Fachbereich Erziehungswissenschaft und Psychologie
der Freien Universität Berlin

Neurophysiological correlates of action monitoring during music performance

Dissertation
zur Erlangung des akademischen Grades
Doktor der Philosophie (Dr. phil.)

vorgelegt von

Magister Artium
Maidhof, Clemens

Berlin, 2015

Erstgutachter

Prof. Dr. Stefan Koelsch

Zweitgutachter

Prof. Dr. Dirk Ostwald

Tag der Disputation: 4. März 2016

List of original publications

This dissertation is based on the following publications:

1. *Experiment 1 and 2*:

   Maidhof, C., Vavatzanidis, N., Prinz, W., Rieger, M., & Koelsch, S. (2010). Processing expectancy violations during music performance and perception: an ERP study. *Journal of Cognitive Neuroscience*, *22*(10), 2401-2413. doi:10.1162/jocn.2009.21332

2. *(Re-analysis of) Experiment 2:*

   Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2009). Nobody is perfect: ERP effects prior to performance errors in musicians indicate fast monitoring processes. *PloS ONE*, *4*(4), 1-7. doi:10.1371/journal.pone.0005032

3. *Experiment 3:*

   Maidhof, C., Pitkäniemi, A., & Tervaniemi, M. (2013). Predictive error detection in pianists: a combined ERP and motion capture study. *Frontiers in Human Neuroscience*, *7*, 1-14. doi:10.3389/fnhum.2013.00587

4. *Methodological work:*

   Maidhof, C., Kästner, T., & Makkonen, T. (2014). Combining EEG, MIDI, and motion capture techniques for investigating musical performance. *Behavior Research Methods*, *46*(1), 185-195. doi:10.3758/s13428-013-0363-9

Contents

1. Introduction

1.1. General Overview

Music performance is regarded as one of the most complex human activities and a challenging endeavor (Zatorre, Chen, & Penhune, 2007). Even when performing a simple melody like *Happy Birthday*, numerous cognitive and motor processes interact and overlap in time. First, a performer has to retrieve musical structures and units, including notes, their relative timing, duration, and intensity, according to some interpretation from memory[1]. Then, these elements have to be planned and the appropriate actions have to be executed.

The generation of these musical actions, i.e., movements directed towards the goal of producing structured sound, results in different forms of feedback such as auditory—arguably the ultimate goal of musical actions—, tactile, proprioceptive, and visual feedback. A fundamental aspect of music making, as well as of any other goal-directed behavior, is highly accurate motor control and ensuring that the intended goals are or will be achieved. For that, musicians have to constantly monitor their performed actions and the resulting sensory effects in order to detect deviations from the intended behavior, for example, when a wrong key on the piano is accidentally pressed or when the intended outcome in terms of sound is not achieved. Whenever there are indications that the behavior is or will not be appropriate and will not lead to the desired effect, the behavior can be adjusted and corrections be initialized.

How do musicians, and more specifically pianists, plan, execute, and monitor their complex, fast, fine-grained, and highly-trained actions and their outcomes during

---

[1] Note that performing music puts high demands on memory, often requiring the production of several thousands of notes, distributed over several tens of minutes. Thus, memory for music performance is considered as a form of *expert memory* with specialized principles for encoding and retrieving information from memory (see Ericsson & Kintsch, 1995; Chaffin & Imreh, 1997, 2002; Chaffin & Logan, 2006).

performance? How do pianists process the auditory and tactile feedback of their keystrokes, and what role does action and perception play for the processing of auditory feedback? How do pianists detect that their musical actions and associated outcomes did not go (or will not go) the way they intended them to go? The present thesis is an attempt of providing some answers to these questions by investigating the neurophysiological correlates of action monitoring processes during music performance.

Even though professional pianists have spent several thousands of hours of deliberate practice (Ericsson, Krampe, & Tesch-Römer, 1993; Sloboda, 2000)[2], they commit errors—defined as the unintended result of an action—occasionally. From the perspective of the performer, errors are usually something that has to be avoided. However, during music learning, errors provide an important source of information. Only if we are aware of our errors, e.g., when we realize that a keystroke sounded wrong (whether in terms of pitch, timing, or intensity), we can learn and try to prevent them in the future, and taking risks and making errors is regarded as an essential part of the learning process and performance (Kruse-Weber & Parncutt, 2014).

From the perspective of the researcher, investigations of errors during skilled human behavior have revealed important insights into planning and execution of complex tasks such as speech, typing, or music performance (for the music domain, see e.g., Palmer & Van de Sande, 1993, 1995; Palmer, 1997, 2005; Palmer & Pfordresher, 2003; Pfordresher, 2006). Early investigations of errors during typing showed for example that people can detect their own errors immediately, possibly before the results of an erroneous action can be perceived (Rabbitt, 1978). Additionally, participants showed slower responses on trials following errors (*post-error slowing*), suggesting that they adjusted their behavioral strategies after error commissions. Consequently, it was assumed that the functioning of an action-monitoring system is responsible for the detection of errors and initiation of performance adjustments when necessary.

In the next sections, I will briefly introduce central concepts from the motor control

---

[2] Deliberate practice refers to a structured, goal-oriented, and effortful activity to improve a specific skill like sight-reading, improvising, memorizing, etc. In addition, high motivation, attention, instruction, close monitoring of skill improvement and supervision by a teacher are characteristics of deliberate practice (see Ericsson, Krampe, & Tesch-Römer, 1993).

literature. Then, accounts of action monitoring and error processing will be summarized, followed by a brief review of literature concerned with how sensory information during music performance is used. At the end of this chapter, I will give an overview of electroencephalography (EEG) and the event-related technique as the primary research method used for this thesis. In the following two chapters, I will summarize the conducted experiments and draw some conclusions as well as discuss avenues for future research.

1.2. Goal-directed Behavior

1.2.1. Structure of the Human Motor System

Movements, which are a necessary but not sufficient condition for actions (which imply some sort of goal or intention), can be performed with the help of approximately 600 muscles controlling around 200 joints. Several neural structures, which can be divided in 4 interactive subsystems, are involved in controlling our movements and voluntary actions (see Purves et al., 2004).

(1) The *local circuitry within the gray matter of the spinal cord and brainstem*, consisting of local circuit neurons and lower motor neurons that form the final pathway for translating neural information into movement. Lower motor neurons directly innervate muscles fibers, forming a *motor unit* (Rosenbaum, 2009). Their major synaptic input is from interneurons, which receive projections from higher motor centers as well as sensory information from receptors within the muscles, joints, and skin. (2) *Descending modulatory pathways* consist of upper motor neurons in motor regions of the cerebral cortex and in the brainstem, from where thick myelinated axons descend in different pathways to lower motor neurons. Cortical motor regions consist mainly of Brodmann's area (BA) 4 and 6 and can be divided into primary motor cortex (M1) and premotor areas consisting of ventral and dorsal premotor areas, supplementary (SMA) and cingulate motor areas (CMA; see, e.g. Dum & Strick, 2002).

By modulating the activity of upper motor neurons, two additional brain structures are important for motor behavior and action monitoring: (3) the *basal ganglia*, which have been suggested to continuously monitor ongoing events and predict whether outcomes of these events are worse or better than expected (Holroyd & Coles, 2002, see also section 1.3). (4) The *cerebellum* plays an important role for integrating motor and sensory information about ongoing movements, which can in turn be used for error detection (e.g. by comparing actions and their outcomes) and correction (also during piano performance, see Pfordresher, Mantell, Brown, Zivadinov, & Cox, 2014). Evidence from brain imaging and lesion studies indicate that the cerebellum (and also the posterior parietal cortex and SMA; Ikudome, Nakamoto, Yotani, Kanehisa, & Mori, 2013) play a crucial role for forward models (e.g., Blakemore, Frith, & Wolpert, 2001; Miall, Christensen, Cain, & Stanley, 2007; Bastian, 2006; Ito, 2008; Imamizu & Kawato, 2012, see section 1.2.2.1): The cerebellum is involved in generating sensory predictions (e.g., motor-to-auditory; Knolle, Schröger, Baess, & Kotz, 2012) and in "signaling the sensory discrepancy between the predicted and actual sensory consequences of movement" (Blakemore et al., 2001, p. 1879).[3] All these structures are also crucial for music performance, and in part for music perception (for overviews, see e.g. Zatorre et al., 2007; Brown, Zatorre, & Penhune, 2015). However, the degree to which each region is engaged and how they interact with each other probably depends on learning- and experience-related structural and functional changes in the brain of performers (for reviews, see Herholz & Zatorre, 2012; Sanes & Donoghue, 2000; Schlaug, 2009; Münte, Altenmüller, & Jäncke, 2002).

1.2.2. Motor Control

To answer the question how we control our movements and why we move the way we do, several concepts and models have been developed (for an overview, see e.g. Schmidt & Lee, 2005). For example, in closed-loop systems for motor control, it is assumed that sensory feedback is used to regulate and modulate our actions. However, processing and evaluating sensory information would be too slow for the control of fast movements occurring for example during piano performance (e.g., Lashley, 1951). Thus, open-loop models assume

---

[3] Note, however, that although these structures are important for generating and controlling movements, they can also influence nonmotor areas of the cortex and appear to perform also other, more cognitive functions (see, e.g., Schubotz, 2007; Strick, Dum, & Fiez, 2009; Ramnani, 2006; Petacchi, Laird, Fox, & Bower, 2005).

that motor control does not depend on sensory input. The influential concept of (generalized) motor programs assumes some sort of pre-structured set of central and abstract representation of movements, independent of the actual effectors executing movements (e.g., Schmidt, 1975). Although the term motor program is still widely used, there is an ongoing debate about what exactly a motor program is and what role it plays for movement control (for reviews, see Morris, Summers, Matyas, & Iansek, 1994; Summers & Anson, 2009). For the context of the present thesis, two related concepts are particularly important: internal models (computational point of view; see Wolpert, 1997; Wolpert & Flanagan, 2001; Wolpert, Ghahramani, & Flanagan, 2001), and the ideomotor approach (cognitive psychology perspective, see Prinz, 1997).

1.2.2.1. Internal Models

An internal model can be described as "a system that mimics the behavior of a natural process" (Wolpert, Ghahramani, & Jordan, 1995, p. 1880). Two kinds of internal models can be differentiated: forward models and inverse models (see also Figure 1.1). Internal inverse models are involved in the transformation from desired consequences (e.g., a specific sound) to the appropriate actions required to achieve these consequences (e.g., striking a specific key on the piano). The input to an internal inverse model includes the desired trajectory (a motor plan) and its output is a motor command which results in the desired trajectory. In contrast, internal forward models relate to the transformation from motor commands to the effects of these motor commands. This transformation depends on the physics of the environment, the musculoskeletal system, and sensory receptors. The function of internal forward models is to "model the causal relationship between actions and their consequences" (Wolpert & Ghahramani, 2000, p.1212) and thus to "mimic the causal flow of a process by predicting its next state (for example, position and velocity) given the current state and the motor command" (Wolpert, Ghahramani, & Jordan, 1995, p. 1880). A forward model reflects thus a model of the neuromuscular system and the external world, which acts as a neural simulator to make predictions about the behavior of our bodies and the external environment the body is interacting with. Internal forward models can be further classified into forward dynamic and forward sensory models. The input to both models consists of the current state estimate of the system and a copy of the current motor outflow, that is, an efference copy. However, the output of a forward dynamic model is an estimation how the current state—with respect to limb position, velocity, etc.—changes, whereas the output of forward sensory models is the prediction of sensory feedback given this estimated new state.

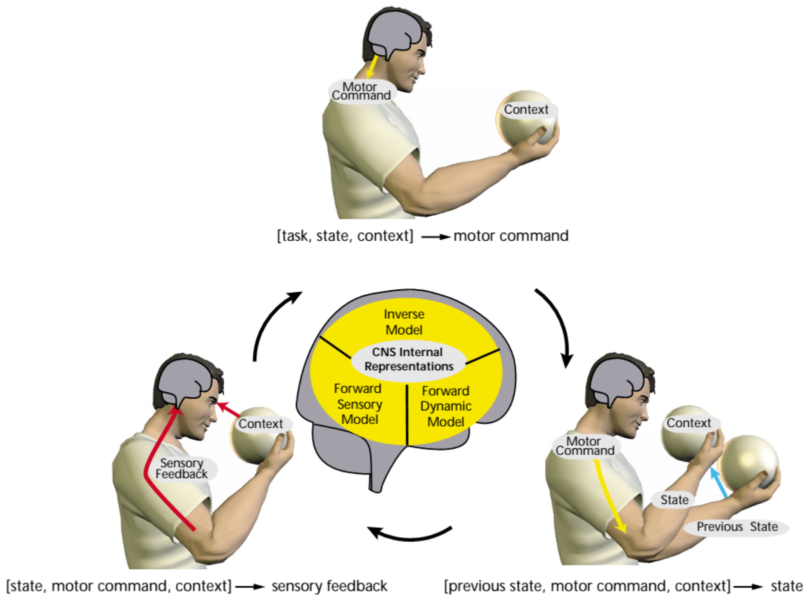

Figure 1.1.: Relationships between internal models. The current task, the current state and the context form the inputs to an inverse model, leading to the generation of a motor command. This motor command, the previous state, and the context are input to a forward dynamic model, which predicts the state transition, i.e., how the state of the system will change depending on the previous state, given motor command, and context. The forward sensory model predicts the sensory feedback given the state, motor command, and context. Figure from Wolpert and Ghahramani (2000). Reprinted by permission from Macmillan Publishers Ltd: Nature Neuroscience, ©2000.

One of the major advantages of forward models is that the problem of time delays—present during sensory transduction, neural conduction, and in muscles responding to motor commands—can be diminished. Although responses to external signals can occur very fast, like the monosynaptic stretch reflex with around 10 to 40 ms, or more complex reflexes even involving cortical processing that vary with task demands around 30 to 70 ms (e.g., Evarts & Tanji, 1976; Evarts & Granit, 1976; Cole, Gracco, & Abbs, 1984; Pruszynski et al., 2011), "we effectively live in the past" (Franklin & Wolpert, 2011, p. 426) because information about our body and the environment is outdated. Because forward sensory models predict the sensory consequences of actions, the predicted feedback is much more rapidly available than actual sensory feedback, and can be directly compared to the intended feedback (Miall & Wolpert, 1996; Wolpert, 1997). Thus, this internal feedback loop is independent of sensory feedback, and discrepancies between predicted and intended outcomes of actions can be detected before any error occurs.

Besides this internal feedback loop, forward models also combine motor outflow (in form of efference copies) with sensory inflow and can thus compare predicted sensory feedback with the actual received sensory feedback. Thus, forward models have a feedforward component, but include also a feedback component (see also Wolpert, 1997; Wolpert & Ghahramani, 2000; Desmurget & Grafton, 2000). For simple arm reaching movements, it has been demonstrated that integrating both types of information can improve the estimate of the state of the system (Wolpert, Ghahramani, & Jordan, 1995), whereby the contribution from each component to motor control depended on the current task and movement stage (Wolpert, Ghahramani, & Jordan, 1995). Therefore, and it is feasible that also motor control during more complex movements (e.g. during music performance) makes use of such a mechanism.

The concept of internal forward models is also used in perception research, for example to explain how we can discriminate between sensory input due to self-generated movements or due to events in the environment. In that view, the sensory prediction derived from the forward model can be used to cancel out the sensory reafference, which in turn can enhance more relevant sensory information. This sensory suppression based on efference copies and internal models has also been reported for the auditory domain, indicating that prediction mechanisms are also relevant for the discrimination of auditory consequences of one's own actions and those of others. Using event-related potentials (ERPs, see also section 1.5), several studies showed that the amplitude of several components of the auditory evoked potential is decreased when sounds are self-produced compared to when sounds are externally generated (Baess, Horváth, Jacobsen, & Schröger, 2011; Baess, Widmann, Roye, Schröger, & Jacobsen, 2009; Knolle, Schröger, & Kotz, 2013).

In order that motor control can remain stable, internal models have to be acquired and modified over time, for example due to growth-, age-, and weight-related changes (e.g., see Franklin & Wolpert, 2011; Wolpert, 1997; Wolpert & Flanagan, 2001; Wolpert, Ghahramani, & Flanagan, 2001). The most important mechanism for learning and adaptation of forward models appears to be driven by the prediction error (difference between desired and actual sensory outcome of actions), which can be used to update the internal model so that motor commands and perceived outcomes match again (for a review, see also Shadmehr, Smith, & Krakauer, 2010). The accuracy of internal models also changes with the level of experience. Thus, it can be assumed that expert musicians have highly developed internal models that allow accurate (spatio-temporal sensory) predictions, as well as the generation of accurate motor commands derived from inverse models, consistent with findings of increased cerebellar volume in male keyboard players compared to non-musicians (Hutchinson, Lee, Gaab, & Schlaug, 2003) and amateur musicians, who showed an intermediate volume increase (Gaser & Schlaug, 2003).

1.2.2.2. Ideomotor and Common Coding Approach

Basically, the ideomotor principle states that actions are planned and controlled by the anticipation of the perceivable effects those actions produce (*action-effects*). Thus, action goals (the anticipatory representation of intended action outcomes) play a central role in this approach. These representations are associated with the corresponding actions, and activating the goal in terms of its sensory consequences can prime the associated action leading to the intended effect. The goal itself can be proximal (such as proprioceptive or tactile feedback when striking a key on the piano), but also more distal (distant in space and delayed in time, for example, when planning a concert tour).

The action-effect principle implies that actions and their consequences must be learned before they can be used for voluntary, goal-directed action: we have to know what movements will lead to which effects. Once actions and their effects are learned, they form bidirectional associations. Therefore, the representation of an action can prime the associated effect and, importantly, the perception of a potential action-effect can prime the associated action that leads to the effect. Based on empirical findings showing that stimuli that share features with planned actions can induce the corresponding action or can interfere with them (according to the ideomotor principle), Prinz (1997) concluded that early stages of actions and late stages of perception share a common representational domain (the *common coding* approach). Thus, a much closer link between action and perception was postulated, in contrast to the largely separate investigations into action and perception in cognitive psychology (see also Hommel, Müsseler, Aschersleben, & Prinz, 2001).

Interestingly, the ideomotor approach as well as the common coding approach share important features with the concept of internal forward and inverse models: inverse models can be used to select the appropriate movements that will achieve the desired sensory consequences, which is similar to the notion that the anticipation of action outcomes guides the selection of the action resulting in these outcomes. Further, forward models can be used to predict the outcome of actions, which is similar to the notion that actions can prime the corresponding action-effects. Thus, both the ideomotor approach and the concept of internal models suggest that actions and perception are tightly linked.

1.3. Action Monitoring and Error Processing

To reach behavioral goals, our cognitive system has to permanently monitor and evaluate our actions. A crucial part of this evaluating is the detection of errors, which can be used to adjust our behavior adequately or immediately so that we can learn from our errors. Research into the neurocognitive basis of error processing and how we monitor our actions has attracted much attention after the finding of an *Error-related negativity* (ERN or Ne; Falkenstein, Hohnsbein, Hoormann, & Blanke, 1990; Gehring, Goss, Coles, Meyer, & Donchin, 1993). The ERN occurs around 50-100 ms after participants responded incorrectly during choice-reaction tasks, usually followed by a slower positive wave around 200-400 ms termed *Error Positivity* (Pe).[4] The ERN can be elicited independently of the response

---

[4] Although the early negativity has been mostly studied in rather simple speeded choice-reaction time tasks, subsequent studies reported this deflection also occurring during other tasks such as mental rotation, monitoring of internal speech production, speech errors, continuous force production and visual tracking tasks (Band & Kok, 2000; Ganushchak & Schiller, 2006, 2008a, 2008b; de Bruijn, Hulstijn, Meulenbroek, & Van Galen, 2003; Krigolson & Holroyd, 2007). The Pe has been much less rigorously studied than the ERN, but presumably reflects different processes than the ERN (for a review, see e.g. Overbeek, Nieuwenhuis, & Ridderinkhof, 2005). Note that the Pe can also be decomposed into two subcomponents: an early Pe, which

modality (hand, foot, eyes, vocal responses; Holroyd, Dien, & Coles, 1998; Endrass et al., 2007; Masaki, Tanaka, Takasawa, & Yamazaki, 2001), and independently from the modality in which the stimulus is presented (auditory or visual; Falkenstein, Hoormann, Christ, & Hohnsbein, 2000).

The ERN is likely the consequence of processing *internal* feedback, because it seems unlikely that external feedback could already have been processed by the peak of the ERN around 50-100 ms. However, several studies reported a similar but later electrophysiological response (around 230 to 270 ms) occurring when there is *external* feedback (in the auditory, visual, or somatosensory domain) signaling that an error has occurred or indicating a loss or punishment in time-estimation tasks, guessing tasks, and gambling tasks (e.g., Miltner, Braun, & Coles, 1997; Gehring & Willoughby, 2002; Hajcak, Holroyd, Moser, & Simons, 2005; Hajcak, Moser, Holroyd, & Simons, 2007; Nieuwenhuis, Yeung, Holroyd, Schurger, & Cohen, 2004; Yeung & Sanfey, 2004).

Based on the available evidence, it has been suggested that this feedback-related negativity (FRN) involved similar neural processes as the ERN, and is part of a monitoring system that evaluates whether action-outcomes are worse (or better) than expected—that is, the FRN is taken to be involved in the processing of expectancy violations. Although it is elicited regardless of whether the outcome is worse or better than expected (Jessup, Busemeyer, & Brown, 2010; Oliveira, McDonald, & Goodman, 2007), its amplitude can be modulated by the degree of expectancy (larger for more unexpected events; Hajcak, Moser, et al., 2007; Holroyd, Nieuwenhuis, Yeung, & Cohen, 2003; Holroyd, Larsen, & Cohen, 2004).[5]

Importantly, mediofrontal negativities such as the FRN appear to be elicited by

---

occurs around 200 to 400 ms and shows a frontocentral scalp distribution, and a late Pe occurring around 300 to 600 ms and with maximal amplitudes at centro-parietal electrodes (Endrass, Reuter, & Kathmann, 2007; Van Veen & Carter, 2002). This distinction is very similar to the distinction between the P3a and P3b (for a review, see e.g. Polich, 2007), and it is still an ongoing debate whether the positivities elicited during error processing and the P300 components might reflect similar or even the same neural processes.

[5] However, it seems that "...participants' expectancies may have to be relatively strong and closely coupled with action outcome pairs to influence feedback negativity amplitude" (Hajcak, Moser, et al., 2007, p. 6). During music performance, it is relatively straightforward to assume that musicians have strong expectancies towards specific tones, and it has been shown that musicians show a strong coupling between musical actions and the resulting auditory effects (e.g., Baumann, Koeneke, Meyer, Lutz, & Jäncke, 2005; Bangert, Peschel, et al., 2006; Lotze, Braun, Birbaumer, Anders, & Cohen, 2003).

unexpected outcomes, even if participants only observe outcomes without producing these outcomes by their own actions (Donkers & van Boxtel, 2005; Donkers, Nieuwenhuis, & van Boxtel, 2005; Yeung, Holroyd, & Cohen, 2005; Tzur & Berger, 2007, 2009). Thus, observation and direct experience of error commission or outcome evaluation elicit similar neural responses, suggesting that "the observer would in some sense be doing the same task as the virtual subject so that the observer's own executive system is engaged when an error is observed" (Miltner, Brauer, Hecht, Trippe, & Coles, 2004, p. 128). Therefore, the FRN seems not dependent on recently executed actions (neither own actions nor observing somebody else's actions; van Schie & Mars, 2004; Miltner, Brauer, et al., 2004; Marco-Pallarés, Krämer, Strehl, Schröder, & Münte, 2010; Itagaki & Katayama, 2008), suggesting that the FRN might reflect a more general evaluation of ongoing events in terms of a conflict or mismatch between expected and experienced events.

There is converging evidence that the error-related negativity and the feedback-related negativity receive major contributions from the rostral cingulate zone (RCZ) of the posterior medial frontal cortex (including the anterior cingulate cortex; for a review, see Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004), based on electrophysiological (e.g., Luu, Tucker, Derryberry, Reed, & Poulsen, 2003; Herrmann, Römmler, Ehlis, Heidrich, & Fallgatter, 2004; Gehring & Willoughby, 2002; Ruchsow, Grothe, Spitzer, & Kiefer, 2002), functional magnetic resonance imaging (fMRI; e.g., Braver, Barch, Gray, Molfese, & Snyder, 2001; Carter, 1998; Kiehl, Liddle, & Hopfinger, 2000; Menon, Adleman, White, Glover, & Reiss, 2001; Debener et al., 2005; Holroyd, Nieuwenhuis, Yeung, Nystrom, et al., 2004; Ullsperger & von Cramon, 2003; Ullsperger, Nittono, & von Cramon, 2007; Bush et al., 2002), and non-human primate studies (e.g., Gemba, Sasaki, & Brooks, 1986).[6]

Error detection hypothesis One of the first theoretical accounts of the ERN was the error detection (or mismatch) hypothesis (Falkenstein, Hohnsbein, et al., 1990; Coles, Scheffers, & Holroyd, 2001). In this view, an assumed error-processing system consists of a

---

[6] Note that there is still an ongoing debate as to whether the ERN and FRN only share overlapping resources (Luu et al., 2003; Müller, Möller, Rodriguez-Fornells, & Münte, 2005) or have the same neural generators (see also Nieuwenhuis, Holroyd, Mol, & Coles, 2004, p. 445).

monitoring system and a remedial action system, which initiates actions that inhibit or correct the error, or changes task priorities to prevent errors in the future (such as post-error slowing). The most important part of this model is the comparator, which detects errors by comparing representations of actual responses with representations of the appropriate or desired responses. Representations of the actual response are assumed to be derived from an efference copy of the ongoing motor command. When the efference copy arrives at the monitoring system, the comparison process is triggered and in case of a mismatch, the comparator sends an error signal to the remedial action system. According to this view, the ERN is generated when the error signal is received by the remedial action system.[7]

Conflict-monitoring hypothesis In contrast, the conflict-monitoring hypothesis states that the ERN is rather a signal indicating that conflict has occurred (reviewed in Botvinick, Braver, Barch, Carter, & Cohen, 2001; Yeung, Botvinick, & Cohen, 2004; Botvinick, Cohen, & Carter, 2004; Carter & van Veen, 2007). According to this view, the anterior cingulate cortex (ACC) monitors for the occurrence of conflict, which can be understood as crosstalk between concurrent cognitive processes and can occur, for example, during response selection when there are "concurrently active incompatible responses" (Yeung, Botvinick, & Cohen, 2004, p. 933). The information about the level of conflict can then be passed to other systems responsible for implementing cognitive control in order to resolve the conflict, e.g. to the dorso-lateral prefrontal cortex (DLPFC) which implements performance adjustments (Miller, 2000; Cavanagh, Cohen, & Allen, 2009).[8]

Reinforcement learning theory of the error-related negativity This theory states that the ACC is the recipient of an error signal generated by the basal ganglia (and conveyed by the mesencephalic dopamine system), which monitor events and continuously predict whether outcomes of events will be better or worse than expected (see Holroyd & Coles, 2002). The predictions can be based on internal (e.g., as occurring during overt responses)

---

[7] Note that there is also the alternative view that the ERN reflects the comparison process itself, and not the outcome of the comparison (Falkenstein, Hoormann, et al., 2000).

[8] Note that if a broader definition of *conflict* is used, in which negative feedback is conceptualized as conflict between actual and expected events, the conflict monitoring hypothesis can also account for the feedback-related negativity.

or on external information (e.g., sensory feedback), whichever provides the first indication on which the system can evaluate the appropriateness of a behavior or, more generally, of events (see also Heldmann, Rüsseler, & Münte, 2008; Müller et al., 2005). By making these assumptions, the reinforcement learning theory can account for the ERN as well as for the FRN, even in the absence of overt responses. Note that also efference copies can be incorporated in this theory, so that the model can “detect task-inappropriate neural activity before it has developed into an overt behavior" (Holroyd & Coles, 2002, p. 699). Similarly to concepts from computational motor control, the theory of reinforcement learning of the ERN also stresses the possibility that the neural signal reflected by the ERN can be used to modify activity of the motor system during ongoing behavior. In other words, the ERN can be interpreted as initiating remedial actions before the completion of an erroneous movement.

1.4. Music Performance

In contrast to studies investigating action monitoring and error processing during simpler and often artificial tasks (see section 1.3), studies addressing these questions during complex time-based sequential behaviors such as music performance are scarce. In this section, I will briefly summarize two particular relevant aspects, namely how is auditory and tactile feedback used during music performance.

1.4.1. Auditory Feedback Processing

For music, auditory feedback seems intuitively to be important because a performer can use it to monitor his or her own actions and to make performance adjustments if the auditory result of the actions deviates from some performance goal. However, the importance and role of auditory feedback for music performance is still unclear, and the available evidence suggests a somewhat limited role of auditory feedback.

A common approach to test this issue is the altered auditory feedback (AAF) paradigm, a variant of which will be also used in the present thesis (see section 2.1). In this paradigm, participants produce musical sequences but hear auditory feedback that is not consistent with the produced actions (and thus unexpected). Two types of feedback alterations (and combinations thereof) have been primarily investigated: alterations of onset synchrony of musical actions and auditory feedback, and alterations of the pitch content of the feedback. Disruptive effects of alterations on action planning and execution are usually quantified in terms of error rates, inter-onset intervals (IOIs), and timing variability of whole sequences, and can be used to investigate if and how performers rely on auditory feedback.

Several studies showed that the complete absence of auditory feedback during the performance of learned melodies has no effects on performances (a) when pianists played melodies from memory (Finney, 1997; Finney & Palmer, 2003; Pfordresher, 2005), (b) when pianists were sight-reading melodies (Banton, 1995), (c) when untrained participants were playing piano music (Pfordresher, 2005), and only small effects during expressive piano performances (Repp, 1999b).[9]

On the other hand, specific alterations of auditory feedback can profoundly disrupt piano performance, and there appears to be a dissociation between pitch and timing of auditory feedback: pitch manipulations result mainly in higher error rates, but influence timing only little (especially in serial shifts, in which feedback matches a pitch event intended at a different location in a sequence; Pfordresher, 2003, 2005; Pfordresher & Palmer, 2006); in contrast, feedback delays increase mainly timing variability and IOIs, but influence error rates only marginally (for a review, see Pfordresher, 2006). More specifically, the amount of disruption from feedback delays seems to depend on the relative phase (rhythmic) relationship between key presses and auditory feedback (Pfordresher & Palmer, 2002; Pfordresher, 2003; Pfordresher & Benitez, 2007) and on the current movement trajectory of the finger (at least during tapping; see Pfordresher & Dalla Bella, 2011). However, auditory feedback manipulations of single key presses disrupt also the production of directly successive keystrokes, indicating an important role of auditory feedback in these cases

---

[9] Note that most studies employing the AAF paradigm investigated piano performances, and results probably do not translate to other instruments, especially to string instruments like violin or cello. On these instruments, proper intonation is required and the musician has to choose from a continuum of possible pitches, whereas there are only discrete pitches, produced by discrete keys, on keyboard instruments. Accordingly, it has been suggested that error detection and correction on these instruments depends crucially on auditory feedback (Chen, Woollacott, Pologe, & Moore, 2008). However, that study used a rather slow tempo of one beat per second, and it remains unclear how important auditory feedback is at faster tempos during string instrument performances.

(Furuya & Soechting, 2010).

It has been suggested that disruptions may reflect the "perturbation of perception/action coordination at a given time scale" (Pfordresher, 2006, p. 194), or "active attempts to counteract the influence of auditory feedback" (Pfordresher, 2006, p. 194), and that the source for disruptions could be expectancy violations induced by altered feedback (Gabrielsson, 2003).[10] However, expectancy violations can also be induced by the processing of the auditory input alone, and based on behavioral data it is difficult to decide what role exactly the performing of musical actions plays in this context. Therefore, one aim of the present thesis was to investigate the neural mechanisms underlying the processing of auditory feedback manipulations (pitch manipulations), which can help to disentangle the roles of perception and action (see section 2.1).

1.4.2. Tactile Feedback Processing

In the context of action monitoring and motor control during music performance, the densely distributed mechanoreceptors in the fingertips are of special interest, because they often provide direct tactile feedback about the manipulation of an instrument. The importance of tactile feedback is evidenced by disruptive effects on performance when tactile feedback is reduced, for example, by anesthesia during grasping, precision grip tasks, or typing (e.g., Gentilucci, Toni, Daprati, & Gangitano, 1997; Fisher, Galea, Brown, & Lemon, 2002; Rabin & Gordon, 2004).[11]

However, only relatively few studies investigated the role of tactile feedback for music performance. Because the available evidence suggests that tactile feedback helps error detection during typing (Gordon & Soechting, 1995; Rabin & Gordon, 2004) and timing accuracy during tapping (Repp, 1999a; Drewing, Hennings, & Aschersleben, 2002; Aschersleben, Gehrke, & Prinz, 2001) and music performance (Goebl & Palmer, 2008), the

---

[10] Consistent with the common coding approach suggesting that action and perception are tightly coupled and share resources and with the concept of internal forward models used for action control and predictions about the sensory consequences of actions (see section 1.2.2.2).

[11] The tight coupling between motor control and tactile feedback from the fingers and hands was also demonstrated by the speed with which people can respond to tactile feedback: when an object held with the fingers started to slightly slip, participants were able to adjust their grip after only 80 ms (Johansson & Westling, 1988).

present thesis also exploratively investigated the role of tactile feedback of keystrokes for error detection mechanisms during music performance (see section 2.3).

1.5. Electroencephalography

When an excitatory neurotransmitter is released at the apical dendrites of a pyramidal cell in the human neocortex, it produces an excitatory postsynaptic potential (EPSP): an ionic current flows inward through the synaptic membrane (leading to a net negativity at the apical dendrites), towards the soma, and outward along the extrasynaptic membrane (leading to a net positivity in the region of the soma). The EPSPs produced by single pyramidal cells summate well over time. If these cells are arranged in open field structures, and if sufficiently enough cells are synchronously active (at least 100000), the resulting electrical activity can be measured extracellularly with electrodes, even from some centimeters distance. The recording of electric activity (i.e. time-varying voltages) of the brain with electrodes on the scalp is called *electroencephalography* (EEG). The EEG reflects mostly EPSPs generated in the cerebral cortex, whereas action potentials only minimally contribute to the EEG, because they hardly summate over time due to their short duration (Kandel, Schwartz, & Jessel, 1991).

The term *event-related potential* (ERP) refers to all electrocortical potentials occurring before, during, or after a sensory, motoric, or psychological event. It is assumed that ERPs reflect the cognitive processing related to such an event. ERPs show relatively small amplitudes (around 1 - 30 μV) compared to the large amplitudes of the spontaneous EEG (around 50 to 100 μV), reflecting the electrical activity elicited by the vast number of operations permanently performed by a brain. To distinguish between event-related activity and spontaneous brain activity, ERPs are usually derived by averaging many epochs of the EEG time-locked to a repeatedly occurring event. The primary assumption for this averaging procedure is that the underlying brain activity is identical or very similar each time the event occurs. In contrast, the spontaneous brain activity and artifacts such as electrical activity of muscles, eye movements, or cardiac artifacts are random and unrelated to the event. Therefore, during averaging, the amplitude of the EEG reflecting spontaneous brain activity and artifacts decreases (the uncorrelated activity is averaged out), while the event-related signal increases.

One crucial advantage of the EEG (and ERPs) is its high temporal resolution in the milliseconds range, together with its ability to record brain activity continuously. Thus, EEG is a method particularly suited for the investigation of phenomena unfolding over time, such as music performance. However, the EEG signals measured at the scalp do not inform about the location of the active neurons generating the EEG signal. This is mainly due to the inverse problem of EEG (and magnetoencephalography, MEG): theoretically, an infinite number of possible charge distributions in a sphere (like the human head) can lead to the same pattern on the surface of this sphere (like the human scalp). By making a priori assumptions on the generation of the signals, various solutions to this problem have been developed, which can be divided in two broad categories: equivalent dipole approaches and linear distributed approaches (for an overview, see e.g. Pizzagalli, 2007). The first approach assumes that EEG/MEG signals are generated by a relatively small number of focal sources, which can be modeled as dipole(s) with certain parameters such as location, orientation and strength positioned in a head model. In an iterative process, the forward solution of each dipole configuration is calculated and compared to the actual EEG measurements, and the approximate solution of the inverse problem is found by identifying the dipole parameters that best explain the observed scalp potentials (i.e., show the smallest difference between predicted and observed EEG signal). However, the solution derived from these approaches strongly depends on the numbers of dipoles modeled, which can often not be determined a priori. In contrast, linear distributed approaches consider all possible source locations simultaneously, and no a priori assumptions about the number of sources have to be made. In addition, distributed source models typically allow to restrict the solution space by anatomical and functional constraints (e.g. gray matter). However, unlike equivalent dipole models, the dipoles have fixed positions, which are determined by anatomical and physiological constraints. One disadvantage of this approach is that the solutions have low spatial resolution, i.e. their solutions are often blurred.

2. Experiments

In this chapter three ERP studies investigating action monitoring and error processing during piano performance, with an emphasis on auditory and tactile feedback processing, are summarized. To this aim, situations were contrasted in which the sensory (auditory) outcome of an action does or does not match the intended outcome. A match between action and outcome occurs when a pianist produces the correct note and receives the corresponding correct feedback. For a mismatch, two scenarios were realized: (a) a pianist performs a correct key press, but perceives the auditory feedback of a nearby (incorrect) key. Thus, the auditory effect of the (correct) action does not match with the indented auditory result (first two ERP studies in section 2.1; Maidhof, Vavatzanidis, Prinz, Rieger, & Koelsch, 2010); (b) a pianist commits an error by accidentally hitting a wrong key, and perceives the auditory feedback of that incorrect key.[1] Thus, the result of the keystroke does not match with the intended auditory result in terms of pitch content, although the auditory effect is consistent with the currently pressed incorrect key (re-analysis of the data of the second ERP study in section 2.2; Maidhof, Rieger, Prinz, & Koelsch, 2009). To exploratively investigate the role of tactile feedback for error detection during music performance, the third study used three-dimensional movement data of pianists' fingers and ERPs were analyzed with respect to different movement stages of (erroneous) key presses (section 2.3; Maidhof, Pitkäniemi, & Tervaniemi, 2013). The novel setup for the combined ERP-motion capture study (Maidhof, Kästner, & Makkonen, 2014) will be shortly described in section 2.4.

---

[1] Although different types of performance errors exist, for example timing or loudness errors, additionally played notes, and omitted notes, I will focus here on so-called pitch errors (or substitution errors). Pitch errors (or hereof simply "errors") occur when a given note in the score is not produced on the piano, but instead a different note is produced; in other words, the correct note is substituted by an incorrect note.

2.1. Processing Expectancy Violations during Music Performance and Perception

Musicians intend to produce specific auditory effects by executing certain actions. Due to previously learned associations between these actions and their auditory consequences, musicians can build specific expectations for the auditory feedback of their actions. This coupling between auditory and motor systems in musicians is evidenced by results of behavioral (Drost, Rieger, Brass, Gunter, & Prinz, 2005a, 2005b; Drost, Rieger, & Prinz, 2007), neurophysiological (Bangert & Altenmüller, 2003), and neuroimaging studies (e.g., Haueisen & Knösche, 2001; Lahav, Saltzman, & Schlaug, 2007; Kleber, Birbaumer, Veit, Trevorrow, & Lotze, 2007).

Nevertheless, there is evidence that the role of auditory feedback for action control seems to be limited, although certain feedback alterations can disrupt piano performance and influence planning (Pfordresher, 2006). However, these disruptive effects might be due to violations of musical expectancies built up during the perception of a specific musical context, and thus not primarily due to a perturbation of action and perception coordination (see section 1.4.1). To further disentangle the roles of action and perception in auditory feedback processing during music performance, the first experiment investigated the neural mechanisms underlying the processing of auditory feedback manipulations during piano performances.

Event-related potentials elicited during music performance (*action condition*) were compared with ERPs when pianists only perceived such stimuli they had to produce in the other condition (*perception condition*). In the action condition, participants performed bimanually sequences on a digital piano from memory without visual feedback, while at random positions (between every 40th to 60th performed note) the auditory feedback of single key presses was lowered by one semitone. Thus, in the action condition, both action-related expectancies toward a tone (based on the performed keystroke and the intention to produce a specific tone) and perception-related expectancies (induced by the preceding musical context) were violated. In the perception condition, pianists listened to the stimuli (including the same manipulations) without performing any actions, but under the instruction to silently count the manipulated tones. Therefore, only perception-related expectancies toward a tone were violated.

Results showed that manipulated tones elicited in both, the action as well as the perception condition, a negativity that showed maximal amplitude values around 200 ms after the tone onsets over fronto-central electrodes, resembling the feedback ERN/N200 component (see Figure 2.1). The negative deflections were followed by later P3a and P3b potentials, usually associated with an automatic shift of attention to deviant stimuli and the conscious detection of target stimuli, respectively (Polich, 2007). The amplitude of the negativity in the action condition was larger compared to the perception condition, indicating that when pianists were performing, pitch manipulations of the auditory feedback were more unexpected than pitch manipulations when pianists were only listening to the stimuli. Therefore, it seems likely that feedback alterations can disrupt piano performance by violating action-related expectancies, and not only perception-related expectancies.

Although I argue that the observed negativities do not only consist of other negative ERP potentials occurring in the same latency range and reflecting acoustic or tonal deviance processing such as the mismatch negativity (MMN; Winkler, 2007) or early right anterior negativity (ERAN; Koelsch, 2005), the results leave open the possibility that the negative potential in the perception condition overlaps with an N2b component, reflecting the controlled and conscious detection of task-relevant deviants (Novak, Ritter, Vaughan, Wiznitzer, & Wiznitzer, 1990). Therefore, the task of the participants was varied in the second ERP study so that pitch manipulations in the perception condition were task-irrelevant (participants performed time interval judgments). The hypothesis was that, if the negativity in the previous experiment reflects expectancy-related processes (as indexed by an FRN/N200) and not only the detection of task-relevant targets (as indexed by an N2b), it should also be elicited by task-irrelevant manipulations.

Results showed, as in the previous experiment, that pitch manipulations elicited a negative deflection around 200 ms, which showed larger amplitudes in the action compared to the perception condition (see Figure 2.1). Because the manipulated tones were not task-relevant in the perception condition, as indicated by the absence of a P3b component, and because the P3b is usually observed in combination with the N2b, it is likely that the observed negativity is not simply an N2b potential. Nevertheless, the possible contributions of MMN and ERAN potentials to the negativities elicited by pitch manipulations in action and perception conditions cannot be disentangled based on the present data.

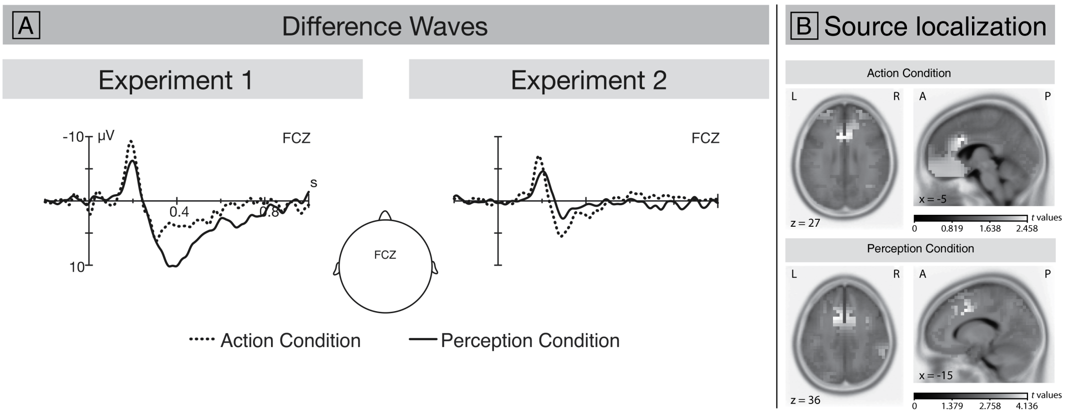

Figure 2.1.: Experiment 1 and 2: ERP and source localization results. (A) Difference waves of Experiments 1 and 2. Difference waves were obtained by subtracting the ERPs elicited by the standard tones from the ERPs elicited by the manipulated tones in the action and perception conditions, respectively. (B) shows the source localization results of the sLORETA analysis, suggesting that the main neural generators of both negativities in Experiment 2 lie within the RCZ (Talairach coordinates: x = -5, y = 16, z = 27 and x = -15, y= 11, z = 36 in the action and perception condition, respectively). Modified with permission from Maidhof, Vavatzanidis, et al. (2010). ©2010 by the Massachusetts Institute of Technology.

