

SCIENTIFIC REPORTS

**OPEN**

Training of Working Memory Impacts Neural Processing of Vocal Pitch Regulation

Received: 02 March 2015
Accepted: 15 October 2015
Published: 10 November 2015

Weifeng Li^{1,*}, Zhiqiang Guo^{2,*}, Jeffery A. Jones³, Xiyang Huang¹, Xi Chen¹, Peng Liu¹, Shaozhen Chen¹ & Hanjun Liu^{1,2}

Working memory training can improve the performance of tasks that were not trained. Whether auditory-motor integration for voice control can benefit from working memory training, however, remains unclear. The present event-related potential (ERP) study examined the impact of working memory training on the auditory-motor processing of vocal pitch. Trained participants underwent adaptive working memory training using a digit span backwards paradigm, while control participants did not receive any training. Before and after training, both trained and control participants were exposed to frequency-altered auditory feedback while producing vocalizations. After training, trained participants exhibited significantly decreased N1 amplitudes and increased P2 amplitudes in response to pitch errors in voice auditory feedback. In addition, there was a significant positive correlation between the degree of improvement in working memory capacity and the post-pre difference in P2 amplitudes. Training-related changes in the vocal compensation, however, were not observed. There was no systematic change in either vocal or cortical responses for control participants. These findings provide evidence that working memory training impacts the cortical processing of feedback errors in vocal pitch regulation. This enhanced cortical processing may be the result of increased neural efficiency in the detection of pitch errors between the intended and actual feedback.

Working memory refers to the neural process by which information is stored and manipulated over a brief period of time¹. Multiple lines of evidence have demonstrated strong associations between working memory and complex cognitive tasks such as fluid reasoning^{2,3}, reading comprehension^{4,5}, and attentional control^{6,7}. Brain damage due to events such as stroke and traumatic brain injury, as well as developmental and psychiatric disorders such as intellectual development disorder and schizophrenia, can cause impairments in working memory that affect quality of life^{8,9}. Given the importance of working memory in facilitating complex cognition, recent years have seen a surge in the development of training programs aimed at not only enhancing working memory capacity, but also producing effects that generalize to untrained tasks.

Previous research has shown that working memory capacity can be improved by training. For example, following computerized working memory training, children with attention-deficit/hyperactivity disorder and patients who have recently suffered a stroke showed a significant improvement in their working memory capacity^{10,11}. Moreover, working memory training can lead to beneficial effects in tasks that were not trained. For example, in a study by Jaeggi *et al.*², participants were trained on a dual *n*-back task where they simultaneously heard a series of single letters while they saw a square sequentially placed at different positions on a screen. Their task was to determine whether each of these stimuli was presented *n* items back in the series. After this training, participants exhibited improved performance

¹Department of Rehabilitation Medicine, The First Affiliated Hospital, Sun Yat-sen University, Guangzhou, 510080, P. R. China. ²Department of Biomedical Engineering, School of Engineering, Sun Yat-sen University, Guangzhou, China, 510006. ³Psychology Department and Laurier Centre for Cognitive Neuroscience, Wilfrid Laurier University, Waterloo, Ontario, N2L 3C5, Canada. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to H.L. (email: lhanjun@mail.sysu.edu.cn)

on an untrained task that involved non-verbal reasoning. Similarly, Dahlin *et al.*¹² reported that, after participating in five weeks of adaptive training that required participants to update single items (letters, digits, colors, and spatial locations) in their working memory, young adults significantly improved their performance on a non-trained 3-back working memory task.

In addition to its influence on the above-mentioned cognitive functions, working memory has been found to be involved in sensorimotor integration in speech processing. In the phonological loop system of Baddeley and Hitch's working memory model¹, verbal information is processed through the interaction between the passive phonological storage and the active articulatory rehearsal. It has been suggested that sensorimotor processes may assist with the representation and manipulation of information, and auditory working memory acts to translate the auditory information into a rehearseable sensorimotor code¹³. Previous research has established a neural network underlying auditory verbal working memory, which includes the dorsolateral prefrontal cortex, premotor cortex, superior temporal gyrus, and area Spt (Sylvian-parietal-temporal)^{14,15}; activity in these structures was also observed for the auditory-motor processing of feedback errors during vocal pitch regulation^{16–18}. In particular, area Spt has been suggested to act as an auditory-motor interface for working memory because of its involvement in the temporary storage of verbal information during working memory tasks¹⁹, and the mapping of perceived speech sounds onto articulatory representations²⁰. It is thus proposed that the internal representation and manipulation of speech sounds stored in working memory rely on auditory-motor integration¹³.

There are theories that postulate that auditory-motor integration in speech processing relies on a subtractive comparison of the expected auditory feedback (corollary discharge), based on a copy of the motor command, with the actual vocal output (re-afference)^{21,22}. In the case of vocal pitch regulation, errors heard in auditory feedback may be detected by comparing the auditory re-afference with a sensory memory trace of the vocal target, with perceived discrepancies eliciting a corrective motor command to regulate the vocal production²³. Working memory functions to integrate incoming information with stored information in existing memory²⁴, and it has been suggested that information related to the speech motor command and sensory re-afference can be stored in short-term memory within a feedback circuit and recalled when needed to adjust the motor activity^{25,26}. On the other hand, activation of the brain's error detection system is affected by characteristics of the working memory system²⁷, and the error detection mechanisms can benefit from training of working memory, as evidenced by increased amplitudes of the error-related negativity (ERN) using electroencephalography (EEG) for participants who received a training of auditory, visual, and cross-modality working memory skills using the CogniFit Personal Coach Program²⁸. Moreover, there is evidence that intensive training of auditory working memory can make the storage, access, updating, and rehearsal of auditory information more efficient²⁹. Thus, training of working memory may increase neural efficiency in the detection of feedback errors during the online monitoring of self-produced vocalization, facilitating the cortical mechanisms of auditory-motor integration for voice control.

There is also evidence that training individuals to recognize speech sounds involves a variety of cognitive abilities such as working memory and acoustic discrimination, and that this perceptual training can drive changes to speech motor control. For example, in one recent speech perceptual learning study, Mandarin speakers were trained to associate pictures with words that were spoken with pitch patterns that resembled five Thai tones. Participants who received the training produced significantly smaller N1 responses and larger P2 responses to pitch errors in their voice auditory feedback, while control participants who did not undergo the training did not³⁰. Given that improving perceptual abilities can directly impact working memory³¹, it is also possible that changes in sensorimotor integration in speech processing caused by perceptual training are related to improvements in working memory capabilities. To date, no study has directly investigated the interactions between working memory training and plasticity in auditory-motor integration.

