

Accepted Manuscript

Auditory prediction cues motor preparation in the absence of movements

Marianne A. Stephan, Carlotta Lega, Virginia Penhune

PII: S1053-8119(18)30254-4

DOI: [10.1016/j.neuroimage.2018.03.044](https://doi.org/10.1016/j.neuroimage.2018.03.044)

Reference: YNIMG 14813

To appear in: *NeuroImage*

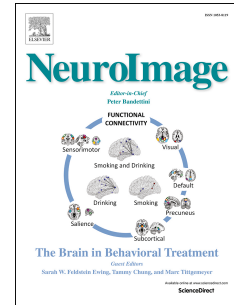
Received Date: 8 September 2017

Revised Date: 28 February 2018

Accepted Date: 18 March 2018

Please cite this article as: Stephan, M.A., Lega, C., Penhune, V., Auditory prediction cues motor preparation in the absence of movements, *NeuroImage* (2018), doi: [10.1016/j.neuroimage.2018.03.044](https://doi.org/10.1016/j.neuroimage.2018.03.044).

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Auditory prediction cues motor preparation in the absence of movements

Marianne A. Stephan^{1,2}, Carlotta Lega³, Virginia Penhune^{1,2}

¹*Department of Psychology, Concordia University, Montreal, QC, Canada H4B 1R6*

²*International Laboratory for Brain, Music and Sound Research (BRAMS), University of Montreal, Montreal, QC, Canada H2V 2J2*

³*Department of Psychology, University of Milano-Bicocca, Milan, Italy [include postal code]*

*Correspondence:

Virginia Penhune

Department of Psychology, SP-A 244

Concordia University

7141 Sherbrooke Street West

Montreal, QC, Canada H4B 1R6

Acknowledgments:

This research was supported by the Swiss Foundation for Grants in Biology and Medicine (SFGBM) and the Swiss National Science Foundation (SNSF) to MS, and the National Science and Engineering Research Council of Canada to VP (2015-04225). We thank Joseph Thibodeau and Bernard Bouchard for their technical assistance. The authors declare no competing financial interests.

Abstract

There is increasing evidence for integrated representation of sensory and motor information in the brain, and that seeing or hearing action-related stimuli may automatically cue the movements required to respond to or produce them. In this study we tested whether anticipation of tones in a known melody automatically activates corresponding motor representations in a predictive way, in preparation for potential upcoming movements. Therefore, we trained 20 non-musicians (8 men, 12 women) to play a simple melody. Then, while they passively listened to the learned or unlearned melodies, we applied single pulse transcranial magnetic stimulation (TMS) over M1 to measure motor evoked potentials from the associated finger muscle either preceding or following the onset of individual tones. Our results show that listening to the learned melody increased corticospinal excitability for specific finger muscles before tone onset. This demonstrates that predictable auditory information can activate motor representations in an anticipatory muscle-specific manner, even in the absence of intention to move. This suggests that the motor system is involved in the prediction of sensory events, likely based on auditory-parietal-prefrontal feedforward/feedback loops that automatically prepare predictable sound-related actions independent of actual execution and the associated auditory feedback. Overall, we propose that multimodal forward models of upcoming sounds and actions support motor preparation, facilitate error detection and correction, and guide perception.

1. Introduction

Predicting sequences of sensory events from the environment plays an important role in our everyday life. Hearing the sounds of approaching footsteps can make us turn around, or, for a concert pianist, imagining the upcoming notes of a frequently played piece can cue the movements required to play them. These phenomena suggest that because there is an integrated representation of sound and action in the brain the prediction of sequences of action-related sounds may automatically cue the movements required to respond to those sounds or to produce them (Schubotz & von Cramon 2002; Schubotz 2007; Keller & Koch 2006; Keller & Koch 2008).

Global evidence that melodies can cue motor responses comes from neuroimaging and neurostimulation studies showing enhanced activation of cortical and subcortical motor regions when people listen to pieces they know how to play (Haueisen & Knösche 2001; Bangert & Altenmüller 2003; Bangert et al. 2006; D'Ausilio et al. 2006; Baumann et al. 2007; Lahav et al. 2007; Lappe et al. 2008). However, none of these studies provide direct evidence that predictable melodic sequences can automatically cue specific upcoming actions. Therefore, in the current experiment we trained non-musicians to play a simple melody. Then, while they listened without playing, we used single pulse transcranial magnetic stimulation (TMS) to test whether hearing the tones of a learned melody automatically activated corresponding motor representations of the associated finger muscles in advance of the specific tones.

Melodies represent a useful framework to study sensory-motor integration and prediction, because of their ordered sequential nature and the strong but flexible coupling between auditory and motor information in the brain (see Zatorre et al. 2007 or Herholz & Zatorre 2012 for review). To play a melody we need to learn which movement produces which tone, the specific order of tones and the time at which they have to be played (Chen et al. 2012). Once learned, the auditory-motor sequence allows us to predict upcoming tones and prepare the appropriate finger movements (Palmer 2005). Such auditory-motor learning has been shown to induce plastic changes within a network of brain areas (see Zatorre et al. 2007 or Herholz & Zatorre 2012 for review) including primary auditory and motor cortex, premotor cortex, and the supplementary motor area. These regions have been found to be active during passive listening to known melodies in musicians, but also in non-musicians after relatively short periods of training (Haueisen & Knösche 2001; Bangert & Altenmüller 2003; Bangert et al. 2006; D'Ausilio et al. 2006; Baumann et al. 2007; Lahav et al. 2007; Lappe et al. 2008). In pianists, who are expected to have pre-existing auditory-motor representations for their instrument, listening to familiar pieces elicited greater effector-specific responses in motor cortex (Haueisen & Knösche 2001). In non-musicians, an

electroencephalography (EEG) study 20 min of piano training induced changes in auditory-sensorimotor co-activity during passive listening to learned pieces (Bangert & Altenmüller 2003), presumably as a result of newly acquired associations between sound and action. Further, using TMS, D'Ausilio et al. (2006) found increased intra-cortical facilitation when musicians listened to a rehearsed piece as compared to a non-rehearsed piece after only 30 min of piano practice and a facilitation of corticospinal excitability after 5 days of practice.

Taken together, there is ample evidence that auditory-motor training leads to global changes in motor system engagement during listening to action-related sounds. However, a direct link between hearing individual notes in a melody and predictive, effector-specific activation of the motor system has not been shown. Thus, in the current experiment we wanted to know whether there would be a muscle-specific increase in corticospinal excitability before the onset of tones that had been previously associated with movement of that muscle - a hallmark of motor preparation (Chen et al. 1998; Chen & Hallett 1999). Using single-pulse TMS over M1 we assessed motor evoked potential (MEP) amplitudes before and after non-musicians were trained to play a simple melody. We hypothesized that MEPs would show muscle-specific increases before the onset of associated tones in a learned melody, thus demonstrating predictive auditory-motor priming.

Our findings support this hypothesis, suggesting that auditory information can activate motor representations in a muscle-specific and anticipatory manner. These findings provide direct evidence for auditory-motor predictive coupling (Schubotz 2007; Novembre & Keller 2014) and shed further light on the putative role of auditory-motor representations in analyzing and predicting sensory cues and supporting motor preparation.