However, it is unlikely that the early negativities were simply MMN or ERAN potentials (for a detailed discussion, see Maidhof, Vavatzanidis, et al., 2010). For example, performance errors committed with one hand are acoustically similar to the feedback manipulations, thus providing an estimate of MMN and ERAN contributions. Because performance errors elicited no negative potential in the time range around 200 ms (section 2.2), it is unlikely that the observed negativities elicited by feedback manipulations are simply an MMN or ERAN. Furthermore, by using a variant of the low-resolution brain electromagnetic tomography (sLORETA; Pascual-Marqui, Michel, & Lehmann, 1994; Pascual-Marqui, 2002), the neural generators of both negativities (in action and perception condition) were estimated to lie within the RCZ (see Figure 2.1), which is consistent with an explanation in terms of feedback ERN/N200 (see section 1.3), but not in terms of MMN or ERAN, whose main neural generators are located in the auditory cortex (Näätänen, Paavilainen, Rinne, & Alho, 2007; Alho, 1995) and the inferior fronto-lateral cortices (Garza Villarreal, Brattico, Leino, Ostergaard, & Vuust, 2011; Koelsch, 2009).

In sum, the results of the first two experiments show that during music performance and perception, pitch manipulations can elicit ERP components such as the FRN/N200 associated with action monitoring and the processing of expectancy violations. Importantly, the amplitude of the FRN/N200 was modulated by the action of an individual. During music performance, pianists expect, on the basis of their intention and their act of performing, to perceive a specific auditory effect. This expectancy is likely based on predictions derived from internal forwards models and on the action-effect principle (see section 1.2.2). While listening to music, expectations are formed based on the underlying regularities of the preceding musical context (Koelsch, 2011). Consequently, pitch manipulations during performance were more unexpected (due to the violation of action-related and perception-related expectancies) than manipulations during perception, resulting in the enlarged FRN/N200 in the action compared to the perception condition.

2.2. ERP Effects Prior to Performance Errors in Musicians Indicate Fast Monitoring Processes

Whereas the previous section dealt with the processing of unexpected external sensory (auditory) information during piano performance, the following two sections deal with the detection of performance errors, as these can provide further insights into the mechanisms of action monitoring during music performance. Questions that arise in this context are at what point in time and how errors are actually detected, whether they are detected prior to execution, and at what point in time potential errors can still be corrected. To address these questions, the data of the action condition of the previous experiment were re-analyzed and ERPs were calculated for correctly performed notes and for incorrectly performed notes.[2]

On a behavioral level, incorrect keystrokes were performed with a lower velocity than the

---

[2] Incorrectly performed notes occurred when one hand was playing an incorrect note while the correct key was pressed with the other hand. Note also that only errors were analyzed that were preceded by a 1 s period during which no feedback manipulations occurred, as well as no other performance errors.

correctly performed keystrokes simultaneously with the other hand and correct keystrokes when no error was present. This indicated that the velocity of erroneous keystrokes in one hand did not influence the velocity of simultaneous correct keystrokes. In addition, the IOIs of incorrect keystrokes were prolonged compared to IOIs of correct keystrokes, indicating that the upcoming error slowed down the key presses (pre-error slowing).

On a neurophysiological level, ERPs elicited during correct and incorrect key presses showed a fronto-central negative difference occurring around 100 ms before the note onset (termed *pre-error negativity*, see Figure 2.2). The pre-error negativity was followed by a later positive deflection around 280 ms after the onset of the incorrect note, resembling the Error Positivity or the P3a. Thus, ERPs differed already before the erroneous movement was fully executed and before auditory feedback of the error was available.

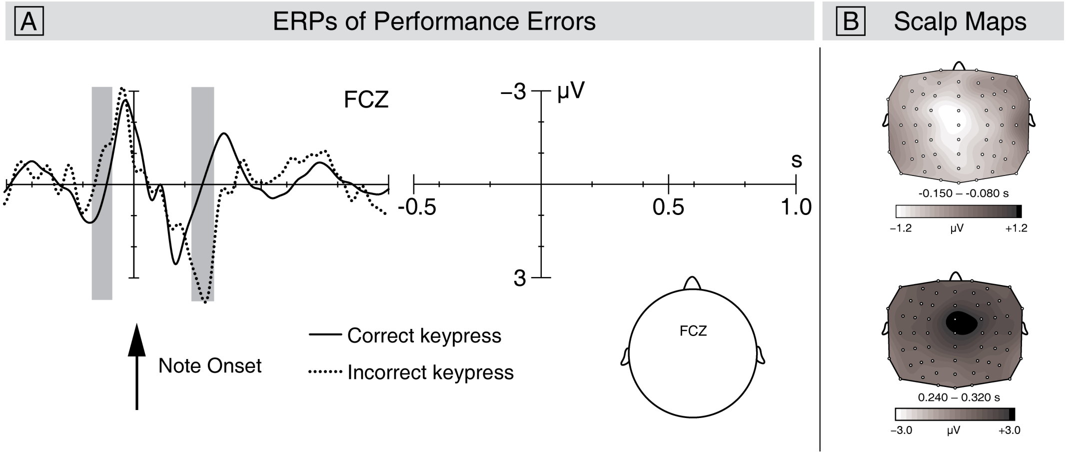

Figure 2.2.: Experiment 2: ERP results and scalp distributions of correctly and incorrectly performed notes. (A) Grand-average ERPs elicited by correctly and incorrectly performed key presses. The gray areas show the time windows chosen for statistical analyses. (B) shows the scalp distributions for the difference potentials for correct key presses subtracted from incorrect key presses. Modified with permission from Maidhof, Rieger, et al. (2009).

Because IOIs were prolonged before errors, the pre-error negativity could reflect lower-level motor-related processes such as adjusting the force of muscles, resulting in the erroneous key press. Because these processes are expected to be lateralized (Shibasaki, Barrett, Halliday, & Halliday, 1980; Colebatch, 2007), and because a separate analysis of errors committed with the right or left hand showed no lateralization effects, an explanation in terms of purely motor-related processes seems rather unlikely.

Therefore, it was assumed that the early ERP difference reflects cognitive processes of error detection, error correction, or both. More specifically, it was assumed that this error detection prior to error commission[3] is based on internal forward models that predict sensory consequences of ongoing movements, such as proprioceptive and tactile information, visual feedback, and auditory feedback (see section 1.2.2.1). Whenever a discrepancy between predicted and intended (or actual) feedback is detected, an error signal is generated that can lead to corrective modulation of motor commands. The lower velocity of erroneous keystrokes might reflect such a corrective modulation, in order to prevent the error or at least reduce its sensory effects (e.g., a decrease of loudness).

However, the comparison of correct and incorrect keystrokes is confounded by their different IOIs. Because ERPs of wrong notes (longer IOIs) and the previous correct notes (shorter IOIs) might have overlapped, the different IOIs possibly influenced the ERP prior to erroneous notes. If, however, the pre-error negativity mainly reflects error-related processes and not tempo differences, it should be elicited even when correct and incorrect notes are performed in a comparable tempo. In addition, EEG data were related to the time-point when the key was almost fully pressed down, (i.e., the point at which the MIDI [Musical Instrument Digital Interface] signal is generated by the digital piano). This point in time is preceded by touching and pressing down the key, which can last several tens of milliseconds (Goebl & Palmer, 2008). Therefore, tactile and proprioceptive feedback could already have contributed to the observed pre-error negativity. These issues were addressed in the next experiment.

2.3. Predictive Error Detection in Pianists: A Combined ERP and Motion Capture Study

In this study, participants performed sequences with the right hand without visual feedback of their hand movements in a high tempo (IOI = 125 ms). In addition to the EEG data, finger movements were recorded with a motion capture system that registered the three-dimensional position data of small markers attached to the fingernails of the

---

[3] If the onset of the audio signal generated by a keystroke is defined as error commission.

participants (every 8.3 ms).

On a behavioral level, results replicated previous findings and showed that incorrect notes were performed with a prolonged IOI and a reduced key press velocity. However, the analysis of the motion capture data provided additional information about the underlying movements (for an example, see Figure 2.3). (1) The onset of tactile feedback of incorrect key presses, i.e. the time when a finger makes first contact with the piano key surface, occurred around 60 ms prior to key depression. In contrast, tactile feedback for correct key presses occurred around 50 ms prior to key depression. (2) Acceleration values of this first finger-key contact (as an indicator of the speed of the movement towards a key) did not differ between correct or incorrect events. Given that IOIs between correct and incorrect notes differed around 13 ms, it appears that not the movement towards the incorrect key itself was executed slower, but rather the phase between touching and complete key depression was prolonged for incorrect keystrokes. Thus, the observed pre-error slowing observed in this study (and probably in similar studies, see Herrojo Ruiz, Jabusch, & Altenmüller, 2009; Herrojo Ruiz, Strübing, Jabusch, & Altenmüller, 2011; Strübing, Herrojo Ruiz, Jabusch, & Altenmüller, 2012) is likely due to the lower velocity with which wrong keys were pressed down (which takes more time), and not due to a slower movement of a finger towards the incorrect key.

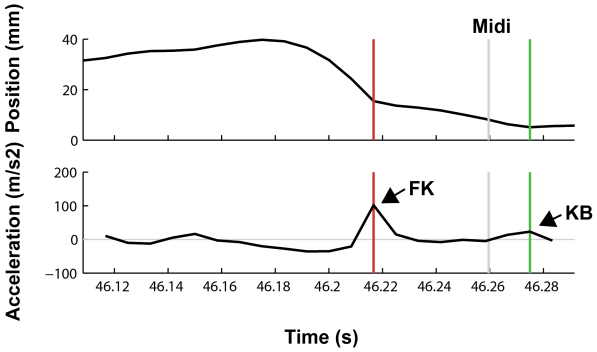

Figure 2.3.: Vertical motion of a fingertip of a pianist's ring finger during a keystroke (G#5 key). Upper panel: finger position, lower panel: acceleration. An acceleration peak can be observed around 40 ms prior to the MIDI note onset, occurring when the fingertip makes initial contact with the key surface (finger-key landmark, FK). Shortly after the MIDI onset, another acceleration peak occurs when the key reaches the key bed after key depression (key-bottom landmark, KB). Modified with permission from Maidhof, Pitkäniemi, and Tervaniemi (2013).

On a neurophysiological level, ERP (computed relative to the MIDI onset, i.e. when a key is almost fully pressed down) results replicated previous findings and showed a pre-error negativity (around 60 ms prior to note onset), followed by a Pe/P3a (see Figure 2.4.A). Importantly, when ERPs of incorrect notes were compared with ERPs of correct notes that showed a comparable tempo, virtually the same ERP pattern was observed. Furthermore, ERPs were exploratively computed based on the onset of tactile feedback for each keystroke, as determined by the movement data (similar to the method used in Goebl & Palmer, 2008, 2009). Results tentatively suggest that ERPs of correct and incorrect keystrokes did not differ prior to tactile feedback, but differed slightly around 40 ms after tactile feedback was available (see Figure 2.4.B).

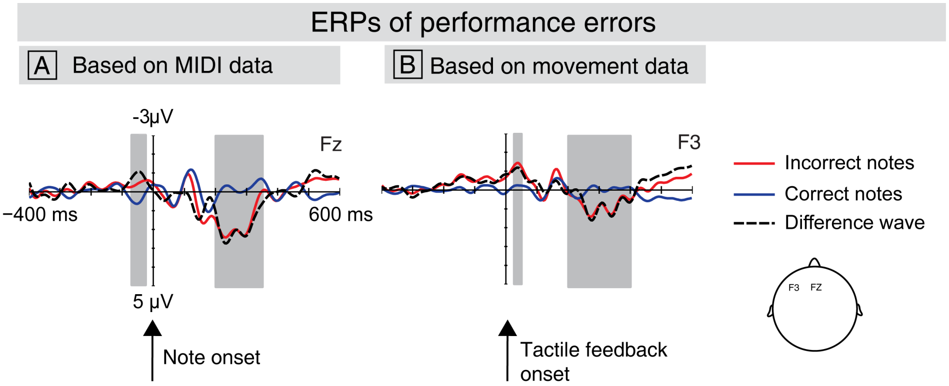

Figure 2.4.: Experiment 3: Grand-average ERP results of correctly and incorrectly performed notes. (A) ERPs elicited by correctly and incorrectly performed key presses, computed relative to the note onset. (B) ERPs elicited by correctly and incorrectly performed key presses, computed relative to the onset of tactile feedback, as determined by the movement data. The gray areas show the time windows chosen for statistical analyses. Modified with permission from Maidhof, Pitkäniemi, and Tervaniemi (2013).

In sum, the findings of this experiment indicate that the pre-error negativity reflects error-related processes and is not strongly influenced by the different IOIs of correct and incorrect key presses. In addition, one might speculate that tactile feedback of the key surface provides an important source of information for the detection of an upcoming error.

2.4. Combining EEG, MIDI, and Motion Capture Techniques for Investigating Musical Performance

In this section, I will briefly describe the setup used for the simultaneous recording of EEG, MIDI, and movement data. Thus, this section does not introduce a new experiment, but rather methodological work that was necessary for the previous experiment.

2.4.1. Motivation

Quantifying musical behaviors became considerably easier by using generally available MIDI instruments like digital pianos. Using these instruments, note on- and offsets can be recorded accurately, as well as an estimate of the force of a key press (the MIDI velocity). However, MIDI data is based on discrete time points (i.e., note events), and much of the information about a musical performance cannot be quantified. Therefore, researchers have used motion capture techniques to directly record the movements of musicians and to investigate the role of auditory and tactile feedback for example in timing accuracy, emotional expressions, or joint action during ensemble coordination (Goebl & Palmer, 2008; Livingstone, Thompson, & Russo, 2009; Keller & Appel, 2010). However, previous studies investigating the neural correlates of music performance (for reviews, see e.g. Zatorre et al., 2007; Brown et al., 2015) could not provide detailed behavioral information (except MIDI- or audio-based information about tone on- and offsets and MIDI keystroke velocity). Thus, a combination of different measurements could lead to a more realistic account of cognitive and brain processes underlying music performance, and to a more behaviorally informed brain research in general (see also Makeig, Gramann, Jung, Sejnowski, & Poizner, 2009).

2.4.2. Setup

The basic concept of simultaneous recording EEG, MIDI, and motion capture data as used in Experiment 3 is depicted in Figure 2.5. All three different data streams are stored on separate computers. However, to synchronize and combine the data streams, the MIDI data recording software (FTAP; Finney, 2001a, 2001b) was modified to send synchronization signals in form of transistor-transistor logic (TTL) pulses synchronously with key press events to the EEG acquisition device. In addition, the motion capture system sends TTL pulses synchronously with each recorded frame to the EEG acquisition device. Thus, the timing information of each captured frame and of each keystroke are included in the EEG data, enabling the analysis of EEG data with respect to events in the movement data (and

MIDI data) and vice versa.

To test the reliability and temporal precision of the setup, several tests were performed, for example with respect to additional timing delays and accuracy of the TTL pulses (for details, see Maidhof, Kästner, & Makkonen, 2014). In sum, results showed that the setup worked reliably and with sufficient temporal precision for most ERP studies (although not in the preferable submillisecond range). For example, all TTL pulses were correctly received by the EEG device, the delays introduced by using MIDI and FTAP, compared to direct audio output of the digital piano, were on average 1.4 ms, and the audio signal started about 3 ms after the onset of the TTL pulse of the corresponding key press event, and this difference did not change over time.

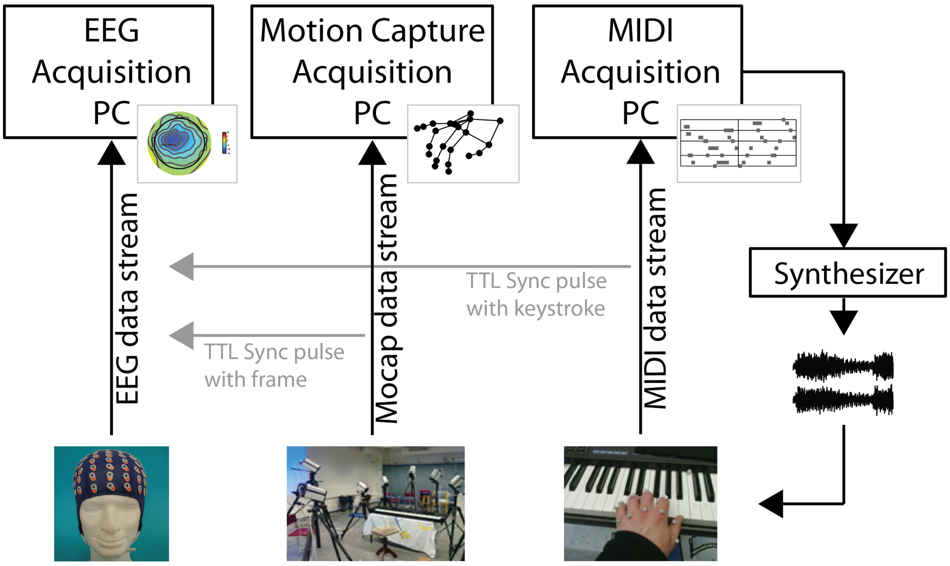

Figure 2.5.: Schematic illustration of the setup combining MIDI, EEG, and motion capture recordings. Figure from Maidhof, Kästner, and Makkonen (2014). ©2013 Psychonomic Society, Inc. With permission of Springer.

3. General Discussion and Further Directions

The main objective of this thesis was the investigation of the neurophysiological correlates of action monitoring and error detection processes during expert music performance. There are several theories and a wealth of empirical findings concerning action monitoring and error processing during simpler and rather artificial (button-pressing) tasks. In contrast, the processes during a more natural, highly complex goal-directed behavior that requires intensive year-long practice such as music performance have remained elusive.

Results of the first two experiments suggested that similar expectancy mechanisms operate during production and perception of music. However, the intention and action of producing a certain auditory effect enhances expectancies and thus influences the processing of the auditory input, as indicated by a larger FRN/N200 component in the action compared to the perception condition (Maidhof, Vavatzanidis, et al., 2010). The results of performance errors in Experiment 2 and 3 showed that error-related processing can occur independently of auditory feedback, as indicated by a pre-error negativity elicited already prior to the onset of erroneous notes, i.e. before participants could hear that they make an error. Additionally, detailed analysis of the underlying movements suggested that the observed pre-error slowing is not due to slower movements towards a wrong key, but to a prolonged key depression phase, which could indicate an attempt to cancel the sensory effects of an error (Maidhof, Rieger, et al., 2009; Maidhof, Pitkäniemi, & Tervaniemi, 2013). Both main findings were explained in terms of predictive mechanisms based on the common coding approach and on internal forward models. In the next two sections, I will discuss open issues and propose ways how these can be addressed in future research.

3.1. Feedback Processing

One issue that could be clarified in future studies is the impact of attention and increased levels of arousal on the ERP amplitude differences (independent of action-related expectancies) between the action and perception conditions. However, a simple attention-based account of the findings seems unlikely: if the FRN/N200 is strongly influenced by attention, it should have been increased in the perception condition (of Experiment 1), because manipulated tones in the perception condition were task-relevant, whereas participants were instructed to continue playing after pitch manipulations in the action condition. This is inconsistent with the present results. Nevertheless, one could argue that due to generally increased attention levels during music performance (and because it is hard to imagine that auditory feedback is completely task-irrelevant for musicians), any unexpected tone during performance would have a stronger attentional focus and consequently would elicit a larger ERP response (as was shown for other ERP components such as the N1, e.g. Hillyard, Teder-Sälejärvi, & Münte, 1998; Hillyard, Hink, Schwent, & Picton, 1973). One possibility to further investigate an attentional influence could be to confront participants with a second task unrelated to piano performance during the action condition, for example watching a silenced movie with the instruction to ignore the acoustic input and focus on the visual stimuli (a widely used task to test attentional influences on auditory processing; see, e.g., Brattico, Tervaniemi, Näätänen, & Peretz, 2006; Koelsch & Jentschke, 2008; Tervaniemi, Castaneda, Knoll, & Uther, 2006). Note that this procedure could also help to disentangle a potential overlap of the N2b with the FRN/N200 during the action condition. Another approach could be to make perception and performing conditions more similar, e.g. by instructing participants in the perception condition to make muted key presses on a piano (i.e., without auditory feedback), in addition to watching a silenced movie and to ignore the presented stimuli.

Another issue that needs more consideration is the potential role of internal action simulation mechanisms. The unexpectedness of manipulated tones during music perception might not only be due to expectancies based on the preceding musical context, but also to internal simulation mechanisms, so that pianists perceived the stimuli as auditory feedback to internally simulated actions (or as feedback to someone else's actions; Loehr, Kourtis, Vesper, Sebanz, & Knoblich, 2013). Such mechanisms should be operating only in musically trained participants, who have acquired bi-directional associations between sounds and actions. Therefore, the processing of manipulated tones should differ between musically trained and non-trained individuals, which could be tested in future studies. Furthermore, if internal simulation influences the FRN/N200, performing a second motor task during music perception should reduce the amplitude of this component more strongly than a non-motor task (e.g., a working memory test), because previous studies have shown that cortical motor areas are involved during internal action simulation (Lotze, Montoya, et al., 1999; Meister et al., 2004). Another way of addressing this issue would be a comparison between a condition where participants are explicitly instructed to imagine they were playing versus a condition where they are not (or instructed to imagine someone else is producing the perceived tones).

Given that the amplitude of the FRN/N200 appears to scale with expectancy (Nieuwenhuis, Holroyd, et al., 2004), this component could be used to investigate whether disruptions of feedback timing (i.e., asynchronies between key presses and auditory feedback) or disruptions of feedback content (i.e., pitch) are more unexpected. In addition, it would be interesting to investigate what influence the musical context and its underlying structure have on auditory feedback processing (e.g., is auditory feedback of metrically strong notes more important than feedback of metrically weak notes?). This could be achieved by comparing neural responses to different types and sizes of manipulations at different positions within a musical sequence.

Finally, the FRN/N200 could be used as an index of the strength of associations between actions and their auditory outcomes, and thus as a tool to investigate musical learning processes. The finding of a negative correlation of the FRN/N200 and duration of musical training in Experiment 2 (Maidhof, Vavatzanidis, et al., 2010), as well as the absence of an FRN/N200 in participants who had only moderate musical training in a similar study (Katahira, Abla, Masuda, & Okanoya, 2008) provided first evidence for an influence of musical training. However, because another study (Lutz, Puorger, Cheetham, & Jäncke,

2013) found no effect on the FRN during the course of auditory-motor learning in non-musicians, future studies have to address this issue in more detail.

3.2. Performance Errors

Considering the results of the performance errors, two main questions can be addressed in the future: (1) when and how exactly are errors detected and corrected, and (2) what are the causes for performance errors. The present results and results of similar studies (Herrojo Ruiz et al., 2009; Strübing et al., 2012; Herrojo Ruiz et al., 2011) showed consistently that error processing can occur independently of auditory feedback. However, the latency of the pre-error negativity appears to be influenced by the performance tempo and occurs earlier when notes were produced with a slower tempo (Maidhof, Rieger, et al., 2009) compared to a higher tempo (Maidhof, Pitkäniemi, & Tervaniemi, 2013; Herrojo Ruiz et al., 2009). Although this is an open research question, one might speculate that during slower performances, there is more time to prepare, initiate, and execute the following keystroke. Hence, error detection, error correction, or both, can occur at an earlier stage, for example because perhaps tactile feedback is earlier available and can be used to detect a discrepancy between predicted and actual feedback. The data of Experiment 3 (Maidhof, Pitkäniemi, & Tervaniemi, 2013) tentatively support the assumption that tactile feedback might be an important source of information for the detection of an upcoming error, which parallels the notion that tactile feedback contributes to the timing accuracy of pianists (Goebl & Palmer, 2008). However, these assumptions need to be further tested. Specifically, the present work leaves open the possibility that error-related processes could occur even earlier than tactile feedback of keystrokes, for example during the movement itself (based on information from muscles and joints), or based on internal information about the ongoing motor command and the predicted outcome of this motor command (similar to the view that the ERN during reaction time tasks is based on an internal signal and can be observed in deafferented patients; see Allain, Hasbroucq, Burle, Grapperon, & Vidal, 2004).

A related issue concerns the question whether the pre-error negativity reflects an error signal itself or is associated with implementation of behavioral adjustments (after error detection). A recent study using a similar paradigm as in this thesis addressed this issue by investigating the neural oscillatory activity during performance errors. Results showed increased theta and beta band oscillations that started already 120 ms before the onset of wrong notes (Herrojo Ruiz et al., 2011). Importantly, phase synchronization in the beta band between electrodes above the posterior frontomedial cortex (pFMC, including the ACC) and above the lateral prefrontal cortex (lPFC) increased around 100 ms prior to errors, and this increase was associated with more efficient corrective mechanisms. Based on accounts of action monitoring and cognitive control, these results suggest that also during music performance, the pFMC is involved in detecting unfavorable events (such as errors; MacDonald, Cohen, Stenger, & Carter, 2000; Botvinick, Braver, et al., 2001; Ridderinkhof et al., 2004), which is then signaled to the lPFC, assumed to be involved in cognitive control and implementing performance adjustments (Miller, 2000; Cavanagh et al., 2009; Wittfoth, Schardt, Fahle, & Herrmann, 2009). Because the neural generators of the pre-error negativity have been localized in the rostral ACC (BA 32; Herrojo Ruiz et al., 2009), it appears that the pre-error negativity reflects rather error detection processes.

Regarding the question why performance errors occur in the first place, it has been shown that the frequency and kind of errors are influenced by the musical structure (e.g., phrase structure) and underlying planning processes (Palmer & Van de Sande, 1993, 1995). In Experiment 3, ERPs elicited by correct notes that immediately preceded incorrect notes and ERPs elicited by correct notes at other positions in the sequences showed also a negative difference prior to note onsets (Maidhof, Pitkäniemi, & Tervaniemi, 2013). This suggests that neural processing differs between these notes, although they are correct in terms of pitch. One could speculate that this reflects that the monitoring system detected some problem already in motor planning, which might in turn foreshadow actual erroneous events (i.e., an incorrect pitch event). Clearly, more research is needed to investigate this issue in more detail.

Furthermore, several studies provided evidence that a temporarily disengagement of the action monitoring or attentional system might lead to incorrect behavior during rather monotonous tasks (Ridderinkhof, 2003; Hajcak, Nieuwenhuis, Ridderinkhof, & Simons, 2005; Eichele et al., 2008). It remains to be seen whether the same deficiencies or other factors contribute to committing performance errors during a highly complex task performed by highly-trained individuals.

Another open question relates to affective responses to errors, or even affective states contributing to performance errors. Previous research has shown that people not only have knowledge that they committed an error, but they also "feel" that they made an error: besides neural indices like ERN and Pe, also error-specific responses of the autonomic nervous system like heart rate deceleration and skin conductance responses can be observed during reaction-time tasks (e.g., Hajcak, McDonald, & Simons, 2003; Fiehler, Ullsperger, Grigutsch, & Cramon, 2003; Critchley, Tang, Glaser, Butterworth, & Dolan, 2005). It would be interesting to replicate these findings during complex tasks like music performance, and to investigate which factors might modulate these responses. On the other hand, stress can increase the neuromotor noise, which in turn can negatively influence ongoing motor behavior, for example in terms of reaction time and movement accuracy (Van Gemmert & Van Galen, 1997). It is thus appealing to examine interactions between factors that might cause error commission and affective reactions to such errors.

3.3. Concluding Remarks

In sum, this work represents a first step towards a better understanding of how musicians monitor their actions and detect or correct their errors. Although this is a relatively new line of research and many future avenues await further investigation, the approach to study expert performance seems promising, for several reasons. For example, music performance can be used as a tool to study action-based effects on perception, and vice versa. This in turn is useful to further test predictions derived from the common coding theory (e.g., with regards to the learning and strength of bidirectional associations between actions and action-effects), and can thereby inform the embodied approach to music cognition (for a recent review, see e.g. Maes, Leman, Palmer, & Wanderley, 2014). Similarly, theories of motor control (e.g. the concept of internal models) and action monitoring can be tested and refined by using music as a highly complex task performed by trained experts. Music performance provides also an excellent model to study joint action and social cognition (for a review, see DAusilio, Novembre, Fadiga, & Keller, 2015): usually, different performers and listeners are participating in a highly interactive situation, in which two or more performers have to coordinate their actions to reach a shared goal (for a recent study using a similar paradigm as in the present work to investigate musical joint action, see Loehr et al., 2013). Music might offer a solution for balancing the trade-off between ecological validity and experimental control in this domain.

Furthermore, knowledge about action monitoring and causes for (and consequences of) errors might improve skill learning, such as in music education (for reviews, see e.g. Parncutt, 2007; Hodges, 2009), or even other domains involving the planning and execution of complex action sequences (sports sciences, dance education, or similar). For that, a fruitful approach could be to draw on parallels to the speech domain. Recently, a *hierarchical state feedback control model* for speech production with multiple interacting levels of control was proposed, in which forward predictions are compared to auditory and somatosensory feedback (Hickok, 2012), and it is likely that a future model of music performance can benefit from such work in the speech domain.

Finally, the combination of detailed behavioral data acquired for example by motion capture systems with neurophysiological measures can help to derive at a more behaviorally informed brain research (see also Makeig et al., 2009), from which the above mentioned directions (action-perception interaction, motor control, skill learning, joint action) as well as novel music-therapeutic approaches could benefit.

4. Research Papers

4.1. Processing Expectancy Violations during Music Performance and Perception: An ERP Study

Maidhof, C., Vavatzanidis, N., Prinz, W., Rieger, M., & Koelsch, S. (2010). Processing expectancy violations during music performance and perception: an ERP study. *Journal of Cognitive Neuroscience*, *22*(10), 2401–2413. doi:10.1162/jocn.2009.21332

Processing Expectancy Violations during Music Performance and Perception: An ERP Study

Clemens Maidhof[1], Niki Vavatzanidis[1], Wolfgang Prinz[1],
Martina Rieger[1], and Stefan Koelsch[1,2]

Abstract

■ Musicians are highly trained motor experts with pronounced associations between musical actions and the corresponding auditory effects. However, the importance of auditory feedback for music performance is controversial, and it is unknown how feedback during music performance is processed. The present study investigated the neural mechanisms underlying the processing of auditory feedback manipulations in pianists. To disentangle effects of action-based and perception-based expectations, we compared feedback manipulations during performance to the mere perception of the same stimulus material. In two experiments, pianists performed bimanually sequences on a piano, while at random positions, the auditory feedback of single notes was manipulated, thereby creating a mismatch between an expected and actually perceived action effect (action condition). In addition, pianists listened to tone sequences containing the same manipulations (perception condition). The manipulations in the perception condition were either task-relevant (Experiment 1) or task-irrelevant (Experiment 2). In action and perception conditions, event-related potentials elicited by manipulated tones showed an early fronto-central negativity around 200 msec, presumably reflecting a feedback ERN/N200, followed by a positive deflection (P3a). The early negativity was more pronounced during the action compared to the perception condition. This shows that during performance, the intention to produce specific auditory effects leads to stronger expectancies than the expectancies built up during music perception. ■

INTRODUCTION

Producing music constitutes a complex interplay between motor, auditory, and somatosensory systems, with multiple processes interacting and overlapping in time. Imagine someone playing from memory a simple melody such as "Happy Birthday" on a piano. First, the notes and the order in which they are produced have to be retrieved from memory. Characteristics of the notes, such as the relative timing, duration, and intensity, also have to be remembered. Then, the appropriate actions have to be planned and executed. While executing the actions, their outcomes have to be monitored, which can in turn influence future actions. One fundamental aspect of music performance is the intention of musicians to produce specific auditory effects by executing certain actions. Thus, they expect to perceive a certain sound, that is, the auditory feedback of their action. However, only very little is known about the time course and the neural mechanisms underlying the processing of feedback during music performance. The present study addressed this issue by investigating the neurophysiological correlates of processing manipulated auditory feedback in skilled pianists in an action and a perception condition.

Skilled piano players have learned in thousands of hours of deliberate practice to produce specific auditory effects with highly accurate movements (see Sloboda, 2000; Palmer, 1997; Ericsson & Lehmann, 1996 for reviews). Accordingly, results of behavioral (Drost, Rieger, Brass, Gunter, & Prinz, 2005a, 2005b), electrophysiological (Bangert & Altenmüller, 2003), and neuroimaging studies (Haueisen & Knösche, 2001; for a review, see Zatorre, Chen, & Penhune, 2007) consistently showed pronounced coupling of auditory and motor systems in musicians. For example, the perception of potential action effects (i.e., tones) can induce the action which normally produces these tones (Drost et al., 2005b), and (pre)motor cortex of pianists exhibits activity during listening to well-known piano melodies (Haueisen & Knösche, 2001). Moreover, musically naïve participants show auditory–sensorimotor EEG coactivity already within 20 min of piano learning (Bangert & Altenmüller, 2003).

In contrast, the importance of auditory feedback for music performance (in terms of fluency of production) is unclear: Although the complete absence of feedback seems to have no effects on performance, manipulations of the synchronicity between a keypress and feedback (i.e., the delay of feedback), or the manipulation of the content of the feedback (i.e., pitch), can have profound effects (e.g., Finney & Palmer, 2003; Pfordresher, 2003; Finney, 1997; for reviews, see Pfordresher, 2006; Finney,

---

[1] Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, [2] University of Sussex, Falmer, UK

1999). Disruptive effects of pitch manipulations occur only when the perceived feedback resembles the intended sequence, but not when the feedback sequence is highly dissimilar to the intended sequence. For example, if feedback is random, it is assumed that pianists perceive the feedback as being unrelated to their planned actions. Thus, it appears that pianists rely on specific mappings of actions and their auditory effects, but that they may not rely on the presence of feedback per se (see Pfordresher, 2006).

However, it is also possible that disruptive effects of pitch manipulations are due to violations of musical expectancies built up during the perception of a specific musical context. This could explain the null effects of random or absent feedback, because if a musical context is lacking, no expectations about forthcoming events can be built up. However, it is difficult to disentangle effects of expectancy violations on the basis of performed action and expectancy violations on the basis of the perception of the preceding musical context on a behavioral level. Therefore, investigating the neural mechanisms underlying the processing of manipulated feedback by means of ERPs can help to clarify this issue.

In the present study, we compared the ERPs elicited during music performance ("action condition") with the ERPs elicited when participants only perceived such stimuli ("perception condition"). In the action condition, pianists produced bimanually fast sequences on a digital piano, while at random positions the auditory feedback of single keypresses was lowered by one semitone (which would normally be produced by the key adjacent to the actually pressed key). Thus, in the action condition, both action-related expectancies toward a tone (based on the performed action and the intention to produce a specific tone) and perception-related expectancies toward a tone (induced by the preceding musical context) were violated. In the perception condition, pianists listened to the material (including the same manipulations) without producing it. Thus, only perception-related expectancies toward a tone were violated. In both conditions (and in the two experiments described below), participants were informed about the occasional wrong pitches.

According to recent theories of action monitoring and cognitive control (Folstein & Petten, 2008; Botvinick, Cohen, & Carter, 2004; Nieuwenhuis, Holroyd, Mol, & Coles, 2004; van Veen, Holroyd, Cohen, Stenger, & Carter, 2004; van Veen & Carter, 2002), the rostral cingulate zone (RCZ) of posterior medial frontal cortex plays a key role in the processing of expectancy violations, performance monitoring, and the adjustment of actions for the improvement of task performance (for a review, see Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004). Different ERP components are taken to receive contributions from neural generators located in the RCZ, foremost the error-related negativity (ERN), the feedback ERN, and the N200 (Ridderinkhof et al., 2004). The feedback ERN and the N200 are particularly relevant for the present study. The feedback ERN is elicited around 250 msec after negative performance feedback (compared to positive feedback), and after feedback stimuli indicating loss (or punishment) in time estimation tasks, guessing tasks, and gambling tasks (e.g., Hajcak, Moser, Holroyd, & Simons, 2007; Hajcak, Holroyd, Moser, & Simons, 2005; Miltner, Braun, & Coles, 1997). Importantly, a feedback ERN-like component can also be observed in the absence of responses on the part of the participants (see Tzur & Berger, 2007, 2009; Donkers, Nieuwenhuis, & van Boxtel, 2005). The N200 component (which is similar to the feedback ERN in latency and scalp distribution) is elicited when a mismatch between an expected and an actual sensory event is detected (see e.g., Ferdinand, Mecklinger, & Kray, 2008; Kopp & Wolff, 2000). However, there is an ongoing debate as to whether the feedback ERN reflects a subcomponent of the N200 (and thus, a different subfunction of ACC; Folstein & van Petten, 2008).

These negativities (ERN, feedback ERN, and N200) are usually followed by P300 potentials which can often be decomposed into an early (P3a) and a later subcomponent (P3b).[1] The P3a has a fronto-central scalp distribution and is considered to reflect the automatic shift of attention to deviant stimuli. The P3b shows maximal amplitude values over parietal leads and reflects the conscious detection of target stimuli (Comerchero & Polich, 1999; Mecklinger & Ullsperger, 1995; Donchin & Coles, 1988).

Based on the findings that unexpected feedback elicits a feedback ERN/N200, followed by P300 potentials (even in the absence of response tasks), we hypothesized that auditory feedback manipulations during piano playing and manipulated tones during the mere perception of the stimuli would elicit such an ERP pattern. We further hypothesized that amplitude values of the negative potential would be increased during piano playing (action condition, i.e., when expectancies for tones are based on both the actions and the preceding musical context).

Because pianists played bimanually (pressing two keys synchronously in octaves), a pitch manipulation of one of the notes violates also the physical regularity of octave intervals and reflects therefore an auditory oddball. Thus, such stimuli are also likely to elicit a mismatch negativity (MMN; indexing the detection of deviant sounds in an otherwise regular stimulus sequence; Winkler, 2007; Näätänen, 1992), which would partially overlap with the feedback ERN/N200. However, if amplitude values of the negative potentials in the present study would differ between the action and perception conditions, it is unlikely that they reflect simply an MMN response, because previous studies found no differences in amplitude of the MMN between conditions in which participants themselves trigger unexpected auditory oddballs (compared to when unexpected tones are presented to the participants; Nittono, 2006), and when participants anticipate a standard tone but trigger instead a deviant tone (compared to when participants anticipate a deviant tone and trigger a deviant tone; Waszak & Herwig, 2007). However, because these studies found that the amplitude of the P3a is modulated by the anticipation of participants (larger when deviant tones were unexpected), we hypothesized that the P3a component would show a larger amplitude in the action compared to the perception condition.

EXPERIMENT 1

Methods

Participants

Eight trained right-handed pianists (4 women, 24.4 ± 2.3 years old) participated in the experiment. Participants had, on average, 18.5 (±2.9) years of formal piano training and were students at the conservatory in Leipzig (Hochschule für Musik und Theater Felix Mendelssohn Bartholdy).

Material and Apparatus

The pianists performed on a Yamaha digital piano (Clavinova CLP 130), and listened to their performances via AKG 240 studio headphones at comfortable listening levels (approximately 55 dB, dependent on the velocity of the keypresses). All tones had the standard Musical Instrument Digital Interface (MIDI) piano timbre generated by a Roland JV-2080 synthesizer (Hamamatsu, Japan).

In the *action condition*, participants had to produce major scales and two sequence patterns bimanually (parallel in octaves; see Figure 1). Each participant performed 12 blocks (4 blocks of Pattern A, 4 blocks of Pattern B, and 4 blocks of musical scales); in each block, scales or patterns had to be produced in different major keys in one of two orders: C-Major/E-Major/D-Major/F#-Major, or G-Major/B-Major/A-Major (in case of scales, these sequences were repeated). The tempo for the scales was 144 beats per minute (bpm), and 69 bpm for the patterns, that is, each note event (consisting of two notes played simultaneously by both hands) had to be produced, on average, every 104 msec in the scales blocks, and every 217 msec in the patterns blocks. Randomly between every 40th and 60th produced note (i.e., randomly at either the left or the right hand), the pitch of one note was lowered by one semitone. That is, the auditory feedback of one key stroke was manipulated, and pianists did not hear the corresponding tone of the pressed key, but a tone with a pitch lowered by one semitone, sounding as if the pianist committed an error with one of the two hands.

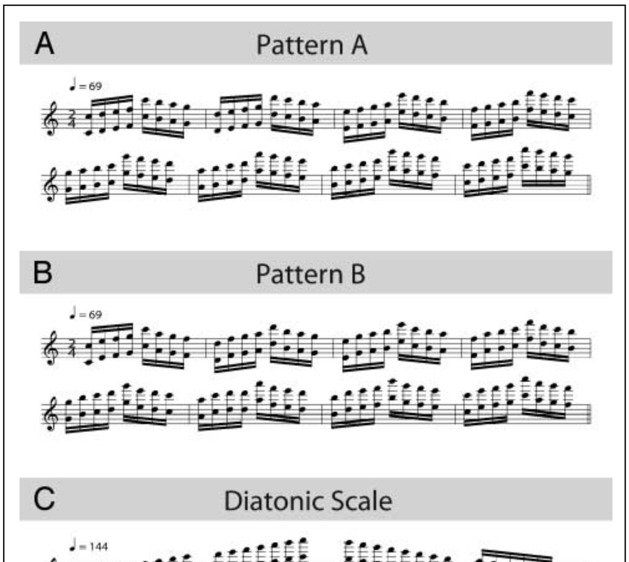

Figure 1. Examples of the stimulus material. (A) Pattern A in C-Major; (B) Pattern B in C-Major; and (C) a diatonic scale in C-Major.

In the *perception condition*, participants listened to prerecorded versions of these stimuli (with the same stimulus types and order of keys), which were performed by a pianist who did not participate in the study. Analogously to the action condition, the pitch of one tone was randomly between every 40th and 60th tone lowered by one semitone.

Procedure

Blocks of the action and the perception conditions occurred in alternating order. The order of blocks was pseudorandomized, with the constraint that no identical stimulus type (scale, Pattern A, Pattern B) occurred in direct succession. In the action condition, pianists were instructed to play as accurately as possible in the given tempo (during their performances, they heard a standard metronome). They were informed about the occasional wrong feedback, but were asked not to stop after feedback manipulations or if they committed an error. They were told that always performing correctly would be very difficult and committing errors was probably sometimes unavoidable. After participants were familiarized with the task and the stimuli, they were blindfolded to exclude visual feedback and to decrease the likelihood of eye artifacts caused by the observation of the hand and finger movements. In the perception condition (in which participants were also blindfolded), their task was to silently count any wrong pitches and to report this number verbally after each block. To detect the targets, participants had to pay attention to all tones.

Data Recording and Analysis

The musical data (in form of MIDI data) were recorded and played back with a modified version of the open source program "aplaymidi" (www.alsa-project.org), which also realized the feedback manipulations. To synchronize MIDI and EEG data, this program sent concurrently with feedback manipulations and every fifth keypress trigger signals to the EEG acquisition computer. The MIDI information (including keypress timing, velocity, and pitch) was saved on a hard disk, so that triggers for all key strokes could be reconstructed off-line for the EEG data evaluation.