Since the training of auditory working memory can lead to increased neural efficiency in the storage and manipulation of auditory information, it is conceivable that auditory working memory training may also lead to beneficial effects for voice control by facilitating the detection of vocal pitch errors in auditory feedback. To answer this question, the present event-related potential (ERP) study recruited one trained group who participated in an adaptive training of auditory working memory using a digit span backwards (DSB) paradigm, and one control group who did not receive any training. At the pre-training and post-training sessions, both trained and control participants were instructed to sustain a vowel phonation while hearing their voice auditory feedback unexpectedly pitch-shifted. Behavioral and cortical responses to pitch perturbations were measured to examine the influence of working memory training on the auditory-motor control of vocal production. Because previous research has shown that working memory training leads to beneficial effects on untrained tasks^{2,12}, and changes in N1 and P2 responses to vocal pitch perturbations following speech perceptual learning³⁰, we predicted that this adaptive DSB training would result in not only improvements in working memory capacity, but also improved auditory-motor processing of vocal pitch. Specifically, we expected larger vocal responses, smaller N1 responses (i.e., less negative), and larger P2 responses (i.e., more positive) to pitch feedback perturbations for the participants who underwent the training.

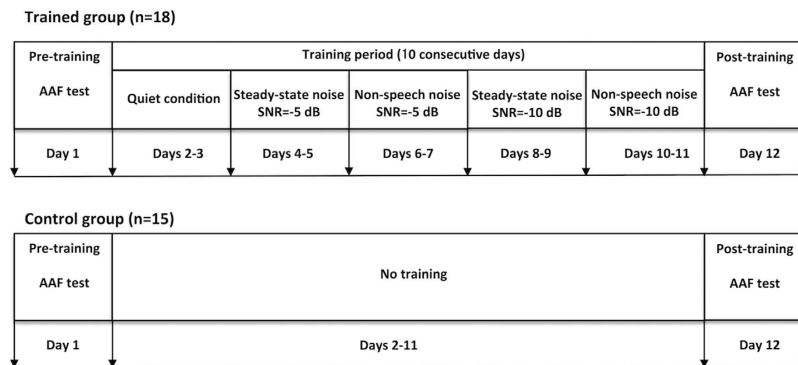


Figure 1. Schematic illustration showing the experimental setup. Trained participants received adaptive training using a digit span backwards paradigm over 10 consecutive days. The digits were presented in quiet, steady-state noise (-5 and -10 dB), and non-speech noise (-5 and -10 dB) conditions. Control participants did not receive training. Both groups participated in the altered auditory feedback (AAF) experiment before and after training, during which they produced vowel sounds while hearing their voice pitch feedback unexpectedly shifted through headphones.

Methods

Subjects. Thirty-three right-handed and native-Mandarin speaking participants, who were students at Sun Yat-sen University in China, were recruited and randomly assigned to the trained or control group. Eighteen participants (8 male and 10 female; mean age = 22 years) were assigned to the trained group and underwent 10 days of working memory training. Fifteen participants (7 male and 8 female; mean age = 21 years) who did not undergo the training were assigned to the control group. The two groups did not differ in terms of age, gender, and education. Participants had no history of language, hearing, or neurological disorders. Participants were included in the experiments only if they had normal hearing (≤ 25 dB hearing level [HL] pure-tone thresholds from 250–4000 Hz). Written informed consent was obtained from all participants, and the research protocol was approved by the Institutional Review Board of The First Affiliated Hospital at Sun Yat-sen University of China in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). The study was carried out in accordance with the approved guidelines.

Working memory training. The training scheme was designed to provide participants with practice in auditory working memory. Working memory training was carried out in a sound-attenuated booth. Training was implemented with E-Prime software (Psychological Testing Services, Pittsburgh, Pennsylvania) using a DSB paradigm, during which participants heard a string of digits and then were asked to type the digits in the reverse order that they were presented (e.g., if the participant heard “1 2 3 4”, then the correct response was “4 3 2 1”). Participants were seated in front of a personal computer and heard the sounds through Sennheiser headphones (HMD 250). Listeners were allowed to choose a preferred loudness level for training stimuli at the outset of the experiment, and this level was maintained throughout training.

The duration of the experiment was 12 days for each participant. Both trained and control participants completed the pre- and post-test on the first and twelfth day, and the trained participants received training on the second through eleventh day (i.e. training days 1–10). The training paradigm was adaptive such that the length of digit span was adjusted based on the listener’s performance. On training day 1, the training session started with 2 digits. Each block consisted of 7 lists of digits, and 10 blocks were completed during each training day. Subjects had to reach 70% accuracy on a given span before the length of the spans was increased. Accuracy of lower than 50% caused the span lengths to decrease on the subsequent trial. Participants began each new training day with the span length they had attained on the last trial of the previous day.

Participants were asked to remember the digits one through nine, which were spoken by a native Mandarin speaker and recorded with a sampling frequency of 44100 Hz. The digit stimuli were then RMS matched to a 70 dB sound pressure level (SPL) pure tone at 1 kHz. Pilot testing revealed that listeners plateaued in performance when all training was presented in quiet, but not when noise was introduced. We therefore presented the digits in noise with varying signal-to-noise (SNR) levels. Steady-state noise matching the long-term average spectrum of the recordings was created using MATLAB and mixed with the target recordings in Adobe Audition to create stimuli with SNRs of -5 and -10 dB. In addition, a set of 27 non-speech sounds were obtained from online databases, including twelve animal sounds (e.g., bird song, dog bark), two mechanical sounds (e.g., clock ticking, gun shot), five non-speech human sounds (e.g., cough, laugh), and eight musical sounds (e.g., violin, flute). These sounds were cropped to 1 s, RMS

matched in amplitude to the calibration tone, and then mixed with all target digits at -5 and -10 dB SNR, resulting in a total of 468 stimulus combinations.

The training paradigm was implemented over 10 consecutive days (see Fig. 1). On training days one and two, the digits were presented in a quiet condition (i.e. without noise). On training days three and four, each digit was presented in steady-state noise at -5 dB SNR. On training days five and six, each digit was presented in non-speech distractors that were randomly chosen and presented at -5 dB SNR. Training days seven through ten replicated training days three through six but at -10 dB SNR.

Behavioral and neurophysiological assessment. The vocal production experiment was carried out using the altered auditory feedback (AAF) paradigm. Both trained and control participants were instructed to produce a steady vowel sound /u/ for about 5–6 s at their habitual and comfortable pitch and loudness level, during which their voice pitch feedback was unexpectedly shifted $+50$ or $+200$ cents (100 cents = one semitone) for 200 ms, and fed back to them instantaneously. The first pitch-shifted stimulus was presented with a delay of 500–1000 ms after vocal onset, and the succeeding stimuli occurred with an inter-stimulus interval of 700–900 ms. There were five pitch-shift stimuli per vocalization and participants produced 40 consecutive vocalizations. In total, 200 trials were collected including 100 trials for $+50$ cents and 100 trials for $+200$ cents.

