

# Sensory, Cognitive, and Sensorimotor Learning Effects in Recognition Memory for Music

Brian Mathias<sup>1</sup>, Barbara Tillmann<sup>2</sup>, and Caroline Palmer<sup>1</sup>

## Abstract

■ Recent research suggests that perception and action are strongly interrelated and that motor experience may aid memory recognition. We investigated the role of motor experience in auditory memory recognition processes by musicians using behavioral, ERP, and neural source current density measures. Skilled pianists learned one set of novel melodies by producing them and another set by perception only. Pianists then completed an auditory memory recognition test during which the previously learned melodies were presented with or without an out-of-key pitch alteration while the EEG was recorded. Pianists indicated whether each melody was altered from or identical to one of the original melodies. Altered pitches elicited a larger N2 ERP component than original pitches, and pitches within previously produced melodies elicited a larger N2 than pitches in previously perceived melodies. Cortical motor plan-

ning regions were more strongly activated within the time frame of the N2 following altered pitches in previously produced melodies compared with previously perceived melodies, and larger N2 amplitudes were associated with greater detection accuracy following production learning than perception learning. Early sensory (N1) and later cognitive (P3a) components elicited by pitch alterations correlated with predictions of sensory echoic and schematic tonality models, respectively, but only for the perception learning condition, suggesting that production experience alters the extent to which performers rely on sensory and tonal recognition cues. These findings provide evidence for distinct time courses of sensory, schematic, and motoric influences within the same recognition task and suggest that learned auditory–motor associations influence responses to out-of-key pitches. ■

## INTRODUCTION

A growing body of research in cognitive neuroscience documents the role of motor experience in forging links between perception, memory, and action. Recent theories have focused on a role of internal motor simulations in the observation of external events, in which the motor system becomes activated during perception and constrains perceptual interpretations of the environment (Wilson & Knoblich, 2005). Other theories such as the motor theory of speech perception hold that the tight coupling of sounds and movements during speech production results in the perception of motor gestures, rather than the perception of acoustic features (Liberman & Mattingly, 1985). Skills such as learning to speak a language or play a musical instrument forge associations between auditory and motor systems, and these associations are reinforced over years of practice (Draganski & May, 2008; Zatorre, Chen, & Penhune, 2007; Palmer, 1997). Reciprocal interactions between auditory and motor networks have been observed in several neuroimaging studies; for example, listening to musical sounds can trigger activation in cortical motor regions when an individual has experience performing an instrument

(Brown et al., 2013; Lahav, Saltzman, & Schlaug, 2007; Haslinger et al., 2005). Auditory–motor integration may be accomplished through dorsal–ventral auditory stream interactions, which permit the transformation of auditory signals into corresponding motor programs (Rauschecker, 2011) as well as through premotor cortex activations (Rizzolatti & Craighero, 2004) during both the perception and production of auditory signals. Memories for auditory sequences that have been produced may therefore encompass both auditory and motor components.

Some studies suggest that auditory–motor experience can strengthen or enhance recognition memory for sounds. Words that have recently been produced tend to be better recognized than words that have recently only been heard (MacDonald & MacLeod, 1998) or mouthed without sound (Gathercole & Conway, 1988). This effect of motor experience on memory recognition has been termed the “production effect” (MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010; Dodson & Schacter, 2001). Similar effects have been documented in the domain of music: Learning melodies by auditory–motor production can lead to improved recognition of the melodies (Brown & Palmer, 2012), enhanced recognition of within-key pitch changes (Mathias, Palmer, Perrin, & Tillmann, 2014), and greater reorganizational changes within auditory cortex (Lappe, Herholz, Trainor, & Pantev,

<sup>1</sup>McGill University, Montreal, Canada, <sup>2</sup>Lyon Neuroscience Research Center, University Lyon 1, France

2008) compared with auditory-only learning. Thus, effects of motor learning on memory for music suggest that the recognition of previously performed music differs from that of music that has only been heard before.

Memory recognition is influenced by both low-level sensory information and high-level cognitive processes, which interact during perception and arise over different timescales (Bigand, Poulin, Tillmann, Madurell, & D'Adamo, 2003; Tekman & Bharucha, 1998). In the case of music, in which pitches are presented in long sequences, both sensory and cognitive aspects of pitch perception play roles in recognition. Short-term sensory information about acoustic features of musical tones may guide low-level recognition processes (Leman, 2000) and long-term schematic information about typical distributional characteristics of tones may guide high-level recognition processes (Bharucha, 1987; Krumhansl & Kessler, 1982). Thus, memory for pitch may consist of echoic memory traces, which linger or resonate following a sensory experience for a short amount of time (0.5–2 sec; Leman, 2000; Huron & Parncutt, 1993), as well as schematic knowledge of tone frequencies within a musical style, acquired through implicit learning of stimulus regularities over extended exposure (Hannon & Trainor, 2007; Tillmann, Bharucha, & Bigand, 2000).

Accounts of pitch perception have distinguished sensory and cognitive processes. At one end of a sensory-cognitive continuum is Leman's (2000) physiological model of auditory STM, which predicts the stimulus-driven "tension" of tones relative to a preceding context based on pitch periodicity information received by the ear. At the other end is Krumhansl and Kessler's (1982) cognitive account of schematic tonal knowledge, which predicts tonal stability based on listeners' judgments of each tone's relatedness to a preceding tonal context. Tonal profiles observed by Krumhansl and Kessler (1982) closely resemble hierarchical tonal structures proposed by music theorists; these profiles have been correlated with simulations of sensory pitch memory (Leman, 2000). Priming studies in which listeners react to musical target chords following tonal priming sequences have shown that listeners are influenced more by auditory sensory cues when tonal priming sequences are presented at rapid rates (about 75 msec per chord and faster) and more by schematic tonal relatedness when the sequences are presented at slower rates (Bigand et al., 2003; Tekman & Bharucha, 1998; see also Collins, Tillmann, Barrett, Delbé, & Janata, 2014). Thus, pitch perception may unfold over multiple representational stages, across which sensory and cognitive information are differentially weighted.

EEG measures of pitch perception have also distinguished between time frames related to sensory and cognitive processing. ERPs that occur quickly following pitch onsets (within 100 msec) are thought to reflect primarily stimulus-driven neural responses, whereas later potentials are thought to be generated by the brain's own cognitive computations (Rugg & Coles, 1995). The N1 ERP component, a negative-going component occurring about

100 msec following auditory onsets, appears to be sensitive to changes in basic acoustic features (Näätänen & Winkler, 1999). Some evidence suggests that the N1 is sensitive to the tonal importance of pitches: Listeners showed a larger N1 amplitude in response to tonally important (dominant) pitches than to less important (subdominant) pitches that occurred equally often within a melodic context (Krohn, Brattico, Välimäki, & Tervaniemi, 2007); the authors interpreted this as evidence of a more accurate neural representation for pitches with a higher frequency of occurrence over long-term learning (see also Marmel, Perrin, & Tillmann, 2011, for related evidence). The N2 component, which occurs about 200 msec post-onset, and the subsequent frontally maximal P3a respond to more abstract musical properties such as whether a pitch matches the musical key context (Brattico, Tervaniemi, Näätänen, & Peretz, 2006; Tervaniemi et al., 2003; Regnault, Bigand, & Besson, 2001; Besson & Faita, 1995; Janata, 1995). Another P3 subcomponent, the P3b, is elicited when participants are asked to respond to expectancy violations; P3b amplitudes are unrelated to the degree of expectancy violation and may reflect instead the updating of working memory representations following an unexpected stimulus (Donchin & Coles, 1988). The early right anterior negativity, which is sensitive to syntactic chord violations within a particular key, peaks within a similar time frame as the N2 (Koelsch & Jentschke, 2010; Koelsch, Jentschke, Sammler, & Mietschen, 2007). Implicit learning of sequential pitch probabilities during listening can also mediate these ERP responses (Loui, Wu, Wessel, & Knight, 2009). Whereas the N1 reflects sensory processing of acoustic information by auditory cortex (Näätänen & Winkler, 1999), the N2 may indicate the detection of mismatch between the expected and perceived pitches (Folstein & Van Petten, 2008), and the P3 the cognitive evaluation of the unexpected event (Polich, 2007). Thus, these ERPs do not relate to recognition memory per se but often follow altered pitches; we expect therefore that modulation of these components via learning manipulations will yield insights into the properties of recognition memory.

The goal of the current study was to investigate the effect of musicians' auditory and motor familiarity with tonal melodies on their auditory recognition of those melodies. Skilled musicians learned melodies either by performing them on a musical keyboard or simply by listening. Following learning, ERP, electric source current density, and behavioral measures were recorded as musicians heard original pitches or pitch alterations (memory violations) in auditorily presented melodies that they had learned. The pitch alterations were outside the key of the musical sequence and, therefore, engaged both sensory memory (acoustic spectra of the altered tones differed from preceding tones) and schematic memory (altered tones did not belong to the musical scale on which the melody was based). Sensory tension induced by pitch alterations was simulated with a physiological model of auditory sensory memory (Leman, 2000), and schematic

tonality of altered pitches was simulated in terms of their tonal relatedness to the preceding melodic context (Krumhansl & Kessler, 1982). Pitch alterations were expected to elicit an early sensory component (N1) associated with measures of sensory memory (Leman, 2000) based on previous findings that suggest the N1 is sensitive to changes in acoustic features such as pitch (Näätänen & Winkler, 1999) as well as the tonal function of pitches based on a preceding context (Marmel et al., 2011; Krohn et al., 2007). Pitch alterations were also expected to elicit later cognitive components (N2 and P3) associated with measures of tonality (Krumhansl & Kessler, 1982) based on findings that these components are sensitive to musical tonality (Brattico et al., 2006; Janata, 1995). Production learning was expected to enhance pianists' sensitivity to altered pitch events for previously produced melodies (compared with previously heard melodies), indexed by the behavioral measure of detection accuracy and the N2 component amplitudes, as well as greater involvement of motor regions in the current densities following production learning compared with perception learning.