The EEG was recorded with 60 Ag/AgCl scalp electrodes placed according to the extended 10–20 system (see Figure 2), referenced to ML (left mastoid). The ground electrode was located on the sternum. The horizontal electrooculogram (HEOG) was recorded bipolarly from electrodes placed on the outer left and right canthus and the vertical EOG (VEOG) from electrodes placed on the tip of the nose and Fpz. Impedance was kept below 5 kΩ. EEG signals were digitized with a sampling frequency of 500 Hz.

After data acquisition, EEG data were re-referenced to the arithmetical mean of both mastoid electrodes, and bandpass filtered (0.25–25 Hz band pass, finite impulse response [FIR]). Artifacts caused by eye movements were rejected off-line whenever the standard deviation within a 200-msec window centered around each sampling point exceeded 50 μV in the EOG. Artifacts caused by drifts and body movements were eliminated by rejecting sampling points whenever the standard deviation within a 200- or 800-msec window exceeded 40 μV at any electrode. Trials with typical eye blinks were marked and corrected by applying EOG correction (EEP software; ANT Software B.V., The Netherlands). ERPs were computed for 1000 msec time-locked to the onset of the keypresses or tones with a baseline ranging from −200 to 0 msec. Importantly, ERPs (and interonset intervals [IOIs]; see below) of manipulated and correct tones were only computed if no self-produced error or manipulation occurred within the preceding or subsequent second of that event (i.e., ERPs were only computed if they occurred in a 2-sec time window in which no self-produced error or manipulation occurred).

For statistical analysis, mean ERP amplitude values were calculated for two ROIs (see Figure 2). Because the feedback ERN/N200 and the P3a show a fronto-central distribution (e.g., Hajcak et al., 2005, 2007; Goldstein, Spencer, & Donchin, 2002; Simons, Graham, Miles, & Chen, 2001; Katayama & Polich, 1998; Miltner et al., 1997), we chose a midline–anterior ROI including the electrodes AFz, Fz, FCz, and Cz. Because the P3b shows a parieto-central distribution (e.g., Goldstein et al., 2002; Simons et al., 2001; Katayama & Polich, 1998), we chose a midline–posterior ROI including the electrodes CPz, Pz, POz, and Oz. Visual inspection of the effects in our study confirmed the selection of electrodes for these ROIs.

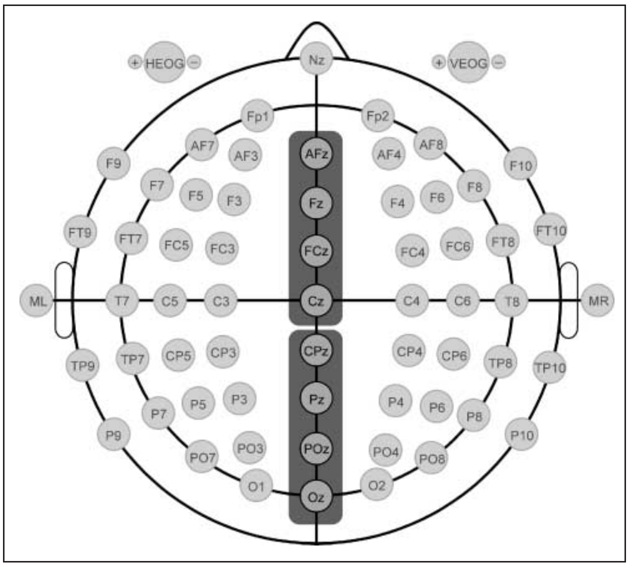

Figure 2. Electrode locations. The highlighted areas show the ROIs used for statistical analyses in the two experiments. Electrode FPz was additionally measured in Experiment 1.

After the rejection procedures, there were, for each participant in the action condition, on average, 33 trials with feedback manipulations and 279 trials with correct feedback. Participants committed, on average, 45 pitch errors (i.e., one hand presses the correct key, while the other hand presses simultaneously an incorrect key). In the perception condition, there were, on average, 71 trials with pitch manipulations and 646 trials with correct pitches. ERPs were statistically evaluated by repeated measures ANOVAs with factors condition (action, perception) and tone (regular, manipulated). Time windows (centered around the grand-average peak latencies) for statistical analyses of ERP data were 140–240 msec (early negativity), 280–330 msec (P3a), and 370–430 msec (P3b). Before calculating the ANOVAs, Kolmogorov–Smirnov tests had shown that all variables in the analyses did not deviate from a standard normal distribution ($.25 < p < .99$ in all tests).

For the analyses of the behavioral data, we calculated the IOIs between the onsets of two succeeding correct notes (played by the same hand) and the IOIs between the onset of a manipulated note and the succeeding note (played by the same hand). Whenever an IOI exceeded 1000 msec, this IOI was not analyzed. To test whether participants showed performance slowing after feedback manipulations, IOIs after manipulations were statistically compared to IOIs between correct tones. Note that the IOI between correct tones is also the estimate of the performed tempo.

Results

Behavioral Results

There was no difference between IOIs succeeding feedback manipulations ($M = 214$ msec, $SD = 16$ msec) and IOIs between correct tones ($M = 216$ msec, $SD = 10$ msec; $p > .4$). The average tempo of 216 msec was slightly slower than the instructed tempo because, for some participants, it was difficult to perform in the demanded tempo (thus reducing the overall tempo).

ERP Results

Figure 3 shows the grand-average waveforms time-locked to the onset of the notes (see Figure 4 for mean amplitudes of ERP effects). In the action condition, feedback manipulations (compared to notes with regular feedback) elicited a negative deflection that was maximal around 200 msec

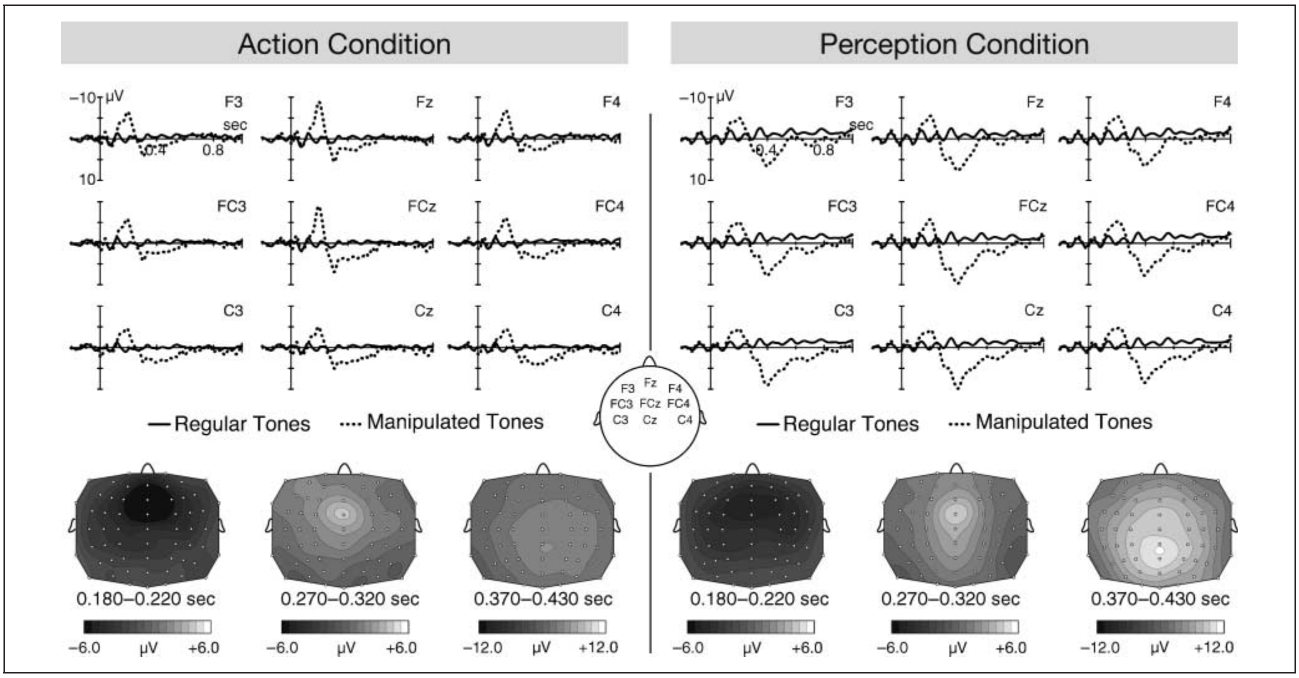

Figure 3. Experiment 1. Grand-average ERPs (and scalp distributions) elicited by the auditory feedback manipulations during the action condition (left) and by the task-relevant manipulated tones during the perception condition (right).

and showed a fronto-central scalp distribution [main effect of tone: $F(1, 7) = 45.43$, $p = .0003$]. This negativity was followed by two subsequent positive components peaking around 300 and 400 msec [main effects of tone: $F(1, 7) = 5.38$, $p = .054$ and $F(1, 7) = 28.98$, $p = .001$, respectively], with the former one showing a slightly more central distribution, and the latter one a centro-parietal distribution (see Figure 3). In the perception condition, manipulated tones (compared to regular tones) elicited a negativity that was maximal around 200 msec [main effect of tone: $F(1, 7) = 112.09$, $p < .0001$]. The negativity was followed by two positive peaks around 300 and 400 msec [main effects of tone: $F(1, 7) = 11.48$, $p = .012$ and $F(1, 7) = 36.54$, $p = .0005$, respectively], showing the same topography as the positive potentials in the action condition (see Figure 3).

Comparison between action and perception condition. The amplitude of the early negative potential was larger in the action condition compared to the perception condition (see Figure 4 for mean amplitude values of ERP effects, and Figure 5 for difference waves): An ANOVA with factors condition (action, perception) and tone (regular, manipulated) showed a main effect of tone [$F(1, 7) = 74.46$, $p < .0001$], and a two-way interaction [$F(1, 7) = 6.1$, $p = .025$]. The amplitude of the early positivity around 300 msec did not differ between the two conditions: An ANOVA with the same factors for the 280–330 msec (P3a) time window showed a main effect of tone [$F(1, 7) = 8.37$, $p = .023$], but no interaction [$F(1, 7) = 0.48$, $p = .51$]. The late positivity (maximal around 400 msec) was clearly larger in the perception condition compared to the action condition: An ANOVA for the 370–430 msec (P3b) time window over the midline–posterior ROI showed main effects of tone [$F(1, 7) = 44.07$, $p = .0003$] and condition [$F(1, 7) = 15.73$, $p = .0054$], as well as a two-way interaction [$F(1, 7) = 16.18$, $p = .0005$].

Discussion

In Experiment 1, pitch manipulations of the auditory feedback during piano performance (action condition) and pitch manipulations during the perception of such stimuli

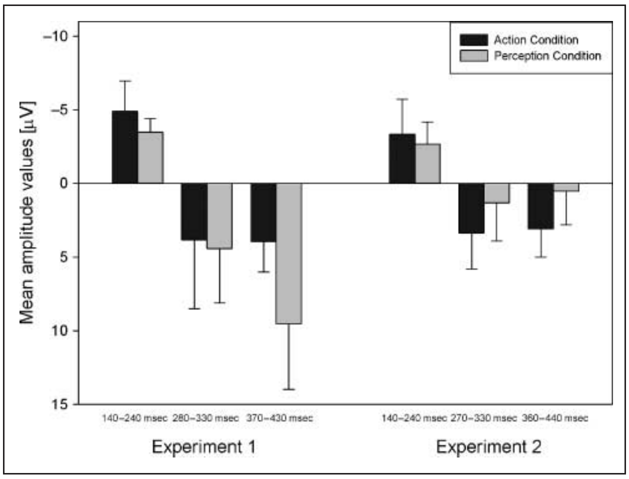

Figure 4. Mean amplitude values (and the standard deviations) for the ERP effects in all three time windows in the two experiments. Note that negative is plotted up.

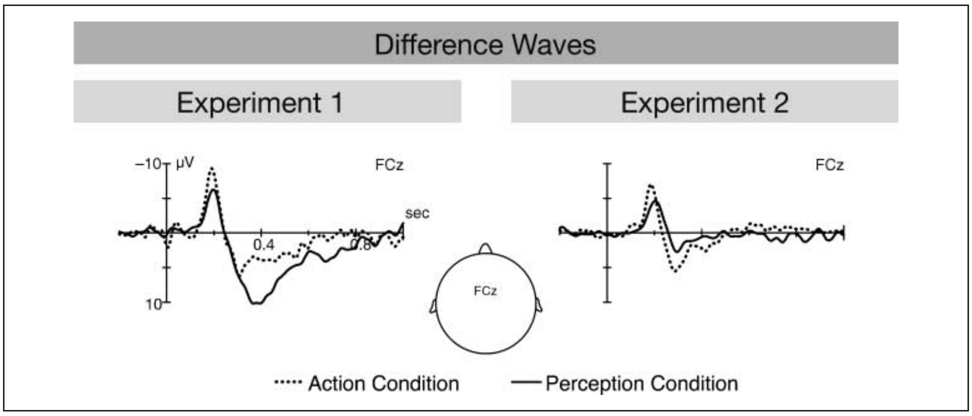

Figure 5. Difference waves of Experiments 1 and 2. Difference waves were obtained by subtracting the ERPs elicited by the standard tones from the ERPs elicited by the manipulated tones in the action and perception conditions, respectively.

(perception condition) elicited a very similar ERP pattern: a negative potential around 200 msec, followed by two positive peaks around 300 (P3a) and 400 msec (P3b), respectively. The scalp distributions of all components in the action condition were also highly similar to the distributions of the components in the perception condition. Because manipulated tones in the action condition violated both action-related expectancies and perception-related expectancies (leading to a supposed overlap of ERPs related to action as well as to perceptual processes), we compared the ERPs elicited in the action condition to those elicited in the perception condition. Note that, in the ERPs of the action condition, any keypress-related effects, and any effects related to the metronome clicks, are cancelled out in the difference waveforms (see Figure 5) because keypresses and metronome clicks were present during the presentation of both manipulated and correct tones.

The early negativity, which was more pronounced during the action compared to the perception condition, resembles the feedback ERN/N200, reflecting general expectancy-related mechanisms, probably irrespective of whether the outcome of an event is worse or better than expected (Ferdinand et al., 2008; Oliveira, McDonald, & Goodman, 2007). It is conceivable that a feedback ERN/N200 was also elicited in the perception condition, because a previous study found a feedback ERN-like waveform also in an experiment which required no actions, or responses, on the part of the participants (Donkers et al., 2005). In addition, two other studies (Tzur & Berger, 2007, 2009) reported feedback ERN-like deflections after rules (i.e., expectations) were violated in tasks without overt responses.

Thus, the results suggest that when pianists were actually performing, pitch manipulations of the auditory feedback were more unexpected than pitch manipulations when pianists were only perceiving the sequences, because in the former case, their expectancy toward a tone was based on the action (or intention) to produce a tone (in addition to the expectancy induced by the preceding musical context), whereas in the latter case, their expectancy was based only on the preceding musical context. This is reflected in the larger early negativity in the action condition compared to the perception condition.

With regards to auditory–perceptual processes, the manipulated tones violated the expectancies of listeners/performers presumably in two ways: (a) with regard to tonal regularity (when an out-of-key note was introduced) and (b) with regard to acoustic regularity, because standard tones formed an octave interval (i.e., a frequency ratio of 2:1), whereas manipulated tones formed a major seventh (i.e., a frequency ratio of about 1.9:1). Such acoustic irregularities usually elicit an MMN/N2b complex (the N2b being due to the controlled and conscious detection of task-relevant deviants; Novak, Ritter, Vaughan, & Wiznitzer, 1990), and tonal regularities are prone to elicit an ERAN/N2b complex (reflecting the processing of structurally unexpected notes within musical contexts; Koelsch, 2005; Koelsch, Gunter, Friederici, & Schröger, 2000). However, it is not plausible to assume that only MMN/N2b (or ERAN/N2b) potentials account entirely for the negativities, because their amplitudes differed between the two conditions, and previous studies found no MMN amplitude differences when participants produced or only listened to unexpected auditory oddballs (Nittono, 2006). Furthermore, the MMN is not influenced by the anticipation of (and thus, expectancy toward) deviant tones (Waszak & Herwig, 2007), nor by prior knowledge of deviant stimuli (Rinne, Antila, & Winkler, 2001). In addition, another recent study using a similar design as the performance condition in the present study (but with a different perception condition) reported no significant differences between feedback manipulations which introduced an auditory oddball (an out-of-key note) and those feedback manipulations which did not (in both cases, a negative potential around 200 msec was elicited; Katahira, Abla, Masuda, & Okanoya, 2008). However, data of Experiment 1 leave open the possibility that the negative potential in the perception condition overlaps with an N2b component; therefore, we conducted Experiment 2, in which the task of the participants in the perception condition was varied. Experiment 2 will also further address possible influences of the MMN on the observed negative potentials in both conditions.

The amplitude of the P3a was, contrary to our hypothesis, not larger in the action condition than in the perception condition, possibly because the P3a was overlapped by the P3b (showing larger amplitudes during the perception condition) elicited by the task-relevant deviant tones in the perception condition.

EXPERIMENT 2

As already mentioned, the early negativity elicited by the manipulated notes in the perception condition was perhaps overlapped, in part, by an N2b, because manipulated tones were task-relevant targets. To estimate the contribution of an N2b component to this negative potential, we conducted another experiment that was identical to Experiment 1 (i.e., it consisted of an action and a perception condition), except that, in the perception condition, manipulated notes were task-irrelevant. If the negative deflection in the perception condition of Experiment 1 reflects expectancy-related mechanisms (as reflected in a feedback ERN/N200), rather than the detection of task-relevant deviant stimuli (as reflected in an N2b potential), then it should be observed irrespective of the task in the perception condition. Because we expected that manipulated tones in the action condition would violate action- and perception-related expectancies (in contrast to the violation of only perception-related expectancies in the perception condition), we further assumed that the negativity would again be more pronounced in the action than in the perception condition. In addition, we hypothesized that the posterior P3b would be smaller in the perception condition relative to the action condition because manipulated tones were task-irrelevant in the perception condition.

Methods

Participants

Twelve right-handed trained pianists (7 women, 24.2 ± 2.6 years old) took part in the second experiment. None of the participants had participated in the first experiment. Participants had, on average, 14.9 (±4.8) years of formal piano training and were current or former students at the conservatory in Leipzig (Hochschule für Musik und Theater Felix Mendelssohn Bartholdy).

Material and Apparatus

Stimulus material, type and frequency of manipulated tones, as well as equipment, were identical to Experiment 1.

Design and Procedure

Experiment 2 was identical to Experiment 1 except that: (1) Manipulated tones were task-irrelevant in the perception condition; (2) stimuli were presented in short blocks (duration ranging from ca. 21 sec to ca. 60 sec), and the task of the participants was to compare the duration of one block with the preceding block, and to give a verbal response after each block (in contrast to the target detection task Experiment 1); (3) a block design was used, and all participants were tested first in the perception condition, and then (after a training phase to familiarize participants with the task and the stimuli) in the action condition; (4) Experiment 2 consisted of twice as many blocks as Experiment 1 (24 blocks in each, perception and action, condition: 8 blocks of Pattern A, 8 blocks of Pattern B, and 8 blocks of musical scales); and (5) participants did not hear a metronome in the action condition, but were instructed to play in the same tempo that they heard in the perception condition. If they were not able to do so, they chose their fastest possible tempo.

Data Recording and Analysis

MIDI data were processed with a modified version of the open source software "FTAP" (Finney, 2001a, 2001b), which sent simultaneously with every fifth keypress and with feedback manipulations trigger signals to the EEG acquisition computer.

EEG recordings were identical to Experiment 1, except that electrode Fpz was excluded and the VEOG was recorded with two electrodes beneath and above the left eye. After data acquisition, EEG data were downsampled to 250 Hz to reduce the data size, re-referenced to the arithmetical mean of both mastoid electrodes, and an independent component analysis with standard parameters for artifact removal as implemented in EEGLAB 4.51 (Swartz Center for Computational Neurosciences, La Jolla, CA; www.sccn.ucsd.edu/eeglab; Delorme & Makeig, 2004) was performed. After calculating the independent components, artifactual components were subtracted from the data. EEG data were filtered (0.25–25 Hz band pass, FIR) and the same rejection procedure was applied as in Experiment 1, except that we lowered the rejection criteria to 30 μV. Criteria for computing the ERPs and ROIs (see Figure 2) were the same as in Experiment 1. In the action condition, there were, for each participant, on average, 137 trials with manipulated feedback and 894 trials with correct feedback. Participants committed, on average, 72 pitch errors. In the perception condition, there were, on average, 113 trials for manipulated tones and 785 trials for correct tones. Time windows (chosen based on the same criteria as in Experiment 1) for statistical analyses were: 140–240 msec (early negativity), 270–330 msec (P3a), and 360–440 msec (P3b). To test the differences between the two conditions (actions and perception) and whether these differed between the two experiments, we conducted ANOVAs with condition (action, perception) and tone (manipulated, correct) as within-subject factors, and experiment (first, second) as between-subjects factor (over the same ROIs as in Experiment 1). Kolmogorov–Smirnov tests had shown, prior to the calculation of the

ANOVAs, that all variables in the analyses did not deviate from the standard normal distribution ($.39 < p < .99$ in all tests). The analysis of the behavioral data was the same as in Experiment 1.

To estimate the localization of the neural generators of the negativities, we used standardized low-resolution electromagnetic tomography (sLORETA; Pascual-Marqui, 2002), which computes the current density for 6239 voxels in the cortical gray matter. This method makes no a priori assumptions about the locus, number, and orientation of sources, only implicating that neighboring voxels should have a maximally similar electrical activity. However, the results of the sLORETA analysis should be considered somewhat cautiously, because we were not able to localize early sensory evoked potentials for control purposes due to the tempo of the performances/stimuli (which prohibited elicitation of clear P1, N1, or P2 components).

Results

Behavioral Results

IOIs succeeding feedback manipulations were nominally longer ($M = 337$ msec, $SD = 93$ msec) than the IOIs between correct notes ($M = 322$ msec, $SD = 77$ msec). However, this difference was not statistically significant [$t(11) = 1.22, p = .25$]. Note also that the overall tempo, as indicated by the IOIs between correct notes (322 msec), was slower than initially instructed (and slower than the average tempo in Experiment 1: Mann–Whitney test: $z = -3.09, p = .002$). This was presumably due to the fact that, in this experiment, participants chose their own fastest possible tempo whenever they were not able to perform in the instructed tempo.

ERP Results

In the action condition, feedback manipulations (compared to notes with correct feedback) elicited a negativity that was maximal around 200 msec [main effect of tone: $F(1, 11) = 23.7, p = .0005$; peak latency at FCz: ca. 188 msec], and showed a fronto-central scalp distribution (see Figure 6; see Figure 4 for mean amplitudes of ERP effects). The negativity was followed by a P3a around 300 msec [main effect of tone: $F(1, 11) = 22.88, p = .0006$] with a slightly more central distribution, and by a P3b around 400 msec [main effect of tone: $F(1, 11) = 30.67, p = .0002$], showing a parietal distribution. In the perception condition, manipulated tones (compared to regular tones) elicited a negativity that was maximal around 200 msec over frontal electrodes [main effect of tone: $F(1, 11) = 37.7, p < .0001$]. The negativity was followed by a fronto-central P3a around 300 msec, which was statistically not significant [main effect of tone: $F(1, 11) = 3.09, p = .106$]. No parietal P3b was elicited (main effect of tone: $F < 1$).

To further investigate possible influences of the MMN and ERAN on the negative potential elicited in the action condition, we also analyzed the ERPs elicited during the generation of self-performed errors. Errors were defined as playing an incorrect key with one hand while pressing the correct key with the other hand (i.e., errors were acoustically similar to the feedback manipulations).[2] Results

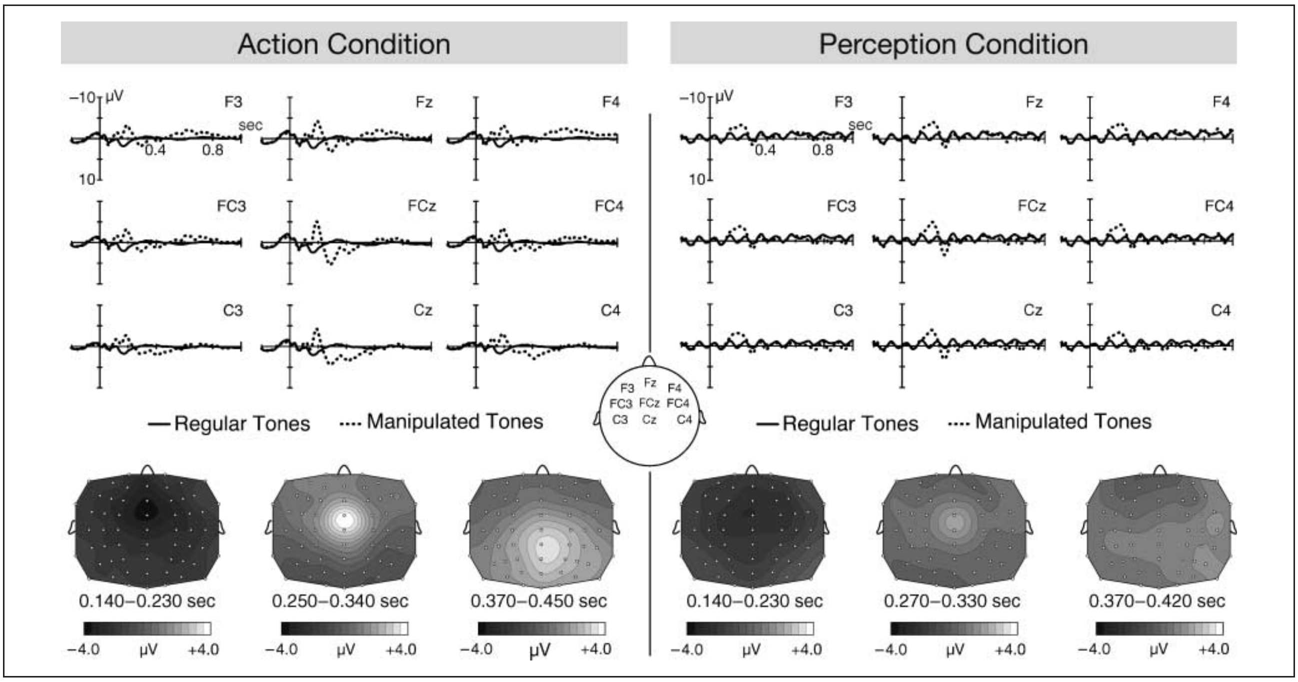

Figure 6. Experiment 2. Grand-average ERPs (and scalp distributions) elicited by the auditory feedback manipulations during the action condition (left) and by the task-irrelevant manipulated tones during the perception condition (right).

(see Figure 7) showed no negative potential around 200 msec ($F < 1$), but a significant difference prior to the onset of the feedback [$F(1, 9) = 8.33, p = .018$], and a positivity around 300 msec after feedback onset [$F(1, 9) = 13.08, p = .0056$].

To examine whether musical expertise (indicated by the duration of training) is related to the observed negativity in the action condition, we calculated the correlation between duration of musical training (in years) and the amplitude of the (negative) difference potential (tones with manipulated feedback minus tones with correct feedback) for electrode Fz in a time window ranging from 140 to 240 msec. Results showed a negative correlation between training and amplitude [Pearson's correlation coefficient: $r(12) = -.577, p = .049$], indicating that pianists with longer training showed a larger negativity.

Comparison between action and perception condition (and between experiments). The amplitude of the early negativity was larger in the action compared to the perception condition, as in Experiment 1: An ANOVA with condition (action, perception) and tone (standard, manipulated) as within-subjects factors, and experiment (first, second) as between-subjects factor, showed a main effect of tone [$F(1, 18) = 105.88, p < .0001$], an interaction between tone and condition [$F(1, 18) = 5.01, p = .038$], an interaction between condition and experiment [$F(1, 18) = 4.86, p = .041$], but no interaction between condition, tone, and experiment [$F(1, 18) = 0.66, p = .43$], indicating that the difference between action and perception conditions did not differ between the two experiments (see also Figure 4 for mean amplitude values). The amplitude of the P3a in Experiment 2 was more pronounced in the action condition than in the perception condition: An analogous ANOVA for the P3a time window showed a main effect of tone [$F(1, 18) = 19.33, p < .0001$], and an interaction between condition, tone, and experiment [$F(1, 18) = 6.8, p = .018$]. Separate ANOVAs with factors condition and tone for each experiment showed main effects of tone in Experiment 1 [$F(1, 7) = 8.37, p = .023$] and in Experiment 2 [$F(1, 11) = 13.23, p = .0039$], but only in Experiment 2 was there an interaction between condition and tone [$F(1, 11) = 9.27, p = .011$], indicating that the amplitude of the P3a elicited in Experiment 2 was larger during the action than during the perception condition. The P3b elicited in Experiment 2 was more pronounced in the action compared to the perception condition: An ANOVA for the P3b time window showed a main effect of tone [$F(1, 18) = 65.1, p < .0001$], and interactions between condition and tone [$F(1, 18) = 5.56, p = .03$], condition and experiment [$F(1, 18) = 39.38, p < .0001$], tone and experiment [$F(1, 18) = 19.4, p < .0001$], and between condition, tone, and experiment [$F(1, 18) = 35.71, p < .0001$]. An ANOVA for the data from Experiment 2 showed main effects of tone [$F(1, 11) = 12.67, p = .0045$] and condition [$F(1, 11) = 20.53, p = .0009$], and an interaction between condition and tone [$F(1, 11) = 14.72, p = .0028$], indicating that the P3b was larger in the action compared to the perception condition (see Results of Experiment 1 for other statistical results).

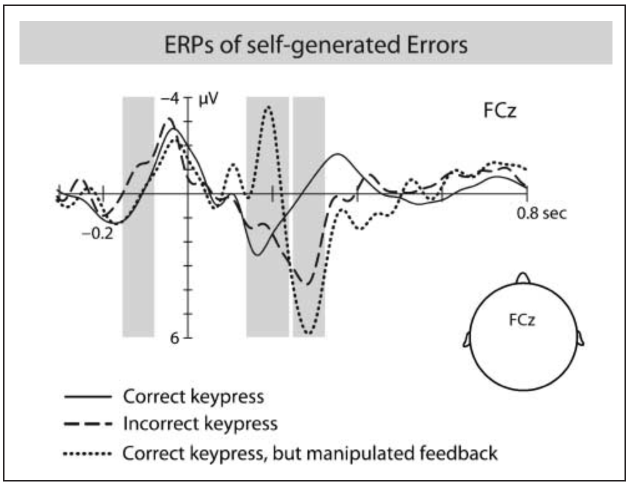

Figure 7. ERPs elicited by correct feedback, feedback manipulations, and self-generated errors. The gray-shaded areas represent the time windows used for statistical analyses. Results showed no negativity after performance errors (compared to correct notes) in the time window of the feedback ERN/N200, elicited by the feedback manipulations. Performance errors only elicited an increased negativity prior to feedback onset and an increased positive potential around 280 msec after feedback onset (for further details on the ERPs of self-generated errors, see Maidhof et al., 2009).

Source Localization

Results of the sLORETA analysis (see Figure 8) suggest that the main neural generators of the negative potential elicited during the action condition (sLORETA time window: 172–184 msec) are located in the RCZ of the posterior medial frontal cortex (Talairach coordinates $x = -5, y = 16, z = 27$; corresponding to Brodmann's area 24). Main generators of the negative potential elicited during the perception condition (sLORETA time window: 208–216 msec) were also located in the RCZ, although slightly more superior–posterior compared to the generators yielded for the action condition (Talairach coordinates $x = -15, y = 11, z = 36$; corresponding to Brodmann's area 24/32).

Discussion

The aim of Experiment 2 was to estimate the influence of an N2b on the negative potential observed in the perception condition. We hypothesized that, if the negativity reflects expectancy-related processes (as indexed by a feedback ERN/N200) and not only the detection of task-relevant targets (as indexed by an N2b), it should also be elicited by task-irrelevant manipulations. Furthermore, we expected (as in Experiment 1) an enlarged negativity during the action condition compared to the

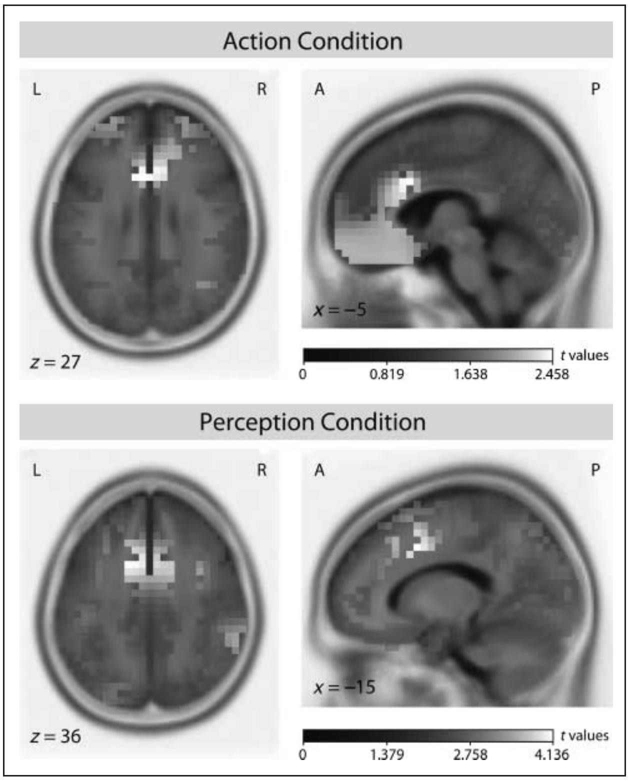

Figure 8. Source localization. Results of the sLORETA analysis suggest that the main neural generators of both negativities lie within the RCZ (Talairach coordinates: $x = -5$, $y = 16$, $z = 27$ in the action condition, and $x = -15$, $y = 11$, $z = 36$ in the perception condition).

perception condition. Results showed that manipulated tones in both conditions elicited early negative potentials with maximal amplitudes around 200 msec, and with larger amplitudes in the action compared to the perception condition (consistent with results of Experiment 1). Similarly, the P3a was more pronounced in the action than in the perception condition. The absence of a P3b during the perception condition reflects that the pitch manipulations were task-irrelevant for the participants (in contrast to Experiment 1, where pitch manipulations were task-relevant). Because the N2b component is usually observed in combination with a P3b (Novak et al., 1990), we therefore conclude that the observed negative potential is not an N2b.

Although the early negativities observed during the action and the perception conditions most presumably reflect, at least in part, a feedback ERN/N200 component, it might well be the case that they overlap with other components, such as the MMN (Winkler, 2007; Näätänen, 1992), the ERAN (Koelsch, 2005, 2009; Koelsch et al., 2000), or—in the action condition—the N2b. Based on the present data, the different contributions of these components cannot be disentangled. However, there are four reasons rendering it unlikely that the early negativities were simply MMN or ERAN potentials: Firstly, the additional analysis of the performance errors of the pianists showed no negative potential in the time range of the feedback ERN/N200, although self-performed errors are acoustically similar to the feedback manipulations. Secondly, the results of the source localization suggest that the neural generators of both negativities (action and perception condition) lie within the RCZ, which is consistent with an explanation in terms of feedback ERN/N200 (for a review, see Ridderinkhof et al., 2004). Interestingly, in another recent study investigating the human action monitoring system during piano performance (Herrojo Ruiz, Jabusch, & Altenmüller, 2009), very similar brain regions (BA 24 of the rostral ACC) generated a negative ERP preceding the onset of performance errors. This corroborates previous findings (see Ridderinkhof et al., 2004) indicating that the RCZ plays a key role in action monitoring, regardless of whether the source of information about an unfavorable outcome is internal (as during self-performed errors; Herrojo Ruiz et al., 2009) or external (i.e., manipulated auditory feedback in the present study). Thirdly, a difference in MMN amplitude between action and perception conditions would be inconsistent with previous studies (see Discussion of Experiment 1; Waszak & Herwig, 2007; Nittono, 2006; Rinne et al., 2001). Fourthly, a recent study (Katahira et al., 2008) reported a negative potential around 200 msec, that did not differ between feedback manipulations introducing an out-of-key tone (i.e., a form of an auditory oddball, which can elicit ERAN-like responses; Brattico, Tervaniemi, Näätänen, & Peretz, 2006) and those that did not introduce an out-of-key tone.

GENERAL DISCUSSION

The present study investigated the neural correlates of processing expectancy violations during the production (action condition) and during the perception of musical sequences (perception condition). Results showed that manipulated tones elicit in both conditions an early negativity, which was more pronounced in the action condition compared to the perception condition, irrespective of whether the manipulations in the perception condition were task-relevant (Experiment 1) or task-irrelevant (Experiment 2). The negativity resembles the feedback ERN/N200, in terms of latency, distribution, and neural generators. The feedback ERN/N200 indexes expectancy-related mechanisms, that is, the detection of a discrepancy between the intended or expected event and the actual event (Ferdinand et al., 2008; Oliveira et al., 2007), and can probably be also elicited in the absence of participants' responses (Tzur & Berger, 2007, 2009; Donkers et al., 2005). Thus, it seems likely that similar expectancy-related mechanisms operated in both the action and the perception conditions. Importantly, results indicate that the feedback ERN/N200 is influenced by the expectancies generated by the intention and action of the pianists to produce a certain auditory effect. In contrast to these action-related expectancies, pianists could build expectations during the perception of the sequences only based on the preceding musical context and its underlying regularities. Consequently, the manipulated tones during piano performance were more unexpected than the manipulated tones during the perception of the sequences, resulting in the enlarged feedback ERN/N200 in the action compared to the perception condition.

An alternative explanation for the increased amplitudes of the feedback ERN/N200 during piano performance is that participants might have recruited more attentional resources than during the perception condition. However, in the perception condition of Experiment 1, participants had to detect the deviant tones (i.e., the tones were task-relevant, as reflected by the P3b), whereas in the action condition, participants were instructed to continue playing after they perceived a feedback manipulation (i.e., the tones were task-irrelevant). Thus, if the feedback ERN/N200 is strongly influenced by attention, it should have been increased in the perception condition, which is inconsistent with the present results. Thus, a simple attention-based account for the amplitude difference seems rather unlikely.

One may criticize that two different tempos were used for the stimulus sequences, possibly influencing the ERP profiles in terms of their latencies. However, it appears that the different tempos of the stimuli have negligible (if any) effects on the latency of the observed ERP components: In a recent study (Katahira et al., 2008), pianists produced melodies with a considerably slower tempo (IOI of around 474 msec) than in the present two experiments [IOI of 216 msec (±10 msec) in Experiment 1 and 322 msec (±77 msec) in Experiment 2], but feedback manipulations in that study (as well as in the present study) elicited negative deflections in the same time range around 200 msec. Note, however, that the study by Katahira et al. (2008) used a different perception condition (including score-reading while listening to the stimuli, and the comparison between action and perception conditions was between-subjects), and no estimation of the neural generators of the negative potentials were reported. In future studies, different manipulations such as the parametric modulation of the frequency of feedback manipulations, the manipulation of the timbre, and the manipulation of the relative musical importance (i.e., different positions in a musical sequence) of feedback alterations would be helpful to learn more about expectancy-related processes and the ERP components involved.

If the feedback ERN/N200 reflects the processing of expectancy violations, how are these expectations during the production and perception of the sequences formed? We assume that during the production of the sequences, pianists anticipated the tone mapped to the particular keypress they were currently performing. After having learned these associations during their extensive training, *performing an action* leads to the prediction of the sensory (auditory) feedback using an internal forward model (Desmurget & Grafton, 2000; Wolpert & Ghahramani, 2000; Wolpert, Ghahramani, & Jordan, 1995; for forward models in the auditory domain, see, e.g., Martikainen, Kaneko, & Hari, 2005). Such a forward model uses an efference copy of the ongoing motor command to compute the sensory consequence of an action. Another possibility is that the expectancies are formed during the *intention to produce a certain effect*, that is, before a motor command is sent. Pianists may have selected their action using an inverse model from the intended effect, also leading to an expectation for a certain effect (the ideomotor principle; see Hommel, Müsseler, Aschersleben, & Prinz, 2001). Importantly, the assumption of these two mechanisms is not mutually exclusive, and it is likely that both mechanisms actually work in parallel. That the expectancy is related to the training of the participants is suggested by the correlation between amplitude and amount of training (pianists with longer training showed larger amplitudes; see Experiment 2). In addition, another study observed a negative potential after manipulated auditory feedback in a musically trained, but not in a nontrained group (Katahira et al., 2008). During the perception of the sequences, we assume that predictive mechanisms extrapolate from the regularities of the preceding auditory input, and thus, generate an expectancy toward a specific tone to follow. This expectancy (or prediction) seems to be a fundamental aspect of perception, which is most likely not under the strategic control of participants (for reviews, see Koelsch, 2009; Schubotz, 2007; Winkler, 2007; Denham & Winkler, 2006).

The data from Experiment 2 also showed an enlarged (fronto-central) P3a component in the action compared to the perception condition. Thus, later processing stages, such as the reorientation of attention (as indexed by the P3a), also seem to be modulated by the expectations built during self-generated actions and during perception. This finding is in accordance with the results of previous studies showing a modulation of deviance processing through effect anticipation (Waszak & Herwig, 2007; Nittono, 2006).

In conclusion, the results of the present study show that the processing of expectancy violations is modulated by the action of an individual. During music performance, pianists expect, on the basis of their intention and their act of performing, to perceive a specific auditory effect. In addition, the preceding musical context induces expectations for specific tones. Hence, when an unexpected tone is encountered following an action, the detection of the violation of these expectancies elicits a brain response similar to the feedback ERN/N200. When pianists only perceive an unexpected tone without performing, the detection of this expectancy violation is only based on the preceding context. This elicits a similar brain response, although with a decreased amplitude. Thus, when a pianist performs "Happy Birthday" for another pianist and produces an unexpected tone (e.g., due to the mistuning of the piano), it is likely that the performer's brain reacts to this event more strongly than the brain of the listener.

Acknowledgments

We thank Sylvia Stasch for help in data acquisition, Sebastian Jentschke and Daniela Sammler for help in data analysis, Kerstin Flake for help with the images, and Nikolaus Steinbeis and Arvid Herwig for helpful comments on earlier versions of this manuscript.

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

Notes

1. Note that the P300 potentials resemble the error positivity (Pe), which can also be decomposed into an early and a late subcomponent. However, whether the P300 and the Pe reflect similar processes is still an open question (for reviews on the Pe, see Overbeek, Nieuwenhuis, & Ridderinkhof, 2005; Falkenstein, Hoormann, Christ, & Hohnsbein, 2000).

2. For this analysis, two participants were excluded due to the small amount of self-performed errors. Furthermore, only the performances of the patterns were analyzed because participants committed an insufficient amount of errors during the playing of the musical scales. ANOVAs were conducted with factor tone (correctly played, incorrectly played) for time windows of −150 to −80 msec, 140 to 240 msec, and 270 to 330 msec over a fronto-central ROI. For further details on the ERPs of self-generated errors, see Maidhof, Prinz, Rieger, & Koelsch (2009).

REFERENCES

Bangert, M., & Altenmüller, E. O. (2003). Mapping perception to action in piano practice: A longitudinal DC-EEG study. *BMC Neuroscience, 4,* 26.

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

Brattico, E., Tervaniemi, M., Näätänen, R., & Peretz, I. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain Research, 1117,* 162–174.

Comerchero, M. D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology, 110,* 24–30.

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

Denham, S. L., & Winkler, I. (2006). The role of predictive models in the formation of auditory streams. *Journal of Physiology (Paris), 100,* 154–170.

Desmurget, M., & Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Sciences, 4,* 423–431.

Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11,* 357–374.

Donkers, F. C. L., Nieuwenhuis, S., & van Boxtel, G. J. M. (2005). Mediofrontal negativities in the absence of responding. *Cognitive Brain Research, 25,* 777–787.

