Aus der Klinik für
Psychiatrie und Psychotherapie
der Medizinischen Fakultät Charité – Universitätsmedizin Berlin

DISSERTATION

Modulating Vocal Pitch Perception and Production with Transcranial Direct Current Stimulation

zur Erlangung des akademischen Grades
Doctor medicinae (Dr. med.)

vorgelegt der Medizinischen Fakultät
Charité – Universitätsmedizin Berlin

von

Anja Hohmann

aus Ludwigshafen am Rhein

Datum der Promotion: 22.06.2014

Meinen Eltern.

Table of Contents

List of Abbreviations

AF	Arcuate fasciculus
ANOVA	Analysis of variance
BOLD	Blood oxygen level dependent
cm	Centimeter
dB	Decibel
DTI	Diffusion tensor imaging
EEG	Electroencephalography
ERP	Event-related potential
F0	Fundamental frequency
fMRI	Functional magnetic resonance imaging
Hz	Hertz
IFG	Inferior frontal gyrus
IQ	Intelligence quotient
lpIFG	Left posterior inferior frontal gyrus
lpSTG	Left posterior superior temporal gyrus
LTD	Long term depression
LTP	Long term potentiation
mA	Milliampere
Max/MSP	Max/Max signal processing
MIDI	Musical instrument digital interface
MIT	Melodic intonation therapy
MRI	Magnetic resonance imaging
ms	Millisecond(s)
NMDA	N-methyl-D-aspartate
PET	Positron emission tomography
RM-ANOVA	Repeated measures analysis of variance
(r)pIFG	(Right) Posterior inferior frontal gyrus
(r)pSTG	(Right) Posterior superior temporal gyrus
s	Second(s)
SD	Standard deviation
STG	Superior temporal gyrus
tDCS	Transcranial direct current stimulation

TMS	Transcranial magnetic stimulation
USB	Universal serial bus

Zusammenfassung

Zielsetzung und Hypothese:

Während mittels bildgebender Verfahren ein Netzwerk von Hirnregionen identifiziert werden konnte, das der Wahrnehmung und Wiedergabe von Lauten dient, ist die genaue Funktion einzelner Knoten innerhalb dieses Netzwerkes unbekannt. Für die willkürliche Kontrolle der menschlichen Stimme postulieren wir eine kausale Rolle des Gyrus temporalis superior posterior (pSTG) und des Gyrus frontalis inferior posterior (pIFG) der linken und rechten Hemisphäre. In dieser Arbeit wird die Funktion dieser Areale durch reversible Läsionsstudien mithilfe der transkraniellen Gleichstromstimulation (tDCS) untersucht.

Methoden:

In zwei Experimenten, bestehend aus je fünf separaten Stimulationen, erhielten die Probanden je eine kathodische Stimulation über dem pSTG und dem pIFG jeder Hemisphäre sowie eine Placebostimulation. Im ersten Experiment sollten die Probanden (n = 10) eine Tonfolge wiedergeben, wobei die Genauigkeit der Tonwiedergabe nach erfolgter tDCS gemessen wurde. Im zweiten Experiment hörten die Probanden (n = 15) sich selbst summen. Dieses Feedback wurde unerwartet in der Tonhöhe verändert. Die Kompensation zum transponierten Feedback nach Hirnstimulation wurde gemessen und mit Daten von Amusikern (n = 8), die keine Hirnstimulation erhielten, verglichen. Alle Daten wurden digital aufgenommen, weiterverarbeitet und ausgewertet.

Ergebnisse:

Nach Stimulation über dem linken pIFG und dem rechten pSTG wird die Tonhöhe einer zu imitierenden Tonfolge weniger exakt wiedergegeben als unter Placebostimulation. Des Weiteren nehmen alle nicht-amusischen Probanden ungewollte Veränderungen im Feedback ihrer eigenen Stimme wahr und kompensieren ohne Stimulation hierfür. Amusiker hingegen zeigen keine Reaktion auf verändertes Feedback. Reduziert man die neuronale Erregbarkeit in dem linken pIFG und dem rechten pSTG, kompensieren musikalisch geschulte Probanden in geringerem Maße als ohne Stimulation. Probanden ohne musikalisches Training zeigen keinen Einfluss der tDCS auf das Ausmaß der Kompensation. Unabhängig von musikalischem Training reagieren jedoch alle

Probanden nach kathodischer tDCS über dem rechten pSTG langsamer auf das veränderte Feedback ihrer eigenen Stimme.

Fazit:

In dieser Arbeit konnte gezeigt werden, dass sowohl der Gyrus frontalis inferior posterior als auch der Gyrus temporalis superior posterior wichtige Knotenpunkte in einem bihemisphäriellen Netzwerk der Stimmkontrolle sind. Die neuronalen Mechanismen einer effizienten Stimmhöhenregulation sind zu einem gewissen Maße trainingsabhängig. tDCS ist eine geeignete Methode, die Lautproduktion zu modulieren und sollte vermehrt in der Therapie von Sprechstörungen zum Einsatz kommen.

Abstract

Objective:

While neuroimaging studies have identified a network of brain regions that are involved in sound perception and production, the roles of each node in this network are unknown. We postulate a causal role of the right and left posterior superior temporal gyrus (pSTG) and posterior inferior frontal gyrus (pIFG) in vocal pitch control. In this study we test this hypothesis by creating temporary reversible lesions using transcranial direct current stimulation (tDCS).

Methods:

Subjects' performance in two experimental tasks was measured over five separate sessions each. They received unilateral cathodal stimulation over pSTG and pIFG for each hemisphere separately and one sham stimulation session. In the first experiment subjects (n = 10) performed a pitch reproduction task and accuracy following tDCS was tested. In the second experiment subjects (n = 15) heard their own voice fed back during hummed vocalization. Feedback was altered unexpectedly in pitch. Subjects' compensation for the transposed feedback after brain stimulation was measured and compared to data from amusic subjects (n = 8) that did not undergo brain stimulation. All data was digitally recorded and analyzed using custom-built software.

Results:

Pitch matching accuracy is impaired after cathodal stimulation over left pIFG and right pSTG compared to sham. In addition, all non-amusic subjects are sensitive to changes in auditory feedback and compensate without brain stimulation. Tone-deaf subjects show a lack of reaction to changes in feedback. When reducing neural excitability in both the left pIFG and the right pSTG, musically trained subjects compensate to a smaller amount when presented with transposed feedback of their own voice. Subjects without musical training do not show an effect of tDCS on magnitude of compensation. Independently from musical training, all subjects react with greater latency to pitch-shifted feedback after stimulation on right pSTG.

Conclusion:

In this study we were able to demonstrate that both the posterior inferior frontal gyrus and the posterior superior temporal gyrus are important nodes in the bi-hemispheric network involved in vocal pitch control. To a certain extent, the neural mechanisms underlying efficient vocal pitch regulation are experience-dependent. TDCS is a viable method to modulate vocal output and should be further employed as a therapeutic tool for speech disorders.

1. Introduction

1.1 The fronto-temporal network of vocalization

1.1.1 Historical background

Producing vocal sounds is an integral part of human communication that requires the simultaneous integration of multiple neural systems for sound perception, cortical representation, subsequent motor command initiation and finally concerted sound production [1, 2].

The basic principles of our current understanding of speech-related neural pathways go back to the 19th century with the pioneering work of Paul Broca and Carl Wernicke. In 1861, Broca described the case of a single patient named Tan (after his incapability to pronounce much more than “tan-tan”) who presented with extremely limited speech production but intact comprehension after a lesion of the frontal lobe [5] (Figure 1.1a).

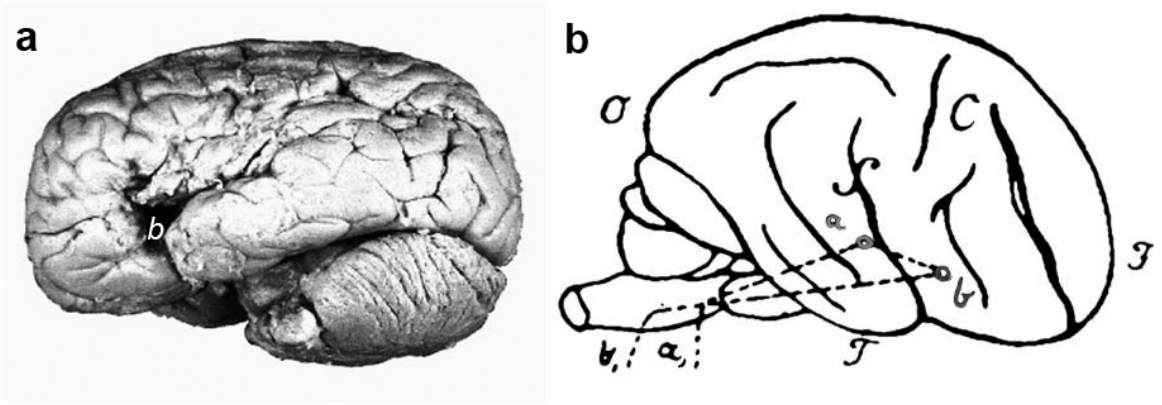

Figure 1.1: Broca's and Wernicke's area

a) Left hemisphere of Broca's famous patient Tan with a lesion in the inferior frontal lobe (labeled *b*) [3].
b) Wernicke's drawing (on the right hemisphere of a monkey brain, however), showing locations of today's Wernicke's (labeled *a*) and Broca's (labeled *b*) area. Dotted lines represent speech pathways - sensory input (*a*-*a'*) and speech output (*b*-*b*'). Modified from [4].

Broca’s report of an isolated production impairment after damage to the left frontal lobe is considered the first account of the specific localization of a cerebral function, here speech articulation [6]. A decade later, Wernicke’s description of conduction aphasia

marks the beginning of neuroanatomical models for an entire language network in the human brain [7] (Figure 1.1b). He describes the role of the left hemisphere arcuate fasciculus (AF) as a connection between motor planning areas in the frontal gyrus and sensory areas in the temporal gyrus [4]. Since then, subsequent lesion studies have supported the assumption that Broca's area or the dominant hemisphere's posterior inferior frontal gyrus (pIFG), encompassing Brodmann's area 44/45, is mainly involved in speech production whereas the corresponding posterior superior temporal gyrus (pSTG) or Wernicke's area subserves speech perception [8-10] (Figure 1.2). The dominant hemisphere has been shown to correspond to the left hemisphere in over 95% of the right-handed population and at least 70% of the left-handed [11, 12].

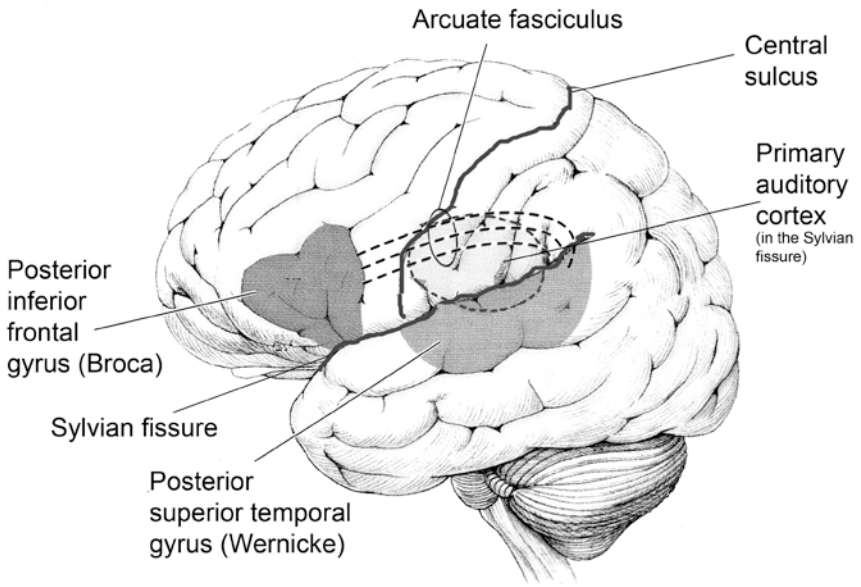

Figure 1.2: Neuroanatomy of the fronto-temporal network
Modified from [13].

Over the past decades, technical advances in, and widespread availability of, functional neuroimaging techniques such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) have enabled researchers to further investigate the neural basis of language processing (for a review, see [14]).

1.1.2 Between speaking and singing

The clinical observation that non-fluent aphasic patients such as Broca's Tan are still able to sing normally despite major speech impairment has raised the question whether speaking and intoned speaking - that is singing - rely on the same neural mechanisms [15, 16]. Reciprocal hemispheric dominance for music and language has been proposed to explain the clinically observed dissociation between speech production and singing. This means the neural correlates of music processing are thought to be localized predominantly in right hemispheric areas that are homologous to language areas on the left side [17]. This assumption is further supported by therapies emphasizing singing as a treatment for post-stroke aphasia such as Melodic Intonation Therapy (MIT) [18, 19]. Observed improvements in speech fluency under MIT have been attributed to the fact that intoned speaking recruited increasingly right hemispheric "singing" equivalents of left hemispheric speech processing regions [20].

Nonetheless, reports of stroke patients that had lesions on either the right or left side of the brain and were still able to sing point more towards a bi-hemispherically organized network [21]. This bi-hemispheric organization is further supported by recent functional neuroimaging studies that suggest a significant overlap in both perception and production areas for singing and speaking [22, 23]. Functional magnetic resonance imaging measures brain activity as an increased hemodynamic response called blood oxygen level dependent (BOLD) signal [24, 25]. A fMRI study by Özdemir et al. that compared singing and speaking found activation in superior temporal gyrus (STG) and inferior frontal gyrus (IFG) in both hemispheres during pitch production (i.e. humming) [23] (Figure 1.3a). Furthermore, during vocal production of intoned speech (i.e. singing pitches with words) additional right-lateralized activation was observed in both STG and IFG as compared to humming (Figure 1.3b). When contrasting activation during singing with speaking, singing showed significantly stronger activation in the mid-portion of the STG, especially in the right hemisphere, supporting the idea of a bi-hemispheric network with a preference for the right side of the brain (Figure 1.3c). However, another fMRI study that investigated the production of pitches in different vocal registers showed bilateral activity in the inferior frontal areas but a stronger activation in the left IFG than in its right hemispheric homologue [26].

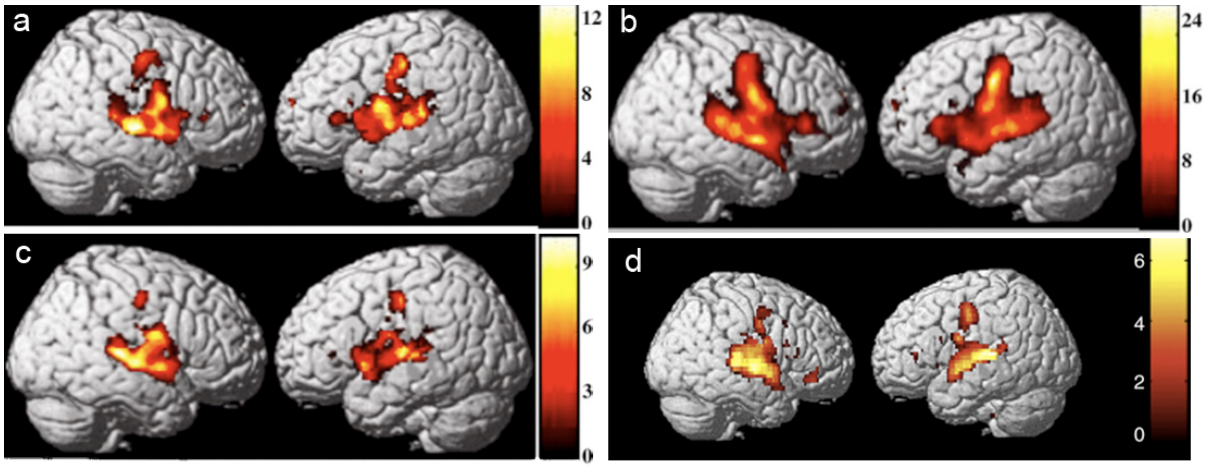

Figure 1.3: fMRI investigating brain activation during vocalization

BOLD responses to a) humming b) singing c) singing more than speaking contrast [23] d) perturbed auditory feedback [27], also see Section 1.2.2.

Hence it can be assumed that vocal production engages the STG and IFG in both hemispheres. Hemispheric laterality, however, remains a matter of debate and might depend on the specific performance task that is being tested.

