

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

Clemens Maidhof¹, Niki Vavatzanidis¹, Wolfgang Prinz¹,
Martina Rieger¹, 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,

¹Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany, ²University 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, $18.5 (\pm 2.9)$ 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.

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 pseudo-randomized, 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.

Figure 1 consists of three musical examples labeled A, B, and C. Example A, titled 'Pattern A', shows a musical score in 3/4 time with a tempo of 69 bpm. It features two staves of music with complex, overlapping rhythmic patterns. Example B, titled 'Pattern B', also shows a musical score in 3/4 time with a tempo of 69 bpm, featuring similar complex rhythmic patterns. Example C, titled 'Diatonic Scale', shows a musical score in 3/4 time with a tempo of 144 bpm, featuring a simple diatonic scale. All examples are in C-Major.

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.

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 k Ω . 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 band-pass 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 μ V 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 μ V 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 inter-onset 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.

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

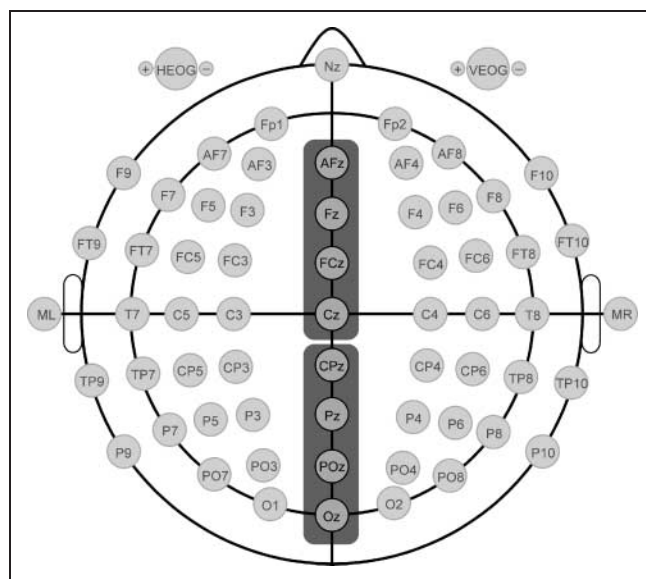


Figure 2. Electrode locations. The highlighted areas show the ROIs used for statistical analyses in the two experiments. Electrode Fpz was additionally measured in Experiment 1.

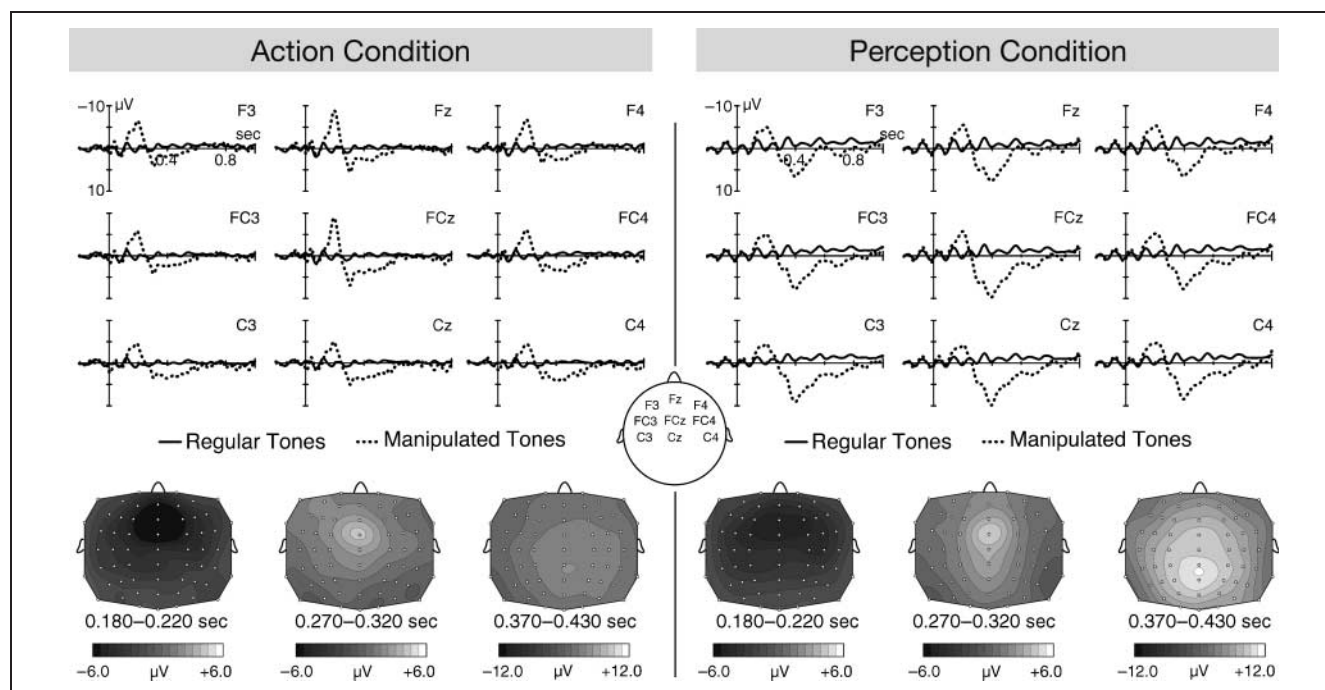


Figure 3. Experiment 1. Grand-average ERPs (and scalp distributions) elicited by the auditory feedback manipulations during the action condition (left) and by the task-relevant manipulated tones during the perception condition (right).