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005a). Action-effect coupling in pianists. *Psychological Research, 69,* 233–241.

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

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47,* 273–305.

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology, 51,* 87–107.

Ferdinand, N., Mecklinger, A., & Kray, J. (2008). Error and deviance processing in implicit and explicit sequence learning. *Journal of Cognitive Neuroscience, 20,* 629–642.

Finney, S. A. (1997). Auditory feedback and musical keyboard performance. *Music Perception, 15,* 153–174.

Finney, S. A. (1999). *Disruptive effects of delayed auditory feedback on motor sequencing*. Providence, RI: Brown University.

Finney, S. A. (2001a). FTAP: A Linux-based program for tapping and music experiments. *Behavior Research Methods, Instruments, & Computers, 33,* 65–72.

Finney, S. A. (2001b). Real-time data collection in Linux: A case study. *Behavior Research Methods, Instruments, & Computers, 33,* 167–173.

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

Folstein, J. R., & van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology, 45,* 152–170.

Goldstein, A., Spencer, K. M., & Donchin, E. (2002). The influence of stimulus deviance and novelty on the P300 and novelty P3. *Psychophysiology, 39,* 781–790.

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology, 42,* 161–170.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology, 44,* 905–912.

Haueisen, J., & Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience, 13,* 786–792.

Herrojo Ruiz, M., Jabusch, H.-C., & Altenmüller, E. (2009). Detecting wrong notes in advance: Neuronal correlates of error monitoring in pianists. *Cerebral Cortex, 19,* 2625–2639.

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

Katahira, K., Abla, D., Masuda, S., & Okanoya, K. (2008). Feedback-based error monitoring processes during musical performance: An ERP study. *Neuroscience Research, 61,* 120–128.

Katayama, J., & Polich, J. (1998). Stimulus context determines P3a and P3b. *Psychophysiology, 35,* 22–23.

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

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

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

Kopp, B., & Wolff, M. (2000). Brain mechanisms of selective learning: Event-related potentials provide evidence for error-driven learning in humans. *Biological Psychology, 51,* 223–246.

Maidhof, C., Prinz, W., Rieger, M., & Koelsch, S. (2009). Nobody is perfect: ERP effects prior to performance errors in musicians indicate fast monitoring processes. *PLoS ONE, 4,* e5032.

Martikainen, M. H., Kaneko, K., & Hari, R. (2005). Suppressed responses to self-triggered sounds in the human auditory cortex. *Cerebral Cortex, 15,* 299–302.

Mecklinger, A., & Ullsperger, P. (1995). The P300 to novel and target events: A spatiotemporal dipole model analysis. *NeuroReport, 7,* 241–245.

Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience, 9,* 788–798.

Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Erlbaum.

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. (2004). Reinforcement-related brain potentials from medial frontal cortex: Origins and functional significance. *Neuroscience and Biobehavioral Reviews, 28,* 441–448.

Nittono, H. (2006). Voluntary stimulus production enhances deviance processing in the brain. *International Journal of Psychophysiology, 59,* 15–21.

Novak, G. P., Ritter, W., Vaughan, H. G., & Wiznitzer, M. L. (1990). Differentiation of negative event-related potentials in an auditory discrimination task. *Electroencephalography and Clinical Neurophysiology, 75,* 255–275.

Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: Expectancy deviation and the representation of action–outcome associations. *Journal of Cognitive Neuroscience, 19,* 1994–2004.

Overbeek, T. J. M., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable components of error processing. On the functional significance of the Pe vis-à-vis the ERN/Ne. *Journal of Psychophysiology, 19,* 319–329.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48,* 115–138.

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details. *Methods and Findings in Experimental and Clinical Pharmacology, 24D,* 5–12.

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

Pfordresher, P. Q. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 949–964.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306,* 443–447.

Rinne, T., Antila, S., & Winkler, I. (2001). Mismatch negativity is unaffected by top–down predictive information. *NeuroReport, 12,* 2209–2213.

Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences, 11,* 211–218.

Simons, R. F., Graham, F. K., Miles, M. A., & Chen, X. (2001). On the relationship of P3a and the novelty-P3. *Biological Psychology, 56,* 207–218.

Sloboda, J. A. (2000). Individual differences in music performance. *Trends in Cognitive Sciences, 4,* 397–403.

Tzur, G., & Berger, A. (2007). When things look wrong: An ERP study of perceived erroneous information. *Neuropsychologia, 45,* 3122–3126.

Tzur, G., & Berger, A. (2009). Fast and slow brain rhythms in rule/expectation violation tasks: Focusing on evaluation processes by excluding motor action. *Behavioural Brain Research, 198,* 420–428.

van Veen, V., & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience, 14,* 593–602.

van Veen, V., Holroyd, C. B., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2004). Errors without conflict: Implications for performance monitoring theories of anterior cingulate cortex. *Brain and Cognition, 56,* 267–276.

Waszak, F., & Herwig, A. (2007). Effect anticipation modulates deviance processing in the brain. *Brain Research, 1183,* 74–82.

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology, 21,* 147–160.

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience, 3(Suppl.),* 1212–1217.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science, 269,* 1880–1882.

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

4.2. Nobody is Perfect: ERP Effects Prior to Performance Errors in Musicians Indicate Fast Monitoring Processes

Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2009). Nobody is perfect: ERP effects prior to performance errors in musicians indicate fast monitoring processes. *PloS ONE*, *4*(4), 1–7. doi:10.1371/journal.pone.0005032

Nobody Is Perfect: ERP Effects Prior to Performance Errors in Musicians Indicate Fast Monitoring Processes

Clemens Maidhof[1]*, **Martina Rieger**[1], **Wolfgang Prinz**[1], **Stefan Koelsch**[1,2]

1 Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, **2** Department of Psychology, University of Sussex, Brighton, United Kingdom

Abstract

Background: One central question in the context of motor control and action monitoring is at what point in time errors can be detected. Previous electrophysiological studies investigating this issue focused on brain potentials elicited *after* erroneous responses, mainly in simple speeded response tasks. In the present study, we investigated brain potentials before the commission of errors in a natural and complex situation.

Methodology/Principal Findings: Expert pianists bimanually played scales and patterns while the electroencephalogram (EEG) was recorded. Event-related potentials (ERPs) were computed for correct and incorrect performances. Results revealed differences already 100 ms prior to the onset of a note (i.e., prior to auditory feedback). We further observed that erroneous keystrokes were delayed in time and pressed more slowly.

Conclusions: Our data reveal neural mechanisms in musicians that are able to detect errors prior to the execution of erroneous movements. The underlying mechanism probably relies on predictive control processes that compare the predicted outcome of an action with the action goal.

Citation: Maidhof C, Rieger M, Prinz W, Koelsch S (2009) Nobody Is Perfect: ERP Effects Prior to Performance Errors in Musicians Indicate Fast Monitoring Processes. PLoS ONE 4(4): e5032. doi:10.1371/journal.pone.0005032

Editor: Naomi Rogers, University of Sydney, Australia

Received July 24, 2008; **Accepted** February 17, 2009; **Published** April 1, 2009

Funding: This study was supported in part by a grant from the German Research Foundation (Deutsche Forschungsgemeinschaft) awarded to S.K. (KO 2266/4-1). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: maidhof@cbs.mpg.de

Introduction

Musical performance is a highly complex and demanding challenge for the human brain [1–3]. For example, a pianist playing a Beethoven sonata has to retrieve from memory which notes have to be played, and in which order this has to be done. Then, the corresponding motor programs have to be activated in order to execute the right movements at the right time with the right intensity. Last but not least, the pianist permanently has to monitor and evaluate the effects of the executed actions for correctness. Importantly, all the processes are constantly overlapping in time. Even though the pianist tries to avoid errors like hitting the wrong key, such errors nevertheless occasionally occur. One question that arises in the context of any kind of motor expertise (in our case piano playing) is at what point in time errors are actually detected by the sensorimotor system. More specifically, in the present study we investigated whether errors are detected before a movement is fully executed.

In the motor control literature, it is assumed that fast movement sequences are controlled without external feedback, because the delays of sensory feedback are too long to have an impact on performance (for a review, see [4]). Accordingly, studies in the music domain showed that auditory feedback is not a prerequisite for a successful performance ([5–7], for a review, see [8]). These studies found that the complete absence of feedback has mostly no effects on piano performance (whereas specific alterations of auditory feedback can profoundly disrupt performance, see [5–7,9]). Hence, it seems possible that monitoring mechanisms in pianists can operate without auditory feedback, i.e. without the perception of an auditory action-effect.

Furthermore, a behavioral study tried to investigate whether motor experts can detect errors before the movement is completed [10]. That study found that incorrect responses of expert typists were less forceful than correct responses. However, it is not clear whether this effect reflects error-specific processing or results from less activation of the incorrect response (see e.g. [11]). In addition, no real-time correlate of electrical brain activity (e.g., EEG) was recorded. Recording EEG is a technique particularly suited to investigate the time course of cognitive processes on a fine-grained time-scale, as for example the time an error is detected.

EEG-studies on error processing (for reviews, see [12–14]) isolated a component of the event-related potential (ERP) appearing shortly *after* participants commit an error in a variety of speeded response tasks (termed the error-related negativity, ERN or Ne [15,16]). The ERN/Ne typically peaks around 50–100 ms after incorrect responses, regardless of the modality in which the stimulus is presented, and regardless of the modality in which the response is made.

Although the ERN/Ne typically appears after the commission of errors, a recent study [17] found increased negativities before participants committed errors in a speech production task. Participants were presented with sequences of word pairs with identical initial phonemes (e.g., “ball doze”, “bash door”, “bean deck”). Every few trials, a word pair was marked for overt articulation. Importantly, in 10% of the sequences the initial phonemes of the last word pair were exchanged (e.g. "darn bore"). When participants are required to vocalize those last word pairs, they are likely to commit errors (e.g. "barn dore"), because two competing speech plans are activated and interfere with each other. This study [17] found an increased negativity after the presentation of the last word pair, and a second negativity after the presentation of the vocalization prompt. However, it remained unclear when exactly participants started to produce speech, and hence the timing of this error response is not evident. Furthermore, participants saw in each trial the stimuli that induced conflict and hence the speech errors. Therefore, the observed ERP effect might have reflected the resolution of conflict in erroneous trials, rather than the detection of an upcoming error. Thus, neural correlates of error detection prior to error execution have remained elusive.

In the present study we investigated expert pianists performing from memory while we recorded the EEG. That is, we investigated highly trained experts committing errors in a complex situation, in which participants did not react to external conflict-inducing stimuli. We compared the brain potentials before and after correct and incorrect keystrokes. More specifically, we hypothesized that differences in the ERP pattern of correct and incorrect keystrokes would occur even before the completion of the movement.

Methods

Participants

Ten highly trained pianists (6 female; mean age 24.3 years, *SD* = 2.8 years) took part in the study. Participants had on average 15.5 years of formal piano training (*SD* = 4.5 years) and were students at the music conservatory in Leipzig. All participants were right-handed according to the Edinburgh Handedness Inventory [18] (mean laterality quotient: 90.5, *SD* = 11.2) and gave informed written consent prior to the experiment. The study was approved by the local ethics committee of the University of Leipzig, and conducted in accordance with the Declaration of Helsinki.

Material and Apparatus

The stimuli consisted of major scales and two similar scale-like patterns in two voices (see Figure 1). In each of 24 experimental blocks, the stimuli had to be produced in different major keys in one of the following two orders: C-Major/E-Major/D-Major/F#-Major, or G-Major/B-Major/A-Major (in case of scales, these sequences were repeated). The order of blocks was randomized with the constraints that no identical stimulus type (scale, pattern A, pattern B) occurred in direct succession and that stimuli with the same order of major keys occurred maximally two times in direct succession.

The instructed tempo for the scales was 144 beats per minute (bpm) and for the patterns 69 bpm, i.e. each note event (= two simultaneous notes) in scales should be produced every 104 ms and in patterns every 217 ms. Randomly between every 40th to 60th produced note, the auditory feedback of a single note was manipulated by lowering the pitch of one note by one semitone. The results of that manipulation will be reported elsewhere.

The pianists performed on a Yamaha digital piano (Clavinova CLP 130), and listened to their performances via AKG 240 studio headphones at comfortable listening levels (approximately 65 dB, dependent on the velocity of a keypress). All tones had the standard MIDI (Musical Instrument Digital Interface) piano timbre generated by a Roland JV-2080 synthesizer (Hamamatsu, Japan).

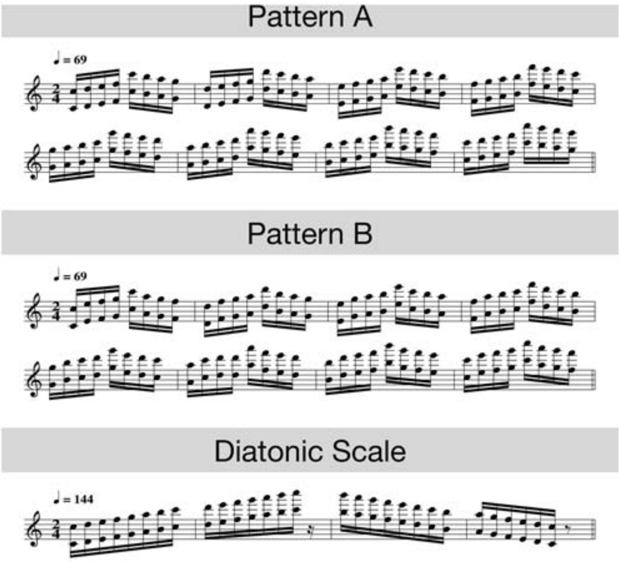

Figure 1. Examples of the stimulus material. A) Pattern A in C-Major. B) Pattern B in C-Major and C) a diatonic scale in C-Major.
doi:10.1371/journal.pone.0005032.g001

Procedure

In the first part of the experiment (ca. 20 min), pianists listened to prerecorded versions of the sequences, which were presented in the same order as the pianists were later required to perform them. Following a practice period with the notation in front of them, participants were blindfolded (to exclude visual feedback and to increase the task difficulty) and instructed to reproduce these stimuli bimanually (parallel in octaves) in the same tempo as they heard them before, i.e. stimuli should be reproduced from memory. If they were not able to perform in the same tempo, they chose their fastest possible tempo. They were informed about the feedback manipulations, and instructed to continue playing, in the event of a feedback manipulation as well as a mistake. When required, participants could rest between two blocks. Before each block, an acoustic instruction was played, informing the participants which scales or patterns they had to produce in the following block. Each performance session lasted approximately 1.5–2 h, and pianists were paid for their participation.

Data Recording and Analysis

Testing was carried out in an acoustically and electrically shielded EEG cabin. Musical data were processed in MIDI format with a modified version of the open source program "FTAP" [19,20]. To synchronize musical and electrophysiological data, this program sent trigger signals concurrently with every 5th keypress (and concurrently with the feedback manipulations) to the EEG acquisition computer. For offline analyses, the MIDI information (including timing information, keypress velocities, and pitch) was saved on a hard disk.

The EEG was recorded with 60 Ag/AgCl scalp electrodes placed according to the extended 10–20 system (FP1, FP2, AF7, AF3, AFZ, AF4, AF8, F9, F7, F5, F3, FZ, F4, F6, F8, F10, FT9, FT7, FC5, FC3, FCZ, FC4, FC6, FT8, FT10, A1, T7, C5, C3, CZ, C4, C6, T8, A2, TP9, TP7, CP5, CP3, CPZ, CP4, CP6, TP8, TP10, P9, P7, P5, P3, PZ, P4, P6, P8, P10, PO7, PO3, POZ, PO4, PO8, O1, OZ, O2), referenced to the electrode at the left mastoid. The ground electrode was placed on the sternum. The horizontal electrooculogram (EOGH) was recorded bipolarly from electrodes placed on the outer left and right canthus and the vertical EOG (EOGV) from electrodes placed below and above the left eye. Impedance was kept below 5 kΩ. EEG signals were digitized with a sampling frequency of 500 Hz.

After data acquisition, EEG data were downsampled to 250 Hz to reduce the data size and re-referenced to the arithmetical mean of both mastoid electrodes. We then performed an independent component analysis (ICA) with standard parameters for artifact removal as implemented in EEGLAB 4.51 (Swartz Center for Computational Neurosciences, La Jolla, CA; http://www.sccn.ucsd.edu/eeglab [21]). After calculating the independent components, artifactual components due to eye movements and blinks were selected based on the following criteria: a component was considered to be artifactual if its topography showed peak activity only over the horizontal or vertical eye electrodes, if it showed a smoothly decreasing power spectrum (which is typical for eye movement artifacts, see [21]), and if the component's activity contributed mainly to the raw EEG signal recorded by the horizontal and vertical eye electrodes. The artifactual components were subtracted from the EEG data, and then the EEG data were filtered with a 0.25–25 Hz bandpass, finite impulse response filter. Subsequently, an automatic rejection procedure was applied: Eye artifacts (which could have still been present after the ICA rejection procedure) were rejected whenever the standard deviation within a 200 ms window centered around each sampling point exceeded 25 µV in the EOG. Artifacts caused by drifts and body movements were eliminated by rejecting sampling points whenever the standard deviation exceeded 25 µV at any electrode either within a 200, or within a 800 ms gliding window.

Performance errors were defined as playing an incorrect key with one hand while pressing the correct key with the other hand. Errors were manually identified off-line. Epochs containing other types of errors like omissions or incorrect keypresses with both hands simultaneously were discarded (on average, there were only 18 trials per participant containing the latter type of error). Only errors that were preceded by a 1 s period of error-free performance (and free of feedback manipulations) were analyzed. Errors were identified separately for the scales and the patterns to take into consideration that the different tempi of both types of stimuli possibly influenced ERP effects. On average, there were only 9 error trials during the performance of the scales, which is insufficient to obtain a reasonable signal-to-noise ratio. Therefore, these data were discarded and we will thus only report the data of the performances of the patterns.

Subsequent to the rejection and filtering procedures, event-related potentials were computed for incorrect ($M = 62$, $SD = 37$) and correct ($M = 682$, $SD = 187$) keypresses for 2000 ms time-locked to the onset of the tones (1000 ms before the onset and 1000 ms after the onset). The baseline was set from 1000 ms to 800 ms before the onset of the tone. For the computation of the signal-to-noise ratio (SNR), we estimated the signal power by determining the highest amplitude in the ERPs between -800 ms and +1000 ms. The noise power was estimated by the standard deviation in the baseline time interval, i.e. between −1000 and −800 ms. The SNR averaged across all participants was 11.1 ($SD = 5.2$).

For statistical analysis, mean ERP amplitude values were calculated for two regions of interest (ROIs) over the midline of the scalp: one anterior with electrodes AFZ, FZ, FCZ, and CZ, and one posterior with electrodes CPZ, PZ, POZ, and OZ. ERPs were statistically analyzed by repeated measures analyses of variance (ANOVAs) with the factors Keypress (correct, incorrect) and AntPos (anterior, posterior). Time windows for statistical analyses of ERP data were chosen based on visual inspection of the grand average and centered around the maximum of the differences between correct and incorrect performed notes. The resulting time windows were −150 to −80 ms (i.e. before the note onset) and 240 to 320 ms (after note onset).

For the behavioral data, we analyzed the MIDI velocities (i.e., the speed at which pianists pressed a key, measured on a scale ranging from 0 to 127; MIDI velocity corresponds to the loudness of the produced tone) of incorrect notes, simultaneous correct notes (played by the other hand), and correct notes when there was no error in either hand. The inter-onset intervals (IOIs) were calculated between the onset of an erroneous note and the onset of the previous note (played by the same hand), between the onset of the simultaneously played correct note and the previous correct note (played by the same hand), and between the onset of successive correct notes (i.e. when there was no error in either hand). Whenever the IOI exceeded 1000 ms, this IOI was discarded. The (signed values of the) asynchronies of keypresses were calculated between errors and the simultaneous correct notes, and between two simultaneous correct notes. All behavioral data were statistically analyzed using repeated measures ANOVAs and paired samples *t*-tests.

Results

Behavioral Results

Pianists pressed incorrect and correct keys with different MIDI velocities. An ANOVA with factor condition (incorrect keypress, simultaneous correct keypress, correct keypress when no error was present) showed a significant main effect of condition ($F(2,18) = 15.18$, $p<.0001$). Contrasts indicated that participants pressed incorrect keys with a lower velocity ($M = 59$, $SD = 8$) than the simultaneous correct keypresses ($M = 63$, $SD = 7$; $p = .003$) and keypresses when there was no error present ($M = 64$, $SD = 7$, $p<.0001$). There was no difference between simultaneous correct keypresses (when an error was present in the other hand) and keypresses when there was no error present ($p = .4$). This pattern of results indicates that the lower velocity of the erroneous keypress did not influence the simultaneous correct keypress of the other hand.

Pianists produced correct and incorrect keypresses with different IOIs. An ANOVA with factor condition (IOIs between incorrect keypress and the previous keypress, IOIs between simultaneous correct keypress and the previous correct keypress, IOIs between two successive correct keypresses) showed a main effect of condition ($F(2,18) = 21.22$, $p = .001$). Contrasts revealed that there was no difference between IOIs between incorrect keypress and the previous keypress by the same hand ($M = 407$ ms, $SD = 106$ ms) and IOIs between the simultaneous correct keypress and the previous correct keypress by the same hand ($M = 404$ ms, $SD = 109$ ms; $p = .24$). However, IOIs between incorrect keypress and the previous keypress were prolonged compared to the IOIs between successive correct keypresses when there was no error present ($M = 367$ ms, $SD = 89$ ms; $p = .001$), indicating that the upcoming error slowed down the keypresses (pre-error slowing). Note that the overall tempo (i.e., the IOIs between correct notes) is slower than initially instructed. This is based on the fact that participants could choose their own (fastest possible) tempo whenever they were not able to perform in the instructed tempo, resulting in a slower mean performance speed.

The asynchronies between two simultaneous correct notes ($M = -2$ ms, $SD = 5$ ms) and between an incorrect and a simultaneous correct note ($M = -4$ ms, $SD = 9$ ms) did not significantly differ from each other ($t(9) = -.71$, $p = .5$).

ERP Results

Figure 2.A shows the grand-averaged waveforms time-locked to the onset of keypresses. Compared to correct keypresses, incorrect

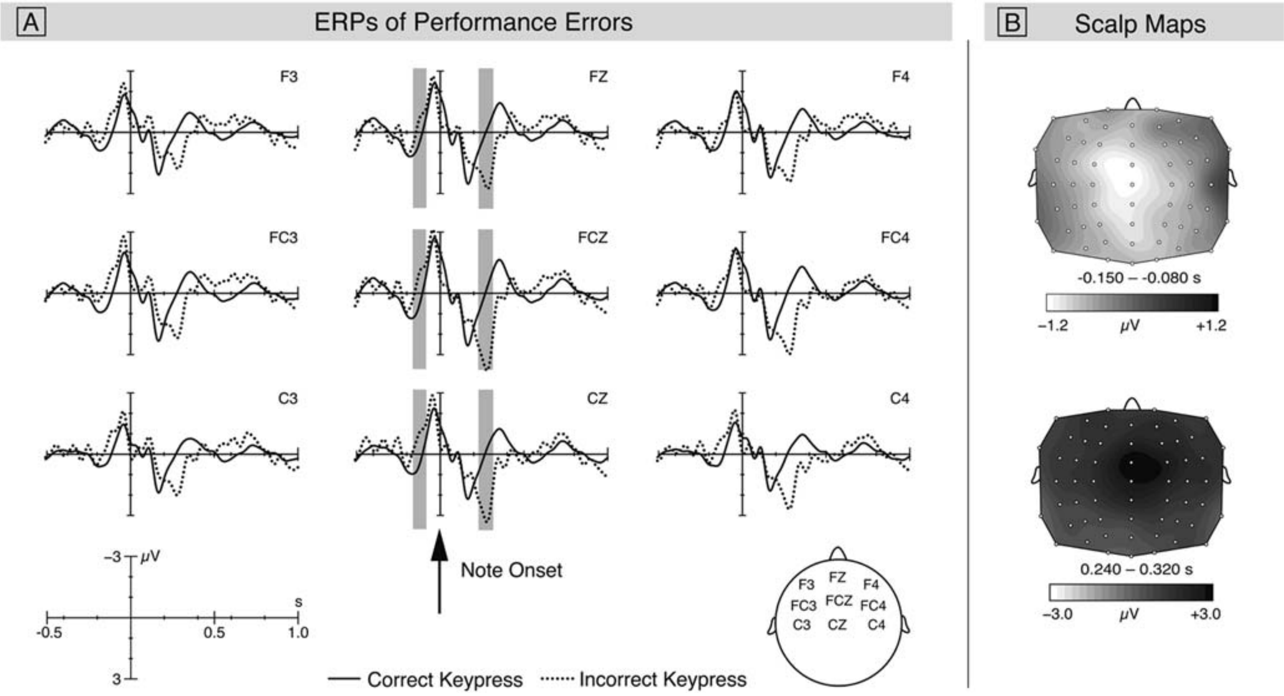

Figure 2. ERP results and scalp distributions of correct and incorrect piano performances. A) Grand-average ERPs elicited by correctly and incorrectly performed keypresses. The arrow indicates the note onset and thus the onset of the auditory feedback. The grey areas show the time windows chosen for statistical analyses for electrodes that were included in the ROIs. Analysis revealed an early increased negative potential prior to the onset of the note (termed pre-error negativity) and a subsequent positive deflection, resembling the early Error positivity (Pe) or the P3a. B) shows the scalp distributions for the difference potentials for correct keypresses subtracted from incorrect keypresses.
doi:10.1371/journal.pone.0005032.g002

keypresses elicited an increased negativity before a wrong key was actually pressed down. The difference was maximal around 100 ms before the onset of the keypresses and showed a central distribution (see Figure 2.B). An ANOVA for a time window ranging from −150 to −80 ms (i.e., before note onset) with factors Keypress and AntPos indicated a significant main effect of Note ($F(1,9) = 8.3$, $p = .018$), but no interaction between Keypress and AntPos ($F<1$). The pre-error negativity was followed by a later positive deflection with an amplitude maximum at around 280 ms after the onset of an incorrect note. This potential showed a fronto-central scalp topography (see Figure 2.A and 2.B). An ANOVA for a time window from 240 ms to 320 ms with factors Keypress and AntPos revealed a main effect of Keypress ($F(1,9) = 9.14$, $p = .014$) and an interaction between factors Keypress and AntPos ($F(1,9) = 6.8$, $p = .028$), indicating that amplitude values were larger over frontal leads than over parietal leads.

Note that IOIs were prolonged before incorrect keypresses and that incorrect keys were pressed with lower velocities. Hence, the ERP difference occurring before the keypress might be due to motor-related processes, such as adjusting the force of the muscles involved in the movement, rather than cognitive processes underlying error monitoring. Such motor-related processes are expected to be lateralized [22,23], whereas cognitive processes of error processing do not show hemispheric differences (for reviews, see [12–14]). To dissociate between a motor and a cognitive explanation, we tested the lateralization of the ERP difference between correct and incorrect keypresses: The ERPs were analyzed separately for left-hand and right-hand errors, with the assumption that motor-related processes of left-hand errors would be reflected in potentials over right-hemispheric motor areas, and vice versa.

Potential maps of ERPs of left-hand errors compared to correct notes (averaged across both hands) are shown in Figure 3.A (difference potential: correct notes subtracted from left-hand errors). The analogous comparison for the right-hand errors is shown in Figure 3.B (correct notes subtracted from right-hand errors). For this analysis, three participants were excluded due to the small number of trials (<10). An ANOVA performed on these difference potentials with factors Hand (left, right), and Hemisphere (left ROI including FC3, FC5, C3, and C5 vs. right ROI including FC4, FC6, C4, and C6) showed no effect of Hand ($F(1,6)<1$, $p = .78$), reflecting that the amplitude of ERP effects did

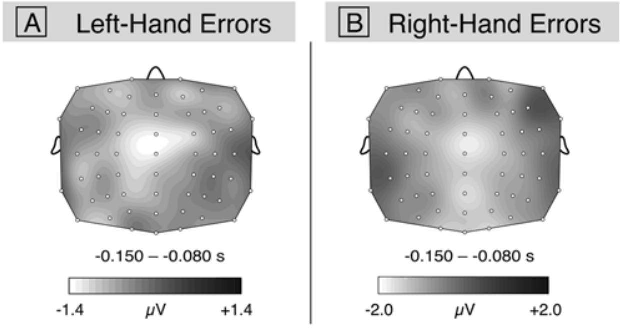

Figure 3. Scalp maps of the difference potentials of left and right-hand errors. A) shows the difference potential for correct keypresses subtracted from left-hand errors and B) the difference potential for correct keypresses subtracted from right-hand errors. Correct keypresses are averaged across both hands.
doi:10.1371/journal.pone.0005032.g003

not differ between left- and right-hand errors, and no interaction between factors Hand and Hemisphere (*F*(1,6)<1, *p* = .88), reflecting that potentials elicited by the errors were not lateralized.

Discussion

Brain potentials elicited by correct and incorrect keypresses of expert pianists differed already 100 ms before keypresses were fully executed, and thus prior to the onsets of erroneous tones (pre-error negativity). The early detection of errors is also observable at the behavioral level: IOIs before erroneous keypresses were prolonged, and erroneous keypresses were executed more slowly. However, the asynchronies between the hands did not increase in erroneous trials. 280 ms after erroneous keypresses a frontocentral positive potential was observed. In the following we will first discuss processes occurring before errors are committed and then turn to the processes occurring after errors are committed.

We assume that the ERPs elicited by incorrect performances reflect neural mechanisms that detect errors before they are actually committed, and before auditory feedback is available. Given the speed of movement sequences in the present study (about 3 keypresses with each hand per second), we suggest that internal forward models predicting the sensory consequences of actions [24–28] are the basis for detecting the errors even before they were fully executed: Monitoring of fast movements, whose control cannot wait for sensory feedback, has to rely mainly on predictive (feedforward) mechanisms that compare internal action goals with the predicted consequences of planned movements.

Studies investigating the activity of neurons in the primary motor cortex (M1) of non-human primates showed that the latency between the first activity in M1 and movement onset is variable and can range up to several hundred milliseconds [29–32], but the typical assumed latency is around 100 to 150 ms (e.g. [33]). At the same time as the motor command is sent from M1 to the periphery, an efference copy (or corollary discharge) is created in brain structures also involved in the generation of the movement. The efference copy is, however, not used to generate the ongoing motor activity, but can be used to predict the outcome of the motor command [24–28] (information of efference copies interact at several levels of the central nervous system, and often modulate sensory processing; for reviews, see [34,35]). The predicted outcome can be compared to the intended outcome, and an error signal is generated whenever there is a mismatch between intended goal and predicted consequence. The error signal can, in turn, modulate the motor command [27].

Accordingly, we assume that the mismatch between the predicted consequence of a planned keypress and the associated internal action goal, as detected by a feedforward control mechanism, is reflected in the pre-error negativity. From the present data we cannot conclude during which part of the movement (planning, initiation, early stages of execution) this feedforward control mechanism exactly occurs. However, it is important to note that a detection mechanism seems to operate before the pianists receive auditory feedback of their errors, i.e. before pianists perceive the auditory results of their actions.

The modulation of the motor command by the error signal of the feedforward mechanism might have resulted in the prolonged IOIs before and the slower velocities of incorrect keypresses, probably reflecting an attempt to avoid the error. In contrast to what one might have assumed, IOIs were not only prolonged for the hand that pressed the incorrect key, but IOIs were also prolonged for the other hand that pressed simultaneously the correct key. This is presumably due to bimanual coupling: studies show that bimanual movements begin and end at similar times, even when they have different parameters (e.g. amplitudes) and movement times differ when the respective movements are performed in isolation by one hand [36–38]. Our task required tight bimanual coupling of the hands in terms of the timing. Correspondingly, asynchronies between the hands did not differ when an error was present or not.

One could argue that the pre-error negativity might reflect an error during memory retrieval and, thus, an even earlier stage than motor control or error monitoring. It is assumed that serial-ordering errors (i.e. notes that are intended at another location in the sequence) reflect the current activation of this erroneous element in memory [39,40]. However, because pianists in our study performed the same tones in parallel with both hands (one octave apart), errors reflecting false memory retrieval should occur in both hands, instead of only in one. Because we only analyzed errors committed by one hand, it is unlikely that the pre-error negativity reflects false retrieval from memory. Moreover, one could also argue that the ERP difference before the note onsets might be due to motor-related processes. Motor execution processes are, however, expected to elicit lateralized EEG potentials [22,23], which is not consistent with our data: The separate analysis of left-hand and right-hand errors did not reveal any lateralization effect. Therefore, it is unlikely that the ERP difference reflects simply motor-related processes, but rather processes operating at a higher cognitive level, associated with monitoring or control. Finally, one could reason that the increased negativity before incorrectly played notes reflects a process that actually *results* in the production of an error. For instance, a recent study [41] showed that lapses in preparatory attention networks can lead to production errors. In that study the amplitude of the Contingent Negative Variation (CNV), a brain potential indexing preparatory attention, was decreased before stimulus presentation when an erroneous response occurred. Therefore, if lapses in preparatory attention were responsible for the errors in our study, one would have expected a similar decrease in ERP amplitude. However, ERPs elicited before incorrect performances had larger (negative) amplitude values than those elicited before correct performances, rendering such an explanation unlikely. Further, we think that the observed ERP difference in our study occurred too late to reflect lapses in attention. Considering the delay of activity in M1 to movement onset (presumably around 100 to 150 ms), lapses of attention should be observable before that time (as it was reported in [41]), i.e. several hundred milliseconds before the button press. Thus, the fact that an increased negativity (instead of a decreased negative amplitude) was observed, in combination with the observed timing of the effect (around 100 ms before movement completion) renders it improbable that lapses in preparatory attention can account for the present findings. A similar explanation for the present results might be a temporal disengagement of the action monitoring system. Two other studies [42,43] found that trials preceding erroneous trials (in Eriksen flanker and Stroop tasks) showed an enhanced positivity (compared to trials preceding correct trials), thereby 'foreshadowing' errors in future trials. This effect (termed the Error-preceding Positivity, EPP) is thought to reflect "transient deficiencies in the functioning of the monitor system prior to actual execution of an error" [42]. These deficits may be associated with failures to activate adaptive control processes, resulting in occasional future errors. Because we observed no enhanced positivity before production errors, it is unlikely that a disengagement of the action monitoring system is reflected in the observed ERP effect.

The expertise of our participants and the characteristics of our task might explain why we did not observe an ERN (a potential frequently observed following the commission of errors, see [12–

14] for reviews) or an EPP component: In contrast to most studies investigating error processing (mostly in simple speeded response tasks, including the aforementioned studies [41–43]) our participants did not react to external stimuli according to pre-defined arbitrary rules. Instead, they had to select the appropriate motor commands according to internal goals that they formed on the basis of instructions and the musical knowledge stored in their long-term memories. In addition, the present experimental situation reflects a task for which musicians are highly trained, compared to the button press responses to stimuli presented in standard error processing paradigms. Consequently, the error could be detected earlier than in choice reaction tasks. Incorrect notes also violated the regularity of the sequences and thus represented auditory oddballs, which are known to elicit a mismatch negativity (MMN; for a review, see [44]). However, no MMN was visible in the ERPs, perhaps because it was overlapped by the positive potential emerging in a similar latency range (see below). Note that the magnitude of the ERPs (around 3 μV) was rather small compared to the amplitude of ERPs elicited in standard error processing paradigms [12–14]. This is probably due to the complexity of our task, involving a range of interacting cognitive processes (e.g., memory retrieval, motor planning, performance monitoring etc., see Introduction). In addition, the simultaneous processing of input from different sensory systems (auditory, tactile, somatosensory) might have influenced the magnitude of the ERPs.

The fronto-central positive potential (emerging around 200 ms and) peaking around 280 ms after incorrect keypresses strongly resembles the Error Positivity (Pe), a potential frequently observed following the ERN in studies of error processing (for reviews, see [45,46]). Although the functional significance of the Pe has remained rather unclear, three hypotheses about the Pe have emerged: The affective-processing hypothesis [45,47] suggests that the Pe reflects affective processing of the committed error or its consequences. According to the behavior-adaption hypothesis [48], the Pe reflects the adaptation of response strategy after an error has been perceived, involving remedial performance adjustments following errors. The error-awareness hypothesis [49,50] proposes that the Pe reflects the conscious recognition of a committed error. There is only little evidence in favor of the first two hypotheses, whereas there are some empirical data supporting the error-awareness hypothesis (e.g. [46,50,51]). Another way of addressing the question about the functional significance of this potential is to consider its similarities to the P300 component, which has led to the suggestion that the Pe could reflect a P3b associated with the motivational significance of an error (for a review on the P300, see [52]). The Pe, however, can be decomposed into an early and a late component, very similar to the distinction between P3a (indexing the involuntarily attention switch to novel and deviant stimuli, e.g. [53,54]) and P3b (taken to reflect memory updating operations after task-relevant stimuli, e.g. [55,56], but see also [57]). However, there are no studies directly comparing the early Pe with the P3a and the late Pe with the P3b, and therefore it remains unclear whether the early Pe reflects similar processes as the P3a. Based on previous studies [50,51,55,58] we suggest that the positive deflection observed in the present study most likely reflects an early Pe or a P3a. Whether this potential is related to later processing stages of tactile and/or auditory feedback of the error, or simply due to the processing of an oddball stimulus (leading to an involuntary reallocation of attention) remains to be clarified. One way to address this would be to investigate performance errors committed in the absence of auditory feedback: if these errors also elicit the positivity, this potential cannot reflect auditory novelty processing.

In conclusion, the method of investigating motor experts in a natural context, accompanied with on-line measures of electrical brain activity (like EEG), can help to answer crucial questions in the domain of motor control and action monitoring. The occurrence of a pre-error negativity indicates that an early error detection mechanism operates in pianists even before an erroneous movement is fully executed. Our data also show that the early detection of errors influences movement execution, resulting in pre-error slowing of both hands and in keypresses with reduced velocity of the erroneous hand only. We assume that the underlying process is the detection of a mismatch between a predicted sensory consequence of an action and the intended action goal. Thus, our results reveal neural mechanisms that are able to detect errors prior to the execution of erroneous movements.

Acknowledgments

The authors would like to thank Sylvia Stasch for help in data acquisition, and Sebastian Jentschke and Daniela Sammler for help in data analysis.

Author Contributions

Conceived and designed the experiments: CM MR SK. Performed the experiments: CM. Analyzed the data: CM. Wrote the paper: CM MR WP SK.

References

1. Münte TF, Altenmüller E, Jäncke L (2002) The musician's brain as a model of neuroplasticity. Nat Rev Neurosci 3: 473–478.
2. Palmer C (2005) Sequence Memory in Music Performance. Curr Dir Psychol Sci 14: 247–250.
3. 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.
4. Kawato M (1999) Internal models for motor control and trajectory planning. Curr Opin Neurobiol 9: 718–727.
5. Finney SA (1997) Auditory Feedback and Musical Keyboard Performance. Music Percept 15: 153–174.
6. Finney SA, Palmer C (2003) Auditory feedback and memory for music performance: Sound evidence for an encoding effect. Mem Cognit 31: 51–64.
7. 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.
8. Pfordresher PQ (2006) Coordination of perception and action in music performance. Advances in Cognitive Psychology 2: 183–198.
9. Pfordresher PQ (2003) Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. J Exp Psychol Hum Percept Perform 29: 949–964.
10. Rabbitt PM (1978) Detection of errors by skilled typists. Ergonomics 21: 945–958.
11. Gehring WJ, Coles MG, Meyer DE, Donchin E (1995) A brain potential manifestation of error-related processing. In: Karmos G, Molnár M, Csépe V, Czigler I, Desmedt JE, eds. Perspectives of event-related potentials research. Amsterdam: Elsevier. pp 267–272.
12. Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. Psychol Rev 108: 624–652.
13. Yeung N, Cohen JD, Botvinick MM (2004) The neural basis of error detection: conflict monitoring and the error-related negativity. Psychol Rev 111: 931–959.
14. van Veen V, Carter CS (2006) Error detection, correction, and prevention in the brain: a brief review of data and theories. Clin EEG Neurosci 37: 330–335.
15. Falkenstein M, Hohnsbein J, Hoormann J, Blanke L (1990) Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In: Brunia CHM, Gaillard AWK, Kok A, eds. Psychophysiological Brain Research. Tilburg, The Netherlands: Tilburg University Press. pp 192–195.
16. Gehring WJ, Goss B, Coles MG, Meyer DE, Donchin E (1993) A Neural System For Error Detection And Compensation. Psychol Sci 4: 385–390.
17. Möller J, Jansma BM, Rodriguez-Fornells A, Münte TF (2007) What the Brain Does before the Tongue Slips. Cereb Cortex 17: 1173–1178.
18. Oldfield RC (1971) The assessment and analysis of handedness: The Edinburgh Inventory. Neuropsychologia 9: 97–113.