1.1.3 Linking structural and behavioral abnormalities

In addition to mapping brain activity in healthy subjects, studies investigating individuals who have exceptional pitch processing abilities are informative of neural mechanisms underlying pitch production and perception. Both production and perception of pitches is significantly impaired in so-called tone-deaf or congenital amusic people, individuals who show substantial music processing difficulties but have otherwise normal peripheral hearing [28, 29]. One common characteristic is that they often have singing difficulties [30, 31]. This behavioral observation correlates with structural abnormalities in gray and white matter structures for both superior temporal and inferior frontal regions [32, 33]. Especially abnormalities in the inferior frontal gyrus (IFG) seem to play a crucial role in poor singing. It has been suggested that the IFG serves as a center for sound-motor mapping that is mapping pitches to corresponding vocal output [33, 34]. While several neuroimaging studies comparing tone-deaf individuals with controls have shown that the aforementioned brain regions are abnormal among tone-deaf people, the hemisphere

most affected in these studies differs between different publications.

At the other end of the spectrum, possessors of absolute pitch are characterized by the ability to name the pitch of any given tone without a reference, indicating unique and increased pitch perception and categorization skills [35, 36]. For both groups, results from diffusion tensor imaging (DTI) have demonstrated a significant increase respectively decrease in connectivity between the hypothesized brain regions subserving pitch perception and production. DTI is a MRI technique that analyzes the course of white matter tracts in the brain from diffusion-weighted images [37]. Using DTI, tone-deaf individuals show a marked decrease in connectivity in the arcuate fasciculus (the white matter tract that connects between superior temporal and inferior frontal areas) as compared to controls, corresponding to difficulties in both pitch perception and production [38]. Meanwhile, absolute pitch possessors have significantly larger volume in tracts connecting the superior temporal to the medial temporal gyrus, an area considered to be mainly involved in pitch categorization [35].

1.1.4 Clinical implications

Although the exact neural mechanisms underlying intoned speaking remain unclear, its use as a therapeutic tool has drawn increasing attention in recent years (for a review, see [16]). Besides the aforementioned speech deficits in post-stroke aphasia, therapeutic singing has been successfully implemented in clinical therapy for various other neurological conditions such as Parkinson's disease [39], stuttering [40, 41] and autism [39-42]. Apart from playing a role in these expressive speech-motor difficulties, vocal pitch processing abilities seem to be fundamental to the development of other communication skills such as reading as well. Pitch perception and production in children has been shown to be closely related to phonemic awareness and rapid auditory processing [43], both of which are reduced in children with developmental dyslexia [44] and illiterate adults [45]. Furthermore, remediation that focused on processing of acoustic stimuli showed significant improvement in language and reading skills in children with dyslexia [44].

Taken together, despite numerous neuroimaging and lesion studies and clinical experience, the exact cortical mechanisms underlying pitch perception and production remain poorly understood: while some research groups propose distinct brain systems

for speaking and singing [46], others advocate the view that both share the same brain network, at least to a certain extent [47]. Therefore, studying the neural mechanisms that subserve vocal pitch control is indispensable to further understand the neural mechanisms involved in everyday human communication. In the following chapter, we will briefly review the current model for a cortical voice control network, and then present a new approach to its investigation.

1.2 Investigating the neural mechanisms of voice control

1.2.1 Detecting self-generated sounds

To monitor one's own vocal production the speaker needs to distinguish between self- and externally-generated sounds. Electrophysiological recordings in humans have demonstrated that activity in the auditory cortex - located in the superior temporal lobe - is suppressed during vocal production [48]. This suggests that the auditory-motor system builds a precise forward model during sound production. Moreover, the auditory cortex modulates its activity as a function of the expected acoustic feedback: When the intended sound is produced, the auditory cortex attenuates its sensitivity to one’s own voice [49, 50]. When deviating from the intended vocal pitch, however, neural responses in the human auditory cortex are enhanced [51]. This might allow immediate correction of the detected production-perception mismatch, most likely via commands from the auditory cortex to the frontal motor cortex [52]. In addition, the concept of an area in the auditory cortex that is selective to one’s own vocal production is not limited to human vocalization but has been confirmed and extended by similar findings in various animal models [53-55].

1.2.2 Perturbed auditory feedback

One method that has been developed to simulate the neural processes underlying vocal feedback control is pitch-shifted or perturbed auditory feedback [56]. When using this method, subjects are asked to vocalize while they constantly hear their own voice

played back through headphones. At some point during the vocalization, this auditory feedback is unexpectedly transposed in pitch, so that the perceived feedback does not match the intended output anymore. Although the subject keeps vocalizing the intended pitch, the auditory system is tricked by the wrong feedback into thinking that they are out of tune. Normal subjects have been shown to react to this mismatch with an involuntary reflex compensation to the opposite direction of pitch shift, even if asked to hold their pitch steady [57, 58]. Experiments that investigated manipulations in loudness [59] or formant frequencies [60] during vocal output also showed automatic adjustments compensatory to the unintended changes. This further supports the theory for a cortical motor-control mechanism involving feedback from the auditory cortex [52].
Confirming this assumption, adaptation to pitch-shifted feedback led to increased activity in bilateral superior temporal cortex when applied during functional MRI [27, 34, 61, 62] (Figure 1.3d). Likewise, delayed auditory feedback was shown to activate the superior temporal area [63]. Especially the posterior temporal gyrus of the right hemisphere showed greater activation in pitch-shifted feedback studies [64, 65]. However, it remains unclear to what extent other parts of the fronto-temporal network are involved in voice-motor control. While some fMRI studies using pitch-shifted feedback have reported increased activation in bilateral prefrontal and premotor areas in addition to the temporal lobe activation [61, 62], these findings could not be repeated by more recent neuroimaging studies [27].

Taken together, convergent results from neuroimaging studies and populations with special pitch processing abilities suggest that the posterior superior temporal gyrus (pSTG) and the posterior inferior frontal gyrus (pIFG) are important nodes in the neural network that enables vocal pitch control. However, these reports have relied upon merely correlational observations of neural activity [66, 67]. No direct causal evidence exists for example from a circumscribed lesion in this network causing a dysfunction of the pitch perception or pitch production network.
Therefore, to be able to investigate the causality between a certain brain region and a particular behavior, we need to go beyond observation to the realm of intervention.

1.2.3 Non-invasive brain stimulation

Studies in cognitive neuroscience have employed mainly two tools to interfere with human brain function: transcranial magnetic stimulation (TMS) and transcranial direct current stimulation (tDCS) [68, 69]. Both are non-invasive focal stimulation techniques that modulate brain regions underlying the scalp being stimulated. Transcranial direct current stimulation (tDCS) alters the excitability of the underlying brain tissue via either hypo- or hyperpolarization [69]. In contrast to TMS, tDCS does not directly lead to neuronal discharge but changes resting membrane potentials which then lead to regional excitability changes [70]. This makes it a relatively safe tool increasingly employed in neuro-rehabilitative settings [71, 72].

For auditory research, tDCS has been suggested as a preferable method to TMS for two reasons: Firstly, tDCS is silent while TMS emits clicking sounds which may stimulate the auditory cortex and thus affect performance on auditory tasks. Secondly, tDCS is relatively painless for the subject, especially when compared to TMS. When applied around the ear, TMS also stimulates temporal muscles and some of the short lateral neck muscles, resulting in more discomfort when stimulation sites include areas such as the superior temporal and inferior frontal cortices [68, 73].

1.2.4 Transcranial direct current stimulation

Transcranial direct current stimulation (tDCS) uses a weak electrical current between two scalp electrodes. Thus, excitability changes are related to current strength, stimulus duration, and the direction of current flow, defined by electrode position and polarity (anodal vs. cathodal) [74-77]. Current strength usually ranges from 1 to 2 mA but it has been shown that current densities up to 25 mA/cm^2 do not cause damage to the brain tissue [78].

Studies have shown that tDCS-induced neural excitability modulations are associated with changes in performance in tasks that draw on the region that is being modulated. Behavioral effects that outlast the duration of stimulation are attributed to NMDA-receptors known to be involved in neuroplasticity through long-term potentation (LTP) and long-term depression (LTD) [79]. Additionally, stimulation with tDCS affects regional cerebral blood flow [80]. Stimulation with anodal tDCS has been demonstrated to

enhance task performance [81], whereas cathodal tDCS hypopolarizes or decreases the excitability of the underlying brain tissue. It can therefore create a virtual lesion which temporarily impairs function of that region [82-85]. To date, studies have shown that cathodal tDCS disrupts reaction time tasks when applied over motor areas, increases auditory frequency-discrimination thresholds when applied over Heschl's gyrus, and causes impairments in pitch memory when applied over the angular gyrus [73, 84, 85]. These results implicate tDCS as a viable method for inducing regional cortical dysfunctions. Furthermore, it has a reliable sham mode that allows it to conduct randomized controlled experiments [86]. However, no study so far has employed tDCS to explore the role of suspected key regions in the auditory-vocal control network.

1.3 Hypotheses and aims

Our aims in the current study were to apply non-invasive brain stimulation to test the causal role of the posterior superior temporal gyrus and posterior inferior frontal gyrus in the neural network that subserves pitch production and perception, and to further investigate the hemispheric laterality of auditory feedback control. For two different experimental tasks, we applied tDCS over the pSTG and pIFG in each hemisphere separately to create temporary reversible lesions. Hereby, we tested the effects of these localized disruptions on pitch matching ability as well as on sensitivity to perturbed auditory feedback.

In the first experiment, we focused on pitch reproduction, employing a simple pitch matching task optimized for each individual subject. The task consisted of eight pure target tones that had to be repeated as accurately as possible, requiring subjects to develop an exact motor plan before each utterance. Following stimulation with tDCS, we expected decreased accuracy in pitch production. We hypothesized that accuracy would be most affected after stimulation over bilateral inferior frontal gyri due to disrupted sound-motor mapping but had no clear hypothesis on hemispheric laterality.

For the second experiment, we employed the pitch-shifted feedback paradigm to investigate neural mechanisms integrating vocal motor control with auditory feedback during an utterance. Humans with intact pitch perception are sensitive to changes in

auditory feedback and compensate when given false feedback. Since electrophysiological and fMRI studies have shown that voice monitoring involves both pIFG and pSTG, brain stimulation was applied over each of these brain regions separately to modulate subjects' response to their own vocal feedback. In addition, preliminary data from tone-deaf individuals, i.e. subjects with impaired pitch perception abilities and structural abnormalities in these areas, showed a lack of sensitivity to perturbed auditory feedback [87]. This resulted in decreased amplitude of compensation. We therefore expected decreased sensitivity to perturbed auditory feedback after non-invasive brain stimulation as compared to sham stimulation, presenting with a tone-deaf-like task performance in normal subjects.

2. Material and Methods

2.1 Experiment 1: Pitch matching

2.1.1 Subjects

Ten right-handed individuals (4 female) from the Greater Boston area were recruited via online advertisements and were compensated for participating in this study. Subjects were aged between 21 to 28 years (mean age: 25). Inclusion criteria for the study comprised: 1) no history of hearing problems or neurological/psychiatric disorders, and 2) a pitch discrimination threshold of less than 5 Hz. This pitch discrimination threshold was assessed at 500 Hz for each subject using a three-up-one-down adaptive staircase procedure [88]. Subjects had a mean of 7.4 years (range: 0 to 21 years) of active music experiences including playing musical instruments. However, none of them was a trained singer or a professional musician. Including the initial practice session, all subjects were required to participate in a total of six experimental sessions on six different days. One subject withdrew from the study after the second session of stimulation, resulting in nine complete datasets being included in the final analysis. Written informed consent was obtained from each subject as approved by the Institutional Review Board of the Beth Israel Deaconess Medical Center.

2.1.2 Procedure

2.1.2.1 Transcranial Direct Current Stimulation

We conducted one practice session of hummed pitch reproduction prior to applying non-invasive brain stimulation to ensure that all subjects were familiar with the experimental procedures. This was followed by applying one tDCS session per day on five subsequent days to avoid carryover effects between stimulation sessions. The order of stimulation was counterbalanced across subjects. We identified the location of four relevant brain areas, whose role in pitch reproduction were of interest, in the international 10-20 system traditionally used for placing electrodes for EEG recordings. Locations were identified using LORETA [89] that allowed us to map the target regions in the 10-20 system onto the model brain (Figure 2.1 a,b).

These four brain regions were:

1) right posterior superior temporal gyrus (pSTG), which was identified in the international 10-20 system for EEG sites as one third of the distance between TP8 and C6;

2) left pSTG, which was one third of the distance from TP7 to C5;

3) right posterior inferior frontal gyrus (pIFG), which was one third of the distance from F8 to C6;

4) left pIFG, which was one third of the distance from F7 to C5.

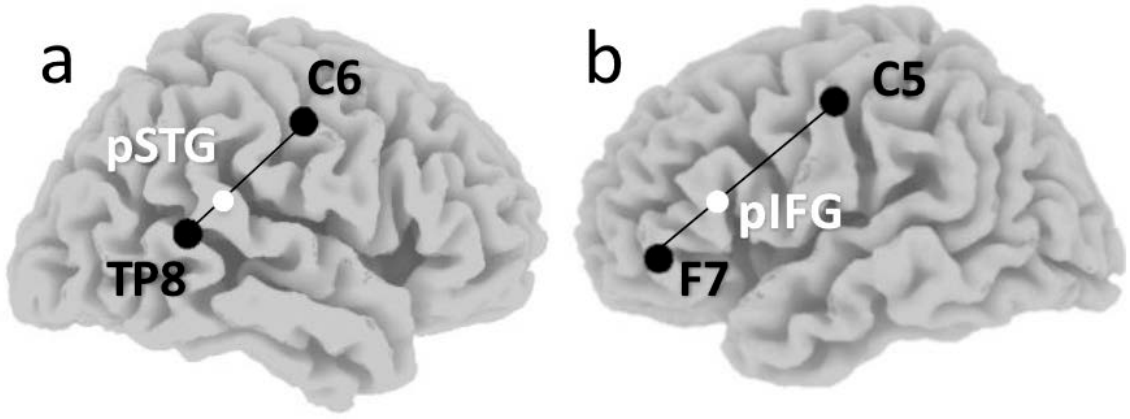

Figure 2.1: Targeted brain regions in the international 10-20 system
a) right pSTG b) left pIFG.

In addition to these four brain regions, sham stimulation (described below) was also conducted on one randomly selected region among the four regions identified above.

To verify that the active electrode was placed over the expected region of the cortex, anatomical T1 images of a subset of subjects were obtained using a 3T GE MRI scanner. The brain regions of interest were identified using the international 10-20 system and a marker was placed on the scalp over each of the regions.

During the application of cathodal tDCS, the saline-dampened active electrode (using an oval electrode size of 16.3 cm^2) was placed over the target region. The reference electrode (a square electrode of 25.0 cm^2) was placed over the contralateral supraorbital region where it was functionally ineffective in this experimental design [90]. A current strength of 2.0 mA was applied for 20 minutes, using a battery-driven,

constant-current stimulator (Phoresor II PM850, Iomed Inc., Salt Lake City, UT, USA). For the sham session, the cathodal electrode was placed over one of the four target regions, which were counterbalanced between subjects, and the reference electrode was placed over the contralateral supraorbital region. To administer the sham stimulation, current was ramped up for 30 seconds until it reached 2.0 mA, and then turned to zero for the next 30 seconds and kept at zero for the remaining time period. All participants reported a tingly sensation under the cathodal and/or reference electrode when ramping up the current at the beginning of the stimulation, which was the same for real as well as for sham stimulation and faded away after approximately 1 minute. Participants were unable to distinguish sham stimulation from real stimulation according to their own verbal report, a finding consistent with similar experimental procedures [91]. During stimulation, subjects read a magazine or a book.

2.1.2.2 Pitch production task

At the start of the first session for each subject, the subject was asked to hum tones that were within their vocal range in order to determine each subject's comfortable vocal range. This ranged from 151 to 262 Hz across all subjects. Subjects were then presented with one target tone within their vocal range and were asked to reproduce that tone as a practice trial. After the initial practice trial, eight pitch reproduction trials were recorded. Each trial consisted of one different sine wave tone. Target tones were centered around each subject's comfortable vocal range as assessed initially and ranged from 2 semitones below to 3 semitones above the center frequency. They were presented in the same ascending order during each experimental session with interval steps of either 1 or 0.5 semitones. Therefore, we assumed that potential learning effects affected all stimulation conditions equally and did not constitute a relevant confounder in this experimental design. All tones were presented with equal amplitude (70 dB) and duration (1000 ms, smooth envelopes with rise and decay times of 50 ms each) through Altec Lansing headphones (AHP512i). Subjects were asked to reproduce the target pitch by humming the perceived pitch for 3 seconds. Vocal production was recorded digitally in Praat [92] via a USB microphone (Logitech 980186-0403 USB Desktop Microphone) for subsequent offline analysis.