2. Materials and Methods

2.1 Subjects

Twenty young, healthy participants without psychiatric or neurological disorders were tested. One participant was excluded from analysis because of too much EMG background activity. Data was thus analyzed for 19 participants (8 men, 11 women; age: mean (M) 22.8, standard deviation (SD) 4.05 years). Participants were selected to have little musical training. They had on average 0.92 (0.82) years of formal or informal musical training, which had ended on average 11.6 (6.66) years before the beginning of the study. They were all right-handed, according to the Flinders Handedness Survey (FLANDERS) (Nicholls et al. 2013). All participants were screened for any contraindications to TMS (Rossi et al. 2009) and gave written informed consent prior to their inclusion in the study. The study conformed to the principles of the Declaration of Helsinki and was approved by the ethics

committee ‘Comité d’éthique de la recherche en santé’ (CERES) of the University of Montreal and the Research Ethics Committee of Concordia University.

2.2 Melodic stimuli

The goal of this experiment was to test whether individual tones in a trained melody could enhance motor corticospinal excitability in a muscle-specific manner, even in the absence of movement. Motor excitability was assessed during passive listening by measuring the motor-evoked potentials (MEPs) elicited by single-pulse TMS in two finger muscles of the right hand: the index (first dorsal interosseus, FDI) and pinky finger (abductor digiti minimi, ADM) muscles. These two fingers were associated with specific tones, thus motor representations activated by these tones could be compared for a learned melody – where the context allows a prediction of upcoming tones – and unlearned melodies – where the context allows no prediction. Further, we tested whether motor corticospinal excitability could be enhanced predictively, before the onset of the tones. The experimental procedure and stimuli are illustrated in Figures 1 and 2.

Figure 1

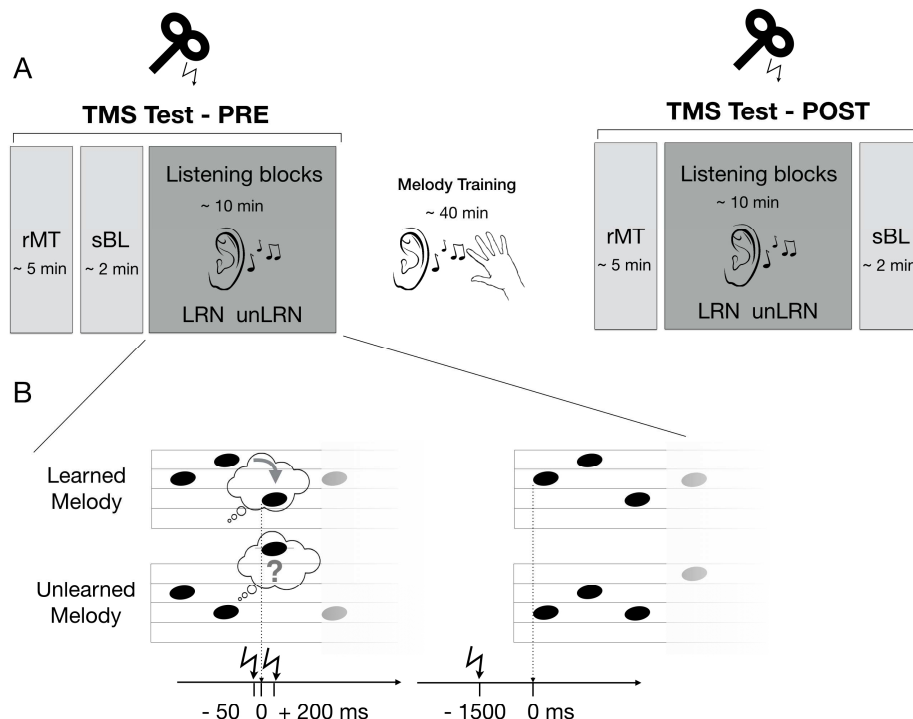


Figure 1: Illustrates the experimental protocol. Panel A illustrates the overall design of the experiment and Panel B illustrates the timing of TMS pulses relative to individual notes. rMT = resting Motor Threshold; sBL = silent Baseline assessment of corticospinal excitability; LRN = the learned melody; unLRN = unlearned melodies.

To do this we created a set of 8-tone melodies in which the four fingers of the right hand were associated with four tones on a piano-like keyboard (See Fig 2: C4 (259 Hz) = index, E4 (329 Hz) = middle; G4 (389 Hz) = ring; C5 (531 Hz) = pinky). Within these melodies, the probe tones to be tested with TMS were the lowest and highest, corresponding to the index and pinky fingers (See Figure 1). All melodies were created according to the following criteria: the same tone did not occur twice in a row, each tone was represented twice, and the number of times that the TMS probe tones followed each other in a row was minimized to avoid both muscles being activated at the same time. That is, transitions from C4 to C5 or C5 to C4 occurred only once. The Learned Melody consisted of the pattern E4 - G4 - C4 - E4 - C5 - G4 - C4 - C5. For this melody, the first two tones were not TMS probe tones (that is, not C4 or C5). For the Unlearned Melodies only the first tone could not be a probe tone and the first two tones could not be the same as the Learned Melody. In addition, each melody had to be unique.

The duration of each of the melodies was 8 s and consisted of quarter notes only. The duration of each tone was 800 ms, including 25 ms fade in and 25 ms fade out. The inter-tone interval (that is, the time between offset of the previous tone and onset of the subsequent tone) was 200 ms. In order to prevent participants from predicting the start of each presentation of the melodies, there was no cue, and a randomly varying time interval of 3.3 to 3.7 s was introduced between presentations. All tones were synthesized with Adobe Audition v.3.0 using a piano timber. Tones were presented over insertable one-way EEG earphones. In addition, to minimize disturbance due to the TMS clicks participants wore noise-cancelling earmuffs.

2.3 Tasks and procedure

The experimental procedure is illustrated in Figure 1A. Before the start of the experiment, each participant's resting motor threshold was determined (rMT-pre), followed by a baseline assessment of corticospinal excitability in silence (sBL-pre). Next, baseline corticospinal excitability was assessed while listening to two blocks of melodies (Listening-pre). These comprised: 1) 24 repetitions of a single melody (Learned Melody) and 2) 24 different predefined melodies (Unlearned Melodies). Next, during Melody Training participants practiced the Learned Melody for 15 blocks of 10 trials by playing along with the melody which was presented over earphones. The Melody Training condition included also 4 additional interleaved catch trial blocks to assess learning (See Figure 2). Catch trial blocks included either 10 trials of a single melody which was the inverse of the Learned melody (Inverted Melody) or 10 trials in which the same melody was performed but with a changed key-press to tone mapping (Switched Mapping). The 4 catch trial blocks occurred after the 1st and 2nd blocks of training and after the 13th and 14th blocks. Melody Training lasted

approximately 40 minutes, after which resting motor threshold was again determined (rMT-post). Then corticospinal excitability was re-assessed during listening to the Learned and Unlearned melodies (Listening-post). Finally, baseline corticospinal excitability was measured during silence to test for potential non-specific changes in corticospinal excitability (sBL-post). The entire experiment lasted approximately 1.5 h, with short breaks between conditions as needed.