19. Finney SA (2001a) FTAP: a Linux-based program for tapping and music experiments. Behav Res Methods Instrum Comput 33: 65–72.
20. Finney SA (2001b) Real-time data collection in Linux: a case study. Behav Res Methods Instrum Comput 33: 167–173.
21. 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.
22. Shibasaki H, Barrett G, Halliday E, Halliday AM (1980) Components of the movement-related cortical potential and their scalp topography. Electroencephalogr Clin Neurophysiol 49: 213–226.
23. Colebatch JG (2007) Bereitschaftspotential and movement-related potentials: origin, significance, and application in disorders of human movement. Mov Disord 22: 601–610.
24. Wolpert DM, Ghahramani Z, Jordan MI (1995) An internal model for sensorimotor integration. Science 269: 1880–1882.
25. Miall RC, Wolpert DM (1996) Forward models for physiological motor control. Neural Netw 9: 1265–1279.
26. Wolpert DM, Miall RC, Kawato M (1998) Internal models in the cerebellum. Trends Cogn Sci 2: 338–347.
27. Desmurget M, Grafton S (2000) Forward modeling allows feedback control for fast reaching movements. Trends Cogn Sci 4: 423–431.
28. Wolpert DM, Ghahramani Z (2000) Computational principles of movement neuroscience. Nat Neurosci 3 Suppl: 1212–1217.
29. Evarts EV (1974) Precentral and postcentral cortical activity in association with visually triggered movement. J Neurophysiol 37: 373–381.
30. Porter R, Lewis MM (1975) Relationship of neuronal discharges in the precentral gyrus of monkeys to the performance of arm movements. Brain Res 98: 21–36.
31. Thach WT (1978) Correlation of neuronal discharge with pattern and force of muscular activity, joint position, and direction of intended next movement in motor cortex and cerebellum. J Neurophysiol 41: 654–676.
32. Holdefer RN, Miller LE (2002) Primary motor cortical neurons encode functional muscle synergies. Exp Brain Res 146: 233–243.
33. Hatsopoulos NG, Xu Q, Amit Y (2007) Encoding of movement fragments in the motor cortex. J Neurosci 27: 5105–5114.
34. Poulet JFA, Hedwig B (2006) New insights into corollary discharges mediated by identified neural pathways. Trends Neurosci 30: 14–21.
35. Crapse TB, Sommer MA (2008) Corollary discharge circuits in the primate brain. Curr Opin Neurobiol 18: 1–6.
36. Marteniuk RG, MacKenzie CL, Baba DM (1984) Bimanual Movement Control: Information Processing and Interaction Effects. Q J Exp Psychol 36A: 335–365.
37. Spijkers W, Heuer H, Kleinsorge T, van der Loo H (1997) Preparation of bimanual movements with same and different amplitudes: specification interference as revealed by reaction time. Acta Psychol 96: 207–227.
38. Swinnen SP, Wenderoth N (2004) Two hands, one brain: cognitive neuroscience of bimanual skill. Trends Cogn Sci 8: 18–25.
39. Palmer C, van de Sande C (1993) Units of knowledge in music performance. J Exp Psychol Learn Mem Cogn 19: 457–470.
40. Palmer C, Pfordresher PQ (2003) Incremental planning in sequence production. Psychol Rev 110: 683–712.
41. Padilla ML, Wood RA, Hale LA, Knight RT (2006) Lapses in a Prefrontal-Extrastriate Preparatory Attention Network Predict Mistakes. J Cogn Neurosci 18: 1477–1487.
42. Ridderinkhof KR, Nieuwenhuis S, Bashore T (2003) Errors are foreshadowed in brain potentials associated with action monitoring in cingulate cortex in humans. Neurosci Lett 348: 1–4.
43. Hajcak G, Nieuwenhuis S, Ridderinkhof KR, Simons RF (2005) Error-preceding brain activity: Robustness, temporal dynamics, and boundary conditions. Biol Psychol 70: 67–78.
44. Näätänen R, Paavilainen P, Rinne T, Alho K (2007) The mismatch negativity (MMN) in basic research of central auditory processing: a review. Clin Neurophysiol 118: 2544–2590.
45. Falkenstein M, Hoormann J, Christ S, Hohnsbein J (2000) ERP components on reaction errors and their functional significance: a tutorial. Biol Psychol 51: 87–107.
46. Overbeek TJM, Nieuwenhuis S, Ridderinkhof KR (2005) Dissociable Components of Error Processing. On the Functional Significance of the Pe Vis-a-vis the ERN/Ne. J Psychophysiology 19: 319–329.
47. van Boxtel GJM, van der Molen MW, Jennings JR (2005) Differential Involvement of the Anterior Cingulate Cortex in Performance Monitoring During a Stop-Signal Task. J Neurophysiol 19: 1–10.
48. Hajcak G, McDonald N, Simons RF (2003) To err is autonomic: Error-related brain potentials, ANS activity, and post-error compensatory behavior. Psychophysiology 40: 895–903.
49. Kaiser J, Barker R, Haenschel C, Baldeweg T, Gruzelier JH (1997) Hypnosis and event-related potential correlates of error processing in a stroop-type paradigm: a test of the frontal hypothesis. Int J Psychophysiol 27: 215–222.
50. Nieuwenhuis S, Ridderinkhof KR, Blom J, Band GP, Kok A (2001) Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. Psychophysiology 38: 752–760.
51. Endrass T, Reuter B, Kathmann N (2007) ERP correlates of conscious error recognition: aware and unaware errors in an antisaccade task. Eur J Neurosci 26: 1714–1720.
52. Polich J (2007) Updating P300: An integrative theory of P3a and P3b. Clin Neurophysiol 118: 2128–2148.
53. Escera C, Alho K, Winkler I, Näätänen R (1998) Neural Mechanisms of Involuntary Attention to Acoustic Novelty and Change. J Cogn Neurosci 10: 590–604.
54. Escera C, Corral MJ (2007) Role of Mismatch Negativity and novelty-P3 in involuntary auditory attention. J Psychophysiol 21: 251–264.
55. Courchesne E, Hillyard SA, Galambos R (1975) Stimulus novelty, task relevance and the visual evoked potential in man. Electroencephalogr Clin Neurophysiol 39: 131–143.
56. Knight R (1996) Contribution of human hippocampal region to novelty detection. Nature 383: 256–259.
57. Verleger R (2008) P3b: Towards some decision about memory. Clin Neurophysiol 119: 968–970.
58. van Veen V, Carter CS (2002) The timing of action-monitoring processes in the anterior cingulate cortex. J Cogn Neurosci 14: 593–602.

4.3. Predictive Error Detection in Pianists: A Combined ERP and Motion Capture Study

Maidhof, C., Pitkäniemi, A., & Tervaniemi, M. (2013). Predictive error detection in pianists: a combined ERP and motion capture study. *Frontiers in Human Neuroscience*, *7*, 1–14. doi:10.3389/fnhum.2013.00587

frontiers in
HUMAN NEUROSCIENCE

ORIGINAL RESEARCH ARTICLE
published: 26 September 2013
doi: 10.3389/fnhum.2013.00587

Predictive error detection in pianists: a combined ERP and motion capture study

***Clemens Maidhof*[1,2]*, *Anni Pitkäniemi*[1,2,3] *and Mari Tervaniemi*[1,2,3]

[1] Cognitive Brain Research Unit, Cognitive Science, Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland
[2] Department of Music, Finnish Centre of Excellence in Interdisciplinary Music Research, University of Jyväskylä, Jyväskylä, Finland
[3] Department of Psychology, University of Jyväskylä, Jyväskylä, Finland

Edited by:
Srikantan S. Nagarajan, University of California, San Francisco, USA

Reviewed by:
Caroline A. Niziolek, University of California, San Francisco, USA
Roozbeh Behroozmand, University of Iowa, USA

****Correspondence:***
Clemens Maidhof, Cognitive Brain Research Unit, Cognitive Science, Institute of Behavioural Sciences, University of Helsinki, PO Box 9 (Siltavuorenpenger 1 B), 00014 Helsinki, Finland
e-mail: clemens.maidhof@gmail.com

Performing a piece of music involves the interplay of several cognitive and motor processes and requires extensive training to achieve a high skill level. However, even professional musicians commit errors occasionally. Previous event-related potential (ERP) studies have investigated the neurophysiological correlates of pitch errors during piano performance, and reported pre-error negativity already occurring approximately 70–100 ms before the error had been committed and audible. It was assumed that this pre-error negativity reflects predictive control processes that compare predicted consequences with actual consequences of one's own actions. However, in previous investigations, correct and incorrect pitch events were confounded by their different tempi. In addition, no data about the underlying movements were available. In the present study, we exploratively recorded the ERPs and 3D movement data of pianists' fingers simultaneously while they performed fingering exercises from memory. Results showed a pre-error negativity for incorrect keystrokes when both correct and incorrect keystrokes were performed with comparable tempi. Interestingly, even correct notes immediately preceding erroneous keystrokes elicited a very similar negativity. In addition, we explored the possibility of computing ERPs time-locked to a kinematic landmark in the finger motion trajectories defined by when a finger makes initial contact with the key surface, that is, at the onset of tactile feedback. Results suggest that incorrect notes elicited a small difference after the onset of tactile feedback, whereas correct notes preceding incorrect ones elicited negativity before the onset of tactile feedback. The results tentatively suggest that tactile feedback plays an important role in error-monitoring during piano performance, because the comparison between predicted and actual sensory (tactile) feedback may provide the information necessary for the detection of an upcoming error.

Keywords: EEG, performance monitoring, music performance, motor control, musical expertise

INTRODUCTION

Performing a piece of music is a highly demanding task, involving several cognitive and motor processes (for reviews, see e.g., Palmer, 1997; Münte et al., 2002; Zatorre et al., 2007). Although professional musicians spend thousands of hours of deliberate practice to master their instrument (Ericsson et al., 1993; Ericsson and Lehmann, 1996; Sloboda et al., 1996), they make errors occasionally. To detect errors, which represent deviations from intended goals and actions and their effects, humans have to constantly monitor their ongoing behavior and its outcomes. The present study aimed at investigating the neurophysiological correlates of error-related processes during music performance. To relate neurophysiological findings to different movement stages, we used a new exploratory paradigm in which EEG and 3D movement data with a motion capture system were concurrently recorded.

Most previous neuroscientific research has focused on errors committed during various choice-reaction time tasks (for reviews, see Nieuwenhuis et al., 2004; Yeung et al., 2004; van Veen and Carter, 2006; Taylor et al., 2007). A prominent finding was a sharp negative deflection in the event-related potential (ERP) peaking around 50–100 ms after an incorrect response, termed the error-related negativity or error negativity (ERN and Ne, respectively; Falkenstein et al., 1990; Gehring et al., 1993). The ERN can be elicited independently of the modality in which the stimulus is presented (Falkenstein et al., 2000), and independently of the effector (hand or foot) with which the incorrect response is made (Holroyd et al., 1998). Evidence from EEG source localization studies, functional neuroimaging studies, as well as from single-unit recordings from primates indicate that the ERN receives major contributions from the dorsal part of the anterior cingulate cortex (dACC; for a review, see Ridderinkhof et al., 2004).

Several theories have been put forward with regard to the functional role of the ERN. According to the comparator theory, the ERN reflects the outcome of a process that compares the neural representation of the actual response with the correct response (Falkenstein et al., 2000). By contrast, the conflict monitoring theory (Carter, 1998; Botvinick et al., 2001; van Veen et al., 2001) assumes that the dACC monitors for cognitive conflict occurring when two competing response representations are activated (as for example with the Stroop effect). Within this framework, errors are a special case of high conflict, and the ERN is elicited when the representation of an incorrect response crosses a threshold so that an actual response is being made. The reinforcement learning theory of the ERN (Holroyd and Coles, 2002; Nieuwenhuis et al., 2004), which can be viewed as an extension of the comparator theory, posits that the ERN is elicited when the outcome of an event is worse than expected. In that view, the basal ganglia monitor ongoing events and predict whether they will be better or worse than expected. If an event is predicted to be worse than expected, the basal ganglia signal this with a phasic decrease in dopaminergic activity in the ACC, which gives rise to the ERN. However, there is an ongoing debate as to the degree to which each theory can account for the existing findings.

Another important finding in the domain of performance monitoring was the feedback ERN (for a review, see Nieuwenhuis et al., 2004). This component is elicited around 250 ms after negative performance feedback indicating loss or punishment in time-estimation, guessing, and gambling tasks (e.g., Miltner et al., 1997; Hajcak et al., 2005, 2007), and is presumably generated in the ACC (Nieuwenhuis et al., 2004). However, there is also evidence that the feedback ERN is elicited not only by negative feedback, but also by unexpected positive feedback (Oliveira et al., 2007; Ferdinand et al., 2012), indicating that the medial frontal cortex is sensitive to violations of expectancy in general, regardless of the specific valence of an event.

The processing of errors and unexpected feedback has also been investigated during speech production and musical performance. In the speech domain, vocal errors committed during, for example, the Stroop color-word task (Masaki et al., 2001), a picture naming task (Riès et al., 2011), or during the monitoring of internal speech in Go/No-go tasks (Ganushchak and Schiller, 2006, 2008) elicited a negative potential shortly after the onset of an incorrect response that highly resembled the ERN observed in non-linguistic tasks. This indicated that the ERN reflects more domain-general response monitoring functions.

In addition to the monitoring of internal speech, auditory feedback provides an important source of information about ongoing speech acts, and can be used to control vocal fundamental frequency. Studies investigating the processing of manipulated auditory feedback (i.e., pitch- or time-shifted feedback) during vocalizations in humans and non-human primates reported that the motor-induced suppression of auditory cortical responses (i.e., the inhibitory effects within the auditory cortex during vocalization, as compared to listening) is decreased during feedback perturbations (e.g., Houde et al., 2002; Heinks-Maldonado et al., 2006; Eliades and Wang, 2008; Behroozmand and Larson, 2011; Behroozmand et al., 2011). It has been argued that the underlying mechanisms are based on internal forward models in the auditory domain, and that the dampening of sensory input can help to differentiate self-produced from externally-generated sounds. In that view, an internal forward model receives information about the ongoing motor command in the form of an efference copy. The forward model can predict the sensory consequences of an action by integrating information about the current state of the system and this efference copy (Wolpert et al., 1995; Desmurget and Grafton, 2000). If the comparison between actual (in the form of reafferent sensory and tactile/proprioceptive information) and predicted feedback yields a match, the resulting small prediction error leads to minimal responses in the auditory cortex. In the event of a mismatch between the predicted and actual consequences, an error signal is generated that can be used to cancel the inhibitory effects within the auditory cortex. Furthermore, it is assumed that the error signal can be used for rapid adjustments of ongoing motor activity.

In the music domain, the processing of occasionally manipulated auditory feedback (compared to normal feedback) during piano performances elicited an increased N100 response and a negative potential around 200 ms that was interpreted as a feedback ERN (Katahira et al., 2008; Maidhof et al., 2010). The feedback ERN elicited by unexpected pitch shifts of the feedback showed larger amplitudes during performance than during listening (Maidhof et al., 2010), resembling the results in the speech domain. This indicates that manipulated notes were more unexpected when they were self-generated, which has been explained—in a similar manner to the findings in the speech domain—in terms of internal forward models (Katahira et al., 2008; Maidhof et al., 2010).

Furthermore, several recent studies investigated rare pitch errors (i.e., playing an incorrect note on the keyboard) in highly-skilled pianists performing pieces of music and fingering exercises from memory. Results showed that pitch errors, compared to correct notes, elicited a negative component in the ERP that already peaked approximately 70–100 ms prior to the onset of errors, and thus prior to the auditory feedback of the wrong note (termed "pre-error negativity" or "preERN," Herrojo Ruiz et al., 2009; Maidhof et al., 2009; Strübing et al., 2012). The negative component showed a frontocentral scalp topography (Herrojo Ruiz et al., 2009), regardless of whether errors were committed with the right or left hand (Maidhof et al., 2009). Furthermore, it was elicited even in the complete absence of auditory feedback (Herrojo Ruiz et al., 2009), can be altered in pianists with focal dystonia (Strübing et al., 2012), and is presumed to be generated in the ACC (Herrojo Ruiz et al., 2009).

On the behavioral level, these studies reported that erroneous keys were pressed with a lower velocity than correct keys, which resulted in decreased intensities of wrong notes (Herrojo Ruiz et al., 2009; Maidhof et al., 2009; Herrojo Ruiz et al., 2011; Strübing et al., 2012). In addition, errors and the following notes were executed slower, that is, the inter-onset intervals (IOIs) from (a) the preceding note to the incorrect note (calculated as n minus $n-1$); and (b) from the incorrect note to the subsequent correct note (calculated as $n+1$ minus n) were prolonged. Interestingly, a recent study also showed that correct notes immediately preceding wrong keystrokes ("pre-error notes") were pressed with decreased velocities, although not to the same degree as errors (Palmer et al., 2012). This latter observation is consistent with the notion that errors can influence surrounding events in a sequence, such that pre-error notes "inherit" some features of the following error (decreased intensity), but are still correct with regards to the pitch property.

The ERP effect prior to pitch errors during piano performance (Herrojo Ruiz et al., 2009, 2011; Maidhof et al., 2009;

Strübing et al., 2012) has been hypothesized to reflect error detection processes, which do not rely on auditory feedback and could be based on internal forward mechanisms. However, the precise movement stages during which the pre-error negativity occurs have remained unclear, as has the role played by tactile feedback during musical performance for error-related processes.

The aim of the present study is to address the following issues:

(1) In previous studies, the direct comparison of correct and incorrect notes was confounded by their different IOIs. Because of the possibility that ERPs of wrong notes (longer IOIs before the incorrect note) and the previous correct notes (shorter IOIs) overlap, the different IOIs could possibly have influenced the ERP effect prior to the erroneous notes. In the present study, we therefore compared incorrect and correct keystrokes that were executed at comparable tempi. If the ERP effect before errors reflects mainly error-related processes and not tempo differences, it should be elicited even when correct and incorrect notes show a similar tempo. IOIs were always calculated as *n* minus *n*-1, thus IOIs refer to the pre-note intervals.

(2) Previous studies related the electrophysiological data to the time-point when the key was almost fully pressed (i.e., the point at which the MIDI [Musical Instrument Digital Interface] signal is generated by a digital piano upon depression of a key). In the present study, we exploratively recorded 3D movement data of participants' fingers simultaneously with the EEG to investigate the underlying movements and the role of tactile feedback for error monitoring. Two previous studies (Goebl and Palmer, 2008; Palmer et al., 2009) showed that kinematic landmarks like acceleration peaks in the finger trajectories provide a measure of the available tactile information. Specifically, a finger-key landmark (FK landmark) can occur when a finger arrives at the piano key surface and changes its acceleration abruptly (before the key is actually pressed down), reflecting the onset of tactile feedback (see also Goebl and Palmer, 2009, 2013). In the present study, we analyzed the finger acceleration profiles of correct and incorrect keystrokes, and additionally computed the ERPs time-locked to the onsets of finger-key landmarks. We hypothesized that if a difference between the ERPs of correct and incorrect keystrokes occurs prior to FK landmarks, tactile information does not contribute to the pre-error negativity. In contrast, an ERP difference after FK landmarks would indicate that tactile feedback might play an important role for error-related processes during music performance.

(3) Based on recent behavioral evidence showing that correct pre-error notes inherit some incorrect properties similar to errors (Palmer et al., 2012), we also investigated the ERPs to pre-error notes and compared them to other correct notes.

MATERIALS AND METHODS

PARTICIPANTS

Nineteen pianists participated in the study. Participants were current or former students at Finnish universities with professional music programs or music conservatories. Based on listening to the performances, one participant was excluded because the performance showed that the stimuli could not be produced fluently and included too many interruptions. Six participants were excluded because preliminary analyses indicated that their performances were too slow (mean IOI clearly above 200 ms). One participant was excluded because the markers were not correctly recorded by the motion capture system, and one participant was excluded because of excessive alpha activity. Thus, the data of 10 pianists (6 females; mean age: 23.7 years, *SD* = 5.5 years) were analyzed. They had, on average, 14.2 years of formal musical training (*SD* = 7.3 years), and had begun playing the piano between 4 and 10 years of age. On average, participants spent 2.7 h (*SD* = 1.2 h) daily on piano practice. Handedness was assessed with a revised version of the Edinburgh Handedness Inventory, in which three original activities were discarded (opening box, broom, drawing), and one new one was added (computer mouse; see http://homepage.ntlworld.com/steve.williams7/A major revision of the Edinburgh Handedness Inventory.pdf). Results showed that nine participants were right-handed (mean laterality quotient: 87, *SD* = 14.2), and one participant was mixed-handed (laterality quotient: −25). Participants reported having normal hearing and no neurological disorders, and gave informed written consent prior to the experiment. The study was approved by the local ethical committee of the Faculty of Behavioral Sciences at the University of Helsinki, and conducted in accordance with the Declaration of Helsinki.

STIMULI

The stimuli consisted of two slightly different fingering patterns (see **Figure 1**). In each of the 56 experimental blocks, one type of pattern had to be produced with the right hand four times, in direct succession, in one of the following major keys: C-Major,

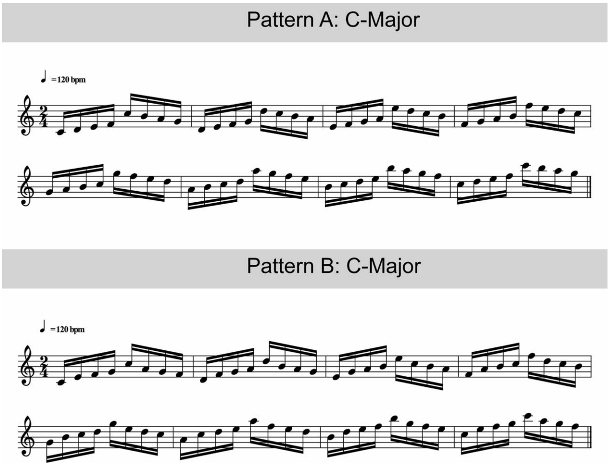

FIGURE 1 | Examples of the stimulus material. Participants were required to perform such sequences with their right hand at an IOI of 125 ms, using the fingering 1-2-3-4-5-4-3-2-1-2-3-4-5-4… etc. (1, thumb; 2, index finger; 3, middle finger; 4, ring finger; 5, pinkie finger).

D-Major, E-Major, F#-Major, G-Major, A-Major, or B-Major. The order of blocks was randomized with the constraints that no sequence in the identical key occurred in direct succession and that the same type of pattern was repeated a maximum of two times. Participants were instructed to use the same fingering throughout the experiment: 1-2-3-4-5-4-3-2-1-2-3… etc. (where 1 = thumb and 5 = pinkie). The instructed tempo for the sequences was 120 beats per minute—or 2 beats per second—for quarter notes. Because the required 16th notes have a duration of one quarter of a quarter note, this tempo resulted in an IOI between notes of 125 ms.

The sequences were performed on a Yamaha digital piano (S90XS; Yamaha Corporation, Japan), and participants listened to their performances with AKG 240 studio headphones at comfortable listening levels (dependent on the velocities of the key depressions). Due to technical difficulties, 4 participants performed on a Yamaha KX88 digital piano. Importantly, results of post-experimental questionnaires showed comparable levels of playing comfort between the two pianos. All tones had a standard MIDI piano timbre, generated by a Roland XV-2020 (Roland Corporation, Japan) synthesizer module.

PROCEDURE

Participants received the musical scores and tempo instructions prior to the experiment and were asked to memorize and rehearse the sequences (without looking at their hands while playing) with their own instrument at home.

After EEG and motion capture preparations, participants sat in front of the piano in a light-dimmed room. Before the experiment, participants could familiarize themselves with the piano, warm up, and perform one practice block. Participants were instructed to play as accurately as possible in the given tempo, but they were unaware of the exact aim of the study. Before each block, an instruction appeared on the screen placed above the keyboard and informed the participant about the type and key of the pattern to be performed. Simultaneously, four metronome beats were played to remind them of the correct tempo. After that, a green fixation circle in the center of the screen signaled that the participant could start playing. After each block, there was a short break and participants were able to continue the experiment by pressing a button whenever they were ready. Participants wore a custom-made visor that prevented them from visually monitoring the keys and their hands while playing but still allowed looking straight ahead at the screen. At the end of the experiment, participants completed questionnaires about their musical backgrounds and about the experiment. The whole experiment, including breaks and preparations, lasted approximately 3–4 h and pianists were paid for their participation.

DATA RECORDING

Musical data

MIDI data were recorded by a modified version of the FTAP software (Finney, 2001a,b). To synchronize MIDI, motion capture, and EEG data, the FTAP software sent synchronization signals concurrently with every fifth key press to the EEG recording device. Similarly, the motion capture system sent synchronization signals simultaneously with each recorded frame to the EEG recording device (for details of this setup, see Maidhof et al., 2013).

EEG data

The EEG was recorded with a Biosemi ActiveTwo system (Biosemi, The Netherlands) from 64 Ag/AgCl active electrodes placed according to the extended 10–20 system (Fp1, AF7, AF3, F1, F3, F5, F7, FT7, FC5, FC3, FC1, C1, C3, C5, T7, TP7, CP5, CP3, CP1, P1, P3, P5, P7, P9, PO7, PO3, O1, Iz, Oz, POz, Pz, CPz, Fpz, Fp2, AF8, AF4, Afz, Fz, F2, F4, F6, F8, FT8, FC6, FC4, FC2, FCz, Cz, C2, C4, C6, T8, TP8, CP6, CP4, CP2, P2, P4, P6, P8, P10, PO8, PO4, and O2). The horizontal electrooculogram (EOGH) was recorded from electrodes placed at the left and right outer canthi, and vertical EOG (EOGV) was recorded from electrodes placed above and below the left eye. Two additional electrodes were placed at the left and right mastoid. EEG signals were digitized with a sampling frequency of 8192 Hz. Low-pass filtering during recording was performed digitally in the ADC's decimation filter, which has a 5th-order sinc response with a -3 dB point at 1638.4 Hz (see also http://www.biosemi.com/faq/adjust_filter.htm).

Motion data

Eight infrared Qualisys ProReflex cameras (Qualisys, Sweden) recorded the three-dimensional position data of 25 reflective markers (4 mm in diameter) with a sampling frequency of 120 Hz. The markers were attached to the fingernails, each finger joint, the wrist, and the back of the right hand of each participant. Additional markers were placed on the C4 and B4 keys. Only the data from the five markers at the fingertips are reported here.

DATA ANALYSIS

Musical data

Performance errors were detected offline by using the MIDI toolbox for MATLAB (Eerola and Toiviainen, 2004) and its extension for matching a musical performance to its corresponding notation (Large, 1993). After identifying errors, only pitch (or substitution) errors were further analyzed; such substitution errors occur when participants play a different note than written in the score. Other error types like note omissions and additional notes were discarded. Furthermore, pitch errors entered the analysis only if they were preceded by at least three correct notes, and correct notes entered the analysis only if they were preceded and followed by at least three correct notes. IOIs of correct and incorrect notes were calculated by subtracting the MIDI onset of the previous note from the MIDI onset of the current note—that is, by calculating n minus $n - 1$ (IOIs thus refer to pre-note intervals). Only notes (correct and incorrect) that showed IOIs between 50 and 300 ms, and whose duration was between 50 and 180 ms, were selected.

Next, we calculated the mean MIDI keystroke velocity and mean IOI of correct and incorrect notes for each participant. Then, a subset of correct notes of each participant was created that included only correct notes that showed the same IOI as the mean IOI of incorrect notes (±5 ms). This selection of notes allowed for a comparison of correct and incorrect notes played with approximately the same tempo.

Motion data

The interaction between the finger and a rigid body (i.e., a key) during a keystroke can be determined by two kinematic landmarks (see Goebl and Palmer, 2008, 2009): a key-bottom contact (KB) landmark occurs when the finger's motion ceases as the key reaches the key bed, that is, when the key is fully pressed down; and a finger-key contact (FK) can occur when the finger initially touches the surface of the key before any pressure to the key is applied, that is, when the finger makes first contact before the key is actually pressed down (key depression). Both landmarks involve an abrupt change in acceleration and are therefore characterized by a peak in acceleration in the height-dimension of the finger's trajectory. Similar to Goebl and Palmer (2008), an FK landmark was identified when an acceleration peak in the finger trajectory was larger than 20 m/s^2 in a time window ranging from −150 to −20 ms (i.e., prior to the MIDI onset). KB landmarks were identified when the acceleration peak in the finger trajectory was larger than 5 m/s^2 in a time window ranging from −10 to 35 ms around MIDI onsets. Note that the kinematic data were analyzed independently from the EEG data.

EEG data

EEG data were processed offline in MATLAB (7.10.0) using the freely available toolbox EEGLAB 10.2.5.8b (Delorme and Makeig, 2004) and custom routines. To reduce the data size, EEG data were down-sampled to 256 Hz. Data were filtered by applying a high-pass filter [0.5 Hz, 3508 points, finite impulse response filter (FIR)], and subsequently a low-pass filter (45 Hz, 164 points, FIR).

Before performing an Independent Component Analysis (ICA), data segments contaminated with untypical and gross-movement artifacts were visually identified and removed. The cleaned datasets were then subjected to an extended Infomax version of ICA as implemented in the runica algorithm in EEGLAB. The resulting independent components (ICs) were screened for artifactual components due to eye blinks and movements, and electrode artifacts. An IC was considered to represent activity from eye blinks and movements if its topography showed peak activity only over the horizontal or vertical eye electrodes, if it showed a smoothly decreasing power spectrum, and if the component's activity contributed primarily to the raw EEG signal recorded by horizontal and vertical eye EOG. Identified artifactual ICs were subtracted from the data, which were subsequently low-pass filtered (25 Hz, 36 points, FIR), and re-referenced to the arithmetical mean of both mastoid electrodes.

Epochs representing single experimental trials time-locked to the MIDI and FK onsets of correct and incorrect keystrokes were extracted from −400 to 600 ms, respectively. To investigate if differences between the ERPs of correct and incorrect notes reflect differences in processing the previous notes rather than error detection mechanisms, two random subsets from the pool of all correct notes were compared. Each subset comprised approximately three times the number of incorrect trials for each participant. This procedure also reduced the amount of correct notes. Epochs were then baseline corrected (from −400 to −200 ms), and subjected to an automatic artifact rejection procedure, which discarded trials that showed larger or smaller amplitude values than $+60\,\mu$V or $-60\,\mu$V, respectively.

Consequently, ERPs time-locked to the MIDI onset were, on average across 10 participants, computed for (a) 105 incorrect keystrokes (±46); (b) a first subset of 300 correct keystrokes (±151); (c) a second subset of 298 correct keystrokes (±154); (d) 387 correct but slow keystrokes (±335); and (e) 105 correct pre-error notes (±48). ERPs time-locked to the onset of FK landmarks were, on average across all participants, computed for (a) 75 incorrect keystrokes (±41); (b) a first subset of 217 correct keystrokes (±116); (c) a second subset of 216 correct notes (±114); and (d) 78 correct pre-error notes (±41). For the comparisons of the two subsets of correct notes and correct pre-error notes, one participant had to be excluded because too few trials could be distributed into the two subsets. Similarly, the number of correct but slow notes was too small for this participant ($n = 21$). Note that the number of trials used for the ERP computation differs depending on whether the ERPs are time-locked to the MIDI onset or to the onset of the FK landmarks. This is because not all keystrokes showed a clear FK landmark (see Results section), and because the artifact rejection procedure resulted in discarding different trials.

Statistical evaluation

Based on visual inspection of the grand-averaged scalp topographies, mean ERP amplitude values were initially calculated for three regions of interest (ROIs): one frontal ROI (including electrodes F3, FZ, and F4), one central ROI (electrodes C3, CZ, and C4), and one parietal ROI (electrodes P3, PZ, P4). Time windows for statistical analysis were chosen based on the visual inspection of the grand-average ERPs and centered around the maximum of the differences between two conditions. Behavioral and movement data were statistically analyzed using paired sample t-tests. ERP data were analyzed with repeated measures ANOVAs with the factors Keystroke (correct, incorrect) and Frontality (frontal, central, parietal). The reported p-values were corrected using the Huynh-Feldt method where appropriate. The significance level for all tests was 0.05. All statistical analyses were performed with the software package PASW Statistics 18. ERPs were low-pass filtered (20 Hz) for presentation purposes only.

RESULTS

BEHAVIORAL RESULTS

Tempo

The mean IOI between two correct successive keystrokes was 129.4 ms ($SD = 31.3$ ms), indicating that participants were able to perform in the instructed tempo of 125 ms. Incorrect keystrokes were performed with a prolonged IOI ($M = 143.8$ ms, $SD = 31.4$ ms) compared to correct keystrokes [$t_{(1,\,9)} = -6.27$, $p < 0.0001$]. In addition, the IOIs of correctly performed notes immediately preceding erroneous keystrokes ("pre-error" notes) were prolonged ($M = 137.9$ ms, $SD = 30.2$ ms), compared to IOIs of correct keystrokes [$t_{(1,\,9)} = -1.31$, $p = 0.025$].

MIDI velocity

The MIDI velocities of incorrect keystrokes ($M = 61.2$, $SD = 7.1$) were decreased compared to correct keystrokes [$M = 68.4$, $SD = 7.2$; $t_{(1,\ 9)} = 5.61$, $p < 0.0001$]. However, MIDI velocities of pre-error notes ($M = 70.2$, $SD = 4.7$) and correct notes did not differ [$t_{(1,\ 9)} = -1.49, p = 0.17$].

Kinematic results

An example of a keystroke showing a finger-key landmark (occurring when the finger touches the surface of the piano key) followed by a key-bottom landmark (occurring when the key reaches the key bed) is shown in **Figure 2**.

Results revealed that the majority of keystrokes showed a key-bottom landmark. In 99.8% of all correct and 97.7% of all incorrect keystrokes, a KB landmark was detected. However, the number of KB landmarks differed significantly between the two conditions [$t_{(1,\ 9)} = 4.99$, $p = 0.001$]. In contrast, percentages of keystrokes with an FK landmark did not differ between correct and incorrect keystrokes (ca. 78% for correct keystrokes and ca. 82% of all incorrect keystrokes; $p = 0.44$). KB landmarks occurred around 13 ms after MIDI onset, regardless of the correctness of keystrokes ($p = 0.62$). However, the mean amplitude of KB landmarks of incorrect keystrokes (27.9 m/s^2, $SD = 3.2$ m/s^2 was decreased compared to correct keystrokes [34.3 m/s^2, $SD = 4.6$ m/s^2; $t_{(1,\ 9)} = 6.21$, $p < 0.0001$]. This is consistent with the decreased MIDI velocity and indicates that erroneous keystrokes were performed with slower downward movements than correct keystrokes.

In contrast, FK landmarks during incorrect keystrokes occurred significantly earlier than FK landmarks during correct keystrokes. On average, FK landmarks of correct notes occurred 51.4 ms ($SD = 8.3$ ms) prior to MIDI onsets, whereas FK landmarks of error notes occurred 59.6 ms ($SD = 8.3$ ms) prior to MIDI onsets [$t_{(1,\ 9)} = 3.4$, $p = 0.008$]. The distances of the FK landmarks with respect to the MIDI onset of the previous note did not differ between correct and incorrect keystrokes [$t_{(1,\ 9)} = -1.207$, $p = 0.258$], and occurred on average around 80 ms after the MIDI onsets. The mean amplitude of FK landmarks was around 48 ms/s^2 and did not differ between correct and incorrect keystrokes ($p = 0.27$).

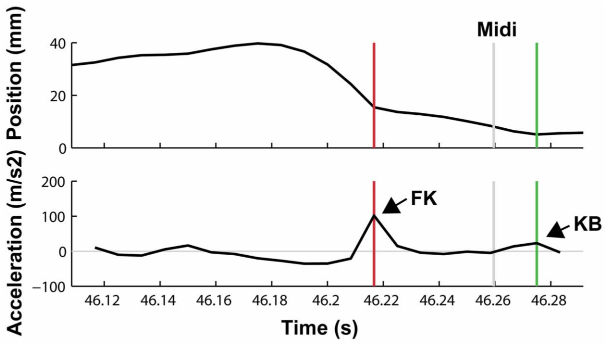

FIGURE 2 | Vertical motion of the fingertip of a pianist's ring finger playing the G#5 key. Upper panel: finger position, lower panel: acceleration. An acceleration peak can be observed around 40 ms prior to the MIDI note onset, occurring when the fingertip makes initial contact with the key surface (finger-key landmark, FK). Shortly after the MIDI onset, another acceleration peak occurs when the key reaches the key bed after key depression (key-bottom landmark, KB).

Although correct notes immediately preceding incorrect notes showed a prolonged IOI, they did not differ from other correct keystrokes in terms of percentage, amplitude, or latency of KB and FK landmarks (p's > 0.2).

ERP RESULTS

MIDI-based

First, we compared the ERPs time-locked to the onset of the MIDI signal of incorrect and correct keystrokes, similar to previous studies (Herrojo Ruiz et al., 2009; Maidhof et al., 2009; Herrojo Ruiz et al., 2011). Compared to correct keystrokes, incorrect keystrokes showed an increased negativity prior to the onset of a keystroke. The difference was maximal around 50 ms prior to the onset of the key press, and showed a frontally distributed scalp topography (see **Figure 3A**). An ANOVA for a time window of −70 to −30 ms showed an interaction between factors Keystroke and Frontality [$F_{(1,\ 9)} = 4.741$, $p = 0.048$], and no main effects (p's > 0.14). A follow-up analysis showed an effect of Keystroke only over the frontal ROI [$F_{(1,\ 9)} = 7.147$, $p = 0.025$]. This early difference was followed by a positive deflection showing maximal amplitudes around 250 ms after the MIDI onset over more central leads. An ANOVA for a time window of 200–350 ms showed main effects of Keystroke [$F_{(1,\ 9)} = 13.632$, $p = 0.005$], Frontality [$F_{(1,\ 9)} = 5.005$, $p = 0.019$], and an interaction between these factors [$F_{(1,\ 9)} = 5.226, p = 0.016$], indicating that amplitude values were larger over central leads.

To investigate whether the observed ERP differences were influenced by overlapping ERP responses from previous notes (e.g., change detection in pitch) and thus to investigate whether the above-mentioned findings are error-specific, we compared the ERPs of two random subsets of correct notes. The rationale was that if ERP effects reflect mainly error-related processes, the comparison of random correct notes should not show any differences. The results of this comparison are shown in **Figure 3B**. ANOVAs in the same time windows of the pre-error negativity and the later positivity (i.e., from −70 to −30 ms, and from 200 to 350 ms, respectively) showed no differences between the ERPs of randomly selected subsets of correct notes (all p's > 0.32).

Next, we investigated the influence of different IOIs of correct and incorrect keystrokes on the ERP effects. **Figure 4A** shows the grand-averaged waveforms time-locked to the onset of the MIDI signal of incorrect keystrokes and a subset of correct keystrokes: only correct keystrokes that matched the mean IOIs of incorrect keystrokes ±5 ms were analyzed (see Method section for details). In line with previous results, incorrect keystrokes elicited an increased negativity prior to the onset of a keystroke compared to (slow, but) correct keystrokes. The difference was maximal around 60 ms prior to the onset of the MIDI signal, and showed a frontal scalp distribution. However, this difference

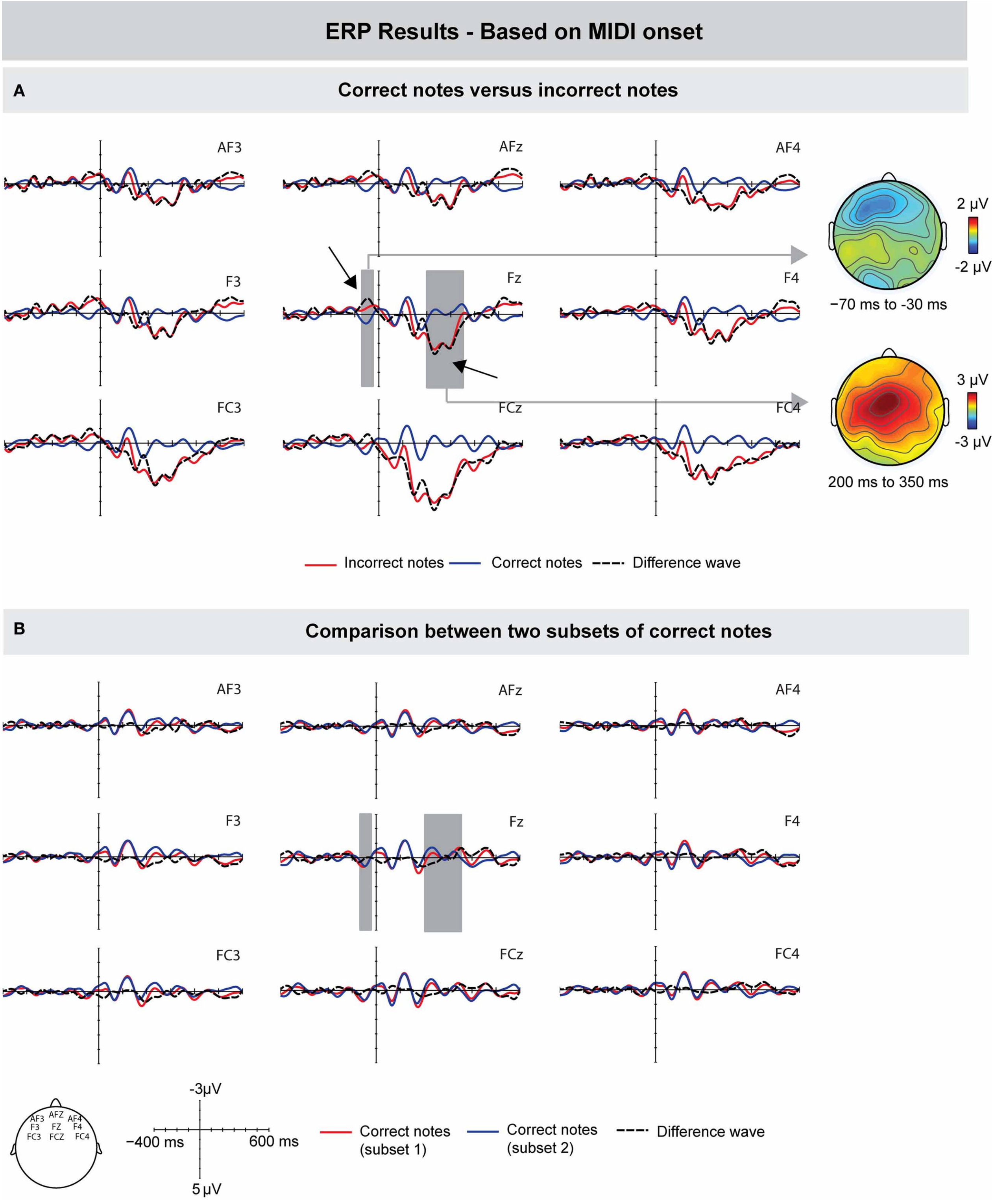

FIGURE 3 | Grand-averaged ERPs (left) and scalp topographies (right) based on the MIDI data. **(A)** Compared to correctly produced keystrokes, incorrectly produced keystrokes elicited a pre-error negativity around 50 ms prior to MIDI onset, and an Error Positivity (Pe) around 250 ms after MIDI onset. **(B)** To investigate whether the difference between correct and incorrect notes is influenced by overlapping ERP responses from previous notes, two subsets of correct notes were compared. However, ERPs did not differ in the time windows used for the comparison between correct and incorrect notes. Gray areas show the time windows used for statistical analysis, exemplarily on one electrode. On the right side, corresponding scalp topographies are shown as the difference potentials between two conditions in the given time windows.

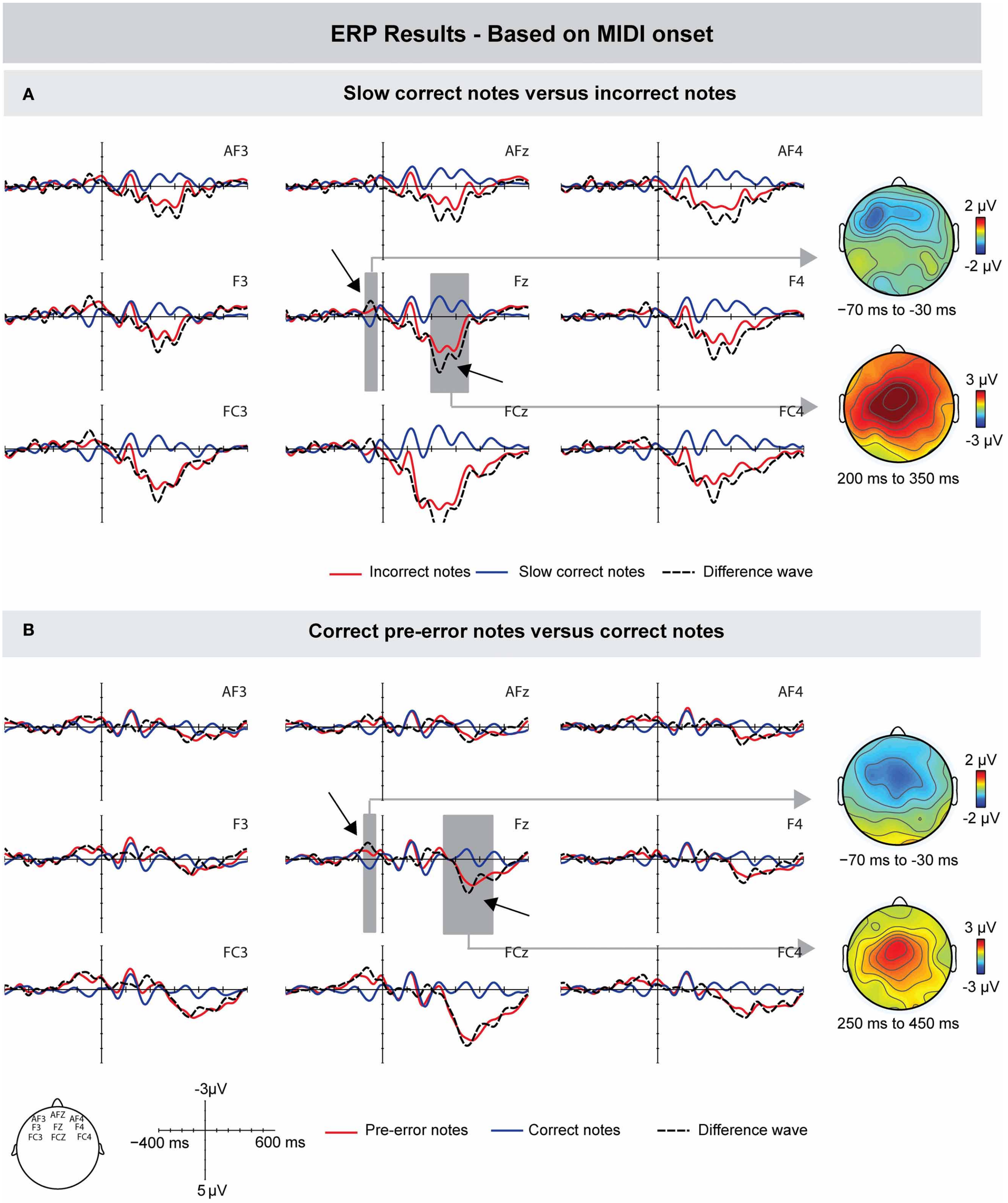

FIGURE 4 | Grand-averaged ERPs (left) and scalp topographies (right) based on the MIDI data. (A) Compared to correct (but slow) keystrokes, incorrect keystrokes elicited a similar (marginally significant) negativity peaking around 50 ms prior to MIDI onset, followed by a Pe around 220 ms after MIDI onset. **(B)** Compared to other correct keystrokes, the correct keystrokes immediately preceding incorrect keystrokes elicited a negativity that preceded MIDI note onsets by about 60 ms, followed by a positivity peaking around 350 ms after MIDI note onsets. Gray areas show the time windows used for statistical analysis, exemplarily on one electrode. On the right side, corresponding scalp topographies are shown as the difference potentials between two conditions in the given time windows.