2.1.3 Data analysis

Subjects' vocal production was recorded and pitch extraction was applied offline using Praat. Since stimulation was predicted to have the greatest effect on initial vocal-motor planning and preparation, only the average of the first 500 ms was analyzed for each of the eight pitches produced per subject. For statistical analysis, all frequencies were converted from absolute frequency in Hertz to relative deviation from target frequency in cents of a semitone (100 cents = one semitone) using the following formula:

$$\text{Cents Deviation} = 1200 * \log_2 (F_{produced}) - \log_2 (F_{target})$$

where F_{target} is the target frequency and $F_{produced}$ is the produced fundamental frequency. Cents deviation scores for each individual trial were exported to SPSS 19 [93] for statistical analysis and tested visually for normal distribution. A repeated measures ANOVA was run on the dependent variable of cents deviation from target with the factor of stimulation condition. If the assumption of sphericity was violated, probability values were corrected using Greenhouse-Geisser estimates and corrected p values were reported along with the original degrees of freedom [94]. An alpha level of .05 was used for all statistical tests.

2.2 Experiment 2: Perturbed auditory feedback

2.2.1 Subjects

Fifteen right-handed subjects (8 female) with a mean age of 25 years (range: 21 to 28) from the Greater Boston area were included in the study. Inclusion and exclusion criteria were the same as in the first experiment. Using a three-up-one-down adaptive staircase procedure, the individual pitch discrimination threshold at 500 Hz was assessed for each subject (mean at 500 Hz: 1.87 Hz, SD = 1.03). Subjects had a mean of 6.9 years of musical training (range: 0 to 21 years) but none of them was a trained singer or a professional musician. For subsequent analysis subjects were divided into two groups: musicians, as defined by 6 or more years of musical training (8 subjects, 4 female, mean = 11 years of musical training) and non-musicians with minimal musical

exposure (7 subjects, 4 female, mean = 2 years). As in the first experiment, subjects were required to attend a total of six experimental sessions on six different days and were compensated for their participation. In addition, 8 tone-deaf subjects (5 female) with a mean age of 25 years (range: 21 to 33) were identified by self-report and verified with the Montreal Battery of Evaluation for Amusia (MBEA) [95]. Normal IQ was ensured using Shipley's abstract and verbal scaled composite score [96] and their individual pitch discrimination threshold was assessed at 500 Hz (mean at 500 Hz: 32.7 Hz, SD = 8.7). These subjects came in for only one session to perform the perturbation experiment and were compensated for their participation. Written informed consent was obtained from each subject as approved by the Institutional Review Board of the Beth Israel Deaconess Medical Center.

2.2.2 Procedure

2.2.2.1 Transcranial Direct Current Stimulation

Similar to experiment 1, each subject underwent four sessions of cathodal stimulation with tDCS – each one for right and left pSTG and right and left pIFG – and one additional sham session on one of those areas. Stimulation sites were identified the same way as in experiment 1 using the international 10-20 system for EEG. They were verified for a subset of subjects by obtaining anatomical T1-weighted MRI images that were correlated with MRI compatible markers on the scalp. To avoid carry-over between regions, stimulation was applied on five different days. Order of stimulation was counterbalanced across subjects and the site for sham stimulation was randomly chosen from one of the four regions identified above.
Tone-deaf subjects did not undergo any stimulation.

2.2.2.2 Altered feedback task

Prior to the first session, the subject was asked to hum tones that were within their vocal range in order to assess each subject's comfortable vocal range. Before the first day of stimulation, one entire session was conducted as a practice trial to familiarize subjects with the experimental procedures.

For the altered feedback task, subjects were presented with a sine wave tone within their vocal range (the range across all subjects was from 151 to 330 Hz), and were asked to reproduce this pitch by humming for at least 5 seconds. They were advised to hum at an amplitude between 65 and 70 dB, resulting in a feedback loudness of 75 to 80 dB which masked most of the bone conduction. No masking noise was used in this study to maximize the perception of vocal sounds as being self-generated. Using a custom-built MIDI software program (Max/MSP 5.0 by Cycling 74 [97]), subjects' vocal production was recorded through a USB microphone (Logitech 980186-0403 USB Desktop Microphone). Fundamental frequency (voice F0) of subjects' produced pitch was extracted online and played back through Altec Lansing headphones (AHP512i) during the vocalization in real time. In 60% of the trials, feedback was shifted in frequency 1.5 – 2.5 seconds after onset of vocalization so as to give participants the impression of being out of tune (Figure 2.2).

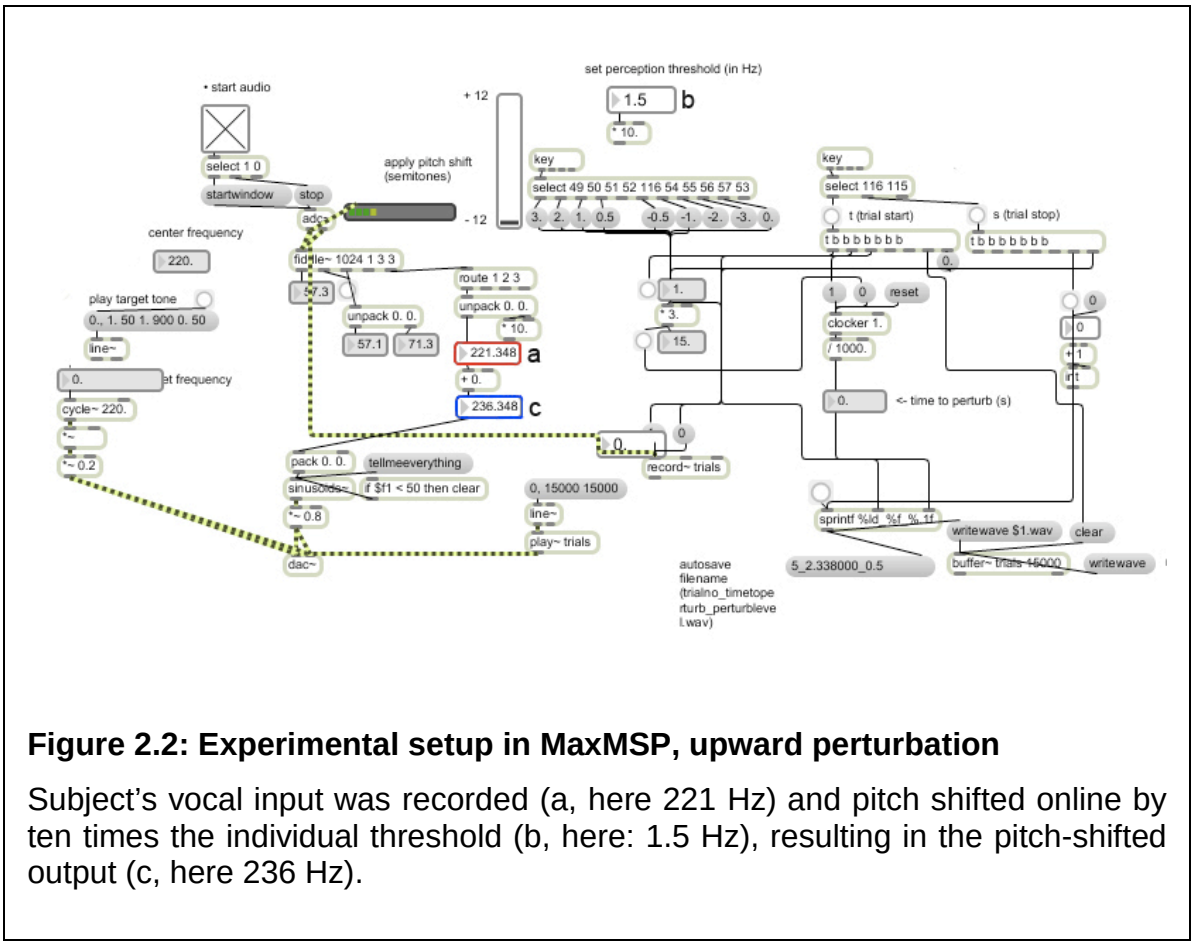

Figure 2.2: Experimental setup in MaxMSP, upward perturbation

Subject's vocal input was recorded (a, here 221 Hz) and pitch shifted online by ten times the individual threshold (b, here: 1.5 Hz), resulting in the pitch-shifted output (c, here 236 Hz).

The transposition was either upward or downward in randomized order and lasted until the end of that trial. The level of transposition was adapted to ten times each participant's individual psychophysically-defined pitch discrimination threshold in Hertz

as assessed initially. Onset time and direction of transposition were automatically marked by the software program for further analysis. Each trial was initiated with the same sine wave tone. Subjects were asked to hum this pitch as they heard it before perturbation onset and to maintain their vocal pitch to the best of their ability throughout each trial (Figure 2.3).

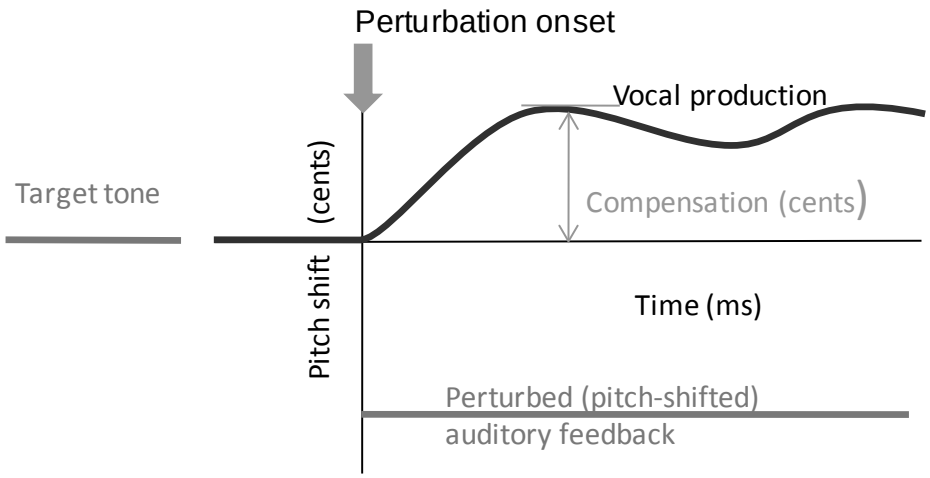

Figure 2.3: Pitch-shifted feedback schematic

The figure illustrates a trial with downward perturbation and compensatory upward shift in subject's vocal production.

Altogether, we recorded responses from 15 subjects for 5 stimulation conditions, that is right and left superior temporal, right and left inferior frontal gyrus as well as sham. In addition, 8 tone-deaf subjects recorded one session without any brain stimulation. Each run included 6 trials per pitch shift direction (6 trials with upward and 6 trials with downward perturbation) resulting in 180 sound recordings with perturbed auditory feedback per stimulation condition. Non-perturbed trials were used as a control condition for adequate vocal production but not included in the further analysis. In a subset of subjects, more than 6 trials per direction were recorded. These were only taken into account when one of the first six trials had to be discarded for technical reasons.

2.2.3 Data analysis

Pitch extraction was performed offline in Praat (using an autocorrelation method with time steps of 10 milliseconds) and then imported into Matlab 2012b [98]. Using custom-built software, the voice F0 values were time-aligned with the onset of the pitch shift stimulus and smoothed with a Gaussian filter (standard deviation $\sigma = 2$). Each subject's average baseline frequency ($F_{baseline}$) within 200 milliseconds before the onset of transposition was calculated. For statistical analysis, all frequencies were converted from absolute frequency in Hertz to relative deviation from that baseline in cents of a semitone (100 cents = one semitone) with the formula:

$$\text{Pitch (cents)} = 1200 * \log_2 (F_{produced}) - \log_2 (F_{baseline})$$

where $F_{produced}$ is the produced fundamental frequency in Hertz and $F_{baseline}$ each trial's pre-perturbation mean F0, averaged over 200 ms.

This way, subject's frequency deviation in vocal pitch was normalized to its pre-perturbation pitch production. For statistical analysis, trials were truncated at 200 ms before and 700 ms after the onset of perturbation [99]. Trials were sorted according to stimulation condition and pitch shift direction and averaged to generate one event-related response for each experimental condition per subject. Response magnitude and latency were calculated using Matlab. Response magnitude was measured as the deviation from pre-perturbation vocal pitch in cents at time point 700 ms.

Latency was defined as the first time point where subject's voice F0 exceeded two standard deviations above or below the mean of the pre-perturbation baseline [100] (Figure 2.4). For statistical analysis on response latency, only valid averaged responses were taken into account. A valid averaged response was defined as a response that deviated by more than 2 standard deviations (SDs) from the pre-stimulus mean F0 ($F_{baseline}$) with a response latency of at least 50 ms and a response duration of at least 60 ms [56, 101]. Using these validity criteria, Burnett et al. reported that between 50 - 75% of their subjects produced valid averaged responses to pitch-shifted feedback [56]. For our experiment, we therefore required at least 15 averaged responses out of possible 30 to be valid when conducted without any stimulation, e.g. for the sham condition. Four subjects that produced invalid trials for the sham condition were excluded from statistical analysis.

Data was then imported to SPSS and tested visually for normal distribution. A repeated measures analysis of variance (RM-ANOVA) was conducted to test for significant differences in response magnitude across conditions. If the assumption of sphericity was violated, probability values were corrected using Greenhouse-Geisser estimates and corrected p values were reported along with the original degrees of freedom.

For latency, a RM-ANOVA could not be conducted because with unequal cell size and missing data the assumptions on which it is based were violated [99]. We refrained from imputation because data had been excluded according to our validity criteria, meaning that we expected missing values to differ from the available data set. Therefore, a two-way ANOVA on response latency (in ms) was performed.

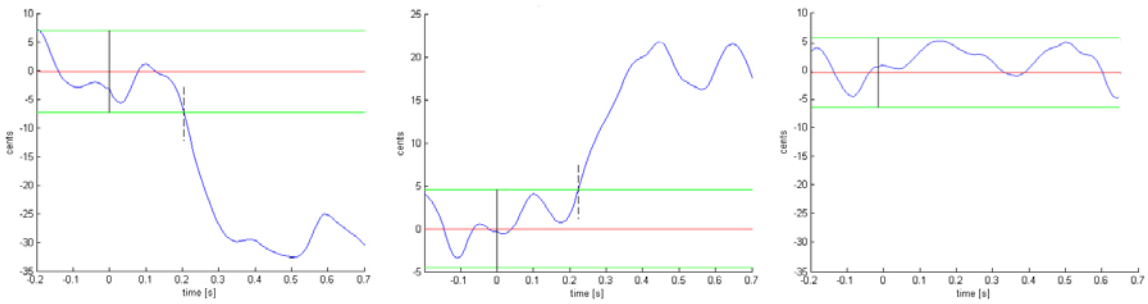

Figure 2.4: Calculating response latency in Matlab

Vocal responses to perturbation in cents for one subject, averaged across trials (n = 6). Upward pitch shift (left), downward pitch shift (middle), invalid trial (right). Green horizontal lines represent 2 SD above/below the pre-perturbation mean vocal pitch (red line), t = 0 indicates onset of perturbation.
Dashed vertical lines illustrate measurement of response latency.

3. Results

3.1 Transcranial direct current stimulation

3.1.1 Verification of stimulation site accuracy

The anatomical T1-weighted MRI images matched our predictions for electrode placements as identified using the 10-20 EEG system and confirmed accurate correspondence between the markers on the scalp surface and the anatomical regions of interest in the brain (Figure 3.1). Therefore, it was assumed that stimulation targeted the superior temporal and inferior frontal gyri on both hemispheres.

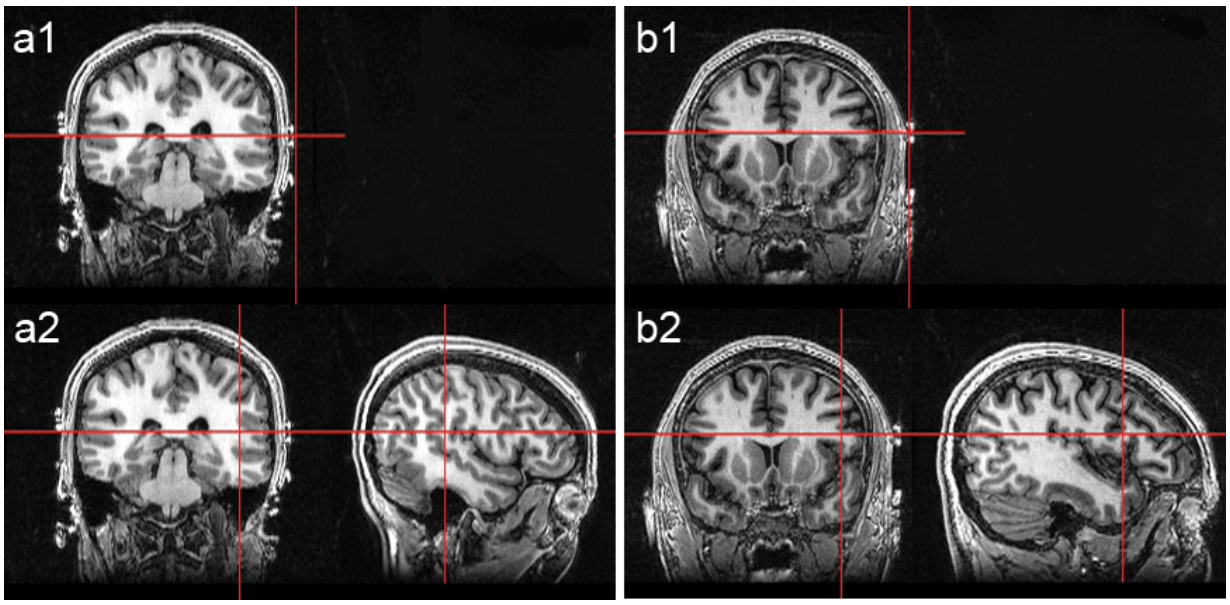

Figure 3.1: Stimulation sites

MRI scans confirming accurate location of stimulation sites:
a1) Marker placed on scalp over right pSTG.
a2) Coronal (left) and sagittal (right) slices showing crosshairs over the region underlying the marker, corresponding to targeted pSTG.
b1) Marker placed on scalp over right pIFG.
b2) Coronal and sagittal slices showing crosshairs over the region underlying the marker, corresponding to targeted pIFG.