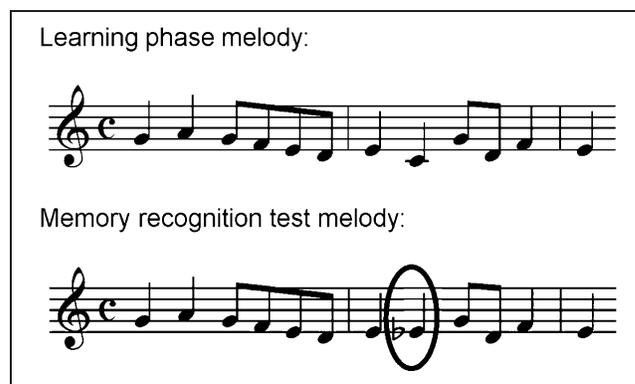
## METHODS

### Participants

Twenty-six right-handed adult pianists from the Montreal community participated in the study. Six were excluded from analyses because of excessive EEG artifacts. The remaining 20 pianists (10 women, age  $M = 21.5$  years,  $SD = 3.1$  years) had between 6 and 17 years of piano instruction ( $M = 10.5$  years,  $SD = 3.0$  years) and currently practiced the piano an average of 4.8 hr per week ( $SD = 3.8$  hr). No participants reported possessing absolute pitch or any hearing problems. Participants provided written informed consent before participating in the study, and the study was reviewed by the McGill University research ethics board.

### Stimulus Materials

Twelve melodies notated in 4/4 time signature, each 12 notes in length and conforming to conventions of Western tonal music, were used in the study (see Figure 1 for an example). The melodies were selected from a larger corpus (Brown & Palmer, 2012) and were assigned to one of two sets for the production and perception learning conditions, to be of equal recognition difficulty, based on previously acquired recognition accuracy scores (Brown & Palmer, 2012). Audio recordings of the melodies, containing natural timing variation, were obtained from two skilled pianists with Cubase 6 software from an M-Audio Keystation 88es MIDI piano keyboard (Cumberland, RI). A 500-msec interonset interval (IOI) metronome, which sounded for eight quarter notes before the start of each recording, set the performance tempo. These recordings



**Figure 1.** Top: One of the notated stimulus melodies. Bottom: The same stimulus melody containing an altered pitch (circled) that was heard during the altered pitch detection test.

were presented to participants during the perception learning condition with a Cubase HALion One piano timbre. The same timbre was used for the auditory feedback heard during the production learning condition.

The two sets of notated melodies presented during the perception and production learning conditions were later presented during a memory recognition test as computer-generated MIDI recordings with 500-msec per quarter note interonset intervals (with no expressive timing variations) and with the same timbre as in the learning conditions. MIDI velocity was constant for all pitches (no expressive intensity variations). Each participant heard every melody with and without pitch alterations. An example of a melody and a pitch alteration is shown in Figure 1. Each altered pitch was a nondiatonic tone (one of five pitch classes from outside the musical “key”) and, therefore, differed from the preceding melodic context, which contained diatonic (in-key) pitches only, in terms of both sensory and schematic characteristics. The altered pitches maintained the melodic contour of the original melody and were close to (within a major third of) the original target pitch. The altered pitches were placed in one of eight different quarter-note locations and never occurred on the first three pitches or the last pitch of a melody. The altered pitches were aligned equally often to weakly accented metrical beats and to strongly accented beats, as determined by a four-tier metrical hierarchy (Lerdahl & Jackendoff, 1983). Finally, altered pitches were designed to be produced by the same right-hand finger that was used to produce the original target pitch during learning, and original and altered pitches were distributed across fingers within the right hand. The sets of melodies assigned to the two learning conditions were matched on each of these features.

### Equipment

Participants completed the experiment in a sound- and electrically attenuated chamber while EEG was recorded, and melodies were presented over EEG-compatible air delivery

headphones (ER-2 TubePhones, Etymotic Research, Inc., Elk Grove Village, IL). During both learning conditions, pianists heard the melodies with Cubase HALion One piano timbre. During the production learning condition, pianists performed the melodies on a Roland RD-700NX keyboard. During the memory recognition test, EEG was recorded with 64 Ag/AgCl electrodes configured according to the international 10–20 system with a BioSemi ActiveTwo system at a resolution of 24 bits and a sampling rate of 1024 Hz (BioSemi, Inc., Amsterdam, The Netherlands). Participants' eyes remained open during EEG recording. Electrodes below and above the right eye monitored vertical eye movements, and two electrodes placed adjacent to the outer canthi of the eyes monitored horizontal eye movements.

## Design

The study used a repeated-measures Learning (perception/production)  $\times$  Target (altered/original pitches) within-participant design. Half of the participants received one set of melodies in the production learning condition and the other set of melodies in the perception learning condition, whereas the other half of participants received the reverse melody-to-learning-condition assignment. The order of the perception and production learning conditions was counterbalanced across participants. In the memory recognition test, the entire set of melodies was presented over five blocks. Within each block, each of the 12 learned melodies was presented once in its original form and once in its altered form, with order of melodies randomized within each block. Each altered pitch occurred only once at a given serial position within the melodic context; thus, each altered melody was unique within the context of the experiment and was therefore heard only once by each participant over the course of the experiment. This resulted in 30 (6 melodies  $\times$  5 blocks) recognition trials per experimental condition (perception/production learning  $\times$  original/altered target pitch), yielding a total of 120 recognition trials.

## Procedure

Participants first completed a musical background questionnaire, followed by a piano performance sight-reading test. Participants who were able to perform a short single-hand notated melody (not used in the experiment) to a note-perfect criterion within two attempts were admitted to the experiment. All pianists who were invited to participate met this criterion. Participants were outfitted with EEG caps and electrodes after completing the sight-reading test.

### *Learning Phase*

Participants learned 12 novel melodies: six melodies in the perception learning condition and six in the production

learning condition, using the same procedure as in Mathias et al. (2015). In the perception learning condition, pianists heard 10 successive renditions of each melody over headphones. In the production learning condition, pianists performed 10 successive renditions of each melody. The musical notation for each melody remained in view during both learning conditions. Fingers used to strike piano keys were notated below the musical staff for melodies in both learning conditions; finger numbers were indicated only for tones for which there were multiple possible fingerings. Each trial in the production learning condition began with an initial metronome sounded at 500 msec per quarter-note beat (the same IOI at which perceived melodies were presented) for eight beats before the start of each performance and stopped when participants began to perform. Auditory feedback triggered by piano key presses was delivered with a piano timbre via headphones during performances. Participants were instructed before the learning conditions that their memory for the melodies would be tested following learning. The learning phase lasted approximately 35 min.

### *Memory Recognition Test*

Following the learning phase, participants were presented over headphones with the computer-generated recordings of the originally learned melodies. EEG was simultaneously recorded, and participants were asked to identify whether or not each melody contained an incorrect pitch. At the beginning of each trial, a fixation cross appeared in the center of the computer screen, and after  $2000 \pm 500$  msec, a melody was presented auditorily (melody notation was not shown during the memory recognition test). The fixation cross remained on the screen for the entire duration of the melody, and participants were instructed to fixate on the cross for the entire duration. Participants were instructed to avoid blinking and moving during the presentation of the melodies. After listening to each melody, participants indicated whether the melody contained an alteration (Yes/No). No time limit was imposed for recognition responses. Participants were told that they could blink and relax before pressing a key to proceed to the next trial. The time interval between the end of the learning phase and the start of the recognition trials was approximately 5 min.

### *Posttest*

Participants then listened to each original melody (with no altered pitches) and indicated whether they had learned the melody by listening to it or performing it in the first phase of the experiment (Listened/Performed). Each melody was presented once, in different random orders for each participant. EEG activity was not recorded during the posttest.

## Data Recording and Analysis

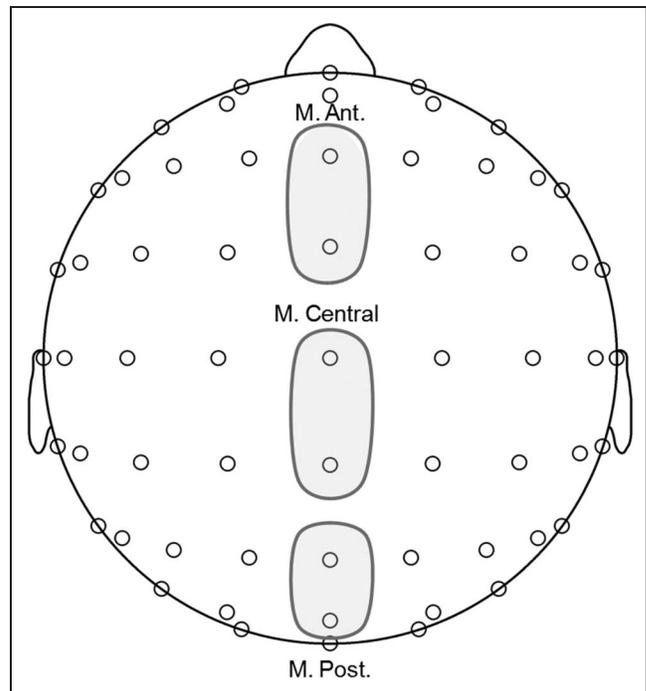
### Behavioral Data

Errors in pitch accuracy during the production learning condition were identified by computer comparison of pianists' performances with the information in the notated musical score (Large, 1993). Corrections (errors in which pianists stopped after an error and corrected) were excluded from error rate computations and analyzed separately. Mean accuracy scores in the memory recognition test were coded categorically as correct or incorrect. For each learning condition (perception, production), response sensitivity ( $d'$ ) and bias ( $c$ ) scores were computed to index participants' sensitivity and bias toward altered and original melodies. Finally, posttest data were analyzed as the proportion of melodies that were correctly identified as learned during production learning or perception learning conditions.

