harmonic and rhythmic influences on musical expectancy. Percept Psychophys. 56:313-325.

Schneider P, Scherg M, Dosch HG, Specht HJ, Gutschalk A, Rupp A. 2002. Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. Nat Neurosci. 5:688-694.

Schön D, Besson M. 2005. Visually induced auditory expectancy in music reading: A behavioral and electrophysiological study. J Cogn Neurosci. 17:694-705.

Schröger E. 1998. Measurement and interpretation of the mismatch negativity. Behav Res Methods Instrum Comput. 30:131-145.

Schröger E, Bendixen A, Denham SL, Mill RW, Bőhm TM, Winkler I. 2014. Predictive regularity representations in violation detection and auditory stream segregation: From conceptual to computational models. Brain Topogr. 27:565-577.

Schröger E, Kotz SA, SanMiguel I. 2015a. Bridging prediction and attention in current research on perception and action. Brain Res. 1626:1-13.

Schröger E, Marzecová A, SanMiguel I. 2015b. Attention and prediction in human audition: A lesson from cognitive psychophysiology. Eur J Neurosci. 41:641-664.

Schürmann M, Raij T, Fujiki N, Hari R. 2002. Mind's ear in a musician: Where and when in the brain. Neuroimage. 16:434-440.

Schwartze M, Tavano A, Schröger E, Kotz SA. 2012. Temporal aspects of prediction in audition: Cortical and subcortical neural mechanisms. Int J Psychophysiol. 83:200-207.

Seger CA. 1994. Implicit learning. Psychol Bull. 115:163-196.

Sheskin D. 2004. Handbook of parametric and nonparametric statistical procedures. CRC Press.

Sommer W, Leuthold H, Matt J. 1998. The expectancies that govern the P300 amplitude are mostly automatic and unconscious. Behav Brain Sci. 21:149-150.

Steinbeis N, Koelsch S, Sloboda JA. 2006. The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. J Cogn Neurosci. 18:1380-1393.

Stock A, Stock C. 2004. A short history of ideo-motor action. Psychol Res. 68:176-188.

Summerfield C, Egner T. 2009. Expectation (and attention) in visual cognition. Trends Cogn Sci. 13:403-409.

Sussman E, Winkler I, Schröger E. 2003. Top-down control over involuntary attention switching in the auditory modality. Psychon Bull Rev. 10:630-637.

Tavano A, Widmann A, Bendixen A, Trujillo-Barreto N, Schröger E. 2014. Temporal regularity facilitates higher-order sensory predictions in fast auditory sequences. Eur J Neurosci. 39:308-318.

Tekman HG, Bharucha JJ. 1998. Implicit knowledge versus psychoacoustic similarity in priming of chords. J Exp Psychol Hum Percept Perform. 24:252-260.

Tervaniemi M, Huotilainen M. 2003. The promises of change-related brain potentials in cognitive neuroscience of music. Ann N Y Acad Sci. 999:29-39.

Tervaniemi M, Tupala T, Brattico E. 2012. Expertise in folk music alters the brain processing of Western harmony. Ann N Y Acad Sci. 1252:147-51.

Thompson F, Bella D, Keller E. 2006. Music Performance. Adv Cogn Psychol. 2:99-102.

Tillmann B. 2005. Implicit investigations of tonal knowledge in nonmusician listeners. Ann N Y Acad Sci. 1060:100-110.

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

Tillmann B, Bharucha JJ. 2002. Effect of harmonic relatedness on the detection of temporal asynchronies. Percept Psychophys. 64:640-649.

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

Tillmann B, Bigand E. 2004. Musical priming: Schematic expectations resist repetition priming. Proceedings of the 8th International Conference:674-676.

Tillmann B, Bigand E. 2010. Musical structure processing after repeated listening: Schematic expectations resist veridical expectations. Music Sci. 14:33-47.

Tillmann B, Janata P, Bharucha JJ. 2003a. Activation of the inferior frontal cortex in musical priming. Cogn Brain Res. 16:145-161.

Tillmann B, Janata P, Birk J, Bharucha JJ. 2003b. The costs and benefits of tonal centers for chord processing. J Exp Psychol Hum Percept Perform. 29:470-482.

Tillmann B, Koelsch S, Escoffier N, Bigand E, Lalitte P, Friederici AD, Von Cramon DY. 2006. Cognitive priming in sung and instrumental music: Activation of inferior frontal cortex. Neuroimage. 31:1771-82.

Tillmann B, Lebrun-Guillaud G. 2006. Influence of tonal and temporal expectations on chord processing and on completion judgments of chord sequences. Psychol Res. 70:345-358.

Tillmann B, Marmel F. 2013. Musical expectations within chord sequences: Facilitation due to tonal stability without closure effects. Psychomusicol Music Mind Brain. 23:1-5.

Tillmann B, Poulin-Charronnat B. 2010. Auditory expectations for newly acquired structures. Q J Exp Psychol (Hove). 63:1646-1664.