In addition, subjects could not distinguish between cathodal stimulation and the sham session according to their own verbal report. They were not able to tell the difference in experiment 1 or in experiment 2, suggesting that differences in task performance after application of cathodal tDCS were effects of stimulation.

3.2 Experiment 1: Pitch matching

3.2.1 Decreased pitch matching accuracy

For the pitch reproduction task, subjects' deviation from the eight different target tones (2 semitones below to 3 semitones above center frequency) was calculated in cents as absolute value. Results were first plotted as a function of target tone, with lower cents deviation values indicating less deviation from target tone, meaning more accurate pitch matching (Figure 3.2).

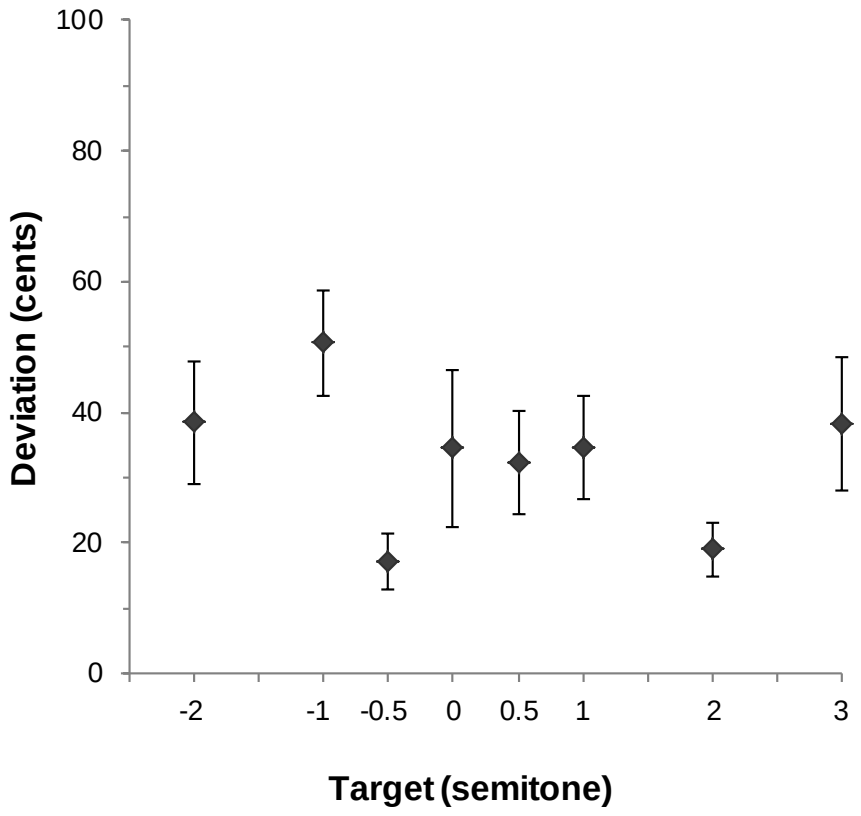

Figure 3.2: Deviation from target pitch (Sham)
Averaged across subjects (n = 9), error bars indicate between subject standard error.

A repeated measures analysis of variance (RM-ANOVA) was run on the dependent variable of deviation from target (in cents of a semitone) with the within-subject factor of stimulation condition (rpSTG, rpIFG, lpSTG, lpIFG, sham).

This revealed a significant effect of transcranial direct current stimulation condition on pitch matching accuracy, F(4,284) = 3.696, p = .009 (Figure 3.3).

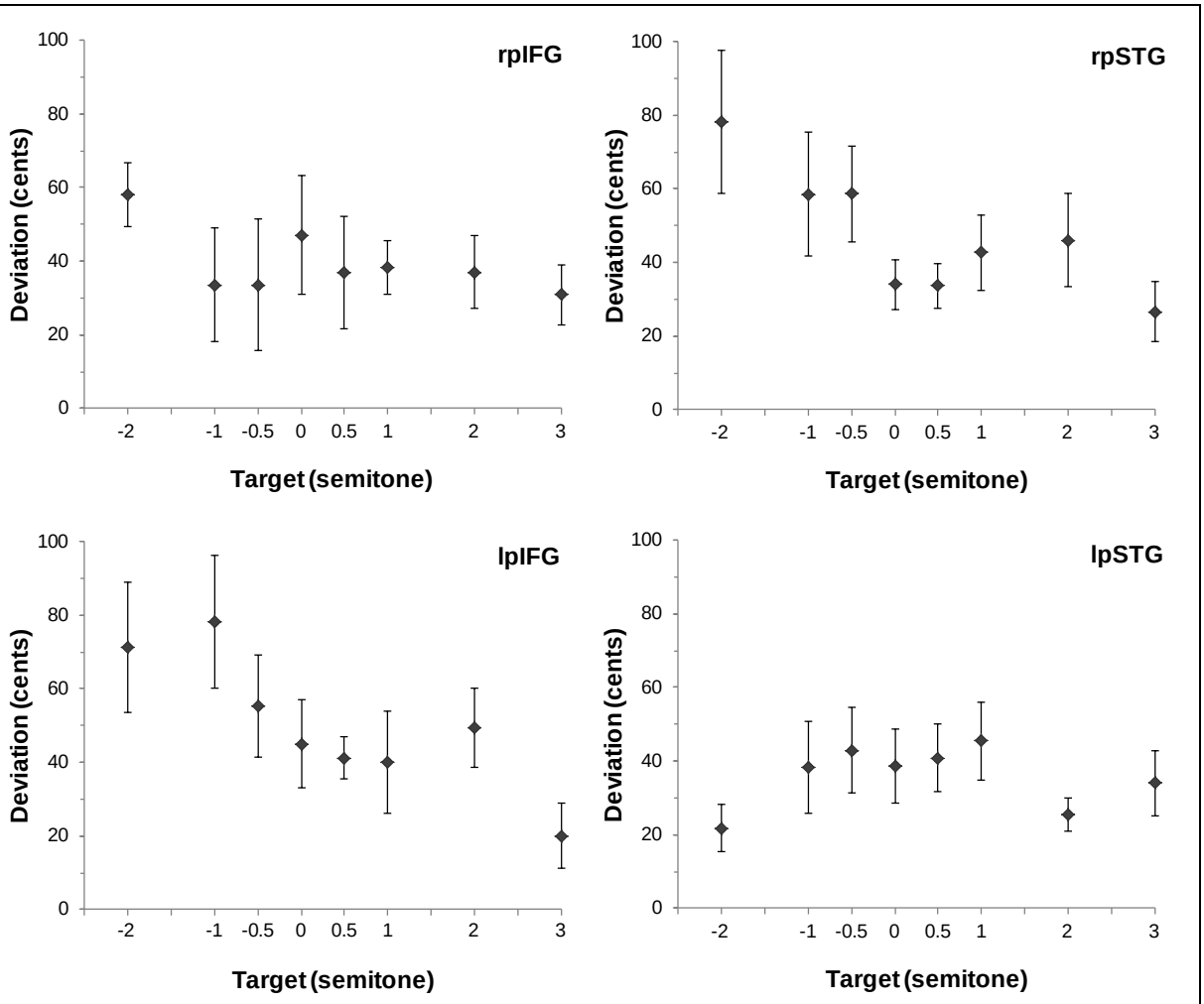

Figure 3.3: Deviation from target pitch after stimulation

Averaged across subjects (n = 9), error bars indicate between subject standard error. Deviation from target pitch after stimulation over right pIFG and pSTG (above) and left pIFG and pSTG (below).

To further evaluate the effect of the factor stimulation condition, post hoc t-tests were carried out, pairwise comparing the mean deviation from target between conditions. These pairwise comparisons revealed a significant decline in task performance after cathodal stimulation over left pIFG, $t(71) = 3.365$, $p = .001$, as compared to sham stimulation. In addition, a significant effect was found for right pSTG stimulation compared to sham stimulation, $t(71) = 2.832$, $p = .006$.

Subjects' mean deviation scores were lowest, i.e. performance was best following sham stimulation (*M* = 33.14, *SD* = 15.51). Mean deviation from target pitch was higher after cathodal stimulation over the left pIFG (*M* = 50.04, *SD* = 22.64) and after cathodal stimulation over the right pSTG (*M* = 47.30, *SD* = 17.98), confirming that tDCS on these two brain areas affected pitch matching accuracy (Figure 3.4).

No other pairwise comparison between cathodal stimulation and sham stimulation was significant, indicating that stimulation over left pSTG (*M* = 35.91, *SD* = 15.38) and right pIFG (*M* = 39.41, *SD* = 28.25) did not affect pitch matching accuracy significantly more than sham stimulation.

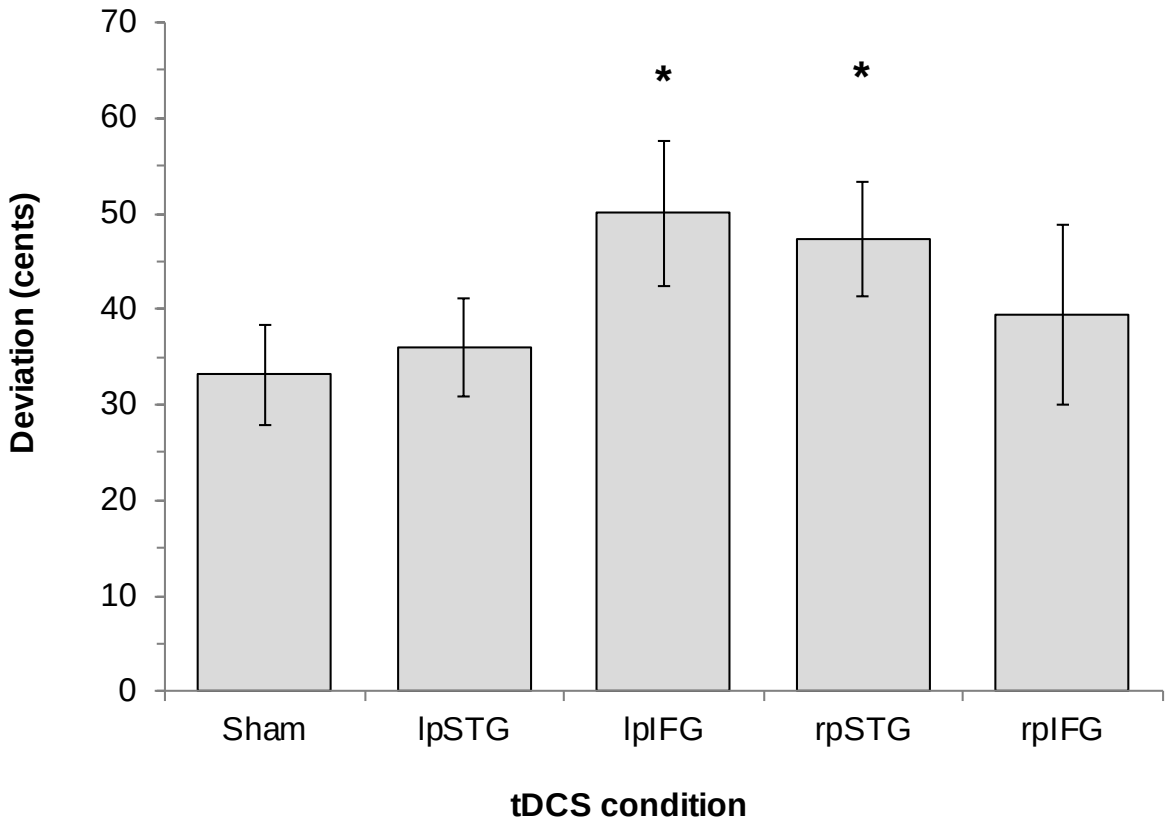

Figure 3.4: Mean deviation from target pitch

Results averaged across subjects (n = 9) and across trials. Error bars indicate between-subject standard error.
* = $p < .05$, paired t-tests with sham stimulation.

In addition, performance after stimulation over left pSTG differed significantly from stimulation over both left pIFG, $t(71) = 2.642$, $p = .01$, and right pSTG, $t(71) = 1.995$, $p = .05$. This suggested that decreased pitch matching accuracy after stimulation on these two brain areas could not be attributed to mere effects of cathodal tDCS. There was no significant difference in deviation from target tone between stimulation over left pIFG and over right pSTG, $t(71) = 0.445$, $p = .657$.

Average cents deviation showed no significant correlation with number of years of musical training, $r = .38$, *n.s.*, suggesting that musical training did not affect pitch matching performance among the subjects in this sample.

3.3 Experiment 2: Perturbed auditory feedback

3.3.1 Responses to perturbed feedback

3.3.1.1 Normal subjects

As expected, without non-invasive brain stimulation (that is, for the sham condition), normal subjects were sensitive to the perturbed auditory feedback. They compensated by producing vocal pitch in the opposite direction of the applied perturbation for the duration of altered feedback (Figure 3.5).

A one-way ANOVA on the dependent variable of pitch produced at 700 ms after onset of perturbation showed a highly significant effect of perturbation for subjects, $F(1,28) = 51.454$, $p < .001$. This compensatory response was on average -28.5 cents ($SD = 15.1$) for upward perturbation and 28.2 cents ($SD = 18.9$) for downward perturbation.

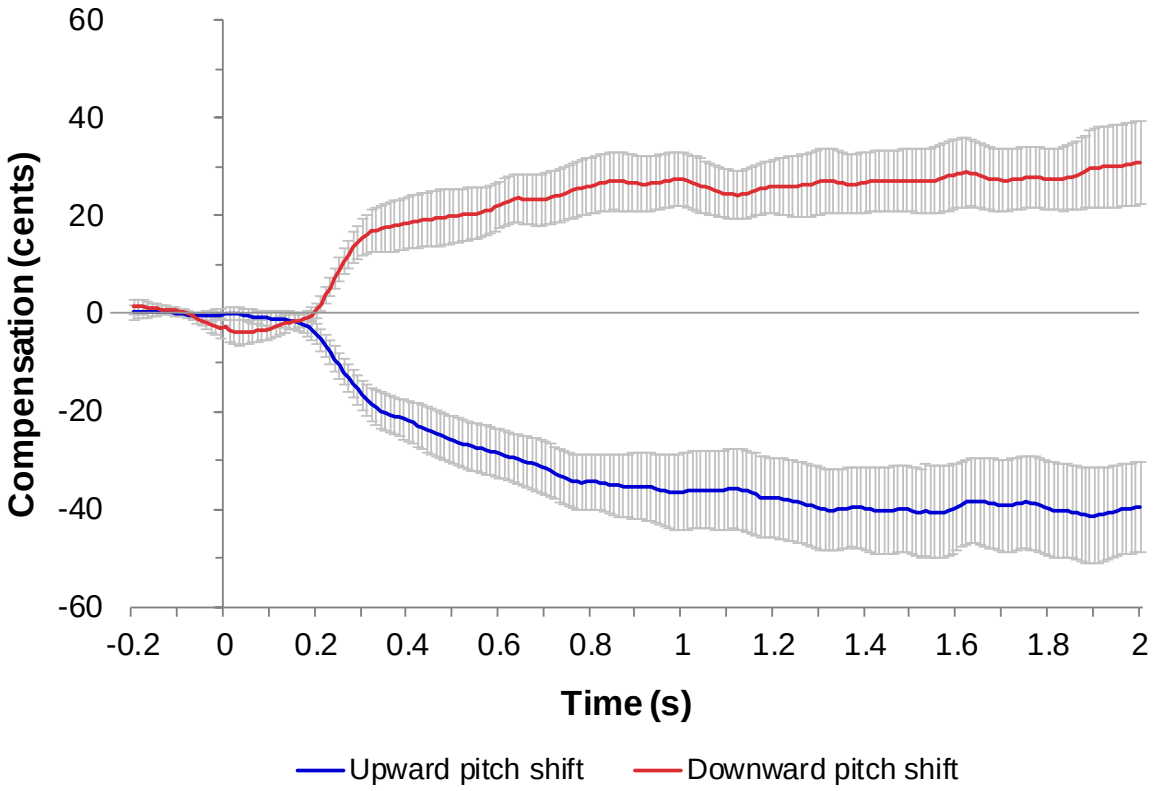

Figure 3.5: Compensation for pitch-shifted feedback (Sham)

Mean response in cents for the sham condition, averaged across subjects (n = 15) with standard error bars. Blue graph: upward shifted feedback, red graph: downward shifted feedback. The pitch shift occurred at t = 0.