Figure 2

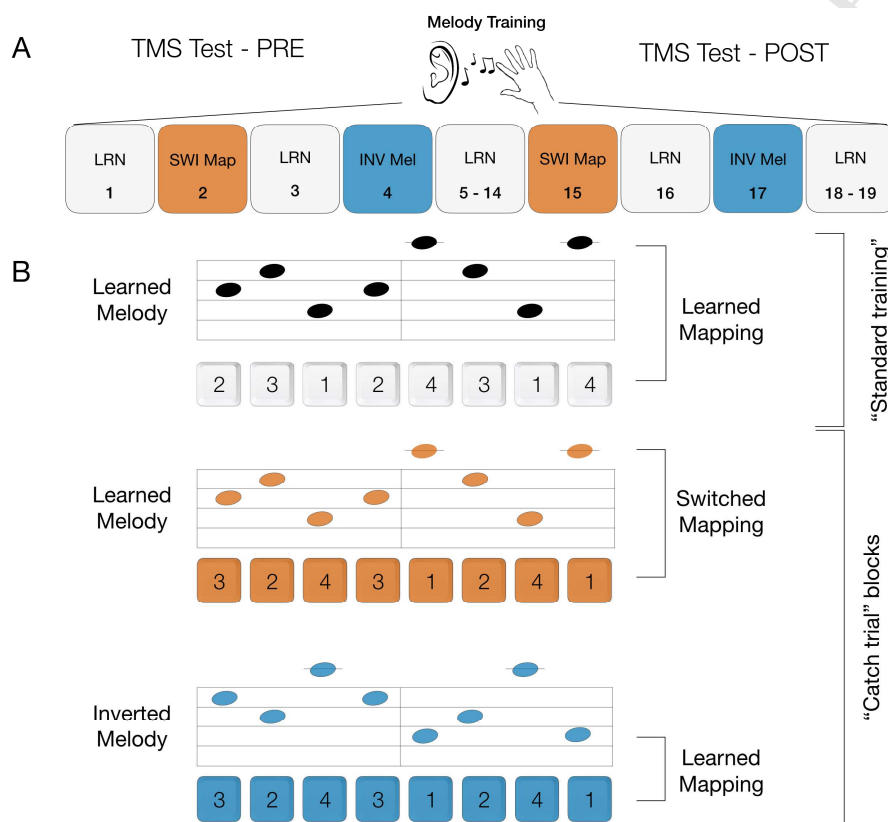


Figure 2: Panel A illustrates the order of blocks in the Melody Training phase of the experiment. Panel B illustrates the melodic stimuli and keypress responses. Learned Melody = LRN (black/white); Switched Mapping = SWI Map (orange); and Inverted Melody = INV Mel (blue). Blocks 1, 3, 5-14, 16, and 18-19 contain 10 trials each of the Learned Melody. Blocks 2, 4, 15 and 17 are catch trial blocks which contain 10 trials of either the Switched Mapping or Inverted Melody.

A laptop computer generated visual displays with instructions and stepped through the experimental procedure using a custom script written in Python. Throughout the experiment, commands were sent via USB to an embedded Beaglebone Black computer with

a Bela real-time module based on the system used by Zappi and McPherson (2014) that executed timing-critical actions: generating triggers to elicit TMS pulses, generating sequences of audio stimuli, recording key presses, and generating audio feedback for the key presses.

Figure 3



Figure 3: Panel A illustrates the keyboard interface used for the experiment. Flexible plastic bars with velcro touches are depressed to actuate switches which output TTL pulses from a parallel port connected to the Beaglebone Black micro computer which recorded the key-presses. The numbers 1-4 are for illustration only to show the relationship between the keys and the sequence to be played. Panel B shows the Learned melody and the order of keys to be played on the device.

Listening blocks

In order to investigate changes in motor system activation based on auditory predictions developed through melody learning, corticospinal excitability was assessed during Listening blocks before and after training (Listening Pre and Post, Figure 1B). During Learned Melody blocks participants listened to the single melody presented 24 times. During

the Unlearned Melodies block participants listened to 24 different melodies composed of the same four tones. Each Listening block lasted approximately 4 min. Half of the participants started with the Learned Melody block, the other half with the Unlearned Melodies block. The order of Listening blocks, as well as the melodies used, were identical before and after training.

In order to focus participants' attention during the Listening blocks on the melodies, a 'sequence detection' task was introduced. Before each Listening block, participants were instructed to listen attentively to the melodies, because they would later be asked to listen to short sequences of notes and to judge whether they had heard them or not. After each Listening block four 5-tone sequences were presented to the participants, two of which had already occurred during the preceding Listening block, and two of which were new. After each tone sequence presentation participants indicated orally to the experimenter whether they had heard it in the preceding listening block or not.

In order to avoid modulations in corticospinal excitability due to observation participants' hands were covered by a styrofoam plate fixed on the chair above the hands. Participants were instructed to relax and to refrain from movement. Movement was closely monitored based on visual observation and surface electromyography (EMG).

Melody training

Melody Training was designed to teach participants both the global mapping between key-presses and specific tones and the specific sequential order of tones in the Learned Melody tested in the Listening blocks (See Figure 2). In this condition, participants were asked to play in synchrony with the melody heard over headphones on a custom made piano-like keyboard (Figure 3). The tones for all melodies were played in a piano timber. To allow participants to distinguish the tones produced by their key-presses from the melody stimuli, the pitch of their key-presses was one octave lower than the target melody (C3 (131 Hz), E3 (165 Hz), G3 (196 Hz), C4 (261 Hz)). To familiarize participants with the tones produced by their key-presses they played each of the four tones one to two times before the start of training. The start of each melody repetition during training was cued by a brief tone with a different timber (woodwind) to help participants getting prepared for playing the subsequent melody.

In order to test learning of the Learned Melody, "catch trial" blocks were introduced at the beginning (after the 1st and 2nd blocks of training) and at the end of the training phase (after the 13th and 14th blocks). Catch trial blocks consisted of either the inverse of the learned melody (Inverted Melody) or the learned melody with a switched key-to-tone mapping (Switched Mapping). In the Inverted-Melody blocks the auditory-motor sequence became: C5-C4-G4-C5-E4-C4-G4-E4). In the Switched-Mapping blocks the tones evoked by the key

presses were inverted such that the keys produced tones in descending order from left to right (1st = C5, 2nd = G4, 3rd = E4, 4th = C4), while the target melody stayed the same. Half of the participants performed the Switched Mapping block first, the other half performed the Switched Melody block first. The order of those catch trial blocks was the same at the beginning and at the end of the training. Before each catch trial block, participants were explicitly instructed either that the target melody would be changed (Inverted Melody), or that the tones evoked by the key presses would be changed (Switched Mapping).

2.3 Transcranial magnetic stimulation and electromyography

In order to assess the effect of auditory-motor training on motor corticospinal excitability during melody listening, single-pulse TMS was applied during Listening blocks before and after training. In order to assess possible changes in corticospinal excitability that are sound-unrelated – for example due to the motor training, or changes in arousal – 12 single TMS pulses were also applied during BL blocks without sound before and after training (sBL-pre and post, see Figure 1A).

To assess motor corticospinal excitability single TMS pulses were applied at an intensity of 120 % of the resting Motor Threshold (rMT) over the ADM-FDI 'hot-spot' of the left hemisphere defined as the site where the largest motor evoked potentials (MEPs) could be evoked simultaneously in the relaxed right ADM and FDI muscles. This location was determined as follows: MEPs of 0.5 to 1 mV were first elicited at an initial estimate 5 cm lateral and 1 cm frontal to Cz. Three additional TMS pulses were then applied at each of four sites around the initial estimate, 1 cm anterior, posterior, medial, and lateral. TMS was applied through a 70 mm figure-of-eight coil, using a Super Rapid Biphasic Stimulator (Magstim, Whitland, UK) with the handle pointing 45° postero-laterally away from the midline. A TMS neuronavigation system (Brainsight 2; Rogue Research Inc., Canada) was used to ensure a constant coil position. The rMT was determined according to standard procedure using the software based 'adaptive method' developed by Awiszus (2003) (Motor Threshold Assessment Tool, MTAT, version 2.0: <http://www.clinicalresearcher.org/software>) (Groppa et al. 2012; Rossini et al. 2015). An MEP ≥ 50 μ V peak-to-peak amplitude in the higher threshold right ADM muscle was fed back to the software as a valid response (Romani et al. 2005). The rMT was assessed immediately before the Listening and sBL corticospinal excitability blocks, before as well as after auditory-motor training (see Figure 1B). For each TMS pulse, EMG recordings were obtained from the right ADM and FDI muscles, with conventional surface electrodes in a belly-tendon montage (LabChart 6.1, ADInstruments Pty Ltd, Australia). EMG was recorded from 200 ms before to 100 ms after the TMS pulse. Signals were amplified, bandpass filtered (1 Hz – 2 kHz) and sampled at a rate of 10 kHz.