### EEG Data

EEG signals were analyzed using BrainVision Analyzer 2.0.2 (Brain Products GmbH, München, Germany). Electrodes were re-referenced offline to the average of all scalp electrodes. The EEG signals were bandpass-filtered between 1 and 30 Hz. Data were segmented into 600-msec epochs beginning 100 msec before the onset of the target pitch (altered pitch or contextually identical original pitch in the presented melodies) and terminating at the onset of the subsequent pitch. Artifact rejection was performed automatically using a  $\pm 40$   $\mu\text{V}$  rejection threshold at electrodes Fz, Cz, Pz, and Oz, as well as the horizontal and vertical electrooculogram, and manually by removing any trials seemingly contaminated with eye movements or muscle activity on any of the electrodes. Trials for which participants' responses were incorrect were excluded from averages, leaving nevertheless an equal number of trials between the four conditions across participants: a mean of 21.7 trials ( $SE = 1.2$ , 72.2% of total trials) in the perception-altered condition, 23.0 trials ( $SE = 1.1$ , 76.5% of total trials) in the production-altered condition, 21.2 trials ( $SE = 1.7$ , 70.5% of total trials) in the perception-original condition, and 21.0 trials ( $SE = 1.4$ , 70.0% of total trials) in the production-original condition.

**ERPs.** Average ERPs for each participant and each of the four experimental conditions were time-locked to the onset of the target pitch using EEG activity occurring up to 100 msec before the target pitch as a baseline. Mean ERP amplitudes were statistically evaluated at three topographical ROIs (see Figure 2), similar to Miranda and Ullman (2007) and Mathias et al. (2015): anterior (Fz, FCz), central (Cz, CPz), and posterior (Pz, POz). Peak amplitude latencies were identified using 100-msec non-overlapping time windows selected on the basis of previous research and visual inspection of the grand averages and calculated by averaging peak amplitude latencies



**Figure 2.** Three topographical ROIs (see Methods for more details): Midline anterior (Fz, FCz), midline central (Cz, CPz), and midline posterior (Pz, POz).

across midline electrodes. Forty-millisecond time windows for statistical analysis of ERP components were then centered on grand average peak amplitude latencies as follows: 130–170 msec (labeled N1), 210–250 msec (labeled N2), and 330–370 msec (labeled P3a).

Mean ERP component amplitudes were assessed by first determining whether effects of independent variables interacted with the scalp location factor (Mathias et al., 2015; Miranda & Ullman, 2007). ERP amplitudes were tested in repeated-measures ANOVAs with factors of Learning condition (perception, production), Target pitch (altered, original), and Scalp location (anterior, central, posterior). When significant interactions involving one or both independent variables (Learning condition, Target pitch) with the Scalp location factor occurred, follow-up ANOVAs were performed only on the ROI in which component amplitudes were statistically maximal.<sup>1</sup> Scalp topographic maps showing ERP component distributions were generated, and activity was averaged across the time window used for the analysis of each component.

**Source localization.** Standardized low-resolution brain electromagnetic tomography (sLORETA) was used to compute cortical activity (source current density in  $\mu\text{A}/\text{mm}^3$ ) corresponding to ERP components that showed effects of learning condition. The sLORETA method is a standardized discrete, three-dimensional distributed, linear, minimum norm solution to the inverse problem (Pascual-Marqui, 2002), which has been validated in several simultaneous EEG/fMRI studies (Mobascher et al., 2009; Olbrich et al.,

2009) and allows accurate localization of deep cortical structures, including the ACC (Pizzagalli et al., 2001).

In the current implementation of sLORETA, computations were made based on a realistic head model (Fuchs, Kastner, Wagner, Hawes, & Ebersole, 2002), using the MNI152 template (Mazziotta et al., 2001), with the 3-D solution space restricted to cortical gray matter, as determined by the probabilistic Talairach atlas (Lancaster et al., 2000). Standard electrode positions on the MNI152 scalp were taken from Jurcak, Tsuzuki, and Dan (2007) and Oostenveld and Praamstra (2001). The intracerebral volume was partitioned in 6239 voxels at a 5-mm spatial resolution. Thus, sLORETA images represented the standardized electric activity at each voxel in neuroanatomical Montreal Neurological Institute (MNI) space as the exact magnitude of the estimated current density. Source current densities for each participant corresponding to the recognition of memory violations in previously produced (production-altered condition) melodies and in previously perceived (perception-altered condition) melodies were compared within the time windows of N1, N2, and P3a ERP components using a voxel-wise randomization test of log  $F$  ratios. sLORETA performed 5000 permutations of the randomized statistical nonparametric mapping, and critical log  $F$  ratios and significance values were corrected for multiple comparisons. Log  $F$  ratio values for each voxel were thresholded based on a corrected significance threshold of  $p < .01$ . Brodmann's areas were identified using the MRIcro Brodmann template (Rorden, 2007, [www.mricro.com](http://www.mricro.com)), and anatomical labels were determined using the Harvard–Oxford cortical and subcortical structural atlases in FSL software ([fsl.fmrib.ox.ac.uk/fsl/](http://fsl.fmrib.ox.ac.uk/fsl/)).

### *Sensory and Cognitive Predictors for Pitch Perception*

Behavioral and ERP measures of target pitches were also compared with predictions based on sensory and cognitive approaches to pitch perception. Theoretical predictors of sensory dissonance arising from each target pitch relative to its preceding pitch context were computed from Leman's (2000) physiological model of auditory STM. Several behavioral and neurophysiological studies on tonality perception have used this model as a means of simulating sensory dissonance (Bigand, Delbé, Poulin-Charronnat, Leman, & Tillmann, 2014; Collins et al., 2014; Marmel, Tillmann, & Delbé, 2010; Koelsch et al., 2007). The acoustic stimulus recordings were first transformed into neural pitch periodicity images by simulating outer and middle ear filtering, basilar membrane resonance, and neural firing rate patterns, and pitch periodicities were analyzed using a windowed autocorrelation function. The resulting information reflected periodicities coded among auditory neurons in the 80–1250 Hz range and was used to generate two pitch images: the local pitch percept, an echoic image of pitch periodicities within a leaky-integrative span of about 0.1 sec, and a global pitch

percept, an echoic image of pitch periodicities within a leaky-integrative span of about 1.5 sec. The echoic pitch images are referred to as “local” and “global” because they include a smaller or greater number of pitch events within the leaky temporal window, respectively. The correlation coefficient between the local and global pitch images is referred to as a contextuality index and represents the amount of echoic memory-based “tension” between the local and global echoic pitch images. Higher contextuality indices (range = 0–1) indicated a better fit (less echoic memory-based tension) between the local pitch percept and its preceding global melodic context. Model simulations were run with global echo parameter values of  $T = 1.5$  sec and local echo values of  $T = 0.1$  sec based on a previous model fit, the parameter combination that best accounted for previous ratings of probe tones in scale and chord contexts (Leman, 2000). Mean ERP amplitudes were then correlated with the simulated contextual indices. To confirm that observed correlations were not dependent on the specific echo parameters, the simulations were repeated, varying the global echo from 1.0 to 4.0 in steps of 0.5 and the local echo from 0.1 to 0.4 in steps of 0.1. The results did not change across the range of global echo values, and smaller correlation values were found for local echos ranging from 0.2 to 0.4. Results for standard parameter values are reported.

Krumhansl and Kessler's (1982, Experiment 1) listener ratings of how well a pitch fit following a major or minor scale context were used as cognitive tonality predictors of target pitch perception, using the same rating scale from 1 (*fits poorly*) to 7 (*fits well*). In their study, musician listeners heard ascending major or minor scales, which were followed by individual “probe tones” from the full set of pitches (major and minor scales). Participants were asked to rate each probe tone in terms of how well, in a musical sense, the tone fit or went with the preceding tonal context.

## **RESULTS**

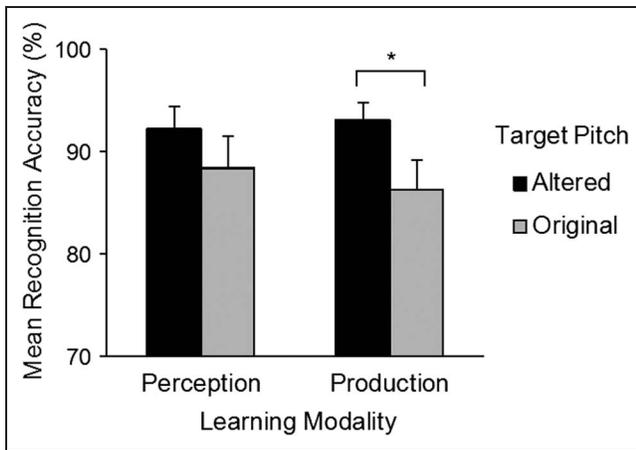
### **Behavioral Results**

#### *Learning Phase*

Pianists were highly accurate in their performances during the production learning condition: Less than 1% of tones per performance were errors ( $M$  pitch error rate per trial = .0079,  $SE = .001$ ) and 93.1% of all performances contained no errors ( $SE = 0.7\%$ ). The mean produced IOI for error-free trials was 508 msec per tone ( $SE = 1.5$  msec), confirming that pianists performed close to the prescribed rate (500 msec).