Tillmann B, Poulin-Charronnat B, Bigand E. 2014. The role of expectation in music: From the score to emotions and the brain. Wiley Interdiscip Rev Cogn Sci. 5:105-113.

Timm J, Schönwiesner M, Schröger E, SanMiguel I. 2016. Sensory suppression of brain responses to self-generated sounds is observed with and without the perception of agency. Cortex:1-16.

Trainor L. 2008. The neural roots of music. Nature. 453:598-599.

Trainor LJ. 2005. Are there critical periods for musical development? Dev Psychobiol. 46:262-278.

Trainor LJ. 2012. Predictive information processing is a fundamental learning mechanism present in early development: Evidence from infants. Int J Psychophysiol. 83:256-258.

Tramo MJ, Cariani PA, Delgutte B, Braida LD. 2001. Neurobiological foundations for the theory of harmony in western tonal music. Ann N Y Acad Sci. 930:92-116.

Van Den Bosch I, Salimpoor VN, Zatorre RJ. 2013. Familiarity mediates the relationship between emotional arousal and pleasure during music listening.

Front Hum Neurosci. 7:534.

Vassena E, Kochman K, Latomme J, Verguts T. 2016. Unimodal and cross-modal prediction is enhanced in musicians. Sci Rep. 6:1-7.

Verleger R. 1990. P3-evoking wrong notes: Unexpected, awaited, or arousing? Int J Neurosci. 55:171-179.

Vespignani F, Canal P, Molinaro N, Fonda S, Cacciari C. 2010. Predictive mechanisms in idiom comprehension. J Cogn Neurosci. 22:1682-1700.

Vuust P, Ostergaard L, Pallesen KJ, Bailey C, Roepstorff A. 2009. Predictive coding of music-Brain responses to rhythmic incongruity. Cortex. 45:80-92.

Vuust P, Witek MA. 2014. Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. Front Psychol. 5:1-14.

Wacongne C, Changeux JP, Dehaene S. 2012. A neuronal model of predictive coding accounting for the mismatch negativity. J Neurosci. 32:3665-3678.

Wacongne C, Labyt E, van Wassenhove V, Bekinschtein T, Naccache L, Dehaene S. 2011. Evidence for a hierarchy of predictions and prediction errors in human cortex. Proc Natl Acad Sci. U.S.A. 108:20754-20759.

Wetzel N, Schröger E. 2007. Cognitive control of involuntary attention and distraction in children and adolescents. Brain Res. 1155:134-46.

Widmann A, Kujala T, Tervaniemi M, Kujala A, Schröger E. 2004. From symbols to sounds: Visual symbolic information activates sound representations. Psychophysiology. 41:709-715.

Winkler I. 2007. Interpreting the mismatch negativity. J Psychophysiol. 21:147-163.

Winkler I, Czigler I. 2012. Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. Int J Psychophysiol. 83:132-143.

Woldorff MG, Hackley SA, Hillyard SA. 1991. The effects of channel-selective attention on the mismatch negativity wave elicited by deviant tones. Psychophysiology. 28:30-42.

Woldorff MG, Hillyard SA, Gallen CC, Hampson SR, Bloom FE. 1998.

Magnetoencephalographic recordings demonstrate attentional modulation of mismatch-related neural activity in human auditory cortex. Psychophysiology. 35:283-292.

Wolpert DM, Ghahramani Z. 2000. Computational principles of movement neuroscience. Nat Neurosci. 3:1212-1217.

Wolpert DM, Ghahramani Z, Jordan MI. 1995. An internal model for sensorimotor integration. Science. 269:1880-1882.

Wong PS, Bernat E, Bunce S, Shevrin H. 1997. Brain indices of nonconscious associative learning. Conscious Cogn. 6:519-544.

Wong PS, Bernat E, Snodgrass M, Shevrin, H. 2004. Event-related brain correlates of associative learning without awareness. Int J Psychophysiol. 53:217-231.

Zatorre RJ, Chen JL, Penhune VB. 2007. When the brain plays music: Auditory-motor interactions in music perception and production. Nat Rev Neurosci. 8:547-558.

Zatorre RJ, Perry DW, Beckett CA, Westbury CF, Evans AC. 1998. Functional anatomy of musical processing in listeners with absolute pitch and relative pitch. Proc Natl Acad Sci. U.S.A. 95:3172-3177.

Zendel BR, Lagrois MÉ, Robitaille N, Peretz I. 2015. Attending to pitch information inhibits processing of pitch information: The curious case of amusia. J Neurosci. 35:3815-3824.

Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig und ohne unzulässige Hilfe verfasst habe. Für die im Rahmen der Promotion durchgeführten Studien und die Erstellung der Fachartikel war ich hauptverantwortlich. Die Arbeit ist in keinem früheren Promotionsverfahren angenommen oder abgelehnt worden.

Berlin, den 30. März 2017

Shuang Guo