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 condi-

tion: 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

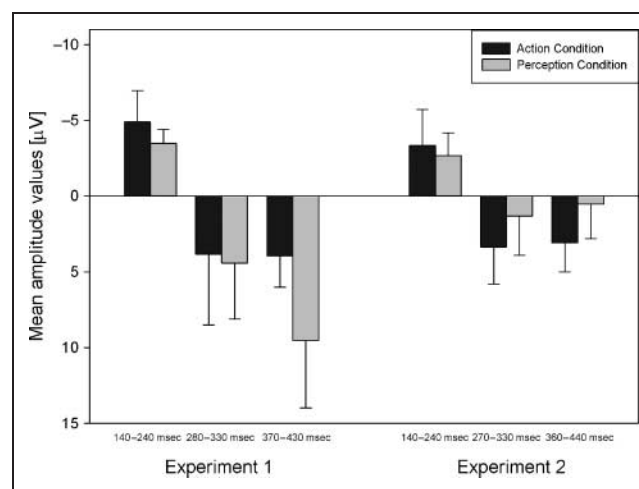
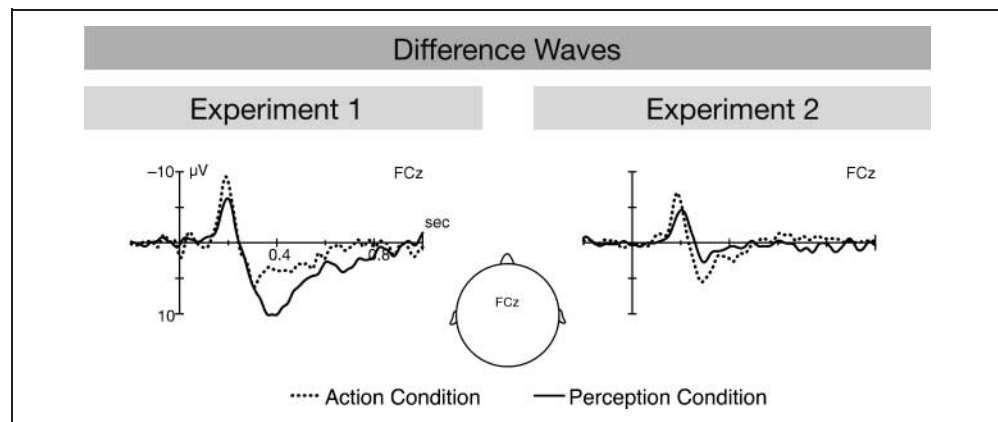


Figure 4. Mean amplitude values (and the standard deviations) for the ERP effects in all three time windows in the two experiments. Note that negative is plotted up.

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, $14.9 (\pm 4.8)$ 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 $30 \mu\text{V}$. 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