#### *Memory Recognition Test*

Percent correct responses in the memory recognition task, shown in Figure 3, were compared for melodies containing altered and original pitches. Although the ANOVA did



**Figure 3.** Mean percentage of correct responses in the altered pitch detection test by learning condition (perception/production) and target pitch (altered/original). Error bars represent 1 *SE*. \* $p < .05$ .

not yield significant main effects or interactions for Target and Learning condition, correct responses to melodies containing altered pitches were marginally greater than correct responses to melodies containing original pitches across learning conditions,  $t(19) = 1.73, p = .10$ , and correct responses to altered pitches were significantly greater than correct responses to original pitches within the production condition,  $t(19) = 2.13, p < .05$ , as expected.

Response sensitivity ( $d'$ ) and bias scores ( $c$ ) are shown in Figure 4. Whereas sensitivity did not differ between production and perception learning conditions, bias was significantly larger for the perception condition compared with the production condition,  $F(1, 19) = 5.79, p < .05$ . Bias scores differed significantly from 0 for both the production condition,  $t(19) = 5.31, p < .001$ , and for the perception condition,  $t(19) = 4.12, p = .001$ . Thus, production learning decreased participants' bias toward identifying melodies as altered, compared with perception learning.

#### Posttest

Whereas posttest accuracy for the production condition ( $M = 60.0\%$ ,  $SE = 4.1\%$ ) exceeded the level expected by chance (50%),  $t(19) = 2.23, p < .05$ , accuracy for the perception condition ( $M = 55.8\%$ ,  $SE = 5.4\%$ ) did not differ from chance. There was no significant effect of Learning condition on posttest accuracy.

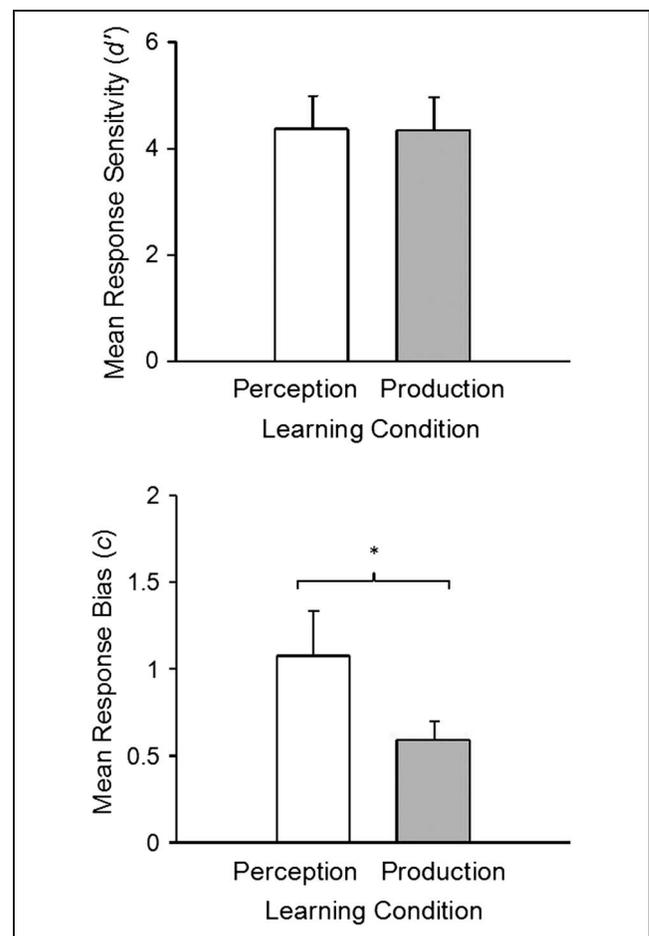
#### ERP Results

Figure 5 shows grand-averaged ERP waveforms time-locked to target pitches averaged across correct response trials. Visual inspection revealed an auditory N1 maximal around 130–170 msec elicited by both the altered and original target pitches in both learning conditions. Subsequent ERP components elicited by altered targets

included an early negative component maximal around 210–250 msec (labeled N2) and a positive component maximal around 330–370 msec (labeled P3a). Scalp topographies corresponding to time ranges for altered pitches are shown in Figure 6.

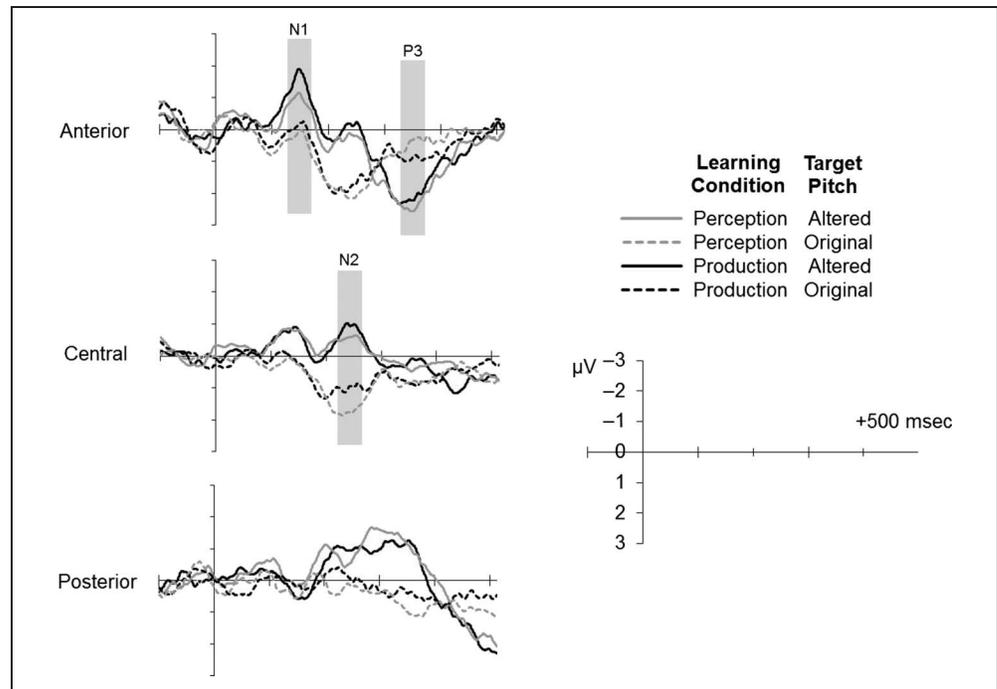
Analysis of amplitudes within the N1 time window at midline ROIs yielded a significant Target  $\times$  Scalp location interaction,  $F(2, 38) = 9.13, p < .001$ . Altered pitches elicited a larger negative potential than original pitches at the anterior ROI than at central and posterior ROIs ( $HSD = .59, \alpha = .05$ ). Analysis of mean amplitudes at only the anterior ROI (see Methods) with the factors Learning condition and Target pitch revealed a significant main effect of Target pitch,  $F(1, 19) = 17.04, p < .001$ . Altered pitches elicited a larger N1 than original pitches (see Figure 7, top). There was no effect of Learning condition and no Learning condition  $\times$  Target pitch interaction.

Analysis of amplitudes within the N2 time window at midline ROIs yielded a significant Target  $\times$  Scalp location interaction,  $F(2, 38) = 6.05, p = .005$ . Altered pitches elicited a larger negative potential than original pitches at the anterior and central ROIs ( $HSD = .56, \alpha = .05$ ).



**Figure 4.** Mean sensitivity (top) and bias (bottom) scores in the memory recognition task following perception learning and production learning. \* $p < .05$ .

**Figure 5.** Grand-averaged ERPs elicited by the four experimental conditions for trials in which participants correctly identified the presented melody as altered or original. Activity with each of the topographical ROIs is shown. Activity within each ROI is averaged across all electrodes contained within the ROI. Negative is plotted upward.



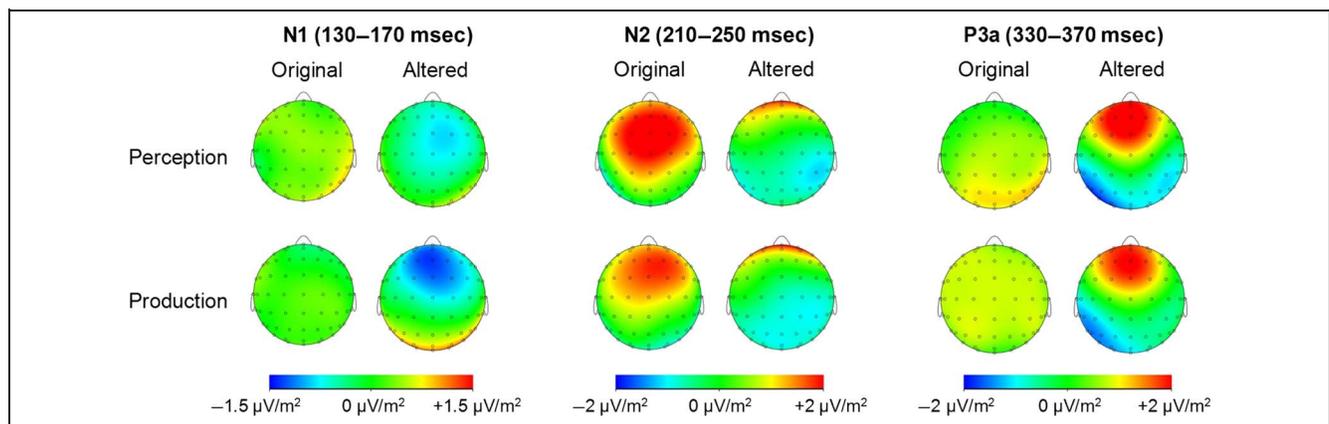
Analysis of mean amplitudes at only the anterior and central ROIs with the factors learning condition and target pitch indicated significant main effects of both learning condition,  $F(1, 19) = 6.27, p < .05$ , and target pitch,  $F(1, 19) = 21.20, p < .001$ . The amplitude of the N2 was larger (more negative) for the production learning condition than for the perception learning and larger for altered pitches than for the original pitches (Figure 7, middle). There was no significant interaction.