The vocal production experiment was conducted in a sound-attenuated booth. Prior to the testing, the experimental system was calibrated to ensure that the intensity of voice feedback heard by the participants was 10 dB SPL higher than that of the subjects' voice output to partially mask any air-borne and bone-conducted feedback³². The voice signals were transduced by a dynamic microphone (DM2200, Takstar Inc.) and amplified by a MOTU Ultralite Mk3 Firewire audio interface. The amplified signals were pitch-shifted through an Eventide Eclipse Harmonizer that was controlled by a custom-developed MIDI software program (Max/MSP, v.5.0 by Cycling 74) running on a Macintosh computer. This program was also used to manipulate parameters including the magnitude, direction, and duration of the pitch shifts and generate transistor-transistor logic (TTL) control pulses to signal the onset and offset of the pitch shifts. Finally, the pitch-shifted voices were amplified by an ICON NeoAmp headphone amplifier and fed back to the participants through insert earphones (ER1-14 A, Etymotic Research Inc.). The voice, feedback, and TTL control pulses were digitized at 10 kHz by a PowerLab A/D converter (ML880, AD Instruments), recorded using LabChart software (v.7.0 by AD Instruments), and saved onto another Macintosh computer.

EEG signals were recorded using a 64-electrode Geodesic Sensor Net, amplified by a Net Amps 300 amplifier (Electrical Geodesics Inc.), digitized at a sampling frequency of 1 kHz, and saved onto a Mac Pro computer. A DIN synch cable was used to send the TTL control pulses generated by Max/MSP to the EEG recording system for the measurement of stimulus-evoked potentials. During the online recording, EEG signals across all channels were referenced to the vertex (Cz). Impedance of individual sensors were adjusted and maintained below $50\text{ k}\Omega$ ³³.

Data analyses. *Vocal Responses.* The measurement of vocal responses to pitch-shifted voice auditory feedback was implemented using a custom-developed IGOR PRO program (v.6.0, Wavemetrics Inc.) using event-related averaging techniques³⁴. Voice F_0 contours in Hertz were extracted from voice signals using Praat³⁵, and converted to the cent scale using the formula: cents = $100 \times (12 \times \log_2(F_0/\text{reference}))$ [reference = 195.997 Hz (G4)]. Voice F_0 contours were then segmented into epochs from 200 ms before and 700 ms after the stimulus onset. All individual trials were visually inspected such that those trials containing vocal interruption or signal processing errors were excluded from analyses. Finally, artifact-free trials for each condition were averaged to generate an overall response. Acceptable responses exceeded a value of two standard deviations (SDs) of the pre-stimulus mean beginning at least 60 ms after the onset of the pitch shift and lasting at least 50 ms. The response magnitude was measured by subtracting the peak value of voice contour following the response onset from the pre-stimulus mean (-200 to 0 ms). The response latency was determined as the time of voice F_0 departure from the pre-stimulus mean by more than 2 SDs.

Evoked Potentials. The EEG signals were analyzed off-line using NetStation software (v.4.5, Electrical Geodesics Inc.). All channels were digitally band-passed filtered with cut-off frequencies of 1 to 20 Hz. Individual trials were segmented into epochs with a window of -200 ms and $+500$ ms relative to the onset of the pitch shift. Segmented trials contaminated by excessive muscular activity, eye blinks, or eye movements were assessed using the Artifact Detection Toolbox in NetStation and excluded from further analyses. An additional visual inspection was also performed on individual trials to ensure that artifacts were being adequately rejected. Individual electrodes were excluded from further analyses if they contained artifacts more than 20% of the segmented trials. Finally, artifact-free segments were averaged, re-referenced to the average of the electrodes on each mastoid, and baseline-corrected across all conditions. Given that cortical responses to pitch-shifted voice auditory feedback are mostly pronounced in the N1-P2 complex^{36,37}, the amplitudes and latencies of N1 and P2 were measured as the negative and positive peaks in the time windows of 80–180 ms and 160–280 ms after the onset of the pitch shift and submitted to statistical analyses.

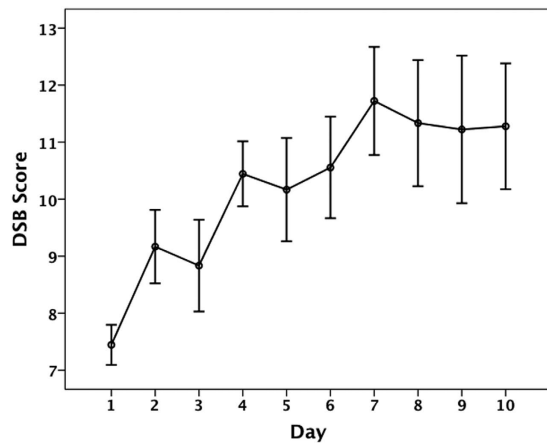


Figure 2. The mean scores of the digit span backwards task across the 10 training days. The error bars represent the standard deviations of the mean scores.

The DSB scores and the magnitudes and latencies of vocal and neurophysiological responses were subjected to repeated-measures mixed analyses of variance (ANOVAs) using SPSS (v.16.0). Testing session (pre- vs. post-training), stimulus magnitude (+50 and +200 cents), and electrode site (FC1, FC3, FCz, FC2, FC4, C1, C3, Cz, C2, C4, P1, P3, Pz, P2, P4) were chosen as within-subject factors, while group (trained vs. control group) was chosen as a between-subject factor. Appropriate subsidiary RM-ANOVAs were calculated if higher-order interactions reached significance. Greenhouse-Geisser was used to correct probability values for multiple degrees of freedom when violations of the sphericity assumption occurred.

Results

DSB measure. Fig. 2 shows participants' performance indexed by the DSB scores throughout training. As can be seen, the DSB score was below 10 (ranging from 7.4 to 9.2) on training days one through three. On training days four through ten, the DSB scores increased to as large as 11.7. A one-way ANOVA revealed a significant main effect of testing day ($F(9, 153) = 22.288, p < 0.001$), and post-hoc Bonferroni comparisons showed that the DSB scores on training days one through three were significantly lower than those on training days seven through ten ($p < 0.03$). In summary, the significant increase in the DSB measure indicates improvements in working memory capacity following training.