Absolute magnitude of compensation did not differ significantly between the two pitch shift directions, $t(14) = 1.42$, $p = .117$, so that for subsequent statistical analyses, absolute values were calculated and data was collapsed across both directions.

3.3.1.2 Tone-deaf subjects

When tested at their individual pitch discrimination threshold, tone-deaf subjects showed no sensitivity to the perturbed auditory feedback (Figure 3.6).

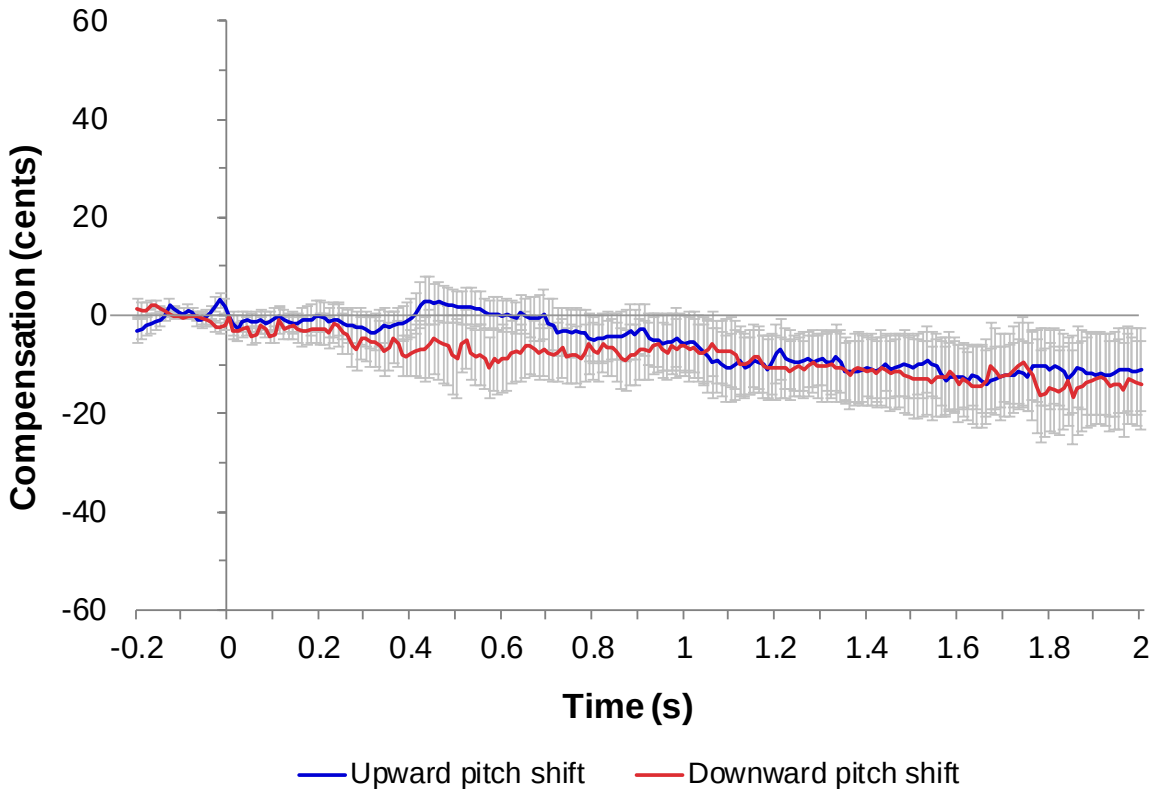

Figure 3.6: Compensation for pitch-shifted feedback (Tone-deaf)

Mean response in cents, averaged across all tone-deaf subjects (n = 8) with standard error bars. Blue graph: upward shifted feedback, red graph: downward shifted feedback. The pitch shift occurred at t = 0.

A one-way ANOVA on the dependent variable of pitch produced at 700 ms after onset of perturbation showed no significant effect of pitch-shifted feedback, $F(1,14) = 0.628$, $p = .441$.

Tone-deaf subjects' absolute response magnitude to pitch-shifted feedback ($M = 11.35$, $SD = 10.5$) was significantly smaller than normal subjects' compensation ($M = 27.6$, $SD = 20.6$) as revealed by an independent sample t-test, $t(43.96) = 3.544$, $p = .001$ (Figure 3.7).

Response latencies could not be calculated for this data set as tone-deaf subjects did not deviate significantly from their pre-perturbation baseline.

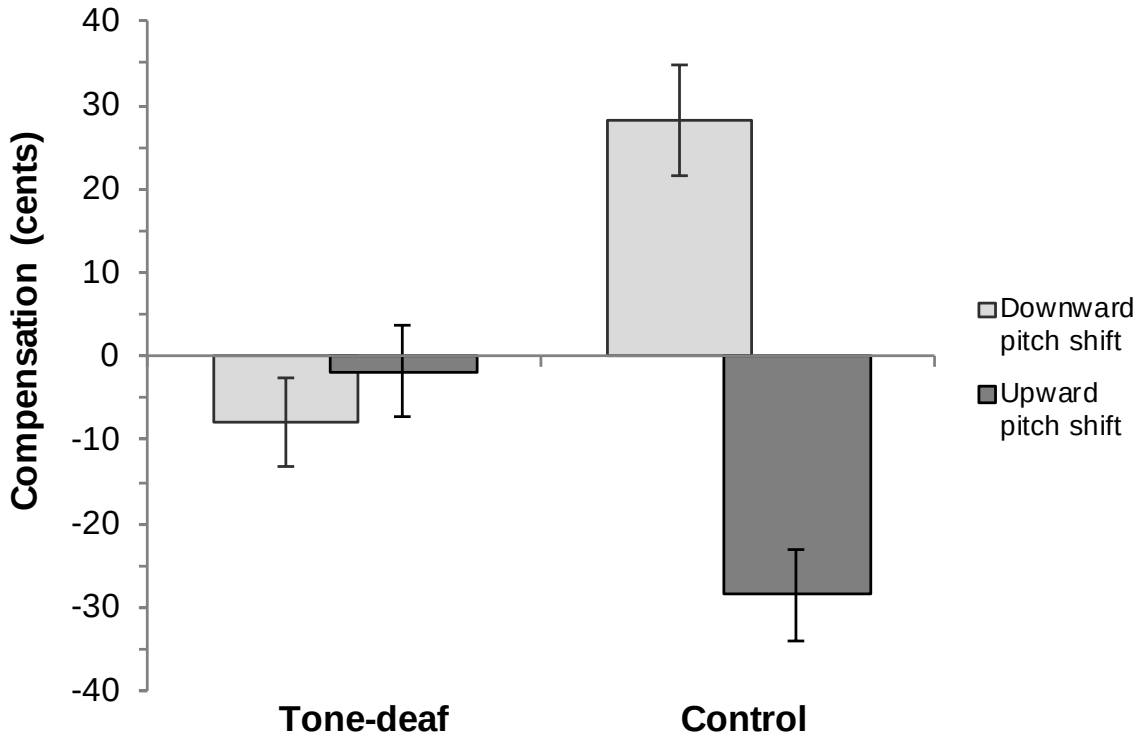

Figure 3.7: Response magnitude

Mean response magnitude after 700 ms in cents, averaged across tone-deaf subjects (n = 8) and normal control subjects (n = 15) with standard error bars, sorted by perturbation direction (upward, downward pitch shift).

3.3.2 Decreased response magnitude

To evaluate the effect of transcranial direct current stimulation on response magnitude in normal subjects, a repeated measures analysis of variance (RM-ANOVA) was run on the dependent variable of deviation from pre-perturbation pitch (in cents of a semitone) with the within-subject factor of stimulation condition (rpSTG, rpIFG, lpSTG, lpIFG, sham) and the between-subject factor of musicianship (musician, non-musician). This revealed no significant main effect of stimulation condition, $F(4,112) = 0.884$, $p = .476$,

but a significant main effect of musicianship, $F(1,28) = 6.626$, $p = .016$, and a significant interaction between stimulation condition and musicianship, $F(4,112) = 3.235$, $p = .015$. Due to significant stimulation x musicianship interaction, separate RM-ANOVAs were performed for each group. This revealed a significant main effect of stimulation condition on response magnitude for the musician group, $F(4,60) = 4.208$, $p = .028$, but not for the non-musician group, $F(4,52) = 0.847$, $p = .502$.

Results are shown in Table 3.1:

	Upward pitch shift		**Downward pitch shift**		**Mean response**	
Stimulation condition	**Musician (n = 8)**	**Non-Musician (n = 7)**	**Musician (n = 8)**	**Non-Musician (n = 7)**	**Musician (n = 8)**	**Non-Musician (n = 7)**
lpIFG	12.87 (10.5)	39.08 (24.1)	14.44 (16.6)	24.70 (18.7)	13.66 (13.44)	31.89 (22.02)
lpSTG	22.14 (12.0)	37.76 (28.1)	15.86 (16.2)	26.54 (24.6)	19.00 (14.16)	32.15 (26.02)
rpIFG	15.82 (10.8)	39.16 (19.5)	17.25 (14.7)	25.95 (18.2)	16.53 (12.51)	32.56 (19.37)
rpSTG	16.08 (9.1)	42.78 (20.8)	14.33 (17.5)	33.64 (23.0)	15.21 (13.54)	38.21 (21.59)
Sham	25.25 (20.0)	38.86 (21.4)	28.88 (23.5)	19.48 (14.8)	27.06 (21.19)	29.17 (20.34)

Table 3.1: Mean response magnitude in cents per condition

Sorted by musicianship and perturbation direction, standard deviation from mean in parentheses. All values are absolute values. For statistical analysis, data was collapsed across directions as mean response.

To further evaluate the effect of the factor stimulation condition, post hoc t-tests were carried out only for the musician group, pairwise comparing the response magnitude between stimulation conditions.

Subjects' mean deviation from pre-perturbation pitch was largest, i.e. compensation for feedback perturbation was most pronounced, following sham stimulation (*M* = 27.1, *SD* = 21.2). A significant decline in response magnitude was found for left pIFG stimulation (*M* = 13.7, *SD* = 13.4) as compared to sham stimulation, *t*(15) = 3.018, *p* = .009 (Figure 3.8).

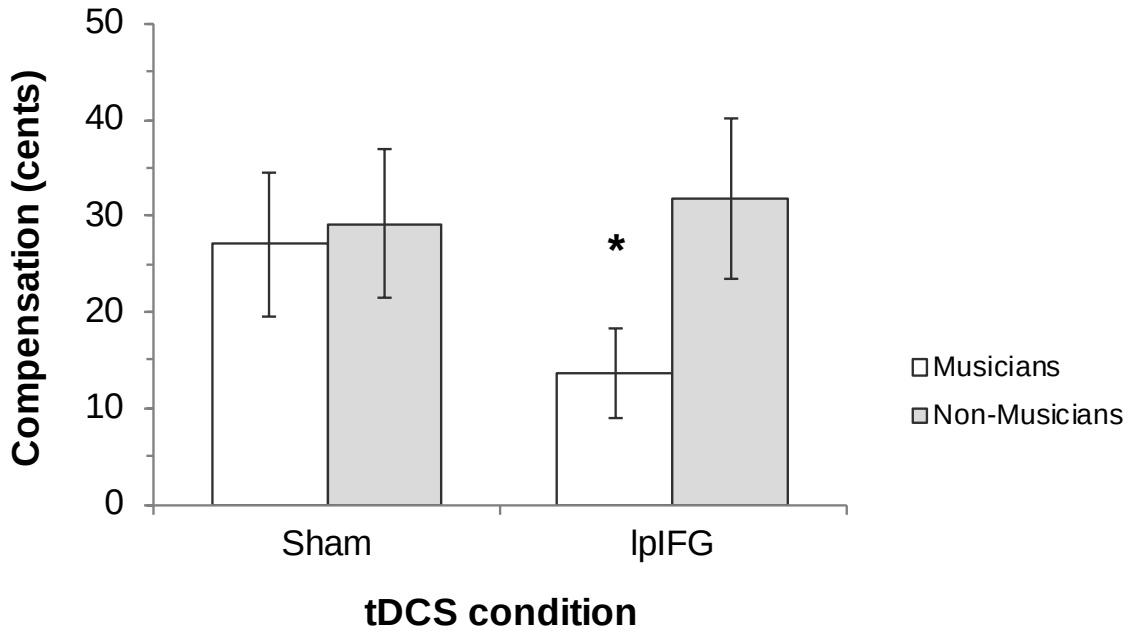

Figure 3.8: Compensation for pitch-shifted feedback after stimulation (lpIFG)

Mean response magnitude after 700ms in cents, averaged across perturbation directions for the musician group (n = 8) and the non-musician group (n = 7) with standard error bars.
Cathodal stimulation over left pIFG compared to sham stimulation.
* $p < .05$, paired t-test with sham stimulation.

Moreover, these pairwise comparisons revealed a significant decline after cathodal stimulation over right pSTG (*M* = 15.2, *SD* = 13.5) as compared to sham stimulation, *t*(15) = 2.214, *p* = .043 (Figure 3.9).

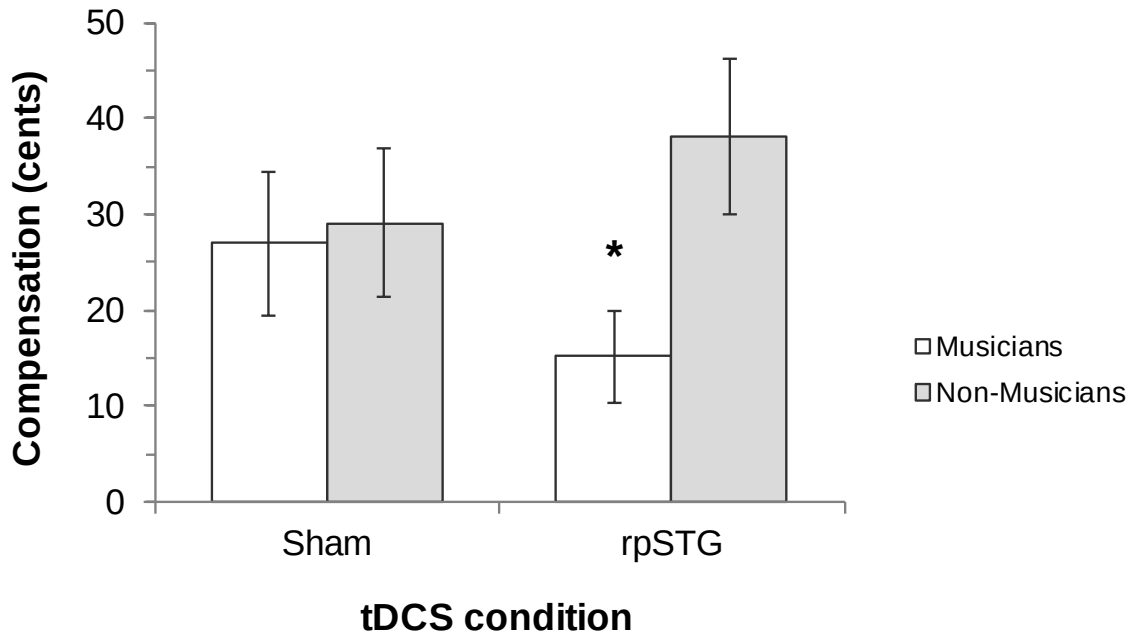

Figure 3.9: Compensation for pitch-shifted feedback after stimulation (rpSTG)

Mean response magnitude after 700ms in cents, averaged across perturbation directions for the musician group (n = 8) and the non-musician group (n = 7) with standard error bars.
Cathodal stimulation over right pSTG compared to sham stimulation.
* $p < .05$, paired t-test with sham stimulation.