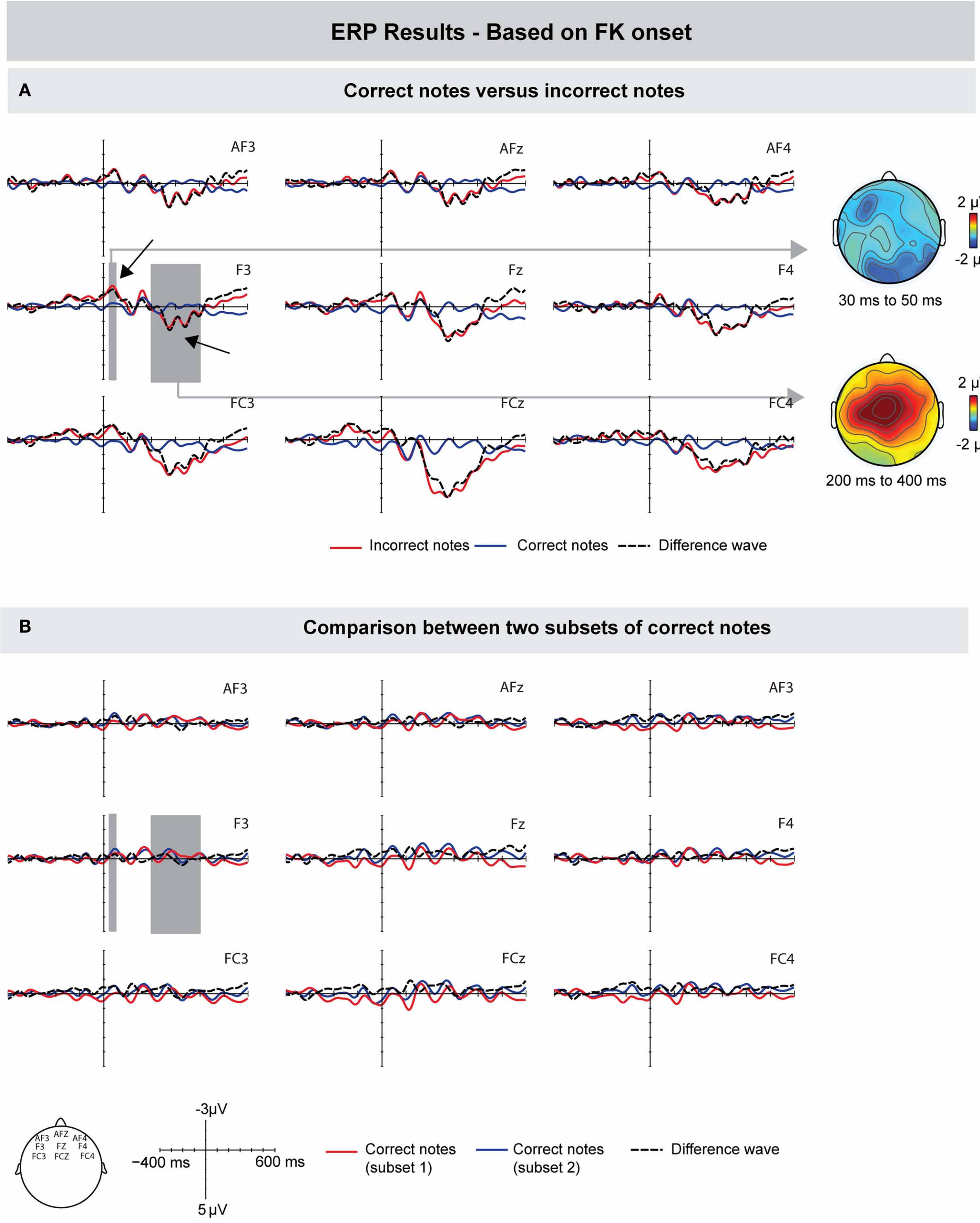

FIGURE 5 | Grand-averaged ERPs (left) and scalp topographies (right) based on the movement data. ERPs are time-locked to the finger-key landmarks (FK) in the motion trajectories, occurring when a finger makes initial contact with the surface of a key. **(A)** Compared to correctly produced keystrokes, incorrectly produced keystrokes showed a marginally significant increased negativity around 40 ms after FK onset, which was focused only to left-frontal electrodes. Incorrect notes elicited a positive deflection peaking around 280 ms after the onset of tactile feedback. **(B)** To investigate whether the difference between correct and incorrect notes is influenced by overlapping ERP responses from previous notes, two subsets of correct notes were compared. ERPs showed marginally significant differences in the time windows used for the comparison of correct and incorrect notes. Therefore, results of the ERPs time-locked to FK onsets can only be tentatively interpreted. Gray areas show the time windows used for statistical analysis, exemplarily on one electrode.

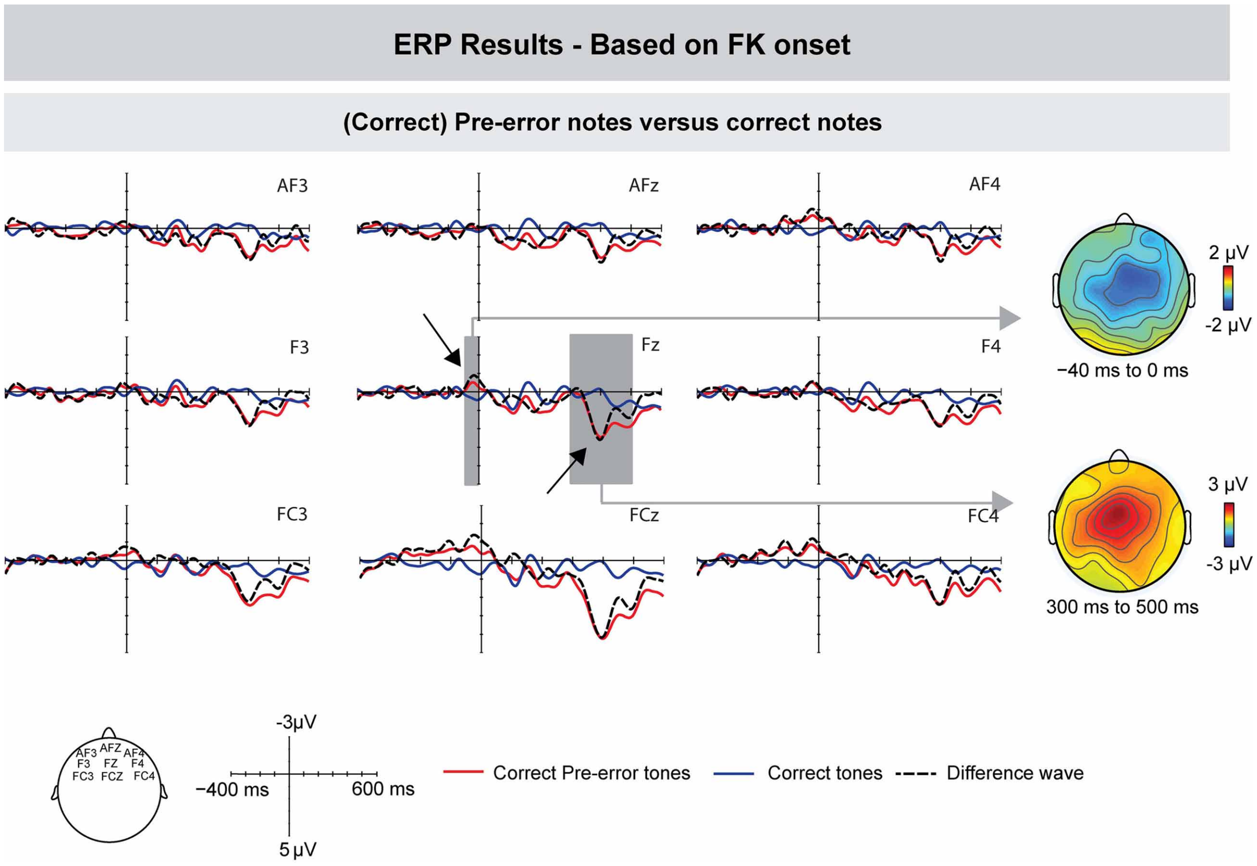

FIGURE 6 | Grand-averaged ERPs (left) and scalp topographies (right) based on the movement data. Compared to other correct keystrokes, pre-error notes showed an increased negativity shortly before FK onset, followed by a positive deflection around 350 ms. Gray areas show the time windows used for statistical analysis, exemplarily on one electrode. On the right side, corresponding scalp topographies are shown as the difference potentials between two conditions in the given time windows.

was only marginally significant: an ANOVA for a time window ranging from −70 to −30 ms showed a marginally significant interaction between factors Keystroke and Frontality [$F_{(1,9)} = 4.248$, $p = 0.058$]. Subsequent ANOVAs computed for the three ROIs separately showed only an effect that was approaching significance over the frontal ROI [$F_{(1,9)} = 3.473$, $p = 0.095$].

The early difference was followed by a positive deflection peaking around 220 ms after key press onset with maximal amplitudes over more central leads. An ANOVA for a time window of 200–350 ms showed main effects of Keystroke [$F_{(1,9)} = 33.277$, $p < 0.0001$], Frontality [$F_{(1,9)} = 4.398$, $p = 0.028$], and an interaction between these factors [$F_{(1,9)} = 6.458$, $p = 0.008$], indicating that this effect was more centrally distributed.

Furthermore, we compared the ERPs elicited during correct events that immediately preceded incorrect pitch events with the ERPs elicited during other correct pitch events (see **Figure 4B**). Results showed that (correct) pre-error notes also elicited, compared with other correct notes, an increased negativity prior to the MIDI onset. This difference peaked around −60 ms and showed a more central scalp topography. An ANOVA for the time window of −70 to −30 ms showed a marginally significant effect of Keystroke [$F_{(1,9)} = 4.852$, $p = 0.055$], and an interaction between the factors Keystroke and Frontality [$F_{(1,9)} = 4.955$, $p = 0.025$]. The early difference was followed by a positivity with a peak latency of around 350 ms (starting around 300 ms): an ANOVA in a time window of 250–450 ms showed a marginally significant effect of Keystroke [$F_{(1,9)} = 4.721$, $p = 0.058$]. Given the mean IOI of ∼130 ms, the latency of 350 ms is consistent with the latency of the positivity elicited by incorrect notes (around 220 ms). Hence, the late positivity seen in the ERPs of pre-error notes is most likely due to the following incorrectly produced note.

Motion-based

To investigate the role of tactile feedback during erroneous keystrokes, we computed the ERPs relative to the onset of FK landmarks, that is, when a finger touches the surface of a key before it is pressed down (instead of relative to the MIDI onset, which occurs when the key is already pressed down). **Figure 5A** shows that incorrect pitch events, compared to correct pitch events, elicited a small negative deflection with a peak latency around 40 ms after the onset of tactile feedback. However, this difference seemed to be focused only to left-frontal electrodes F3 and FC3. An ANOVA for a time window of 30–50 ms after FK landmark onset with the same ROIs used for the other analyses showed no effect of Keystroke and no interaction with this factor (p's > 0.19). Upon close visual inspection of the grand-average waveform, we conducted an additional ANOVA for the means of electrodes F3 and FC3, which showed a marginally significant effect of Keystroke [$F_{(1,\ 9)} = 4.469\ p = 0.059$]. Importantly, there was no difference prior to the onset of tactile feedback (ANOVA for the same time window as in the MIDI-based ERPs, i.e., −70 to −30 ms: p's > 0.18). Around 280 ms, incorrect notes elicited a positive deflection with a central scalp distribution. An ANOVA for a time window of 200–400 ms over frontal, central, and parietal ROIs showed main effects of Keystroke [$F_{(1,\ 9)} = 7.752$, $p = 0.021$], Frontality [$F_{(1,\ 9)} = 4.382$, $p = 0.03$], and an interaction between Keystroke and Frontality [$F_{(1,\ 9)} = 4.578$, $p = 0.025$]. Note that the latency of this positivity is consistent with the peak latency of around 220 ms after MIDI onset, given that FK landmarks occurred around 50 ms prior to MIDI onsets.

To investigate whether this small difference was influenced by overlapping ERP responses from previous notes, we again compared the ERPs of two random subsets of correct notes (analogous to the comparison for the ERPs based on the MIDI signal; **Figure 5B**). Results showed that ERPs of only correct notes differed marginally significantly in the same time window as the difference between correct and incorrect keystrokes was observed [ANOVA for a time window of 30–50 ms over electrodes F3 and FC3: $F_{(1,\ 8)} = 4.932$, $p = 0.057$]. Thus, it seems that overlapping ERP responses from previous notes influenced the difference between correct and incorrect notes, and therefore that results of ERPs time-locked to the onset of FK landmarks can only be interpreted tentatively.

The ERPs of correct notes immediately preceding incorrect notes and ERPs of other correct notes are depicted in **Figure 6** (for this analysis, one participant was excluded due to there being too few trials). In contrast to correctly produced pitch events (elsewhere in the sequences), pre-error notes elicited no negative deflection shortly after the onset of tactile feedback (ANOVA for a time window of 0–50 ms: F's < 1). However, pre-error notes showed an increased negativity prior to the onset of tactile feedback: an ANOVA for a time window of −70 to −30 ms showed a marginally significant effect of Keystroke [$F_{(1,\ 8)} = 3.851$, $p = 0.085$] and an interaction between Keystroke and Frontality [$F_{(1,\ 8)} = 4.88$, $p = 0.028$], indicating that the effect was more pronounced over central leads.

Around 350 ms after FK onset, pre-error notes elicited an increased positive deflection, compared to other correct notes; an ANOVA for a time window of 250–450 ms showed a marginally significant effect of Keystroke [$F_{(1,\ 8)} = 4.721, p = 0.058$].

DISCUSSION

The present study investigated errors during the performance of musical sequences by using a novel setup combining the recording of EEG, MIDI, and 3D movement data (for details of this setup, see Maidhof et al., 2013; for a non-musical setup aiming at mobile EEG recording of freely moving participants, see Makeig et al., 2009). This allowed us to investigate the underlying movements of pitch errors during piano performances, but also to exploratively relate the neurophysiological findings to different movement phases.

Replicating previous behavioral findings (Herrojo Ruiz et al., 2009; Maidhof et al., 2009; Herrojo Ruiz et al., 2011; Strübing et al., 2012), erroneous keystrokes were executed with reduced intensities (in terms of MIDI velocities) compared to correct keystrokes. Furthermore, the IOIs of incorrectly produced pitch events were prolonged compared to correct keystrokes (by ca. 13 ms). However, the analyses of kinematic landmarks during keystrokes provided additional information about the underlying movements. Importantly, the onset of tactile feedback for incorrect pitch events (FK landmarks, occurring when a finger makes first contact with the key surface; Goebl and Palmer, 2008) occurred ca. 60 ms prior to key depression, whereas tactile feedback for correct pitch events started ca. 50 ms prior to key depression. Considering that incorrect events were produced ca. 13 ms slower than correct events, we assume that it was not the movement toward the next incorrect key itself that was executed slower, but rather that the phase between touching the key surface and complete key depression was prolonged for incorrect keystrokes. This assumption is supported by the finding that acceleration values did not differ when a finger made initial contact with the key surface during correct or incorrect pitch events. Thus, it is likely that the prolonged IOIs of incorrect notes are mainly due to the decreased velocity of those wrong keystrokes.

Interestingly, IOIs (but not MIDI velocities) of correctly produced notes immediately preceding incorrect notes were also prolonged, compared to other correct events elsewhere in a sequence. However, these pre-error notes did not differ in latency or amplitude of FK landmarks from other correct events. Thus, it remains unclear what exactly caused this prolongation, and future studies should investigate this effect in more detail.

ERPs based on the MIDI data replicated previous findings showing that incorrect keystrokes, compared to correct key presses, elicited an increased negativity already ∼60 ms before a key was fully pressed down, and before auditory feedback of the error was available (pre-error negativity; Herrojo Ruiz et al., 2009; Maidhof et al., 2009; Strübing et al., 2012). This early negative potential was followed by a centrally distributed positive deflection, resembling the Error Positivity (Pe) or P300 component (note that Pe might actually reflect the same neural mechanisms as the P300; for reviews of the Pe, see Falkenstein et al., 2000; Overbeek et al., 2005).

However, these effects could be due to the different IOIs before correct and incorrect notes, and hence these findings could be confounded by the factor tempo. To exclude this possibility, we calculated the ERPs for a subset of correct notes that were closely matched to the tempo of incorrect notes. Results showed virtually the same ERP pattern, that is, an increased negativity prior to the MIDI onsets of incorrect notes compared to correct notes, and a subsequent P300 component.

Furthermore, these effects could also have been caused or influenced by overlapping ERPs (due to the short IOIs of around 125 ms) from other sensory-motor processes related to previous notes, irrespective of the correctness of the played notes. However, there are several reasons rendering it unlikely that the pre-error negativity and the Pe reflect simply overlapping sensory-motor processes and that they are not being elicited by pitch errors: first, the results of the comparison of two random subsets of correct notes showed no significant differences in the time windows of the pre-error negativity and the Pe, indicating that the effects are rather error-specific. Second, a previous study also reported that a pre-error negativity was also elicited in the complete absence of auditory feedback (Herrojo Ruiz et al., 2009), suggesting that the processing of previous pitch information did not significantly contribute to this effect. Third, a symbolic resonance analysis aimed at disentangling overlapping ERPs during piano performance validated the previous ERP results (Herrojo Ruiz et al., 2009).

In sum, these results, in combination with the kinematic data, suggest that the pre-error negativity and the Pe during piano errors are unlikely only due to tempo differences between correctly and incorrectly produced pitch events, and unlikely only due to overlapping ERPs elicited by previous notes. Instead, it appears that the early negative potential and the following positivity reflect error-related processes.

To investigate the role of tactile feedback for error-related processes during piano performance, we exploratively computed ERPs based on the onset of tactile feedback for each keystroke, which was determined based on the movement data. Results suggest that incorrectly produced pitch events did not differ prior to the tactile feedback of keystrokes, but that they were associated with a slightly increased negativity peaking around 40 ms after tactile feedback was available.

However, these findings can only be interpreted tentatively, because the comparison between the ERPs of two random subsets of correct notes revealed a difference in the time window of 0–50 ms. This indicates that, when ERPs were time-locked to the onset of tactile feedback, the difference between correct and incorrect keystrokes might have been influenced by overlapping ERP responses elicited by previous notes. In addition, the motion capture system had a relatively large sampling interval of 8.3 ms, and therefore the detection of FK landmarks was not as accurate as in previous studies (Goebl and Palmer, 2008; Palmer et al., 2009). Therefore, averaging across FK landmarks might have "smeared" the ERPs considerably, which could also have contributed to the difference between correct notes. It would be interesting to see whether the same results would be obtained with a motion capture system with a higher sample rate.

Nevertheless, one might speculate about the current finding of a negative difference potential after the onset of tactile feedback. One possible interpretation is that tactile and proprioceptive feedback may play an important role in the detection of upcoming performance errors: it is conceivable that, based on the tactile feedback of the key surface (but not earlier, for instance during the movement toward the key), the monitoring system can compare the predicted with the actual sensory consequences of a movement in the form of reafferent tactile information. By contrast, when the auditory feedback of keystrokes or of speech acts is externally manipulated, the earliest information with which the predicted consequences can be compared is the auditory feedback. Thus, brain responses after the onset of the auditory feedback are increased (or inhibitory effects are canceled) when a mismatch is detected (e.g., Houde et al., 2002; Heinks-Maldonado et al., 2006; Eliades and Wang, 2008; Katahira et al., 2008; Maidhof et al., 2013, 2010; Behroozmand and Larson, 2011; Behroozmand et al., 2011). However, when participants commit errors, information available even earlier, including tactile feedback, could be used for error detection. In this regard, predictions regarding the consequences of movements would include predictions of incoming tactile as well as auditory information, which is consistent with the notion that efference copies can interact at several stages of sensory processing (Crapse and Sommer, 2008). Furthermore, corrective modulations of the ongoing motor act can be initiated whenever a mismatch is detected. These corrective modulations can include the slowing of the ongoing keystroke resulting in a decreased loudness of the incorrect pitch, as indexed by the decreased MIDI velocity (and the corresponding lower acceleration values of finger trajectories during incorrect pitch events).

Note that the negativity peaking shortly after the touch of a key could also be interpreted as an error-related negativity in the context of existing theories of action monitoring. For example, the mismatch detection hypothesis holds that the ERN is elicited when a comparator detects a mismatch between the correct response and the actual response, and subsequently triggers an error signal (Falkenstein et al., 2000; Coles et al., 2001). Similarly, the Reinforcement-Learning theory of the ERN (for a review, see Nieuwenhuis et al., 2004), posits that the earliest available information about incorrect performance will generate an error signal. In the case of piano performance, one can speculate that the tactile feedback from the finger arriving at a key (and the proprioceptive feedback about the position of the finger) could represent the first indication of an incorrect performance.

However, from the present data it is difficult to conclude whether the negative difference reflects the error signal itself, or the initiation of behavioral adjustments to prevent the error (or at least to decrease the sensory effects caused by the error, i.e., reduce the loudness).

When time-locked to MIDI onsets, (correct) pre-error notes elicited a similar ERP pattern as incorrect notes, although the early negativity showed a slightly more central scalp topography. However, when time-locked to tactile feedback, the tentative ERP results of pre-error notes did not differ from other correct notes after the onset of tactile feedback, but only shortly before tactile information was available. Therefore, one might speculate that different and/or additional neural processes are operating during the execution of correct notes preceding wrong pitch events. These results might indicate that the monitoring system detected some problem in motor execution or planning (possibly resulting in an increased IOI), which, however, had not yet resulted in an incorrect pitch event. Hence, there is no mismatch between predicted and actual tactile feedback, and no correction has to be initiated, and thus no negativity is elicited after tactile feedback (of the correct key). Clearly, more research is needed to investigate this issue in more detail.

Taken together, the present study provided some further insights into the neural mechanisms of error processing during music performance, by combining electrophysiological with detailed three-dimensional movement data in an exploratory approach. Furthermore, the results tentatively suggest that the tactile feedback of piano keys plays a major role for predictive error processes, although future studies are needed to validate this interpretation. Interestingly, correct notes preceding errors showed similar neural activity to pitch errors themselves, although only pitch errors elicited increased neural activity after the onset of tactile feedback. In the future, we believe that the combined acquisition of electrophysiological and movement data can lead to a more behaviorally-driven analysis of brain activity during the performance of music. This approach also offers interesting possibilities in terms of conducting studies in other domains such as in music learning and education, music therapy, action-perception interactions, and musical expressivity in cross-sectional and—importantly—longitudinal paradigms, to reveal the time course and sensitivity of processes involved in learning and therapy.

ACKNOWLEDGMENTS

We are especially thankful to Birgitta Burger, Tommi Himberg, and Mikko Leimu for their help with the motion capture system, as well as Tommi Makkonen and Miika Leminen for their help with the data analysis, and Ben Gold for improving the English. We would also like to thank Werner Goebl for his support and providing us with the Matlab code of his kinematic data analysis. The study was financially supported by the Academy of Finland and the Center for International Mobility (CIMO), Finland.

REFERENCES

Behroozmand, R., and Larson, C. R. (2011). Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neurosci.* 12:54. doi: 10.1186/1471-2202-12-54

Behroozmand, R., Liu, H., and Larson, C. R. (2011). Time-dependent neural processing of auditory feedback during voice pitch error detection. *J. Cogn. Neurosci.* 23, 1205–1217. doi: 10.1162/jocn.2010.21447

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., and Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol. Rev.* 108, 624–652. doi: 10.1037/0033-295X.108.3.624

Carter, C. S. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* 280, 747–749. doi: 10.1126/science.280.5364.747

Coles, M. G., Scheffers, M. K., and Holroyd, C. B. (2001). Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. *Biol. Psychol.* 56, 173–189. doi: 10.1016/S0301-0511(01)00076-X

Crapse, T. B., and Sommer, M. A. (2008). Corollary discharge circuits in the primate brain. *Curr. Opin. Neurobiol.* 18, 552–557. doi: 10.1016/j.conb.2008.09.017

Delorme, A., and 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. doi: 10.1016/j.jneumeth.2003.10.009

Desmurget, M., and Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends Cogn. Sci.* 4, 423–431. doi: 10.1016/S1364-6613(00)01537-0

Eerola, T., and Toiviainen, P. (2004). *MIDI Toolbox: MATLAB Tools for Music Research.* Jyväskylä: University of Jyväskylä.

Eliades, S. J., and Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* 453, 1102–1106. doi: 10.1038/nature06910

Ericsson, K. A., Krampe, R. T., and Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychol. Rev.* 100, 363–406. doi: 10.1037/0033-295X.100.3.363

Ericsson, K. A., and Lehmann, A. C. (1996). Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annu. Rev. Psychol.* 47, 273–305. doi: 10.1146/annurev.psych.47.1.273

Falkenstein, M., Hohnsbein, J., Hoormann, J., and Blanke, L. (1990). "Effects of errors in choice reaction tasks on the ERP under focused and divided attention," in *Psychophysiological Brain Research*, eds C. H. M. Brunia, A. W. K. Gaillard, and A. Kok (Tilburg: Tilburg University Press), 192–195.

Falkenstein, M., Hoormann, J., Christ, S., and Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biol. Psychol.* 51, 87–107. doi: 10.1016/S0301-0511(99)00031-9

Ferdinand, N. K., Mecklinger, A., Kray, J., and Gehring, W. J. (2012). The processing of unexpected positive response outcomes in the mediofrontal cortex. *J. Neurosci.* 32, 12087–12092. doi: 10.1523/JNEUROSCI.1410-12.2012

Finney, S. A. (2001a). FTAP: a Linux-based program for tapping and music experiments. *Behav. Res. Methods Instrum. Comput.* 33, 65–72. doi: 10.3758/BF03195348

Finney, S. A. (2001b). Real-time data collection in Linux: a case study. *Behav. Res. Methods Instrum. Comput.* 33, 167–173. doi: 10.3758/BF03195362

Ganushchak, L. Y., and Schiller, N. O. (2006). Effects of time pressure on verbal self-monitoring: an ERP study. *Brain Res.* 1125, 104–115. doi: 10.1016/j.brainres.2006.09.096

Ganushchak, L. Y., and Schiller, N. O. (2008). Brain error-monitoring activity is affected by semantic relatedness: an event-related brain potentials study. *J. Cogn. Neurosci.* 20, 927–940. doi: 10.1162/jocn.2008.20514

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., and Donchin, E. (1993). A neural system for error detection and compensation. *Psychol. Sci.* 4, 385–390. doi: 10.1111/j.1467-9280.1993.tb00586.x

Goebl, W., and Palmer, C. (2008). Tactile feedback and timing accuracy in piano performance. *Exp. Brain Res.* 186, 471–479. doi: 10.1007/s00221-007-1252-1

Goebl, W., and Palmer, C. (2009). Synchronization of timing and motion among performance musicians. *Music Percept.* 26, 427–439. doi: 10.1525/mp.2009.26.5.427

Goebl, W., and Palmer, C. (2013). Temporal control and hand movement efficiency in skilled music performance. *PLoS ONE* 8:e50901. doi: 10.1371/journal.pone.0050901

Hajcak, G., Holroyd, C. B., Moser, J. S., and Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology* 42, 161–170. doi: 10.1111/j.1469-8986.2005.00278.x

Hajcak, G., Moser, J. S., Holroyd, C. B., and Simons, R. F. (2007). It's worse than you thought: the feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology* 44, 905–912. doi: 10.1111/j.1469-8986.2007.00567.x

Heinks-Maldonado, T. H., Nagarajan, S. S., and Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *Neuroreport* 17, 1375–1379. doi: 10.1097/01.wnr.0000233102.43526.e9

Herrojo Ruiz, M., Jabusch, H.-C., and Altenmüller, E. (2009). Detecting wrong notes in advance: neuronal correlates of error monitoring in pianists. *Cereb. Cortex* 19, 2625–2639. doi: 10.1093/cercor/bhp021

Herrojo Ruiz, M., Strübing, F., Jabusch, H.-C., and Altenmüller, E. (2011). EEG oscillatory patterns are associated with error prediction during music performance and are altered in musician's dystonia. *Neuroimage* 55, 1791–1803. doi: 10.1016/j.neuroimage.2010.12.050

Holroyd, C. B., and Coles, M. G. H. (2002). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* 109, 679–709. doi: 10.1037/0033-295X.109.4.679

Holroyd, C. B., Dien, J., and Coles, M. G. (1998). Error-related scalp potentials elicited by hand and foot movements: evidence for an output-independent error-processing system in humans. *Neurosci. Lett.* 242, 65–68. doi: 10.1016/S0304-3940(98)00035-4

Houde, J. F., Nagarajan, S. S., Sekihara, K., and Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: an MEG study. *J. Cogn. Neurosci.* 14, 1125–1138. doi: 10.1162/089892902760807140

Katahira, K., Abla, D., Masuda, S., and Okanoya, K. (2008). Feedback-based error monitoring processes during musical performance: an ERP study. *Neurosci. Res.* 61, 120–128. doi: 10.1016/j.neures.2008.02.001

Large, E. W. (1993). Dynamic programming for the analysis of serial behaviors. *Behav. Res. Methods Instrum. Comput.* 25, 238–241. doi: 10.3758/BF03204504

Maidhof, C., Kästner, T., and Makkonen, T. (2013). Combining EEG, MIDI, and motion capture techniques for investigating musical performance. *Behav. Res. Methods*. doi: 10.3758/s13428-013-0363-9. [Epub ahead of print].

Maidhof, C., Rieger, M., Prinz, W., and Koelsch, S. (2009). Nobody is perfect: ERP effects prior to performance errors in musicians indicate fast monitoring processes. *PLoS ONE* 4:e5032. doi: 10.1371/journal.pone.0005032

Maidhof, C., Vavatzanidis, N., Prinz, W., Rieger, M., and Koelsch, S. (2010). Processing expectancy violations during music performance and perception: an ERP study. *J. Cogn. Neurosci.* 22, 2401–2413. doi: 10.1162/jocn.2009.21332

Makeig, S., Gramann, K., Jung, T.-P., Sejnowski, T. J., and Poizner, H. (2009). Linking brain, mind and behavior. *Int. J. Psychophysiol.* 73, 95–100. doi: 10.1016/j.ijpsycho.2008.11.008

Masaki, H., Tanaka, H., Takasawa, N., and Yamazaki, K. (2001). Error-related brain potentials elicited by vocal errors. *Neuroreport* 12, 1851–1855. doi: 10.1097/00001756-200107030-00018

Miltner, W. H. R., Braun, C. H., and Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "Generic" neural system for error detection. *J. Cogn. Neurosci.* 9, 788–798. doi: 10.1162/jocn.1997.9.6.788

Münte, T. F., Altenmüller, E. O., and Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nat. Rev. Neurosci.* 3, 473–478.

Nieuwenhuis, S., Holroyd, C. B., Mol, N., and Coles, M. G. H. (2004). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neurosci. Biobehav. Rev.* 28, 441–448. doi: 10.1016/j.neubiorev.2004.05.003

Oliveira, F. T. P., McDonald, J. J., and Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *J. Cogn. Neurosci.* 19, 1994–2004. doi: 10.1162/jocn.2007.19.12.1994

Overbeek, T. J. M., Nieuwenhuis, S., and Ridderinkhof, K. R. (2005). Dissociable components of error processing. *J. Psychophysiol.* 19, 319–329. doi: 10.1027/0269-8803.19.4.319

Palmer, C. (1997). Music performance. *Annu. Rev. Psychol.* 48, 115–138. doi: 10.1146/annurev.psych.48.1.115

Palmer, C., Koopmans, E., Loehr, J. D., and Carter, C. (2009). Movement-related feedback and temporal accuracy in clarinet performance. *Music Percept.* 26, 439–449. doi: 10.1525/mp.2009.26.5.439

Palmer, C., Mathias, B., and Anderson, M. (2012). Sensorimotor mechanisms in music performance: actions that go partially wrong. *Ann. N.Y. Acad. Sci.* 1252, 185–191. doi: 10.1111/j.1749-6632.2011.06427.x

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447. doi: 10.1126/science.1100301

Riès, S., Janssen, N., Dufau, S., Alario, F.-X., and Burle, B. (2011). General-purpose monitoring during speech production. *J. Cogn. Neurosci.* 23, 1419–1436. doi: 10.1162/jocn.2010.21467

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., and Moore, D. G. (1996). The role of practice in the development of performing musicians. *Br. J. Psychol.* 87, 287–309. doi: 10.1111/j.2044-8295.1996.tb02591.x

Strübing, F., Herrojo Ruiz, M., Jabusch, H. C., and Altenmüller, E. O. (2012). Error monitoring is altered in musician's dystonia: evidence from ERP-based studies. *Ann. N.Y. Acad. Sci.* 1252, 192–199. doi: 10.1111/j.1749-6632.2011.06417.x

Taylor, S. F., Stern, E. R., and Gehring, W. J. (2007). Neural systems for error monitoring: recent findings and theoretical perspectives. *Neuroscientist* 13, 160–172. doi: 10.1177/1073858406298184

van Veen, V., and Carter, C. (2006). Error detection, correction, and prevention in the brain: a brief review of data and theories. *Clin. EEG Neurosci.* 37, 330–335. doi: 10.1177/155005940603700411

van Veen, V., Cohen, J. D., Botvinick, M. M., Stenger, V. A., and Carter, C. S. (2001). Anterior cingulate cortex, conflict monitoring, and levels of processing. *Neuroimage* 14, 1302–1308. doi: 10.1006/nimg.2001.0923

Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882. doi: 10.1126/science.7569931

Yeung, N., Botvinick, M. M., and Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychol. Rev.* 111, 931–959. doi: 10.1037/0033-295X.111.4.931

Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8, 547–558. doi: 10.1038/nrn2152

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2013; accepted: 30 August 2013; published online: 26 September 2013.

Citation: Maidhof C, Pitkäniemi A and Tervaniemi M (2013) Predictive error detection in pianists: a combined ERP and motion capture study. Front. Hum. Neurosci. 7:587. doi: 10.3389/fnhum.2013.00587

This article was submitted to the journal Frontiers in Human Neuroscience.

4.4. Combining EEG, MIDI, and Motion Capture Techniques for Investigating Musical Performance

Maidhof, C., Kästner, T., & Makkonen, T. (2014). Combining EEG, MIDI, and motion capture techniques for investigating musical performance. *Behavior Research Methods*, *46*(1), 185–95. doi:10.3758/s13428-013-0363-9

Behav Res (2014) 46:185–195
DOI 10.3758/s13428-013-0363-9

Combining EEG, MIDI, and motion capture techniques for investigating musical performance

Clemens Maidhof · Torsten Kästner · Tommi Makkonen

Abstract This article describes a setup for the simultaneous recording of electrophysiological data (EEG), musical data (MIDI), and three-dimensional movement data. Previously, each of these three different kinds of measurements, conducted sequentially, has been proven to provide important information about different aspects of music performance as an example of a demanding multisensory motor skill. With the method described here, it is possible to record brain-related activity and movement data simultaneously, with accurate timing resolution and at relatively low costs. EEG and MIDI data were synchronized with a modified version of the FTAP software, sending synchronization signals to the EEG recording device simultaneously with keypress events. Similarly, a motion capture system sent synchronization signals simultaneously with each recorded frame. The setup can be used for studies investigating cognitive and motor processes during music performance and music-like tasks—for example, in the domains of motor control, learning, music therapy, or musical emotions. Thus, this setup offers a promising possibility of a more behaviorally driven analysis of brain activity.

Keywords EEG · Motion capture · MIDI · Music performance · Simultaneous data acquisition

C. Maidhof (✉) · T. Makkonen
Cognitive Brain Research Unit, Cognitive Science, Institute of Behavioural Sciences, University of Helsinki,
Siltavuorenpenger 1 B, 00017 Helsinki, Finland
e-mail: clemens.maidhof@gmail.com

C. Maidhof · T. Makkonen
Finnish Centre of Excellence in Interdisciplinary Music Research, University of Jyväskylä, Jyväskylä, Finland

T. Kästner
Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

Performing music is a remarkable human activity, providing rich information about different motor and cognitive processes and the interactions between perception and action. Motions during music performance can be regarded as being shaped by psychological processes and task demands, which can include anatomical and biomechanical adjustments of the body and sensorimotor adaptations to increase sensory information, ultimately to improve performance (for a summary, see Palmer, 2013). In addition, musicians can use body movements to convey their expressive intentions or emotions, as well as to highlight structural aspects of the music (see, e.g., Thompson & Luck, 2012).

Quantifying music behaviors has become considerably easier by using generally available MIDI (musical instrument digital interface) equipment. MIDI instruments—for example, digital pianos—allow for accurate recording of the on- and offsets of keypresses, in addition to the so-called "MIDI velocity," which corresponds to the force of a keypress.[1] Previous studies using behavioral measures of timing, such as the interonset intervals (IOIs) of keypresses, have focused, for example, on the planning of musical sequences (for a review, see Palmer, 1997), the role of auditory feedback in music performance (for reviews, see Pfordresher, 2006, 2012), and skill learning (Jabusch, Alpers, Kopiez, Vauth, & Altenmüller, 2009). In addition, timing information has been used to quantify the effects of focal dystonia in musicians (Jabusch, Vauth, & Altenmüller, 2004).

To record and/or manipulate MIDI data, a number of different studies have used the FTAP software (Finney, 2001a, 2001b), a freely available program for tapping and music experiments, published under the GNU public license. This software runs on standard PC hardware under a Linux operating system, has reliable millisecond-timing abilities, and is especially flexible in the area of auditory feedback manipulation (for

---

[1] However, there is no standard unit for keypress velocity, and MIDI velocity values can differ between different keyboards.

studies using FTAP, see, e.g., Finney & Palmer, 2003; Fischinger, 2011; Goebl & Palmer, 2008; Loehr & Palmer, 2011; Pfordresher & Benitez, 2007; Pfordresher & Dalla Bella, 2011; Pfordresher, Keller, Koch, Palmer, & Yildirim, 2011; Pfordresher & Palmer, 2002). We describe here an update of the FTAP software that can be used for the simultaneous recording of MIDI and electroencephalographic (EEG) data (as had previously been done in Maidhof, Rieger, Prinz, & Koelsch, 2009; Maidhof, Vavatzanidis, Prinz, Rieger, & Koelsch, 2010). It can also be used in any setup in which the recording of MIDI and additional data is required and synchronization between devices via the parallel port is necessary.

However, MIDI data are based on note events—that is, on discrete time points. Therefore, much of the information about a musical performance cannot be quantified, such as the performer's motion or the kinematics of, for example, finger movements. To circumvent this problem, researchers have directly investigated the movements during music performance or during music-like sequential tasks with motion capture techniques (often combined with MIDI or audio recordings). Motion capture systems allow for the recording and analysis of natural human movements with high spatial (up to 0.1 mm) and temporal (typically 4–8 ms) resolution. Studies employing motion capture techniques have investigated, for instance, the role of tactile and auditory feedback in timing accuracy (Goebl & Palmer, 2008; Palmer, Koopmans, Loehr, & Carter, 2009; Pfordresher & Dalla Bella, 2011), the role of anticipatory auditory imagery (Keller, Dalla Bella, & Koch, 2010), relationships between kinematics of body movement and musical performance (G. Luck & Toiviainen, 2008), emotional expressions (Livingstone, Thompson, & Russo, 2009), or ensemble coordination (e.g., Keller & Appel, 2010).

More recently, interest has also increased in identifying the neural processes involved in music performance (for a review, see Zatorre, Chen, & Penhune, 2007). Research in this domain has employed neuroimaging methods, including functional magnetic resonance imaging (fMRI) and positron-emission tomography (PET), often in combination with digital musical instruments or instrument-like devices (Bengtsson, Csíkszentmihályi, & Ullén, 2007; Berkowitz & Ansari, 2008; Chen, Penhune, & Zatorre, 2008; Limb & Braun, 2008; Lotze et al., 1999; Meister et al., 2004; Parsons, Sergent, Hodges, & Fox, 2005). In these studies, participants are typically required to perform on these devices while lying inside the scanner without moving their head. However, another neuroimaging method, called "near-infrared spectroscopy" (NIRS), might prove useful for music performance studies in the future. This method measures changes in the concentration of oxygenated and deoxygenated hemoglobin in blood that occur within several seconds after increased brain activity. NIRS has the advantage that measurements could be performed under more natural conditions, including music performance.[2] However, head movements can apparently pose problems during NIRS recordings (see, e.g., Izzetoglu, Chitrapu, Bunce, & Onaral, 2010).

Because music and its production are phenomena unfolding over time, methods with a high temporal resolution like EEG (or magnetoencephalography, MEG) are particularly suited for music performance research. In addition, the EEG method requires participants only to wear an EEG cap, and experiments can thus be conducted in more ecologically valid situations; for example, participants can sit in front of a keyboard.

Unfortunately, movement-related artifacts pose a serious problem for EEG recordings. The electromyographic (EMG) activity of mainly the head, neck, and shoulder muscles can be picked up by the EEG sensors (movements below the neck are usually not that important, because the EMG does not propagate very far), and the recorded signal can thus be a mixture of brain-related signals (electrical brain activity) and several classes of non-brain-related artifacts, such as the electrical activity of muscles, eye movements, cardiac artifacts, or line noise. However, several approaches allow researchers to deal with artifacts, such as averaging (as is used in the event-related potential technique), filtering (to suppress EMG artifacts, which primarily consist of high-frequency components above 100 Hz; S. J. Luck, 2005), or independent component analysis (ICA; Bell & Sejnowski, 1995). It has been shown that ICA can be especially effective in separating different classes of artifacts from the EEG data, such as muscle and eye movement artifacts (Jung et al., 2001). By taking additional precautions to prevent excessive head and neck movements of participants, especially in combination with applying ICA to EEG data and recording a sufficient number of trials, the problem of movement artifacts can be diminished.

The possibility to record EEG during musical movements and behaviors, at least during rapid finger movements, has also been demonstrated by recent EEG studies on piano performance. These studies have investigated, for example, the neural correlates of feedback processing (and of unexpected feedback alterations) during the performance of scales, fingering exercises, and simple melodies (Katahira, Abla, Masuda, & Okanoya, 2008; Maidhof, Vavatzanidis, Prinz, Rieger, & Koelsch, 2010); error monitoring during piano performance in highly trained musicians (Herrojo Ruiz, Jabusch, & Altenmüller, 2009; Maidhof et al., 2009) and in musicians with focal-task-specific dystonia (Herrojo Ruiz, Strübing, Jabusch, & Altenmüller, 2011; Strübing, Herrojo Ruiz, Jabusch, & Altenmüller, 2012); and the neural

---

[2] Note that the modified FTAP version could in principle also be used with the NIRS technology.

correlates of ensemble performances (Babiloni et al., 2012; Babiloni et al., 2011; Lindenberger et al., 2009).