Analysis of amplitudes at midline ROIs within the time range of the P3a yielded a significant Target  $\times$  Scalp location interaction,  $F(2, 38) = 18.28, p < .001$ . Altered pitches elicited a larger positive potential than original pitches at the anterior ROI (HSD = .96,  $\alpha = .05$ ). Analysis of mean amplitudes at the anterior ROI with the factors

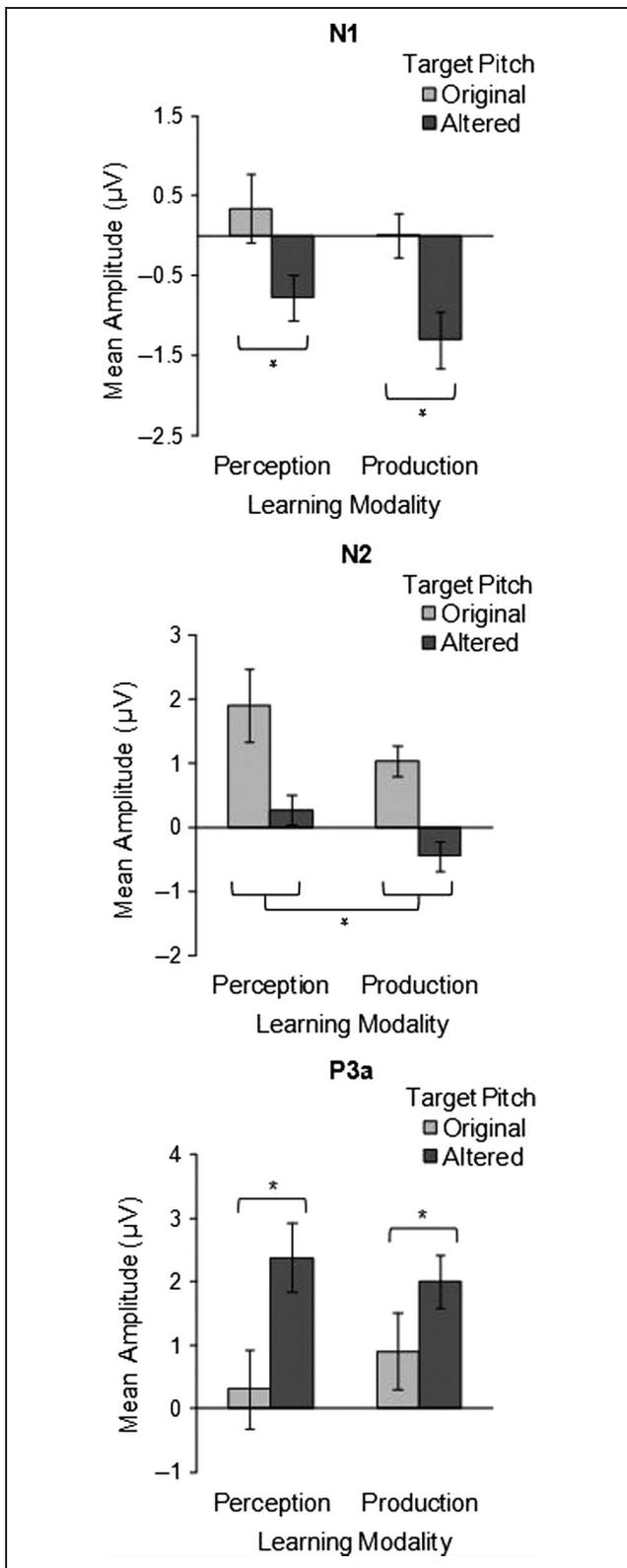
Learning condition and Target pitch indicated a significant main effect of Target pitch,  $F(1, 19) = 5.80, p < .05$ , and a significant Learning condition  $\times$  Target pitch interaction,  $F(1, 19) = 6.25, p < .05$ . A larger positivity was elicited by altered target pitches compared with original target pitches for the perception learning condition than for the production condition (Figure 7, bottom). There was no main effect of Learning condition.

### Source Localization Results

Figure 8 shows differences in source current density activity elicited by altered target pitches in previously produced melodies compared with previously perceived melodies. Differences are shown in terms of log  $F$  ratios



**Figure 6.** Voltage (in  $\mu\text{V}$ ) scalp topographies for altered pitches by learning condition (perception/production) and target pitch (altered/original). Activity averaged over 40 msec surrounding each component's grand-averaged peak is shown (N1, 130–170 msec; N2, 210–250 msec; P3a, 330–370 msec).



**Figure 7.** Mean amplitude values of correct response grand-averaged N1, N2, and P3a ERPs that were elicited by target pitches. Amplitudes from third-level ERP analysis are shown. These amplitudes were pooled across electrodes within a priori ROIs for which the component was statistically determined to be most prominent (the anterior and central ROIs for the N1, central ROI for the N2, and the anterior ROI for the P3a). \* $p < .05$ .

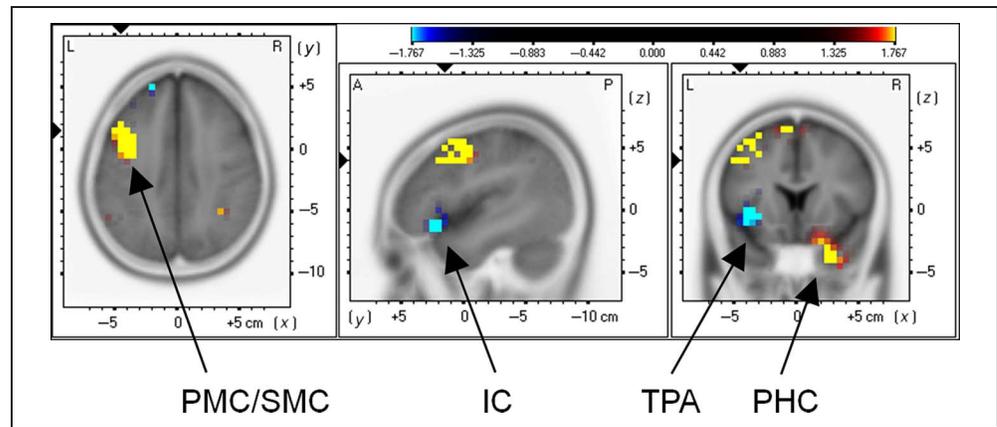
corresponding to the time range of the N2 component, the ERP response that was sensitive to learning condition. Source current density activity elicited by altered target pitches in previously produced and perceived melodies did not reveal significant differences within the time ranges of the N1 and P3a components. Brain regions showing increased activity in the production-altered condition compared with the perception-altered as well as regions showing increased activity for the perception-altered condition compared with the production-altered condition within the time range of the N2 component are listed in Table 1. Motor preparation areas in the middle frontal gyri showed stronger activation for altered target pitches in the production learning condition compared with the perception learning condition within the left hemisphere. The superior parietal lobule, parahippocampal cortex, and precuneus, as well as frontal regions including medial pFC and frontal pole, also showed stronger activation for altered target pitches in the production learning condition compared with the perception learning condition. The insular cortex, the paracingulate gyrus, and the temporopolar area were more strongly activated for altered target pitch in the perception learning condition compared with the production learning condition.

### Correlations of Sensory and Cognitive Predictions with Behavioral and ERP Components

Table 2 shows the correlations between the recognition accuracy measures, the sensory and cognitive tonality predictions, and the three ERP components elicited by altered pitches (N1, N2, and P3a mean amplitudes within their corresponding time regions). Correlations of ERP amplitudes elicited by altered pitches were evaluated because ERP responses to the memory-violating altered pitches were expected to provide insight into properties of recognition memory. Recognition accuracy scores for altered pitches were negatively correlated with mean N2 amplitudes following the production-altered condition: Participants who showed a larger N2 response following altered pitches were more accurate at detecting those altered pitches in melodies they had performed earlier (Table 2). No other component amplitudes correlated with accuracy scores in the memory task for either perception or production conditions.

Leman's (2000) auditory STM model predictions, based on contextuality indices for the altered target pitches within each melodic context, were also compared with the ERP amplitudes. Also shown in Table 2, the correlation of the contextuality index values with N1 amplitudes were positive for the perception-altered condition only. Altered target pitches characterized by a greater amount of echoic memory-based tension (smaller value indicates more echoic tension between target pitch and melodic context) elicited larger N1 amplitudes in melodies (Table 2) that were learned by perception only. No other

**Figure 8.** sLORETA images depicting brain voxels that differed in standardized current density responses to altered target pitches in previously produced versus previously perceived melodies within the time windows of the N2. Voxels that showed the largest increases in standardized current density for the production condition compared with the perception condition are indexed in yellow, and voxels showing largest increases for the perception condition compared with the production condition are indexed in blue. Brighter colors indicate larger differences in terms of statistical log *F* ratios.  $x = -45, y = 15, z = 40$ . PMC = premotor cortex; SMC = supplementary motor cortex; IC = insular cortex; TPA = temporopolar area; PHC = parahippocampal cortex.



component amplitudes correlated with Leman’s echoic STM predictions for either perception or production learning conditions.