This confirmed that tDCS on these two brain areas affected compensation for perturbed auditory feedback. No other pairwise comparison between cathodal stimulation and sham stimulation was significant, indicating that stimulation over left pSTG and right pIFG did not affect compensation for perturbed auditory feedback significantly more than sham stimulation.

In addition, response magnitude after stimulation over left pSTG (M = 19.0, SD = 14.2) was larger than after stimulation over left pIFG, $t(15) = 2.511$, $p = .024$. This suggested that decreased sensitivity after stimulation could not be attributed to mere effects of cathodal tDCS. There was no significant difference in compensation for perturbed auditory feedback between stimulation over left pIFG and over right pSTG, $t(15) = 0.573$, $p = .575$.

To further explore the differences between the musician and the non-musician groups, independent sample t-tests were conducted between groups for each stimulation condition. The two groups showed no significant differences in response magnitude when measured without any stimulation, that is for the sham condition, $t(28) = 0.277$, $p = .784$.

But when comparing performance after stimulation over right pSTG, musicians (M = 15.2, SD = 13.5) compensated significantly less than the non-musician group (M = 38.2, SD = 21.6), $t(21.299) = 3.439$, $p = .002$. Likewise, after stimulation over left pIFG, the musician group (M = 13.7, SD = 13.4) showed a significantly smaller compensation for the pitch-shifted feedback than the non-musicians (M = 31.9, SD = 22.0) after stimulation, $t(20.927) = 2.778$, $p = .014$.

3.3.3 Increased response latency

In addition to measuring the overall amount of compensation, we investigated whether it took subjects longer to react to the pitch-shifted feedback following cathodal stimulation with tDCS over the posterior inferior frontal gyrus and the posterior superior temporal gyrus. Following established validity criteria as explained in the method section, we identified 22 valid averaged responses for sham stimulation, 22 valid responses for cathodal stimulation over right pSTG, 21 for left pSTG, 18 for right pIFG and 19 for left pIFG, with the number of valid responses varying between subjects. There was no subject that produced only valid responses or one that produced only invalid responses and no difference in number of invalid trials was found between musicians and non-musicians, $t(4) = 0.451$, $p = .675$.

However, the downward perturbation direction yielded significantly more invalid trials (M = 5.8, SD = 1.4) across all five stimulation conditions than the upward perturbation (M = 3.8, SD = 0.8), $t(4) = 2.828$, $p = .047$. Therefore, two-way ANOVAs on the dependent variable of response latency in milliseconds with the factors of stimulation condition (rpSTG, rpIFG, lpSTG, lpIFG, sham) and musicianship (musician, non-musician) were conducted for each perturbation direction separately.

Results are shown in Table 3.2:

Stimulation condition	**Upward pitch shift**	**Subjects (included)**	**Downward pitch shift**	**Subjects (included)**
rpSTG	259.5 (140 – 654)	12 (9)	196.1 (96 – 512)	10 (8)
rpIFG	157.5 (50 – 255)	11 (9)	222.6 (60 – 652)	7 (6)
lpSTG	171.6 (60 – 248)	12 (9)	196.3 (60 – 320)	9 (7)
lpIFG	230.3 (81 – 485)	10 (8)	164.1 (56 – 412)	9 (6)
Sham	159.7 (60 – 259)	11 (11)	190.5 (50 – 305)	11 (11)

Table 3.2: Mean response latency in milliseconds per condition

Range across all subjects in parentheses. Only subjects that produced valid trials for the sham condition were included in the statistical analysis.

For the upward perturbation, this revealed a significant main effect of stimulation condition on response latency, $F(4,36) = 3.201$, $p = .024$, but no significant main effect of musicianship, $F(1,36) = 2.082$, $p = .158$, and no significant interaction between musicianship and stimulation condition, $F(4,36) = 0.933$, $p = .456$.

For the downward perturbation, we found no significant main effect of stimulation condition, $F(4,29) = 0.182$, $p = .946$, nor musicianship, $F(1,29) = 0.647$, $p = .428$, and no significant interaction between stimulation condition and musicianship, $F(4,29) = 0.303$, $p = .823$.

To further investigate the effect of transcranial direct current stimulation on the response latency to upward pitch-shifted feedback, independent sample t-tests were carried out. These revealed a significant increase in response latency for cathodal stimulation over the right pSTG ($M = 297.4$, $SD = 152.2$) as compared to sham stimulation ($M = 159.7$, $SD = 77.2$), $t(18) = 2.626$, $p = .017$. No other pairwise comparison between cathodal stimulation and sham stimulation was significant.

In addition, response latency after stimulation over right pSTG was significantly larger than right pIFG ($M = 151.6$, $SD = 72.5$), $t(16) = 2.596$, $p = .02$, and left pSTG ($M = 165.8$, $SD = 72.4$), $t(16) = 2.343$, $p = .032$, indicating that the increase in response latency after stimulation did not result from main effects of cathodal tDCS.

4. Discussion

In this study, we used non-invasive brain stimulation as a tool to investigate the causal role of four important nodes in the hypothesized brain network that subserves vocal pitch production and perception. We applied cathodal tDCS over the posterior superior temporal gyrus (pSTG) and the posterior inferior frontal gyrus (pIFG) in both hemispheres and investigated performance in two experimental tasks.

For the first experiment that focused on pitch reproduction we expected pitch matching accuracy to be most affected after temporarily blocking the function of the pIFG. Using perturbed auditory feedback in the second experiment, we hypothesized that sensitivity to unintended changes in vocal feedback would be decreased after stimulation with cathodal tDCS on both pSTG and pIFG.
Both initial hypotheses were supported by our experimental results: confirming findings from neuroimaging studies, both the pSTG and the pIFG are crucially involved in vocal pitch control. Applying cathodal stimulation over the left pIFG and the right pSTG resulted in decreased pitch matching accuracy and reduced compensation for perturbed auditory feedback as compared to sham stimulation.
Previous fMRI studies investigating the cortical network of vocalization have relied upon merely correlational observations. We are hereby able to demonstrate that intact function of the fronto-temporal brain network, centered around the pSTG and the pIFG, is required for efficient vocal pitch control.

In the following chapters, we shall discuss the results of our two experiments, including some general methodological aspects of the experimental design. Then we shall relate the results to the proposed model for a neural network of vocalization and finally present the potential clinical applications of our findings.

4.1 Experiment 1: Pitch matching

Pitch production ability was measured by mean deviation between produced pitch and target pitch. This measure, derived from acoustic analyses of recorded pitch productions, served as a reliable index of how far subjects' vocal production deviated from a given target pitch. Mean deviation from target pitch was largest, meaning that the effects of cathodal stimulation were strongest, over the left posterior inferior frontal gyrus (pIFG) and right posterior superior temporal gyrus (pSTG). Reducing excitability in those two areas independently impaired subjects' pitch matching accuracy compared to sham stimulation. In addition, deviation from target pitch was significantly larger after cathodal stimulation over the left pIFG and the right pSTG than after stimulation over the left pSTG, confirming that decreased accuracy in pitch production was not a main effect of cathodal tDCS itself.

These results provide causal evidence for a bi-hemispheric role in the execution and sensorimotor control of vocal production. The role of the right pSTG in pitch production, be it singing or humming, has been proposed by various neuroimaging studies before [23, 102, 103]. Some authors suggest that posterior auditory regions around the pSTG are involved in auditory-motor coupling that maps incoming sounds onto corresponding vocal motor representations [104]. We show that temporarily blocking the function of the right pSTG but not the left pSTG impairs the correct reproduction of perceived pitches, indicating a crucial role of this brain area in sound-motor coupling.

In addition, we found a significant decrease in pitch matching accuracy after reducing neural activity in the pIFG of the left hemisphere. The pIFG of the right hemisphere did not show a significant effect of stimulation. Since all our subjects were right-handed, not only was their language dominance located in the left hemisphere but they were also especially using their dominant left hemisphere when it came to fine adjustments in their motor apparatus [105]. We therefore hypothesize a similar lateralization for fine-grained motor control towards the left during vocalization, which would explain the decrease of precision in the pitch matching task after blocking the left pIFG but not its right homologue. Another important aspect is that the ability to map sounds correctly to vocal motor actions seems to be generally more of a left-hemisphere function [106]. Additionally, the left hemisphere's crucial role for short-term pitch memory might have

had a confounding influence on our subjects' performance [85]. Broca's area has also been shown, in addition to its traditional language function, as a sensorimotor integrator of sequential actions which would include a task like listening to and reproduction of several pitches within a short time-period [107, 108]. This significant activity of Broca's area during action-listening tasks has been attributed to mirror neurons in this region [106, 109]. Since both singing and speech acquisition develops through imitation of adult role models, we suggest that those mirror neurons play an important role during speech and music development [102].

In conclusion, our experimental results indicate that intact function of both the left posterior inferior frontal gyrus or Broca's area and the right posterior superior temporal gyrus contributes significantly to vocal production accuracy.

4.2 Experiment 2: Perturbed auditory feedback

Compensation for perturbed auditory feedback was measured both by deviation from pre-perturbation vocal pitch and by latency of compensation.

Our first finding was that naïve subjects without non-invasive brain stimulation compensated by producing vocal pitch in the opposite direction of perturbation. This indicates that subjects with intact pitch perception are sensitive to perturbed auditory feedback and adjust if given the impression of being out of tune. Tone-deaf or amusic subjects - subjects that are known for their inability to sing in tune - did not show any sensitivity to changes of their own vocal feedback, as demonstrated by the lack of reaction to the pitch-shifted feedback.

Our study was the first study that used non-invasive brain stimulation to interfere with the neural mechanisms integrating auditory feedback with vocal motor control. Therefore, our primary assumptions for changes in task performance after stimulation were mainly based on recordings from tone-deaf subjects. These subjects show structural abnormalities in the brain regions we stimulated that involve both reduced gray matter in the left posterior inferior frontal gyrus [33] and reduced white matter connectivity for the right posterior superior gyrus [38], structural differences that are highly correlated with impaired performance in pitch processing tasks.

The fact that our non-amusic subjects performed like tone-deaf subjects after cathodal

stimulation over the right pSTG and left pIFG provides a causal role for these two brain areas in the difficulties of music processing that accompany congenital amusia. Furthermore, it indicates that tDCS is a viable method for simulating disorders that have a suspected cortical dysfunction.

Our second finding was that stimulation with tDCS significantly affected subjects' response to their own vocal feedback. These effects were dependent on musical training and varied between upward and downward perturbation. Firstly, non-invasive brain stimulation with tDCS resulted in a significant reduction in magnitude of compensation for the pitch-shifted feedback in musically trained subjects. Secondly, cathodal stimulation increased subject's reaction time to the pitch shift independently from musical training.

While none of our subjects was a trained singer, various neuroimaging studies provide evidence that musical training in general is associated with changes in brain anatomy in both auditory and motor regions [110-112]. Therefore, we divided our subjects according to their years of musical training into two groups, musicians and non-musicians. This revealed that musical training significantly influenced the effects of transcranial direct current stimulation on compensation magnitude: The two groups showed no differences in response magnitude when measured without stimulation. However, stimulation with cathodal tDCS reduced the amplitude of compensation for both upward and downward pitch-shifted feedback only in the musician group but not in the non-musician group. Reducing excitability in both the left inferior frontal gyrus (pIFG) and the right posterior superior temporal gyrus (pSTG) independently impaired musicians' compensation for the transposed feedback as compared to sham stimulation. In addition, response magnitude was significantly smaller after cathodal stimulation over left pIFG and right pSTG than after stimulation over left pSTG, suggesting that decreased sensitivity could not be attributed to mere effects of cathodal tDCS.

Our experimental results are consistent with previous neuroimaging studies that showed greater activation in right superior temporal areas in musicians than in non-musicians during pitch-shifted feedback, indicating an enhanced recruitment of areas involved in audio-vocal integration [62]. Especially the STG seems to be increasingly employed with practice, that is when subjects learn to monitor their auditory feedback more closely in order to adjust their vocal output accordingly. In addition, frontal cortical areas have

been shown to be more engaged during musical tasks in musicians than in non-musicians, which is assumed to represent top-down influences of pre-existing knowledge [110, 113]. Our study was able to show that brain stimulation over both areas interfered with performance in musicians but not in musically untrained subjects providing causal evidence for an experience-dependent contribution of these two brain areas to vocal pitch control.
Furthermore, our findings support the assumption that musicians in general show a higher sensitivity to brain stimulation-induced changes in excitability [114] due to increased coupling between the auditory and motor systems [110]. While none of our musicians was a professional singer, the ability to integrate fine motor adjustments with auditory feedback is also important when playing an instrument such as violin or piano.

We therefore suggest that while both non-musicians and musicians are sensitive to their own vocal production and exhibit compensatory responses when giving the impression of being out of tune, the brain areas involved in monitoring their vocal pitch differ. Our study shows that musically trained subjects especially use the pIFG of the left hemisphere and the pSTG of the right hemisphere in order to regulate their vocal output. Thus, consistent with observed anatomical changes following musical training [115], our results provide further evidence that musical training induces changes in cortical activity. As it remains unclear what brain areas are predominantly employed in non-musicians, further research will have to be carried out to address this question.

In addition to measuring response magnitude, we measured response latency to pitch-shifted feedback. This revealed a significant increase in response latency after cathodal stimulation over right pSTG as compared to sham stimulation but no differences between musicians and non-musicians. Response latency was significantly larger after cathodal stimulation over right pSTG than after stimulation over left pSTG and right pIFG, indicating that increased reaction time was not a main effect of cathodal tDCS itself. These results extend recent findings by Chang et al. who recorded intracranial EEG during a perturbed auditory feedback task [52]. They found that neuronal activation in the right pSTG right before responding to the pitch-shifted feedback predicted subsequent compensation. Our study shows that responses to perturbed feedback are not only correlated with activity in the right pSTG but that disrupting this neural activity can significantly delay the compensatory response. This shows that tDCS is a viable

research tool to modulate the neural mechanisms underlying auditory feedback control. The fact that we found increased response latencies only for the upward pitch shift is in agreement with studies recording event-related potentials (ERPs) during pitch-shifted feedback that showed different neural responses depending on the direction of voice feedback perturbation [99]. We suggest that the audio-vocal system has separate ways to detect errors of voice being either too high or too low in pitch. While the exact neural mechanisms that enable speakers to discriminate in which direction they deviate from their intended vocal output have not been investigated so far, our results indicate a crucial role for the right pSTG in detecting deviations toward a higher pitch.

Taken together, we found that the amount of compensation for pitch-shifted feedback is regulated by both left pIFG and right pSTG and depends on musical training. Reaction time increased only after temporarily blocking the function of the right pSTG. A possible explanation might be that all subjects used the right pSTG to detect unintended changes in their own vocal feedback as indicated by increased response latency after tDCS. We suggest a common pathway for auditory error detection that is independent from musical training. However, the actual process of stabilizing one’s voice and compensating for potential pitch deviation - as measured by response magnitude in our experiment - is dependent on musical experience and was carried out via both right pSTG and left pIFG.

4.3 Methodological limitations

We identified several methodological limitations to our experimental design. Firstly, the number of subjects that participated in our experiments was the same or larger than in comparable studies that successfully employed transcranial direct current stimulation (tDCS) on pitch discrimination performance [73], and functional imaging studies that investigated pitch production [26] and perturbation [27]. However, this is the first study that investigates effects of tDCS on these tasks. In addition to the reported findings, there might have been subtle changes in performance that did not reach a level of statistical significance in our current sample size.
Secondly, although the underlying mechanisms remain as of yet unknown, parameters

such as age [116], gender [94] or native language [117] have been shown to play a role in voice feedback processing in healthy subjects. This is further evidenced by the fact that responses to altered auditory feedback in normally developing children depend on age and sex, suggesting a sex-specific development of the neural network involved in voice-motor control [118]. Besides controlling for musical training our subjects were therefore all chosen from the same age group and balanced for gender. However, future research might reveal additional parameters that we did not control for.

Furthermore, when analyzing response latency to perturbed auditory feedback, a crucial limitation might be the high number of invalid trials especially for the downward perturbation direction. We used established validity criteria that have been successfully employed for different subject groups and under varying experimental conditions in pitch-shifted feedback studies. Notwithstanding this, future stimulation studies employing perturbed auditory feedback might have to establish new validity criteria that take into account that by eliminating “bad” trials, some effects of stimulation are potentially eliminated at the same time.

With regard to non-invasive brain stimulation sites, accurate correspondence between electrode placement and targeted brain regions was verified using MRI scans of subjects' brains. But since we employed only unilateral stimulation in the current study, the contralateral pSTG and pIFG might have been able to compensate partly for the virtual lesions. Future studies could avoid this compensation by stimulating simultaneously over both hemispheres.