During the Listening blocks, TMS pulses were applied pseudorandomly at predefined time points during melody repetitions, - 50 ms before and + 200 ms after the onset of probe tones (Figure 1B). Probe tones were the lowest and highest tones in the melodies that were associated with index and pinky finger movements, respectively. That is, there were four possible probe tones during each melody presentation. In order to determine whether motor corticospinal excitability related to probe tones was increased or decreased relative to a time point where we expected less or no motor system activation, additional TMS pulses were applied between melody presentations that is, - 1500 ms before the onset of the first tone of a melody (between melodies Baseline, bmBL). Since the time intervals between melody presentations varied randomly, melody onset was not predictable. During each Listening block 36 TMS pulses were administered, 12 pulses - 1500 ms before the first tone of a melody, 12 pulses - 50 ms before tone onset within melodies and 12 pulses + 200 ms after tone onset within melodies (6 pulses on each of the 4 possible probe tones). During the entire experiment 168 TMS pulses were administered ('hot-spot' and rMT determinations not included). TMS pulses were applied with a mean inter-pulse interval of 7.76 sec (SD: 2.09, range: 3.88 - 12.8 sec). During all TMS administrations participants were instructed to relax and to refrain from movement, which was closely monitored based on visual observation and EMG.

2.4 Analysis

Statistical analyses were performed with R (R Core Team 2014). For all outcome measures, normal distribution and homogeneity of variances were tested. When appropriate, non-parametric permutation-based analogs of the mixed factorial ANOVAs (R package 'ez', Lawrence 2013) and Wilcoxon rank sum tests were performed. For the calculation of the effect sizes we also performed parametric ANOVAs which we report in addition to the non-parametric ANOVAs.

TMS

EMG recording was performed from 200 ms before to 100 ms after each TMS pulse. Peak-to-peak MEP amplitudes were calculated from each recording. Trials with EMG background activity exceeding 50 μ V in the 200 ms preceding the TMS pulse or with a MEP amplitude smaller than 50 μ V were discarded (3.15% of the performed EMG recordings). From this dataset, MEPs with an amplitude less or greater than $Q1-1.5*IQR$ or $Q3+1.5*IQR$ respectively (Q1, first quartile; Q3, third quartile; IQR, Interquartile Range), at the single subject level, were discarded (2.97%; that is, overall, 6.03% of the originally performed EMG recordings were excluded). In order to account for sound-unrelated changes in motor corticospinal excitability, MEP amplitudes of the Listening blocks before training were

divided by the mean MEP amplitude of the corresponding participant and muscle of the sBL-pre; in the same way, MEP amplitudes of the Listening blocks after training were normalized with the MEP amplitudes of the sBL-post block after training. Those normalized MEPs were then averaged per Subject, Congruence (C, InC), Stimulation time (pre-Tone, during-Tone), and Listening block (Learned Melody, Unlearned Melodies) before and after training. In addition, the bmBL was calculated per Subject and Listening block before and after training. Congruent recordings (C) refer to MEPs recorded from muscles of the fingers that matched the associated tones (the FDI muscle when the lowest tone was presented or recordings from ADM when the highest tone was presented). Incongruent recordings (InC) refer to recordings from muscles of the fingers which did not match the tones (FDI when the highest tone was presented or from ADM when the lowest tone was presented).

Melody training performance

To assess participants' performance on the auditory-motor sequence the percentage of correct key presses as well as the mean of the absolute key-press time differences relative to target tones' onset were calculated per participant and Melody Training block. One participant didn't make any correct key presses in the first training block and was thus excluded from analyses including this first block.

To test whether participants had learned the auditory-motor sequence overall, performance at the beginning of the training phase (blocks 1-5) was compared with overall performance at the end of the training phase (blocks 14-18). In order to determine more specifically whether participants had learned the key-to-tone mapping and the serial order of the sequence, performance in "catch trial" blocks was compared with the mean performance in the adjacent two "standard training" blocks (see Figure 2A). That is, mixed factorial ANOVAs for percent correct key-presses and the absolute key-press time difference to tone onset were performed, with Time (beginning, end of training phase), Block (standard training blocks, catch trial block), and Type (Switched mapping, Switched sequence) as within-subject factors.

Correlations TMS - Auditory-motor performance

To further investigate whether there was a correlation between the strength of the learned key-to-tone mapping and the serial order of the sequence, respectively, and motor corticospinal excitability we performed correlation analyses between specific calculated motor performance measures at the end of the training phase and normalized MEP amplitudes after training. Motor performance measures were calculated as follows. As a measure for the learning of the key-to-tone mapping, performance on the Switched mapping block was subtracted from the mean performance on the adjacent two standard training blocks (e.g., $(\text{block 14} + \text{block 16})/2 - \text{block 15}$). In the same way, performance in Switched sequence

blocks was subtracted from the mean performance of the adjacent two standard training blocks as a measure for the learning of the serial order of the sequence.

Spearman's rank-order correlations were run to determine the relationship between the strength of the key-to-tone mapping and the mean MEP amplitudes of C trials during-Tone in both Listening blocks after training; moreover, between the strength of the serial order of the auditory-motor sequence and mean MEP amplitudes of C trials pre-Tone for the Learned Melody (Figure 1B). Further, we assessed with the median key-press differences to tone onset of the last 19th block whether participants anticipated the tones and whether this correlated with mean MEP amplitudes of C trials pre-Tone in the Learned Melody.

2.5 Results

TMS

ANOVA for the mean normalized MEP amplitudes revealed a significant interaction Time (Before, After Training) x Congruence (C, InC) x Stimulation time (pre-Tone, during-Tone) x Listening block (Learned Melody, Unlearned Melodies) (permutation-based ANOVA: $p = 0.026$; mixed factorial ANOVA: $F(1, 18) = 5.50$, $p = 0.031$, $\eta_G^2 = 0.003$).

After training: To test our hypothesis that, after auditory-motor training, motor corticospinal excitability pre- and during-Tone would be differentially modulated depending on whether a tone is motorically associated and predictable or not, an ANOVA with Stimulation time, Congruence, and Listening block as within-subject factors was performed for the mean normalized MEP amplitudes after training. There was a significant interaction Congruence x Stimulation time x Listening block (permutation-based ANOVA: $p = 0.026$; mixed factorial ANOVA: $F(1, 18) = 5.41$, $p = 0.032$, $\eta_G^2 = 0.008$; main effect of Congruence: $p = 0.035$; $F(1, 18) = 4.78$, $p = 0.042$, $\eta_G^2 = 0.006$) (Figure 4).