Finally, predictions based on the tonal relatedness of target pitches to the preceding tonal context (Krumhansl & Kessler, 1982) correlated negatively with P3a amplitudes for the perception-altered condition only. Pitches that listeners judged as less related to a preceding scale context elicited larger P3a amplitudes in melodies that were learned by perception only (Table 2). This correlation was not significant for melodies learned in the production condition or for the original target pitch conditions, and the P3a was

the only component that correlated with the tonal relatedness ratings for the perception learning condition.

## DISCUSSION

We examined effects of musicians’ auditory and motor experience on the detection of pitch violations in melodies that they learned either through perception or production. The study yielded three main findings. First, production learning affected melody recognition, reflected in both a decreased bias toward identifying melodies as altered, and a larger negativity (N2) elicited by

**Table 1.** sLORETA Results: Brain Regions Showing Significantly Increased Activity during Pitch Alterations in Previously Produced Melodies Compared with Previously Perceived Melodies (Left Column) and in Previously Perceived Melodies Compared with Previously Produced Melodies (Right Column), within the N2 ERP Component Time Range

Brain Region (Brodmann’s Area)	Production–Perception		Perception–Production	
	( <i>x, y, z</i> )	log <i>F</i> ratio	( <i>x, y, z</i> )	log <i>F</i> ratio
PMC/SMC (6)	(−40, 10, 55)	<b>2.85</b>		
MFG (9)	(−40, 15, 45)	<b>2.78</b>		
PHC (36)	(20, 10, −40)	<b>1.94</b>		
SPL (40)	(35, −50, 45)	<b>1.81</b>		
ITG (20)	(35, 0, −45)	<b>1.74</b>		
FP (11)	(−5, 65, −15)	<b>1.73</b>		
Precuneus (7)	(20, −70, 35)	<b>1.54</b>		
IC			(−35, 15, 0)	<b>−2.02</b>
PCG (32)			(−10, 30, 30)	<b>−1.52</b>
TPA (38)			(−50, 20, −20)	<b>−1.47</b>

MNI coordinates of peak increases in standardized current density activity elicited by altered pitches for the production condition compared with the perception condition and for the perception condition compared with the production condition, within the time range of the N2 and **bolded** peak log *F* ratio values significant at  $p < .01$ , corrected. PMC = premotor cortex; SMC = supplementary motor cortex; MFG = middle frontal gyrus; PHC = parahippocampal cortex; SPL = superior parietal lobule; ITG = inferior temporal gyrus; FP = frontal pole; IC = insular cortex; PCG = paracingulate gyrus; TPA = temporopolar area.

**Table 2.** Correlation of Mean N1, N2, and P3a Component Amplitudes Elicited by Altered Target Pitches with Recognition Accuracy Scores (Left Column), Sensory Predictions (Leman, 2000; Middle Column), and Cognitive Predictions (Krumhansl & Kessler, 1982; Right Column)

ERP Component	Recognition Accuracy Scores		Echoic STM Predictions		Tonal Schema Predictions	
	Production	Perception	Production	Perception	Production	Perception
N1	-.33	-.20	.08	<b>.36**</b>	.06	-.21
N2	<b>-.53*</b>	-.13	.03	.24	.20	-.31
P3a	.16	.18	.01	.02	-.42	<b>-.68*</b>

Correlation values for participants' accuracy scores in the memory recognition task ( $n = 20$  participants,  $df = 18$ ), model predictions for echoic memory-based tension of target pitches ( $n = 60$  unique simulated contextual indices,  $df = 58$ ), and predictions for perception of schema-based tonality of target pitches ( $n = 9$  unique contextual relatedness ratings,  $df = 7$ ). **Bolded** values are statistically significant.

\* $p < .05$ .

\*\* $p < .01$ .

altered pitches, within previously produced melodies than in previously perceived melodies. Second, the N1 ERP component amplitude elicited by altered tones correlated with the amount of sensory tension (Leman, 2000) induced by pitch alterations in previously perceived melodies, and the P3a component amplitude correlated with the tonal stability of the altered pitch (measured using the tonal hierarchy profiles of Krumhansl & Kessler, 1982). Third, brain potentials associated with production learning, sensory memory, and schematic tonal memory followed distinct neural time courses during melody recognition. Echoic (sensory) memory processes were associated with an early neural potential (~150 msec following altered pitch onsets) in previously perceived melodies; production-based learning processes were associated with a later potential (~250 msec) in previously produced melodies; and schematic tonal memory processes were associated with later potentials (~350 msec) in previously perceived melodies. This is the first study, to our knowledge, to present evidence for influences of sensory, schematic, and production-based memories on the processing of pitches within melodies.

### Behavioral Findings

Pianists identified pitch changes in melodies following both production learning and perception learning with high accuracy; their near-ceiling performance may have arisen from the combined sensory and schematic salience of the out-of-key altered pitches. Production learning decreased participants' bias toward identifying melodies as altered. Pianists' posttest responses also revealed a production effect: Their accuracy in identifying how the pianists learned the melody exceeded chance levels following production learning, but not perception learning. Thus, pianists possessed greater knowledge regarding the modality by which they had learned a particular melody following production learning. This finding fits with previous studies on the production effect in the language

domain: Memory recognition for whether a word has been studied by producing it aloud versus silently is more accurate for words that are learned by production (Ozubko, Hourihan, & MacLeod, 2012).

### ERP Findings

#### *N2 Component Modulated by Production Learning*

Although altered pitches elicited an N2 component for melodies learned by both perception and production, the N2 amplitude was larger following production learning. A larger N2 amplitude also correlated with greater behavioral accuracy in detecting altered pitches following production learning. The N2 component has been taken to reflect the degree of mismatch between incoming auditory information and auditory information stored in memory (Folstein & Van Petten, 2008). Thus, the current findings suggest a greater mismatch of the perceived pitch alterations with production-based memory traces than with perception-based memory traces. In the production condition, activation was increased within the time frame of the N2 in motor preparation regions, that is, premotor/supplementary motor cortices as well as the superior parietal lobule. Motor preparation regions have been associated with the generation (Deiber et al., 1998), learning (Pau, Jahn, Sakreida, Domin, & Lotze, 2013), and imagery (Lotze et al., 1999) of movement sequences. Additionally, the parietal lobe may play a role in sensorimotor integration within the dorsal auditory stream (Rauschecker, 2011; Wolpert, Goodbody, & Husain, 1998) and in action understanding and simulation (Fogassi et al., 2005). The specificity of these changes in activation to the N2 component, along with the relationship between N2 amplitudes and recognition accuracy, highlights a possible role of the motor network in memory-based mismatch detection (Folstein & Van Petten, 2008).

Memory recognition following the perception learning condition was associated with increased activation in the paracingulate gyrus, a key area involved in a general

neural system of error detection (Gehring, Liu, Orr, & Carp, 2012). The anterior cingulate has been shown to contribute to the perception of deviant pitches in musical scales (Maidhof, Vavatzanidis, Prinz, Rieger, & Koelsch, 2010) and may be involved in the detection of incorrect, in-key pitches in previously learned melodies (Mathias et al., 2015). Recognition following the perception learning condition was also associated with increased activation in the insular cortex. This region has been shown to demonstrate increases in activation associated with the predictability of musical and linguistic stimuli based on a preceding context (Osnes, Hugdahl, Hjelmervik, & Specht, 2012) and to participate in auditory recognition processes (Fiebach & Schubotz, 2006; Bamiou, Musiek, & Luxon, 2003). In summary, these findings suggest that memory recognition processes can take into account learned auditory–motor associations, consistent with three sensorimotor integration frameworks: (1) motor simulation and/or prediction during auditory perception (Schubotz, 2007); (2) neural sensorimotor integration, including a mirror/echo system (Rizzolatti & Craighero, 2004); and (3) dorsal–ventral auditory stream interactions (Rauschecker, 2011).

It is possible that multiple subcomponents contributed to the negativity observed within the N2 time range, as the out-of-key altered pitches in the current study presumably violated sensory, veridical, and schematic expectations for upcoming pitch events. The N2a subcomponent of the N2 ERP, often referred to as the MMN (Patel & Azzam, 2005; Näätänen & Picton, 1986), is thought to index preattentive sensory memory mechanisms (Näätänen, Paavilainen, Rinne, & Alho, 2007), whereas N2b and N2c subcomponents are thought to be related to the violation of higher-level expectations (Folstein & Van Petten, 2008). MMN potentials are generated predominantly by auditory sensory areas (Koelsch, 2009), as well as the inferior frontal gyrus (Schönwiesner et al., 2007). The N2 component has a similar time course to the early right anterior negativity, which is sensitive to syntactic chord violations in musical sequences and consists of two subcomponents: N125 and N180. The N125 may be related to the relationship of individual tones to a preceding auditory context, and the N180 is related to the syntactic processing of tonal chord functions (Koelsch & Jentschke, 2010). The use of single-tone melodies in the current study likely led to the single early negative peak within the 100–200 msec time range, instead of the two N125 and N180 peaks observed previously with musical chords (Koelsch & Jentschke, 2010). In summary, although multiple negative ERP components are known to peak within an early 100–250 msec time range, the cognitive mechanisms underlying these components may differ (Koelsch, 2009). Although future studies will continue to explore the relationship between sensory, schematic, and syntactic expectations in music, the current findings suggest that production experience can influence expectations for upcoming pitches during auditory perception, even when the pitches also violate sensory and schematic expectations.