And finally, in addition to stimulating both pSTG and pIFG, one would ideally want to disrupt the arcuate fasciculus as the connection between these two areas. However, stimulating over the subcortically located arcuate fasciculus would also affect overlying cortical structures, including primary motor and sensory areas for articulation. We therefore refrained from applying tDCS over the arcuate fasciculus.

4.4 A cross-hemispheric fronto-temporal network

Taken together, results from both experiments provide support for a multi-regional network of brain areas in pitch perception and production, centering around the superior temporal gyrus and the inferior frontal gyrus.

Our first experimental task required vocal-motor planning and preparation before vocalization. Necessary steps involved initially perceiving the target frequency, generating a mental representation of the perceived pitch and finally mapping that pitch onto a motor plan for accurate reproduction. The second experiment focused on vocal pitch regulation during pitch production. That required the simultaneous integration of auditory feedback with motor control to implement appropriate motor adjustments under continuous self-monitoring. In addition, both tasks overlapped in the brain functions they targeted. The pitch matching task consisted of several pitches in ascending order. Therefore, subjects might have been able to adjust their subsequent productions by using the auditory feedback from their currently produced pitch. The perturbation task on the other hand might also have required sound-motor mapping to plan the accurate amount of compensation for the applied perturbation, using a mental representation of the originally produced pitch. Hence it is not surprising that non-invasive brain stimulation affected similar brain areas in both tasks: reducing neural excitability in the right pSTG and the left pIFG significantly interfered with performance in experiments 1 and 2. Recent functional neuroimaging studies on language processing showed inter-hemispheric connectivity between left and right STG in both directions but found only unidirectional transfer of information from right STG to left IFG. No functional connectivity was found from left STG to right IFG or from right STG to right IFG [119]. Likewise, we found performance in both tasks to be impaired following stimulation over right pSTG and left pIFG but not over left pSTG or right pIFG.

While current models for a shared pitch network across the two hemispheres have hypothesized that the right pSTG reflects perceptual processes for pitch information and the left pIFG represents performance-related processes [120], our study is the first one that allows establishing a causal contribution from both areas.

One explanation why different hemispheres were affected by the stimulation can be derived from established models of speech motor control. These models postulate different streams for feedback and feedforward control processes [10] and a similar

control network can be assumed for vocal pitch control as well [23]. More precisely, they state that during the production of sounds, a stream of feedforward commands is sent to frontal motor areas [34]. These feedforward processes have been shown to be mainly lateralized to the left hemisphere of the brain [61]. During vocalization, a constant perceptual feedback control stream towards the temporal lobe compares the actual auditory feedback with predictions about the intended sound and corrects them if needed [121]. This auditory feedback stream has been found to be located in the right hemisphere [61], an observation that is supported by similar findings from animal studies that showed sensitivity to self-generated vocalizations to be lateralized toward the right side of the brain [54]. We suggest that this dichotomy between feedforward and feedback streams in the fronto-temporal network might account for the fact that stimulation affected the pSTG of the right hemisphere but the pIFG of the left hemisphere.

However, at this point our experimental results do not allow us to differentiate between particular sub-functions of these two brain areas. Future research will have to be conducted to further explore the specific contributions of the pSTG and the pIFG to vocal pitch regulation.

4.5 Summary and outlook

Taken together, the results suggest that non-invasive brain stimulation is a viable method to modulate vocal output. We significantly interfered with performance in two pitch production and perception tasks after inducing virtual lesions with cathodal transcranial direct current stimulation. Therefore, the present experiments provide causal evidence for the role of the posterior superior temporal gyrus and the posterior inferior frontal gyrus in the neural network that controls vocal production and communication.

Our results have shed light on the neural correlates of voice-motor control and provide a target for future rehabilitative strategies in populations with communication disorders. In recent years, tDCS has increasingly been employed in clinical setting for rehabilitation, namely to facilitate recovery from post-stroke aphasia [122, 123] and dysphagia [124, 125]. In addition, delayed or frequency shifted feedback has successfully been used as

a therapeutic tool to improve speech output in conditions such as stuttering [126] or Parkinson's disease [127].

Particularly fMRI studies that investigated pitch-shifted vocal production in stuttering showed increased right hemispheric activation as compared to normal speakers, suggesting an increased reliance on auditory feedback control due to poor feedforward motor commands [61, 128]. This is consistent with the finding that successful fluency-inducing treatment for stuttering showed a shift toward more normal, left-lateralized brain activation [129]. Establishing a combination of perturbed auditory feedback and tDCS may serve as a new therapeutic approach by simultaneously reducing pathological right hemisphere overactivity and further training vocal motor skills.

In conclusion, we recommend future clinical research to employ non-invasive brain stimulation as a therapeutic tool in speech-motor disorders. We hereby aim to improve vocal production especially in populations with an underlying dysfunction in vocal feedback control.

5. References

1. Hickok, G. and D. Poeppel, *The cortical organization of speech processing.* Nat Rev Neurosci, 2007. **8**(5): p. 393-402.
2. Pulvermuller, F. and L. Fadiga, *Active perception: sensorimotor circuits as a cortical basis for language.* Nat Rev Neurosci, 2010. **11**(5): p. 351-60.
3. Rorden, C. and H.O. Karnath, *Using human brain lesions to infer function: a relic from a past era in the fMRI age?* Nat Rev Neurosci, 2004. **5**(10): p. 813-9.
4. Wernicke, C., *Der aphasische Symptomencomplex: Eine psychologische Studie auf anatomischer Basis* 1874, Breslau: M. Cohn & Weigert.
5. Broca, P., *Remarques sur le siège de la faculté du language articulé: suivies d'une observation d'aphémie (perte de la parole).* Bull. Soc. Anat. Paris 1861. **6**: p. 330-357.
6. Berker, E.A., A.H. Berker, and A. Smith, *Translation of Broca's 1865 report. Localization of speech in the third left frontal convolution.* Arch Neurol, 1986. **43**(10): p. 1065-72.
7. Catani, M. and M. Mesulam, *The arcuate fasciculus and the disconnection theme in language and aphasia: history and current state.* Cortex, 2008. **44**(8): p. 953-61.
8. Geschwind, N., *Current concepts: aphasia.* N Engl J Med, 1971. **284**(12): p. 654-6.
9. Kertesz, A., D. Lesk, and P. McCabe, *Isotope localization of infarcts in aphasia.* Arch Neurol, 1977. **34**(10): p. 590-601.
10. Rauschecker, J.P. and S.K. Scott, *Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing.* Nat Neurosci, 2009. **12**(6): p. 718-24.
11. Corballis, M.C., *From mouth to hand: gesture, speech, and the evolution of right-handedness.* Behav Brain Sci, 2003. **26**(2): p. 199-208; discussion 208-60.
12. Knecht, S., et al., *Handedness and hemispheric language dominance in healthy humans.* Brain, 2000. **123 Pt 12**: p. 2512-8.
13. Blumenfeld, H., *Neuroanatomy Through Clinical Cases* 2010, Sunderland, MA: Sinauer Associates.
14. Price, C.J., *A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading.* Neuroimage, 2012. **62**(2): p. 816-47.
15. Hebert, S., et al., *Revisiting the dissociation between singing and speaking in expressive aphasia.* Brain, 2003. **126**(Pt 8): p. 1838-50.
16. Wan, C.Y., et al., *The Therapeutic Effects of Singing in Neurological Disorders.* Music Percept, 2010. **27**(4): p. 287-295.
17. Zatorre, R.J., A.C. Evans, and E. Meyer, *Neural mechanisms underlying melodic perception and memory for pitch.* J Neurosci, 1994. **14**(4): p. 1908-19.

18. Albert, M.L., R.W. Sparks, and N.A. Helm, *Melodic intonation therapy for aphasia.* Arch Neurol, 1973. **29**(2): p. 130-1.
19. Schlaug, G., et al., *From singing to speaking: facilitating recovery from nonfluent aphasia.* Future Neurol, 2010. **5**(5): p. 657-665.
20. Saur, D., et al., *Dynamics of language reorganization after stroke.* Brain, 2006. **129**(Pt 6): p. 1371-84.
21. Kinsella, G., M.R. Prior, and G. Murray, *Singing ability after right and left sided brain damage. A research note.* Cortex, 1988. **24**(1): p. 165-9.
22. Brown, S., M.J. Martinez, and L.M. Parsons, *Music and language side by side in the brain: a PET study of the generation of melodies and sentences.* Eur J Neurosci, 2006. **23**(10): p. 2791-803.
23. Ozdemir, E., A. Norton, and G. Schlaug, *Shared and distinct neural correlates of singing and speaking.* Neuroimage, 2006. **33**(2): p. 628-35.
24. Ogawa, S. and T.M. Lee, *Magnetic resonance imaging of blood vessels at high fields: in vivo and in vitro measurements and image simulation.* Magn Reson Med, 1990. **16**(1): p. 9-18.
25. Logothetis, N.K., et al., *Neurophysiological investigation of the basis of the fMRI signal.* Nature, 2001. **412**(6843): p. 150-7.
26. Peck, K.K., et al., *Event-related functional MRI investigation of vocal pitch variation.* Neuroimage, 2009. **44**(1): p. 175-181.
27. Parkinson, A.L., et al., *Understanding the neural mechanisms involved in sensory control of voice production.* Neuroimage, 2012. **61**(1): p. 314-22.
28. Loui, P., et al., *Action-perception mismatch in tone-deafness.* Current Biology, 2008. **18**(8): p. R331-2.
29. Stewart, L., *Congenital amusia.* Curr Biol, 2006. **16**(21): p. R904-6.
30. Dalla Bella, S., J.F. Giguere, and I. Peretz, *Singing in congenital amusia.* J Acoust Soc Am, 2009. **126**(1): p. 414-24.
31. Dalla Bella, S. and M. Berkowska, *Singing proficiency in the majority: normality and "phenotypes" of poor singing.* Ann N Y Acad Sci, 2009. **1169**: p. 99-107.
32. Hyde, K.L., et al., *Cortical thickness in congenital amusia: when less is better than more.* J Neurosci, 2007. **27**(47): p. 13028-32.
33. Mandell, J., K. Schulze, and G. Schlaug, *Congenital amusia: An auditory-motor feedback disorder?* Restorative Neurology and Neuroscience, 2007. **25**: p. 323-334.
34. Guenther, F.H., *Cortical interactions underlying the production of speech sounds.* J Commun Disord, 2006. **39**(5): p. 350-65.
35. Loui, P., et al., *Enhanced cortical connectivity in absolute pitch musicians: a model for local hyperconnectivity.* J Cogn Neurosci, 2011. **23**(4): p. 1015-26.
36. Miyazaki, K., *Musical pitch identification by absolute pitch possessors.* Percept Psychophys, 1988. **44**(6): p. 501-12.
37. Mori, S. and J. Zhang, *Principles of diffusion tensor imaging and its applications to basic neuroscience research.* Neuron, 2006. **51**(5): p. 527-39.

38. Loui, P., D. Alsop, and G. Schlaug, *Tone-Deafness: a Disconnection Syndrome?* Journal of Neuroscience, 2009. **29**(33): p. 10215-10220.

39. de Swart, B.J., et al., *Improvement of voicing in patients with Parkinson's disease by speech therapy.* Neurology, 2003. **60**(3): p. 498-500.

40. Stager, S.V., K.J. Jeffries, and A.R. Braun, *Common features of fluency-evoking conditions studied in stuttering subjects and controls: an H(2)15O PET study.* J Fluency Disord, 2003. **28**(4): p. 319-35; quiz 336.

41. Glover, H., et al., *Effect of instruction to sing on stuttering frequency at normal and fast rates.* Percept Mot Skills, 1996. **83**(2): p. 511-22.

42. Wan, C.Y., et al., *Auditory-motor mapping training as an intervention to facilitate speech output in non-verbal children with autism: a proof of concept study.* PLoS One, 2011. **6**(9): p. e25505.

43. Loui, P., et al., *Relating pitch awareness to phonemic awareness in children: implications for tone-deafness and dyslexia.* Front Psychol, 2011. **2**: p. 111.

44. Gaab, N., et al., *Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: an fMRI study.* Restor Neurol Neurosci, 2007. **25**(3-4): p. 295-310.

45. Schaadt, G., A. Pannekamp, and E. van der Meer, *Auditory Phoneme Discrimination in Illiterates: Mismatch Negativity-A Question of Literacy?* Dev Psychol, 2013.

46. Peretz, I. and M. Coltheart, *Modularity of music processing.* Nat Neurosci, 2003. **6**(7): p. 688-91.

47. Patel, A.D., *Language, music, syntax and the brain.* Nat Neurosci, 2003. **6**(7): p. 674-81.

48. Heinks-Maldonado, T.H., et al., *Fine-tuning of auditory cortex during speech production.* Psychophysiology, 2005. **42**(2): p. 180-90.

49. Aliu, S.O., J.F. Houde, and S.S. Nagarajan, *Motor-induced suppression of the auditory cortex.* J Cogn Neurosci, 2009. **21**(4): p. 791-802.

50. Houde, J.F., et al., *Modulation of the auditory cortex during speech: an MEG study.* J Cogn Neurosci, 2002. **14**(8): p. 1125-38.

51. Behroozmand, R., et al., *Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation.* Clin Neurophysiol, 2009. **120**(7): p. 1303-12.

52. Chang, E.F., et al., *Human cortical sensorimotor network underlying feedback control of vocal pitch.* Proc Natl Acad Sci U S A, 2013. **110**(7): p. 2653-8.

53. Eliades, S.J. and X. Wang, *Neural substrates of vocalization feedback monitoring in primate auditory cortex.* Nature, 2008. **453**(7198): p. 1102-6.

54. Poirier, C., et al., *Own-song recognition in the songbird auditory pathway: selectivity and lateralization.* J Neurosci, 2009. **29**(7): p. 2252-8.

55. Remedios, R., N.K. Logothetis, and C. Kayser, *An Auditory Region in the Primate Insular Cortex Responding Preferentially to Vocal Communication Sounds.* J. Neurosci., 2009. **29**(4): p. 1034-1045.

56. Burnett, T.A., et al., *Voice F0 responses to manipulations in pitch feedback.* J Acoust Soc Am, 1998. **103**(6): p. 3153-61.

57. Jones, J.A. and K.G. Munhall, *Perceptual calibration of F0 production: evidence from feedback perturbation.* J Acoust Soc Am, 2000. **108**(3 Pt 1): p. 1246-51.

58. Munhall, K.G., et al., *Talkers alter vowel production in response to real-time formant perturbation even when instructed not to compensate.* J Acoust Soc Am, 2009. **125**(1): p. 384-90.

59. Bauer, J.J., et al., *Vocal responses to unanticipated perturbations in voice loudness feedback: an automatic mechanism for stabilizing voice amplitude.* J Acoust Soc Am, 2006. **119**(4): p. 2363-71.

60. Houde, J.F. and M.I. Jordan, *Sensorimotor adaptation in speech production.* Science, 1998. **279**(5354): p. 1213-6.

61. Tourville, J.A., K.J. Reilly, and F.H. Guenther, *Neural mechanisms underlying auditory feedback control of speech.* Neuroimage, 2008. **39**(3): p. 1429-43.

62. Zarate, J.M. and R.J. Zatorre, *Experience-dependent neural substrates involved in vocal pitch regulation during singing.* Neuroimage, 2008. **40**(4): p. 1871-87.

63. Hashimoto, Y. and K.L. Sakai, *Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: an fMRI study.* Hum Brain Mapp, 2003. **20**(1): p. 22-8.

64. Fu, C.H., et al., *An fMRI study of verbal self-monitoring: neural correlates of auditory verbal feedback.* Cereb Cortex, 2006. **16**(7): p. 969-77.

65. Toyomura, A., et al., *Neural correlates of auditory feedback control in human.* Neuroscience, 2007. **146**(2): p. 499-503.

66. Vul, E., et al., *Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition.* Perspectives on Psychological Science, 2009. **4**(3): p. 274-290.

67. Logothetis, N.K., *What we can do and what we cannot do with fMRI.* Nature, 2008. **453**(7197): p. 869-78.

68. Loui, P., A. Hohmann, and G. Schlaug, *Inducing Disorders in Pitch Perception and Production: a Reverse-Engineering Approach.* Proc Meet Acoust, 2010. **9**(1): p. 50002.

69. Wagner, T., A. Valero-Cabre, and A. Pascual-Leone, *Noninvasive human brain stimulation.* Annu Rev Biomed Eng, 2007. **9**: p. 527-65.

70. Ardolino, G., et al., *Non-synaptic mechanisms underlie the after-effects of cathodal transcutaneous direct current stimulation of the human brain.* J Physiol, 2005. **568**(Pt 2): p. 653-63.

71. Feng, W.W., M.G. Bowden, and S. Kautz, *Review of transcranial direct current stimulation in poststroke recovery.* Top Stroke Rehabil, 2013. **20**(1): p. 68-77.