However, most of these studies (Herrojo Ruiz et al., 2009; Herrojo Ruiz et al., 2011; Katahira et al., 2008; Lindenberger et al., 2009; Maidhof et al., 2009; Maidhof et al., 2010; Strübing et al., 2012) could not provide detailed behavioral information, except MIDI-based (or audio-based) information, such as tone on- and offsets and estimates of keypress velocities. Therefore, detailed information about the motions underlying music performance was lacking, and it was not possible to relate neurophysiological findings to specific movement phases, like preparatory movements or the beginning of tactile feedback. In contrast, the combined recording of electrophysiological and movement data could lead to a better understanding of ongoing brain processes during music performance—for example, by investigating the neural correlates with respect to ongoing motor actions. A combination of different measurements could thus lead to a more realistic account of the cognitive and brain processes underlying music performance, and to more behaviorally informed brain research in general. Considering the notion that cortical brain functioning evolved to organize our interactions with the environment, such an integrated approach of music performance would be desirable. Recording MIDI, movements, and EEG at the same time also has considerably methodological benefits: Even if it were possible to obtain the same data in separate recording sessions from the same participants, the simultaneous recording would decrease the error unavoidably caused by repeated testing—for example, by minimizing measurement error and preventing effects caused by learning and fatigue. An alternative way to minimize the learning and fatigue effects would be to recruit different participants in different/subsequent recordings. As compared with such traditional setups, in which measurements are taken separately and the same performance thus has to be repeated by the same or different performers, data acquisition at multiple levels for the same music performance at once would considerably reduce intra- and interindividual variability. In the following sections, we describe a setup that combines the recording of MIDI and EEG data with the simultaneous recording of 3-D movement data with the use of a motion capture system.

Introduction of the setup

The basic concept of simultaneously recording of EEG, MIDI, and motion capture data is depicted in Fig. 1. The data collected from different sources are stored on separate computers (as when recorded separately). To synchronize and combine the different data streams, the FTAP software and the motion capture system each send synchronization signals in the form of transistor–transistor logic (TTL) pulses to the EEG acquisition device. With this kind of setup, the EEG recording has to be started first in order to record all TTL pulses sent from the motion capture system (or FTAP), which then can be used, for example, to mark the start of a trial. By doing so, the information about the timing of each captured frame and the information about the timing of a MIDI event are included in the EEG data. Thus, the electrophysiological data can be analyzed with respect to events in the movement data (and the movement data can be analyzed with respect to the EEG data).

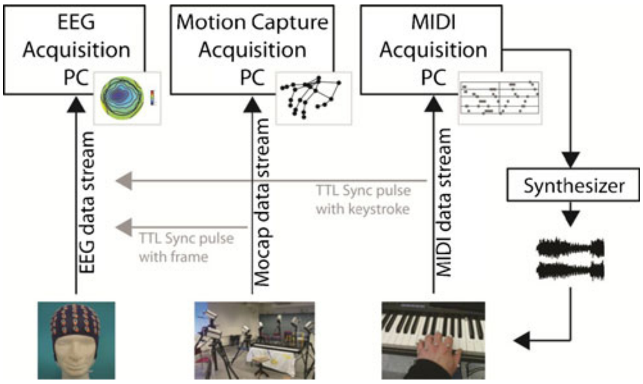

Fig. 1 Schematic illustration of the setup combining MIDI, EEG, and motion capture recordings

Hard- and software specifications

MIDI

A Yamaha KX 88 digital piano (Yamaha Corp., Japan) was connected via the MIDI port to a PC (Intel 4 CPU with 3 GHz and 500 MB RAM) running the modified version of FTAP (for details, see below). The soundcard was a SB Audigy 2 Platinum (Creative Technology Ltd., Singapore). An Ubuntu Linux operation system was installed (Version 9.04), with Kernel Version 2.6.31.4-rt14-rev01 and the preemption patch (https://rt.wiki.kernel.org). We used the freely available ALSA drivers ("advanced Linux sound architecture," Version 1.0.21.). The soundcard transmitted the MIDI messages to a Roland XV-2020 (Roland Corp., Japan) synthesizer module for audio output.

Motion capture

We used a ProReflex passive infrared-light motion capture system with a fixed sampling rate of 120 Hz (Qualisys, Sweden). The system can be synchronized with external devices by sending a TTL pulse with a length of 0.4 ms, coinciding with each exposure initialization. The movement data were recorded with a PC (specifications as recommended by the manufacturer) running the Qualisys Track Manager software (Version 2.6).

EEG

The EEG device was a BioSemi ActiveTwo Mk2 (BioSemi B.V., The Netherlands) AD box (sampling rate: 8192 Hz), sending optical data to a Biosemi USB2 receiver box. For trigger input, this USB2 box receives data from a 37-pin Sub-D connector (16 bits for trigger input). The EEG data were recorded on a laptop (specifications as recommended by BioSemi) with the ActiView EEG acquisition software from BioSemi.

Synchronization of EEG, MIDI, and motion capture systems

A custom-made 37-pin Sub-D connector received TTL pulses from both the motion capture system and the parallel port controlled by FTAP. Because FTAP connects to different pins of the connector than does the motion capture system, separation of the two different inputs is based on the value of the arriving signal at the connector (see supplementary materials available at www.cbru.helsinki.fi//music/MocapEEG/supplement.html and www.cbru.helsinki.fi/music/MocapEEG/ for the pin layout). When connected to the trigger port of the USB2 receiver box, digital data from the AD box (i.e., the neurophysiological and/or other physiological data, such as electromyography data) and TTL pulses are mixed, and the onsets of the TTL pulses are stored as triggers in the EEG data (in an extra channel, with the same sample rate as the electrode channels).

Modifications to FTAP

Synchronization with external devices

In general, FTAP reads MIDI messages from a connected MIDI instrument, processes these data, and sends out MIDI messages to a tone generator. FTAP is especially flexible in changing its behavior during the course of a trial: On the basis of keystroke number, elapsed time, or metronome count, the specific parameters controlling, for instance, auditory feedback can be modified. The parameters are defined in a simple ASCII text configuration file that controls an FTAP experiment. The values of these parameters can be changed by defining so-called "trigger events" in the configuration file, thereby—for instance—muting the auditory feedback of keypresses or delaying it for a defined amount of time (for more details, see Finney, 2001a, 2001b).

The most important modification to FTAP enabled it to communicate with another device or computer via the parallel port. For that, new trigger events were implemented that can activate specific pins of the parallel port for a specific time.[3] Thereby, on the basis of keystroke number, elapsed time, or metronome beat, an output at the parallel port with values ranging from 1 to 128 can be defined. By using the parallel port as an interface to an external device, it is possible to synchronize information from FTAP with, for instance, an EEG recording device (see the Synchronization of EEG, MIDI, and Motion Capture Systems section above for more details). This feature was implemented by using the "parapin" library for Linux PCs (http://parapin.sourceforge.net/), which provides a set of C functions for controlling pins of the parallel port (although other possibilities to control the parallel port do exist—e.g., the "Universal Parallel Port I/O library," or directly using the input/output control with the Linux kernel parport library). For C code examples, see the supplementary materials available at www.cbru.helsinki.fi/music/MocapEEG/supplement.html.

The user only has to specify in the configuration file when and how these new functions will be called. In general, the syntax for evoking trigger events in a configuration file is

```
["TRIGGER"] [Trigger ID] [Trigger type]
[Trigger count] [Parameter name] [Parameter value]
```

In the example configuration file below, the function activating Pin 2 of the parallel port is called 100 ms after trial start (indicated by "T 100"), and then again when the fifth key is pressed on a keyboard (or when FTAP reads the fifth "MIDI on" message, indicated by "K 5").

```
TRIGGER     1     T     100     FUNC     1
TRIGGER     2     K     5       FUNC     1
```

Thus, it is easy to define that synchronously with each keypress, FTAP should send a signal via the parallel port to an external device. However, two keys pressed in close temporal proximity (e.g., when two keys are accidentally pressed at almost the same time) can lead to overlapping output signals at the parallel port (resulting in a pulse length of 7 ms instead of 4 ms), or in the activation of multiple pins, resulting in wrong output values.[4]

---

[3] The optimal length of the signal will depend on the device receiving these signals. For example, if the parallel port is connected to an EEG acquisition device, the signal length has to be adjusted according to the sampling rate of this device.

[4] For example, when the first and second data pins (2^0 and 2^1 in decimal numbers) are activated simultaneously, the output value will be 11_2 (in decimal system: $2^0 + 2^1 = 3$), instead of 01_2 and 10_2 successively ($2^0 = 1$ and $2^1 = 2$, respectively, in the decimal system). Note that this is a general problem for stimulation software programs communicating with other computers via the parallel port—for example, in the widely used Presentation commercial stimulation software (Neurobehavioral Systems, Inc., Albany, CA).

To circumvent this problem, it is easy to define in FTAP's configuration file that, for example, when participants produce melodies, only every fifth keystroke will activate the parallel port, reducing the risk of overlapping digital outputs. However, this problem can still occur if more than five keys are pressed in close temporal proximity. The Presentation commercial stimulation software offers the option to delay the second pulse as long as the parallel port is still activated by a previous pulse, and the user is informed whenever this situation occurs. It is conceivable that a similar mechanism could be implemented in FTAP in the future. Another possibility could be to send regular synchronization pulses from FTAP: For example, the user could define in a parameter file that FTAP should send a synchronization pulse every 100 ms, regardless of any other MIDI events (note, however, that this possibility has not been tested).

In addition, a function was implemented that increases the output value of the parallel port each time that this trigger event is called, by a power of 2 (ranging from 1 to 128; see the supplementary materials at www.cbru.helsinki.fi/music/MocapEEG/supplement.html for the C code). Thus, the receiving device can readily be tested for whether all signals were correctly sent and received. In addition to that function, nine new functions were implemented, each sending an output value of 1, 2, 3, 4, 8, 16, 32, 64, or 128 via the parallel port.

Auditory feedback manipulations

Furthermore, FTAP's capabilities for manipulating MIDI streams were extended. Besides the originally implemented manipulations (Finney, 2001a, b), FTAP can now lower the pitch of the auditory feedback by one semitone. For that, the user has to set the value of the FEED_PMODE parameter to 50 (see the supplementary materials at www.cbru.helsinki.fi/music/MocapEEG/supplement.html for further details). The modified source code of FTAP is available for download at www.cbru.helsinki.fi/music/MocapEEG.

Testing

To test whether the setup combining motion capture, EEG, and MIDI data worked reliably and with sufficient temporal precision, several tests were conducted.

Delays introduced by FTAP

First, we tested the delays that were introduced by using MIDI and FTAP: For that, the onsets of the audio signals generated by the digital piano were compared with the onsets of the audio signals generated by the synthesizer (receiving MIDI input from FTAP).[5] Because the KX 88 has no internal sound generator, a Clavinova CLP-330 (Yamaha) was used for this test. Any differences in the onsets would show how the MIDI setup (i.e., FTAP recording the MIDI input, processing it, and sending it to the synthesizer module; see Fig. 1) influenced the timing of the audio output, as compared with when the audio output was directly generated by the digital piano. Note that any potential internal delays in the digital piano (from keystroke to sound generation, possibly contributing to the measured difference) were not tested, because we were interested in delays that are important from the viewpoint of a participant. To make the onset of the audio signal better identifiable, a drum sound with a high attack was used as a preset of the synthesizer module. Audio signals were recorded from the line out of the digital piano (left channel) and the line out of the synthesizer module (right channel) as a stereo file with a 44100-Hz sampling rate and 16-bit resolution.[6] This test was repeated 90 times (i.e., audio signals generated by 91 keypresses, randomly performed within about 2 min, were recorded).

Parallel port output

Second, we tested the implemented feature of FTAP to send signals via the parallel port (synchronously with MIDI events/keypresses) that can be used to synchronize external devices. For that test, we recorded the audio signals generated by the synthesizer module, as well as the TTL pulses from one data pin directly from the parallel port, as a stereo .wav file (44100 Hz, 16-bit resolution) using a custom-made stereo mini-jack connector.[7] Differences in the onsets would indicate how accurate the timing of the parallel port output was with respect to the audio signal. Audio signals were elicited either by manual keypresses on the digital piano, or by defining a pacing signal (i.e., metronome) in FTAP's configuration file. In addition, we defined that synchronously with each keypress or pacing signal, a trigger event would be called that would activate the first data pin of the parallel port for 4 ms. Several scenarios were tested: 100 keypresses, randomly performed within (1) about 5 min, (2) about 9 min, (3) 200 metronome beats with an IOI of 100 ms, and (4) 200 metronome beats with an IOI of 5,000 ms.

---

[5] For all analyses involving audio data, we used Adobe Audition 3.0 (Adobe Systems Corp., California, USA).

[6] For audio and TTL pulse recordings, we used a MacPro computer with the Logic Pro software (Apple Inc., USA).

[7] The left channel was connected to the line out of the synthesizer, and the right channel was a wire soldered to the first data pin of an LPT connector, which was attached to the parallel port of the computer running FTAP (both channels shared the ground).

Reliability

To check the reliability of the synchronization pulses from the motion capture system and from FTAP, we verified that the number of synchronization pulses received by the EEG device matched the number of sent synchronization pulses. In addition, we computed the distances between two successive synchronization pulses, as recorded by the EEG device, and compared these to the distances between the two corresponding pulses, as documented by the sending devices (motion capture system and FTAP, respectively). Because the available motion capture system had a fixed sampling rate of 120 Hz (and because the rate of synchronization pulses was coupled to the sampling rate in that system), we were not able to specifically test the reliability or accuracy with any higher sampling or synchronization rate.

Pilot experiment

Finally, we recorded an EEG data set with the setup described above with one participant. The task was simply to perform keystrokes with the thumb and pinkie finger of the right hand in alternating order every 2–3 s. The EEG was recorded with 64 scalp electrodes placed according to the extended 10–20 system, in addition of the movement data of two markers attached to the thumb and little finger of the participant. For more details, see the supplementary materials (www.cbru.helsinki.fi/music/MocapEEG/supplement.html).

Results

Delays introduced by FTAP

Figure 2a shows an example of the audio output from the digital piano as compared with the audio output from the synthesizer module, both generated by one keypress on the digital piano. Manual analysis of the tone onsets revealed that on average, the audio onset produced by the synthesizer module was delayed by 1.4 ms (*SD* = 0.58 ms) relative to the onset of the audio output directly from the digital piano. Importantly, the maximal delay was 3 ms (which occurred two out of 91 times).

Parallel port output

Figure 2b shows an example of the recorded audio signal and the TTL pulse sent by FTAP synchronously with a received MIDI event (from the digital piano). Analyses showed that the TTL pulse had a duration of about 4 ms (as programmed), and more importantly, that the audio signal started about 3 ms after the onset of the TTL pulse. This result was found for all of the different scenarios—that is, for 100 keypresses randomly performed within about 5 and 9 min, for 200 metronome beats with an IOI of 100 ms (i.e., test duration of about 25 s), and for 200 metronome beats with an IOI of 5,000 ms (i.e., test duration of about 17 min). Note that the 3-ms difference in the onsets of the TTL pulses and the audio signals did not change over time, even when the recording lasted about 17 min (as compared to 25 s). Similarly, the difference in onsets did not change with different numbers of MIDI signals (100 keypresses vs. 200 metronome beats). Also note that the TTL pulse seemed to be generated before the audio signal of the digital piano was generated. This could be due to the digital-to-analog conversion in the piano, which, however, might differ between keyboards.

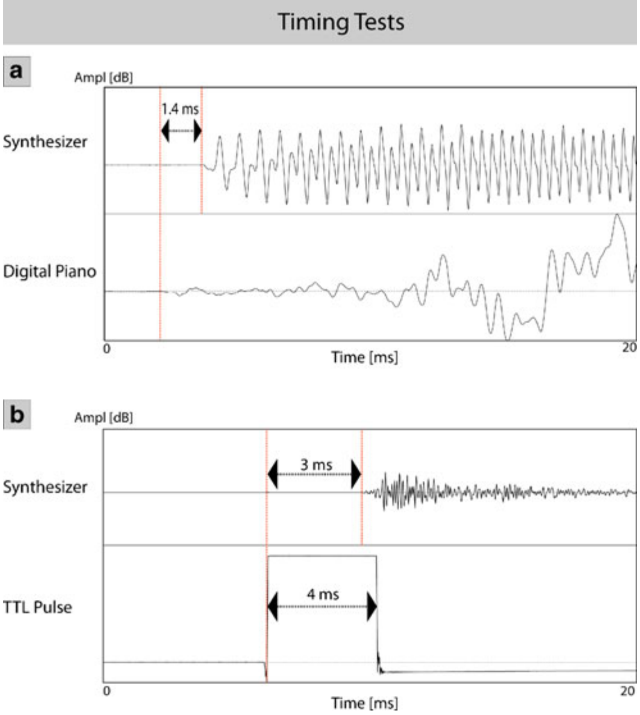

Fig. 2 Illustration of the performed timing tests. (**a**) Results showed that, on average, the onset of the audio output from the synthesizer was delayed by 1.4 ms relative to the onset of the audio output from the digital piano. (**b**) Results showed that the audio output from the synthesizer started around 3 ms after the onset of the TTL pulses, which showed durations of 4 ms

Reliability

All synchronization pulses sent by the motion capture system and FTAP, respectively, were correctly received by the EEG device. Inspection of the distances between two successive pulses recorded in the EEG data showed that (a) the distances between two pulses from the motion capture system as recorded by the EEG and as recorded by the motion capture system itself differed on average less than 1 ms (ranging from 0.09 to –0.04 ms), and (b) the distances between two pulses from FTAP as recorded by the EEG and as logged by FTAP itself differed on average less than 1 ms (ranging from about 1 to –1.5 ms; see Fig. 3 for the distributions).

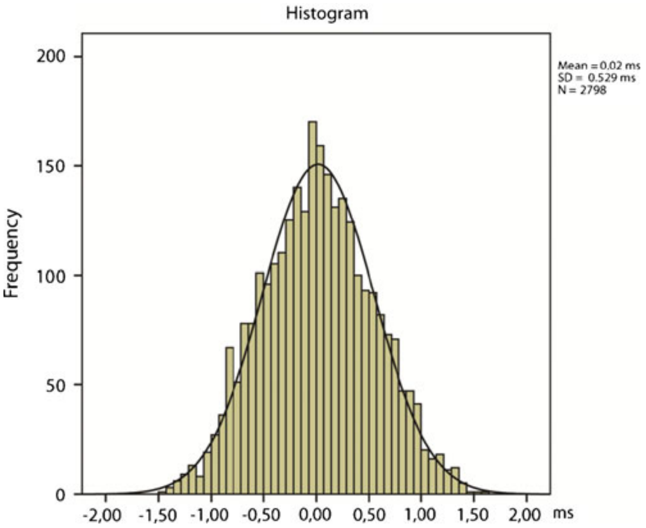

Fig. 3 Distribution of the differences between two successive FTAP trigger signals, as recorded by the EEG device and by FTAP, respectively

Pilot experiment

Figure 4a shows the event-related potentials (ERPs) and their scalp distributions for keypresses performed with the thumb and the pinkie finger on electrode FC1. The ERPs showed an increasing negativity starting around 100 ms prior to and peaking around the time of keystroke onset. This deflection highly resembles the late stage of the readiness potential, indicating processes of motor preparation and/or execution during voluntary movements (for a review, see Shibasaki & Hallett, 2006). The scalp distributions of both potentials showed a left-lateralized central peak, consistent with neural generators in motor areas of the brain. After keypress (and thus sound) onset, an N100 and P200 are elicited by the feedback tones of the keypress, both reflecting different stages of auditory processing (for a review, see, e.g., Näätänen & Picton, 1987). The lower part of Fig. 4 shows the position, velocity, and acceleration data (in the z-axis) of the markers attached to the fingernail of the thumb and pinkie finger. Both fingers reached a minimum in height at the keypress onset, which is also characterized by a peak in acceleration (when the fingers move downward and stop when the keybed is reached).

Discussion

Music performance research provides a rich source of information about various motor, cognitive, and emotional processes. The presented setup for the simultaneous recording of MIDI, EEG, and three-dimensional movement data can be used to combine the measurement of different aspects of music performance that have previously not been studied in combination.

One of the major advantages of recording EEG, MIDI, or movement data with motion capture techniques is that data can be measured with relatively high temporal resolution, ranging from around 1 ms or less (EEG, MIDI) to several milliseconds (8.3 ms for motion capture data recorded with the present setup; note that higher resolutions are possible with different systems). A major challenge for any setup combining these different measures is that the combination does not introduce any additional timing delays, which could lessen the advantage of the temporal resolutions of the measures used. The results of the timing tests confirmed that the setup described above still shows a sufficiently good temporal precision for most purposes (the largest timing inaccuracy, in terms of additional delays, that we found in all of the tests was 3 ms, most likely due partly to delays during the digital-to-analog conversion in the synthesizer; see the Method section for details), and that the combination works reliably. However, because each component of the setup could alter the results (e.g., the timing results that we obtained might be keyboard-dependent), careful testing of any setup is highly recommended prior to any data collection, to ensure that timing accuracy and reliability are as high as needed.

For the synchronization of EEG and MIDI data, we used a modified version of FTAP, a freely available program for the implementation of tapping and music performance experiments (Finney, 2001a, 2001b). Modifications to FTAP included the new functionality of using the parallel port as an interface to other devices. This feature of FTAP offers a simple, low-cost, but flexible solution to synchronize data from digital instruments (MIDI-based) with EEG data for neuroscientific music performance experiments. Importantly, the modified version of FTAP is also useful for synchronizing MIDI recordings with other devices—for example, with fMRI, MEG, NIRS, or any other device capable of receiving TTL pulses.

However, several issues should be taken into account when using this setup.

Timing

Although the temporal accuracy of the setup in terms of additional delays should be sufficient for many experiments, it is not in the preferable submillisecond range. However, it should be noted that the hardware of the Linux computer running FTAP was not optimized in terms of efficiency, except for the soundcard. Thus, it awaits further testing whether our timing results could be improved with a more optimized hardware setup—for example, in terms of increased CPU speed and increased system clock frequency (bus speed).

Time-stamping of events

In the experimental setup described here, we used three computers running different software: (1) FTAP software, (2)

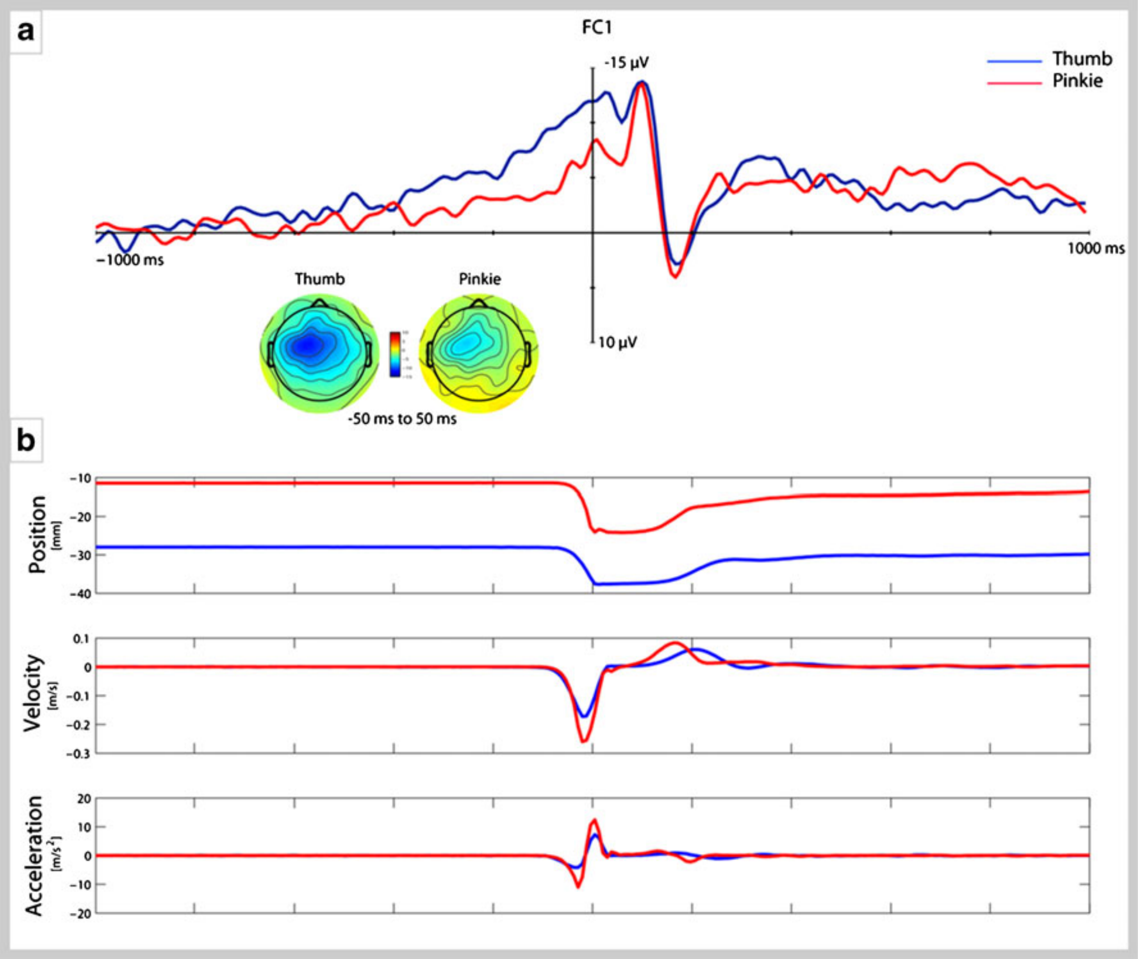

Fig. 4 Event-related potentials of thumb and pinkie-finger keystrokes, with their scalp distributions and the simultaneously recorded position, velocity, and acceleration of both fingers. Time point 0 corresponds to the synchronization pulse sent from FTAP synchronously with keypresses

Qualisys Track Manager software (motion capture), and (3) ActiView EEG acquisition software. To have reliable timing information about MIDI events and recorded frames in the EEG data, the hardware has to be synchronized in time. This was realized by sending trigger signals from FTAP and the motion capture system to the EEG device at the beginning of each recording. However, the clocks of different computers do not run in synchrony over time (i.e., there is a slight time drift in each computer clock), and it is thus necessary to resynchronize during a recording. Millisecond accuracy was achieved by sending triggers to the EEG device with a sufficient frequency: For the motion capture system, this was done automatically, because it sent a synchronization pulse concurrently with each recorded frame[8]; for FTAP, resynchronization was done by sending a synchronization pulse concurrently with every fifth keystroke. We confirmed that the difference between the computer clocks never exceeded 1.5 ms (<1 ms on average) within those time intervals of five consecutive keystrokes. Thus, during offline data analysis, it was possible to calculate the relative latencies of events of interest within those time windows in which the synchrony was not remarkably disrupted. These events could then be added easily in the EEG data, by using the timing information of all MIDI events and synchronization pulses stored in FTAP's result file. To achieve optimal results, we recommend adjusting the rate of resynchronization signals according to the experimental stimuli used (i.e., according to the rate of the produced MIDI events). Another possibility could be to send regular synchronization pulses from FTAP (e.g., every 100 ms), independently from other MIDI events. This procedure would also help to avoid the problem of overlapping signals at the parallel port. Additionally, using central processing units of the same kind in both the FTAP and EEG acquisition computers could help maximize the synchronicity of different computer clocks.

---

[8] Note that the frequency of the synchronization pulse is adjustable with other motion capture systems—for example, with the Qualisys Oqus system or with Vicon motion capture systems (Vicon Motion Systems Limited, UK).

Number of events stored within EEG data

Because each frame recorded by the motion capture system (circa every 8.3 ms) is marked in the EEG data, the number of events stored within the EEG data is large, which might cause problems for some analysis software during offline processing of the data. A convenient way to circumvent this problem would be, offline, to store only a subset of events representing recorded frames in the EEG data—for example, the most important frames surrounding events representing MIDI keypresses, or even only the first recorded synchronization signal from the motion capture system. If the synchronization pulse sent by the motion capture system is very regular (as it was in this case; the deviation of the synchronization signal from the motion capture system, recorded with the EEG device, was maximally 122 ns, corresponding to one sample point with the given sampling frequency of 8192 Hz), it is possible to add new EEG events relative to the first or a few existing EEG events during offline processing.

EEG data and artifacts

A major challenge of EEG experiments investigating music performance (and of any experiments requiring more movements on the part of participants than simple buttonpresses) is that the signal recorded by the EEG sensors is a mixture of brain-related activity and several classes of non-brain-related artifacts originating, for instance, from eye movements, head and neck muscle movements, electrocardiographic activity, and line noise. However, several approaches allow researchers to minimize the effects of movement artifacts, such as averaging (used in ERP studies), filtering, and applying ICA as a method of identifying and removing several classes of artifacts from EEG data. These approaches, in addition to preventing excessive head and neck movements of the participants, have already been used successfully in studies (Herrojo Ruiz et al., 2009; Herrojo Ruiz et al., 2011; Maidhof et al., 2009; Maidhof et al., 2010; Strübing et al., 2012). However, limiting movement during musical performances might lead to strategic behavioral adjustments that are not typical for music performance. Thus, there is currently a trade-off between motions that are representative of musical performance and motions that do not result in large movement artifacts in the EEG data. Interestingly, ICA as a signal-processing approach for EEG data is under active development and has been applied to data recorded during running and walking (Gramann, Gwin, Bigdely-Shamlo, Ferris, & Makeig, 2010; Gwin, Gramann, Makeig, & Ferris, 2010). Therefore, it would be interesting to extend these new approaches to music performance research.

Different motion capture sample rates and MIDI instruments

The present setup was tested with a digital piano, and movements were recorded with a relatively low sample rate of 120 Hz. However, the recording of fast movements during piano playing or during fast performances on other MIDI instruments, such as drums, might require a higher sampling to capture all relevant aspects of the executed movements. Although it awaits further testing, the setup should in principle also work if higher sampling rates were used. As was already mentioned, if a synchronization pulse from the motion capture system were to be sent with each recorded frame, the number of events in the EEG data might cause problems during offline analyses. Therefore, we recommend lowering the rate of synchronization signals during recording, which should be possible with other motion capture systems, such as the Oqus system from Qualisys or the Giganet system from Vicon (see also the Number of Events Stored Within EEG Data section above). The setup is assumed to also work with other MIDI instruments, as long as FTAP is able to correctly receive and process the information from that MIDI instrument. However, different MIDI instruments might require larger movements, especially of the neck, shoulder, and head, potentially resulting in larger EEG movement artifacts. In addition, the rate of FTAP synchronization pulses should be adjusted accordingly (see also the Time-Stamping of Events section above).

Conclusions

The present setup, with a relatively simple technical implementation, can be regarded as one step towards a more behaviorally driven analysis of brain activity. Applications of this setup are for example in the domain of motor control and error monitoring during piano performance (Herrojo Ruiz et al., 2009; Herrojo Ruiz et al., 2011; Maidhof et al., 2009; Strübing et al., 2012). By using kinematic landmarks of finger movements that reflect different movement stages (e.g. the onset of tactile feedback of keystrokes; Goebl & Palmer, 2008), it is possible to analyze event-related brain activity with respect to such motor actions. Other applications are investigations into musical skill learning and associated brain activity changes, music therapy, musical expressivity, and interactions between perception and action in general. Further research and development of new analysis methodologies like the application of ICA for EEG data, possibly with the development of new sensor technologies, could pave the way for studies investigating brain activity of natural musical behaviors without those limitations described above (see also Makeig, Gramann, Jung, Sejnowski, & Poizner, 2009).

Author note C.M. is now at the Institute for Systematic Musicology at the University of Hamburg. The study was financially supported by the Academy of Finland and the Center for International Mobility (CIMO), Finland. We are especially thankful to Birgitta Burger, Tommi Himberg, and Miika Leminen for their help with the motion capture system. We also thank Peter Pfordresher and Mari Tervaniemi for helpful comments on an earlier draft of this article.

References

Babiloni, C., Buffo, P., Vecchio, F., Marzano, N., Del Percio, C., Spada, D., & Perani, D. (2012). Brains “in concert”: Frontal oscillatory alpha rhythms and empathy in professional musicians. *NeuroImage, 60,* 105–116. doi:10.1016/j.neuroimage.2011.12.008

Babiloni, C., Vecchio, F., Infarinato, F., Buffo, P., Marzano, N., Spada, D., & Perani, D. (2011). Simultaneous recording of electroencephalographic data in musicians playing in ensemble. *Cortex, 47,* 1082–1090. doi:10.1016/j.cortex.2011.05.006

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation, 7,* 1129–1159.

Bengtsson, S. L., Csíkszentmihályi, M., & Ullén, F. (2007). Cortical regions involved in the generation of musical structures during improvisation in pianists. *Journal of Cognitive Neuroscience, 19,* 830–842. doi:10.1162/jocn.2007.19.5.830

Berkowitz, A. L., & Ansari, D. (2008). Generation of novel motor sequences: the neural correlates of musical improvisation. *NeuroImage, 41,* 535–543. doi:10.1016/j.neuroimage.2008.02.028

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Moving on time: brain network for auditory–motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience, 20,* 226–239. doi:10.1162/jocn.2008.20018

Finney, S. A. (2001a). FTAP: A Linux-based program for tapping and music experiments. *Behavior Research Methods, Instruments, & Computers, 33,* 65–72.

Finney, S. A. (2001b). Real-time data collection in Linux: A case study. *Behavior Research Methods, Instruments, & Computers, 33,* 167–173.

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

Fischinger, T. (2011). An integrative dual-route model of rhythm perception and production. *Musicae Scientiae, 15,* 97–105. doi:10.1177/1029864910393330

Goebl, W., & Palmer, C. (2008). Tactile feedback and timing accuracy in piano performance. *Experimental Brain Research, 186,* 471–479. doi:10.1007/s00221-007-1252-1

Gramann, K., Gwin, J. T., Bigdely-Shamlo, N., Ferris, D. P., & Makeig, S. (2010). Visual evoked responses during standing and walking. *Frontiers in Human Neuroscience, 4,* 202. doi:10.3389/fnhum.2010.00202

Gwin, J. T., Gramann, K., Makeig, S., & Ferris, D. P. (2010). Removal of movement artifact from high-density EEG recorded during walking and running. *Journal of Neurophysiology, 103,* 3526–3534. doi:10.1152/jn.00105.2010

Herrojo Ruiz, M., Jabusch, H.-C., & Altenmüller, E. (2009). Detecting wrong notes in advance: neuronal correlates of error monitoring in pianists. *Cerebral Cortex, 19,* 2625–2639. doi:10.1093/cercor/bhp021

Herrojo Ruiz, M., Strübing, F., Jabusch, H.-C., & Altenmüller, E. (2011). EEG oscillatory patterns are associated with error prediction during music performance and are altered in musician’s dystonia. *NeuroImage, 55,* 1791–1803. doi:10.1016/j.neuroimage.2010.12.050

Izzetoglu, M., Chitrapu, P., Bunce, S., & Onaral, B. (2010). Motion artifact cancellation in NIR spectroscopy using discrete Kalman filtering. *Biomedical Engineering Online, 9,* 16. doi:10.1186/1475-925X-9-16

Jabusch, H.-C., Alpers, H., Kopiez, R., Vauth, H., & Altenmüller, E. O. (2009). The influence of practice on the development of motor skills in pianists: a longitudinal study in a selected motor task. *Human Movement Science, 28,* 74–84. doi:10.1016/j.humov.2008.08.001

Jabusch, H.-C., Vauth, H., & Altenmüller, E. O. (2004). Quantification of focal dystonia in pianists using scale analysis. *Movement Disorders, 19,* 171–180. doi:10.1002/mds.10671

Jung, T.-P., Makeig, S., McKeown, M. J., Bell, A. J., Lee, T.-W., & Sejnowski, T. J. (2001). Imaging brain dynamics using independent component analysis. *Proceedings of the IEEE, 89,* 1107–1122. doi:10.1109/5.939827

Katahira, K., Abla, D., Masuda, S., & Okanoya, K. (2008). Feedback-based error monitoring processes during musical performance: an ERP study. *Neuroscience Research, 61,* 120–128. doi:10.1016/j.neures.2008.02.001

Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception, 28,* 27–46. doi:10.1525/mp.2010.28.1.27

Keller, P. E., Dalla Bella, S., & Koch, I. (2010). Auditory imagery shapes movement timing and kinematics: evidence from a musical task. *Journal of Experimental Psychology. Human Perception and Performance, 36,* 508–513. doi:10.1037/a0017604

Limb, C. J., & Braun, A. R. (2008). Neural substrates of spontaneous musical performance: An fMRI study of jazz improvisation. *PloS One, 3,* e1679. doi:10.1371/journal.pone.0001679

Lindenberger, U., Li, S.-C., Gruber, W., & Müller, V. (2009). Brains swinging in concert: Cortical phase synchronization while playing guitar. *BMC Neuroscience, 10,* 22. doi:10.1186/1471-2202-10-22

Livingstone, S. R., Thompson, W. F., & Russo, F. A. (2009). Facial expressions and emotional singing: A study of perception and production with motion capture and electromyography. *Music Perception, 26*(5), 475–488.

Loehr, J. D., & Palmer, C. (2011). Temporal coordination between performing musicians. *Quarterly Journal of Experimental Psychology, 64,* 2153–2167.

Lotze, M., Montoya, P., Erb, M., Hülsmann, E., Flor, H., Klose, U., & Grodd, W. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: An fMRI study. *Journal of Cognitive Neuroscience, 11,* 491–501. doi:10.1162/089892999563553

Luck, G., & Toiviainen, P. (2008). Exploring relationships between the kinematics of a singer’s body movement and the quality of their voice. *Journal of Interdisciplinary Music Studies, 2,* 173–186.

Luck, S. J. (2005). Filtering. In *An introduction to the event-related potential technique* (pp. 175–224). Cambridge, MA: MIT Press.

Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2009). Nobody is perfect: ERP effects prior to performance errors in musicians indicate fast monitoring processes. *PloS One, 4,* e5032. doi:10.1371/journal.pone.0005032

Maidhof, C., Vavatzanidis, N., Prinz, W., Rieger, M., & Koelsch, S. (2010). Processing expectancy violations during music performance and perception: An ERP study. *Journal of Cognitive Neuroscience, 22,* 2401–2413. doi:10.1162/jocn.2009.21332

Makeig, S., Gramann, K., Jung, T.-P., Sejnowski, T. J., & Poizner, H. (2009). Linking brain, mind and behavior. *International Journal of Psychophysiology, 73,* 95–100. doi:10.1016/j.ijpsycho.2008.11.008

Meister, I. G., Krings, T., Foltys, H., Boroojerdi, B., Müller, M. M., Töpper, R., & Thron, A. (2004). Playing piano in the mind—An fMRI study on music imagery and performance in pianists. *Cognitive Brain Research, 19,* 219–228. doi:10.1016/j.cogbrainres.2003.12.005

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24,* 375–425.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48,* 115–138. doi:10.1146/annurev.psych.48.1.115

Palmer, C. (2013). Music performance: Movement and coordination. In D. Deutsch (Ed.), *The psychology of music, 3rd ed.* (pp. 405–422). Amsterdam, The Netherlands: Elsevier.

Palmer, C., Koopmans, E., Loehr, J. D., & Carter, C. (2009). Movement-related feedback and temporal accuracy in clarinet performance. *Music Perception, 26*(5), 439–449. doi:10.1525/mp.2009.26.5.439

Parsons, L. M., Sergent, J., Hodges, D. A., & Fox, P. T. (2005). The brain basis of piano performance. *Neuropsychologia, 43,* 199–215. doi:10.1016/j.neuropsychologia.2004.11.007

Pfordresher, P. Q. (2006). Coordination of perception and action in music performance. *Advances in Cognitive Psychology, 2,* 183–198. doi:10.2478/v10053-008-0054-8

Pfordresher, P. Q. (2012). Musical training and the role of auditory feedback during performance. *Annals of the New York Academy of Sciences, 1252,* 171–178. doi:10.1111/j.1749-6632.2011.06408.x

Pfordresher, P. Q., & Benitez, B. (2007). Temporal coordination between actions and sound during sequence production. *Human Movement Science, 26,* 742–756. doi:10.1016/j.humov.2007.07.006

Pfordresher, P. Q., & Dalla Bella, S. (2011). Delayed auditory feedback and movement. *Journal of Experimental Psychology. Human Perception and Performance, 37,* 566–579. doi:10.1037/a0021487

Pfordresher, P. Q., Keller, P. E., Koch, I., Palmer, C., & Yildirim, E. (2011). Activation of learned action sequences by auditory feedback. *Psychonomic Bulletin & Review, 18,* 544–549. doi:10.3758/s13423-011-0077-x

Pfordresher, P. Q., & Palmer, C. (2002). Effects of delayed auditory feedback on timing of music performance. *Psychological Research, 66,* 71–79. doi:10.1007/s004260100075

Shibasaki, H., & Hallett, M. (2006). What is the Bereitschaftspotential? *Clinical Neurophysiology, 117,* 2341–2356. doi:10.1016/j.clinph.2006.04.025

Strübing, F., Herrojo Ruiz, M., Jabusch, H. C., & Altenmüller, E. O. (2012). Error monitoring is altered in musician's dystonia: Evidence from ERP-based studies. *Annals of the New York Academy of Sciences, 1252,* 192–199. doi:10.1111/j.1749-6632.2011.06417.x

Thompson, M. R., & Luck, G. (2012). Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae, 16,* 19–40. doi:10.1177/1029864911423457

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory–motor interactions in music perception and production. *Nature Reviews Neuroscience, 8,* 547–558. doi:10.1038/nrn2152

Bibliography

Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear and Hearing*, *16*(1), 38–51.

Allain, S., Hasbroucq, T., Burle, B., Grapperon, J., & Vidal, F. (2004). Response monitoring without sensory feedback. *Clinical Neurophysiology*, *115*(9), 2014–2020. doi:10.1016/j.clinph.2004.04.013

Aschersleben, G., Gehrke, J., & Prinz, W. (2001). Tapping with peripheral nerve block. *Experimental Brain Research*, *136*(3), 331–339. doi:10.1007/s002210000562

Baess, P., Horváth, J., Jacobsen, T., & Schröger, E. (2011). Selective suppression of self-initiated sounds in an auditory stream: An ERP study. *Psychophysiology*, *48*(9), 1276–1283. doi:10.1111/j.1469-8986.2011.01196.x

Baess, P., Widmann, A., Roye, A., Schröger, E., & Jacobsen, T. (2009). Attenuated human auditory middle latency response and evoked 40-Hz response to self-initiated sounds. *The European Journal of Neuroscience*, *29*(7), 1514–1521. doi:10.1111/j.1460-9568.2009.06683.x

Band, G. P. H., & Kok, A. (2000). Age effects on response monitoring in a mental-rotation task. *Biological Psychology*, *51*(2-3), 201–221.