Vocal response. One two-way mixed ANOVA conducted on the vocal response magnitude revealed no significant main effects of testing session ($F(1, 31) = 0.045, p = 0.834$) (pre-training: 14.7 ± 6.3 cents; post-training: 13.6 ± 5.8 cents), stimulus magnitude ($F(1, 31) = 0.963, p = 0.334$) (+50 cents: 14.0 ± 5.6 cents; +200 cents: 14.2 ± 6.4 cents), or group ($F(1, 31) = 1.055, p = 0.312$) (trained: 13.6 ± 5.6 cents; control: 14.7 ± 6.5 cents). Similarly, response latencies did not differ as a function of testing session ($F(1, 31) = 0.425, p = 0.519$) (pre-training: 77 ± 24 ms; post-training: 75 ± 18 ms), stimulus magnitude ($F(1, 31) = 0.234, p = 0.632$) (+50 cents: 75 ± 19 ms; +200 cents: 77 ± 23 ms), or group ($F(1, 31) = 0.054, p = 0.817$) (trained: 76 ± 21 ms; control: 77 ± 21 ms). In addition, no significant interactions between these factors were observed for response magnitude or latency ($p > 0.05$).

ERP findings. Fig. 3 and 4 show the grand-averaged ERP waveforms and topographical distributions of N1 and P2 amplitudes in response to pitch shifts of +50 and +200 cents across all control and trained participants before (blue solid lines) and after (red solid lines) working memory training. As can be seen, the cortical responses to pitch shifts produced by the control participants during the pre-training session did not differ from the responses produced during the post-training session. By contrast, following training N1 amplitudes decreased and P2 amplitudes increased in response to both the +50 and +200 cents pitch shifts heard by the trained participants, and these changes were primarily observed in the frontocentral electrodes.

One four-way mixed ANOVA conducted on the N1 amplitudes revealed significant main effects of testing session ($F(1, 31) = 4.846, p = 0.035$) and electrode site ($F(14, 434) = 13.265, p < 0.001$), whereas the main effects of stimulus magnitude ($F(1, 31) = 3.209, p = 0.083$) and group ($F(1, 31) = 1.134, p = 0.295$) failed to reach significance. There was also a significant testing session \times electrode site \times group interaction ($F(14, 434) = 2.705, p = 0.038$). Follow-up three-way mixed ANOVAs conducted on the data from the trained group revealed a significant decrease in N1 amplitudes (less negative) following training ($F(1, 17) = 5.039, p = 0.038$), whereas there was no systematic change in N1 amplitudes as a function of testing session for the control group ($F(1, 14) = 0.650, p = 0.434$) (see Fig. 5A).

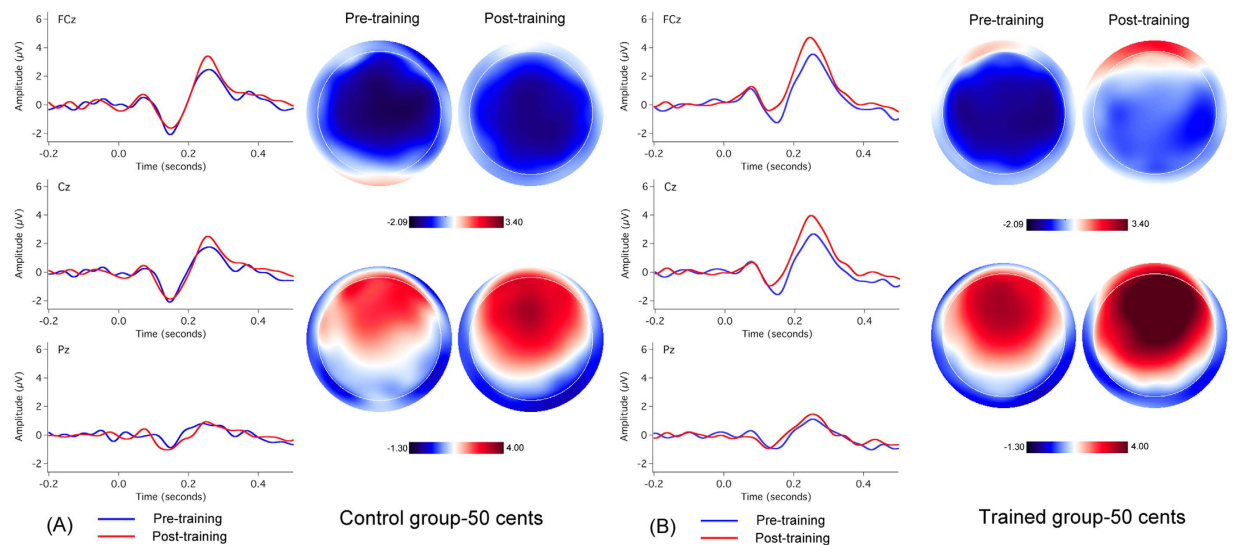


Figure 3. Grand-averaged ERP waveforms (FCz, Cz, Pz) and topographical distributions of N1 and P2 amplitudes in response to pitch shifts of +50 cents across all control (A) and trained participants (B) before (blue solid lines) and after (red solid lines) training.

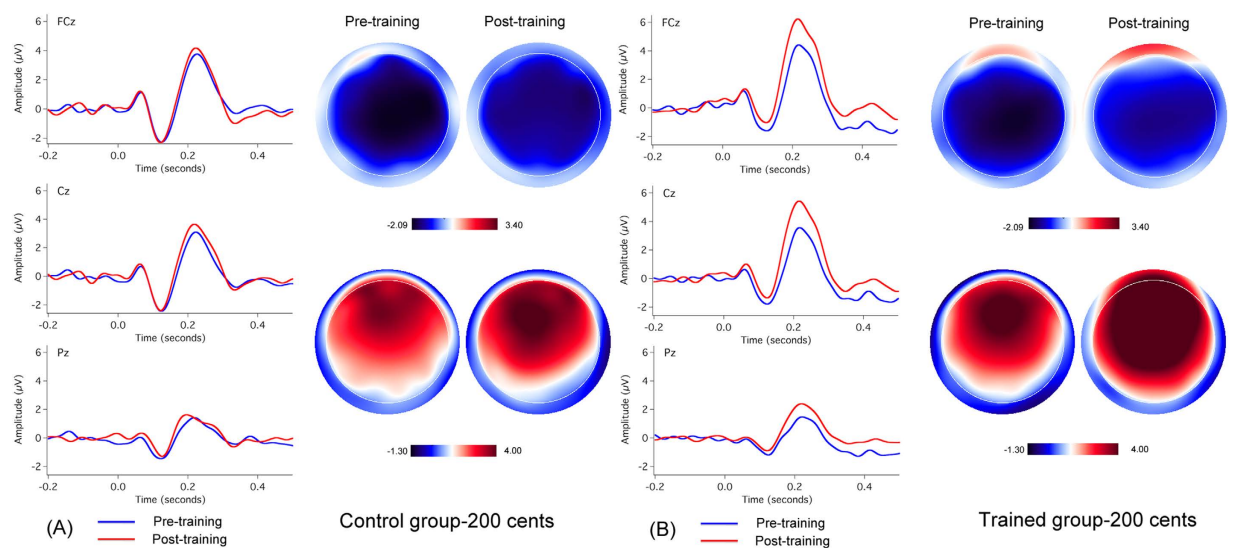


Figure 4. Grand-averaged ERP waveforms (FCz, Cz, Pz) and topographical distributions of N1 and P2 amplitudes in response to pitch shifts of +200 cents across all control (A) and trained participants (B) before (blue solid lines) and after (red solid lines) training.