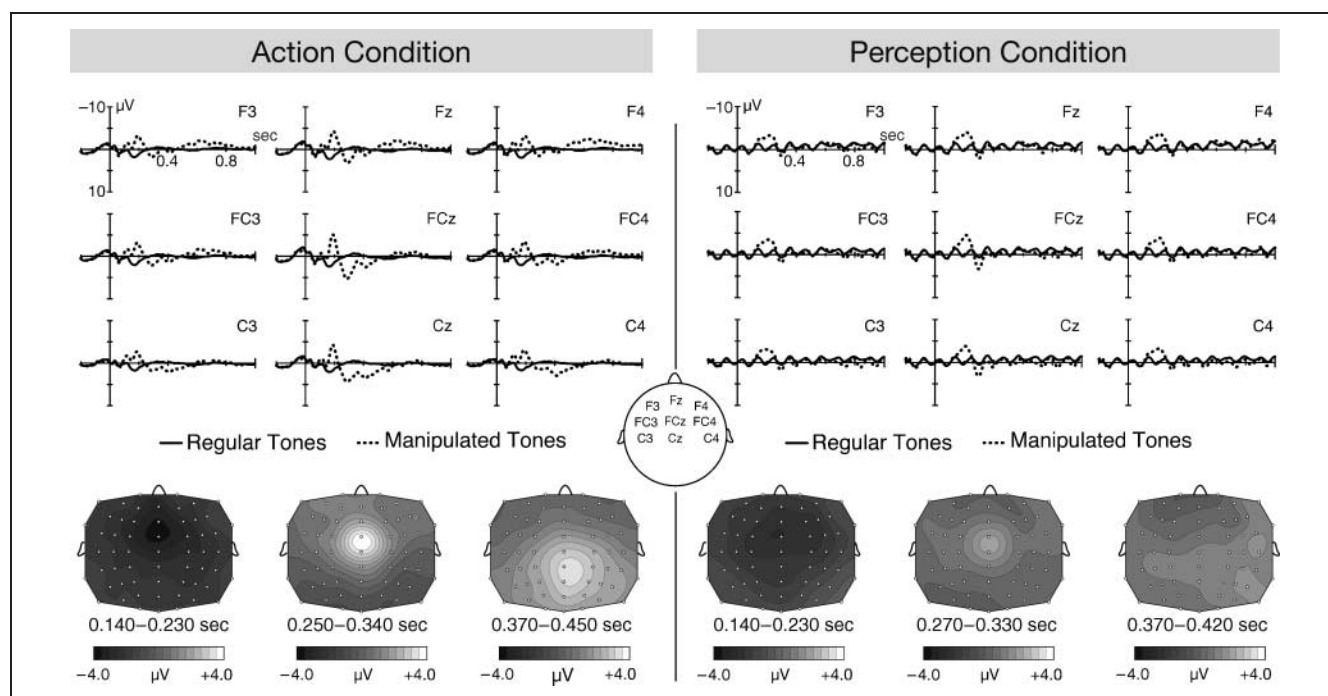


Figure 6. Experiment 2. Grand-average ERPs (and scalp distributions) elicited by the auditory feedback manipulations during the action condition (left) and by the task-irrelevant manipulated tones during the perception condition (right).

(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

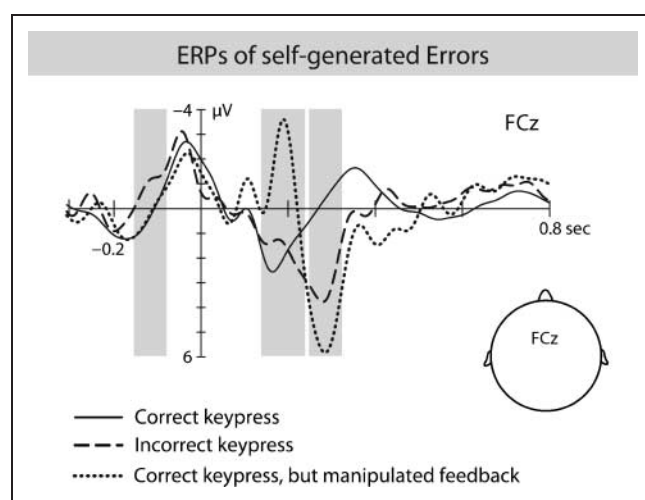


Figure 7. ERPs elicited by correct feedback, feedback manipulations, and self-generated errors. The gray-shaded areas represent the time windows used for statistical analyses. Results showed no negativity after performance errors (compared to correct notes) in the time window of the feedback ERN/N200, 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).

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).

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

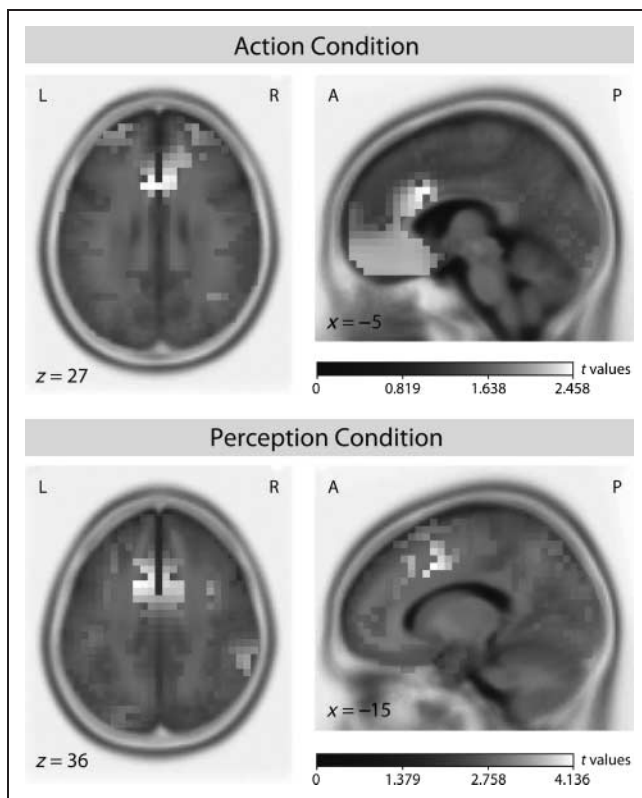


Figure 8. Source localization. Results of the sLORETA analysis suggest that the main neural generators of both negativities lie within the RCZ (Talairach coordinates: $x = -5$, $y = 16$, $z = 27$ in the action condition, and $x = -15$, $y = 11$, $z = 36$ in the perception condition).

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, Müssele, 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

- 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).
- 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., Müssele, 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.