Figure 4

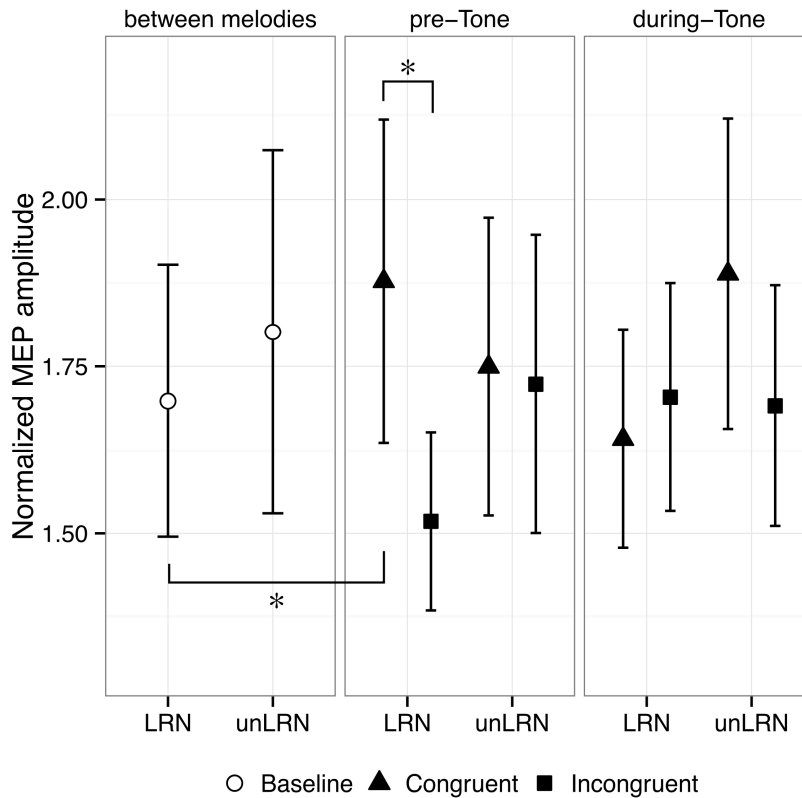


Figure 4: The graph shows the averaged normalized MEP amplitudes for the Learned (LRN) and Unlearned Melodies (unLRN) after training in the between melodies Baseline; for pre-Tone onset and post-Tone onset. The Congruent condition (triangles) included trials where the recorded muscle was that of the fingers which matched the associated tones, and the Incongruent condition (squares) included trials where the recorded muscles did not match the tones. Error bars represent the standard error of the mean (SEM).

Pre-Tone: To further explore this interaction after training, a separate ANOVA for pre-Tone excitability was performed. Results revealed a significant interaction Listening block x Congruence ($p = 0.04$; $F(1, 18) = 4.32$, $p = 0.052$, $\eta_G^2 = 0.009$; main effect of Congruence: $p = 0.026$; $F(1, 18) = 5.08$, $p = 0.037$, $\eta_G^2 = 0.012$). This interaction was due to a significant difference between C and InC trials in the Learned Melody ($W = 156$, $p = 0.012$) but not in the Unlearned Melodies ($W = 116$, $p = 0.42$) as revealed by Wilcoxon signed rank tests. To investigate whether the C and InC recording conditions in the Learned Melody represented an actual increase or decrease in corticospinal excitability they were compared to the bmBL calculated per Subject and Listening block. C trials showed increased excitability compared to bmBL ($W = 44$, $p = 0.040$), while there was no significant difference to bmBL for InC trials ($W = 119$, $p = 0.35$) (Figure 4).

During-Tone: An ANOVA for during-Tone excitability after training revealed as well a significant interaction Listening block x Congruence ($p = 0.037$; $F(1, 18) = 4.46$, $p = 0.049$,

$\eta_G^2 = 0.007$, no other main effects or interactions). This interaction was due to a trend for a difference between C and InC trials in the Unlearned Melodies ($W = 142$, $p = 0.060$) but not in the Learned Melody ($W = 71$, $p = 0.35$). There was however no significant difference between bmBL and C ($W = 65$, $p = 0.24$) or InC trials in the Unlearned Melodies ($W = 83$, $p = 0.65$).

Before training: In order to verify that the used set of tones was not motorically associated before training, the same ANOVA was performed for the mean normalized MEP amplitudes as after training. As expected, there were no significant main effects or interactions before training (trend for an interaction Congruence x Stimulation time: $p = 0.081$; $F(1, 18) = 3.53$, $p = 0.076$, $\eta_G^2 = 0.005$). MEPs before training were thus not analyzed any further. There was no significant difference between rMTs before and after auditory-motor training ($W = 39$, $p = 0.079$, before M: 58.4, SD: 10.1; after: M: 60.1, SD: 9.91).

Melody training performance

The percent of correct key-presses as well as participants' ability to synchronize key-presses with the tones of the melody improved significantly from the beginning to the end of training as revealed by a significant main effect of Time (Figure 5, percentage correct key-presses: $p < 0.001$; $F(1, 17) = 33.3$, $p < 0.001$, $\eta_G^2 = 0.22$; absolute key-press time difference to tone onset: $p < 0.001$; $F(1, 17) = 26.2$, $p < 0.001$, $\eta_G^2 = 0.16$).

Participants learned the key-to-tone mapping and the serial order of the sequence as revealed by a significantly higher percentage of correct key-presses in standard training blocks than in adjacent catch trial blocks (Block, $p < 0.001$; $F(1, 17) = 30.0$, $p < 0.001$, $\eta_G^2 = 0.062$; Block x Time, $p = 0.045$; $F(1, 17) = 5.05$, $p = 0.038$, $\eta_G^2 = 0.020$; no other interactions or main effects). Post-hoc ANOVAs revealed a significant difference between catch trial blocks and adjacent standard training blocks before ($p < 0.001$; $F(1, 17) = 30.0$, $p < 0.001$, $\eta_G^2 = 0.084$) and after training ($p < 0.001$; $F(1, 18) = 17.1$, $p < 0.001$, $\eta_G^2 = 0.098$).

Furthermore, participants showed a lower absolute key-press time difference to tone onset in standard training blocks than in adjacent catch trial blocks (Block, $p = 0.011$; $F(1, 17) = 6.84$, $p = 0.018$, $\eta_G^2 = 0.013$; Type x Block, $p = 0.001$; $F(1, 17) = 12.2$, $p = 0.003$, $\eta_G^2 = 0.019$; Type x Time, $p = 0.012$; $F(1, 17) = 6.68$, $p = 0.019$, $\eta_G^2 = 0.015$; Type, $p = 0.002$; $F(1, 17) = 10.4$, $p = 0.005$, $\eta_G^2 = 0.055$). Post-hoc ANOVAs revealed a significant difference between catch trial blocks and adjacent standard training blocks after training regardless of whether the mapping or sequence was switched (Block, $p = 0.041$; $F(1, 18) = 4.57$, $p = 0.047$, $\eta_G^2 = 0.033$). Before training, there was no significant overall effect of Block, but a significant interaction Block x Type ($p < 0.001$, $F(1, 17) = 13.8$, $p = 0.002$, $\eta_G^2 = 0.030$) and a main effect of Type ($p < 0.001$, $F(1, 17) = 15.6$, $p = 0.001$, $\eta_G^2 = 0.088$), due to the fact that the Switched sequence interfered with performance at the beginning of training.