As for N1 latency, pitch shifts of +50 cents elicited significantly longer N1 latencies than pitch shifts of +200 cents (143 ± 25 vs. 122 ± 20 ms) ($F(1, 20) = 81.329$, $p < 0.001$). There was also a significant main effect of electrode site ($F(14, 434) = 3.785$, $p = 0.008$), which was primarily caused by longer N1 latencies at electrode C4 as compared to Pz ($p = 0.010$) and P2 ($p = 0.004$). A main effect of group ($F(1, 31) = 1.014$, $p = 0.322$) and testing session ($F(1, 31) = 1.486$, $p = 0.232$), however, failed to reach significance. No significant interactive effects between these factors were found either ($p > 0.05$).

One four-way mixed ANOVA conducted on the P2 amplitudes revealed a significant main effect of stimulus magnitude ($F(1, 31) = 20.102$, $p < 0.001$), indicating that pitch shifts of +50 cents elicited significantly smaller P2 amplitudes than pitch shifts of +200 cents. There was also a significant main effect of electrode site ($F(14, 434) = 73.727$, $p < 0.001$), which was primarily driven by larger P2 amplitudes at frontocentral electrodes than parietal electrodes ($p < 0.05$) (see Figs 3 and 4). The main effect of testing session ($F(1, 31) = 15.256$, $p < 0.001$) reached significance, whereas there was no main effect of group ($F(1, 31) = 2.928$, $p = 0.097$). A significant interaction, however, was found between testing session and group ($F(1, 31) = 4.917$, $p = 0.034$). A follow-up three-way mixed ANOVA conducted on the data from

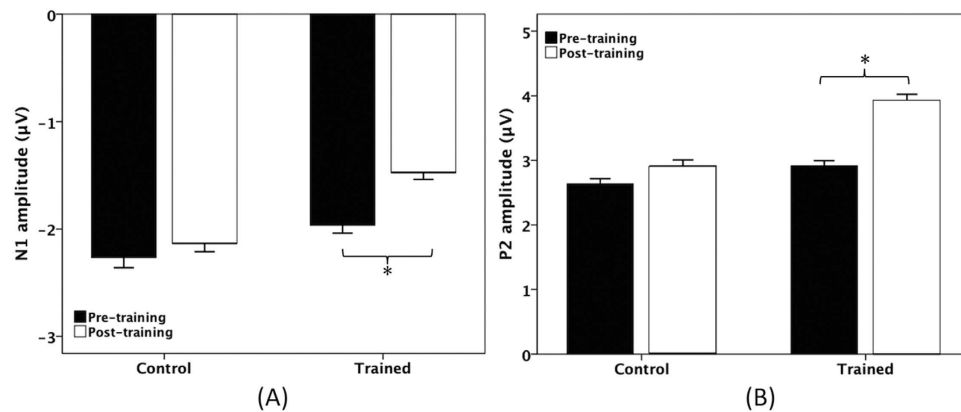


Figure 5. T-bar plots (means and standard errors) of N1 (A) and P2 (B) amplitude as a function of testing session and group. The black and the blank bars denote the cortical responses at the pre-training and the post-training session, respectively. The asterisks indicate significant differences between conditions.

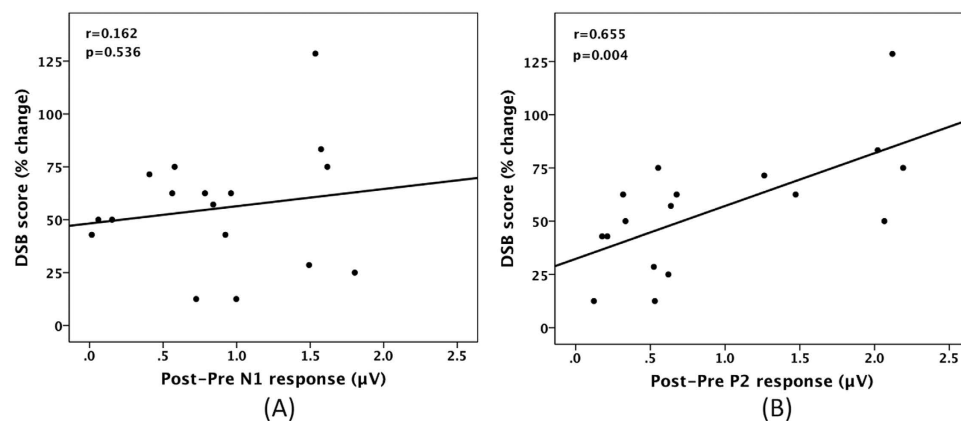


Figure 6. The percentage change of DSB scores is plotted against the post-pre difference for the mean N1 (A) and P2 (B) responses to pitch feedback perturbations. There was a significant positive correlation between the post-pre difference for the mean P2 responses and the percentage change in DSB scores ($r=0.655$, $p=0.004$), whereas the post-pre difference for the mean N1 responses was not significantly correlated with the percentage change in DSB scores ($r=0.162$, $p=0.536$).

the trained group revealed a significant main effect of testing session ($F(1, 17) = 17.828$, $p = 0.001$), indicating that P2 amplitudes significantly increased following training (see Fig. 5B). By contrast, P2 amplitudes did not differ as a function of testing session ($F(1, 14) = 1.614$, $p = 0.225$) for the control group.

Regarding P2 latency, there was a significant main effect of stimulus magnitude ($F(1, 31) = 97.498$, $p < 0.001$) as reflected by significantly shorter P2 latencies elicited by pitch shifts of +200 cents, as compared to pitch shifts of +50 cents (225 ± 24 vs. 252 ± 25 ms). However, the main effects of testing session ($F(1, 31) = 3.774$, $p = 0.061$), electrode site ($F(14, 434) = 1.817$, $p = 0.148$), and group ($F(1, 31) = 0.037$, $p = 0.849$) failed to reach significance. There also were no significant interactions between these factors ($p > 0.05$).