The current production learning effects on memory recognition replicate and extend findings of Mathias et al. (2015), who showed increased N2 amplitudes and motor source current densities elicited by in-key pitch violations following production learning. However, no change in ERP responses to pitch violations compared with original pitches was observed before the time frame of the N2. This might be due to the fact that the introduced in-key pitch alterations closely resembled the surrounding melodic context in terms of sensory and schematic content. The current study extends the earlier findings to demonstrate that the same neural signatures of production experience are elicited by pitches that are highly dissimilar to the melodic context and that production experience may modulate earlier (N1) and later (P3a) ERP components associated with sensory and schematic aspects of pitch memory. These auditory–motor effects on altered pitch processing are also related to the finding that pitch violations in chord sequences can be communicated in the visual domain via action observation when viewers have already established visual–motor associations for those chord sequences (Sammler, Novembre, Koelsch, & Keller, 2013).

#### *Early Time Course of Sensory Memory Predictions*

Enlarged N1 components were elicited following altered pitches in both the perception and production learning conditions. Amplitudes of the enlarged N1 component correlated with measures of sensory memory (Leman, 2000) for the perception learning condition only. Although previous studies have used Leman's (2000) model to simulate sensory dissonance (Bigand et al., 2014; Collins et al., 2014; Marmel et al., 2010; Koelsch et al., 2007), the current study is the first to fit the model to ERP data. The auditory N1 is a preattentive sensory component generated on the supratemporal plane of auditory cortex (Näätänen & Picton, 1987) and has been shown to be sensitive to the relationship of a tone to a preceding tone context and to tonal stability (Marmel et al., 2011; Krohn et al., 2007). Enhanced N1 amplitudes following altered pitches in the current study suggest that auditory cortices may engage in comparison of incoming pitch information with frequency information in immediate echoic memory. That is, the auditory cortex may detect violations of sensory memory at an early sensory stage of pitch processing.

#### *Later Time Course of Schematic Tonal Predictions*

P3 component amplitudes were elicited following altered pitches in perception and production learning conditions. The scalp distribution of the positive component observed in the current study is consistent with that of the “novelty P3a” subcomponent, rather than the response-related P3b, which is maximal at posterior electrode sites at later latencies (Friedman, Cycowicz, & Gaeta, 2001). P3a amplitudes were negatively correlated with the tonal stability of the

altered pitch, as measured using the probe-tone profiles (Krumhansl & Kessler, 1982) for the perception learning condition only. The P3a component has been shown to index the degree of tonal relatedness of pitches in musical sequences to a preceding context (Regnault et al., 2001; Besson & Ffita, 1995; Janata, 1995). The P3a is also thought to coincide with the cognitive evaluation of a perceived stimulus (Rinne, Särkkä, Degerman, Schröger, & Alho, 2006), shifts of attention toward the unexpected stimulus (Schröger & Wolff, 1998), and novelty processing (Polich, 2007).

Interestingly, the correspondence between ERP amplitudes and the sensory and schematic memory predictions in response to altered pitches held for the perception learning condition only. Perhaps production learning reduces the extent to which performers rely on sensory and schematic processes to identify previously learned music. The enhanced memory that listeners exhibited for previously produced melodies may have decreased their neural sensitivity to schematic pitch alterations while increasing veridical knowledge for intended pitches. Performers may rely on veridical knowledge or production-based cues such as memory for physical movements, an interpretation supported by increased motor activation regions during listening to production-learning melodies. This hypothesis is consistent with studies showing that higher motor imagery abilities predict musicians' enhanced auditory recognition of previously learned music (Brown & Palmer, 2012) and that auditory feedback is not essential for successful performance of music (Finney, 1997; Gates & Bradshaw, 1974).

It might be argued that listeners could have used a strategy of detecting out-of-key pitches to complete the experimental task and could therefore accomplish the task without prior knowledge of specific melodies. This is unlikely, however, because participants in the current study were not instructed that the task be accomplished just by detecting out-of-key pitches; they may have thought that some melodies contained in-key changes, which invoke tonal knowledge, evidenced by RTs in priming tasks (Marmel et al., 2010; Marmel & Tillmann, 2009) and by ERP responses that are sensitive to diatonic scale degree (Marmel et al., 2011; Krohn et al., 2007; Poulin-Charronnat, Bigand, & Koelsch, 2006). Furthermore, behavioral and neural differences observed during the recognition task suggest that the psychological and neural processes used to detect the altered pitches differed for melodies learned by perception compared with production. In particular, the reduction in response bias following production learning compared with perception learning suggests that the type of learning, rather than the type of expected pitch change, accounted for participants' responses. The altered out-of-key pitches (which were chosen to be produced by similar finger movements to those that produced original pitches) could be easily integrated, based on movement similarity with original pitches, within the melodic context. Thus, motor activations during recognition of previously produced melodies may have engaged similar neural networks in response to out-of-key pitch

changes as to the original pitches. This interpretation is consistent with studies showing that movements in novel situations are affected by prior motor learning (Malfait, Gribble, & Ostry, 2005; Goodbody & Wolpert, 1998).

## Conclusions

The current study provides new evidence for effects of production learning on memory recognition for pitch sequences, including both low-level sensory and higher-level cognitive influences. Production experience modified electrophysiological responses to pitch alterations in previously performed music within the time frame of the N2 ERP component. Cortical regions for motor planning showed greater involvement in processing pitch alterations in previously produced melodies compared with perceived melodies within the N2 time frame, and larger N2 responses corresponded to enhanced detection accuracy following production. Correlations of N2 amplitudes with recognition accuracy and of N1 and P3a amplitudes with measures of sensory and schematic pitch memory, respectively, underscore the benefit of ERPs in investigating memory-based predictive processing during pitch perception. These findings also support the use of music, a complex auditory stimulus whose organization is based on both low-level acoustic features and higher-level relationships between tones, for experimental manipulation and testing of listeners' memory and expectations. In summary, memory recognition for sound sequences may rely not only on sensory and cognitive auditory information but also on the sensorimotor modality through which the sequence was encoded.

## Acknowledgments

We thank Sasha Ilnyckyj and Frances Spidle of the Sequence Production Lab for their assistance and William Gehring and Fabien Perrin for helpful comments. This work was funded by a National Science Foundation Graduate Research Fellowship and ERASMUS MUNDUS Auditory Cognitive Neuroscience exchange grant to B. M., Centre National de la Recherche Scientifique (CNRS UMR5292) to B. T., and a Canada Research Chair and Natural Sciences and Engineering Research Council of Canada grant 298173 to C. P.

Reprint requests should be sent to Caroline Palmer, Department of Psychology, McGill University, 1205 Dr. Penfield Ave., Montreal QC H3A 1B1 Canada, or via e-mail: caroline.palmer@mcgill.ca.

## Note

1. As in Mathias et al. (2015), we calculated mean ERP amplitudes for midline and lateral ROIs; we report here results for the midline ROIs only, as ERP voltages were larger at midline ROIs compared with lateral ROIs for all components.

## REFERENCES

- Bamiou, D. E., Musiek, F. E., & Luxon, L. M. (2003). The insula (Island of Reil) and its role in auditory processing: Literature review. *Brain Research Reviews*, 42, 143–154.