72. Hummel, F.C. and L.G. Cohen, *Non-invasive brain stimulation: a new strategy to improve neurorehabilitation after stroke?* Lancet Neurol, 2006. **5**(8): p. 708-12.

73. Mathys, C., et al., *Non-invasive brain-stimulation applied to Heschl's gyrus modulates pitch discrimination.* Frontiers in Psychology, 2010. **1**: p. 12.

74. Antal, A., M.A. Nitsche, and W. Paulus, *Transcranial direct current stimulation and the visual cortex.* Brain Res Bull, 2006. **68**(6): p. 459-63.

75. Hummel, F., et al., *Effects of non-invasive cortical stimulation on skilled motor function in chronic stroke.* Brain, 2005. **128**(Pt 3): p. 490-9.

76. Nitsche, M.A. and W. Paulus, *Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans.* Neurology, 2001. **57**(10): p. 1899-901.

77. Schlaug, G., V. Renga, and D. Nair, *Transcranial direct current stimulation in stroke recovery.* Arch Neurol, 2008. **65**(12): p. 1571-6.

78. McCreery, D.B., et al., *Charge density and charge per phase as cofactors in neural injury induced by electrical stimulation.* IEEE Trans Biomed Eng, 1990. **37**(10): p. 996-1001.

79. Liebetanz, D., et al., *Pharmacological approach to the mechanisms of transcranial DC-stimulation-induced after-effects of human motor cortex excitability.* Brain, 2002. **125**(Pt 10): p. 2238-47.

80. Zheng, X., D.C. Alsop, and G. Schlaug, *Effects of transcranial direct current stimulation (tDCS) on human regional cerebral blood flow.* Neuroimage, 2011. **58**(1): p. 26-33.

81. Cerruti, C. and G. Schlaug, *Anodal transcranial direct current stimulation of the prefrontal cortex enhances complex verbal associative thought.* J Cogn Neurosci, 2009. **21**(10): p. 1980-7.

82. Antal, A., et al., *Direct current stimulation over V5 enhances visuomotor coordination by improving motion perception in humans.* J Cogn Neurosci, 2004. **16**(4): p. 521-7.

83. Nitsche, M.A., et al., *Shaping the effects of transcranial direct current stimulation of the human motor cortex.* J Neurophysiol, 2007. **97**(4): p. 3109-3117.

84. Vines, B.W., D.G. Nair, and G. Schlaug, *Contralateral and ipsilateral motor effects after transcranial direct current stimulation.* Neuroreport, 2006. **17**(6): p. 671-4.

85. Vines, B.W., N.M. Schnider, and G. Schlaug, *Testing for causality with transcranial direct current stimulation: pitch memory and the left supramarginal gyrus.* Neuroreport, 2006. **17**(10): p. 1047-50.

86. Gandiga, P.C., F.C. Hummel, and L.G. Cohen, *Transcranial DC stimulation (tDCS): a tool for double-blind sham-controlled clinical studies in brain stimulation.* Clin Neurophysiol, 2006. **117**(4): p. 845-50.

87. Hohmann A., L.P., Schlaug G., *Inducing Tone-deafness via Non-invasive Focal Brain Stimulation*, in *Organization for Human Brain Mapping Annual Meeting* 2009: San Francisco, CA. p. S107.

88. Cornsweet, T.N., *The Staircase-Method in Psychophysics.* The American Journal of Psychology, 1962. **75**(3): p. 485-491.

89. Pascual-Marqui, R.D., C.M. Michel, and D. Lehmann, *Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain.* Int J Psychophysiol, 1994. **18**(1): p. 49-65.

90. Nitsche, M.A., et al., *Facilitation of implicit motor learning by weak transcranial direct current stimulation of the primary motor cortex in the human.* J Cogn Neurosci, 2003. **15**(4): p. 619-26.

91. Cattaneo, Z., A. Pisoni, and C. Papagno, *Transcranial direct current stimulation over Broca's region improves phonemic and semantic fluency in healthy individuals.* Neuroscience, 2011. **183**: p. 64-70.

92. Boersma, P. and D. Weenink, *Praat: doing phonetics by computer*, 2005: http://www.praat.org.

93. *IBM SPSS Statistics for Windows* 2010, IBM Corp.: Armonk, NY.

94. Chen, Z., et al., *Sex-related differences in vocal responses to pitch feedback perturbations during sustained vocalization.* J Acoust Soc Am, 2010. **128**(6): p. EL355-60.

95. Peretz, I., A.S. Champod, and K. Hyde, *Varieties of musical disorders. The Montreal Battery of Evaluation of Amusia.* Ann N Y Acad Sci, 2003. **999**: p. 58-75.

96. Shipley, W.C., *A self-administering scale for measuring intellectual impairment and deterioration.* Journal of Psychology, 1940. **9**: p. 371-377.

97. Zicarelli, D., *An Extensible Real-Time Signal Processing Environment for Max.* Proceedings of the International Computer Music Conference, 1998: p. 463-466.

98. *MATLAB*, 2012b, The MathWorks Inc.,: Natick, Massachusetts, United States.

99. Liu, H., et al., *Differential effects of perturbation direction and magnitude on the neural processing of voice pitch feedback.* Clin Neurophysiol, 2011. **122**(5): p. 951-7.

100. Liu, P., et al., *Auditory feedback control of vocal pitch during sustained vocalization: a cross-sectional study of adult aging.* PLoS One, 2011. **6**(7): p. e22791.

101. Sivasankar, M., et al., *Voice responses to changes in pitch of voice or tone auditory feedback.* J Acoust Soc Am, 2005. **117**(2): p. 850-7.

102. Brown, S., et al., *The song system of the human brain.* Brain Res Cogn Brain Res, 2004. **20**(3): p. 363-75.

103. Perry, D.W., et al., *Localization of cerebral activity during simple singing.* Neuroreport, 1999. **10**(18): p. 3979-84.

104. Hickok, G., et al., *Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt.* J Cogn Neurosci, 2003. **15**(5): p. 673-82.

105. Sun, T. and C.A. Walsh, *Molecular approaches to brain asymmetry and handedness.* Nat Rev Neurosci, 2006. **7**(8): p. 655-62.

106. Lahav, A., E. Saltzman, and G. Schlaug, *Action representation of sound: audiomotor recognition network while listening to newly acquired actions.* J Neurosci, 2007. **27**(2): p. 308-14.

107. Binkofski, F. and G. Buccino, *The role of ventral premotor cortex in action execution and action understanding.* J Physiol Paris, 2006. **99**(4-6): p. 396-405.

108. Binkofski, F. and G. Buccino, *Motor functions of the Broca's region.* Brain Lang, 2004. **89**(2): p. 362-9.

109. Cattaneo, L. and G. Rizzolatti, *The mirror neuron system.* Arch Neurol, 2009. **66**(5): p. 557-60.

110. Zatorre, R.J., J.L. Chen, and V.B. Penhune, *When the brain plays music: auditory-motor interactions in music perception and production.* Nat Rev Neurosci, 2007. **8**(7): p. 547-58.

111. Gaser, C. and G. Schlaug, *Brain structures differ between musicians and non-musicians.* J Neurosci, 2003. **23**(27): p. 9240-5.

112. Schneider, P., et al., *Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians.* Nat Neurosci, 2002. **5**(7): p. 688-94.

113. Chen, J.L., V.B. Penhune, and R.J. Zatorre, *Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training.* J Cogn Neurosci, 2008. **20**(2): p. 226-39.

114. Rosenkranz, K., A. Williamon, and J.C. Rothwell, *Motorcortical excitability and synaptic plasticity is enhanced in professional musicians.* J Neurosci, 2007. **27**(19): p. 5200-6.

115. Schlaug, G., *The brain of musicians. A model for functional and structural adaptation.* Ann N Y Acad Sci, 2001. **930**: p. 281-99.

116. Liu, H., N.M. Russo, and C.R. Larson, *Age-related differences in vocal responses to pitch feedback perturbations: a preliminary study.* J Acoust Soc Am, 2010. **127**(2): p. 1042-6.

117. Liu, H., et al., *Effect of tonal native language on voice fundamental frequency responses to pitch feedback perturbations during sustained vocalizations.* J Acoust Soc Am, 2010. **128**(6): p. 3739-46.

118. Liu, P., et al., *Developmental sex-specific change in auditory-vocal integration: ERP evidence in children.* Clin Neurophysiol, 2013. **124**(3): p. 503-13.

119. Bitan, T., et al., *Bidirectional connectivity between hemispheres occurs at multiple levels in language processing but depends on sex.* J Neurosci, 2010. **30**(35): p. 11576-85.

120. Nan, Y. and A.D. Friederici, *Differential roles of right temporal cortex and Broca's area in pitch processing: Evidence from music and Mandarin.* Hum Brain Mapp, 2012.

121. Golfinopoulos, E., J.A. Tourville, and F.H. Guenther, *The integration of large-scale neural network modeling and functional brain imaging in speech motor control.* Neuroimage, 2010. **52**(3): p. 862-74.

122. Vines, B.W., A.C. Norton, and G. Schlaug, *Non-invasive brain stimulation enhances the effects of melodic intonation therapy.* Front Psychol, 2011. **2**: p. 230.

123. Schlaug, G., S. Marchina, and C.Y. Wan, *The use of non-invasive brain stimulation techniques to facilitate recovery from post-stroke aphasia.* Neuropsychol Rev, 2011. **21**(3): p. 288-301.

124. Shigematsu, T., I. Fujishima, and K. Ohno, *Transcranial direct current stimulation improves swallowing function in stroke patients.* Neurorehabil Neural Repair, 2013. **27**(4): p. 363-9.

125. Kumar, S., et al., *Noninvasive brain stimulation may improve stroke-related dysphagia: a pilot study.* Stroke, 2011. **42**(4): p. 1035-40.

126. Reed, P. and P. Howell, *Suggestions for improving the long-term effects of treatments for stuttering: A Review and synthesis of frequency-shifted feedback and operant techniques.* Eur J Behav Anal, 2000. **1**(2): p. 89-106.

127. Kiran, S. and C.R. Larson, *Effect of duration of pitch-shifted feedback on vocal responses in patients with Parkinson's disease.* J Speech Lang Hear Res, 2001. **44**(5): p. 975-87.

128. Brown, S., et al., *Stuttered and fluent speech production: an ALE meta-analysis of functional neuroimaging studies.* Hum Brain Mapp, 2005. **25**(1): p. 105-17.

129. De Nil, L.F., et al., *A positron emission tomography study of short- and long-term treatment effects on functional brain activation in adults who stutter.* J Fluency Disord, 2003. **28**(4): p. 357-79; quiz 379-80.

List of Tables and Figures

Figures

Tables

Eidesstattliche Versicherung

Ich, Anja Hohmann, versichere an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorgelegte Dissertation mit dem Thema *Modulating Vocal Pitch Perception and Production with Transcranial Direct Current Stimulation* selbstständig und ohne nicht offengelegte Hilfe Dritter verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel genutzt habe.
Alle Stellen, die wörtlich oder dem Sinne nach auf Publikationen oder Vorträgen anderer Autoren beruhen, sind als solche in korrekter Zitierung (siehe „Uniform Requirements for Manuscripts (URM)“ des ICMJE -*www.**icmje**.org*) kenntlich gemacht. Die Abschnitte zu Methodik (insbesondere praktische Arbeiten, Laborbestimmungen, statistische Aufarbeitung) und Resultaten (insbesondere Abbildungen, Graphiken und Tabellen) entsprechen den URM (s.o) und werden von mir verantwortet.
Meine Anteile an etwaigen Publikationen zu dieser Dissertation entsprechen denen, die in der untenstehenden gemeinsamen Erklärung mit dem/der Betreuer/in, angegeben sind. Sämtliche Publikationen, die aus dieser Dissertation hervorgegangen sind und bei denen ich Autor bin, entsprechen den URM (s.o) und werden von mir verantwortet.

Die Bedeutung dieser eidesstattlichen Versicherung und die strafrechtlichen Folgen einer unwahren eidesstattlichen Versicherung (§156,161 des Strafgesetzbuches) sind mir bekannt und bewusst.

Datum Unterschrift

Anteilserklärung an etwaigen erfolgten Publikationen

Anja Hohmann hatte folgenden Anteil an den folgenden Publikationen:

Publikation 1: Loui P, Li HC, Hohmann A, Schlaug G: *Enhanced Cortical Connectivity in Absolute Pitch Musicians: A Model for Local Hyperconnectivity*. Journal of Cognitive Neuroscience (2011)
Beitrag im Einzelnen (bitte kurz ausführen):
Datenerhebung (DTI) und Datenauswertung, Konzeption und Verfassen der Publikation.

Publikation 2: Loui P, Hohmann A, Schlaug G: *Inducing Disorders in Pitch Perception and Production: a Reverse-Engineering Approach.* Proceedings of Meetings on Acoustics (2010)
Beitrag im Einzelnen (bitte kurz ausführen):
Planung und Durchführung der Experimente; Datenauswertung und Statistik, Konzeption und Verfassen der Publikation.

Publikation 3: Wan CY, Rueber T, Hohmann A, Schlaug G. *The Therapeutic Effects of Singing in Neurological Disorders.* Music Perception (2010)
Beitrag im Einzelnen (bitte kurz ausführen):
Literaturrecherche, Konzeption und Verfassen der Publikation, vor allem Abschnitt über Stottern.

Unterschrift der Doktorandin

Lebenslauf

Mein Lebenslauf wird aus datenschutzrechtlichen Gründen in der elektronischen Version meiner Arbeit nicht veröffentlicht.

Publikationen

Publikationen

Loui P, Li HC, **Hohmann A**, Schlaug G. *Enhanced Cortical Connectivity in Absolute Pitch Musicians: A Model for Local Hyperconnectivity.* J Cog Neurosci 2011;23(4):1015-1026. PMID: 20515408.

Loui P, **Hohmann A**, Schlaug G. *Inducing Disorders in Pitch Perception and Production: a Reverse-Engineering Approach*. POMA 2010;9:1-8. PMID: 20725606.

Wan CY, Rueber T, **Hohmann A**, Schlaug G. *The Therapeutic Effects of Singing in Neurological Disorders.* Music Perception 2010;27(4):287-295. PMID: 21152359.

Posterpräsentationen

Hohmann A, Loui P, Schlaug G*. Inducing Tone-deafness via Non-invasive Focal Brain Stimulation*. Organization for Human Brain Mapping Annual Meeting, San Francisco, Juni 2009. Abstract in: NeuroImage 2009;47:S107.

Loui P, Li HC, **Hohmann A,** Schlaug G. *Absolute Pitch – a Group Study in Cortical Hyperconnectivity.* Organization for Human Brain Mapping Annual Meeting, San Francisco, Juni 2009. Abstract in: NeuroImage 2009;47:S106.

Hohmann A, Loui P, Schlaug G. *Modeling Vocal Pitch Production with Perturbed Auditory Feedback.* Cognitive Neuroscience Society Annual Meeting, San Francisco, März 2009. Abstract in: J Cog Neurosci Suppl. 2009; (ISSN 1096-885):155.

Loui P, Li HC, **Hohmann A**, Schlaug G. *Cortical Hyperconnectivity in Absolute Pitch Musicians.* Cognitive Neuroscience Society Annual Meeting, San Francisco, März 2009. Abstract in: J Cog Neurosci Suppl. 2009; (ISSN 1096-885):34.

Hohmann A, Loui P, Schlaug G. *Modeling Neurological Disorders of Musical Ability with Vocal Feedback*. Neurosciences and the Arts Symposium, Harvard Medical School Boston, Januar 2009.

Danksagung

Besonderer Dank gilt meinem Doktorvater Herrn Professor Felix Bermpohl für die exzellente Betreuung beim Verfassen dieser Arbeit. Seine Bereitschaft, jederzeit Fragen und Probleme gemeinsam zu erörtern, sowie die konstruktive Kritik haben ganz entscheidend zum Gelingen dieser Arbeit beigetragen.

Desweiteren möchte ich mich bei Herrn Professor Gottfried Schlaug vom Department of Neurology des Beth Israel Deaconess Medical Centers für die Überlassung des Themas und das damit entgegengebrachte Vertrauen sowie die Betreuung durch ihn und seine Arbeitsgruppe während des Jahres an der Harvard Medical School bedanken. Vor allem danke ich hierbei meiner direkten Betreuerin in Boston, Frau Dr. Psyche Loui, für ihr großes Engagement und ihren unerschütterlichen Optimismus.
Zudem möchte ich Steffi Spengler, Moritz Bächer und Léa Steinacker für ihre kompetente Unterstützung bei der Analyse der Arbeit danken.

Von ganzem Herzen danke ich meinen Eltern, die mich all die Jahre bedingungslos unterstützt und mir damit dieses Studium ermöglicht haben, sowie meinem Bruder Stephan.