In addition, regression analyses were performed to examine the relationship between working memory training and cortical responses to pitch-shifted voice auditory feedback. The percentage change of DSB scores (i.e. post-pre DSB scores divided by pre-training DSB scores) is plotted against the post-pre difference for the mean N1 and P2 responses in Fig. 6. The results showed a significant positive correlation between the post-pre difference for the mean P2 responses and the percentage change in DSB scores ($r=0.655$, $p=0.004$), indicating that the degree of improvement in working memory capacity was predictive of the training-related enhancement of cortical responses to pitch feedback perturbations. The post-pre difference for the mean N1 responses, however, was not significantly correlated with the percentage change in DSB scores ($r=0.162$, $p=0.536$).

Discussion

The present study investigated whether working memory training can lead to improved processing of feedback errors during vocal pitch regulation. Before and after an adaptive DSB training procedure, participants' vocal and cortical responses to errors detected in their auditory feedback were measured. As hypothesized, in addition to significant improvements in participants' working memory capacity (i.e. the DSB score), we observed an impact of working memory training on the cortical processing of vocal pitch errors. The trained participants' N1 responses to pitch errors in voice auditory feedback decreased, whereas the N1 responses for the first and second AAF sessions of control participants' were not significantly different. Conversely, P2 responses increased after working memory training, but remained the same across the vocal production tasks for control participants. Moreover, the percentage change in DSB scores was significantly correlated with the post-pre difference for the mean P2 responses to pitch feedback perturbations, indicating that improvement in working memory capacity was predictive of training-related enhancement of cortical responses to voice feedback errors. Neither the trained participants nor the control participants, however, showed a systematic change in their vocal compensation magnitude or latency across the pre- and post-training sessions. Taken together, these findings demonstrate that training-related improvements in working memory capacities generalize to other tasks, and modify the cortical processing of mismatches between intended and actual auditory feedback.

Training-related effects on the N1-P2 complex. A primary finding in this study is that the training of working memory led to decreased N1 responses when participants heard pitch-shifted auditory feedback. This finding is in line with other studies that showed decreased N1 amplitudes in response to pitch feedback perturbations after speech perceptual learning³⁰, and decreased brain activity in the frontoparietal regions after working memory training^{29,38,39}. The N1 component is generally thought to reflect pre-attentive detection of a mismatch between incoming auditory stimuli and the memory trace of previous sensory input into the auditory system⁴⁰. We hypothesize that increased neural efficiency⁴¹ accounts for the decreased N1 responses to pitch errors in voice auditory feedback that we observed. This hypothesis is supported by work showing increased efficiency in the processing of auditory information following training of auditory working memory as reflected by decreases in brain activation. For example, in a recent study an adaptive *n*-back training with tonal sequences led to not only improved performance on an auditory 2-back task, but also decreased activation in the right inferior frontal gyrus and posterior parietal regions²⁹. Thus, following intensive training of auditory working memory, fewer neural resources may be required for encoding the same level of auditory information such that the integration of incoming information with information stored in memory becomes more efficient. In the case of voice control then, the detection and comparison of pitch errors in voice auditory feedback may become more efficient, and this increased efficiency may be reflected by the significantly decreased N1 responses observed after training in the present study.

In addition to the changes in N1 responses, enhancement of P2 responses to pitch-shifted voice auditory feedback was also observed for trained participants. Similarly in a previous study, P2 responses to pitch feedback perturbations were significantly increased after a sound-to-word learning of lexical tones by associating speech stimuli with pictures of objects³⁰. The present finding is also in line with other findings that showed increased theta power⁴² and ERN²⁸ using EEG, and increased blood oxygenation level dependent (BOLD) responses measured with fMRI^{43–45} following working memory training. Since P2 responses to pitch feedback perturbations are associated with relatively later stages of cortical processing, it has been suggested that the P2 component may reflect functional mechanisms that demand higher-level cognitive processes of auditory-motor integration⁴⁶. In particular, there is evidence that P2 receives contributions from the planum temporale⁴⁷, and one function of this region is to support sensorimotor integration in speech processing²². For example, the posterior region of the planum temporale, area Spt, has been found to function as an interface between auditory and motor representations^{20,48}. Taken together, increased P2 responses to pitch feedback errors may index the training-induced enhancement of the coordinated neural activity for the interaction between the auditory and motor systems in voice control.

Interestingly, the post-pre difference in the mean P2 responses was positively correlated with the percentage change in DSB scores, suggesting a relationship between the degree of improvement in working memory capacity and training-related enhancement of cortical responses to pitch feedback perturbations. This finding provides further evidence that working memory in the phonological loop is dependent on the operation of an auditory-motor interface system in the posterior planum temporale¹⁹. Also this correlation lends support to the idea that the P2 component reflects higher-level cognitive processing of feedback errors during the auditory-motor integration for voice control.

Since it has been previously shown that P2 responses become larger when participants attend to pitch feedback perturbations as compared to when they do not^{49,50}, and that attentional control can be improved by working memory training⁵¹, one may argue that the training-related changes we observed in the N1-P2 complex were not the result of improved working memory capacities per se, but the result of an increased capacity for attentional control. That is, after working memory training, our trained participants may have paid more attention to their voice auditory feedback, and this increased attention caused the training-related changes in the N1-P2 complex. However, in addition to enhancing P2 responses, attending to pitch feedback perturbations has been shown to increase N1 amplitudes and increase the

size of vocal compensations, as compared to when they are not attended^{49,52}. In the present study, we observed *decreased* N1 responses and vocal compensation remained unchanged after training. Moreover, the correlation observed between the improvements in working memory capacity and the enhanced P2 responses to pitch perturbations suggests that the training gains in auditory-motor integration for voice control are most likely due to increased working memory capacity.

The lack of training effect on the vocal compensation. Unlike the N1-P2 complex, the vocal compensations for the pitch errors heard in voice auditory feedback were not affected by the working memory training. Similarly, Chen *et al.*³⁰ did not find a systematic change in the vocal responses to pitch feedback perturbations for participants who underwent training to associate unfamiliar speech sounds with words. These results are also in line with studies on visuomotor adaptation, which showed that improving working memory capacity using dual *n*-back training does not alter the rates of visuomotor adaptation⁵³, although individual differences in spatial working memory capacity are predictive of the rate of early visuomotor adaptation⁵⁴. In terms of the present study, there are a number of potential explanations for the lack of transfer to vocal compensation responses. For example, it is possible that working memory training may only benefit the detection of pitch errors in auditory feedback. It may also be that if participants were presented with feedback perturbations that were much smaller in magnitude than those presented in the present study, and near the threshold for detection, differences between a trained and untrained group of participants would emerge. In addition, there are several other factors that contribute to the underlying mechanisms engaged in the modulation of vocal compensation for voice feedback errors, such as stimulus features^{55,56} and task demands^{57,58}. Future studies that manipulate these factors could provide a more thorough characterization of transfer effect of working memory training to auditory-motor integration.