Figure 5

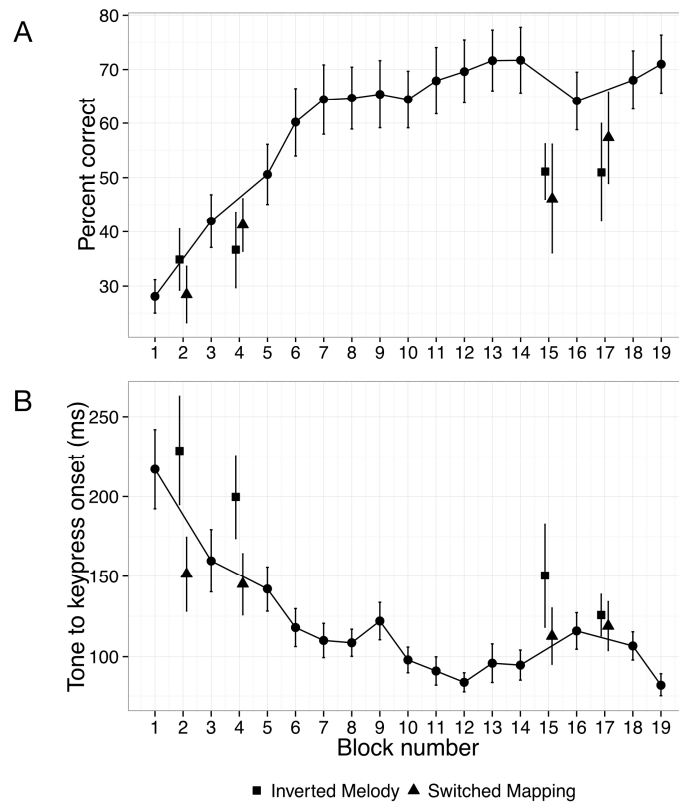


Figure 5: These graphs show the change in performance across trials for the Learned Melody (circles) compared to the Inverted Melody (squares) and Switched Mapping (triangles) conditions. Panel A shows average percent correct for each block. Panel B shows the average absolute value of the tone to keypress onset for each block.

Correlations TMS - Melody Training performance

There was no correlation between the strength of the key-to-tone mapping and the mean MEP amplitudes of C trials during-Tone after training (absolute key-press time difference to tone onset: $r_s(932) = 0.18$, $p = 0.45$; percentage correct key-presses: $r_s(1483) = -0.30$, $p = 0.21$). There was no correlation between the strength of the serial order of the sequence and mean MEP amplitudes of C trials pre-Tone in the Learned Melody block after training (absolute key-press time difference to tone onset: $r_s(818) = 0.28$, $p = 0.24$; percentage correct key-presses: $r_s(1408) = -0.24$, $p = 0.33$).

Participants anticipated the tones as revealed by key-presses occurring before the onset of tones, that is, median key-press times were significantly lower than zero in the final 19th block ($W = 47$, $p = 0.027$, 19th block: average Mdn -17.1 ms, Range -83.5 to 121 ms; 1st block: average Mdn 136 ms, Range -73.1 to 430 ms). There was no correlation between median key-press differences to tone onset of the last 19th block and mean MEP amplitudes of C trials pre-Tone in the Learned Melody block ($r_s(1154) = -0.012$, $p = 0.96$).

Discussion

Our findings demonstrate that predictable auditory information can cue motor preparation in a muscle-specific manner. We showed that after training anticipation of a tone in the learned melody automatically cues the movement that would produce that tone, as demonstrated by increased corticospinal excitability for the specific finger muscle before tone onset and in the absence of movement execution. This modulation of corticospinal excitability preceding tone onset was present after, but not before training, suggesting that 40 minutes of practice on an auditory-motor sequence is sufficient to induce a coupling between neural perception and action processes. Finally, we found a trend for muscle-specific MEP amplitudes to increase after tone onset, but only for unlearned melodies.

These findings provide neurophysiological evidence for a common neural code of action and perception, and for involvement of the motor system in prediction of sensory events (Schubotz 2007; Novembre & Keller 2014). Based on work from our lab and others, we hypothesize that during learning auditory information and motor responses are linked to form a joint auditory-motor representation, likely encoded in an auditory-parietal-motor network (Lahav et al. 2007; Chen et al. 2009; Brown et al. 2015; Herholz et al. 2016; Lega et al. 2016). Once an auditory-motor representation is established, the predicted or imagined tones in the learned melody activate motor representations, even in the absence of movement. This is consistent with the concept of forward models for speech and music that postulate that the motor system relies on joint representations of sensory targets and their motor implementation to plan upcoming actions (Tourville & Guenther 2013; Novembre & Keller 2014; Hickok & Poeppel 2015).

Our results show that MEPs elicited before tone onset were enhanced when participants listened passively to melodies that they knew how to play. This is consistent with previous brain imaging studies showing greater activity in motor networks when listening to learned melodies (Bangert et al. 2006; Baumann et al. 2007; Lahav et al. 2007; Chen et al. 2012). Importantly, however, we demonstrate for the first time that individual tones in a melody activate effector-specific motor representations even before perception of the tone occurs. Evidence for anticipatory activity in the motor system to predictable sensory information comes from single-cell recordings in monkeys showing that cells in primary motor cortex are active during anticipation of memorized sequences of reaching movements

that were previously associated with color codes (Lu & Ashe 2005). Further, EEG studies of piano playing have shown that oscillatory markers of movement error precede actual keystroke mistakes (Maidhof et al. 2009; Ruiz et al. 2009).

These results may be partly explained based on the classical feedforward model of motor control, postulating that predictions of action effects are automatically generated based on the efference copy of the ongoing motor command (Wolpert et al. 1995). Once a link between an action and its effect is learned the intended effect can be anticipated and modifies action planning (Keller & Koch 2006; Keller & Koch 2008). Further, the intended action effect can be compared with the automatically generated action effect prediction. Thereby, errors may be detected in advance of their execution (Keller & Koch 2006; Keller & Koch 2008; Maidhof et al. 2009). Models of auditory-motor integration for music and speech postulate that auditory information relevant for movement is processed in posterior auditory cortex (STG) and passed to parietal cortex where it is transformed into an auditory-motor/spatial code. This information is then passed to premotor cortex which encodes auditory-motor representations and to prefrontal regions which underwrite motor planning and sequencing (Tourville & Guenther 2013; Novembre & Keller 2014; Hickok & Poeppel 2015). This auditory-parietal-prefrontal network forms the basis for a feedforward/feedback loop that allows the movements required for speech or music to be implemented in a predictive manner without waiting for auditory feedback. Further, Schubotz (2007) has postulated that once a sensorimotor model is established, the motor system, in particular the PMC, can function in a kind of simulation mode, in which the prediction of the effect of action is automatically generated based on the efference copy of the motor command, but without actual movement execution. Interpreting our data in light of these models, we hypothesize that during the learning phase joint auditory-motor representations of the tone-to-keypress responses are formed, likely in PMC. Then, when listening to the learned melody, feedforward sensory predictions are generated involving an activation of linked motor representations. Tone-related motor representations are also activated in the context of the unlearned melody, but not as strongly, and only after the onset of the associated tone.

The notion that actions are triggered automatically by the anticipation of their sensory effects was first postulated in the *ideo-motor principle of action* (see Koch et al. 2004 or Stock & Stock 2004 for review). Evidence for this principle come from studies manipulating response-effect compatibility showing that target stimuli can bring distal effects of a response to mind, before movement is initiated, and that these associated distal effects can modulate motor performance (Keller & Koch 2006). For example, selective responses to colored stimuli at higher spatial locations are faster if they reliably trigger high-pitched tones than low-pitched tones and vice versa (Hommel et al. 2001; Keller & Koch 2006; Keller & Koch 2008).