Bangert, M., & Altenmüller, E. O. (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neuroscience*, *4*(26), 1–14. doi:10.1186/1471-2202-4-26

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., … Altenmüller, E. O. (2006). Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *NeuroImage*, *30*(3), 917–926. doi:10.1016/j.neuroimage.2005.10.044

Banton, L. J. (1995). The role of visual and auditory feedback during the sight-reading of music. *Psychology of Music*, *23*(3), 3–16. doi:10.1177/0305735695231001

Bastian, A. J. (2006). Learning to predict the future: the cerebellum adapts feedforward movement control. *Current Opinion in Neurobiology*, *16*(6), 645–649. doi:10.1016/j.conb.2006.08.016

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? *Annals of the New York Academy of Sciences*, *1060*, 186–188. doi:10.1196/annals.1360.038

Blakemore, S. J., Frith, C., & Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport*, *12*(9), 1879–1884.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, *8*(12), 539–546. doi:10.1016/j.tics.2004.10.003

Brattico, E., Tervaniemi, M., Näätänen, R., & Peretz, I. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain Research*, *1117*(1), 162–174. doi:10.1016/j.brainres.2006.08.023

Braver, T. S., Barch, D. M., Gray, J. R., Molfese, D. L., & Snyder, A. (2001). Anterior cingulate cortex and response conflict: effects of frequency, inhibition and errors. *Cerebral Cortex*, *11*(9), 825–836.

Brown, R. M., Zatorre, R. J., & Penhune, V. B. (2015). Expert music performance: cognitive, neural, and developmental bases. *Progress in Brain Research*, *217*, 57–86. doi:10.1016/bs.pbr.2014.11.021

Bush, G., Vogt, B. A., Holmes, J., Dale, A. M., Greve, D. N., Jenike, M. A., & Rosen, B. R. (2002). Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(1), 523–528. doi:10.1073/pnas.012470999

Carter, C. S. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, *280*(5364), 747–749. doi:10.1126/science.280.5364.747

Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive, Affective & Behavioral Neuroscience*, *7*(4), 367–379.

Cavanagh, J. F., Cohen, M. X., & Allen, J. J. B. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *Journal of Neuroscience*, *29*(1), 98–105. doi:10.1523/JNEUROSCI.4137-08.2009

Chaffin, R., & Imreh, G. (1997). "Pulling teeth and torture": musical memory and problem solving. *Thinking and Reasoning*, *3*(4), 315–336.

Chaffin, R., & Imreh, G. (2002). Practicing perfection: piano performance as expert memory. *Psychological Science*, *13*(4), 342–9.

Chaffin, R., & Logan, T. R. (2006). Practicing perfection: How concert soloists prepare for performance. *Advances in Cognitive Psychology*, *2*(2), 113–130. doi:10.2478/v10053-008-0050-z

Chen, J., Woollacott, M. H., Pologe, S., & Moore, G. P. (2008). Pitch and space maps of skilled cellists: accuracy, variability, and error correction. *Experimental Brain Research*, *188*(4), 493–503. doi:10.1007/s00221-008-1380-2

Cole, K., Gracco, V., & Abbs, J. (1984). Autogenic and nonautogenic sensorimotor actions in the control of multiarticulate hand movements. *Experimental Brain Research*, *56*, 582–585.

Colebatch, J. G. (2007). Bereitschaftspotential and movement-related potentials: origin, significance, and application in disorders of human movement. *Movement Disorders*, *22*(5), 601–610. doi:10.1002/mds.21323

Coles, M. G., Scheffers, M. K., & Holroyd, C. B. (2001). Why is there an ERN/Ne on correct trials? Response representations, stimulus-related components, and the theory of error-processing. *Biological Psychology*, *56*(3), 173–189.

Critchley, H. D., Tang, J., Glaser, D., Butterworth, B., & Dolan, R. J. (2005). Anterior cingulate activity during error and autonomic response. *NeuroImage*, *27*(4), 885–895. doi:10.1016/j.neuroimage.2005.05.047

DAusilio, A., Novembre, G., Fadiga, L., & Keller, P. E. (2015). What can music tell us about social interaction? *Trends in Cognitive Sciences*, 1–4. doi:10.1016/j.tics.2015.01.005

de Bruijn, E., Hulstijn, W., Meulenbroek, R. G. J., & Van Galen, G. P. (2003). Action monitoring in motor control: ERPs following selection and execution errors in a force production task. *Psychophysiology*, *40*(5), 786–795.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., von Cramon, D. Y., & Engel, A. K. (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *Journal of Neuroscience*, *25*(50), 11730–11737. doi:10.1523/JNEUROSCI.3286-05.2005

Desmurget, M., & Grafton, S. T. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Sciences*, *4*(11), 423–431.

Donkers, F. C., Nieuwenhuis, S., & van Boxtel, G. J. (2005). Mediofrontal negativities in the absence of responding. *Cognitive Brain Research*, *25*(3), 777–787. doi:10.1016/j.cogbrainres.2005.09.007

Donkers, F. C., & van Boxtel, G. J. (2005). Mediofrontal negativities to averted gains and losses in the slot-machine task. *Journal of Psychophysiology*, *19*(4), 256–262. doi:10.1027/0269-8803.19.4.256

Drewing, K., Hennings, M., & Aschersleben, G. (2002). The contribution of tactile reafference to temporal regularity during bimanual finger tapping. *Psychological Research*, *66*(1), 60–70. doi:10.1007/s004260100074

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005a). Action-effect coupling in pianists. *Psychological Research*, *69*(4), 233–241. doi:10.1007/s00426-004-0175-8

Drost, U. C., Rieger, M., Brass, M., Gunter, T. C., & Prinz, W. (2005b). When hearing turns into playing: Movement induction by auditory stimuli in pianists. *The Quarterly*

Journal of Experimental Psychology, *58*(8), 1376–1389. doi:10.1080/02724980443000610

Drost, U. C., Rieger, M., & Prinz, W. (2007). Instrument specificity in experienced musicians. *Quarterly Journal of Experimental Psychology*, *60*(4), 527–533. doi:10.1080/17470210601154388

Dum, R. P., & Strick, P. L. (2002). Motor areas in the frontal lobe of the primate. *Physiology & Behavior*, *77*(4-5), 677–682.

Eichele, T., Debener, S., Calhoun, V. D., Specht, K., Engel, A. K., Hugdahl, K., ... Ullsperger, M. (2008). Prediction of human errors by maladaptive changes in event-related brain networks. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(16), 6173–6178. doi:10.1073/pnas.0708965105

Endrass, T., Reuter, B., & Kathmann, N. (2007). ERP correlates of conscious error recognition: aware and unaware errors in an antisaccade task. *The European Journal of Neuroscience*, *26*(6), 1714–20. doi:10.1111/j.1460-9568.2007.05785.x

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, *102*(2), 211–245.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*(3), 363–406.

Evarts, E. V., & Granit, R. (1976). Relations of reflexes and intended movements. *Progress in Brain Research*, *44*, 1–14.

Evarts, E. V., & Tanji, J. (1976). Reflex and intended responses in motor cortex pyramidal tract neurons of monkey. *Journal of Neurophysiology*, *39*, 1069–1080.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1990). Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In C. H. M. Brunia, A. W. K. Gaillard & A. Kok (Eds.), *Psychophysiological brain research* (Vol. 1, pp. 192–195). Tilburg, The Netherlands: Tilburg University Press.

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: a tutorial. *Biological Psychology*, *51*(2-3), 87–107.

Fiehler, K., Ullsperger, M., Grigutsch, M., & Cramon, D. Y. V. (2003). Cardiac responses to error processing and response conflict. In M. Ullsperger & M. Falkenstein (Eds.), *Errors, conflict, and the brain. current opinions on performance monitoring* (pp. 135–147). Leipzig: Max Planck Institute for Human Cognitive and Brain Sciences.

Finney, S. A. (1997). Auditory feedback and musical keyboard performance. *Music Perception*, *15*(2), 153–174.

Finney, S. A. (2001a). FTAP: a Linux-based program for tapping and music experiments. *Behavior Research Methods, Instruments, & Computers*, *33*(1), 65–72.

Finney, S. A. (2001b). Real-time data collection in Linux: a case study. *Behavior Research Methods, Instruments, & Computers*, *33*(2), 167–173.

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

Fisher, R. J., Galea, M. P., Brown, P., & Lemon, R. N. (2002). Digital nerve anaesthesia decreases EMG-EMG coherence in a human precision grip task. *Experimental Brain Research*, *145*(2), 207–214. doi:10.1007/s00221-002-1113-x

Franklin, D. W., & Wolpert, D. M. (2011). Computational mechanisms of sensorimotor control. *Neuron*, *72*(3), 425–442. doi:10.1016/j.neuron.2011.10.006

Furuya, S., & Soechting, J. F. (2010). Role of auditory feedback in the control of successive keystrokes during piano playing. *Experimental Brain Research*, *204*(2), 223–237. doi:10.1007/s00221-010-2307-2

Gabrielsson, A. (2003). Music performance research at the millennium. *Psychology of Music*, *31*(3), 221–272. doi:10.1177/03057356030313002

Ganushchak, L. Y., & Schiller, N. O. (2006). Effects of time pressure on verbal self-monitoring: an ERP study. *Brain Research*, *1125*(1), 104–115. doi:10.1016/j.brainres.2006.09.096

Ganushchak, L. Y., & Schiller, N. O. (2008a). Brain error-monitoring activity is affected by semantic relatedness: an event-related brain potentials study. *Journal of Cognitive Neuroscience*, *20*(5), 927–940. doi:10.1162/jocn.2008.20514

Ganushchak, L. Y., & Schiller, N. O. (2008b). Motivation and semantic context affect brain error-monitoring activity: an event-related brain potentials study. *NeuroImage*, *39*(1), 395–405. doi:10.1016/j.neuroimage.2007.09.001

Garza Villarreal, E. a., Brattico, E., Leino, S., Ostergaard, L., & Vuust, P. (2011). Distinct neural responses to chord violations: a multiple source analysis study. *Brain Research*, *1389*, 103–14. doi:10.1016/j.brainres.2011.02.089

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, *23*(27), 9240–9245.

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*(6), 385–390.

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*(5563), 2279–2282. doi:10.1126/science.1066893

Gemba, H., Sasaki, K., & Brooks, V. B. (1986). 'Error' potentials in limbic cortex (anterior cingulate area 24) of monkeys during motor learning. *Neuroscience Letters*, *70*(2), 223–227.

Gentilucci, M., Toni, I., Daprati, E., & Gangitano, M. (1997). Tactile input of the hand and the control of reaching to grasp movements. *Experimental Brain Research*, *114*(1), 130–137.

Goebl, W., & Palmer, C. (2008). Tactile feedback and timing accuracy in piano performance. *Experimental Brain Research*, *186*(3), 471–479. doi:10.1007/s00221-007-1252-1

Goebl, W., & Palmer, C. (2009). Synchronization of timing and motion among performance musicians. *Music Perception*, *26*(5), 427–439.

Gordon, A. M., & Soechting, J. F. (1995). Use of tactile afferent information in sequential finger movements. *Experimental Brain Research*, *107*(2), 281–292.

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*(2), 161–170. doi:10.1111/j.1469-8986.2005.00278.x

Hajcak, G., McDonald, N., & Simons, R. F. (2003). To err is autonomic: Error-related brain potentials, ANS activity, and post-error compensatory behavior. *Psychophysiology*, *40*(6), 895–903. doi:10.1111/1469-8986.00107

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: the feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905–912. doi:10.1111/j.1469-8986.2007.00567.x

Hajcak, G., Nieuwenhuis, S., Ridderinkhof, K. R., & Simons, R. F. (2005). Error-preceding brain activity: robustness, temporal dynamics, and boundary conditions. *Biological Psychology*, *70*(2), 67–78. doi:10.1016/j.biopsycho.2004.12.001

Haueisen, J., & Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience*, *13*(6), 786–792. doi:10.1162/08989290152541449

Heldmann, M., Rüsseler, J., & Münte, T. F. (2008). Internal and external information in error processing. *BMC Neuroscience*, *9*, 33. doi:10.1186/1471-2202-9-33

Herholz, S. C., & Zatorre, R. J. (2012). Musical training as a framework for brain plasticity: behavior, function, and structure. *Neuron*, *76*(3), 486–502. doi:10.1016/j.neuron.2012.10.011

Herrmann, M. J., Römmler, J., Ehlis, A.-C., Heidrich, A., & Fallgatter, A. J. (2004). Source localization (LORETA) of the error-related-negativity (ERN/Ne) and positivity (Pe). *Cognitive Brain Research*, *20*(2), 294–299. doi:10.1016/j.cogbrainres.2004.02.013

Herrojo Ruiz, M. H., Strübing, F., Jabusch, H.-C., & Altenmüller, E. O. (2011). EEG oscillatory patterns are associated with error prediction during music performance and are altered in musician's dystonia. *NeuroImage*, *55*(4), 1791–1803. doi:10.1016/j.neuroimage.2010.12.050

Herrojo Ruiz, M., Jabusch, H.-C., & Altenmüller, E. O. (2009). Detecting wrong notes in advance: neuronal correlates of error monitoring in pianists. *Cerebral Cortex*, *19*(11), 2625–2639. doi:10.1093/cercor/bhp021

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*, 135–145.

Hillyard, S. A., Hink, R., Schwent, V., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*(4108), 177–180.

Hillyard, S. A., Teder-Sälejärvi, W., & Münte, T. F. (1998). Temporal dynamics of early perceptual processing. *Current Opinion in Neurobiology*, *8*, 202–210.

Hodges, D. A. (2009). Can neuroscience help us do a better job of teaching music? *General Music Today*, *23*(2), 3–12. doi:10.1177/1048371309349569

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. doi:10.1037//0033-295X.109.4.679

Holroyd, C. B., Dien, J., & Coles, M. G. (1998). Error-related scalp potentials elicited by hand and foot movements: evidence for an output-independent error-processing system in humans. *Neuroscience Letters*, *242*(2), 65–68.

Holroyd, C. B., Larsen, J. T., & Cohen, J. D. (2004). Context dependence of the event-related brain potential associated with reward and punishment. *Psychophysiology*, *41*(2), 245–253. doi:10.1111/j.1469-8986.2004.00152.x

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., & Cohen, J. D. (2003). Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport*, *14*(18), 2481–2484. doi:10.1097/01.wnr.0000099601.41403.a5

Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G., & Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, *7*(5), 497–498. doi:10.1038/nn1238

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

Hutchinson, S., Lee, L. H.-L., Gaab, N., & Schlaug, G. (2003). Cerebellar volume of musicians. *Cerebral Cortex*, *13*(9), 943–9.

Ikudome, S., Nakamoto, H., Yotani, K., Kanehisa, H., & Mori, S. (2013). Visuomotor process in movement correction: role of internal feedback loop. *Neuroreport*, *24*(16), 879–883. doi:10.1097/WNR.0000000000000010

Imamizu, H., & Kawato, M. (2012). Cerebellar internal models: implications for the dexterous use of tools. *Cerebellum*, *11*(2), 325–335. doi:10.1007/s12311-010-0241-2

Itagaki, S., & Katayama, J. (2008). Self-relevant criteria determine the evaluation of outcomes induced by others. *Neuroreport*, *19*(3), 383–7. doi:10.1097/WNR.0b013e3282f556e8

Ito, M. (2008). Control of mental activities by internal models in the cerebellum. *Nature Reviews Neuroscience*, *9*(4), 304–313. doi:10.1038/nrn2332

Jessup, R. K., Busemeyer, J. R., & Brown, J. W. (2010). Error effects in anterior cingulate cortex reverse when error likelihood is high. *Journal of Neuroscience*, *30*(9), 3467–3472. doi:10.1523/JNEUROSCI.4130-09.2010

Johansson, R. S., & Westling, G. (1988). Programmed and triggered actions to rapid load changes during precision grip. *Experimental Brain Research*, *71*, 72–86.

Kandel, E. R., Schwartz, J. H., & Jessel, T. M. (Eds.). (1991). *Principles of Neural Science* (3rd ed.). Norwalk, Conneticut: Appleton and Lange.

Katahira, K., Abla, D., Masuda, S., & Okanoya, K. (2008). Feedback-based error monitoring processes during musical performance: an ERP study. *Neuroscience Research*, *61*(1), 120–128. doi:10.1016/j.neures.2008.02.001

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

Kiehl, K. A., Liddle, P. F., & Hopfinger, J. B. (2000). Error processing and the rostral anterior cingulate: an event-related fMRI study. *Psychophysiology*, *37*(2), 216–23.

Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., & Lotze, M. (2007). Overt and imagined singing of an Italian aria. *NeuroImage*, *36*(3), 889–900. doi:10.1016/j.neuroimage.2007.02.053

Knolle, F., Schröger, E., Baess, P., & Kotz, S. A. (2012). The cerebellum generates motor-to-auditory predictions: ERP lesion evidence. *Journal of Cognitive Neuroscience*, *24*(3), 698–706. doi:10.1162/jocn_a_00167

Knolle, F., Schröger, E., & Kotz, S. a. (2013). Prediction errors in self- and externally-generated deviants. *Biological Psychology*, *92*(2), 410–416. doi:10.1016/j.biopsycho.2012.11.017

Koelsch, S. (2005). Neural substrates of processing syntax and semantics in music. *Current Opinion in Neurobiology*, *15*(2), 207–212. doi:10.1016/j.conb.2005.03.005

Koelsch, S. (2009). Music-syntactic processing and auditory memory: similarities and differences between ERAN and MMN. *Psychophysiology*, *46*(1), 179–90. doi:10.1111/j.1469-8986.2008.00752.x

Koelsch, S. (2011). Toward a neural basis of music perception - a review and updated model. *Frontiers in Psychology*, *2*, 1–20. doi:10.3389/fpsyg.2011.00110

Koelsch, S., & Jentschke, S. (2008). Short-term effects of processing musical syntax: an ERP study. *Brain Research*, *1212*, 55–62. doi:10.1016/j.brainres.2007.10.078

Krigolson, O. E., & Holroyd, C. B. (2007). Predictive information and error processing: the role of medial-frontal cortex during motor control. *Psychophysiology*, *44*(4), 586–595. doi:10.1111/j.1469-8986.2007.00523.x

Kruse-Weber, S., & Parncutt, R. (2014). Error management for musicians: an interdisciplinary conceptual framework. *Frontiers in Psychology*, *5*, 1–14. doi:10.3389/fpsyg.2014.00777

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience*, *27*(2), 308–14. doi:10.1523/JNEUROSCI.4822-06.2007

Lashley, K. S. (1951). The Problem of Serial Order in Behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (7, pp. 112–136). New York: Wiley.

Livingstone, S. R., Thompson, W. F., & Russo, F. A. (2009). Facial expressions and emotional singing: A study of perception and production with motion capture and electromyography. *Music Perception*, *26*(5), 475–488.

Loehr, J. D., Kourtis, D., Vesper, C., Sebanz, N., & Knoblich, G. (2013). Monitoring individual and joint action outcomes in duet music performance. *Journal of Cognitive Neuroscience*, *25*(7), 1049–1061. doi:10.1162/jocn

Lotze, M., Braun, C., Birbaumer, N., Anders, S., & Cohen, L. G. (2003). Motor learning elicited by voluntary drive. *Brain*, *126*(4), 866–872. doi:10.1093/brain/awg079

Lotze, M., Montoya, P., Erb, M., Hülsmann, E., Flor, H., Klose, U., … Grodd, W. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *Journal of Cognitive Neuroscience*, *11*(5), 491–501.

Lutz, K., Puorger, R., Cheetham, M., & Jäncke, L. (2013). Development of ERN together with an internal model of audio-motor associations. *Frontiers in Human Neuroscience*, *7*, 1–9. doi:10.3389/fnhum.2013.00471

Luu, P., Tucker, D. M., Derryberry, D., Reed, M., & Poulsen, C. (2003). Electrophysiological responses to errors and feedback in the process of action regulation. *Psychological Science*, *14*(1), 47–53.

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*(5472), 1835–1838. doi:10.1126/science.288.5472.1835

Maes, P.-J., Leman, M., Palmer, C., & Wanderley, M. M. (2014). Action-based effects on music perception. *Frontiers in Psychology*, *4*, 1008. doi:10.3389/fpsyg.2013.01008

Maidhof, C. (2013). Error monitoring in musicians. *Frontiers in Human Neuroscience*, *7*, 1–8. doi:10.3389/fnhum.2013.00401

Maidhof, C., Kästner, T., & Makkonen, T. (2014). Combining EEG, MIDI, and motion capture techniques for investigating musical performance. *Behavior Research Methods*, *46*(1), 185–95. doi:10.3758/s13428-013-0363-9

Maidhof, C., & Koelsch, S. (2011). Effects of selective attention on syntax processing in music and language. *Journal of Cognitive Neuroscience*, *23*(9), 2252–2267. doi:10.1162/jocn.2010.21542

Maidhof, C., Pitkäniemi, A., & Tervaniemi, M. (2013). Predictive error detection in pianists: a combined ERP and motion capture study. *Frontiers in Human Neuroscience*, *7*, 1–14. doi:10.3389/fnhum.2013.00587

Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2009). Nobody is perfect: ERP effects prior to performance errors in musicians indicate fast monitoring processes. *PloS ONE*, *4*(4), 1–7. doi:10.1371/journal.pone.0005032

Maidhof, C., Vavatzanidis, N., Prinz, W., Rieger, M., & Koelsch, S. (2010). Processing expectancy violations during music performance and perception: an ERP study. *Journal of Cognitive Neuroscience*, *22*(10), 2401–2413. doi:10.1162/jocn.2009.21332

Makeig, S., Gramann, K., Jung, T.-P., Sejnowski, T. J., & Poizner, H. (2009). Linking brain, mind and behavior. *International Journal of Psychophysiology*, *73*(2), 95–100. doi:10.1016/j.ijpsycho.2008.11.008

Marco-Pallarés, J., Krämer, U. M., Strehl, S., Schröder, A., & Münte, T. F. (2010). When decisions of others matter to me: an electrophysiological analysis. *BMC Neuroscience*, *11*(1). doi:10.1186/1471-2202-11-86

Masaki, H., Tanaka, H., Takasawa, N., & Yamazaki, K. (2001). Error-related brain potentials elicited by vocal errors. *Neuroreport*, *12*(9), 1851–1855.

Meister, I., Krings, T., Foltys, H., Boroojerdi, B., Müller, M. M., Töpper, R., & Thron, A. (2004). Playing piano in the mind–an fMRI study on music imagery and performance in pianists. *Cognitive Brain Research*, *19*(3), 219–228. doi:10.1016/j.cogbrainres.2003.12.005

Menon, V., Adleman, N. E., White, C. D., Glover, G. H., & Reiss, A. L. (2001). Error-related brain activation during a Go/NoGo response inhibition task. *Human Brain Mapping*, *12*(3), 131–43.

Miall, R. C., Christensen, L. O. D., Cain, O., & Stanley, J. (2007). Disruption of state estimation in the human lateral cerebellum. *PLoS Biology*, *5*(11), 1–12. doi:10.1371/journal.pbio.0050316

Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, *9*(8), 1265–1279.

Miller, E. K. (2000). The prefrontal cortex and cognitive control. *Nature Reviews Neuroscience*, *1*(1), 59–65. doi:10.1038/35036228

Miltner, W. H., Brauer, J., Hecht, H., Trippe, R., & Coles, M. G. (2004). Parallel brain activity for self-generated and observed errors. In M. Ullsperger & M. Falkenstein (Eds.), *Errors, conflict, and the brain. current opinions on performance monitoring* (pp. 124–129). Max Planck Institute for Human Cognitive and Brain Sciences.

Miltner, W. H., Braun, C., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798.

Morris, M., Summers, J. J., Matyas, T., & Iansek, R. (1994). Current status of the motor program. *Physical Therapy*, *74*, 738–748.

Müller, S. V., Möller, J., Rodriguez-Fornells, A., & Münte, T. F. (2005). Brain potentials related to self-generated and external information used for performance monitoring. *Clinical Neurophysiology*, *116*(1), 63–74. doi:10.1016/j.clinph.2004.07.009

Münte, T. F., Altenmüller, E. O., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience*, *3*(6), 473–478. doi:10.1038/nrn843

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, *118*(12), 2544–2590. doi:10.1016/j.clinph.2007.04.026

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. (2004). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neuroscience and Biobehavioral Reviews*, *28*(4), 441–448. doi:10.1016/j.neubiorev.2004.05.003

Nieuwenhuis, S., Yeung, N., Holroyd, C. B., Schurger, A., & Cohen, J. D. (2004). Sensitivity of electrophysiological activity from medial frontal cortex to utilitarian and performance feedback. *Cerebral Cortex*, *14*(7), 741–747. doi:10.1093/cercor/bhh034

Novak, G. P., Ritter, W., Vaughan, H. G., Wiznitzer, J., & Wiznitzer, M. L. (1990). Differentiation of negative event-related potentials in an auditory discrimination task. *Electroencephalography and clinical Neurophysiology*, *75*, 225–275.

Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. *Journal of Cognitive Neuroscience*, *19*(12), 1994–2004. doi:10.1162/jocn.2007.19.12.1994

Overbeek, T. J., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable components of error processing. *Journal of Psychophysiology*, *19*(4), 319–329. doi:10.1027/0269-8803.19.4.319

Palmer, C. (1997). Music performance. *Annual Review of Psychology*, *48*, 115–138. doi:10.1146/annurev.psych.48.1.115

Palmer, C. (2005). Sequence memory in music performance. *Current Directions in Psychological Science*, *14*(5), 247–250.

Palmer, C., & Pfordresher, P. Q. (2003). Incremental planning in sequence production. *Psychological Review*, *110*(4), 683–712. doi:10.1037/0033-295X.110.4.683

Palmer, C., & Van de Sande, C. (1993). Units of knowledge in music performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(2), 457–470.

Palmer, C., & Van de Sande, C. (1995). Range of planning in music performance. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(5), 947–962.

Parncutt, R. (2007). Can researchers help artists? Music performance research for music students. *Music Performance Research*, *1*(1), 1–25.

Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods and Findings in Experimental and Clinical Pharmacology*, *24 Suppl D*, 5–12.

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1994). Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *International Journal of Psychophysiology*, *18*, 49–65.

Petacchi, A., Laird, A. R., Fox, P. T., & Bower, J. M. (2005). Cerebellum and auditory function: an ALE meta-analysis of functional neuroimaging studies. *Human Brain Mapping*, *25*(1), 118–128. doi:10.1002/hbm.20137

Pfordresher, P. Q. (2003). Auditory feedback in music performance: Evidence for a dissociation of sequencing and timing. *Journal of Experimental Psychology: Human Perception and Performance*, *29*(5), 949–964.

Pfordresher, P. Q. (2005). Auditory feedback in music performance: the role of melodic structure and musical skill. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1331–1345. doi:10.1037/0096-1523.31.6.1331

Pfordresher, P. Q. (2006). Coordination of perception and action in music performance. *Advances in Cognitive Psychology*, *2*(2), 183–198. doi:10.2478/v10053-008-0054-8

Pfordresher, P. Q., & Benitez, B. (2007). Temporal coordination between actions and sound during sequence production. *Human Movement Science*, *26*(5), 742–756. doi:10.1016/j.humov.2007.07.006

Pfordresher, P. Q., & Dalla Bella, S. (2011). Delayed auditory feedback and movement. *Journal of Experimental Psychology. Human Perception and Performance*, *37*(2), 566–579. doi:10.1037/a0021487

Pfordresher, P. Q., Mantell, J. T., Brown, S., Zivadinov, R., & Cox, J. L. (2014). Brain responses to altered auditory feedback during musical keyboard production: An fMRI study. *Brain Research*, *1556*, 28–37. doi:10.1016/j.brainres.2014.02.004

Pfordresher, P. Q., & Palmer, C. (2002). Effects of delayed auditory feedback on timing of music performance. *Psychological Research*, *66*(1), 71–79. doi:10.1007/s004260100075

Pfordresher, P. Q., & Palmer, C. (2006). Effects of hearing the past, present, or future during music performance. *Perception & Psychophysics*, *68*(3), 362–376.

Pizzagalli, D. (2007). Electroencephalography and high-density electrophysiological source localization. In J. T. Cacioppo, L. G. Tassinary & G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 56–84). Cambridge: Cambridge University Press.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*(10), 2128–2148. doi:10.1016/j.clinph.2007.04.019

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, *9*(2), 129–154. doi:10.1080/713752551

Pruszynski, J. A., Kurtzer, I., Nashed, J. Y., Omrani, M., Brouwer, B., & Scott, S. H. (2011). Primary motor cortex underlies multi-joint integration for fast feedback control. *Nature*, *478*(7369), 387–390. doi:10.1038/nature10436

Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A.-S., McNamara, J. O., & Williams, S. M. (Eds.). (2004). *Neuroscience*. Sunderland, Massachusetts: Sinauer Associates, Inc.

Rabbitt, P. (1978). Detection of errors by skilled typists. *Ergonomics*, *21*(11), 945–958.

Rabin, E., & Gordon, A. M. (2004). Tactile feedback contributes to consistency of finger movements during typing. *Experimental Brain Research*, *155*(3), 362–369. doi:10.1007/s00221-003-1736-6

Ramnani, N. (2006). The primate cortico-cerebellar system: anatomy and function. *Nature Reviews Neuroscience*, *7*(7), 511–522. doi:10.1038/nrn1953

Repp, B. H. (1999a). Control of expressive and metronomic timing in pianists. *Journal of Motor Behavior*, *31*(2), 145–164.

Repp, B. H. (1999b). Effects of auditory feedback deprivation on expressive piano performance. *Music Perception*, *16*(4), 409–438.

Ridderinkhof, K. R. (2003). Errors are foreshadowed in brain potentials associated with action monitoring in cingulate cortex in humans. *Neuroscience Letters*, *348*(1), 1–4. doi:10.1016/S0304-3940(03)00566-4

Ridderinkhof, K. R., Ullsperger, M., Crone, E. a., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*(5695), 443–447. doi:10.1126/science.1100301

Rosenbaum, D. A. (2009). *Human Motor Control* (2nd ed.). Amsterdam, The Netherlands: Elsevier.

Ruchsow, M., Grothe, J., Spitzer, M., & Kiefer, M. (2002). Human anterior cingulate cortex is activated by negative feedback: evidence from event-related potentials in a guessing task. *Neuroscience Letters*, *325*(3), 203–206.

Sanes, J., & Donoghue, J. (2000). Plasticity and primary motor cortex. *Annual Review of Neuroscience*, *23*, 393–415.

Schlaug, G. (2009). Music, musicians, and brain plasticity. In S. Hallam, I. Cross & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 197–207). Oxford.

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

Schmidt, R. A., & Lee, T. D. (2005). *Motor control and learning: A behavioral emphasis* (Fourth edi). Champaign, IL: Human Kinetics Books.

Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, *11*(5), 211–218. doi:10.1016/j.tics.2007.02.006

Shadmehr, R., Smith, M. a., & Krakauer, J. W. (2010). Error correction, sensory prediction, and adaptation in motor control. *Annual Review of Neuroscience*, *33*, 89–108. doi:10.1146/annurev-neuro-060909-153135

Shibasaki, H., Barrett, G., Halliday, E., & Halliday, A. M. (1980). Components of the movement-related cortical potential and their scalp topography. *Electroencephalography and clinical Neurophysiology*, *49*(3-4), 213–226.

Sloboda, J. A. (2000). Individual differences in music performance. *Trends in Cognitive Sciences*, *4*(10), 397–403.

Strick, P. L., Dum, R. P., & Fiez, J. a. (2009). Cerebellum and nonmotor function. *Annual Review of Neuroscience*, *32*, 413–434. doi:10.1146/annurev.neuro.31.060407.125606

Strübing, F., Herrojo Ruiz, M., Jabusch, H.-C., & Altenmüller, E. O. (2012). Error monitoring is altered in musician's dystonia: evidence from ERP-based studies. *Annals of the New York Academy of Sciences*, *1252*, 192–199. doi:10.1111/j.1749-6632.2011.06417.x

Summers, J. J., & Anson, J. G. (2009). Current status of the motor program: revisited. *Human Movement Science*, *28*(5), 566–577. doi:10.1016/j.humov.2009.01.002

Tervaniemi, M., Castaneda, A., Knoll, M., & Uther, M. (2006). Sound processing in amateur musicians and nonmusicians: event-related potential and behavioral indices. *Neuroreport*, *17*(11), 1225–1228. doi:10.1097/01.wnr.0000230510.55596.8b

Tzur, G., & Berger, A. (2007). When things look wrong: theta activity in rule violation. *Neuropsychologia*, *45*(13), 3122–3126. doi:10.1016/j.neuropsychologia.2007.05.004

Tzur, G., & Berger, A. (2009). Fast and slow brain rhythms in rule/expectation violation tasks: focusing on evaluation processes by excluding motor action. *Behavioural Brain Research*, *198*(2), 420–428. doi:10.1016/j.bbr.2008.11.041

Ullsperger, M., Nittono, H., & von Cramon, D. Y. (2007). When goals are missed: dealing with self-generated and externally induced failure. *NeuroImage*, *35*(3), 1356–1364. doi:10.1016/j.neuroimage.2007.01.026

Ullsperger, M., & von Cramon, D. Y. (2003). Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *Journal of Neuroscience*, *23*(10), 4308–4314.

Van Gemmert, a. W., & Van Galen, G. P. (1997). Stress, neuromotor noise, and human performance: a theoretical perspective. *Journal of Experimental Psychology. Human Perception and Performance*, *23*(5), 1299–1313.

Van Veen, V., & Carter, C. S. (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience*, *14*(4), 593–602. doi:10.1162/08989290260045837

van Schie, H., & Mars, R. B. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, *7*(5), 549–554. doi:10.1038/nn1239

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology*, *21*(1992), 147–163. doi:10.1027/0269-8803.21.3.147

Wittfoth, M., Schardt, D. M., Fahle, M., & Herrmann, M. (2009). How the brain resolves high conflict situations: double conflict involvement of dorsolateral prefrontal cortex. *NeuroImage*, *44*(3), 1201–1209. doi:10.1016/j.neuroimage.2008.09.026

Wolpert, D. M. (1997). Computational approaches to motor control. *Trends in Cognitive Sciences*, *1*(6), 209–216.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), R729–732.

Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, *3*(supplement), 1212–1217. doi:10.1038/81497

Wolpert, D. M., Ghahramani, Z., & Flanagan, J. R. (2001). Perspectives and problems in motor learning. *Trends in Cognitive Sciences*, *5*(11), 487–494.

Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, *269*(5232), 1880–1882.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review*, *111*(4), 931–959. doi:10.1037/0033-295X.111.4.931

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex*, *15*(5), 535–544. doi:10.1093/cercor/bhh153

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, *24*(28), 6258–64. doi:10.1523/JNEUROSCI.4537-03.2004

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, *8*(7), 547–558. doi:10.1038/nrn2152

List of abbreviations

AAF	Altered auditory feedback
ACC	Anterior cingulate cortex
ANOVA	Analysis of Variance
BA	Brodman's area
BPM	Beats per minute
CMA	Cingulate motor areas
CNV	Contingent Negative Variation
DLPFC	Dorso-lateral prefrontal cortex
EEG	Electroencephalography, or Electroencephalogram
EMG	Electromyography, or Electromyogram
EPP	Error-preceding Positivity
EPSP	Excitatory postsynaptic potential
ERAN	Early right anterior negativity
ERN	Error-related negativity
ERP	Event-related potential
FIR	Finite impulse response
fMRI	Functional magnetic resonance imaging
FRN	Feedback-related negativity
ICA	Independent Component Analysis
lPFC	lateral prefrontal cortex
MMN	Mismatch negativity
Pe	Error positivity

pFMC	posterior frontomedial cortex
PMd	dorsal premotor area
PMv	ventral premotor area
sLORETA	standardized low resolution brain electromagnetic tomography
SMA	Supplementary motor area
SNR	Signal-to-noise ratio

A. Short Summary of the Findings

Music performance is regarded as one of the most complex human behaviors. A fundamental aspect of music making is highly accurate motor control and ensuring that the intended goals are or will be achieved. For that, musicians have to constantly monitor their performed actions and the resulting sensory effects in order to detect deviations from the intended behavior. However, the cognitive and neural mechanisms of action monitoring including error processing during complex time-based sequential behaviors such as music performance have remained elusive. This dissertation describes three experiments that investigated different aspects of action monitoring during piano performances of skilled musicians by using event-related potentials (ERPs). In Experiment 1 and 2, ERPs elicited by correctly produced key presses and their corresponding auditory feedback were compared to ERPs elicited by correctly produced key presses whose auditory feedback (pitch content) was at random positions manipulated (action condition). Additionally, ERPs were recorded while pianists merely listened to those stimuli they had to perform in the action condition (perception condition). Results suggest that the intention and action of producing a certain auditory effect enhances expectancies and influences the processing of the auditory input (indicated by a larger feedback-related negativity in the action compared to the perception condition). In Experiment 3, the comparison of ERPs elicited by correctly and incorrectly produced key presses showed that error-related processes occur before auditory feedback of an error is available. In addition, a newly developed setup allowed for the simultaneous recording of detailed movement and EEG data. Results of the combined analysis of pianists' finger movements and ERP results tentatively indicate that tactile feedback might play an important role for error detection. Both main findings are explained in terms of predictive mechanisms based on the common coding approach and internal forward models.

B. Kurzzusammenfassung der Ergebnisse

Musizieren ist ein komplexes Unterfangen und erfordert präzise motorische Kontrolle und Handlungsüberwachung. So kann festgestellt werden, ob ausgeführte Handlungen und ihre sensorischen Effekte dem intendierten Verhalten entsprechen. Die neurokognitiven Grundlagen der Handlungsüberwachung während eines komplexen Verhaltens wie dem Musizieren sind jedoch weitgehend noch unerforscht. Im Rahmen dieser Arbeit wurden drei Experimente durchgeführt, die mittels Ereignis-korrelierten Potentialen (EKPs) Aspekte der Handlungsüberwachung (inklusive Fehlerverarbeitung) bei Musikern während des Klavierspielens aus dem Gedächtnis untersuchten. In den Experimenten 1 und 2 wurden EKPs verglichen, die einmal während korrekt gespielter Noten samt korrekter auditorischer Rückmeldung, und einmal während korrekt gespielter Noten, die eine manipulierte auditorische Rückmeldung enthielten, evoziert wurden (Handlungsbedingung). Zusätzlich wurden EKPs in einer Bedingung gemessen, in der Versuchsteilnehmer die Stimuli lediglich hörten, die sie in der Handlungsbedingung produzieren mussten (Wahrnehmungsbedingung). Die Ergebnisse deuten darauf hin, dass die Intention und die ausgeführte Handlung Erwartungen bezüglich der auditorischen Rückmeldung verstärken, was sich in einer grösseren Amplitude der Feedback-related negativity in der Handlungs- im Vergleich zur Wahrnehmungsbedingung widerspiegelte. In Experiment 3 wurden EKPs miteinander verglichen, die während des Spielens korrekter und inkorrekter Noten evoziert wurden. Die Ergebnisse lassen darauf schließen, dass Prozesse der Fehlerverarbeitung schon vor der auditorischen Rückmeldung des Fehlers ablaufen, und somit unabhängig von dieser sind. Eine kombinierte Betrachtung der Fingertrajektorien und EKPs lässt vorläufig vermuten, dass der taktilen Rückmeldung von Tastendrücken eine wichtige Rolle für frühe fehlerbezogene Prozesse während des Klavierspiels zukommt. Diese Befunde werden innerhalb des Ansatzes des *common coding* und interner Vorwärtsmodelle erklärt.

C. Lebenslauf

The curriculum vitae is not included in the online version.

Der Lebenslauf ist in der Online-Version aus Gründen des Datenschutzes nicht enthalten.

D. Publikationsliste

Forschungsartikel

- Maidhof, C., Kästner, T., & Makkonen, T. (2014). Combining EEG, MIDI, and motion capture techniques for investigating musical performance. *Behavior Research Methods*, *46*(1), 185–95. doi:10.3758/s13428-013-0363-9
- Maidhof, C., Pitkäniemi, A., & Tervaniemi, M. (2013). Predictive error detection in pianists: a combined ERP and motion capture study. *Frontiers in Human Neuroscience*, *7*, 1–14. doi:10.3389/fnhum.2013.00587
- Maidhof, C. (2013). Error monitoring in musicians. *Frontiers in Human Neuroscience*, *7*, 1–8. doi:10.3389/fnhum.2013.00401
- Maidhof, C. & Koelsch, S. (2011). Effects of selective attention on syntax processing in music and language. *Journal of Cognitive Neuroscience*, *23*(9), 2252–2267. doi:10.1162/jocn.2010.21542
- Maidhof, C., Vavatzanidis, N., Prinz, W., Rieger, M., & Koelsch, S. (2010). Processing expectancy violations during music performance and perception: an ERP study. *Journal of Cognitive Neuroscience*, *22*(10), 2401–2413. doi:10.1162/jocn.2009.21332
- Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2009). Nobody is perfect: ERP effects prior to performance errors in musicians indicate fast monitoring processes. *PloS ONE*, *4*(4), 1–7. doi:10.1371/journal.pone.0005032

Konferenzbeiträge (Auswahl)

- Maidhof, C., Pitkäniemi, A., & Tervaniemi, M. (2014). Investigation of music performance by combining EEG, motion capture, and MIDI data. **Posterpräsentation** auf der Konferenz *The Neurosciences and Music - V*, Dijon, Frankreich, 29. Mai–1. Juni 2014.
- Maidhof, C., Pitkäniemi, A., & Tervaniemi, M. (2012). Investigating the neural correlates of action monitoring in pianists by combining MIDI, EEG, and movement data. **Vortrag** auf der Konferenz *Annual Symposium for Music Scholars in Finland*, 21.–23. März 2012.

- Maidhof, C. & Tervaniemi, M. (2011). Combining EEG, motion capture, and MIDI data for the investigation of music performance. **Posterpräsentation** auf der Konferenz *The Neurosciences and Music - IV*, Edinburgh, Schottland, Großbritannien, 9.–12. Juni 2011.

- Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2009). Neural correlates of action monitoring and error processing during music performance and perception. **Vortrag** auf der *7th Triennial Conference of the European Society for the Cognitive Sciences of Music*, Jyväskylä, Finnland, 12.–16. August 2009.

- Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2008). Neural correlates of expectancy violations during music performance and perception. An ERP study. **Posterpräsentation** auf der Konferenz *The Neurosciences and Music – III: Disorders and Plasticity*, Montreal, Kanada, 25.–28. Juni 2008.

- Maidhof, C., Rieger, M., Prinz, W., & Koelsch, S. (2007). Nobody's perfect: Neural correlates of performance errors in musicians. **Posterpräsentation** auf der Konferenz *48th Annual Meeting of the Psychonomic Society*, Long Beach, USA, 15.–18. November 2007.

- Maidhof, C., Rieger, M., & Koelsch, S. (2006). Processing of self-made and induced errors in musicians. **Posterpräsentation** auf der *9th International Conference on Music Perception and Cognition*, Bologna, Italien, 22.–26. August 2006.

- Maidhof, C., Rieger, M., & Koelsch, S. (2006). Die Verarbeitung von eigenen und fremd-induzierten Fehlern bei Musikern. **Posterpräsentation** auf der *48. Tagung experimentell arbeitender Psychologen*, Johannes Gutenberg-Universität, Mainz, 26.–29. März 2006.

E. Selbstständigkeitserklärung

Hiermit versichere ich, dass ich die Dissertation selbstständig verfasst habe und keine anderen als die angegebenen Hilfsmittel und Hilfen verwendet habe. Zudem versichere ich, dass ich die Dissertation an keiner anderen Universität eingereicht habe, und sie dementsprechend auch nicht in einem früheren Promotionsverfahren angenommen oder abgelehnt wurde. Die Gemeinsame Promotionsordnung zum Dr. phil. der Freien Universität Berlin, veröffentlicht im Amtsblatt der Freien Universität Berlin (60/2008, 2. Dezember 2008), ist mir bekannt.

Ich hatte überwiegenden Anteil bei der Entwicklung der Konzeption, Methodenentwicklung, Versuchsdesign, Datenerhebung, Datenauswertung, Ergebnisdiskussion und Erstellung der Manuskripte. Die Literaturrecherche wurde von mir (vollständig) alleine ausgeführt. Diese Angaben beziehen sich auf alle Veröffentlichungen, auf denen diese Dissertation basiert.

Berlin, 2. Juli 2015