Limitation. One limitation of the present study was the use of a passive control group that was used to control for pre-/post-test effects, but did not receive any form of training between these tests. Thus, confounding factors such as motivational and psychological effects, which have been shown to influence the effectiveness of working memory training⁵⁹, were not controlled using the present experimental design. Therefore, it is possible that transfer effects we observed are not solely a result of training, but may in part be related to different levels in motivation and/or practice. Positive transfer effects of working memory training will be more convincing when compared against an active control group with a cognitively engaging alternative intervention. Another limitation is the inclusion of a perception-in-noise component, which makes it difficult to determine how much of the transfer effect was a result of the working memory training gains or an improved ability to discriminate the acoustic stimuli in noise. In future studies, the present training paradigm should be compared to an adaptive DSB training without noise to determine the contributions of perception in noise vs. working memory capacity to the transfer effects we observed.

Conclusions

Overall, we found neurophysiological evidence that training of auditory working memory impacts the auditory-motor processing of vocal pitch regulation at the cortical level. This includes decreased N1 and increased P2 responses to pitch errors in voice auditory feedback following training. Our results extend previous findings regarding the transfer of training gains in working memory, demonstrating that the cortical processing of vocal pitch regulation can benefit from working memory training.

References

1. Baddeley, A. Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* **4**, 829–39 (2003).
2. Jaeggi, S. M., Buschkuhl, M., Jonides, J. & Perrig, W. J. Improving fluid intelligence with training on working memory. *Proc. Natl. Acad. Sci. USA.* **105**, 6829–33 (2008).
3. Borella, E., Carretti, B. & Mammarella, I. C. Do working memory and susceptibility to interference predict individual differences in fluid intelligence? *Eur. J. Cogn. Psychol.* **18**, 51–69 (2006).
4. DeDe, G., Caplan, D., Kemtes, K. & Waters, G. The relationship between age, verbal working memory, and language comprehension. *Psychol. Aging* **19**, 601–16 (2004).
5. De Beni, R., Borella, E. & Carretti, B. Reading comprehension in aging: the role of working memory and metacomprehension. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* **14**, 189–212 (2007).
6. Bherer, L. *et al.* Testing the limits of cognitive plasticity in older adults: application to attentional control. *Acta Psychol. (Amst)* **123**, 261–78 (2006).
7. Kramer, A. F., Hahn, S. & Gopher, D. Task coordination and aging: explorations of executive control processes in the task switching paradigm. *Acta Psychol. (Amst)* **101**, 339–78 (1999).
8. Robertson, I. H. & Murre, J. M. Rehabilitation of brain damage: brain plasticity and principles of guided recovery. *Psychol. Bull.* **125**, 544–75 (1999).
9. Goldman-Rakic, P. S. Working memory dysfunction in schizophrenia. *J. Neuropsychiatry Clin. Neurosci.* **6**, 348–57 (1994).
10. Klingberg, T. *et al.* Computerized training of working memory in children with ADHD—a randomized, controlled trial. *J. Am. Acad. Child. Psy.* **44**, 177–86 (2005).
11. Westerberg, H. *et al.* Computerized working memory training after stroke—a pilot study. *Brain Inj.* **21**, 21–9 (2007).
12. Dahlin, E., Nyberg, L., Backman, L. & Neely, A. S. Plasticity of executive functioning in young and older adults: immediate training gains, transfer, and long-term maintenance. *Psychol. Aging* **23**, 720–30 (2008).
13. Schulze, K. & Koelsch, S. Working memory for speech and music. *Ann. N. Y. Acad. Sci.* **1252**, 229–36 (2012).