- Besson, M., & Faita, F. (1995). An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1278–1296.
- Bharucha, J. J. (1987). Music cognition and perceptual facilitation: A connectionist framework. *Music Perception*, *5*, 1–30.
- Bigand, E., Delbé, C., Poulin-Charronnat, B., Leman, M., & Tillmann, B. (2014). Empirical evidence for musical syntax processing? Computer simulations reveal the contribution of auditory short-term memory. *Frontiers in Systems Neuroscience*, *8*, 1–27.
- Bigand, E., Poulin, B., Tillmann, B., Madurell, F., & D'Adamo, D. A. (2003). Sensory versus cognitive components in harmonic priming. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 159.
- 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.
- Brown, R. M., Chen, J. L., Hollinger, A., Penhune, V. B., Palmer, C., & Zatorre, R. J. (2013). Repetition suppression in auditory–motor regions to pitch and temporal structure in music. *Journal of Cognitive Neuroscience*, *25*, 313–328.
- Brown, R. M., & Palmer, C. (2012). Auditory–motor learning influences auditory memory for music. *Memory & Cognition*, *40*, 567–578.
- Collins, T., Tillmann, B., Barrett, F. S., Delbé, C., & Janata, P. (2014). A combined model of sensory and cognitive representations underlying tonal expectations in music: From audio signals to behavior. *Psychological Review*, *121*, 33–65.
- Deiber, M. P., Ibanez, V., Honda, M., Sadato, N., Raman, R., & Hallett, M. (1998). Cerebral processes related to visuomotor imagery and generation of simple finger movements studied with positron emission tomography. *Neuroimage*, *7*, 73–85.
- Dodson, C. S., & Schacter, D. L. (2001). “If I had said it I would have remembered it”: Reducing false memories with a distinctiveness heuristic. *Psychonomic Bulletin & Review*, *8*, 155–161.
- Donchin, E., & Coles, M. G. (1988). Is the P300 component a manifestation of context updating?. *Behavioral and Brain Sciences*, *11*, 357–374.
- Draganski, B., & May, A. (2008). Training-induced structural changes in the adult human brain. *Behavioural Brain Research*, *192*, 137–142.
- Fiebach, C. J., & Schubotz, R. I. (2006). Dynamic anticipatory processing of hierarchical sequential events: A common role for Broca’s area and ventral premotor cortex across domains?. *Cortex*, *42*, 499–502.
- Finney, S. A. (1997). Auditory feedback and musical keyboard performance. *Music Perception*, *15*, 153–174.
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F., & Rizzolatti, G. (2005). Parietal lobe: From action organization to intention understanding. *Science*, *308*, 662–667.
- 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.
- Friedman, D., Cycowicz, Y. M., & Gaeta, H. (2001). The novelty P3: An event-related brain potential (ERP) sign of the brain’s evaluation of novelty. *Neuroscience & Biobehavioral Reviews*, *25*, 355–373.
- Fuchs, M., Kastner, J., Wagner, M., Hawes, S., & Ebersole, J. S. (2002). A standardized boundary element method volume conductor model. *Clinical Neurophysiology*, *113*, 702–712.
- Gates, A., & Bradshaw, J. L. (1974). Effects of auditory feedback on a musical performance task. *Perception & Psychophysics*, *16*, 105–109.
- Gathercole, S. E., & Conway, M. A. (1988). Exploring long-term modality effects: Vocalization leads to best retention. *Memory & Cognition*, *16*, 110–119.
- Gehring, W. J., Liu, Y., Orr, J. M., & Carp, J. (2012). The error-related negativity (ERN/Ne). In S. J. Luck & E. S. Kappenman (Eds.), *Oxford handbook of event-related potential components* (pp. 231–291). Oxford: Oxford University Press.
- Goodbody, S. J., & Wolpert, D. M. (1998). Temporal and amplitude generalization in motor learning. *Journal of Neurophysiology*, *79*, 1825–1838.
- Hannon, E. E., & Trainor, L. J. (2007). Music acquisition: Effects of enculturation and formal training on development. *Trends in Cognitive Sciences*, *11*, 466–472.
- Haslinger, B., Erhard, P., Altenmüller, E., Schroeder, U., Boecker, H., & Ceballos-Baumann, A. O. (2005). Transmodal sensorimotor networks during action observation in professional pianists. *Journal of Cognitive Neuroscience*, *17*, 282–293.
- Huron, D., & Parncutt, R. (1993). An improved model of tonality perception incorporating pitch salience and echoic memory. *Psychomusicology: Music, Mind & Brain*, *12*, 154–171.
- Janata, P. (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*, *7*, 153–164.
- Jurcak, V., Tsuzuki, D., & Dan, I. (2007). 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. *Neuroimage*, *34*, 1600–1611.
- Koelsch, S. (2009). Music syntactic processing and auditory memory: Similarities and differences between ERAN and MMN. *Psychophysiology*, *46*, 179–190.
- Koelsch, S., & Jentschke, S. (2010). Differences in electric brain responses to melodies and chords. *Journal of Cognitive Neuroscience*, *22*, 2251–2262.
- Koelsch, S., Jentschke, S., Sammler, D., & Mietschen, D. (2007). Untangling syntactic and sensory processing: An ERP study of music perception. *Psychophysiology*, *44*, 476–490.
- Krohn, K. I., Brattico, E., Välimäki, V., & Tervaniemi, M. (2007). Neural representations of the hierarchical scale pitch structure. *Music Perception*, *24*, 281–296.
- Krumhansl, C. L., & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, *89*, 334–368.
- Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience*, *27*, 308–314.
- Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., et al. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, *10*, 120–131.
- Lappe, C., Herholz, S. C., Trainor, L. J., & Pantev, C. (2008). Cortical plasticity induced by short-term unimodal and multimodal musical training. *Journal of Neuroscience*, *28*, 9632–9639.
- Large, E. W. (1993). Dynamic programming for the analysis of serial behaviors. *Behavior Research Methods, Instruments, & Computers*, *25*, 238–241.
- Leman, M. (2000). An auditory model of the role of short-term memory in probe-tone ratings. *Music Perception*, *17*, 481–509.
- Lerdahl, F., & Jackendoff, R. (1983). An overview of hierarchical structure in music. *Music Perception*, *2*, 229–252.

- Lieberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.
- Lotze, M., Montoya, P., Erb, M., Hülsmann, E., Flor, H., Klose, U., et al. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: An fMRI study. *Journal of Cognitive Neuroscience*, *11*, 491–501.
- Loui, P., Wu, E. H., Wessel, D. L., & Knight, R. T. (2009). A generalized mechanism for perception of pitch patterns. *Journal of Neuroscience*, *29*, 454–459.
- MacDonald, P. A., & MacLeod, C. M. (1998). The influence of attention at encoding on direct and indirect remembering. *Acta Psychologica*, *98*, 291–310.
- MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 671–685.
- 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.
- Malfait, N., Gribble, P. L., & Ostry, D. J. (2005). Generalization of motor learning based on multiple field exposures and local adaptation. *Journal of Neurophysiology*, *93*, 3327–3338.
- Marmel, F., Perrin, F., & Tillmann, B. (2011). Tonal expectations influence early pitch processing. *Journal of Cognitive Neuroscience*, *23*, 3095–3104.
- Marmel, F., & Tillmann, B. (2009). Tonal priming beyond tonics. *Music Perception*, *26*, 211–221.
- Marmel, F., Tillmann, B., & Delbé, C. (2010). Priming in melody perception: Tracking down the strength of cognitive expectations. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 1016.
- Mathias, B., Palmer, C., Perrin, F., & Tillmann, B. (2015). Sensorimotor learning enhances expectations during auditory perception. *Cerebral Cortex*, *25*, 2238–2254.
- Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., et al. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *356*, 1293–1322.
- Miranda, R. A., & Ullman, M. T. (2007). Double dissociation between rules and memory in music: An event-related potential study. *Neuroimage*, *38*, 331–345.
- Mobascher, A., Brinkmeyer, J., Warbrick, T., Musso, F., Wittsack, H. J., Stoermer, R., et al. (2009). Fluctuations in electrodermal activity reveal variations in single trial brain responses to painful laser stimuli: A fMRI/EEG study. *Neuroimage*, *44*, 1081–1092.
- 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*, 2544–2590.
- Näätänen, R., & Picton, T. W. (1986). N2 and automatic versus controlled processes. *Electroencephalography and Clinical Neurophysiology Supplement*, *38*, 169–186.
- 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.
- Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, *125*, 826–859.
- Olbrich, S., Mulert, C., Karch, S., Trenner, M., Leicht, G., Pogarell, O., et al. (2009). EEG-vigilance and BOLD effect during simultaneous EEG/fMRI measurement. *Neuroimage*, *45*, 319–332.
- Oostenveld, R., & Praamstra, P. (2001). The five percent electrode system for high-resolution EEG and ERP measurements. *Clinical Neurophysiology*, *112*, 713–719.
- Osnes, B., Hugdahl, K., Hjelmervik, H., & Specht, K. (2012). Stimulus expectancy modulates inferior frontal gyrus and premotor cortex activity in auditory perception. *Brain and Language*, *121*, 65–69.
- Ozubko, J. D., Hourihan, K. L., & MacLeod, C. M. (2012). Production benefits learning: The production effect endures and improves memory for text. *Memory*, *20*, 717–727.
- 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 & Findings in Experimental & Clinical Pharmacology*, *24D*, 5–12.
- Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: Selected studies of the event-related potential. *International Journal of Medical Sciences*, *2*, 147–154.
- Pau, S., Jahn, G., Sakreida, K., Domin, M., & Lotze, M. (2013). Encoding and recall of finger sequences in experienced pianists compared with musically naive controls: A combined behavioral and functional imaging study. *Neuroimage*, *64*, 379–387.
- Pizzagalli, D., Pascual-Marqui, R. D., Nitschke, J. B., Oakes, T. R., Larson, C. L., Abercrombie, H. C., et al. (2001). Anterior cingulate activity as a predictor of degree of treatment response in major depression: Evidence from brain electrical tomography analysis. *American Journal of Psychiatry*, *158*, 405–415.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology*, *118*, 2128–2148.
- Poulin-Charronnat, B., Bigand, E., & Koelsch, S. (2006). Processing of musical syntax tonic versus subdominant: An event-related potential study. *Journal of Cognitive Neuroscience*, *18*, 1545–1554.
- Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing Research*, *271*, 16–25.
- Regnault, P., Bigand, E., & Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: Evidence from auditory event-related brain potentials. *Journal of Cognitive Neuroscience*, *13*, 241–255.
- Rinne, T., Särkkä, A., Degerman, A., Schröger, E., & Alho, K. (2006). Two separate mechanisms underlie auditory change detection and involuntary control of attention. *Brain Research*, *1077*, 135–143.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, *27*, 169–192.
- Rorden, C. (2007). MRICroN.
- Rugg, M. D., & Coles, M. G. (1995). *Electrophysiology of mind: Event-related brain potentials and cognition*. New York: Oxford University Press.
- Sammler, D., Novembre, G., Koelsch, S., & Keller, P. E. (2013). Syntax in a pianist's hand: ERP signatures of “embodied” syntax processing in music. *Cortex*, *49*, 1325–1339.
- Schönwiesner, M., Novitski, N., Pakarinen, S., Carlson, S., Tervaniemi, M., & Näätänen, R. (2007). Heschl's gyrus, posterior superior temporal gyrus, and mid-ventrolateral prefrontal cortex have different roles in the detection of acoustic changes. *Journal of Neurophysiology*, *97*, 2075–2082.
- Schröger, E., & Wolff, C. (1998). Behavioral and electrophysiological effects of task-irrelevant sound change: A new distraction paradigm. *Cognitive Brain Research*, *7*, 71–87.
- Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences*, *11*, 211–218.

- Tekman, H. G., & Bharucha, J. J. (1998). Implicit knowledge versus psychoacoustic similarity in priming of chords. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 252–260.
- Tervaniemi, M., Huotllainen, M., Bratiico, E., Ilmoniemi, R. J., Reinkainen, K., & Alho, K. (2003). Event-related potentials to expectancy violation in musical context. *Musicae Scientiae*, *7*, 241–261.
- Tillmann, B., Bharucha, J. J., & Bigand, E. (2000). Implicit learning of tonality: A self-organizing approach. *Psychological Review*, *107*, 885–913.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*, 460–473.
- Wolpert, D. M., Goodbody, S. J., & Husain, M. (1998). Maintaining internal representations: The role of the human superior parietal lobe. *Nature Neuroscience*, *1*, 529–533.
- 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.