These phenomena are consistent with our results and highlight the strong interaction between action perception and execution, further supporting the idea of a common neural code. Neurophysiological studies of the action observation or “mirror neuron” system based on single-cell recording studies in nonhuman primates (see Rizzolatti & Craighero 2004 for review) have shown that viewing or hearing another agent's actions triggers activity in pre-frontal and parietal neurons, similar to actual movement execution (di Pellegrino et al. 1992; Rizzolatti et al. 2001; Keysers et al. 2003; Fogassi et al. 2005). Indirect evidence for a similar fronto-parietal network has also been found in humans (Gazzola et al. 2006; Galati et al. 2008; Aglioti & Pazzaglia 2010), as evidenced by measures of corticospinal excitability (D’Ausilio et al. 2006), blood-oxygen-level-dependent (BOLD) signal (Bangert et al. 2006; Lahav et al. 2007), EEG potentials (Bangert & Altenmüller 2003) and MEG fields (Hauelsen & Knösche 2001). In particular, TMS studies assessing corticospinal excitability by measuring MEPs have found that mere observation or listening to the sound of an action increased MEP amplitudes recorded from the same muscles that would be active during performance (Fadiga et al. 1995; Strafella & Paus 2000; Gangitano et al. 2001; Aziz-Zadeh et al. 2004; Romani et al. 2005; Urgesi et al. 2006; Candidi et al. 2010; Urgesi et al. 2010; Ticini et al. 2012).

The action-observation network has considerable overlap with the auditory-motor network thought to be involved in speech and language production, and is thought to play a role in learning of both domains (Gazzola et al. 2006; Lahav et al. 2007; Galati et al. 2008; Chen et al. 2009; Aglioti & Pazzaglia 2010; Brown et al. 2015; Herholz et al. 2016; Lega et al. 2016). Moreover, there is evidence that frontal and premotor regions that parallel the mirror system in monkeys are engaged in sequence prediction (Maess et al. 2001; Kilner et al. 2004; Iacoboni et al. 2005) and internal simulation of sequential actions (Platel et al. 1997; Nishitani & Hari 2000; Schubotz & Von Cramon 2004). In line with those suggestions, functional magnetic resonance imaging (fMRI) studies from our lab have shown engagement of premotor regions when non-musicians passively listened to musical rhythms with no intent to move (Chen et al. 2008), pointing to a possible automatic motor involvement in analyzing and predicting temporal patterns of sensory sequences (Chen et al. 2008; Vuust & Witek 2014).

Conclusion

Our study provides evidence that predictable auditory information can activate motor representations in an anticipatory and muscle-specific manner. This may indicate an involvement of the motor system in the prediction of sensory events even in the absence of movement. This is likely based on auditory-parietal-prefrontal feedforward/feedback loops that automatically prepare predictable sound-related actions independent of actual execution

and the associated auditory feedback. Overall, we propose that multimodal forward models of upcoming sounds and actions support motor preparation, facilitate error detection and correction, and guide perception.

ACCEPTED MANUSCRIPT

References

- Aglioti, S.M. & Pazzaglia, M., 2010. Representing actions through their sound. *Exp Brain Res*, 206(2), pp.141–151.
- Awiszus, F., 2003. TMS and threshold hunting. *Suppl Clin Neurophysiol*, 56, pp.13–23.
- Aziz-Zadeh, L. et al., 2004. Left hemisphere motor facilitation in response to manual action sounds. *European Journal of Neuroscience*, 19(9), pp.2609–2612.
- Bangert, M. et al., 2006. Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage*, 30(3), pp.917–926.
- Bangert, M. & Altenmüller, E.O., 2003. Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neurosci*, 4, p.26.
- Baumann, S. et al., 2007. A network for audio-motor coordination in skilled pianists and non-musicians. *Brain Res*, 1161, pp.65–78.
- Brown, R.M., Zatorre, R.J. & Penhune, V.B., 2015. Expert music performance: Cognitive, neural, and developmental bases. *Progress in Brain Research*, 217(October), pp.57–86.
- Candidi, M. et al., 2010. Competing mechanisms for mapping action-related categorical knowledge and observed actions. *Cerebral Cortex*, 20(12), pp.2832–2841.
- Chen, J.L., Penhune, V.B. & Zatorre, R.J., 2008. Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, 18(12), pp.2844–2854.
- Chen, J.L., Penhune, V.B. & Zatorre, R.J., 2009. The role of auditory and premotor cortex in sensorimotor transformations. *Annals of the New York Academy of Sciences*, 1169, pp.15–34.
- Chen, J.L., Rae, C. & Watkins, K.E., 2012. Learning to play a melody: An fMRI study examining the formation of auditory-motor associations. *NeuroImage*, 59(2), pp.1200–1208. Available at: <http://dx.doi.org/10.1016/j.neuroimage.2011.08.012>.
- Chen, R. et al., 1998. Time course of corticospinal excitability in reaction time and self-paced movements. *Ann Neurol*, 44(3), pp.317–325. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9749597.
- Chen, R. & Hallett, M., 1999. The time course of changes in motor cortex excitability associated with voluntary movement. *Can J Neurol Sci*, 26(3), pp.163–169. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10451737.
- D'Ausilio, A. et al., 2006. Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *European Journal of Neuroscience*, 24(3), pp.955–958.
- Fadiga, L. et al., 1995. Motor facilitation during action observation: a magnetic stimulation study. *Journal of Neurophysiology*, 73(6), pp.2608–2611. Available at:

- <http://jn.physiology.org/content/73/6/2608>
<http://jn.physiology.org/content/73/6/2608.full.pdf>
<http://www.ncbi.nlm.nih.gov/pubmed/7666169>.
- Fogassi, L. et al., 2005. Parietal Lobe: From Action Organization to Intention Understanding. *Science*, 308(April), p.662–667.
- Galati, G. et al., 2008. A selective representation of the meaning of actions in the auditory mirror system. *NeuroImage*, 40(3), pp.1274–1286.
- Gangitano, M., Mottaghy, F.M. & Pascual-Leone, A., 2001. Phase-specific modulation of cortical motor output during movement observation. *Neuroreport*, 12(7), pp.1489–92. Available at:
http://journals.lww.com/neuroreport/Abstract/2001/05250/Phase_specific_modulation_of_cortical_motor_output.38.aspx
<http://www.ncbi.nlm.nih.gov/pubmed/11388435>.
- Gazzola, V., Aziz-Zadeh, L. & Keysers, C., 2006. Empathy and the somatotopic auditory mirror system in humans. *Curr Biol*, 16(18), pp.1824–1829.
- Groppa, S. et al., 2012. A practical guide to diagnostic transcranial magnetic stimulation: report of an IFCN committee. *Clin Neurophysiol*, 123(5), pp.858–882.
- Hauelsen, J. & Knösche, T.R., 2001. Involuntary motor activity in pianists evoked by music perception. *J Cogn Neurosci*, 13(6), pp.786–792.
- Herholz, S.C. et al., 2016. Dissociation of neural networks for predisposition and for training-related plasticity in auditory-motor learning. *Cerebral Cortex*, 26(7), pp.3125–3134.
- Herholz, S.C. & Zatorre, R.J., 2012. Musical Training as a Framework for Brain Plasticity: Behavior, Function, and Structure. *Neuron*, 76(3), pp.486–502. Available at:
<http://dx.doi.org/10.1016/j.neuron.2012.10.011>.
- Hickok, G. & Poeppel, D., 2015. *Neural basis of speech perception* 1st ed., Elsevier B.V. Available at: <http://dx.doi.org/10.1016/B978-0-444-62630-1.00008-1>.
- Hommel, B. et al., 2001. The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences*, 24(5), pp.849–878. Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/12239891>
http://www.journals.cambridge.org/abstract_S0140525X01000103.
- Iacoboni, M. et al., 2005. Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology*, 3(3), pp.0529–0535.
- Keller, P.E. & Koch, I., 2008. Action planning in sequential skills: relations to music performance. *Q J Exp Psychol (Hove)*, 61(2), pp.275–291.
- Keller, P.E. & Koch, I., 2006. Exogenous and endogenous response priming with auditory stimuli. *Advances in Cognitive Psychology*, 2(4), pp.269–276.
- Keysers, C. et al., 2003. Audiovisual mirror neurons and action recognition. *Exp Brain Res*, 153(4), pp.628–636.