14. Buchsbaum, B. R., Olsen, R. K., Koch, P. & Berman, K. F. Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. *Neuron* **48**, 687–97 (2005).
15. Koelsch, S. *et al.* Functional architecture of verbal and tonal working memory: an fMRI study. *Hum. Brain Mapp.* **30**, 859–73 (2009).
16. Behroozmand, R. *et al.* Sensory-motor networks involved in speech production and motor control: an fMRI study. *Neuroimage* **109**, 418–28 (2015).
17. Parkinson, A. L. *et al.* Understanding the neural mechanisms involved in sensory control of voice production. *Neuroimage* **61**, 314–322 (2012).
18. Zarate, J. M. & Zatorre, R. J. Experience-dependent neural substrates involved in vocal pitch regulation during singing. *Neuroimage* **40**, 1871–87 (2008).
19. Buchsbaum, B. R. & D'Esposito, M. The search for the phonological store: from loop to convolution. *J. Cogn. Neurosci.* **20**, 762–78 (2008).
20. Hickok, G., Buchsbaum, B., Humphries, C. & Muftuler, T. Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *J. Cogn. Neurosci.* **15**, 673–82 (2003).
21. Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M. & Ford, J. M. Fine-tuning of auditory cortex during speech production. *Psychophysiology* **42**, 180–190 (2005).
22. Hickok, G., Houde, J. F. & Rong, F. Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* **69**, 407–22 (2011).
23. Hawco, C. S. & Jones, J. A. Control of vocalization at utterance onset and mid-utterance: Different mechanisms for different goals. *Brain Res.* **1276**, 131–139 (2009).
24. Brumback, C. R., Low, K. A., Gratton, G. & Fabiani, M. Putting things into perspective: individual differences in working-memory span and the integration of information. *Exp. psychol.* **52**, 21–30 (2005).
25. Troyer, T. W. & Doupe, A. J. An associational model of birdsong sensorimotor learning I. Efference copy and the learning of song syllables. *J. Neurophysiol.* **84**, 1204–23 (2000).
26. Gonzalez, C. C. & Burke, M. R. The brain uses efference copy information to optimise spatial memory. *Exp. Brain Res.* **224**, 189–97 (2013).
27. Hochman, E. Y. & Meiran, N. Central interference in error processing. *Mem. Cognit.* **33**, 635–43 (2005).
28. Horowitz-Kraus, T. & Breznitz, Z. Can the error detection mechanism benefit from training the working memory? A comparison between dyslexics and controls—an ERP study. *PLoS ONE* **4**, e7141 (2009).
29. Schneiders, J. A. *et al.* The impact of auditory working memory training on the fronto-parietal working memory network. *Front. Hum. Neurosci.* **6**, 173 (2012).
30. Chen, Z. *et al.* Transfer effect of speech-sound learning on auditory-motor processing of perceived vocal pitch errors. *Sci. Rep.* **5**, 13134 (2015).
31. Berry, A. S. *et al.* The influence of perceptual training on working memory in older adults. *PLoS ONE* **5**, e11537 (2010).
32. Behroozmand, R., Karvelis, L., Liu, H. & Larson, C. R. Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. *Clin. Neurophysiol.* **120**, 1303–1312 (2009).
33. Ferree, T. C., Luu, P., Russell, G. S. & Tucker, D. M. Scalp electrode impedance, infection risk, and EEG data quality. *Clin. Neurophysiol.* **112**, 536–44 (2001).
34. Li, W. *et al.* Neurophysiological evidence of differential mechanisms involved in producing opposing and following responses to altered auditory feedback. *Clin. Neurophysiol.* **124**, 2161–2171 (2013).
35. Boersma, P. Praat, a system for doing phonetics by computer. *Glott International* **5**, 341–345 (2001).
36. Behroozmand, R., Liu, H. & Larson, C. R. Time-dependent neural processing of auditory feedback during voice pitch error detection. *J. Cogn. Neurosci.* **23**, 1205–1217 (2011).
37. Chen, Z. *et al.* ERP correlates of language-specific processing of auditory pitch feedback during self-vocalization. *Brain Lang.* **121**, 25–34 (2012).
38. Sayala, S., Sala, J. B. & Courtney, S. M. Increased neural efficiency with repeated performance of a working memory task is information-type dependent. *Cereb. Cortex.* **16**, 609–17 (2006).
39. Dahlin, E., Neely, A. S., Larsson, A., Backman, L. & Nyberg, L. Transfer of learning after updating training mediated by the striatum. *Science* **320**, 1510–2 (2008).
40. Nääätänen, R. & Picton, T. 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 (1987).
41. Kelly, C., Foxe, J. J. & Garavan, H. Patterns of normal human brain plasticity after practice and their implications for neurorehabilitation. *Arch. Phys. Med. Rehab.* **87**, S20–9 (2006).
42. Langer, N., von Bastian, C. C., Wirz, H., Oberauer, K. & Jancke, L. The effects of working memory training on functional brain network efficiency. *Cortex* **49**, 2424–38 (2013).
43. Olesen, P. J., Westerberg, H. & Klingberg, T. Increased prefrontal and parietal activity after training of working memory. *Nat. Neurosci.* **7**, 75–9 (2004).
44. Wexler, B. E., Anderson, M., Fulbright, R. K. & Gore, J. C. Preliminary evidence of improved verbal working memory performance and normalization of task-related frontal lobe activation in schizophrenia following cognitive exercises. *Am. J. Psychiat.* **157**, 1694–7 (2000).
45. Moore, C. D., Cohen, M. X. & Ranganath, C. Neural mechanisms of expert skills in visual working memory. *J. Neurosci.* **26**, 11187–96 (2006).
46. Behroozmand, R., Ibrahim, N., Korzyukov, O., Robin, D. A. & Larson, C. R. Left-hemisphere activation is associated with enhanced vocal pitch error detection in musicians with absolute pitch. *Brain Cogn.* **84**, 97–108 (2014).
47. Godey, B., Schwartz, D., de Graaf, J. B., Chauvel, P. & Liegeois-Chauvel, C. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clin. Neurophysiol.* **112**, 1850–9 (2001).
48. Hickok, G. & Poeppel, D. The cortical organization of speech processing. *Nat. Rev. Neurosci.* **8**, 393–402 (2007).
49. Liu, Y. *et al.* Selective and divided attention modulates auditory-vocal integration in the processing of pitch feedback errors. *Eur. J. Neurosci.* **42**, 1895–1904 (2015).
50. Hu, H. *et al.* Attention modulates cortical processing of pitch feedback errors in voice control. *Sci. Rep.* **5**, 7812 (2015).
51. Salminen, T., Strobach, T. & Schubert, T. On the impacts of working memory training on executive functioning. *Front. Hum. Neurosci.* **6**, 166 (2012).
52. Tumber, A. K., Scheerer, N. E. & Jones, J. A. Attentional Demands Influence Vocal Compensations to Pitch Errors Heard in Auditory Feedback. *PLoS ONE* **9**, e109968 (2014).
53. Anguera, J. A. *et al.* The effects of working memory resource depletion and training on sensorimotor adaptation. *Behav. Brain Res.* **228**, 107–15 (2012).
54. Anguera, J. A., Reuter-Lorenz, P. A., Willingham, D. T. & Seidler, R. D. Contributions of spatial working memory to visuomotor learning. *J. Cogn. Neurosci.* **22**, 1917–30 (2010).

55. Liu, H. & Larson, C. R. Effects of perturbation magnitude and voice F0 level on the pitch-shift reflex. *J. Acoust. Soc. Am.* **122**, 3671–3677 (2007).
56. Larson, C. R., Sun, J. & Hain, T. C. Effects of simultaneous perturbations of voice pitch and loudness feedback on voice F0 and amplitude control. *J. Acoust. Soc. Am.* **121**, 2862–2872 (2007).
57. Natke, U., Donath, T. M. & Kalveram, K. T. Control of voice fundamental frequency in speaking versus singing. *J. Acoust. Soc. Am.* **113**, 1587–1593 (2003).
58. Liu, H., Xu, Y. & Larson, C. R. Attenuation of vocal responses to pitch perturbations during Mandarin speech. *J. Acoust. Soc. Am.* **125**, 2299–2306 (2009).
59. von Bastian, C. C. & Oberauer, K. Effects and mechanisms of working memory training: a review. *Psychol. Res.* **78**, 803–20 (2014).

Acknowledgements

The authors would like to thank Drs. Patrick C. M. Wong and Erin M. Ingvalson for their help with data collection. This study was funded by grants from National Natural Science Foundation of China (Nos. 31070990, 31371135, 81301675 and 81472154), Guangdong Natural Science Funds for Distinguished Young Scholar (No. S2013050014470), and the Fundamental Research Funds for the Central Universities (No. 13ykzd05).

Author Contributions

W.L. and H.L. designed the project; W.L., Z.G. and X.H. performed the experiment and analyzed the data; W.L., J.J., X.C., P.L., S.C. and H.L. interpreted the results; W.L., J.J. and H.L. wrote the manuscript; all authors read and approved the final manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Li, W. *et al.* Training of Working Memory Impacts Neural Processing of Vocal Pitch Regulation. *Sci. Rep.* **5**, 16562; doi: 10.1038/srep16562 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>