- Kilner, J.M. et al., 2004. Motor activation prior to observation of a predicted movement. *Nature neuroscience*, 7(12), pp.1299–1301.
- Koch, I., Keller, P. & Prinz, W., 2004. *The ideomotor approach to action control: Implications for skilled performance*,
- Lahav, A., Saltzman, E. & Schlaug, G., 2007. Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(2), pp.308–314.
- Lappe, C. et al., 2008. Cortical plasticity induced by short-term unimodal and multimodal musical training. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 28(39), pp.9632–9639.
- Lawrence, M.A., 2013. ez: Easy analysis and visualization of factorial experiments. *R package version 4.2-2*. Available at: <http://cran.r-project.org/package=ez>.
- Lega, C. et al., 2016. Testing the role of dorsal premotor cortex in auditory-motor association learning using Transcranial Magnetic Stimulation (TMS). *PLoS ONE*, 11(9), pp.1–16.
- Lu, X. & Ashe, J., 2005. Anticipatory activity in primary motor cortex codes memorized movement sequences. *Neuron*, 45(6), pp.967–973. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15797556> [Accessed August 4, 2016].
- Maess, B. et al., 2001. Musical syntax is processed in Broca's area: an MEG study. *Nature Neuroscience*, 4(5), pp.540–545. Available at: http://www.nature.com/neuro/journal/v4/n5/abs/nn0501_540.html http://www.nature.com/neuro/journal/v4/n5/pdf/nn0501_540.pdf.
- Maidhof, C. et al., 2009. Nobody is perfect: ERP effects prior performance errors in musicians indicate fast monitoring processes. *PLoS ONE*, 4(4).
- Nicholls, M.E.R. et al., 2013. The flinders handedness survey (FLANDERS): A brief measure of skilled hand preference. *Cortex*, 49(10), pp.2914–2926. Available at: <http://dx.doi.org/10.1016/j.cortex.2013.02.002>.
- Nishitani, N. & Hari, R., 2000. Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America*, 97(2), pp.913–8. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=15430&tool=pmcentrez&rendertype=abstract>.
- Novembre, G. & Keller, P.E., 2014. A conceptual review on action-perception coupling in the musicians brain: what is it good for? *Frontiers in Human Neuroscience*, 8(August), p.603. Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4139714&tool=pmcentrez&rendertype=abstract> <http://journal.frontiersin.org/article/10.3389/fnhum.2014.00603/abstract>.

- Palmer, C., 2005. Time course of retrieval and movement preparation in music performance. *Annals of the New York Academy of Sciences*, 1060, pp.360–367.
- di Pellegrino, G. et al., 1992. Understanding motor events: a neurophysiological study. *Experimental Brain Research*, 91(1), pp.176–180.
- Platel, H. et al., 1997. The structural components of music perception. A functional anatomical study. *Brain : a journal of neurology*, 120 (Pt 2, pp.229–243.
- R Core Team, 2014. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, Vienna, Au. Available at: <http://www.r-project.org/>.
- Rizzolatti, G. & Craighero, L., 2004. The Mirror-Neuron System. *Annual Review of Neuroscience*, 27(1), pp.169–192. Available at: <http://www.annualreviews.org/doi/abs/10.1146/annurev.neuro.27.070203.144230>.
- Rizzolatti, G., Fogassi, L. & Gallese, V., 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience*, 2(September), pp.661–70.
- Romani, M. et al., 2005. Motor facilitation of the human cortico-spinal system during observation of bio-mechanically impossible movements. *NeuroImage*, 26(3), pp.755–763.
- Rossi, S. et al., 2009. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology*, 120(12), pp.2008–2039.
- Rossini, P.M. et al., 2015. Non-invasive electrical and magnetic stimulation of the brain, spinal cord, roots and peripheral nerves: basic principles and procedures for routine clinical and research application. An updated report from an I.F.C.N. Committee. *Clinical Neurophysiology*, 126(6), pp.1071–1107. Available at: <http://linkinghub.elsevier.com/retrieve/pii/S1388245715000711>.
- Ruiz, M.H., Jabusch, H.C. & Altenmüller, E., 2009. Detecting wrong notes in advance: Neuronal correlates of error monitoring in pianists. *Cerebral Cortex*, 19(11), pp.2625–2639.
- Schubotz, R.I., 2007. Prediction of external events with our motor system: towards a new framework. *Trends in Cognitive Sciences*, 11(5), pp.211–218.
- Schubotz, R.I. & von Cramon, D.Y., 2002. Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: an fMRI study. *NeuroImage*, 15(4), pp.787–796.
- Schubotz, R.I. & Von Cramon, D.Y., 2004. Sequences of Abstract Nonbiological Stimuli Share Ventral Premotor Cortex with Action Observation and Imagery. *Journal of Neuroscience*, 24(24), pp.5467–5474. Available at:

- <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1169-04.2004>.
- Stock, A. & Stock, C., 2004. A short history of ideo-motor action. *Psychological Research*, 68(2–3), pp.176–188.
- Strafella, A.P. & Paus, T., 2000. Modulation of cortical excitability during action observation: a transcranial magnetic stimulation study.[Miscellaneous Article]. *Neuroreport*, 11(10), pp.2289–2292.
- Ticini, L.F. et al., 2012. When Sounds Become Actions: Higher-order Representation of Newly Learned Action Sounds in the Human Motor System. *Journal of Cognitive Neuroscience*, 24(2), pp.464–474.
- Tourville, J.A. & Guenther, F.H., 2013. DIVA model for speech acquisition. , 26(7), pp.1–27.
- Urgesi, C. et al., 2006. Motor facilitation during action observation: Topographic mapping of the target muscle and influence of the onlooker's posture. *European Journal of Neuroscience*, 23(9), pp.2522–2530.
- Urgesi, C. et al., 2010. Simulating the future of actions in the human corticospinal system. *Cerebral Cortex*, 20(11), pp.2511–2521.
- Vuust, P. & Witek, M.A.G., 2014. Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology*, 5(OCT), pp.1–14.
- Wolpert, D.M., Ghahramani, Z. & Jordan, M.I., 1995. An internal model for sensorimotor integration. *Science-AAAS-Weekly Paper Edition*, 269(5232), pp.1880–1882.
- Zappi, V. & McPherson, A., 2014. Design and Use of a Hackable Digital Instrument. *Proceedings of the 2nd International Conference on Live Interfaces*, pp.208–219.
- Zatorre, R.J., Chen, J.L. & Penhune, V.B., 2007. When the brain plays music: auditory-motor interactions in music perception and production. *Nature reviews. Neuroscience*, 8(7), pp.547–558.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT