The role of somatosensation in vocal motor control for singing

Dissertation

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

der Mathematisch-Naturwissenschaftlichen Fakultät

und

der Medizinischen Fakultät

der Eberhard-Karls-Universität Tübingen

vorgelegt

von

Sebastian Finkel

aus Reutlingen, Deutschland

Februar - 2016

Tag der mündlichen Prüfung:	19. Dezember 2016
Dekan der Math.-Nat. Fakultät:	Prof. Dr. W. Rosenstiel
Dekan der Medizinischen Fakultät:	Prof. Dr. I. B. Autenrieth
1. Berichterstatter:	Prof. Dr. Niels Birbaumer
2. Berichterstatter:	Prof. Dr. Christoph Braun
Prüngskommission:	Prof. Dr. Martin Hautzinger Prof. Dr. Christoph Braun PD Dr. Steffen Hage Dr. Andreas Bartels

Eidesstattliche Erklärung

Ich erkläre, dass ich die zur Promotion eingereichte Arbeit mit dem Titel: *"The role of somatosensation in vocal motor control for singing"* selbstständig verfasst, nur die angegeben Quellen und Hilfsmittel benutzt und wörtlich oder inhaltlich übernommene Stellen als solche gekennzeichnet habe. Ich versichere an Eides statt, dass diese Angaben wahr sind und dass ich nichts verschwiegen habe. Mir ist bekannt, dass die falsche Angabe einer Versicherung an Eides statt mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft wird.

Reutlingen, den 5. Februar 2017

Contents

Abbreviations

ACC	anterior cingulate cortex
ADM	abductor digiti minimi
CT	cricothyroid muscle
cTBS	continous theta burst stimulation
F_0	fundamental frequency
FDI	first dorsal interessous
fMRI	functional magnetic resonance imaging
Hz	hertz
IA	interarytenoid muscle
iTBS	intermittend theta burst stimulation
imTBS	intermediate theta burst stimulation
kHz	kilohertz
LCA	lateral cricoarytenoid muscle
LMC	laryngeal motor cortex
LTP	long-term potentiation
LTD	long-term depression
MEP	motor evoked potential
ms	milliseconds
MSO	maximum stimulator output

MRI	magnetic resonance imaging
M1	primary motor cortex
NMDA	N-methyl-D-aspartat
PAG	periaqueductal gray
PCA	posterior cricoarytenoid muscle
RLN	recurrent laryngeal nerve
s	seconds
S1	primary somatosensory cortex
SAI	short-latency afferent inhibition
SEP	somatosensory evoked potential
SLN	superior laryngeal nerve
T1	longitudinal relaxation time
T2	transverse relaxation time
TA	thyroarytenoid muscle
TR	repetition time
TE	echo time
TBS	theta burst stimulation
TMS	transcranial magnetic stimulation

Abstract – Zusammenfassung

Extensive research on the human voice with its sensory and motor systems has converged on the idea that the auditory system is critical for vocal production, yet recent reports suggest that the somatosensory system contributes more substantially to vocal motor-control than currently recognized. This thesis assessed the modulator influence of primary somatosensory cortex (S1) on vocal pitch-matching with transcranial magnetic stimulation, applied to right larynx-S1 and a dorsal-S1 control area in untrained singers. In experiment I, participants sang before and after TMS with normal auditory feedback whereas in experiment II, auditory feedback was masked with noise. TMS showed no effects on singing in experiment I. However, when auditory feedback was masked, larynx-S1 stimulation significantly improved both initial pitch accuracy and final pitch stability in contrast to dorsal-S1 stimulation. Positive effects of larynx S1 stimulation on initial and final pitch accuracy were more pronounced in participants who sang less accurately prior to iTBS. Moreover, masking showed more adverse effects on pitch-control in participants with higher pitch-discrimination thresholds. Conversely, these participants also profited more from larynx-S1 stimulation in initial and final-pitch accuracy. These data provide first evidence for a critical involvement of larynx-S1 in pitch motor-control independent from prior singing experience.

Umfangreiche Untersuchungen am motorischen und sensorischen Kontrollsystem der menschlichen Stimme basieren zumeist auf der Idee, dass das Gehör entscheidend für die Stimmkontrolle ist. Doch neuere Forschungsergenisse deuten darauf hin, dass das somatosensible System wesentlicher zur exakten Steuerung der Stimme beiträgt als bisher anerkannt wird. Diese Arbeit untersucht den modulatorischen Einfluss des primären somatosensorischen Kortex (S1) auf die Tonhöhenanpassungsfähigkeit beim Singen mit Hilfe von transkranieller Magnetstimulation (TMS). Diese wird angewandt auf die rechtsseitige somatosensorische Repräsentation der Kehlkopfmuskulatur und einer dorsalen S1 Kontrollregion. Getestet wurden ungeübte Sänger. In Experiment I sangen die Teilnehmer vor und nach der TMS mit normalen akustischem Feedback während im Experiment II das auditive Feedback mit 'pink noise' maskiert wurde. Die TMS zeigte keine Auswirkungen auf die Gesangsgenauigkeit in Experiment I. Wenn jedoch das akustische Feedback maskiert wurde führte die Stimulation der Kehlkopfrepräsentation zu einer deutlichen Verbesserung, sowohl der initialen Tonhöhengenauigkeit, als auch finalen Tonhöhenstabilität des gesungen Tones, im Gegensatz zur Stimulation des dorsalen Kontrollareals. Positive Auswirkungen der Stimulation des Kehlkopfareals auf die initiale und finale Tonhöhengenauigkeit waren stärker ausgeprägt je ungenauer die Teilnehmer vor der TMS gesungen haben. Darüber hinaus zeigte die Maskierung des Hören stärkere negative Auswirkungen auf Tonhöhenkontrolle bei Teilnehmern mit schlechterer Hörgenauigkeit. Umgekehrt profitierten diese Teilnehmer auch mehr von der TMS der Kehlkopfregion im Bereich der initialen wie auch finalen Tonhöhengenauigkeit. Diese Daten liefern erste Hinweise auf eine notwendige Beteiligung der S1 Kehlkopfregion in der Stimmkontrolle unabhängig von vorhergehender Gesangserfahrung.

Overview

Chapter 1 – Introduction

This chapter introduces the reader to the main idea of this thesis and gives a short overview over the current literature. It highlights the knowledge gap in the current research and exhibits how these two experiments can help to close this gap.

Chapter 2 – The Human Voice

Here, I depict the key aspects of vocal motor control and its control systems from different point of views. (e.g. anatomy, physiology, neuroanatomy)

Chapter 3 – Transcranial Magnetic Stimulation

Transcranial Magnetic Stimulation (TMS) is a widely used, non-invasive brain stimulation technique. Besides the technical and physical background, the experimental setup and procedure are explained. Moreover, stimulation protocols and their effects are discussed.

Chapter 4 – Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is a modern and powerful non-invasive imaging technique making it possible to record in vivo high resolution brain images. This chapter describes the basic physical principles of MRI and the applied technique e.g. the event related sparse sampling design.

Chapter 5 – Behavioural Testing and Statistics

In both studies I used behavioural measures to investigate pitch discrimination and pitch reproduction performance. Moreover, several questionnaires are utilised to effectively acquire subjects personal and musical background.

Chapter 6 – Experiment I

In the first experiment I used TMS over the right laryngeal S1 and a dorsal S1 control area to investigate the role of somatosensation on vocal motor control while auditory feedback remained unchanged. Methods and results are depicted and preliminary discussed.

Chapter 7 – Experiment II

This chapter describes the second experiment where I used TMS over the right laryngeal S1 and a dorsal S1 control area to investigate the role of somatosensation on vocal motor control while auditory feedback was removed during singing with noise. Methods and results are explained in detail and a first discussion of the results is started

Chapter 8 – Discussion

Finally, results of experiment I and II are reviewed and put into relation with respect to the recent literature. Furthermore, open questions, limitations of both experiments, and possible future research are discussed.

Part I

Theory and Background

CHAPTER 1

Introduction

Human vocal production is a highly complex motor task that develops over an extended time period and reaches high levels of motor consistency for speech sounds only around the age of 14 (Smith and Zelaznik 2004; Smith 2006). The neural control over this dynamic system is demanding, as it requires the precise coordination of respiratory, laryngeal, and articulatory muscle groups and the simultaneous integration of multimodal sensory information (Titze 1993; Ludlow 2005; Eickhoff et al. 2009; Kleber et al. 2013). Within the sensory modalities, auditory feedback from the own voice is thought to play a major role in the development of voluntary control over vocal utterances, as it provides precise information about the acoustic consequences of goal-directed movements within vocal-tract structures with respect to intended auditory trajectories (Perkell et al. 2007; Zarate 2013). Several behavioral and neuroimaging studies have investigated auditory-motor interactions, mostly by demonstrating corrective vocal motor responses to auditory perturbations such as real-time shifting of vocal fundamental frequency (i.e., pitch) or masking hearing with noise (for example, see Burnett et al. 1998; Larson 1998; Hain et al. 2000; Zarate and Zatorre 2008). While these data support the widely held view that acoustically monitoring ones own voice is the main tool for guiding human vocalizations, a growing body of evidence also suggests that somatosensory feedback from the vocal tract plays a greater role in fine vocal-motor control than previously thought (Tremblay et al. 2003; Mürbe et al. 2004; Nasir and Ostry 2006; Kleber et al. 2010; Lametti et al. 2012; Kleber et al. 2013).

Somatosensory feedback that accompanies vocalizations reflects feelings of touch, stretch, vibration, and position. They originate from sensory receptors located in the chest wall, vocal tract muscles, joints, and mucosa and contribute to the vocal motor management by providing information about

the current state of the vocal system (Jürgens 2002; Smith and Zelaznik 2004; Hickok 2012). Signals generated by these receptors ascent via brainstem nuclei and the thalamus to insular and primary somatosensory cortices (S1). From there they project to primary motor (M1) and parietal sensory integration areas respectively (Jürgens 2002; Ackermann and Riecker 2004; Eickhoff et al. 2009). At the highest level of motor control, these signals are somatotopically organized for lip, jaw, and tongue movements and the larynx in M1 (Brown et al. 2008, 2009; Grabski et al. 2012), which reflects also the exceptional ability of humans to learn new vocal motor patterns due to direct connections between primary motor cortex and vocal motor neurons in the brainstem (Simonyan and Horwitz 2011).

The precision requirements of movements for articulation and sound production depend critically on the functional integration of somatosensory input. For example, mucosal mechanoreceptors in the larynx contribute substantially to sustained vocal fold oscillations by coordinating sub-glottal air pressure and intrinsic laryngeal muscular tension via a reflexogenic control system (Wyke 1974a,b; Titze and Hunter 2004; Gozaine and Clark 2005). The repeated coupling of successful motor commands and corresponding kinesthetic sensations eventually generates stable somatosensory-motor mappings that provide reliable information for meeting the precision requirement of vocal production even in the absence of audition (Nasir and Ostry 2008). With respect to vocal motor learning, singing represents a particularly useful model for exploring experience-dependent differences in the neural control of vocalization. Compared to speech, music provides a clearly predefined set of tonal and rhythmical relationships that necessitate the development of a more fine-grained motor concept and corresponding vocal motor control (Natke et al. 2003; Zatorre and Baum 2012). In line with this, trained singers

were found to show superior pitch matching skills compared to non-singers (Hutchins and Peretz 2012; Kleber et al. 2013), a difference that may be related to training induced facilitation of somatosensory-motor interactions. Recent neuroimaging studies have repeatedly found increased activation in laryngeal somatosensory and insular cortex as a function of experience in singing (Zarate and Zatorre 2008; Kleber et al. 2010, 2013). These results correspond with behavioral data suggesting that somatosensory feedback continuously contributes to vocal pitch-matching accuracy, even after an extended period of vocal training, whereas auditory feedback did not (Mürbe et al. 2002, 2004). Conversely, pitch discrimination training in non-singers did not lead to improvement in pitch accuracy (Zarate et al. 2010), despite improved perceptual skills (Zatorre et al. 2012).

Taken together, these data point towards a stronger role of somatosensory feedback in vocal motor control, yet they do not allow assumptions regarding causal relationships. Therefore, I applied repetitive transcranial magnetic stimulation (rTMS) to right ventral (i.e. laryngeal) primary somatosensory cortex in order to assess causal effects on singing accuracy in a pitch-reproduction paradigm with non-singers. I used intermittent and continuous theta burst stimulation (iTBS and cTBS). iTBS is a TMS protocol that has shown to facilitate neural processing in both motor (Huang et al. 2005) and somatosensory cortices (Ragert et al. 2008; Morley et al. 2007; Katayama and Rothwell 2007), while cTBS has shown the opposite effect (Conte et al. 2012; Rai et al. 2012). Based on the role of somatosensation in experienced singers, I expected that during normal (i.e. unperturbed) auditory perception iTBS to the right S1 larynx area would enhance pitch accuracy and stability in non-singers, while cTBS would impair performance. As non-singers may not yet use somatosensation based pitch control strategies, I additionally

masked participants hearing with pink noise during singing in the subsequent second iTBS experiment. Therefore, I proposed that positive effects of iTBS on pitch accuracy and stability may be enhanced in the masking compared to the normal singing condition. However, under both conditions, I expect subjects to have improved pitch reproduction performance after stimulation of the right larynx S1 area, but not after stimulation over a dorsal control area.

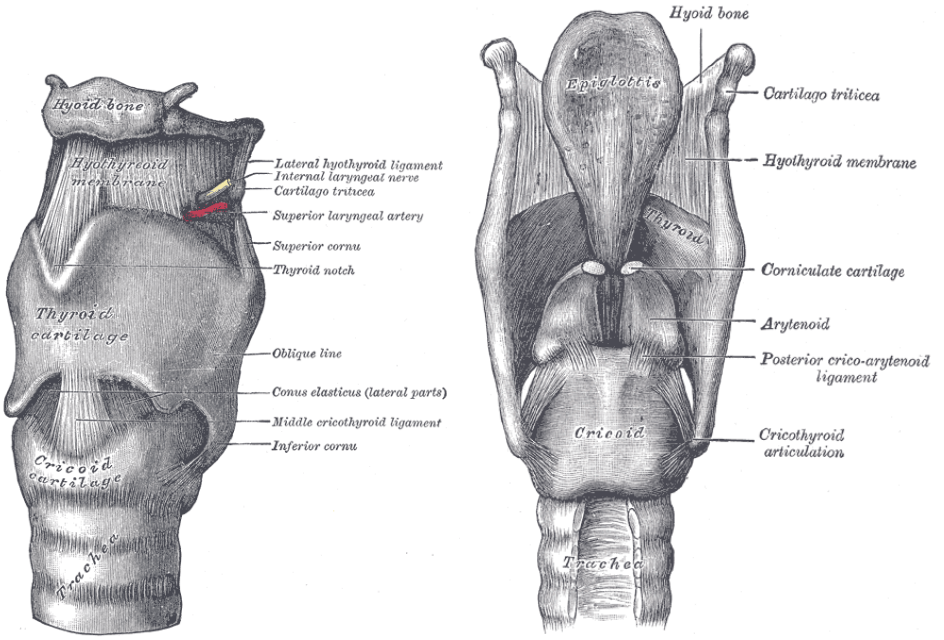
CHAPTER 2

The Human Voice

In the following chapter the basic anatomy and physiology of the human voice is described in detail. Starting with the anatomical features of the laryngeal system and moving towards the functional principles sound production. Later the neuroanatomical basis for vocal motor control and its feedback mechanism are reviewed before finally giving a short overview about neuroscientific and behavioral research of the vocal motor control and its sensory feedback systems.

2.1 Anatomy of the larynx

The larynx, which comprises the main source of vocal sound production is situated inferior to the root of the tongue and superior to the trachea. It is comprised of five cartilages (thyriod, cricoid, epiglottis, arytenoid and corniculate) and a range of external and internal laryngeal muscles, connecting cartilages with each other (internal) as well as to outer structures (external).

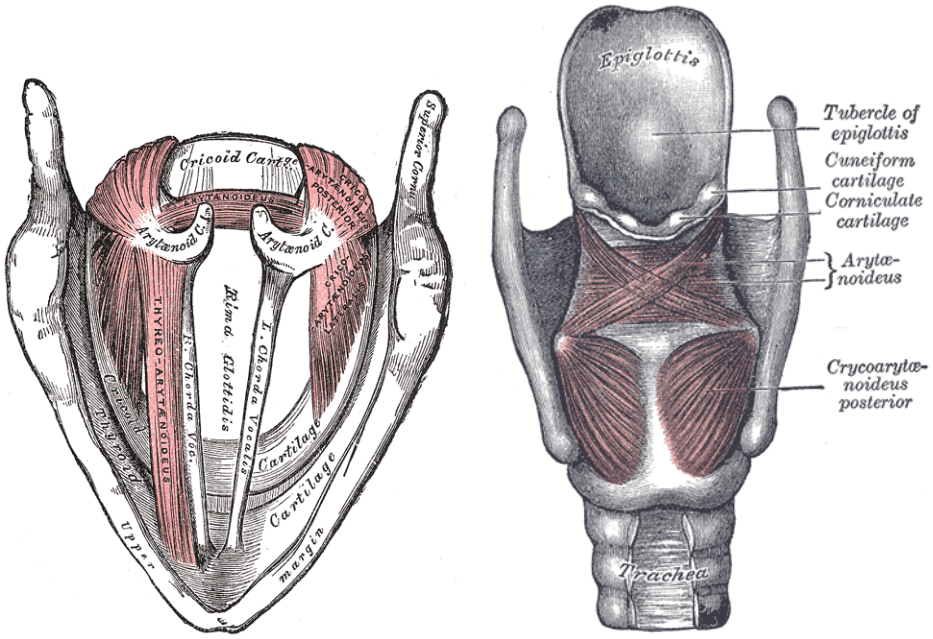


antero-lateral view posterior view

Figure 2.1: Antero-lateral and posterior view of the larynx with its different cartilages and ligaments. [Source: Gray (1918), accessed online via http://www.bartleby.com/107: 02.07.2013]

Figure 2.1 shows the antero-lateral and the posterior view in the larynx describing the different cartilages and ligaments while figure 2.2 shows the intrinsic and extrinsic laryngeal muscles and how they connect to the cartilages. The posterior cricoarytenoid muscle (PCA) is the only intrinsic

abductor muscle within the larynx and can open the glottis, while the lateral cricoarytenoid (LCA), the interarytenoid (IA) and the thyroarytenoid (TA) are adductor muscles involved in vocal fold closing. The cricothyronoid (CT) can stretch the vocal folds while moving the cricoid towards the thyroid. As the intrinsic laryngeal muscles connect the different laryngeal cartilages with each other, the extrinsic laryngeal muscles connect the larynx with external structures (e.g. the hyoid bone, sternum, pharynx). There are five external laryngeal muscles, but only the thyrohoid, the stenothyroid and the stenohyoid play an important role in the process of vocalization.

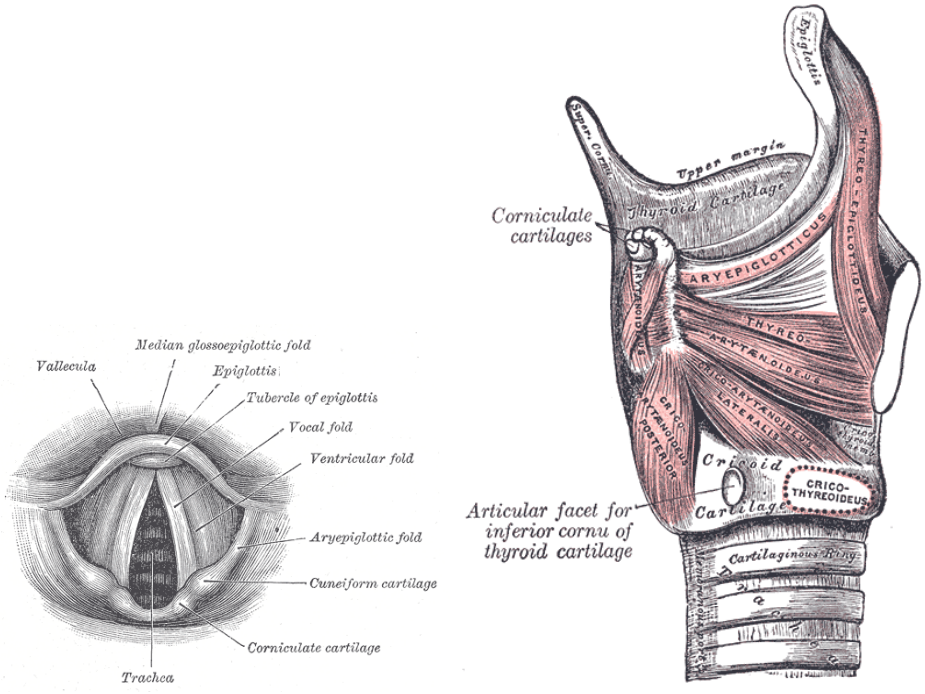


above view posterior view

Figure 2.2: The different intrinsic laryngeal muscles are shown and how they are connected to the cartilages. CT = cricoarytenoid muscle, TA = thyroarytenoid muscle, IA = interarytenoid muscle, PCA = posterior cricoarytenoid, LCA = lateral cricoarytenoid. [Source: Gray (1918), accessed online via http://www.bartleby.com/107: 13.04.2015]

They can alter the position of the larynx and therefore also influence the

tension of the vocal folds (Simonyan and Horwitz 2011), but mainly have the function to stabilize the larynx while the internal laryngeal muscles are active (Jürgens 2002). The external laryngeal muscles are innervated by branches of the ansa cervicalis profunda which arises from the ventral horn of C1-2.



above view posterior view

Figure 2.3: The different intrinsic laryngeal muscles are shown and how they are connected to the cartilages. CT = cricoarytenoid muscle, TA = thyroarytenoid muscle, IA = interarytenoid muscle, PCA = posterior cricoarytenoid, LCA = lateral cricoarytenoid. [Source: Gray (1918), accessed online via http://www.bartleby.com/107: 13.04.2015]

In the antero-lateral view of figure 2.1 one can see, that the internal branch of the superior laryngeal nerve (SLN) is entering the larynx supero-poterior to the superior laryngeal artery. The smaller external branch is not shown

in this picture. Besides the internal and external branch of the SLN, the recurrent laryngeal nerve (RLN) is the second nerve innervating the intrinsic laryngeal muscles and the mucous membranes of the vocal chords (see figure 2.1). Both nerve fibres (SLN and RLN) are bilateral branches from the vagus nerve (cranial nerve X).

The superior laryngeal nerve arises from the ganglion inferius and is descending along the pharynx before arborizing into the internal and external branch. The internal SLN is the principal sensory nerve that carries fibers from laryngeal chemo- and mechanoreceptors superior of the glottis (Trepel 2008; Sulica 2004). The external SLN therefore innervates the cricothyroid muscle (CT), which is the only tensor muscle of the larynx influencing intonation by tilting the inferior part of the thyroid outwards (see 2.2).

The course of the recurrent laryngeal nerve is different to each side of the body. Both nerve fibres leave the vagus nerve shortly after entering the thorax and continue descending. The right RLN turns left below the bow of the aorta, while the right RLN turns below the subclavian artery. Finally, both return between trachea and oesophagus back to the larynx. The RLN innervates all internal laryngeal muscles, thyroarytoid (TA), lateral and posterior cricoarythenoid (LCA, PCA), interarythenoid (IA), apart from the m. cricothyroidus (which is innervated by the external SLN). Moreover, the RLN transports afferent sensory information from the laryngeal mucous membrane below the glottis (Sanders et al. 1993; Trepel 2008; Simonyan and Horwitz 2011; Jürgens 2002).

Within the five intrinsic muscles, the thyroarytenoid (see figure 2.2 and figure 2.3), also known as vocalis muscle or vocal fold, plays a special role. It is the high frequency vibration of these two muscles which makes vocal utterances possible. For this reason, it is covered by a mucosa which contains

a large number of corpuscular mechanoreceptors, plexiform and free unmyelinated nerve endings (Adzaku and Wyke 1979), which makes the precise control of vibration possible. At the peripheral level, vocal fold muscular control is sub-served by a somatosensory reflexogenic system that integrates signals from those receptors and nerve endings to adjust vocal fold tension (Wyke 1974a,b; Sanders et al. 1998). Phylogenetically, the primary function of the vocal folds was to protect the airways. However, the distinctive and comprehensive amount of sensory information, relayed via brain stem nuclei and the thalamus to insular, somatosensory and motor cortices, makes a precise control for speech an singing possible. Especially, as some of these receptors react directly and frequency specific to vocal fold vibration, up to a frequency of 600 *Hz* (Davis and Nail 1987). Noteworthy, this information is based on animal research it is not clear if it is one to one transferable to the human. Nevertheless, receptors detecting vibration frequencies up to 600 *Hz* would cover the whole range (60 - 500 *Hz*) of fundamental frequency (F_0) necessary for speech (Traunmüller and Eriksson 1994; Standring et al. 2005). Physically, a stretch or increased tension in the vocal folds will results in a higher pitch to be produced when singing or speaking. Moreover, a change of the larynx position also results in a different formant structure of vocal utterances (Sundberg 1974).

2.2 Principles of vocal sound production

Apart from the laryngeal muscular and nervous system, there are two further 'components' important for vocal production:

1. A sub-laryngeal or respiratory component which includes the abdominal and thoracic muscles for creating a pressure in the lungs producing an air-stream through the larynx.

2. A supra-laryngeal component including lip muscles, jaw muscles, velum, and tongue muscles. They are modulating the air pressure waves created by the vocal folds by influencing the resonances.

These components and their muscles and nervous innervations are not explained in detail here, but a review can be found in Sundberg (1977) and Jürges (2002). However, in the following part their role in the production of vocal sounds is described.
The production of vocal sounds involves three physiological distinctive functions, which can be divided into *power source*, *oscillator* and *resonator* (Sundberg 1977). As power source functions a system of air pressure modulating muscles, organs and cartilages, while the oscillating or resonating system is reduced to the larynx and its muscles. The sound is finally shaped by a resonating and modulating system including for example the tongue and the oral cavity (see figure 2.4). Below the functions and the corresponding physiological elements are explained in detail.

1. Lungs (power source): The volume lungs can be compressed or extended through the contraction of the diaphragm and the intercostal muscles. The compressions leads to an pressure gradient between the lungs and the outside. This pressure gradient subsequently leads to an airflow

through the trachea and the larynx into the oral and nasal cavity. Finally the airflow passes out through mouth (or nose). Therefore, the lungs function as power source creating the airflow for the vibration of the vocal folds.

2. Vocal folds (generator/oscillator): The vocal folds are located at the base of the larynx. While completely open during breathing, they are initially closed during vocal production. With increased air pressure from the lungs, the vocal folds open for a short moment. The opening leads to an air flow and therefore local decrease in pressure (Bernoulli effect) in the larynx and the vocal folds close again. This process is repeated again and again in a very high rate creating periodic fluctuations in the air pressure. The peak to peak time of these pressure variations is inversely proportional to the fundamental frequency of the sound being produced by the air pressure changes. The vibration rate of the vocal folds are determined by their tension influenced by the intrinsic and extrinsic laryngeal muscles.

3. Oral and nasal cavities (resonator) and lips and tongue (modulator): The nasal and and oral cavities act as resonator, amplifying the sound waves generated by the vocal folds. Additionally, they function as a acoustic filter with a specific frequency curve. This frequency curve is created by the geometry of the cavities and mainly determine the perceived character of a voice. Moreover, they also influence the so called *formants*. Formants describe the frequency and bandwidth of specific amplitude peaks in the spectrum of spoken sounds e.g. a vowel. In voice research typically the first and second formant (f_1 and f_2) are used. Their relationship is specific for each vowel but different (in a

certain range) for each person and a result of the natural resonance behaviour of the vocal tract (Juslin and Scherer 2008). However, oral and nasal cavities are fixed and have a low influence in actively modulating the formants, despite lips and tongue. They are necessary to generate different vowels and consonants and therefore actively influence the formant structure compared to the oral and nasal cavities which have a mainly passive influence (due to the size and shape).

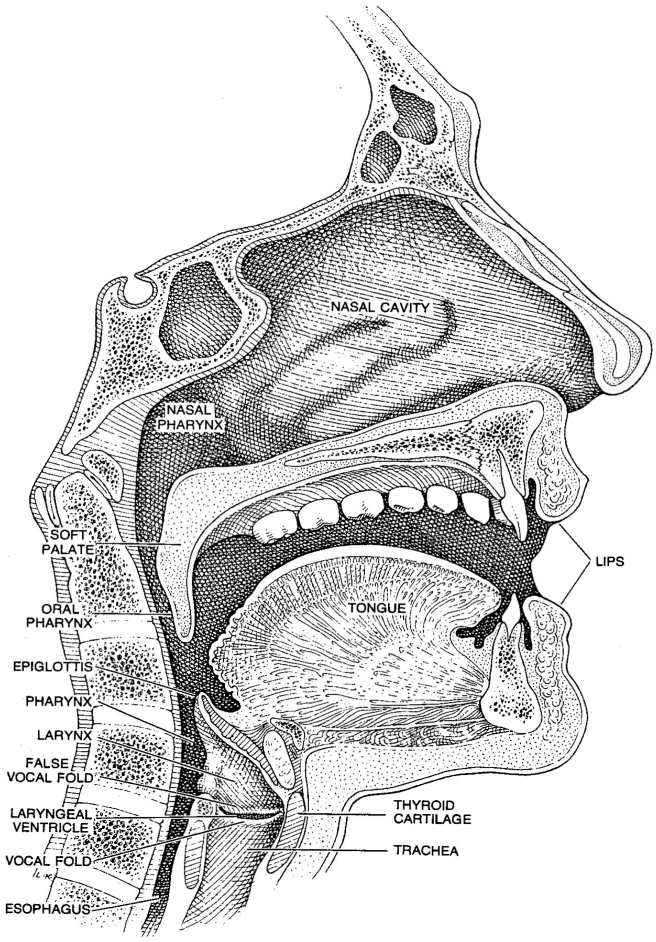


Figure 2.4: shows the mid-sagittal view of the front head and the neck. It depicts the position of relevant muscles, cavities and structures involved in vocal production. Images has been taken from Sundberg (1977).

2.3 Neuroanatomy of the vocal motor control and feedback system

The anatomical structures involved in vocal motor control can generally be divided into three subsystems (Simonyan and Horwitz 2011). These subsystems are hierarchical organised due to their phylogenetic development. For this reason, the lowest subsystem is located in the brain stem and is mainly responsible for laryngeal motor control during innate vocalisation. It includes the reticular formation and several phonatory and sensory nuclei. The second subsystem comprises the periaqueductal gray (PAG), the anterior cingulate cortex (ACC) and several limbic structures (e.g. hypothalamus, midline thalamus, amygdala, red nucleus and septum). It is involved in the initiation of vocalisation and the control of voluntary emotional vocal utterances.

The final and for humans most important vocal subsystem is the laryngeal motor cortex (LMC)[1], located at the ventral part of the primary motor cortex. It is essential for any voluntary vocalisation. Within the LMC, the organisation is topographical (Rödel et al. 2004). Lesions in these regions prevent speech production in humans, but have only minor effect on the vocalisation of monkeys (Jürgens 2002). It underlines the importance of the LMC in voluntary vocalisation and vocal learning. Interestingly, apart from the IA, all intrinsic laryngeal muscles are innervated bilaterally (Simonyan and Horwitz 2011). This means, that each hemisphere has direct connection to both sides of the larynx building a backup system for unilateral cortical damage (Jürgens 2002). However, in animal studies electric stimulation to the

[1] Laryngeal motor cortex (LMC) is now further used throughout the thesis to identify the region in the motor cortex which contains the representation of the laryngeal muscles

LMC effected vocal utterances differently depending on the stimulation side (Jürgens and Zwirner 2000). This is supported by several studies underlining the hemispheric asymmetry of the vocal motor control system. Lesions of the left hemispheric LMC influence speech production more strongly than singing and vice versa (Alcock et al. 2000; Riecker et al. 2000). These results support the notion, that there is a functional specialisation of left and right LMC (Simonyan et al. 2009). Nevertheless, the complexity of the human motor control system for vocalization is hardly comparable with those of even closely related mammals (Smotherman 2007).

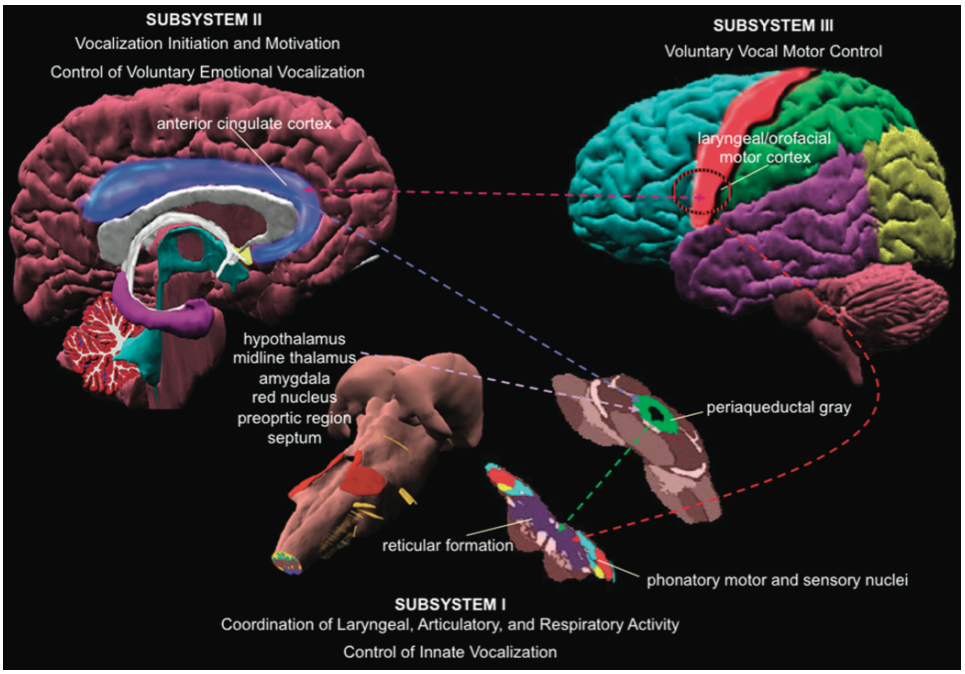


Figure 2.5: The image depicts the three brain areas important for vocal motor control and the resulting two distinct neural pathways. Image has been taken from Simonyan and Horwitz (2011).

Finally, the above described three subsystems result in two distinct vocal motor control pathways. An ACC-PAG path and the direct LMC pathway to the reticular formation (see figure 2.5). In the first pathway, the PAG

functions as an important relay, which is essential in eliciting vocal utterances from the ACC, while deactivation of the ACC can still relate in vocalisation initiated by the reticular formation (Jürgens 2002). The second pathway runs directly from the LMC to the reticular formation (Hannig and Jürgens 2006) and is essential in learning and execution of fine vocal motor control (Simonyan and Horwitz 2011).

However, motor-learning always involves sensory feedback mechanism to precisely calibrate motor sequences for prospective actions. In vocal production, like speech or singing, which are one of the most complex human motor behaviours, learning requires the highly precise feedback of timing of respiratory, laryngeal, and articulatory muscle groups and simultaneous integration of multi-modal sensory information (Titze 1993; Ludlow 2005; Kleber et al. 2013). It is known, that dynamic coordination of these subsystems for verbal communication develops over a protracted time period, in which oral motor patterns become consistent only around the age of 14 (Smith and Zelaznik 2004; Smith 2006).

In trained singers, for example, the somatosensory-motor loop contributes increasingly to pitch accuracy with developing singing expertise (Mürbe et al. 2004). This behavioural effect is in line with recent neuroimaging data, suggesting that **a)** experience in classical singing predicts cortical activation of areas processing somatosensory information from the larynx, and **b)** that enhanced pitch accuracy of professional singers is related to increased control of areas related to somatosensory feedback integration (Kleber et al. 2010, 2013). Based on theories from neural networks of speech acquisition, only in early stages of development the acoustical feedback is the dominant modality for sensorimotor speech learning and control (Tian and Poeppel 2010; Hickok et al. 2011; Houde and Nagarajan 2011; Guenther and Vladusich 2012). A

precise relationship between motor activity and acoustical consequence is learned until a stage is reached where almost no acoustical error in speech production occurs any more (Smith and Zelaznik 2004). Taken together, vocal production in speech and music is one of the most complex human motor behaviours and requires the highly accurate timing of several muscle groups and the continuous integration of multi-sensory information (Titze 1993; Ludlow 2005; Kleber et al. 2013). It is known that these systems develop in a specific time frame, reaching a ceiling level in the beginning of the teenage (Smith and Zelaznik 2004; Smith 2006). However, this might only be true for speech production, yet there is no research supporting this view for singing which involves a more accurate control of pitch. Possibly, pitch-control in speech reaches a ceiling, as demands are low compared to singing and therefore acoustical feedback is no longer needed for precise control of speech intonation.

Pitch, which is the fundamental vibration frequency of the vocal folds, is one of the perceptually most salient features of vocal production. It depends on the ability to control vocal fold tension (via intrinsic and extrinsic laryngeal muscles) and the release of air-pressure from the lungs (Titze 1993). Pitch sub-serves the production of continuous intonation contours in speech, which follow relative frequency relationships. In addition, in musical melodies it follows clearly prescribed interval relationships. This difference exhibits the specific demands of a control system for singing.

In music, even small deviations from target pitches will be perceived as errors by listeners and thus require higher accuracy in vocal motor control and sensory feedback integration for singing (Natke et al. 2003; Zatorre and Baum 2012). Singing errors are predominantly related to inaccuracies in pitch production (Dalla Bella et al. 2007). About 50 % of untrained singers

fail to match a target pitch to within half a semitone (equalling 50 cents pitch deviation) on half of their attempts whereas trained singers match pitches at significantly higher levels of accuracy (Hutchins and Peretz 2011, 2012).

The importance of auditory feedback for vocal pitch control has been studied extensively, mostly by examining compensatory responses to auditory perturbations, such as a pitch-shifting reflex in the opposite direction to change in voice feedback (Jones and Munhall 2000; Burnett and Larson 2002; Zarate and Zatorre 2008) or the decreased pitch-stability and accuracy when auditory feedback was masked by noise (Ternstrom et al. 1988; Mürbe et al. 2002). Yet auditory feedback is not the only way to control pitch (Amir et al. 2003; Zarate et al. 2010) but somatosensory feedback may also contribute significantly to vocal motor control (Lametti et al. 2012).

At the peripheral level, vocal fold muscular control is subserved by a somatosensory reflexogenic system that integrates signals from laryngeal myotactic mechanoreceptors to adjust vocal fold tension (Wyke 1974a,b; Sanders et al. 1998). These signals are then relayed via brain stem nuclei and the thalamus to insular and somatosensory cortices respectively, where acoustic goals become associated with kinesthetic representations of accurate motor commands (Jürgens 2002; Guenther and Vladusich 2012; Hickok 2012). Studies have shown that somatosensory feedback alone provides already sufficient information to produce precise vocal tract movements (Tremblay et al. 2003; Nasir and Ostry 2006, 2008; Lametti et al. 2012).

Overall, results from neuroscientific and behavioral studies point towards a model which underlines the importance of auditory and somatosensory feedback mechanism for precise vocal utterances. However, it might be, that the importance of each feedback channel (auditory or somatosensory)

changes with training. While in speech, the literature suggests a decrease in significance of the auditory feedback around a certain age (Smith and Zelaznik 2004), there is no study supporting this for pitch control in singing. Moreover, in singing auditory feedback only looses importance after a long time of professional training (Kleber et al. 2010; Mürbe et al. 2004).

CHAPTER 3

TMS - Transcranial Magnetic Stimulation

Transcranial magnetic stimulation is a non-invasive brain stimulation technique using rapidly changing magnetic fields to induce electrical currents in the cerebral cortex. These currents can elicit direct neural activation in a circumscribed area. The altered neural activation can than be measured in behavioural or neurophysiological tests. Over the last years a large number of stimulation types have been developed, depending on the effects to be achieved. Below, the history, principle mechanisms and more advanced technical details are explained, finishing with the stimulation protocols used in my experiments and their specific effects.

3.1 History

The history of transcranial brain stimulation goes back to the end of the 19th century. D'Arsonval (around 1896) placed subjects whole head in a rapidly changing magnetic field (110 V, 30A, 42 Hz), reporting phosphenes, vertigo and even syncopes. Further, rather simple experiments with exposure to changing magnetic fields have also been continued by Thompson and Dunlap (both around 1910/11). However, these were very basic attempts in influencing the neuronal activity of the brain. They could certainly not distinguish between effects on different areas in the brain or even influence on peripheral nerves or the vestibular system. (Barker and Freeston 2007). The introduction of the modern TMS machine, the way we use it today, goes back to 1985 on Barker and colleagues (Barker et al. 1985; Hess 2007) at Sheffield University. They developed the first portable stimulation system which was able to evoke specific responses from circumscribed cortical areas. This made it possible to test almost painless the functions of different cortical regions. It opened an interesting possibility to causally investigate brain functions. Moreover, it was a painless alternative to methods like electroconvulsive stimulation (Hess 2007). Today, TMS is widely used in experimental and clinical research.

3.2 Physics of TMS

Transcranial magnetic stimulation builds on the simple principles of electromagnetic induction. This means, every electric current is inducing a magnetic field. If the strength of the electric current changes than the induced magnetic field strength changes too. On the other hand, a changing magnetic field induces an electrical current in materials having freely moving electrons

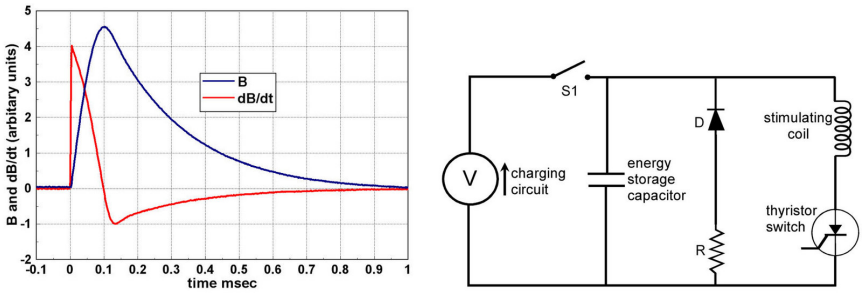
or ions. This bidirectional principle is in details described by *Maxwell's equations* and *Faraday's law* (Stöcker 2004).

In a TMS coil, an increasing and a decreasing, strong electric current flows circular through the coils and therefore induces a changing magnetic field around the coil. The magnetic field, which is induced by the electric current can be described by the *Biot-Savart law* (see equation 3.1). Its strength depends on the amount of current flowing through the coils and the distance from the coils (Stöcker 2004).

$$B = \frac{\mu_0}{4\pi} I \int_C \frac{dI \times \hat{r}}{r^2} \tag{3.1}$$

The magnetic field is not static but rapidly changing while the current flow in the coil is not steady. Figure 3.1 (a) shows the magnetic field and the corresponding electric field. The figure displays a mono-phasic pulse. A mono-phasic pulse has a current flow only in one direction (for details see below). Equation 3.2 demonstrates, how a changing magnetic field leads to a electric field gradient. This gradient ultimately leads to an ion flow in the brain perpendicular to the direction of the magnetic field vector, creating a depolarisation of neurons.

$$\Delta \times E = -\frac{\delta B}{\delta t} \tag{3.2}$$



(a) (b)

Figure 3.1: Image (a) shows how the current flowing in the coils relates to the created magnetic field of a single pulse. Diagram (b) depicts schematically the circuit of simple TMS system. Both images are taken from Barker and Freeston (2007)

Pulse types

The most simple form of magnetic stimulation is done by a mono-phasic single pulse as shown in figure 3.1 (a). In this case, the saved current from the capacitor flows only once through the coil before it is finally dissipated by a resistor and a diode. A mono-phasic pulse has a specific wave with only one peak. The advantage is, that effects caused by the pulse are more easily explained (Barker and Freeston 2007). However, most modern TMS stimulators, e.g. the Magstim Rapid^2 used in my studies, typically use biphasic stimulation pulses. They are even more efficient as they 'recycle' a part of the energy which than flows in the opposite direction. This allows faster pulse frequencies, which is especially important for fast repetitive stimulation protocols like theta burst stimulation (TBS). Further, overheating is less problematic with a biphasic stimulator as the remaining current after a pulse has not to be dissipated by a resistor. Nevertheless, biphasic stimulation has the drawback of having a current flow in both directions which comes with the cost of effects being harder to interpret (Siebner and Ziemann 2007).

3.3 Figure-of-eight coil and neuro-navigation

There are two different kinds of coils commonly used in TMS research. On the one hand, a simple round coil which has a broader circular area stimulating in the cortex and on the other hand the figure-of-eight coil. The latter one has the advantage of having a steep field gradient with one focal field maximum (see figure 3.2). The idea of the figure-of-eight has first been described by Ueno et al. (1988). The focality of the figure-of-eight coil makes it especially useful if only a circumscribed area should be stimulated. This is of general interest for testing a specific function of a cortical area as mostly done in research (Weyh and Siebner 2007).

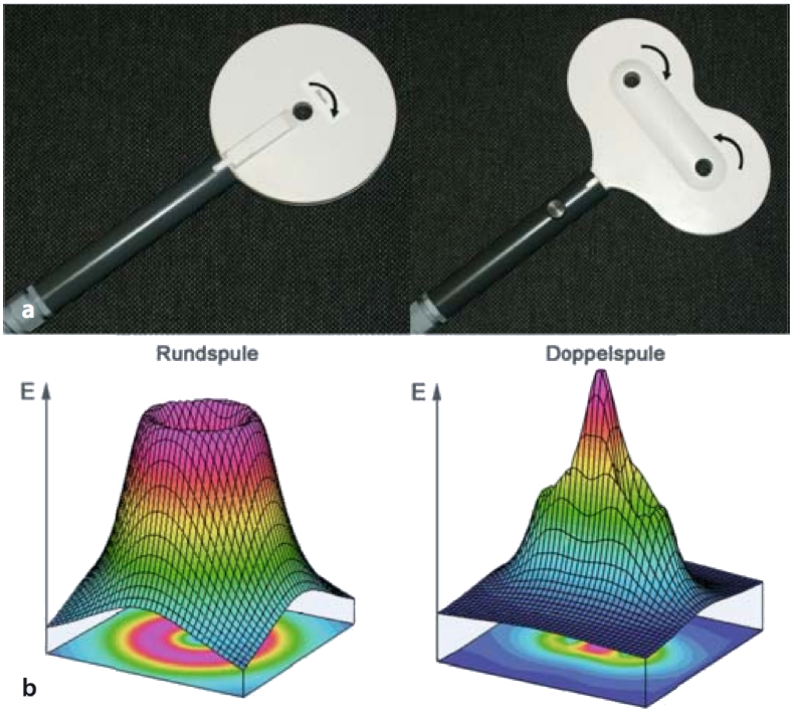


Figure 3.2: The two most common types of TMS coils and the electric field they induce. (Left) A round coil and (right) a figure-of-eight coil. Figure has been adapted from Weyh and Siebner (2007).

In my studies I used a Magstim figure-of-eight coil (The Magstim Company Limited, UK) with a diameter of 70 mm, similar to the one shown on the right in figure 3.2. While this coil ensures a focal stimulation area, I further needed to guarantee to stimulate on the right cortical area. Therefore, I used a stereotactic neuro-navigation system from Localite (Localite GmbH, Germany). This system uses MRI based structural images and a three-dimensional camera system to identify the exact position of the brain in the head. Together with optical reference marker on the head and on the coil it is possible to target a specific point in space, e.g. in the brain. The basic elements of the system are depicted in figure 3.3. A detailed description of the system can be found in Herwig and Schönfeldt-Lecuona (2007).

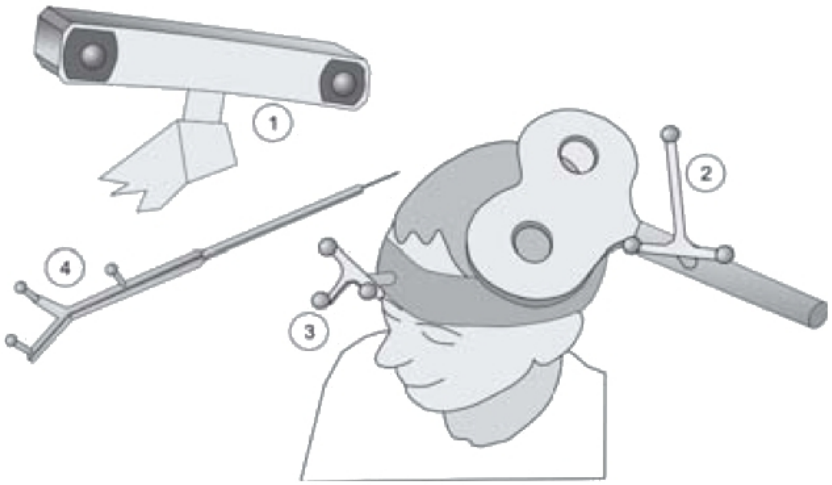


Figure 3.3: Elements involved in the usage of a neuro-navigation system. (1) The stereotactic camera, (2) reference marker for the coil, (3) head marker with tape, (4) registration pointer. Figure had been adapted from Herwig and Schönfeldt-Lecuona (2007).

3.4 Measuring motor threshold

To define the output intensity of the TMS system it has become a standard procedure to determine the motor threshold of a peripheral muscle, e.g. the first dorsal interosseous (FDI) or the abductor digiti minimi (ADM). This threshold is later used as a reference for the stimulation intensity applied over the cortical area to investigate. One can measure the threshold either with relaxed muscle (passiv motor threshold) or with the muscle contracted to about 10 - 20 % maximum strength (active motor threshold) (Kaelin-Lang 2007). The motor threshold (MT) is defined as the minimum stimulation intensity which is sufficient to receive a motor evoked response (MEP) in 5 out of 10 trials. Additionally, the muscle derived potential must be in the range of 100μV or above (Rossini et al. 1994; Rothwell et al. 1999). In the following the standard procedure, which has been used to determine the active motor threshold for the first experiment, will be described. It is based on a detailed description by (Kaelin-Lang 2007; Kaelin-Lang and Cohen 2000) as well as on personal correspondence with experienced colleagues in the field (Thielscher, Andoh).

(a) **Placing the electrodes using the 'belly tendon montage':** One starts in placing two electrodes on the chosen muscles. One electrode has to be placed on the belly of the muscle, while the second is placed at the end of the muscle close to the tendon. I took the FDI for deriving the muscle potentials as it is easy accessible and widely used in TMS research for the determination of the MT. Before placing the electrodes, the skin had been cleaned to optimise conduction. I used the 'Electrode Prep Pad'

(PDI, Canada) with 70 % isopropyl alcohol. A third electrode is used as a reference which has been placed on the collar bone, contra-lateral to the side the TMS stimulation is applied. I used self-adhesive ECG electrodes (Tyco Healthcare, Germany) with 24 mm diameter.

(b) **Visualization of signals:** Once the placement is done, electrodes had been connected via a signal amplifier to the computer to make the muscle potential visible using the BrainVision software (Brain Products, Germany). To receive a good signal, a 10 Hz high-pass filter and a 2000 Hz low-pass filter were used, as well as a 50 Hz notch filter, to remove possible influences from the power line. I chose a sampling rate of 5 kHz to certainly overcome problem of sub-sampling as described by the Nyquist theorem (Stöcker 2004).

(c) **Determination of hot spot:** To determine the region, which is closest to the cortical representation, of the muscle to be stimulated, a try and error approach had been used. The TMS coils had been placed perpendicular to the pre-central gyrus. This means that the induced current direction was 90 degrees to the spatial pathway of the gyrus (Thielscher et al. 2010). As starting point I chose a spot about 5 cm dorsal to the cross-section of midline and the thought line between the left an right mastoid bone. Before, I covered the head with a 'swim' cap which is overlaid with white tape on the interesting region. This gives the possibility to later mark the optimal position for the coil. Stimulator output was set to 60 % MSO for determination of the hot spot.

(d) **Determination of threshold:** After the best coil position is found the next step is to lower the stimulator output until only 5 out of 10 stimulus pulses produce a MEP with at least 100 μV. If this point was reached one starts a few percent lower this point and raises the output level until again 5 out of 10 stimuli produce MEPs. Importantly one has to keep a time interval of at least 5 seconds between each pulse to achieve a complete relaxation of the cortical potential to a normal level (Kiers et al. 1993). Therefore, the inter-stimulus-interval has been varied between 5 to 8 s (Kaelin-Lang and Cohen 2000).

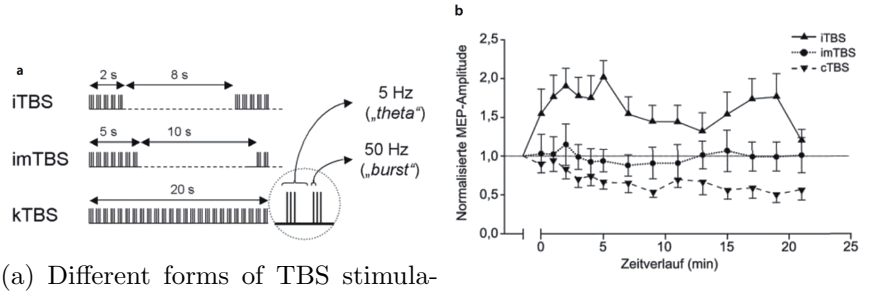
For TBS stimulation it is common to use 80 % of the stimulation output of the threshold found for the active MT (Lang and Siebner 2007). The AMT is, due to the pre-tension of the muscle, significant lower than the resting MT. In recent publications (Kalla et al. 2009; Leveque et al. 2013) a more time effective way to establish a stimulation threshold had been presented. These studies demonstrate, that 40 % of maximum simulator output (MSO) is a reasonable guideline for a stimulation threshold. This is in line with findings from my first study, where I found a mean of 41.4 % MSO for the 80 % AMT with a range of 38 % to 46 % MSO. Therefore, a MSO of 40 % seems to be a good guideline and in most of the cases threshold was even slightly below the 80 % AMT in my results.

Another critical point in using the MT for defining the strength of the stimulation output is the a priori justification of the transferability into other cortical regions (Castro-Alamancos et al. 1995; Tsang et al. 2014). There is no clear reason for using a motor cortex related threshold for example in the sensory cortex. Based on these arguments I used the 80 % AMT only in the first experiment and decided the used the fixed 40 % MSO in experiment 2.

3.5 TBS - Theta burst stimulation protocols

I used the theta burst stimulation (TBS) protocols to alter processing of the larynx area in primary somatosensory cortex. TBS stimulation protocols have been established by Huang and colleagues (2005) and became a popular tool, that allows off-line investigations, because they have long lasting effects and short stimulation time. This can be especially helpful if the task would make subsequent and precise targeting for stimulation impossible, like in the case in my experiment, where stimulation during singing would lead to facial muscle artefacts and therefore bias the task. Moreover, small head movements during singing would make it necessary to re-target the exact spot of stimulation each time, which is probably unlikely to achieve. Constriction of the head would be a solution, but lead to an biased singing performance. Theta burst stimulation is based on LTP and LTD induction protocols used in animal research (Larson and Lynch 1986). It contains of 50 Hz triple pulse bursts every 200 ms. Therefore, the burst repetition frequency is 5 Hz, which is in the range of the theta band (4 - 7 Hz) and gives the name to the protocol.

There are different forms of TBS. The intermittent form of theta burst stimulation (iTBS) leads to long lasting and long term potentiation (LTP) like effects, while continuous theta burst stimulation (cTBS) leads to long term depression (LTD). For example, Huang et al. (2005) demonstrated that cTBS leads to decreased MEP amplitude, while iTBS increased amplitudes. Both effects have been shown to be NMDA receptor dependent Huang et al. (2007). As NMDA receptor are important for learning (Riedel et al. 2003), TBS induced plasticity might reflect short-term learning like effects.



(a) Different forms of TBS stimulations protocols

(b) rTMS protocols and their effects

Figure 3.4: (a) Figurative sequence of three forms of TBS. (b) Effects of different TBS protocols on MEP amplitudes - difference between pre and post TBS. Figures have been adapted from Lang and Siebner (2007).[1]

More recently, several studies started investigating effects of TBS on primary somatosensory areas. Ragert et al. (2008) demonstrated successfully, that iTBS on the primary somatosensory representation of the index finger leads to improved two-point discrimination thresholds. The effect was equivalent to the one shown by Tegenthoff et al. (2005), using 5 Hz repetitive TMS with a similar paradigm. Morley and colleagues (2007), for example, deployed vibrotactile stimuli while using TMS to alter S1 functions, illustrating a decrease in discrimination threshold after iTBS stimulation. Additionally, Katayama and Rothwell (2007) showed that iTBS can alter in somatosensory evoked potentials (SEPs).

On the other hand, continuous theta burst stimulation (cTBS) leads to decreased temporal and spatial tactile acuity after stimulation over the hand area of the primary somatosensory cortex (Conte et al. 2012; Rai et al. 2012). Nevertheless, effects of theta burst stimulation are controversially discussed and might not always be consistent. For example, Hamada et al. (2012) demonstrated, that a cTBS protocol can have different effects even

[1]Captions of Figure 3.4 are in German. (a) kTBS is equivalent to cTBS. (b) x-axis: time after TBS administration, y-axis: changes in MEP amplitude.

within the same setup and might be individually different. Nevertheless, it is questionable if the results can be transferred to other stimulation protocols like iTBS or if this behaviour is specific for cTBS intervention. Furthermore, a voluntary tonic activation of the target muscle shortly after cTBS (about 1 minute) leads to LTP instead of LTD like effects (Huang et al. 2007). An adapted version of cTBS (30 Hz triple pulses with 6 Hz frequency) had been proposed by Goldworthy and colleagues (2012) which seems to have a stronger and more consistent effect.

Studies from animal research suggest that effects of theta burst stimulation might reflect cytoarchitectonically differences in the cortex and be different between granular and agranular areas, for example between M1 and S1 (Castro-Alamancos et al. 1995). Tsang et al. (2014) used a modulated cTBS version (Goldsworthy et al. 2012) testing its effect on M1 and S1. They measured MEP and short-latency afferent inhibition (SAI) and demonstrated a opposite effects depending on the stimulation site. While cTBS on M1 decreased MEPs, it had no effect on SAI, while cTBS on S1 led to increased MEPs and shortened SAI.

In conclusion, TBS stimulation has been successfully used in several studies without reporting considerable problems. There is a growing amount of studies demonstrating the effectiveness of TBS on S1. The short administering time and the effective time range up to 20 minutes is a clear advantage. It makes it possible to study TMS related effects off-line in a easy and effective way. However, this comes with the cost of an incomplete understanding of the cortical effects of TBS. Therefore, results have to be considered carefully.

CHAPTER 4

Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is a modern and powerful non-invasive imaging technique making it possible to record in vivo high resolution brain images. With the beginning of the 1990s it has rapidly been developed to standard tool in neuroscientific research. This chapter describes the basic physical principles of MRI and the applied standard technique e.g. event related sparse sampling design.

4.1 History & Basic Principles

Today magnetic resonance imaging is broadly used in clinical diagnosis and research. It allows to acquired high resolution structural images of internal parts of the body, e.g. the brain. Moreover, it enables the mapping of neural activation pattern onto those structural images. This is a non trivial technical process based on changes in cerebral blood flow and proportion of different types of haemoglobin.

It were Charles Roy and Charles Sherrington from Cambridge University, who already in the 1930s experimentally showed that brain function is related to cerebral blood flow. However, it took many more years to make it technically feasible to measure blood flow experimentally in the brain. It were Ogawa an colleagues (1990) who could demonstrate that changes in the concentration of oxygenated (Hb) and deoxygenated (dHb) haemoglobin can be measured in vivo using magnetic resonance imaging. Deoxyhaemoglobin is paramagnetic and can therefore been used as natural contrast agent in MRI measurements. Using a gradient-echo imaging technique they measured the blood oxygen level dependent (BOLD). BOLD is a direct measure for the ratio of Hb and dHb and in indirect measure of blood flow. So finally after about 60 years things have come full circle.

With the spread of MRI scanners and technical advances in magnetic field technology today it is possible to acquire brain activation even in real-time. The advantage the BOLD signal acquisition using gradient echo is its sensitivity which detects changes in neural activation between 0.5 % and 3 % (Ogawa and Sung 2007). However, the spatial resolution is limited to a range of a few millimetre, which makes it only possible to identify activity of large neuronal ensembles. The largest drawback might be the very slow temporal

resolution, as the signal is based on changes in the vascular system, which is relatively slow (in the range of seconds). Therefore, it is easily possible to distinct between several temporal clearly separated events, but makes it harder to study the timely dynamics of neural systems (Faro and Mohamed 2006).

4.2 Physics of fMRI

Magnetic resonance imaging makes use of the basic atomic properties called spin. The spin is a quantum mechanic phenomenon which describes an angular momentum. It can be conceptualised as the rotation of a proton on a specific axis and is described by the spin number s and the angular spin momentum S. It is a vector with a specific direction and value (length). This characteristic creates what one can measure as magnetic dipole moment (Faro and Mohamed 2006).

$$s = n/2 \tag{4.1}$$

where n is a non-negative integer

$$S = \frac{h}{2\pi}\sqrt{s(s+1)} = \frac{h}{4\pi}\sqrt{n(n+2)} \tag{4.2}$$

where h is the Planck constant

Magnetic resonance imaging (MRI)

The angle of the spin can be influenced by applying an external magnetic field. This is the case inside an MRI scanner, which has a static magnetic field with a typical strength between 1.5 and 3 Tesla. Inside this field, the

spin direction aligns with the magnetic field lines leading to a homogeneous spin direction for all atoms. The alignment can than be either parallel or anti-parallel to the outer magnetic field. In reality, there is no exact 50:50 ratio in the split of parallel and anti-parallel alignment, but always a slight surplus of parallel aligned spins, which is energetic more efficient. The alignment also depends on other external factors like body temperature and magnetic field strength of the scanner and is in the range of 6 to 1.000.000. Although this surplus is relatively low, the large amount of protons, for example in hydrogen atoms, leads to strong enough magnetisation within a specific volume element (voxel).

If one now applies an electro-magnetic (radio) pulse to the protons it will lead to a short displacement of the spin orientation. Returning back to the alignment direction, forced by the outer magnetic field, all protons precess along the longitudinal axis back to the parallel or anti-parallel orientation. This excitation is only possible with a specific radio-frequency called *Larmor* frequency, which is given by the magnetic field strength B and the properties of the molecule γ (see: Faro and Mohamed (2006); Stöcker (2004); Weißhaupt et al. (2014); Siemens (2003).

$$f_{Larmor} = \frac{\gamma}{2\pi} B \tag{4.3}$$

where B is the magnetic field strength and γ a molecule specific constant.

The *Lamor* frequency is specific for each molecule. This means, the strongest excitation is elicit with the molecule specific radio frequency. Different frequencies have only a low or no impact on the atomic spin. While the molecule returns to the optimal energy state, e.g. in alignment with the out magnetic

field, it precess round emits energy in the form of radio wave. This radio waves can now be detected from coils with the MRI. For magnetic resonance imaging it is of interest to measure the time from the point of maximum excitation to the time point the spin has returned to a certain degree to the low energy state. Those are called relaxation times. The time constant for the longitudinal relaxation to reach the initial longitudinal alignment of the spins along the B-field lines is T1 or spin-lattice relaxation time. For example, the relaxation time T1 is defined as the point of maximum longitudinal excitation to around 63 % while the time T2 represents a measure for the transversal relaxation time. T1 is depending on the type of tissue and is between 300 ms (fat) and 2000 ms (water). The transverse relaxation time T2 is much smaller and usually around 30-150 ms. The time depends on the type of chemical bonding of the hydrogen atom and is therefore different for each type of tissue. The relaxation times play an important role in tomography since they are different for each tissue type. In T1-weighted images, the structure of the tissue is particularly well seen. Since the momentum of the protons is very small, they are repeatedly excited and several measurements are averaged (averaging using $Fourier$ analysis). In this way the signal-to-noise ratio of the image is improved. The result of several high-frequency pulses is called sequence. The time between excitations is the repetition time (TR) the time between excitation and signal acquisition is the echo time (TE).

The image contrast of different tissues depends on the tissue parameters (T1, T2) and the sequence parameters (TR, TE). For example, a T1-weighted sequence has a short TR and TE. Tissues with short T1 appear more white (white matter, which contains fat), while tissues with a long T1 appear darker (gray matter, which contains high amount of water). These kind of sequences are mainly used to capture a high-resolution structural image

of the brain/head. Images of that kind can be used as a later reference for functional images, which are then overlaid to the T1-weighted image to exactly identify for example neural activation in a region of interest. However, they do not have to be overlaid with functional images, but also can be used as stand-alone images. Then one can mark specific cortical landmarks, for example for later TMS stimulation (see: Faro and Mohamed (2006); Dawson and Lauterbur (2008); Weißhaupt et al. (2014); Ogawa and Sung (2007)).

Functional resonance imaging (fMRI)

So far, I have discussed the general mechanism of recording signals from atom using magnetic resonance imaging. However, to make interference about the changes in neural activation this previously explained technique has to be adapted. To track neural activation, fMRI measurements rely on the BOLD (blood oxygen level dependent). For this reason, I am coming back to the role of oxy- and deoxyhaemoglobin. In case of haemoglobin, the spin differs if it is oxygenated (s = 0) having paramagnetic properties or deoxygenated (s = 2) having diamagnetic properties (Pauling and Coryell 1936). This is useful as neural activation leads to local changes in the ratio of oxy- and deoxyhaemoglobin. With each form of haemoglobin having different magnetic features, it is now possible to track these changes in the ratio of both. This ratio is best identified using the T2* relaxation time. T2* is a measure of magnetic field homogeneity and depends on the presence of deoxyhaemoglobin (Faro and Mohamed 2006; Ogawa and Sung 2007). However, not only the strength of the change in the homogeneity has to be captured, but also the exact spot where this change happens. For this reason, three additional magnetic gradient fields have to be used. While the static MRI magnetic field is homogeneous in the centre of the scanner, a gradient field can slightly change the field strength. As we have seen before

4.3, the Lamor frequency depends of the magnetic field strength. Having a gradient field in z-direction [1] only a specific plane inside the field will respond to the excitation pulse. Since we now have chosen a single plane we need two more spacial information to separate the brain into defined voxels. For this reason another gradient field is applied in y-direction shortly after the excitation of the spins. This leads to a gradient in the Lamor frequency from top to bottom and subsequently to a phase difference. Lastly, a gradient i x-direction is applied and lead to a frequency difference from left to right or vice versa. The data recorded with the head coil generate a data matrix called the k-space which represents the phase- and frequency distribution in 2-dimensional space. Only after applying a Fourier transformation the original brain data is recovered (see Siemens (2003); Weißhaupt et al. (2014)).

4.3 Experimental designs - sparse sampling

In neuroscientific research basic fMRI paradigms have been established as standard procedure. In the first experiment I have used the so called *sparse sampling design*. The sparse sampling design, which only acquires brain data at certain defined time points after a specific task, can avoid interferences of scanner noise with the task. This is especially useful in tasks that include auditory stimulation or production of vocal utterances as auditory stimulation leads to unwanted brain activation and movement during the scanning leads to movement artefact of the acquired image data. However, as there is less data acquired, it goes along with a cost in statistical power and therefore one might need more trials in an experiment to gather statistical meaningful

[1] In a thought 3-dimensional orthogonal grid inside the scanner the z-direction is in line with the static magnetic field (front to back), while x is left to right and y is up and down. The intersection of all three point is in the centre of the scanner.

data (Lazar 2008).

I decided to use the sparse sampling technique in the localiser task to reduced task-related and -unrelated movement artefacts. The technique was based on those used in previous studies by Kleber and colleagues (2007, 2010). The paradigm consisted of one block with 72 volumes (whole head scans), acquired by a 3 Tesla whole body Scanner (Magnetom Tim Trio, Siemens, Germany) employing echo planar imaging (EPI, echo time [TE]: 30 ms, repetition time [TR]: 10 s, acquisition time [TA]: 3s, 40 transversal slices, 3.4 mm thickness, no gap).

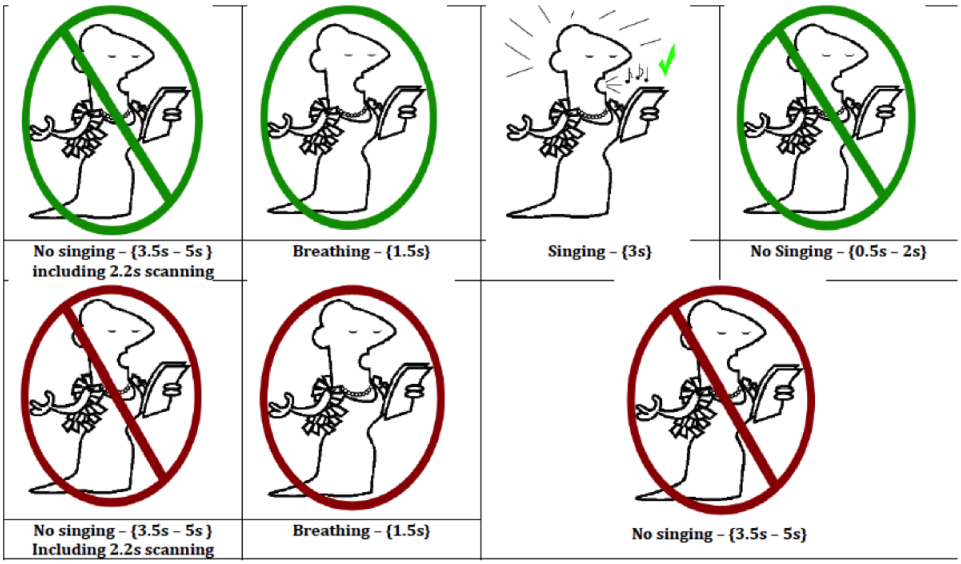


Figure 4.1: Schematic display of the fMRI design deployed. Green circle indicate singing trial and red circle rest trial. Paradigm had been adapted from Kleber et al. (2010).

Subjects received visual cues (see figure 4.1) indicating them either to sing the steady vowel /a/ in a comfortable pitch (green circled) or to exhale without sound production (red circled). It is important to acquire the brain data at the point of maximum BOLD response, as the TA of 2.2 seconds is

relatively short in respect to a TR of 10 s. The maximum BOLD response can be expected about 3 s to 5 s after task onset (Faro and Mohamed 2006; Ogawa and Sung 2007). Therefore, the start of fMRI data acquisition was varied between 3.5 seconds and 5 seconds after singing onset to grasp a larger timespan and efficiently cover the maximum BOLD response.

4.4 Data analysis procedure

The analysis of fMRI data comprises a series of steps to pre-process the recorded data and finally to identify regions with task-specific activation. These steps are the *realignment and co-registration* of the time-series being recorded during the scan. In this step six movement regressors (*x, y, z, pitch, roll, yaw*), simultaneously recorded during the fMRI scan, are used to correct for small movements the participant did during the scans. Further, each voxel of each time-series is 'connected' (*co-registered*) to voxels of a high resolution structural T1-weighted image. This enables a later activation tracking to the exact cortical area. For later group analysis, the next analysis step would incorporate the so called spatial *normalisation*, which is a morphological procedure to deform each subjects' brain that it reflects most cosily a given standard brain. However, this step was omitted in this case, as the analysed brain data needed still to fit the subjects brain, for the TMS neuronavigation procedure. Group results were not relevant as individual activation pattern were only used for the exact TMS targeting of activation spots. Skipping the *normalisation*, the next step is called *smoothing*. Here a gaussian kernel of usually 6 mm (two times the image voxel size) is applied. Primarily this improves the signal to noise ratio as neighbouring data points are average in

signal intensity, while reducing the spatial resolution of the images.

With these steps, the spatial preprocessing is finished and is followed by first-level statistical analysis. Here, the different experimental conditions (singing, listening) are modelled and a general linear model is applied to identify statistical differences between conditions within one subject (Penny et al. 2011). Statistical difference can be seen as changes in neural activation and overlaid on a high-resolution anatomical scan as colour-coded maps. Activation peaks can also been identified and their coordinates can subsequently been used for precise targeting during an TMS experiment.

CHAPTER 5

Behavioral Testing and Statistics

In this chapter behavioural tests being used in the course of the study for assessing pitch perception and pitch reproduction performance are explained in more detail. Further, statistical procedures, especially how finally the dependent variables had been extracted from the raw signal, are illustrated. Lastly, standard and self developed questionnaires are described.

5.1 Pitch discrimination task

All subjects underwent a pitch discrimination task to rule out that poor pitch perception influences pitch reproduction accuracy. A two-tone forced choice test had been used to estimate hearing thresholds. Subjects were presented with two pure tones (250 ms duration, 600 ms gap between tones) with a starting frequency difference of Δf = 7 %. The lower standard tone was fixed at 500 Hz while the order of both tones was randomized. Δf was adaptively changed using a two-down one-up rule, which tracks 70.7 % correct thresholds on the psychometric function (Levitt 1971). After two consecutive correct responses Δf was decreased and increased after one incorrect responses by the factor β. Initially, β was set to 2 but changed to 1.25 after the second reversal. After 15 reversals the test was terminated. The final threshold was calculated by the geometric mean of Δf of the last eight reversals. Participants were not included in the study when their pitch perception cut-off exceeded a 2 % frequency discrimination difference, which corresponds to a 35 cents interval (100 cents = 1 semitone). Typical frequency discrimination thresholds are 0.86 % in healthy non-musicians and 0.13 % in trained musicians (Micheyl et al. 2006). The selected criterion in the study thus guaranteed that participants' perceptual thresholds were at least 65 % lower than the smallest pitch interval used for singing in this study[1].

[1]The smallest interval used was one semitone, which corresponds to 100 cents which is about 3 times the cut-off value.

5.2 Pitch reproduction task

Primary goal of both experiments was to test changes in pitch reproduction accuracy and stability due to TMS intervention. Therefore, subjects had to perform a singing task. They were presented with a series of musical intervals (consecutive tone pairs; 105 intervals in experiment 1; 60 intervals in experiment 2). The total number of intervals was reduced, as 60 trials provided sufficient statistical information to detect reliable differences in pitch performance between sessions. Subjects had to reproduce the intervals in singing on the syllable 'na' or 'la'. Tones were played over Bose QuiteComfort 15 noise cancelling headphones (Bose, Germany) and pitch reproduction was recorded via head mounted microphones AKG C 477 WR L/p (AKG, Austria) and a m-audio firewire audio system using custom tailored presentation and recording script for the MAX/MSP software ('74 Cycling, USA).

Each pair of tones (900 ms duration for each tone, 200 ms gap between tones) was followed by a period of 3 seconds in which subjects had to reproduce the two tones. In experiment 2 this time period was filled with pink noise which was played over the headphones. In this case, pink noise was set at highest just tolerable level for each participant individually, to efficiently mask subjects' own voice.

The first tone always started at 311.11 Hz (D#4) for female subjects and 155.565 Hz (D#3) for male subjects, respectively. The second tone differed between ± 9 semitones from the first tone. Target tones were complex waves on the syllable /a/ as previously used by Hutchins and Peretz (2011), designed to reflect the timbre of the human voice. Each subject produced the series of intervals before and after theta burst stimulation containing 19

difference intervals ranging from 0 to ± 9 semitones. The exact distribution of the intervals can be found in appendix A.

Pink noise

Pink noise is a modification of white noise (Ward and Greenwood 2007; Gelfand 2009). White noise is a artificially designed sound with a constant spectral density function $S(f) = f$. While the frequency range (usually 20 Hz to 20 kHz) covers the the whole human hearing range, it is perceived as high frequency stressed noise. The reason for this is, that the loudness of all frequencies is equal. However, as the human perception is not linear, different frequency need to have different acoustics pressures to be perceived as equally loud. (For details see ISO 226 2003) To correct for this mismatch between physical sound pressure and perceived loudness pink nose is used. It follows the spectral density $S(f) = 1/f^{\alpha}$ with $\alpha = 1$ which reflects more accurately a equally balanced loudness perception for each frequency. Physically spoken, the density function compensates that higher octaves contain a larger number of frequencies with each octave the frequency doubles. This is important to efficiently mask all frequencies, also lower ones, which are 'under-represented' in white noise. However, a complete masking will probably not be achieved due to the bone conduction of the skull.

Normal vocal range and interval sizes

To ensure that subjects are able to reproduce the tone pairs played to them, they have to be in a adequate range. Subjects have to be able to reproduce the tones, while the range of intervals has to be challenging enough to detect a possible effect of improvement. In the literature one can find a vocal range of 28 to 37 semitones for women and 29 to 37 semitones for man (for a review

see Hacki 1999). Moreover, Hacki (1999) reports for the lower end of the male vocal range a frequency 55 - 103 Hz and for the upper end a frequency of 415 - 880 Hz. For females, respectively, he states 97 - 207 Hz to 659 - 1566 Hz. A similar range has been reported by Heylen et al. (2002). In the above described pitch reproduction task procedure I only used tones within the reported range. Therefore, it can be expected that all subjects are able to reproduce the given tones.

5.3 Audio data analysis

Audio recordings have been analysed using a custom-made script within the CUEX performance analysis system (Friberg et al. 2007) run under Matlab.[2] Due to weak phonation in some subjects the analysis was customised according to the specific examples regarding pitch range and analysis methods. The analysis was separated for the group of female and male subjects and the corresponding frequency parameters were adapted correspondingly. A high-pass filter (Butterworth, order 4 or 8) was set to the lowest expected sung frequency and used to filter out the low frequency components of noise. The remaining noise floor was estimated and onset and offset times of the two notes were determined from the crossing points of the sound level envelope. The same envelope were low-pass filtered and with a level offset. See Friberg et al. (2005) for details. The pitch was estimated using the YIN algorithm (de Cheveigne and Kawahara 2002) with the frequency range limits restricted to the range used in the experiment. This was C#3 to D5 for the females and F2 to B3 for the males. Additionally, the test tones, which were played to

[2]Audio data analysis has been done in cooperation with Dr. Anders Friberg from KTH Stockholm.

the participants, were analysed. The results showed that the analysed pitch deviated in most cases less than one cent from the intended pitches, thus a negligible deviation within this context considering that the just noticeably deviation is about 10 cents in good conditions.

There were a small number of examples that were omitted due to either that the subjects did not sing, they just provided one note, or they started too late. Also, a few examples could not be analysed due to pitch tracking problems. Since drop-outs were relatively few they were considered unproblematic and had been omitted.

Two fixed time constants were used for the pitch analysis. The analysis started 50 ms from the detected onset time, thus omitting the pitch onset which typically exhibited very large pitch variations. The **initial pitch** section was constituted by the following 200 ms (50 ms to 250 ms after onset). This section corresponded approximately to the initial pitch adjustment by the subject (Grell et al. 2009). The remaining part of the note starting at 250 ms after the onset to the end of production constituted the **final pitch**. For both the initial pitch and the final pitch the median and the upper and lower quartiles were computed. The median represents a measure for the overall singing accuracy within the initial or final pitch component if compared to a given reference. Further, the range was computed as the difference between upper and lower quartile, which reflects a measure for the **stability** within a pitch component.

Further, the slope of the initial pitch was estimated using a linear least square polynomial fit. In addition, the median vibrato rate and extent were estimated from the entire tone, as well as the onset, offset, sound level, onset rise time and a spectrum parameter. However, these parameters had not been used for statistical analysis as they were not related to my a priori

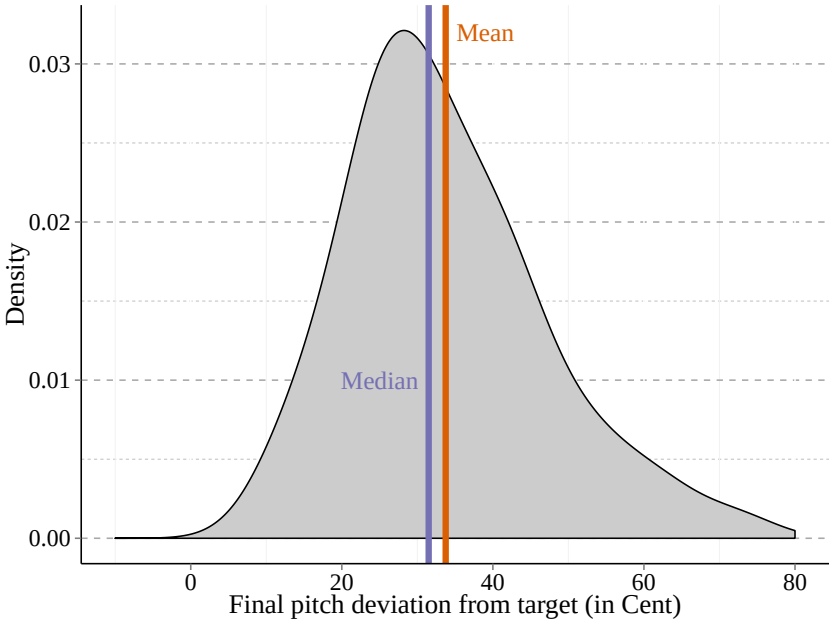
hypotheses. Figure 5.2 represents a recorded audio signal and shows visually the computed pitch variables for a two semi-tone interval.
Starting with the extracted data from the raw signal I created three variables: pitch accuracy, interval accuracy and pitch stability, which had been computed for each component (initial and final) and all trials. Their creation is explained in every detail below.

(A) **Pitch accuracy:** For each tone component (e.g. initial pitch) the median of the estimated pitch level contour had been computed, which represents a measure for tone height of the sung tone. The median had been used, as it is relatively unaffected by variations in the pitch level signal (see figure 5.1). This is especially important as subjects are untrained singers and show larger variations within the sung tone (Zarate and Zatorre 2008; Kleber et al. 2013). Further, the difference ΔA_{ij} between presented tone and the median of the recorded tone was calculated. Where i is the tone number (either tone 1 or tone 2) and j is the trial number. This difference is a relative measure and could be positive or negative, depending if the subject sings flat (lower than intended = negative deviation) or sharp (higher than intended = positive deviation). For this reason, the absolute value of the difference $|\Delta A_{ij}|$ had been calculated to later prevent cancelling out of positive and negative values. As participants had to sing two tones, I finally averaged across both tones, giving one absolute value $|\Delta A_{mean,j}|$ corresponding to each trial for the initial and for the final component, respectively. These values denote the deviation from the target pitches presented to the subjects and reflect their overall singing accuracy. The higher the value, the poorer the accuracy is.

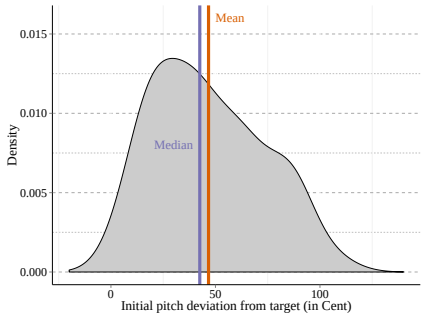
(B) **Pitch stability:** The tone height of the estimated pitch level varies within each tone component depending on the singing performance of the subject. Therefore, I used the range of the estimated upper and lower quartile of a component as attribute for the stability of the singing. A smaller pitch level range signifies a higher performance stability. Stability values had been averaged across tone one and two and resulting in one value for each trial for the initial and the final component, respectively. A lower value reflects a higher pitch stability.

(C) **Interval accuracy:** This measure reflects how accurate the presented musical interval was reproduced, independent from the tone height the subjects starts and ends with. For example, if a subjects sings the first and the second tone flat, the accuracy of each tone might be poor. However, the interval could still be correct. Therefore, the interval accuracy is defined by the difference: $\Delta A_{2j} - \Delta A_{1j}$, which is the pitch deviation from tone 2 minus tone 1. Here, a negative value is the result of a interval sung to small, while a positive value is a interval sung to large. The greater the dispersion from 0, the poorer the interval accuracy is.

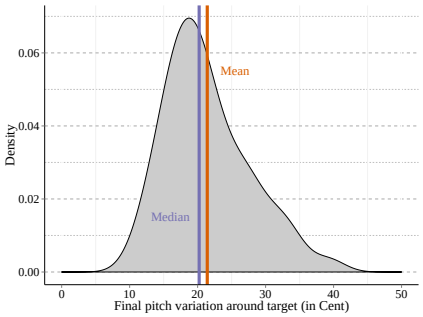
In the next step I looked on the distribution of pitch accuracy and stability and could see, that they were skewed to the lower end. This can be seen in figure 5.1. I also included the mean (orange) and the median (purple) of this distribution as vertical lines. Obviously, the median represents much better the maximum of the distribution. Therefore, I decided to use the median of all trials as the definite measure for each subject.



(a) Final pitch accuracy distribution



(b) Initial pitch accuracy distribution



(c) Final pitch stability distrbution

Figure 5.1: These three plots show density distributions for three different dependent measures from three different subjects. The vertical lines represent the mean (orange) and median (purple) demonstrating, that in all cases the median more closely relates to the maximum of the distribution. The y-axis denotes the calculated density and the x-axis the measures for deviation from (a, b) or stability of (c) the target pitch.

In summary, the above explained computation leads to six dependent variables for each subject which give information about the absolute (accuracy and

interval accuracy) and relative (stability) pitch reproduction performance. Initial and final pitch accuracy (and interval accuracy) describe how accurate a subject reproduced the two tones in comparison to the given tones (or interval), while initial and final pitch stability depicts how much variation is on average in the pitch component without referring to an given absolute reference. Below there is a short summary of all dependent variables.

(1) **Initial pitch accuracy:** The absolute pitch difference of the initial pitch component between presented and recorded tones, being averaged across both tone components of the interval. Finally, the median of all trials had been taken.

(2) **Final pitch accuracy:** The absolute pitch difference of the final pitch component between presented and recorded tones, being averaged across both tone components of the interval. Finally, the median of all trials had been taken.

(3) **Initial pitch stability:** The inter quartile range of the pitch height variation the early pitch component in absolute values, being averaged across both tones. Finally, the median of all trials had been taken.

(4) **Final pitch stability:** The inter quartile range of the pitch height variation the final pitch component in absolute values, being averaged across both tones. Finally, the median of all trials had been taken.

(5) **Initial pitch interval accuracy:** The difference of the initial pitch component between presented and recorded tone 1 minus the difference of the initial pitch component between presented and recorded tone 2. From This value the median of all trials has been taken.

(6) **Final pitch interval accuracy:** The difference of the final pitch component between presented and recorded tone 1 minus the difference of the final pitch component between presented and recorded tone 2. From

This value the median of all trials has been taken.

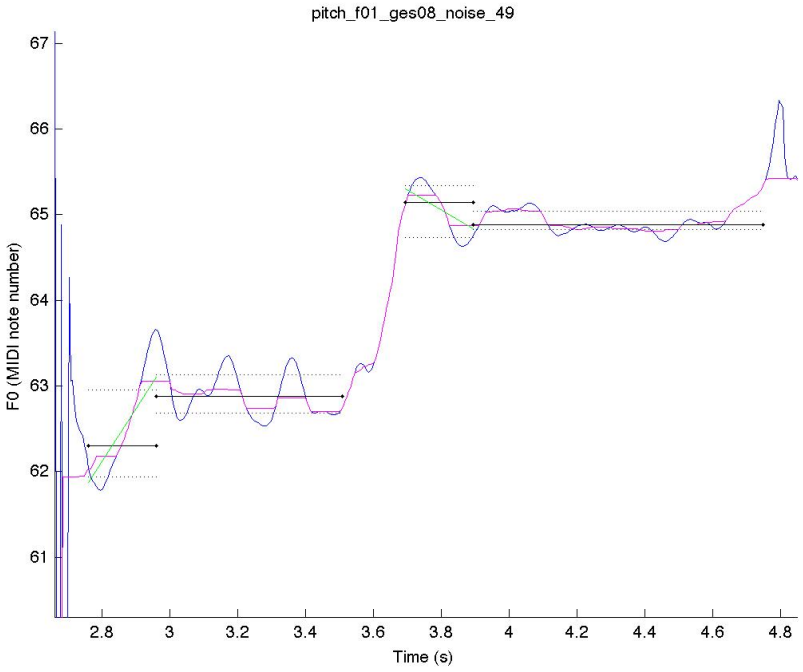


Figure 5.2: shows the recorded audio signal (blue line) and the corresponding estimated pitch level (pink line) for a typical used interval from both studies. On the x-axis the time in seconds is annotated and on the y-axis the pitch height in MIDI notion is labeled. the difference of one MIDI tone is equal to a halftone or a pitch difference of 100 cents. MIDI tone 63 corresponds to the musical D#4 (311.11 Hz). Moreover, the computed median of the pitch (black line) and the upper and lower borders of the range (dotted black lines) are depicted. The green line describes the pitch direction from initial to final component.

5.4 Statistical analysis

Statistical analyses were carried out using R (R Core Team 2014). Shapiro-Wilk[3] test revealed in both studies non-normal distributions of the data. Hence, non-parametric tests were applied throughout all analyses. To iden-

[3]According to Ghasemi and Zahediasl (2012) the Shapiro-Wilk test proves to be the most powerful and reliable test for evaluating normality.

tify a main effect Friedman's ANOVA had been used to compare the effect of stimulation type (independent variable) on the pitch reproduction accuracy and stability (dependent variable; see section *Audio Data Analysis*). Friedman's ANOVA is the non-parametric version of a repeated measures ANOVA.

Post-hoc comparison were done using the Wilcoxon sign rank test. In case of multiple comparisons, p-values had been adjusted using the FDR-method (false discovery rate method). However, post-hoc comparisons were based on carefully chosen a priori hypotheses to avoid unnecessary correction for multiple comparisons. Further, one-sided post-hoc test had been applied as the effect direction on iTBS and cTBS had been defined before.

5.5 Questionnaires

Both studies included a range of different questionnaires to examine safety issues, demographics and musical background, and the subjectively perceived singing performance. Below, all used questionnaires are listed and their content is shortly explained. Copies of each questionnaire can be found in appendix C.

(a) **fMRI safety questionnaire:** This standard questionnaire had been used in the both studies as required by the standard safety procedure. It asks participants for any possible contraindications (e.g. metal implants, pacemaker, pregnancy) which would lead to exclusion of the study for safety reasons.

(b) **TMS safety questionnaire:** Comparable to the fMRI safety questionnaires possible contraindications for participation in an TMS experiment are controlled. Although fMRI and TMS both include magnetic fields, contraindications for TMS are partly different to an fMRi experiment, while a specific questionnaire is need.

(c) **Edinburgh Handeness Inventory:** A widely used questionnaire in psychological research for testing the handeness of participants (Oldfield 1971).

(d) **Musical background:** This had been developed for earlier studies (Kleber et al. 2009; 2010). It asks for the singing and instrumental experience as well as the general musical background and music traditions in the family.

Part II

Experiments and Discussions

CHAPTER 6

Experiment I: Continuous and intermittent transcranial magnetic theta burst stimulation on S1 and its effect on pitch reproduction when auditory feedback is available

6.1 Introduction

In this first study I investigated the role of the laryngeal representation in the right ventral post-central gyrus for vocal control during a two-tone pitch reproduction task in non-singers using repetitive transcranial magnetic stimulation. An fMRI localiser had been used to identify the corresponding cortical region in S1 functionally representing the larynx. Further, an image-guided neuro-navigation system had been utilised for precise targeting. iTBS and cTBS stimulation had been applied to the target site and additionally, iTBS on a dorsal S1 control site. I hypothesised, that iTBS stimulation on the target site will lead to improvement in the reproduction task, while cTBS will impair performance, and stimulation in the control site will have no effect. Each condition had been compared to a neutral baseline which had been recorded prior to stimulation in each session.

6.2 Material and Methods

6.2.1 Participants

10 subjects (5 male) with a mean age of 23.9 years (range: 20 - 33 years) from the body of students of the University of Tübingen participated in the study. None of them had academic training in music nor professional vocal training. All subjects reported normal hearing and no cases of pathological hearing problems. Hearing abilities had been controlled using a pitch discrimination task (see *6.3.2*). All subject were right handed as controlled by the Edinburgh Handeness Inventory (Oldfield 1971). The inventory revealed a mean score of 80.7, with a range from 66 to 100. Moreover, musical behaviour and background was accessed using a self developed custom tailored questionnaire (see Kleber et al., 2010). No subject had previous singing training One participants was removed before final analysis as its pitch reproduction performance was highly inaccurate and unstable within and between different sessions (see *6.3.3*) Another subject had not completed the study and was already discarded before analysis.

6.2.2 Ethics Statement

Written consent was obtained from all subjects prior participation. The study was approved by the ethics committee of the University of Tübingen. Inclusion criteria for TMS experiments were adhered according to Wassermann (1998).

6.2.3 fMRI Technique

Structural whole head images were acquired using T1-weighted images (magnetisation prepared rapid acquisition gradient echo; 176 sagittal slices, 1 mm effective thickness) for anatomical reference. Functional images had been recorded using a sparse sampling method. The technique was based on those used in previous studies by Kleber and colleagues (2007, 2010). The paradigm consisted of one block with 72 volumes (whole head scans), acquired by a 3 Tesla whole body Scanner (Siemens Magnetom Tim Trio, Germany) employing echo planar imaging (EPI, echo time [TE]: 30 ms, repetition time [TR]: 10 s, acquisition time [TA]: 3s, 40 transversal slices, 3.4 mm thickness, no gap). Subjects received visual cues indicating them either to sing the steady vowel /a/ in a comfortable pitch or to exhale without sound production(See *4.3* for more details including a figure of the visuals cues).

6.2.4 Transcranial Magnetic Stimulation

The TMS setup consisted of a Magstim Rapid^2 stimulator with a 70 mm biphasic figure-of-eight coil (The Magstim Company Limited, UK) and a Localite stereo-tactic neuro-navigation system (Localite GmbH, Germany). High resolution MRI scans (1mm^3) had been acquired for all participants as basis for precise targeting. For this reason, structural MRI scans were overlaid with functional activation maps.

I used the intermittent theta burst (iTBS) and continuous theta burst (cTBS) stimulation protocols introduced by Huang et al. (2005). iTBS consists of 50 Hz triple pulses every 200 ms for a total time of 200 s and cTBS of 50 Hz

triple pulses for 40 s continuously. Stimulator output was set to 80% of AMT for each subject individually and kept constant for all TMS sessions. AMT was measured in the first TMS session using the procedure described in *3.4*. TMS stimulation was applied over the right somatosensory vocal fold representation as identified using the fMRI paradigm. Control stimulation was applied over the most dorsal part of right S1. Subsequently, target stimulation over somatosensory vocal fold representation are called *larynx-iTBS* and *larynx-cTBS*, while stimulation over the dorsal S1 control area will be called *dorsal-iTBS*.

Coil orientation was anterior-posterior and perpendicular to the post-central gyrus for the larynx-iTBS and larynx-cTBS. However, based on previous findings (Thielscher et al. 2010) the coil was tilted by 45 degree for dorsal-iTBS control stimulation to further reduce the effect on the cortex and create a sham like stimulation with similar skin sensation.

6.2.5 Behavioural Testing

Behavioural testing included the pitch discrimination task and the pitch reproduction task which are explained in detail in chapter 5.

6.2.6 Procedure

The experiment consisted of five separate sessions which were carried out at 5 separate days.

Day 1: On the first day, subjects were screened for being qualified to take part in a fMRI and TMS experiment, to rule out any possible health

issues. I used standardised safety questionnaires from the lab. They can be found in appendix C. Subsequently subjects performed the perceptual screening task (see chapter 4). Hereby it was possible to check beforehand if perception abilities were with a adequate range. Otherwise subjects had been excluded from further participation.

Day 2: On day two subjects underwent an fMRI scan to identify relevant cortical regions involved in simple tone production. Therefore, they had to do a simple voice production task. Additionally, I recorded a high resolution structural brain image as basis for the neuro-navigation system.

Day 3 – 5: Session 3 –5 included finally the main experiment, the pitch reproduction task. At the beginning each session subjects were equipped with noise cancelling headphones and a head mounted microphone. Subjects trained the paradigm with 2-5 trials depending on individual needs then they sang a series of 105 two-tone intervals before and after TMS. During the singing task the investigator left the room to not create an uncomfortable situation which might influence performance. Subjects were naive to what effect the the TMS stimulation should have on their singing performance.

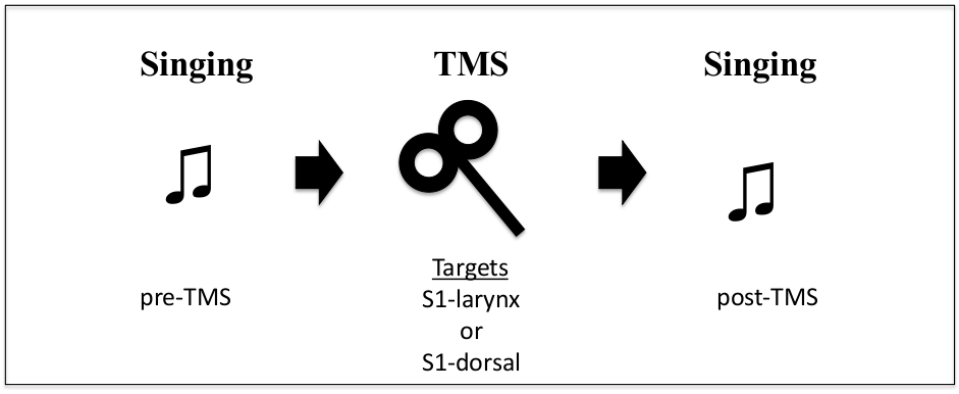


Figure 6.1: Plot shows the pitch reproduction test procedure of each session. Subjects did a singing task pre-TMS, then received the TMS stimulation, and did again the singing task post-TMS. Stimulation was in a pseudo-randomised order with either iTBS on larynx-S1, cTBS on larynx-S1 or iTBS on dorsal-S1.

6.2.7 Audio Data Analysis and Statistics

Audio data analysis and statistical interference were computed using software and methods presented in chapter 5.

6.3 Results

6.3.1 Pitch discrimination task

All participants had normal to very good hearing abilities as controlled by the pitch discrimination task. The average frequency perception difference was 0.5% (8.6 Cent) with a range from 0.17% (2.9 Cent) to 1.08% (18.6 Cent). Therefore, pitch discrimination thresholds are much smaller than the smallest interval (one semitone = 100 Cent) used in the pitch reproduction task.

6.3.2 Pitch reproduction task

In this subsection I present the results from the pitch reproduction task. However, before revealing the effects of of TMS stimulation I are going to investigate the individual singing performance to make sure all subjects had normal singing abilities. As I want to facilitate statistical procedures, in the next step I test for differences between the pre-TMS performances of each session.

Checking for normal singing performance

I initially investigated pitch reproduction performance of each subject to make sure all participants had stable and normal pitch reproduction abilities. For this reason I explored the variance of the final pitch variable for the first tone of the pre-TMS condition. As the first tone (311.11Hz = D#4) is kept equal for all intervals it should be easier to be reproduced than the

second tone which varies in pitch. Moreover, subjects adapt to the first tone during the task. Therefore, variance should be low, if singing performance is normal. Figure 6.2 shows box-plots for pre-TMS final pitch reproduction performance of all subjects.

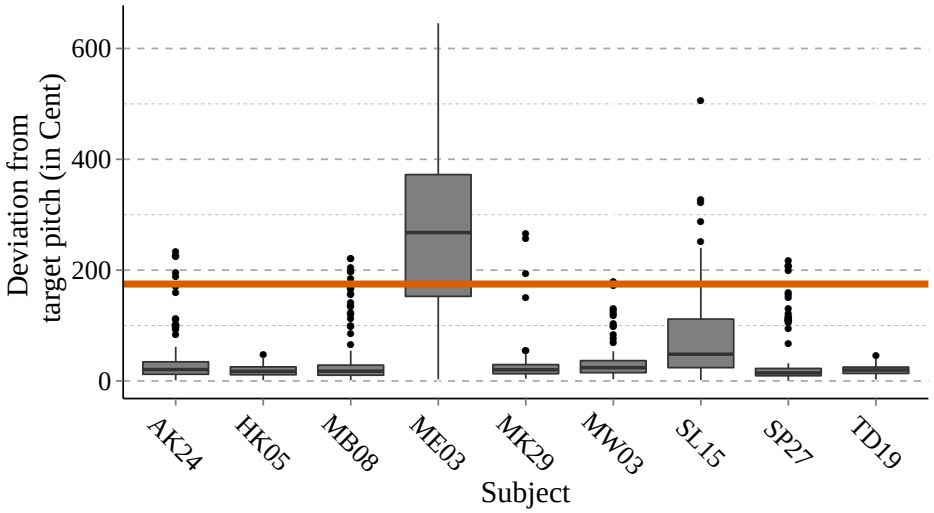


Figure 6.2: Plot shows box-and-whisker diagrams for final pitch accuracy for the first tone of the pre-TMS condition for each subject. Y-axis represents the deviation from target pitch and x-axis shows the participant code for each individual subject. The orange line marks the defined cut-off for the exclusion criteria, which is one standard deviation above mean group performance.

It can be seen in figure 6.2 that subject ME03 had a highly unstable performance compared to the remaining subjects. The red line marks the one standard deviation boundary (1 SD) which was defined as cut-off criteria for inclusion in the further analysis procedure. Therefore, subject ME03 had been removed from the study before further analysis to prevent any bias due to this unstable performance.

Combining pre-TMS performances

To ensure that pre-TMS performances were stable over all sessions I compared each session to each other using Friedman's ANOVA. Table 7.1 shows the results for the baseline comparison of the four dependent variables tested. Results demonstrate no significant differences between the pitch reproduction performances of each pre-TMS. Therefore, the combining of all pre-TMS performances to one overall pre-TMS measure is justified and will be used for later analysis and comparison with intervention pitch reproduction performances (S1-larynx-iTBS, S1-larynx-cTBS, S1-dorsal-iTBS).

Table 6.1
Comparing baselines

Variable	Results		
	df	χ^2	p-value
Initial pitch accuracy	2	3.25	0.20
Final pitch accuracy	2	0	1
Initial pitch stability	2	1	0.61
Final pitch stability	2	0.75	0.69

Table 6.1: Friedman's ANOVA test results for comparison between baselines of different sessions. $N = 8$.

Performance after TMS

Pitch reproduction performance has been compared for overall interval accuracy, pitch accuracy and pitch stability. Main effects of TMS intervention had been investigated using Friedman's ANOVA. Comparing the pitch accuracy variables reveals no significant main effect, neither for the initial component, χ^2 (3, $N = 8$) = 5.10, $p = .17$, nor for the final pitch component, χ^2 (3, $N = 8$) = 1.65, $p = .65$. Figure 6.3 shows the results for the pitch accuracy measure. In the next step, I compared the stability variables, which also

revealed no significant main effect, neither for the initial pitch component χ^2 (3, $N = 8$) = 2.40, $p = .49$, nor for the pitch component, χ^2 (3, $N = 8$) = 1.05, $p = .79$. In figure 6.4 one can see the results for the pitch stability measure.

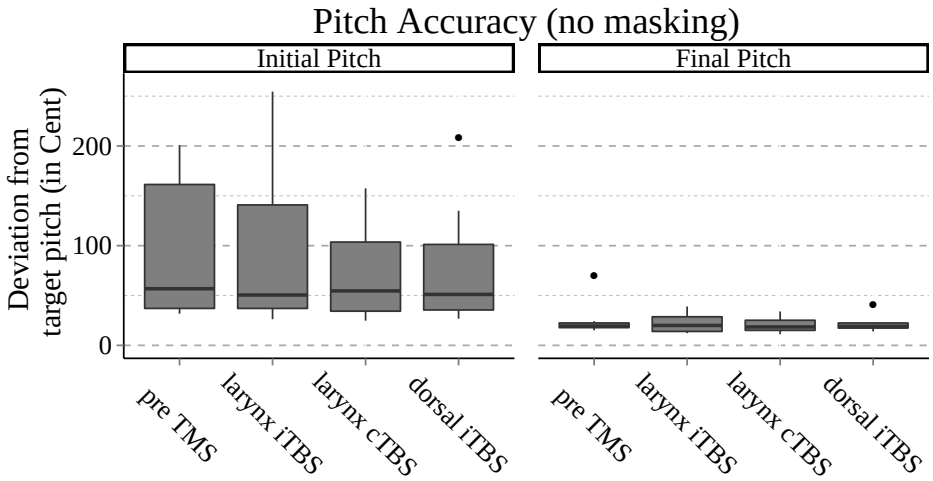


Figure 6.3: Box-and-whisker plots display deviation of sung tones from target pitch for pre-TMS and all three post-TMS conditions (S1-larynx-iTBS, S1-larynx-cTBS and S1-dorsal-iTBS). $N = 8$, where N is the number of subjects.
Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

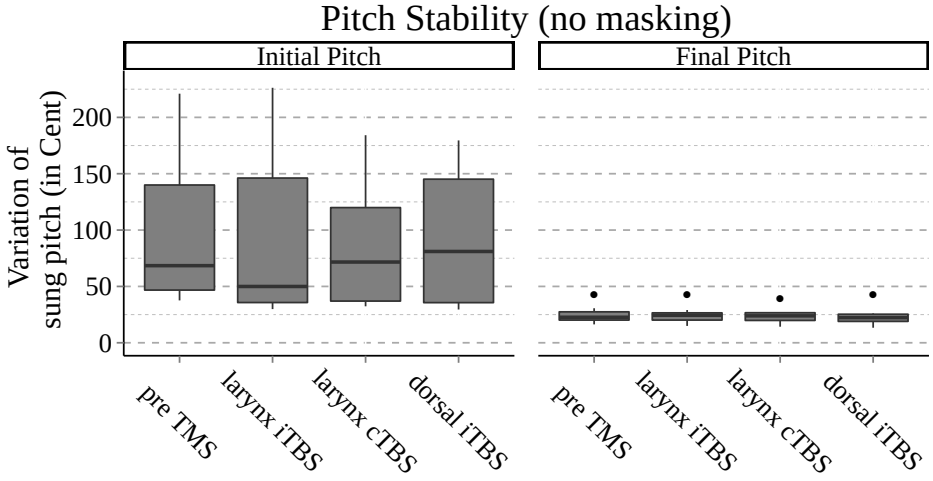


Figure 6.4: Box-and-whisker plots display deviation of sung tones from target pitch for pre-TMS and all three post-TMS conditions (S1-larynx-iTBS, S1-larynx-cTBS and S1-dorsal-iTBS). $N = 8$, where N is the number of subjects.
Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Additionally, I also compared the interval accuracy for initial and final pitch component. Again, it revealed no significant main effect, neither for the initial pitch component χ^2 (3, $N = 8$) = 0.45, $p = .93$, nor for the final pitch component, χ^2 (3, $N = 8$) = 1.65, $p = .65$ using Friedman's ANOVA. Figure 6.5 depicts the results for the interval accuracy measure.

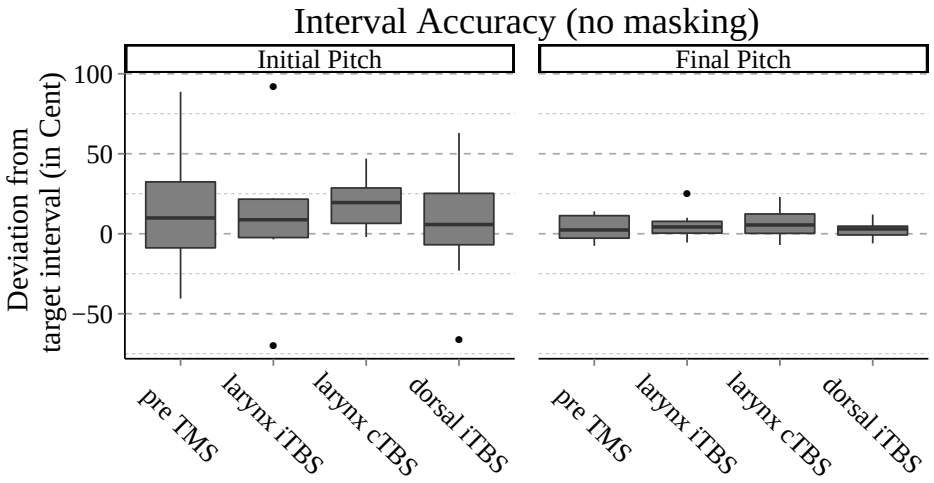


Figure 6.5: Box-and-whisker plots display deviation of sung tones from target pitch for pre-TMS and all three post-TMS conditions (larynx-iTBS, larynx-cTBS and dorsal-iTBS). $N = 8$, where N is the number of subjects.
Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Overall, there were no significant effects of TBS stimulation on pitch reproduction, which could been illuminated by investigating pitch accuracy, pitch stability or interval accuracy. Possible implications of the results are discussed in the following section (6.4) and in "Chapter 8 - Discussion".

6.4 Preliminary conclusion

In this first experiment I used continuous and intermittent theta burst stimulation protocols to alter the accuracy and stability in a pitch reproduction paradigm testing untrained singers. I expected theta burst stimulation to change the ability to vocally reproduce musical two-tone intervals. Based on previous studies (Ragert et al. 2008; Morley et al. 2007; Premji et al. 2010) I hypothesised that iTBS will lead to improvement and cTBS to deterioration of singing performance compared to a neutral baseline and stimulation on a control site (dorsal part of post-central gyrus).

However, my results revealed no significant effect of TMS stimulation on singing performance compared to baseline performance. I could not demonstrate that applying theta burst stimulation for altering the processing of the somatosensory representation of the laryngeal muscles leads to changes in vocal motor performance. I expected these behavioural changes manifest itself in improved or impaired pitch reproduction accuracy and stability after TMS stimulation. Moreover, performance and changes in performance due to stimulation are not correlated with pitch perception skills.

As shown in previous studies, theta burst stimulation had been successfully applied to the primary somatosensory cortex (Ragert et al. 2008; Katayama and Rothwell 2007; Morley et al. 2007; Premji et al. 2010; Rai et al. 2012). Neither iTBS nor cTBS stimulation had an measurable effect on the behavioural performance. I did use standard settings which have been established in the literature (e.g. 80 % active motor threshold for stimulation intensity) and used a modern neuro-navigated targeting system, which minimises errors in targeting the exact spot to a few millimetres.

Nevertheless, effects after theta burst stimulation on S1 have only been test

in the field of sensory related performance measures (e.g. two-point discrimination on the finger). So far, no study demonstrated any cross-modal effect of the motor system after altering sensory processing. For the same reason it could be that short term functional changes in the primary somatosensory system do not or only marginally transfer to other modalities. Moreover, I cannot rule out that a unilateral stimulation will be compensated by the contralateral hemisphere. This is especially reasonable in the case of vocal control, as there is evidence that each hemisphere controls both sites of the vocal folds (Simonyan and Horwitz 2011; Jürgens 2002).

In my view the lack of any significant effect of TMS on singing performance might be explained by the complex interactions of two feedback systems being used to monitor pitch during vocal production – the auditory and the somatosensory system. It has been shown that non-singers rely more on auditory feedback for voice control compared to professional opera singers (Kleber et al. 2010; Zarate and Zatorre 2008). Therefore I assume the auditory system playing a greater role in non-professional singers despite the TMS intervention, which might only have a small or no effect on lay singers. Amateur singers might have not learned to use the somatosensory system. Further, it has been shown that singers rely more on internal models for vocal control, while amateurs have to build on real-time feedback (Jones and Keough 2008). Possibly, the auditory feedback is preferred real-time feedback for lay singers. Then, any effect on the somatosensory system would than be covered by the auditory feedback system strategy subjects had been used or just covered by the general variability in singing performance.

In summary these results do not rule out any importance role for somatosensory feedback in vocal control. However, to clarify the role of the ventral area of S1 on vocal motor control, a design is needed eliminated the confound

of auditory feedback.

CHAPTER 7

Experiment II: Intermittent transcranial magnetic theta burst stimulation on S1 and its effect on pitch reproduction when auditory feedback is masked

Introduction

To further investigated the role of sensory feedback on vocal motor control in singing I implemented a second experiment. This time, to reduce the influence of the auditory feedback loop, pink noise was played to the participants via headphones during the period of pitch reproduction. Similar to the first experiment, participants had to reproduce a series of two-tone intervals before and after administration of transcranial magnetic stimulation. Only S1-larynx-iTBS stimulation has been used to reduce complexity of the design. This protocol is has been proven to be effective in improving cortical processing in S1 in other studies (Ragert et al. 2008; Morley et al. 2007; Katayama and Rothwell 2007). Moreover, it seemed more likely to see an improvement effect after noise masking, which already leads to less accurate singing.

7.1 Material and Methods

7.1.1 Participants

14 subjects (7 male) with a mean age of 27.3 years (range: 22 - 35 years) participated in the study. Only Participants with no or very low previous vocal or formal musical training and without reported normal hearing were included in this study. 4 Subjects from study 1 also participated in this experiment. Musical experience was accessed using a custom tailored questionnaire (see Kleber et al., 2010). This questionnaire investigated subjects' musical background, e.g. asking for possible instrumental and vocal lessons, as well as general music related behaviour. All participants were right handed according to the Edinburgh Handeness Inventory (see appendix A). Only subjects without reported neurological or psychiatric diseases participated in this study. Written consent was obtained from all subjects prior participation. The study was approved by the ethics committee of the University of Tübingen. Inclusion criteria were adhered according to Wassermann (1998).

7.1.2 Behavioral Testing

Prior to participation, all participants underwent a perceptual screening to rule out that poor pitch perception can be accounted for pitch reproduction accuracy. A two-tone forced choice test was applied that had been used previously (Kleber et al., 2013) to estimate frequency discrimination thresholds (see 5.1 for a detailed description). The main task incorporated that participants listened to 60 trials of two-tone musical intervals via headphones

and subsequently sang them back using the syllable 'na' or 'la'. Auditory feedback was masked with pink noise during pitch reproduction to perturb acoustic feedback. Vocal production was recorded and stored digitally for off-line analysis (a detailed description can be found in section 5.2).

7.1.3 MRI Technique

Structural whole head MRI images were acquired using T1-weighted images (magnetisation prepared rapid acquisition gradient echo; 176 sagittal slices, 1 mm effective thickness) for anatomical reference for precise targeting. In three cases this data could be obtained from a previous study as subjects participated again.

7.1.4 Transcranial Magnetic Stimulation

The TMS setup consisted of a Magstim Rapid^2 stimulator with a 70 mm biphasic figure-of-eight coil (The Magstim Company Limited, UK) and a Localite stereotactic neuronavigation system (Localite GmbH, Germany). I used the intermittent theta burst stimulation protocol (iTBS) introduced by Huang et al. (2005) which consists of 50 Hz triple pulses every 200 ms for a total time of 200 s. Coil orientation was anterior-posterior (AP) in both cases and perpendicular to the post-central gyrus for the larynx-iTBS stimulation of the vocal fold representation but only 45° tilted for the dorsal-iTBS stimulation on the most dorsal part of right S1. TMS stimulation was applied over the right somatosensory (S1) vocal fold representation adjunct to motor representation (M1) [54, -4, 38; Tailarach] as report by Brown et al. (2009) in a meta analysis for vocal control. Tailarach coordinates were

recomputed for each individual brain. Sham stimulation was applied over the most dorsal part of right S1 close to the vertex. Stimulator output was fixed 40 % of maximum stimulator output for both conditions (larynx, dorsal) and all participants. Coil orientation was anterior-posterior (AP) in both cases and perpendicular to the post-central gyrus for the iTBS stimulation of the vocal fold representation but only 45 degree tilted for the sham stimulation.

7.1.5 Procedure

Subjects who did not participate in study 1, first did the perceptual screening before they performed the pitch reproduction task prior and post TMS intervention on two separate days, while a minimum of 48 hours was kept between both sessions. At the beginning each session subjects were equipped with noise cancelling headphones (Bose QuiteComfort 15; Bose, Germany) and a head mounted microphone(AKG C 477 WR L/p; AKG, Austria). Noise level was set to maximum tolerable level and subjects trained the paradigm with 2-5 trials depending on individual needs. During the singing task the investigator left the room to not create an uncomfortable situation which might influence performance. Subjects were naive to what effect the TMS stimulation should have on their singing performance.

7.1.6 Audio Data Analysis and Statistics

Audio data analysis and statistical interference had been computed using YIN under Matlab and R. Details about analysis an statistical methods are presented in chapter 5.

7.2 Results

7.2.1 Pitch discrimination task

The average frequency perception difference in the pitch discrimination task was 0.7 % with a range of 0.2 % (very good) to 1.25 % (good). The cut-off criterion for participation in the TMS experiment was 2 % (35 Cent). Within this range there is no effect on pitch production to expect (Kleber et al., 2013). All subjects fulfilled the inclusion criteria and continued the study at this point. Detailed information about the discrimination thresholds for each participant in $\Delta\%$ and in ΔCent differences can be found in appendix A.

7.2.2 Pitch reproduction task

Checking for normal singing performance

Initially, I investigated pitch reproduction performance of each subject to ensure all participants had stable and normal pitch reproduction abilities. For this purpose I looked at the variance of the final pitch accuracy measure of first tone of baseline performance only, using box-whisker-plots. As the first tone (311.11 Hz = D#4 for men and 622.22 Hz = D#5 for women, respectively) is kept equal for all intervals, variance should be low, if singing performance is normal. I expect subjects to habituate to the recurrent stimulus. Moreover, target tone was in a comfortable voice range. Figure 7.1 shows box-whisker-plots for final pitch accuracy performance of all subjects. It can be seen at figure 7.1 that subjects ME03 and NB63 had a highly unstable performance compared to the remaining subjects. The red line

marks the one standard deviation boundary (1 SD) above the mean, which was defined as cut-off criteria for inclusion in the further analysis procedure. Hence, I removed subjects ME03 and NB63 from the study as the median of their performance was above threshold. This has been done to prevent any bias due to this unstable performances.

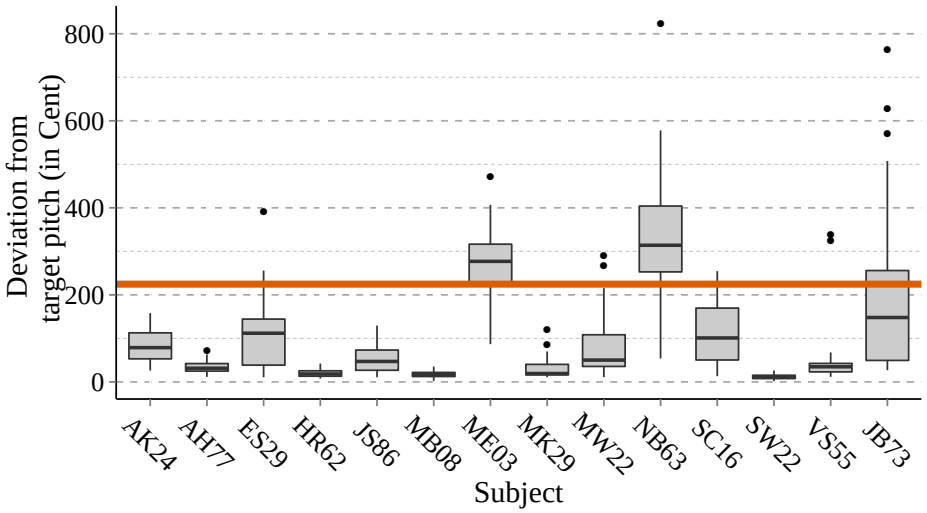


Figure 7.1: Plot shows box-whisker-plot for final pitch reproduction accuracy for the first tone of the baseline for each subject. Y-axis represents the deviation from target pitch and x-axis shows the participant code for each individual subject. Red line marks the defined cut-off for the exclusion criteria, which is one standard deviation above mean group performance.

Combining masking pre-TMS performances

Equal to the procedure in the first experiment I controlled for performance difference between both sessions in the pre-TMS measure. Again, I found no significant differences between both sessions. Results are shown in table 7.1. Consequently, both pre-TMS recordings have been averaged to receive one single baseline measure for later comparison.

Table 7.2
Comparing pre-TMS performances

Variable	Results		
	df	χ^2	p-value
Initial pitch accuracy	2	3.25	0.20
Final pitch accuracy	2	0	1
Initial pitch stability	2	1	0.61
Final pitch stability	2	0.75	0.69

Table 7.1: Friedman's ANOVA test results for comparison between pre-TMS pitch reproduction performances (with masking) of both sessions. $N = 12$.

Effects of auditory masking

In order to demonstrate that pink noise is efficiently masking the auditory feedback, I compared pitch reproduction accuracy and stability with auditory masking (mask) and without pink noise interference (no-mask), but without any TMS interference. Results demonstrated that pink noise masking impairs the auditory feedback and leads to less accurate singing performance. However, changes in pitch accuracy and pitch stability are only apparent for final pitch measures, but not for initial pitch measures (see figure 7.2, figure 7.3 and table 7.2). Possible reasons for this difference are discussed later.

Table 7.2
Effects of auditory masking

Variable	mask vs. no-mask	
	z-value	p-value
Initial pitch accuracy	1.16	0.124
Final pitch accuracy	2.40	0.016*
Initial pitch stability	-0.89	0.813
Final pitch stability	2.76	0.006**

Table 7.2: Wilcoxon sign rank test results (one-sided) for comparison of pre TMS mask and no-mask singing performance. $N = 11$. All p-values have been corrected for multiple comparison using the FDR-method.
Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

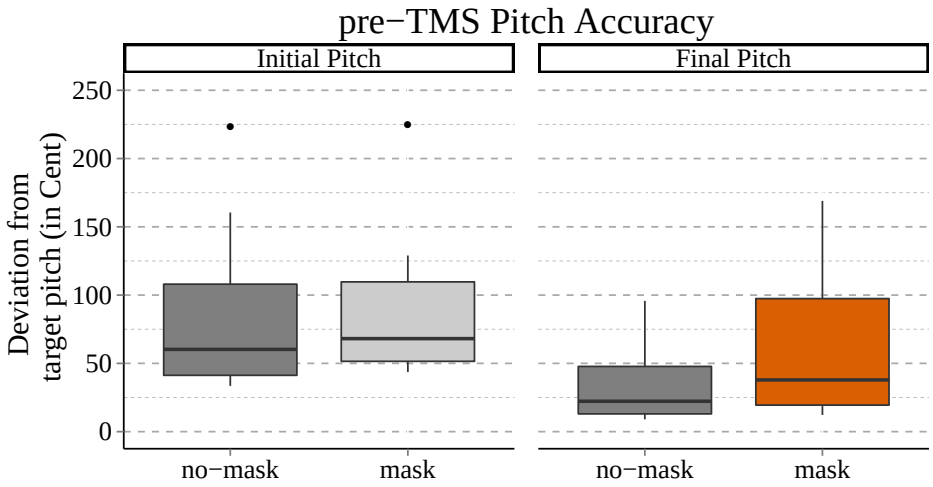


Figure 7.2: Box and Whisker plots display deviation of sung tones from target pitch. $N = 11$, where N is the number of subjects. Orange coloured box plots signalise a significant higher degree of deviation from target pitch (decreased accuracy). Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

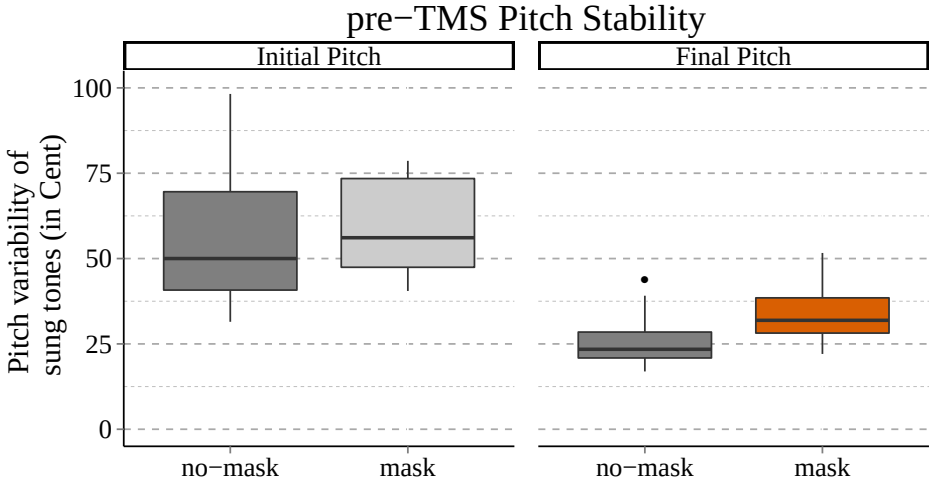


Figure 7.3: Box and Whisker plots display variation around sung tones from target pitch. $N = 11$, where N is the number of subjects. Orange coloured box plots signalise a significant higher degree of variation around target pitch (decreased stability). Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Effects of TMS on Pitch Accuracy

For determining the effect of iTBS stimulation on pitch reproduction performance I compared the pitch accuracy variables for the different condition (*larynx* and *dorsal*)s. Figure 7.4 shows the results for the pitch accuracy measure for the initial and final pitch component for all three conditions as box-and-whisker plot. It can be seen, that the initial pitch performance is less accurate in all conditions compared to the final pitch performance, while the average group-performance for all conditions is below one semitone.

However, these plots do not give a clear idea of a possible effects of TMS. Therefore, figure 7.4 shows the differences between pre-TMS and post-TMS as 'change box-plots'. Here, one can see the effect of TMS and the dependent variables. A change into negative direction marks an improvement after stimulation, while 0 reflects the point without any change between pre-TMS and post-TMS. Each box plot represents the post-TMS *'minus'* pre-TMS difference. Significant changes have been highlighted in green.

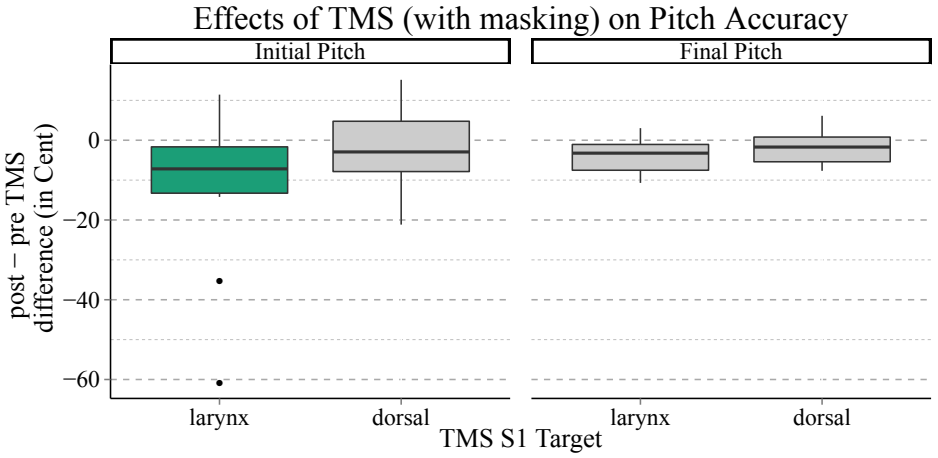


Figure 7.4: Box-and-whisker plots, displaying the effect of S1-larynx-iTBS and S1-dorsal-iTBS (post – pre TMS differences) on singing accuracy when auditory feedback was masked with pink noise. $N = 12$. Green colour indicates significantly increased pitch accuracy. Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Comparing the accuracy measure reveals a significant effect for initial pitch, χ^2 (2, $N = 12$) = 6.17, $p < .05$*, but not for the final pitch component, χ^2 (2, $N = 12$) = 3.87, $p = .14$, using Friedman's ANOVA. Table 7.3 shows results for the post-hoc comparison of initial pitch accuracy.

Table 7.3
Pairwise post-hoc comparison

	Initial pitch accuracy		
	Z	p-value	r
S1-larynx TMS vs. pre TMS	-2.35	0.028*	0.48
S1-dorsal TMS vs. pre TMS	-0.08	0.479	-
S1-larynx TMS vs. S1-dorsal TMS	-2.00	0.034*	0.41

Table 7.3: Wilcoxon sign rank test (one-sided) has been used for post-hoc testing iTBS effects on initial pitch accuracy. All results are corrected for multiple comparisons using the FDR-method. $N = 12$.
Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Pairwise post-hoc comparison results demonstrated a significant effect of iTBS on larynx area, but not for iTBS on dorsal control area compared to baseline pre-TMS performance and also in direct comparison. This results was underlined by a medium effect size.

Effects of TMS on Pitch Stability

To investigate how well subjects kept the pitch production steady I compared the pitch stability measure for the initial and final component of each condition. Figure 7.5 depicts the change box-and-whisker plots for the corresponding pitch stability results.

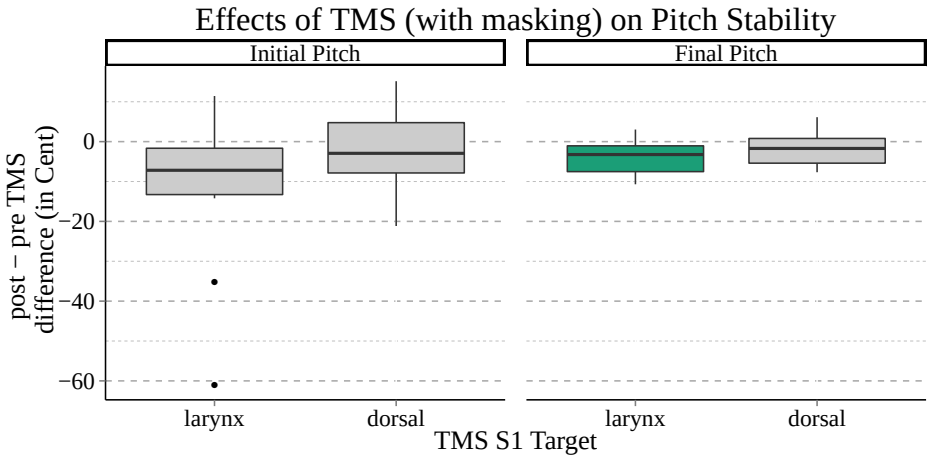


Figure 7.5: Box-and-whisker plots, displays the effect of S1-larynx-iTBS and S1-dorsal-iTBS (post – pre TMS differences) on singing stability when auditory feedback was masked with pink noise. $N = 12$. Green colour indicates significantly increased pitch stability. Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Friedmans's ANOVA revealed a significant main effect for final pitch stability, χ^2 (2, $N = 12$) = 7.17, $p < .05^*$, but not for initial pitch stability, χ^2 (2, $N = 13$) = 5.17, $p = .08$. Post-hoc comparisons revealed that this effect was due to target (S1-larynx) but not control (S1-dorsal) stimulation (see table 7.4).

Table 7.4
Pairwise post-hoc comparison

	Final pitch stability		
	Z	p-value	r
S1-larynx TMS vs. pre TMS	-2.51	0.018*	0.51
S1-dorsal TMS vs. pre TMS	-0.08	0.091	-
S1-larynx TMS vs. S1-dorsal TMS	-2.04	0.031*	0.42

Table 7.4: Wilcoxon sign rank test (one-sided) has been used for post-hoc testing effects of iTBS on final pitch stability. All results are corrected for multiple comparisons using fdr method. $N = 12$.
Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Effects of TMS on Interval Accuracy

So far, only measures had been taken into account which averages both sound tones to judge the changes in pitch reproduction accuracy. Furthermore, I also compared the interval accuracy for initial and final pitch component. However, I found no significant effect, neither for the initial pitch χ^2 (3, N = 12) = 3.17, p = .21, nor for the final pitch interval accuracy, χ^2 (3, N = 12) = 1.67, p = .56 using Friedman's ANOVA.

Correlation Analysis

Finally, exploratory correlation analyses revealed a complex relationship between perceptual pitch-discrimination and pitch reproduction skills with respect to production changes due to masking and post-TMS effects.

Pitch accuracy improvement vs. baseline performance

I was interested if the improvement in singing performance after iTBS stimulation is correlated with participants' performance (pre-TMS). I found a significant negative correlation between post-TMS pitch accuracy improvement and pre-TMS pitch accuracy levels (initial pitch: $r = -.58, p < .05$*; final pitch: $r = -.66, p < .01$**). This suggests a tendency that participants with lower pitch reproduction accuracy prior to TMS also showed larger improvements after S1 larynx-iTBS stimulation in both initial and final pitch accuracy (Figure 4). Results for the correlation analysis are shown in table 7.5. Figure 7.9 and figure 7.10 show a point plot for the significant pitch measures.

Table 7.5
Correlation analysis results

Variable	r-value	p-value
Initial pitch accuracy	-0.58	0.025*
Final pitch accuracy	-0.66	0.009**
Initial pitch stability	-0.26	0.41
Final pitch stability	-0.39	0.21

Table 7.5: Pearson correlation has been computed between the pre-TMS performance and the improvement after iTBS-larynx-S1 stimulation. $N = 12$.
Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

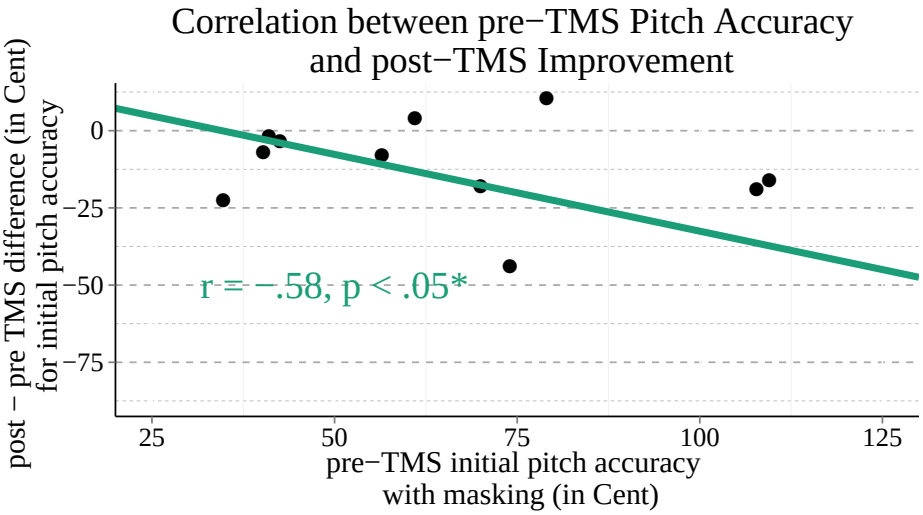


Figure 7.6: Point plot showing the relationship between pre-TMS deviation from initial target pitch and improvement after larynx-iTBS in initial pitch accuracy. $N = 12$. Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

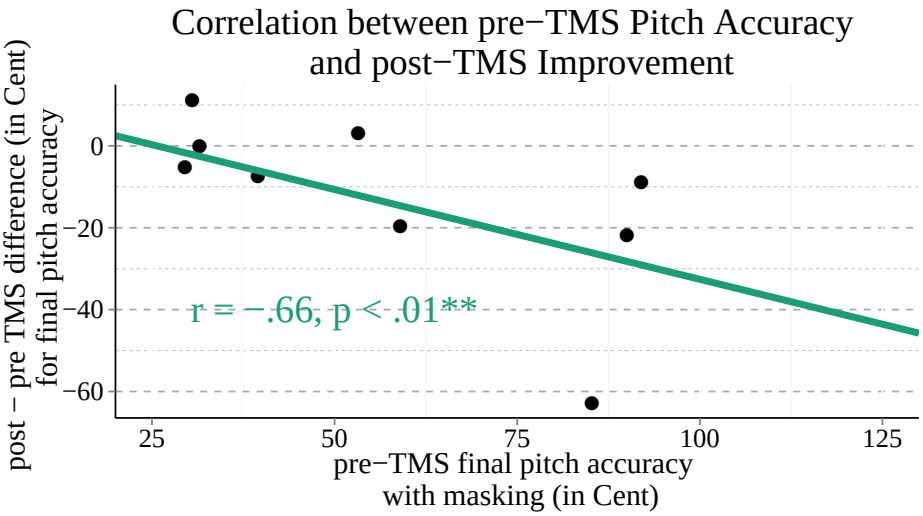


Figure 7.7: Point plot showing the relationship between pre-TMS deviation from final target pitch and improvement after larynx-iTBS in final pitch accuracy. $N = 12$. Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Pitch discrimination skills vs. masking effect and TMS

So far I could proof that pink noise is effectively impairing pitch production accuracy. An additional correlation analyses revealed a relationship between perceptual pitch discrimination and pitch reproduction skills with respect to production changes due to masking. That is, participants perceptual skills revealed a significant positive correlation ($r = .85$, $p < .001$***) with the strength of final pitch accuracy impairment due to masking (see figure 7.8). This indicates that participants with higher frequency discrimination thresholds (lower perceptual skills) also seem to be more affected by pink noise masking during singing.

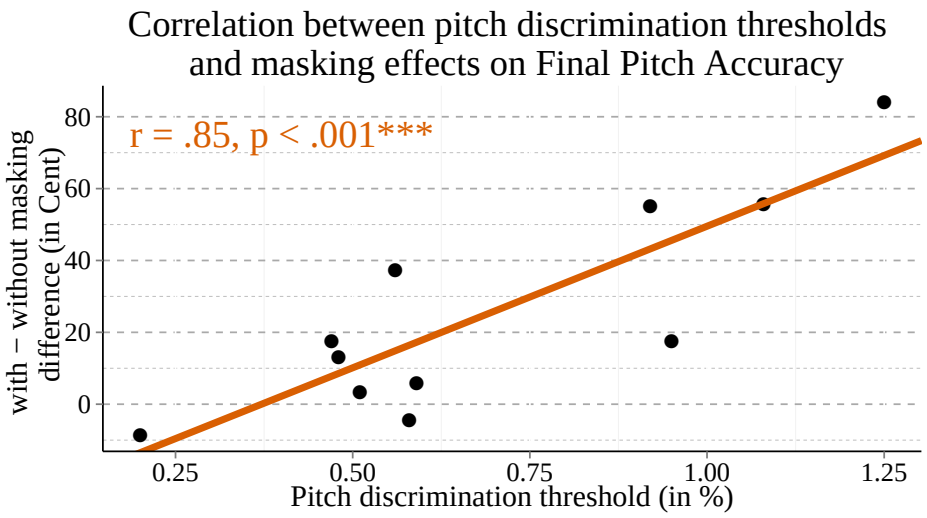


Figure 7.8: Point-plot shows a significant relationship between the effect of masking (with - without masking difference) on final pitch accuracy and subjects' pitch discrimination thresholds. $N = 12$.
Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

Furthermore, participants frequency discrimination thresholds were negatively correlated with the extent of pitch accuracy improvement for the initial ($r = -0.73$, $p < .01$**) and final ($r = -0.66$, $p < .05$*) components after S1-larynx iTBS (see figure 7.6 and figure 7.7). This effect suggests that participants with larger frequency discrimination thresholds have a higher potential to improve their performance after S1-larynx iTBS. Together, results from exploratory correlation analyses suggest a complex interaction between sensory feedback modalities (auditory and somatosensory) and vocal motor control depending on whether or not untrained singers could employ auditory monitoring strategies during singing.

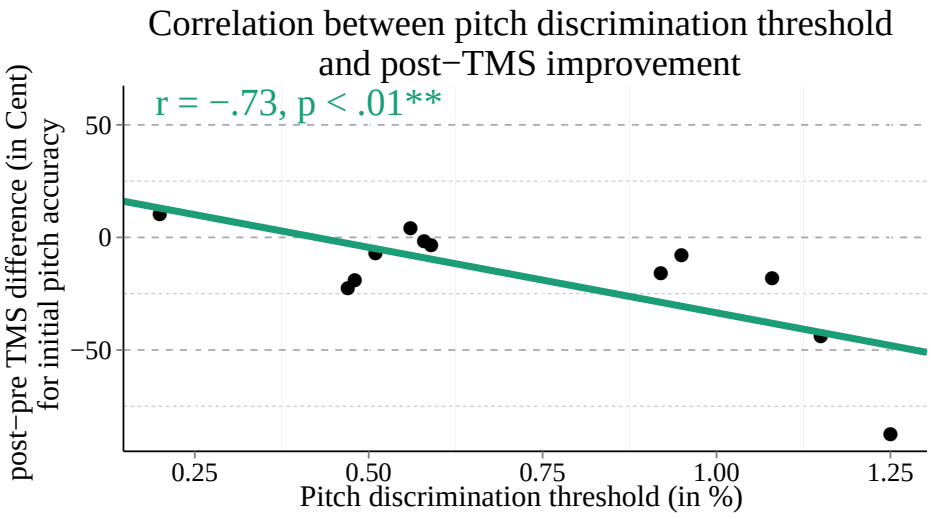


Figure 7.9: Point plot shows the relationship subjects' pitch discrimination thresholds and reproduction improvement for initial pitch accuracy after S1-larynx iTBS. $N = 12$.
Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

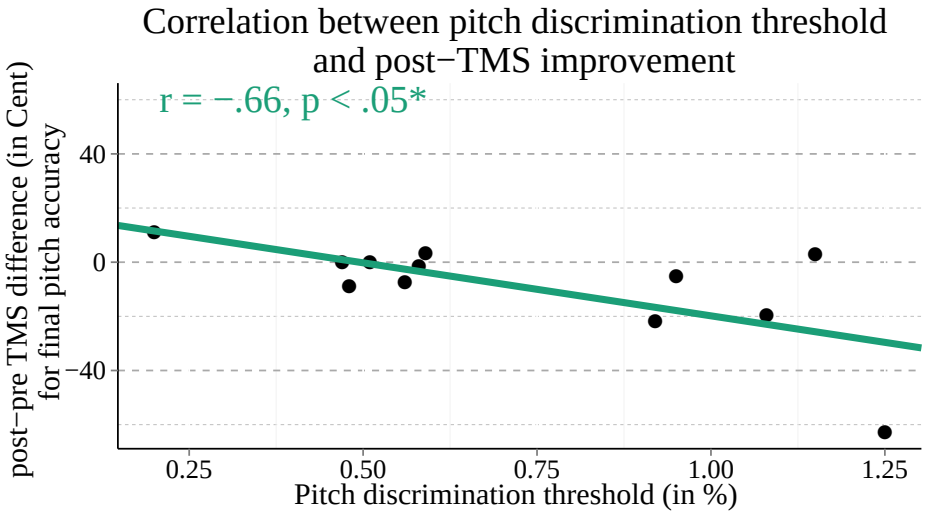


Figure 7.10: Point plot shows the relationship subjects' pitch discrimination thresholds and reproduction improvement for final pitch accuracy after S1-larynx iTBS. $N = 12$.
Significances: * = $p < .05$, ** = $p < .01$, *** = $p < .001$

7.3 Preliminary Conclusion

In order to prove the role of the right ventral primary somatosensory cortex for vocal control I used intermittent theta burst stimulation in a pitch reproduction paradigm testing non-singers while their auditory feedback was masked using pink noise. I hypothesised that iTBS improves pitch performances measures compared to a neutral baseline and stimulation on a control site.

I could demonstrated that auditory masking with pink noise effectively impairs auditory perception during a singing task, leading to a significant reduction in pitch reproduction accuracy and stability. These findings are in line with the current literature (Zarate and Zatorre 2008; Mürbe et al. 2002). However, additionally I found a clear difference in the masking effect between early and late pitch components. Only the late pitch component was effected, which suggests that initial an final pitch might be controlled by different feedback systems.

In the main experiment I found, that larynx-iTBS significantly improved singing, while TMS over a dorsal control area did not effect performance. Moreover, pitch accuracy was enhanced for the initial pitch, but stability improved for final pitch. Further, I found several interesting correlations showing that subjects with less accurate baseline performance improved more after stimulation. In addition, I found a strong relationship between hearing abilities and the effect of larynx-iTBS on singing, which might suggest an interaction between auditory and somatosensory abilities.

Overall, the results underlines to involvement of the primary somatosensory system in accurate singing in the absent of auditory feedback. However, stimulation did not effect singing performance an all measured parameters,

which will be further explained in the discussion section.

CHAPTER 8

Discussion

8.1 Summary

Several studies previously pointed towards an increased importance of somatosensory feelings from the vocal tract for the control of complex motor tasks such as singing. Kleber et al. (2010), for example, found that trained singers as compared to non-singers showed increased activation during singing within primary somatosensory cortex, a finding that confirmed behavioural results suggesting that the kinesthetic feedback loop becomes increasingly important with the development of singing skills (Mürbe et al. 2002, 2004; Jones and Keough 2008). However, as these studies measured predominantly correlations between vocal skill level and respective dependent variables they do not allow making inferences regarding causality. In the presented work I investigated the role of somatosensation in pitch motor-control with theta burst stimulation, applied to right laryngeal primary somatosensory cortex and dorsal S1 respectively

I used a neuro-navigated system to apply iTBS and cTBS over primary laryngeal somatosensory cortex to probe the role of somatosensory feedback in vocal motor control. In two related experiments, effects of TBS on pitch-production (accuracy and stability) were assessed in a pitch-matching paradigm with untrained singers. In experiment I, participants were able to hear their own voice while they sang musical intervals. In experiment II, auditory feedback was masked with pink noise during singing. Analogous to previous studies (Zarate and Zatorre 2008; Mürbe et al. 2002), the data of the second experiment confirmed a significant effect of auditory masking on singing accuracy in general. Moreover, my main results show that with normal auditory feedback, TMS showed no effect on vocal performance. However, when auditory feedback was masked with noise, iTBS significantly improved

initial pitch accuracy and final pitch stability after larynx S1 but not after dorsal S1 stimulation. Interestingly, positive effects of larynx S1 stimulation on pitch accuracy (i.e., initial and final pitch components) were pronounced in participants with lower pitch-matching skills prior to iTBS. Furthermore, participants with lower perceptual skills (i.e., higher pitch-discrimination thresholds) were also more adversely affected when auditory feedback was masked and improved more after larynx-S1 stimulation. These novel results lend further support to the idea that somatosensory feedback from the vocal tract plays a greater role in fine vocal-motor control than previously thought (Tremblay et al. 2003; Mürbe et al. 2004; Nasir and Ostry 2006; Kleber et al. 2010; Lametti et al. 2012; Kleber et al. 2013). To my knowledge, this provides first evidence for a causal relationship between neural facilitation of larynx-S1 and temporary improvement of pitch motor-control in musically untrained participants.

8.2 Mechanisms of vocal pitch motor control

One of the perceptually most salient features of vocal production is pitch, the fundamental frequency of sound. It is determined by the vibration frequency of the vocal folds and controlled via vocal fold tension and length in interaction with air-pressure from the lungs (Titze 1993). Pitch relationships are relative and continuous in speech whereas in music, they are discrete and follow precise interval relations, where listeners would perceive even small deviations from target frequencies as error (Natke et al. 2003; Zatorre and Baum 2012). This poses higher demands on vocal-motor control and is one reason why singing errors are mostly related to inaccuracies in pitch

production: 50 % of untrained singers fail to match a target pitch to within 50 cents (100 cents = one semitone) deviation (Dalla Bella et al. 2007). In contrast, trained singers show significantly higher levels of pitch-matching accuracy, which points towards an experience dependent improvement of vocal motor control (Nikjeh et al. 2009; Hutchins and Peretz 2011, 2012). Vocal sensorimotor interactions are commonly examined by compensatory responses to auditory perturbations, such as a pitch-shifting reflex in the opposite direction to pitch change in auditory feedback (Jones and Munhall 2000; Burnett and Larson 2002; Zarate and Zatorre 2008) or decreased pitch-stability and accuracy when auditory feedback is masked with noise (Ternstrom et al. 1988; Mürbe et al. 2002). According to these data, communication between auditory feedback processing and vocal motor-control supports correct pitch-production (Zarate et al. 2010). Moreover, positive correlations between pitch-discrimination and pitch-performance accuracy imply that auditory perception may be the main tool for controlling vocal production (Amir et al. 2003; Watts et al. 2005; Moore et al. 2010; Estis et al. 2011). In contrast, studies have also shown that even extremely inaccurate singers may possess good pitch perception (Bradshaw and McHenry 2005; Estis et al. 2011), while training of pitch-discrimination skills leaves pitch-reproduction accuracy unaffected (Zarate et al. 2010). Further evidence for other mechanisms being involved in vocal control stems from research in congenital amusia, a neuro-degenerative disorder estimated to affect about 4 % of the population. Amusia is associated with impaired music perception despite normal cognitive and hearing abilities (Peretz et al. 2007; Stewart 2011). Congenital amusics are unable to detect pitch differences less than a semitone (Hyde and Peretz 2004) or distinguish between rising and falling pitches (Foxton et al. 2004). This deficit is typically reported to coincide

with impaired singing abilities (Ayotte et al. 2002), such as increased pitch errors in familiar tunes (Dalla Bella et al. 2009; Tremblay-Champoux et al. 2010) and single tones compared to non-amusics (Hutchins and Peretz 2011). Yet the existence of production-preserved congenital amusics, in which the conscious perception of pitch is independent from vocal action, suggests a possible dissociation between auditory perceptual and vocal production skills (Dalla Bella et al. 2009; Loui et al. 2009; Hutchins et al. 2010; Hutchins and Peretz 2013).

Research in vocal sensorimotor interactions often neglects that auditory feedback becomes highly correlated with somatosensation during production. Indeed, growing evidence suggests that somatosensory feedback plays a larger role in vocal motor control than commonly thought (Tremblay et al. 2003; Mürbe et al. 2004; Nasir and Ostry 2006, 2008; Kleber et al. 2010; Golfinopoulos et al. 2011; Lametti et al. 2012; Kleber et al. 2013). When comparing pitch-shift paradigms with somatosensory perturbations of jaw movements, both sensory perturbations trigger compensatory responses yet auditory feedback is not dominant for all subjects (Lametti et al. 2012). Even post-lingually deaf patients show accurate compensation to somatosensory perturbation regardless if their cochlea implant was switched on or off (Nasir and Ostry 2008). Mürbe and colleagues (2004) found that in contrast to auditory feedback, kinesthesia contributes increasingly to pitch-control with singing experience, which may explain why trained singers are better at ignoring pitch-shifts (Zarate and Zatorre 2008) and compensate less compared to non-singers (Jones and Keough 2008). This experience-dependent role of kinesthesia corresponds to neuro-imaging results, revealing increased activation of laryngeal primary somatosensory cortex in opera singers compared to non-singers when singing an Italian aria (Kleber et al. 2010).

8.3 Effects of iTBS on vocal motor control

Intermittent theta burst stimulation is a form of transcranial magnetic stimulation that requires only short stimulation time and produces long lasting changes in cortico-spinal excitability based on long-term potentiation-like effects on cortical synapses (Huang et al. 2011). Originally introduced by Huang and colleagues (2005) iTBS related changes had been referred to long-term potentiation like changes and the dependency on NMDA receptors (Huang et al. 2007). Studies have demonstrated that iTBS over primary somatosensory cortex improves vibrotactile (Morley et al. 2007) and two-point discrimination skills (Ragert et al. 2008), equivalent to improvements shown with a 5 Hz repetitive TMS paradigm (Tegenthoff et al. 2005). iTBS over S1 also facilitates somatosensory evoked potentials, with a peak effect found about 15 minutes after stimulation (Katayama and Rothwell 2007).

In this thesis, I applied cTBS and iTBS over the right larynx S1 and found improved pitch reproduction accuracy and pitch stability when auditory feedback from the own voice was masked with noise but not when auditory feedback was available. A possible explanation for this difference may be derived from a well-known neural network model of vocal motor skill acquisition and production (Directions Into Velocities of Articulators, DIVA) introduced by Guenther (Guenther et al. 2006). The DIVA model proposes that neural mappings for vocal tract movements are initially tuned by auditory feedback from self-generated sounds to learn the relationship between motor actions and their acoustic consequences (Guenther and Vladusich 2012). Once the relationship is established, activation of a speech sound map (corresponding to a small population of neurons in pre-motor cortex) leads to motor commands from primary motor cortex, which is mediated by both sensory

feedback and predictive feedforward mechanisms. Auditory targets for new sounds become encoded in the neural projections of the sound map to an auditory error map within auditory cortex, whereas feedforward predictions are tuned by employing the acoustic feedback to transform auditory errors in corrective motor commands until the model is sufficiently refined to produce sounds without error (see figure 8.1). The repeated production, however, eventually leads to the creation of somatosensory targets, which become so strongly correlated with the acoustic target (Golfinopoulos et al. 2010, 2011) that they may even be sufficient to match the precision requirements for vocal production in the absence of sound (Nasir and Ostry 2008).

Although both somatosensory and auditory feedback likely act together for vocal motor-control, the creation of precise somatosensory targets may require more time for development (Mürbe et al. 2004). Lametti and colleagues (2012) found that subjects who failed to adapt to auditory perturbation adapted to somatosensory perturbation and vice versa. They proposed that individuals might be more sensitive to a particular sensory modality depending on their individual sensory experience, whereas singing experience has been associated with enhancement of the somatosensory feedback system (Kleber et al. 2010, 2013). In my experiments, participants were inexperienced singers and may not have developed somatosensory pitch-targets. It is likely that they adopted a predominantly auditory error-detection strategy, which might explain why iTBS was ineffective as long as auditory feedback was available whereas iTBS was more effective when masking forced their attention towards somatosensory feedback.

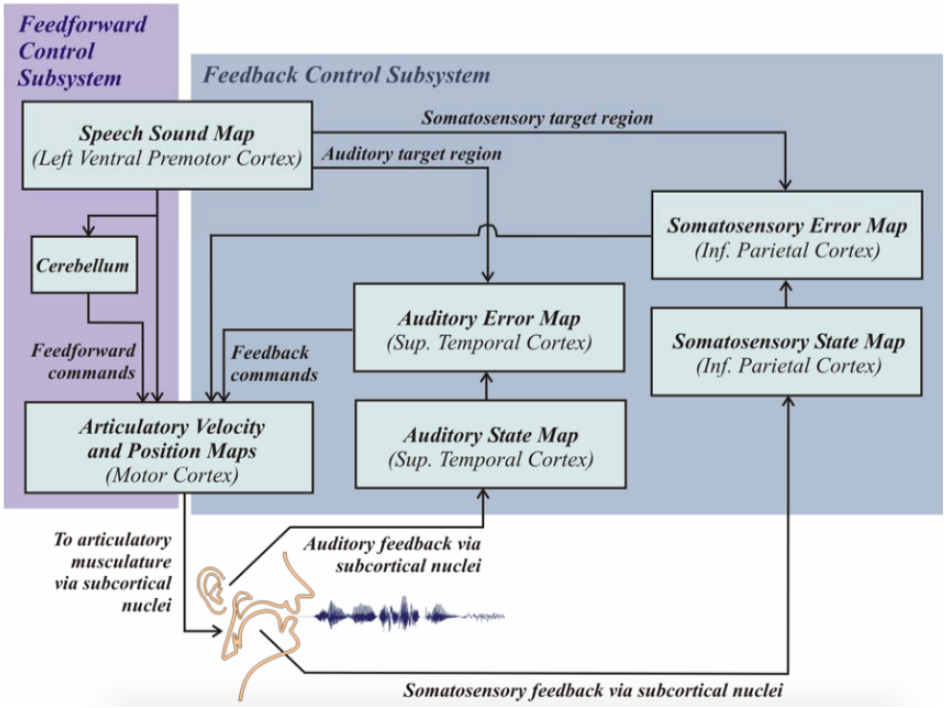


Figure 8.1: Schematic display of the neural precessing stages as hypothezised by the DIVA model. Image has been taken from Guenther et al. (2006).

A possible explanation can be derived from a model for vocal motor learning introduced by Guenther et al. (2006). Based on the DIVA model, auditory feedback will be dominant in non-singers as their vocal control system for singing is relatively low developed (compared for example to speech). Only with training, when people become professional singers, the somatosensory feedback gains importance (Mürbe et al. 2004; Kleber et al. 2013; Jones and Keough 2008). I hypothesise that non-singers primarily rely on an auditory feedback strategy in a singing task as long as this feedback channel is available. If subjects solely have to rely on somatosensory feedback, TMS can influence singing performance. This is in line with my results showing that iTBS improves pitch reproduction accuracy and stability only when

auditory feedback is disabled. However, one must also take the difference between early and late aspects of pitch-production into account.

8.4 Early versus late effects of iTBS on pitch performance

Perturbations of auditory feedback typically show auditory-motor responses in the opposite direction, indicating a closed-loop feedback system that plays an important role in phonatory control during singing and speaking (Larson et al. 2000; Natke et al. 2003). These responses include a small magnitude short latency ($\tilde{5}$0-150 ms after auditory stimulation) and larger magnitude long latency (350 ms) component. The long latency component is thought to show more voluntary control by engaging transcortical pathways whereas the early component may be mediated by subcortical structures, reflecting lower-level automatic mechanisms of pitch control during phonation to stabilise voice pitch around an intentional pitch target (Sapir et al. 1983; Burnett et al. 1998; Hain et al. 2000; Burnett and Larson 2002; Grell et al. 2009). Interestingly, also the short latency component is modulated by intentions, depending on the instructions given prior to testing (Hain et al. 2000).
Prior work has emphasised a critical role of kinaesthesia in early pitch-control. Mechanical perturbations of the larynx can elicit long and short latency vocal-responses (Sapir et al. 2000; Loucks et al. 2005) whereas pitch-shift response latencies are typically prolonged when laryngeal mechanoreceptors in the mucosa are anaesthetized (Grell et al. 2009). Larson and colleagues (2008) investigated pitch-shift responses with and without laryngeal anaesthesia.

Based on data and mathematical models they concluded that kinaesthetic error acts with a shorter latency (≈ 20 ms) compared to auditory error, which implies that somatosensory feedback is weighted more heavily during the initial part of voice production whereas after ≈ 100 ms, auditory feedback also participates. Therefore, iTBS may also affects the more automatically controlled short latency pitch component. The finding that iTBS improved initial pitch accuracy when auditory feedback was masked is consistent with a vocal motor control system that monitors both auditory and somatosensory feedback (Golfinopoulos et al. 2010; Hickok 2012; Katseff et al. 2012), yet this effect suggests that the relative weighting of sensory feedback may not only be driven by training (Kleber et al. 2010, 2013) or sensory preference (Lametti et al. 2012) but also by attention (Larson et al. 2008). A change in feedback-strategy could thus account for a higher gain of kinaesthetic feedback when auditory control is no longer available, which may explain why effects of iTBS were only seen under masking.

Unlike auditory feedback, somatosensory feedback is also available for pre-phonatory tuning, that is the adjustment of vocal fold position, length, and tension prior to voice onset (Wyke 1974b; Loucks et al. 2005). Accurate pre-phonatory tuning is typical for trained singers (Murry 1990) but is also found in subjects without singing experience (Watts et al. 2003). Amorosa and colleagues (1986) demonstrated that children with unintelligible speech show significantly more signs of abnormal pre-phonatory tuning and phonatory modulation than control children. This suggests that iTBS could improve initial pitch accuracy also by facilitating pre-phonatory tuning whereas the former implies that somatosensory integration may be a pre-requisite for singing skill independent from training.

In contrast, why was final pitch accuracy unaffected despite a non-significant

trend towards improvement? As mentioned before, the late pitch-response latency reflects a cortically mediated conscious process, whereas the early latency represents automatic and less consciously controlled processes (Hain et al. 2001). The laryngeal primary motor cortex of humans possesses direct connections with the brainstem laryngeal motor-neurons, which is one of the major evolutionary developments in humans toward the ability to speak and vocalise voluntarily and represents the human advantage to vocal learning (Simonyan and Horwitz 2011). Trained singers appear to have better voluntary control over the late pitch-response (Zarate and Zatorre 2008) and show greater involvement of primary sensorimotor areas during singing (Kleber et al. 2010, 2013). It might be that non-singers do not possess somatosensory pitch memory due to the lack of singing experience and therefore no somatosensory error detection, which is why S1 facilitation has no strong effect on final pitch control even in the absence of hearing. However, this speak against the finding, that final pitch stability is significantly improved after S1-larynx stimulation. Pitch stability refers to fluctuation about the median of the extracted fundamental frequency. A stable pitch production contains less variance and indicates more even vocal fold oscillations (Titze 1988, 2008). Lower pitch stability is a typical characteristic of the adolescent speaking voice (Boltez?ar et al. 1997) and improves during the course of speech- motor development (Smith 2006). As pitch stability may therefore have been extensively trained in speech tasks, I suggest that larynx-S1 stimulation enhanced this feature of phonation control by improving the mechanisms of vocal production (Wyke 1974a,b).

8.5 Interactions with auditory feedback

Only non-singers with good-to-average pitch-discrimination skills participated in this study to rule out that any perceptual deficit could be accountable for deficits in singing performance. Unexpectedly, however, I found significant interactions between auditory perceptual skills and masking and iTBS results respectively. One the one hand, participants with lower perceptual skills (i.e. higher discrimination thresholds) showed also lower pitch accuracy scores during masked singing and, on the other hand, improved more in initial and final pitch accuracy after iTBS larynx-S1 stimulation. It is not uncommon that participants with lower entry-level skills also benefit more from interventions aiming at improving them (Ladda et al. 2014). Nevertheless, these results are indicative of a more inaccurate tuning of somatosensory-motor mappings due to lower auditory perceptual skills as suggested by the DIVA model (Guenther et al. 2006). Alternatively, the opposite route is also conceivable since lower production accuracy could shape less precise auditory perception based on the observation that action and perception are inherently interwoven (Hickok et al. 2011). The latter has recently also been impressively demonstrated for the somatosensory domain (Ito and Ostry 2010, 2012).

Taken together, these results point to the idea, that dependent relationship exist between pitch discrimination skills and pitch-matching accuracy, and this relationship may help to explain why some persons, without training, have an accurate singing voice, whereas others do not (Watts et al. 2005). To sing accurately, then, a person most likely must have both accurate perceptual abilities and precise vocal motor control (Estis et al. 2011). My data provides strong evidence for the hypothesis that vocal accuracy in singing (and more

generally in vocal production) is constrained by the somatosensory feedback system and only partially dependent on the auditory domain. In summary, I speculate that facilitation of somatosensory processing leads to general improvements in pitch accuracy when dominant auditory feedback is not accessible, which emphasizes the role of kinesthesia in pitch control.

8.6 Limitations and Outlook

I think that this study added an important new components to the understanding of the human voice and its control systems. However, there are still open questions my experiments could not answer, partly due to the limitations I had to adhere based on the experimental setup. In both experiments I could only focus on one hemisphere. This was due to technical reason but also methodological uncertainties. On the one hand it was impossible to stimulate both hemisphere at the same time, as the system allowed only the stimulation with one coil at a time. On the other hand, there are no experimental experiences with simultaneous TBS stimulation on two sites. Therefore, I decided to stimulate only the right hemisphere. One reason for this decision was, that other areas (e.g. right auditory cortex) of this hemisphere are reported as being more important in fine pitch discrimination (Zatorre et al. 2002; Hyde et al. 2008) and motor control in singing (Alcock et al. 2000; Riecker et al. 2000; Kleber et al. 2013). In my opinion, this makes it also more likely that stimulation of the right hemispheric somatosensory representation of the laryngeal muscles has an effect on precise pitch control for singing. Nevertheless, I think future work might focus on the inter-hemispheric differences to further disentangle the relationship of

auditory and somatosensory feedback in voice control. It might be that, comparable to the auditory system and motor system, the somatosensory feedback system is different for each hemisphere, too. This might show up in a similar differentiation as in the auditory system, with the left hemisphere more specialised for time related aspects of a signal and the right hemisphere more specialised for pitch related information.

The extensive connections between laryngeal S1 and M1 as well as S1 and parietal areas suggest that S1 stimulation also activated the motor network via involvement of M1. Secondly, I evaluated behavioural effects in response to a control area but not in response to an area that is actively involved in vocal motor control (e.g., inferior parietal lobe). Therefore, a double-dissociation was not possible, which would have strengthened my claim. Additional methods such as fMRI pre- and post-TMS are required to confirm assumed network-effects with respect to interhemispheric specialization lateralization and the involvement of other areas of the singing network due to stimulation. Although the effect sizes are moderate to good, a larger sample could clarify if missing significances for final pitch accuracy and initial pitch stability are due to the limited number of subjects.

Further, I would suggest to include professional singers of which one knows that they have a stronger somatosensory representation (Kleber et al. 2010) and rely more in kinaesthetic feedback (Mürbe et al. 2004) in future studies. A system which relies more on somatosensory information should also react stronger to changes in this modality. Therefore, I would expect that a cTBS stimulation might have a deteriorating effect on professional singers, while their auditory feedback is masked, similar to the effect Kleber et al. (2013) could demonstrated in anaesthetising the vocal folds. For arguable reasons (see above), one could not see a effect of cTBS on non-singers. However,

a worsening effect of cTBS would even more strongly support my theory that the S1 representation of the larynx is causally involved in precise vocal motor control. I believe that deploying a inhibitory TMS protocol could lead to impairment of singing performance in processional singers and therefore would further establish the role of the primary somatosensory laryngeal muscle representation for complex vocal motor control feedback.

Additionally, I suggest to clarify the relationship between pitch discrimination and pitch production performance. My results suggest a strong coupling between auditory perceptual skills and pitch reproduction skills. It might be interesting to investigate further possible cross-modal effects. One could examine if stimulation on the laryngeal S1 area also leads to changes in auditory skills, e.g. pitch discrimination. Recent studies support the idea of a bidirectional cross-modal relationship between auditory and somatosensory areas (Leveque et al. 2013).

Finally, I would propose to investigate in detail the short term functional changes caused by the TMS stimulation using a fMRI design. Using a similar setup which has been used in both presented experiments, however, conducting the singing task inside an MRI scanner could yield information about the functional and connectivity changes caused by TMS (Andoh and Zatorre 2011). Of special interest would be the exact location of the changes correlated with improved singing performance.

8.7 Conclusion

Here I report first evidence for a causal relationship between neural facilitation of laryngeal somatosensory cortex and behavioural effects on fine vocal motor-control. Based on this data I propose that kinaesthetic feedback is more fundamentally involved in respiratory-laryngeal motor coordination during vocal pitch-control than acknowledged by current models of speech motor control (Guenther et al. 2006; Hickok et al. 2011; Houde and Nagarajan 2011; Tian and Poeppel 2010). These findings might help to further theories on vocal pitch control in the general population (Hutchins and Peretz 2013) and production preserved congenital amusics (Dalla Bella et al. 2009; Hutchins et al. 2010; Hutchins and Peretz 2012; Loui et al. 2009). Additional studies are required to test TMS effects on interhemispheric connectivity (Andoh and Zatorre 2011) and on the neural networks underlying vocal motor control.

List of Figures

List of Tables

Appendices

A - Additional Data and Statistical Results

Table 8.1
Results of MEP measuring
(Experiment I)

Subject	Stimulator output	
	AMT	80 % AMT
AK24	55	44
HK05	56	46
MB08	46	38
ME03	52	42
MK29	52	42
MW03	48	40
SL15	51	42
SP27	48	40
TD19	48	40
group mean	**50.67**	**41.56**

Table 8.1: Results display the lower bound threshold when minimum 5 out of 10 MEPs had a magnitude of at leas 100 μV (AMT) after applying single pulse TMS to the left FDI muscle representation (left hemisphere and right hand). Later stimulation had been done with 80 % AMT as stated in previous studies (Huang et al. 2005). Result are from experiment I.

Table 8.2
Pitch discrimination task performance
(Experiment I)

Subject	Thresholds	
	Δ %	Δ Cent
AK24	1.08	18.59
HK05	0.47	10.18
MB08	0.58	10.01
ME03	0.43	07.43
MK29	0.20	03.46
MW03	0.48	08.29
SL15	0.54	16.37
SP27	0.58	08.17
TD19	0.17	08.81
group mean	**0.50**	**12.08**

Table 8.2: Pitch discrimination task thresholds for all subjects individually (experiment I). Results are displayed in % difference from tone one to tone two and in Cent. All subjects fulfilled inclusion criteria for participation in the study. Results are from experiment I.

Table 8.3
Peak fMRI activation for singing (Experiment I)

Subject	Coordinates		
	x	y	z
AK24	51	-7	44
HK05	54	-5	40
MB08	66	-25	44
ME03	66	-10	36
MK29	54	2	44
MW03	51	-7	48
SL15	51	-4	44
SP27	54	-7	44
TD19	54	-10	40
group mean	**55.7**	**-8.1**	**42.7**

Table 8.3: Above the main activation peak (right hemisphere) in MNI-coordinates are displayed for the singing fMRI localizer task in experiment I.

Table 8.5
Pitch discrimination task performance (Experiment II)

Subject	Thresholds	
	Δ %	Δ Cent
JB73	1.25	21.51
HR62	0.59	10.18
NB63	0.68	11.73
JS86	0.56	09.67
ES29	0.29	05.01
SC99	0.48	08.29
VS55	0.95	16.37
AH77	0.47	08.17
SW22	0.51	08.81
AK24	1.08	18.59
MK29	0.20	03.46
MW22	1.15	19.79
ME03	0.43	07.43
MB08	0.58	10.01
group mean	**0.70**	**12.08**

Table 8.5: Pitch discrimination task thresholds for all subjects individually. Results are displayed in % difference from tone one to tone two and in Cent. All subjects fulfilled inclusion criteria for participation in the study. Results are from experiment II.

Table 8.4
Demographic information
(Experiment I)

Subject	Sex	Age	EHI score
AK24	m	44	88
HK05	m	46	79
MB08	f	38	100
ME03	m	42	66
MK29	m	42	78
MW03	f	40	69
SL15	m	42	60
SP27	f	40	100
TD19	f	40	67
group mean	**m = 5 f = 4**	**23.67**	**78.56**

Table 8.4: Detailed demographic information including subjects' sex, age and Edinburgh Handiness Inventory (EHI) score. m = male, f = female. Data stems from experiment I.

Table 8.6
Demographic information
(Experiment II)

Subject	Sex	Age	EHI score
AK24	m	34	88
AH77	f	24	63
ES29	m	25	60
HR62	f	35	45
JB73	m	23	100
JS86	m	26	89
MB08	f	26	100
ME03	m	26	66
MK29	m	32	78
MW22	f	24	67
NB63	m	22	71
SC16	f	29	100
SW22	f	28	100
VS55	f	28	86
group mean	**m = 7**	**27.29**	**73.07**

Table 8.6: Detailed demographic information including subjects' sex, age and Edinburgh Handiness Inventory (EHI) score. m = male, f = female. Data stems from experiment II.

Table 8.7
Singing trials overview
(Experiment I)

Trial	$F_0(start)$ in Hz	$F_0(target)$ in Hz	Interval step	Start tone	Target tone
1	311.13	349.23	2	D#4	F4
2	311.13	196.00	-8	D#4	G3
3	311.13	311.13	0	D#4	D#4
4	311.13	440.00	6	D#4	A4
5	311.13	246.94	-4	D#4	B3
6	311.13	207.65	-7	D#4	G#3
7	311.13	261.63	-3	D#4	C4
8	311.13	523.26	9	D#4	C5
9	311.13	207.65	-7	D#4	G#3
10	311.13	523.26	9	D#4	C5
11	311.13	185.00	-9	D#4	F#3
12	311.13	311.13	0	D#4	D#4
13	311.13	370.00	3	D#4	F#4
14	311.13	261.63	-3	D#4	C4
15	311.13	185.00	-9	D#4	F#3
16	311.13	311.13	0	D#4	D#4
17	311.13	293.67	-1	D#4	D4
18	311.13	329.63	1	D#4	E4
19	311.13	185.00	-9	D#4	F#3
20	311.13	233.08	-5	D#4	A#3
21	311.13	293.67	-1	D#4	D4
22	311.13	466.17	7	D#4	A#4
23	311.13	293.67	-1	D#4	D4
24	311.13	466.17	7	D#4	A#4
25	311.13	207.65	-7	D#4	G#3
26	311.13	261.63	-3	D#4	C4
27	311.13	466.17	7	D#4	A#4
28	311.13	261.63	-3	D#4	C4
29	311.13	185.00	-9	D#4	F#3
30	311.13	466.17	7	D#4	A#4
31	311.13	311.13	0	D#4	D#4
32	311.13	233.08	-5	D#4	A#3
33	311.13	261.63	-3	D#4	C4
34	311.13	311.13	0	D#4	D#4
35	311.13	466.17	7	D#4	A#4
36	311.13	233.08	-5	D#4	A#3
37	311.13	370.00	3	D#4	F#4

38	311.13	185.00	-9	D#4	F#3
39	311.13	415.31	5	D#4	G#4
40	311.13	329.63	1	D#4	E4
41	311.13	523.26	9	D#4	C5
42	311.13	233.08	-5	D#4	A#3
43	311.13	523.26	9	D#4	C5
44	311.13	329.63	1	D#4	E4
45	311.13	370.00	3	D#4	F#4
46	311.13	293.67	-1	D#4	D4
47	311.13	185.00	-9	D#4	F#3
48	311.13	293.67	-1	D#4	D4
49	311.13	415.31	5	D#4	G#4
50	311.13	329.63	1	D#4	E4
51	311.13	523.26	9	D#4	C5
52	311.13	293.67	-1	D#4	D4
53	311.13	311.13	0	D#4	D#4
54	311.13	207.65	-7	D#4	G#3
55	311.13	261.63	-3	D#4	C4
56	311.13	233.08	-5	D#4	A#3
57	311.13	370.00	3	D#4	F#4
58	311.13	261.63	-3	D#4	C4
59	311.13	415.31	5	D#4	G#4
60	311.13	311.13	0	D#4	D#4
61	311.13	415.31	5	D#4	G#4
62	311.13	466.17	7	D#4	A#4
63	311.13	207.65	-7	D#4	G#3
64	311.13	311.13	0	D#4	D#4
65	311.13	261.63	-3	D#4	C4
66	311.13	370.00	3	D#4	F#4
67	311.13	233.08	-5	D#4	A#3
68	311.13	261.63	-3	D#4	C4
69	311.13	207.65	-7	D#4	G#3
70	311.13	329.63	1	D#4	E4
71	311.13	466.17	7	D#4	A#4
72	311.13	185.00	-9	D#4	F#3
73	311.13	293.67	-1	D#4	D4
74	311.13	311.13	0	D#4	D#4
75	311.13	415.31	5	D#4	G#4
76	311.13	293.67	-1	D#4	D4
77	311.13	466.17	7	D#4	A#4
78	311.13	523.26	9	D#4	C5
79	311.13	370.00	3	D#4	F#4
80	311.13	185.00	-9	D#4	F#3
81	311.13	523.26	9	D#4	C5

82	311.13	207.65	-7	D#4	G#3
83	311.13	370.00	3	D#4	F#4
84	311.13	207.65	-7	D#4	G#3
85	311.13	523.26	9	D#4	C5
86	311.13	370.00	3	D#4	F#4
87	311.13	185.00	-9	D#4	F#3
88	311.13	329.63	1	D#4	E4
89	311.13	233.08	-5	D#4	A#3
90	311.13	261.63	-3	D#4	C4
91	311.13	207.65	-7	D#4	G#3
92	311.13	329.63	1	D#4	E4
93	311.13	523.26	9	D#4	C5
94	311.13	466.17	7	D#4	A#4
95	311.13	523.26	9	D#4	C5
96	311.13	293.67	-1	D#4	D4
97	311.13	466.17	7	D#4	A#4
98	311.13	185.00	-9	D#4	F#3
99	311.13	415.31	5	D#4	G#4
100	311.13	207.65	-7	D#4	G#3
101	311.13	370.00	3	D#4	F#4
102	311.13	311.13	0	D#4	D#4
103	311.13	329.63	1	D#4	E4
104	311.13	311.13	0	D#4	D#4
105	311.13	415.31	5	D#4	G#4

Table 8.7: All musical tones and their frequencies which had been used in experiment I, as well as the corresponding interval size in semitones.

Table 8.8
Singing trials overview
(Experiment II)

Trial	$F_0(start)$ in Hz	$F_0(target)$ in Hz	Interval	Start tone	Target tone
1	311.13	349.23	2	D#4	F4
2	311.13	196.00	-8	D#4	G3
3	311.13	311.13	0	D#4	D#4
4	311.13	440.00	6	D#4	A4
5	311.13	246.94	-4	D#4	B3
6	311.13	207.65	-7	D#4	G#3
7	311.13	185.00	-9	D#4	F#3
8	311.13	466.17	7	D#4	A#4
9	311.13	523.26	9	D#4	C5
10	311.13	415.31	5	D#4	G#4
11	311.13	466.17	7	D#4	A#4
12	311.13	523.26	9	D#4	C5
13	311.13	329.63	1	D#4	E4
14	311.13	293.67	-1	D#4	D4
15	311.13	261.63	-3	D#4	C4
16	311.13	207.65	-7	D#4	G#3
17	311.13	233.08	-5	D#4	A#3
18	311.13	466.17	7	D#4	A#4
19	311.13	293.67	-1	D#4	D4
20	311.13	233.08	-5	D#4	A#3
21	311.13	415.31	5	D#4	G#4
22	311.13	293.67	-1	D#4	D4
23	311.13	233.08	-5	D#4	A#3
24	311.13	466.17	7	D#4	A#4
25	311.13	311.13	0	D#4	D#4
26	311.13	523.26	9	D#4	C5
27	311.13	207.65	-7	D#4	G#3
28	311.13	370.00	3	D#4	F#4
29	311.13	329.63	1	D#4	E4
30	311.13	523.26	9	D#4	C5
31	311.13	185.00	-9	D#4	F#3
32	311.13	293.67	-1	D#4	D4
33	311.13	261.63	-3	D#4	C4
34	311.13	311.13	0	D#4	D#4
35	311.13	329.63	1	D#4	E4
36	311.13	415.31	5	D#4	G#4
37	311.13	233.08	-5	D#4	A#3

38	311.13	185.00	-9	D#4	F#3
39	311.13	329.63	1	D#4	E4
40	311.13	311.13	0	D#4	D#4
41	311.13	207.65	-7	D#4	G#3
42	311.13	415.31	5	D#4	G#4
43	311.13	233.08	-5	D#4	A#3
44	311.13	261.63	-3	D#4	C4
45	311.13	370.00	3	D#4	F#4
46	311.13	261.63	-3	D#4	C4
47	311.13	185.00	-9	D#4	F#3
48	311.13	370.00	3	D#4	F#4
49	311.13	185.00	-9	D#4	F#3
50	311.13	293.67	-1	D#4	D4
51	311.13	523.26	9	D#4	C5
52	311.13	311.13	0	D#4	D#4
53	311.13	466.17	7	D#4	A#4
54	311.13	207.65	-7	D#4	G#3
55	311.13	370.00	3	D#4	F#4
56	311.13	329.63	1	D#4	E4
57	311.13	370.00	3	D#4	F#4
58	311.13	415.31	5	D#4	G#4
59	311.13	261.63	-3	D#4	C4
60	311.13	311.13	0	D#4	D#4

Table 8.8: All musical tones and their frequencies which had been used in experiment II, as well as the corresponding interval size in semitones.

B - Additional Figures

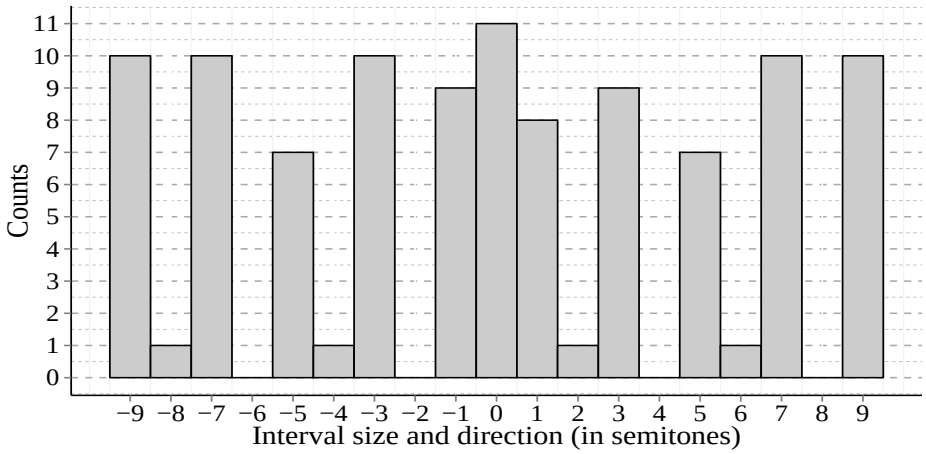


Figure 8.2: Histogram shows how often the different interval sizes had been used in the pitch reproduction task in experiment I. Interval direction could be positive (target tone was higher than the start tone) or negative (target tone was lower the the start tone).

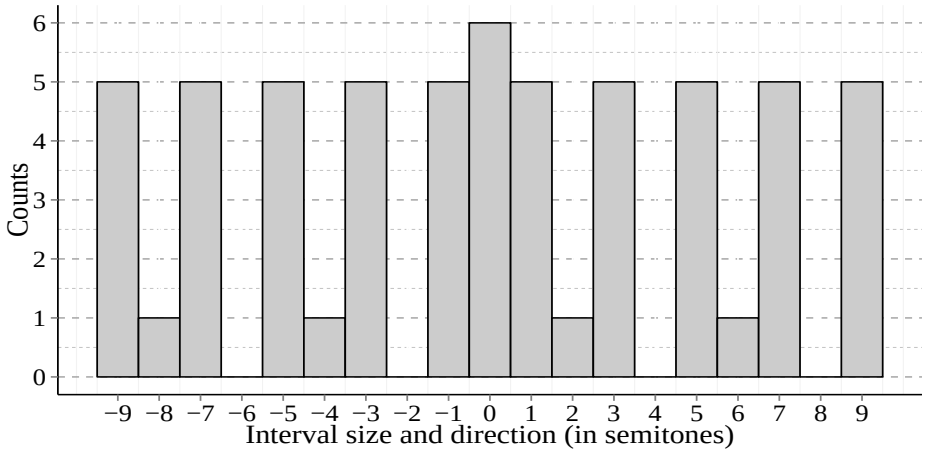


Figure 8.3: Histogram shows how often the different interval sizes had been used in the pitch reproduction task in experiment II. Interval direction could be positive (target tone was higher than the start tone) or negative (target tone was lower the the start tone).

C - Questionnaires

Edinburgh Handeness Inventory

Versuchspersonencode:

Händigkeitsfragebogen
Edinburgh-Handedness Inventory

Bitte geben Sie an, **welche Hand** (links oder rechts) **Sie bei folgenden Aktivitäten bevorzugen**, indem Sie in die entsprechende Spalte ein Kreuz machen.

- Wenn Sie also eine Hand bei der jeweiligen Tätigkeit bevorzugen machen Sie bitte in die entsprechende Spalte **ein Kreuz**.
- Wenn die Handbevorzugung so groß ist, das Sie nie versuchen würden, die andere Hand zu benutzen, machen Sie bitte **2 Kreuze** in die entsprechende Spalte.
- Falls Sie sich nicht schlüssig sind, ob Sie bei der Tätigkeit die linke oder die rechte Hand bevorzugen, machen Sie ein **Kreuz in beiden Spalten**.

Einige Aktivitäten erfordern beide Hände. In diesem Fall ist der Teil der Aufgabe, für den die Bevorzugung gefragt ist, in Klammern gesetzt.

Bitte beantworten Sie alle Fragen. Nur wenn Sie keinerlei Erfahrung mit dem Objekt oder der Aufgabe haben, lassen Sie die Spalte frei.

Falls Sie noch Fragen haben sollten, stellen Sie diese bitte dem Leiter der Untersuchung.

		L	R
1.	Schreiben		
2.	Zeichnen		
3.	Werfen		
4.	Schere		
5.	Zahnbürste		
6.	Messer (ohne Gabel)		
7.	Löffel		
8.	Besen (obere Hand)		
9.	Zündholz anzünden (Zündholz)		
10.	Schachtel öffnen (Deckel)		
i.	Mit welchem Fuß treten Sie bevorzugt einen Gegenstand?		
ii.	Welches Auge benutzen Sie, wenn Sie nur eines benutzen?		

Untersucher(in) Untersuchungsabschnitt

Musical Background Questionnaire

VPN-Code:

Fragen zur musikalischen Erfahrung

1. Ausbildung und Beruf

Schulabschluss: _____

Ausbildung/Beruf: _____

2. Musikalische Erfahrung

A: Gesang

1. Hatten Sie schon einmal Gesangsunterricht? Ja ☐ Nein ☐

   Wenn ja, in welchem Alter? Jahr: _____ bzw. von:_____bis: _____

   Wenn ja, wie oft? _____(h) Woche _____(Tage) Monat

2. Singen Sie alleine oder im Chor? Allein ☐ Chor ☐ Keines von beiden ☐

3. Wie viele Jahre singen Sie insgesamt? _____

4. Singen Sie regelmäßig? Ja ☐ Nein ☐

   Wenn ja, wie viele Tage/Monat? _____

   Wenn ja, wie viele Stunden/Woche? _____

5. Sangen Sie in einem Kinderchor? Ja ☐ Nein ☐

   Wenn ja, in welchem Alter? Jahr: _____ bzw. von:_____bis: _____

   Wenn ja, wie regelmäßig? _____

6. Wurde in Ihrer Familie gesungen? Ja ☐ Nein ☐

7. Wenn Sie Gesang studieren:

   Wann haben Sie damit begonnen (Alter)? _____

   In welchem Semester sind Sie? _____

   In welchem Studiengang sind sie? _____

   Wie viele Auftritte singen Sie pro Semester? _____

8. Haben Sie Freude am Singen?

   Keine ☐ Ein Wenig ☐ Durchschnittlich ☐ Viel ☐ Sehr viel ☐

B: Instrument

9. Spielen Sie ein Instrument? Ja ☐ Nein ☐ (Hauptinstrument)
= → Frage 11

Wenn ja, welches? _____

Wenn ja, seit wann (Alter)? Jahr: _____ bzw. von:_____bis: _____

Wenn ja, wie oft? _____(h) Woche _____(Tage) Monat

Hatten Sie darin Instrumentalunterricht? Ja ☐ Nein ☐

Wenn ja, seit wann (Alter)? Jahr: _____ bzw. von:_____bis: _____

Wenn ja, wie oft? _____(h) Woche _____(Tage) Monat

Hatten Sie an einer Hochschule unterricht? Ja ☐ Nein ☐

Wenn ja, in welchem Alter? Jahr: _____ bzw. von:_____bis: _____

Hatten Sie Auftritte mit dem Instrument? Ja ☐ Nein ☐

Wenn ja, seit wann (Alter)? Jahr: _____ bzw. von:_____bis: _____

Wie oft pro Semester? _____

10. Spielen Sie ein zweites Instrument? Ja ☐ Nein ☐ (Nebeninstrument)

Wenn ja, welches? _____

Wenn ja, in welchem Alter? Jahr: _____ bzw. von:_____bis: _____

Wenn ja, wie oft? _____(h) Woche _____(Tage) Monat

Hatten Sie Instrumentalunterricht? Ja ☐ Nein ☐

Wenn ja, seit wann Alter? Jahr: _____ bzw. von:_____bis: _____

Wenn ja, wie oft? _____(h) Woche _____(Tage) Monat

Ergänzung für weiterer Instrumente:_____

_____

11. Wurde in Ihrer Familie ein Instrument gespielt? ☐ Ja Nein ☐

Wenn ja, welche _____

12. Welche Musiktraditionen gab es Zuhause? _____

_____

C: Einfluss von Musik

13. Hatten Sie irgendwelchen Musiktheorieunterricht oder Hörtraining? Wenn ja, wie lange und wie alt waren Sie als Sie anfingen?_____

_____

14. Schätzen Sie Ihre Fähigkeit im Notenlesen ein:

Keine ☐ Ausreichend ☐ Durchschnittlich ☐ Gut ☐ Sehr gut ☐

15. Welchen Musikstil hören Sie am meisten? _____

16. Welchen Musikstil hören/hörten Ihre Eltern am meisten? _____

17. Mit welchem Alter entwickelten Sie Ihren eigenen Musikgeschmack? _____

18. Wie viele Stunden pro Woche hören Sie aufmerksam Musik? _____

3. Demographische Daten

19. Wie viele Sprachen sprechen Sie? Bitte zählen Sie diese in Reihenfolge beginnend mit der Muttersprache auf:_____

_____

20. In welchem Land sind Sie geboren und bzw. oder aufgewachsen?_____

_____

21. Haben Sie diagnostizierte Hörschäden? Ja ☐ Nein ☐

Wenn ja, welche? _____

22. Haben Sie ein absolutes Gehör? Ja ☐ Nein ☐
(Die Fähigkeit die Höhe eines beliebigen, gehörten Tones zu bestimmen)

23. Dürfen wir Sie für zukünftige Studien kontaktieren? Ja ☐ Nein ☐

TMS Safety Questionnaire

Versuchspersonencode:

Bitte beantworten Sie diese Fragen vor Durchführung der TMS-Untersuchung:

Traten bei Ihnen jemals Anfallsleiden oder unklare Bewusstseinseintrübungen auf?

O Nein **O Ja** _____
Welche

Sind in Ihrer Familie epileptische Erkrankungen bekannt?

O Nein **O Ja** _____

Bestehen oder bestanden andere neurologische oder psychiatrische Erkrankungen?
(z.B. Psychosen, Angstanfälle, Tics, Krampfanfälle)

O Nein **O Ja** _____

Hatten Sie jemals einen schweren Unfall unter Mitleidenschaft des Kopfes?
(z.B. Schädel-Hirn-Trauma)

O Nein **O Ja** _____

Hatten Sie jemals eine Operation am Kopf ?

O Nein **O Ja** _____

Wenn ja, befinden sich seit einer Operation Metallteile in Ihrem Kopf ?
(z.B. chirurgische Metallteile wie Platten, Drähte oder Klammern)

O Nein **O Ja** _____

Befindet sich durch einen Unfall oder eine Verletzung Metall in Ihren Kopf ?
(z.B. durch einen Auto-, Freizeit- oder Berufsunfall, oder eine Schussverletzung)

O Nein **O Ja** _____

Tragen Sie elektrische Stimulationsgeräte am Körper?
(zum Beispiel einen Herzschrittmacher, TENS- Einheit oder tiefe Hirnstimulation)

O Nein **O Ja** _____

Bitte wenden

Nehmen Sie regelmäßig Medikamente?

O Nein O Ja _____

z.B. Aspirin in hoher Dosis (> 2g/d) – Schleifendiuretika – Chemotherapie – Neuroleptika - regelmäßige Einnahme von sedierenden Mitteln (z.B. Benzodiazepinen)

Bei Frauen: könnten sie schwanger sein?

O Nein O Ja _____

Einwilligungserklärung

Mit dieser Unterschrift bestätige ich, dass ich über die Ziele, die Dauer, den Ablauf, den Nutzen sowie sämtliche Risiken der Studienteilnahme aufgeklärt worden bin. Ich bin darüber informiert worden, dass die Teilnahme an dieser Untersuchung von einem Arzt supervisiert werden muss, vollkommen freiwillig ist und dass das Einverständnis jederzeit ohne Angaben von Gründen und ohne Nachteile für mich widerrufen werden kann. Die Weitergabe von personenspezifischen Daten erfolgt gemäß Datenschutzrichtlinien ausschließlich in anonymisierter Form.

Name:_____ Geburtsdatum: _____

_____

(Ort) (Datum) (Unterschrift) Versuchsleiter

fMRI Safety Questionnaire

Sicherheitsfragen vor Durchführung einer MRT-Studie

Ist durch einen Unfall oder eine Verletzung Metall in Ihren Körper gekommen? Möglich wäre dies z.B.

durch einen Berufsunfall (Metallarbeiter, Schweißer, etc.)
durch Schussverletzungen oder Granatsplitter
bei Verletzungen im Gesicht v.a. am Auge

☐ Ja ☐ Nein

Befinden sich seit einer Operation Metallteile bzw. andere Implantate/Prothesen in Ihrem Körper wie

• Herzschrittmacher oder andere Geräte ?	□ ja	□ nein
• Implantate an Herz oder Blutgefäßen, z.B. künstliche Herzklappe, Stent, Shunt oder Port, Clips, Coils, Filter, Katheder, Defibrilator etc. ?	□ ja	□ nein
• orthopädische oder chirurgische Metallteile (Clips, Platten, Nägel, Drähte, Klammern, Nähte, etc.) ?	□ ja	□ nein
• irgendwelche Prothesen (Einfache Zahnplomben sind ohne Belang) ?	□ ja	□ nein
• andere Implantate, z.B. Gelenkimplantate, Intrauterinpessar (Spirale), Insulinpumpe etc. ?	□ ja	□ nein
• metallhaltige Teile anderer Art (z.B. abgebrochene Biopsienadeln, Dauerakupunktur-nadeln) ?	□ ja	□ nein
• Tätowierungen	□ ja	□ nein
• Piercing	□ ja	□ nein
• Permanent make-up	□ ja	□ nein

Einige Erkrankungen und besondere Umstände sollten bei der Untersuchung besonders berücksichtigt werden. Trifft einer der folgenden Punkte bei Ihnen zu?

• Zuckerkrankheit	□ ja	□ nein
• Bluthochdruck	□ ja	□ nein
• Lungen- oder Herzerkrankung	□ ja	□ nein
• Nieren- oder Lebererkrankung	□ ja	□ nein
• Blutarmut	□ ja	□ nein
• neurologische Erkrankungen wie Epilepsie, Schlaganfall, etc.	□ ja	□ nein
• sonstige Erkrankungen, die eine regelmäßige ärztliche Behandlung oder die Einnahme von Medikamenten erfordern	□ ja	□ nein
• Abhängigkeit (auch frühere) von Alkohol, Drogen oder Medikamenten	□ ja	□ nein
• bei Frauen: mögliche Schwangerschaft	□ ja	□ nein
• Klaustrophobie (Angstzuständen in engen Räumen)?	□ ja	□ nein

Bitte beachten!

Legen Sie bitte alle der im Folgenden aufgelisteten Gegenstände ab, bevor Sie den MR-Raum betreten, damit diese nicht beschädigt werden bzw. keine Unfälle verursachen können. Sie können Ihre Sachen vor dem Untersuchungsraum in Schließfächern deponieren. Ein Raum zum Umziehen ist ebenfalls vorhanden.

Vorher abzulegende Gegenstände:

Kreditkarten u.a. Karten mit Magnetstreifen
Brieftasche, Geldbeutel incl. Kleingeld in den Taschen !
Uhr
Hörgerät
Sicherheitsnadeln
Haarklammern, -nadeln
Schmuck (Ringe, Halsketten, Ohrringe, etc.)
Stifte
Schlüssel (außer dem Schließfachschlüssel)
Taschenmesser
Gürtel
Kugelschreiber
sonstige Gegenstände aus Metall, die Sie bei sich tragen

Brille und **Schuhe** können Sie im MR-Raum ablegen. Den **Schließfachschlüssel** sollten Sie bei den Untersuchern abgeben, die ihn sicher verwahren.
Bitte durchsuchen Sie alle Ihre Taschen und vergewissern Sie sich, dass Sie sicher keine Gegenstände mehr bei sich haben, die Metall enthalten könnten, bevor Sie den MR-Raum betreten.

Ich bestätige, dass Herr/Frau _____die Sicherheitsfragen mit mir besprochen hat, und die Richtigkeit meiner Angaben.

_____
(Name desProbanden)

_____ _____
Ort, Datum (Unterschrift des Probanden)

Bibliography

Ackermann, H. and Riecker, A. (2004). The contribution of the insula to motor aspects of speech production: a review and a hypothesis. *Brain Lang*, 89(2):320–8.

Adzaku, F. and Wyke, B. (1979). Innervation of the subglottic mucosa of the larynx, and its significance. *Folia Phoniatrica et Logopaedica*, 31(4):271–283.

Alcock, K. J., Wade, D., Anslow, P., and Passingham, R. E. (2000). Pitch and timing abilities in adult left-hemisphere-dysphasic and right-hemisphere-damaged subjects. *Brain and language*, 75(1):47–65.

Amir, O., Amir, N., and Kishon-Rabin, L. (2003). The effect of superior auditory skills on vocal accuracy. *The Journal of the Acoustical Society of America*, 113:1102 – 1108.

Amorosa, H., von Benda, U., Dames, M., and Schaferskupper, P. (1986). Deficits in fine motor coordination in children with unintelligible speech. *European archives of psychiatry and neurological sciences*, 236(1):26 – 30.

Andoh, J. and Zatorre, R. J. (2011). Interhemispheric connectivity influences the degree of modulation of tms-induced effects during auditory processing. *Front Psychol*, 2:161–161.

Ayotte, J., Peretz, I., and Hyde, K. (2002). Congenital amusia: a group study of adults afflicted with a music-specific disorder. *Brain*, 125(Pt 2):238–51.

Barker, A. T. and Freeston, I. (2007). Transcranial magnetic stimulation. *Scholarpedia*, 2(10):2936.

Barker, A. T., Jalinous, R., and Freeston, I. L. (1985). Non-invasive magnetic stimulation of human motor cortex. *The Lancet*, 325:1106 – 1107.

Boltez?ar, I. H., Burger, Z. R., and Z?argi, M. (1997). Instability of voice in adolescence: Pathologic condition or normal developmental variation? *The Journal of pediatrics*, 130(2):185–190.

Bradshaw, E. and McHenry, M. A. (2005). Pitch discrimination and pitch matching abilities of adults who sing inaccurately. *Journal of Voice*, 19(3):431–9.

Brown, S., Laird, A. R., Pfordresher, P. Q., Thelen, S. M., Turkeltaub, P., and Liotti, M. (2009). The somatotopy of speech: Phonation and articulation in the human motor cortex. *Brain and Cognition*, 70(1):31 – 41.

Brown, S., Ngan, E., and Liotti, M. (2008). A larynx area in the human motor cortex. *Cerebral Cortex*, 18(4):837–845.

Burnett, T. A., Freedland, M. B., Larson, C. R., and Hain, T. C. (1998). Voice f0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America*, 103(6):3153–61.

Burnett, T. A. and Larson, C. R. (2002). Early pitch-shift response is active in both steady and dynamic pitch control. *The Journal of the Acoustical Society of America*, 112:1058–163.

Castro-Alamancos, M. A., Donoghue, J. P., and Connors, B. W. (1995). Different forms of synaptic plasticity in somatosensory and motor areas of the neocortex. *J Neurosci*, 15(7 Pt 2):5324–33.

Conte, A., Rocchi, L., Nardella, A., Dispenza, S., Scontrini, A., Khan, N., and Berardelli, A. (2012). Theta-burst stimulation-induced plasticity over primary somatosensory cortex changes somatosensory temporal discrimination in healthy humans. *PLoS ONE*, 7(3):e32979.

Dalla Bella, S., Giguere, J. F., and Peretz, I. (2007). Singing proficiency in the general population. *The Journal of the Acoustical Society of America*, 121:1182 – 1189.

Dalla Bella, S., Giguere, J. F., and Peretz, I. (2009). Singing in congenital amusia. *J Acoust Soc Am*, 126(1):414–24.

Davis, P. and Nail, B. (1987). Quantitative analysis of laryngeal mechanosensitivity in the cat and rabbit. *The Journal of physiology*, 388(1):467–485.

Dawson, J. and Lauterbur, P. C. (2008). Magnetic resonance imaging. *Scholarpedia*, 3(7):3381.

de Cheveigne, A. and Kawahara, H. (2002). Yin, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111(4):1917–1930.

Eickhoff, S. B., Heim, S., Zilles, K., and Amunts, K. (2009). A systems perspective on the effective connectivity of overt speech production. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 367(1896):2399–421.

Estis, J. M., Dean-Claytor, A., Moore, R. E., and Rowell, T. L. (2011). Pitch-matching accuracy in trained singers and untrained individuals: The impact of musical interference and noise. *Journal of Voice*, 25(2):173 – 180.

Faro, S. H. and Mohamed, F. B. (2006). *Functional MRI: basic principles and clinical applications*. Springer Science & Business Media.

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., and Griffiths, T. D. (2004). Characterization of deficits in pitch perception underlying 'tone deafness'. *Brain: a journal of neurology*, 127(4):801–810.

Friberg, A., Schoonderwaldt, E., and Juslin, P. N. (2005). Cuex: An algorithm for extracting expressive tone variables from audio recordings. *Acustica united with Acta Acustica*, 93(3):411–420.

Friberg, A., Schoonderwaldt, E., and Juslin, P. N. (2007). Cuex: An algorithm for automatic extraction of expressive tone parameters in music performance from acoustic signals. *Acta Acustica united with Acustica*, 93(3):411–420.

Gelfand, S. A. (2009). *Essentials of Audiology*. Thieme Medical Publishers, New York.

Ghasemi, A. and Zahediasl, S. (2012). Normality tests for statistical analysis: a guide for non-statisticians. *International journal of endocrinology and metabolism*, 10(2):486.

Goldsworthy, M. R., Pitcher, J. B., and C, R. M. (2012). A comparison of two different continuous theta burst stimulation paradigms applied to the human primary motor cortex. *Clinical Neurophysiology*, 123(11):2256 – 2263.

Golfinopoulos, E., Tourville, J. A., Bohland, J. W., Ghosh, S. S., Nieto-Castanon, A., and Guenther, F. H. (2011). fmri investigation of unexpected somatosensory feedback perturbation during speech. *NeuroImage*, 55(3):1324–38.

Golfinopoulos, E., Tourville, J. A., and Guenther, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *NeuroImage*, 52(3):862 – 874.

Gozaine, T. C. and Clark, K. F. (2005). Function of the laryngeal mechanoreceptors during vocalization. *The Laryngoscope*, 115(1):81–8.

Grabski, K., Lamalle, L., Vilain, C., Schwartz, J. L., Vallee, N., Tropres, I., Baciu, M., Le Bas, J. F., and Sato, M. (2012). Functional mri assessment of orofacial articulators: Neural correlates of lip, jaw, larynx, and tongue movements. *Human brain mapping*, 33(10):2306–21.

Gray, H. (1918). *Anatomy of the Human Body*. Lea & Febiger, Philadephia.

Grell, A., Sundberg, J., Ternstrom, S., Ptok, M., and Altenmuller, E. (2009). Rapid pitch correction in choir singers. *The Journal of the Acoustical Society of America*, 126(1):407–13.

Guenther, F. H., Ghosh, S. S., and Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang*, 96(3):280–301.

Guenther, F. H. and Vladusich, T. (2012). A neural theory of speech acquisition and production. *Journal of neurolinguistics*, 25(5):408–422.

Hacki, T. (1999). Vocal capabilities of nonprofessional singers evaluated by measurement and superimposition of their speaking, shouting and singing voice range profiles. *HNO*, 47(9):809–15.

Hain, T. C., Burnett, T. A., Kiran, S., Larson, C. R., Singh, S., and Kennedy, M. K. (2000). Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Experimental Brain Research*, 130:133 – 141.

Hain, T. C., Burnett, T. A., Larson, C. R., and Kiran, S. (2001). Effects of delayed auditory feedback (daf) on the pitch-shift reflex. *The Journal of the Acoustical Society of America*, 109(5 Pt 1):2146–52.

Hamada, M., Murase, N., Hasan, A., Balaratnam, M., and Rothwell, J. C. (2012). The role of interneurons networks in driving human motor cortical plasticity. *Cerebral Cortex*, 23(7):1593 – 605.

Hannig, S. and Jürgens, U. (2006). Projections of the ventrolateral pontine vocalization area in the squirrel monkey. *Experimental Brain Research*, 169(1):92–105.

Herwig, U. and Schönfeldt-Lecuona (2007). Neuronavigation der transkraniellen magnetstimulation. In Siebner, H. R. and Ziemann, R., editors, *Das TMS Buch - Transkranielle Magnetstimulation*. Springer Medizin Verlag, Heidelberg.

Hess, C. W. (2007). Hirnstimulation - historischer überblick. In Siebner, H. R. and Ziemann, R., editors, *Das TMS Buch - Transkranielle Magnetstimulation*. Springer Medizin Verlag, Heidelberg.

Heylen, L., Wuyts, F., Mertens, F., De Bodt, M., and Van de Heyning, P. (2002). Normative voice range profiles of male and female professional voice users. *Journal of voice*, 16(1):1–7.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, 13:135 – 145.

Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron*, 69(3):407–22.

Houde, J. F. and Nagarajan, S. S. (2011). Speech production as state feedback control. *Frontiers in human neuroscience*, 5:82.

Huang, Y.-Z., Chen, R.-S., Rothwell, J. C., and Wen, H.-Y. (2007). The after-effect of human theta burst stimulation is nmda receptor dependent. *Clinical Neurophysiology*, 118(5):1028–32+.

Huang, Y.-Z., Edwards, M. J., Rounis, E., Bhatia, K. P., and Rothwell, J. C. (2005). Theta burst stimulation of the human motor cortex. *Neuron*, 45(2):201–206.

Huang, Y. Z., Rothwell, J. C., Chen, R. S., Lu, C. S., and Chuang, W. L. (2011). The theoretical model of theta burst form of repetitive transcranial magnetic stimulation. *Clinical Neurophysiology*, 122(5):1011 – 8.

Hutchins, S., Gosselin, N., and Peretz, I. (2010). Identification of changes along a continuum of speech intonation is impaired in congenital amusia. *Frontiers in psychology*, 1:236.

Hutchins, S. and Peretz, I. (2011). Perception and action in singing. *Progress in Brain Research*, 191:103–18.

Hutchins, S. and Peretz, I. (2013). Vocal pitch shift in congenital amusia (pitch deafness). *Brain and language*, 125(1):106 – 117.

Hutchins, S. M. and Peretz, I. (2012). A frog in your throat or in your ear? searching for the causes of poor singing. *Journal of Experimental Psychology General*, 141:76 – 97.

Hyde, K. L. and Peretz, I. (2004). Brains that are out of tune but in time. *Psychological science*, 15(5):356–60.

Hyde, K. L., Peretz, I., and Zatorre, R. J. (2008). Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia*, 46(2).

ISO 226 (2003). Acoustics – normal equal-loudness-level contours. Technical report, International Organisation for Standardisation.

Ito, T. and Ostry, D. J. (2010). Somatosensory contribution to motor learning due to facial skin deformation. *Journal of neurophysiology*, 104(3):1230–1238.

Ito, T. and Ostry, D. J. (2012). Speech sounds alter facial skin sensation. *Journal of neurophysiology*, 107(1):442–447.

Jones, J. A. and Keough, D. (2008). Auditory-motor mapping for pitch control in singers and nonsingers. *Exp Brain Res*.

Jones, J. A. and Munhall, K. G. (2000). Perceptual calibration of f0 production: evidence from feedback perturbation. *The Journal of the Acoustical Society of America*, 108:1246 – 1251.

Jürgens, U. (2002). Neural pathways underlying vocal control. *Neurosci Biobehav Rev*, 26:235 – 258.

Jürgens, U. and Zwirner, P. (2000). Individual hemispheric asymmetry in vocal fold control of the squirrel monkey. *Behavioural Brain Research*, 109(2):213–217.

Juslin, P. N. and Scherer, K. R. (2008). Speech emotion analysis. *Scholarpedia*, 3(10):4240. revision 90367.

Kaelin-Lang, A. (2007). Motorisch evozierte potenziale - eine einfhrung. In Siebner, H. R. and Ziemann, R., editors, *Das TMS Buch - Transkranielle Magnetstimulation*. Springer Medizin Verlag, Heidelberg.

Kaelin-Lang, A. and Cohen, L. G. (2000). Enhancing the quality of studies using transcranial magnetic and electric stimulation with a new computer-controlled system. *J Neurosci Methods*, pages 81 – 89.

Kalla, R., Muggleton, N. G., Cowey, A., and Walsh, V. (2009). Human dorsolateral prefrontal cortex is involved in visual search for conjunctions but not features: A theta tms study. *Cortex*, 45(9):1085–1090.

Katayama, T. and Rothwell, J. C. (2007). Modulation of somatosensory evoked potentials using transcranial magnetic intermittent theta burst stimulation. *Clinical Neurophysiology*, 118(11):2506 – 2511.

Katseff, S., Houde, J., and Johnson, K. (2012). Partial compensation for altered auditory feedback: A tradeoff with somatosensory feedback? *Language and speech*, 55(2):295–308.

Kiers, L., Cros, D., Chiappa, K. H., and Fang, J. (1993). Variability of motor potentials evoked by transcranial magnetic stimulation. *Electroencephalogr Clin Neurophysiol*, 89:415 – 423.

Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J., and Lotze, M. (2010). The brain of opera singers: experience-dependent changes in functional activation. *Cereb Cortex*, 20(5):1144 – 52.

Kleber, B., Zeitouni, A. G., Friberg, A., and J, Z. R. (2013). Experience-dependent modulation of feedback integration during singing: the role of the right anterior insula. *The Journal of Neuroscience*, 33:6070 – 6080.

Ladda, A. M., Pfannmoeller, J. P., Kalisch, T., Roschka, S., Platz, T., Dinse, H. R., and Lotze, M. (2014). Effects of combining 2 weeks of passive sensory stimulation with active hand motor training in healthy adults. *PloS one*, 9(1):e84402.

Lametti, D. R., Nasir, S. M., and Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *The Journal of Neuroscience*, 32:9351 – 9358.

Lang, N. and Siebner, H. R. (2007). Repetitive transkranielle magnetstimulation. In Siebner, H. R. and Ziemann, R., editors, *Das TMS Buch - Transkranielle Magnetstimulation*. Springer Medizin Verlag, Heidelberg.

Larson, C. R. (1998). Cross-modality influences in speech motor control: the use of pitch shifting for the study of f0 control. *Journal of communication disorders*, 31(6):489–502; quiz 502–3; 553.

Larson, C. R., Altman, K. W., Liu, H., and Hain, T. C. (2008). Interactions between auditory and somatosensory feedback for voice f0 control. *Experimental brain research*, 187(4):613–21.

Larson, C. R., Burnett, T. A., Kiran, S., and Hain, T. C. (2000). Effects of pitch-shift velocity on voice fo responses. *The Journal of the Acoustical Society of America*, 107(1):559–64.

Larson, J. and Lynch, G. (1986). Induction of synaptic potentiation i hippocampus by patterned stimulation involves two events. *Science*, 232:985 – 988.

Lazar, N. (2008). *The statistical analysis of functional MRI data*. Springer Science & Business Media.

Leveque, Y., Muggleton, N., Stewart, L., and Schön, D. (2013). Involvement of the larynx motor area in singing-voice perception: a tms study. *Frontiers in Psychology*, 4(418).

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49:467+.

Loucks, T. M., Poletto, C. J., Saxon, K. G., and Ludlow, C. L. (2005). Laryngeal muscle responses to mechanical displacement of the thyroid cartilage in humans. *Journal of Applied Physiology*, 99(3):922 – 930.

Loui, P., Alsop, D., and Schlaug, G. (2009). Tone deafness: a new disconnection syndrome? *The Journal of neuroscience*, 29(33):10215–20.

Ludlow, C. L. (2005). Central nervous system control of the laryngeal muscles in humans. *Respir Physiol Neurobiol*, 147(2-3):205 – 222.

Micheyl, C., Delhommeau, K., Perrot, X., and J, O. A. (2006). Influence of musical and psychoacoustical training on pitch discrimination. *Hearing Research*, 219(1):36 – 47.

Moore, R. E., Estis, J., Gordon-Hickey, S., and Watts, C. (2010). Pitch discrimination and pitch matching abilities with vocal and nonvocal stimuli. *Journal of voice*, 22(4):399 – 407.

Morley, J. W., Vickery, R. M., Stuart, M., and Turman, A. B. (2007). Suppression of vibrotactile discrimination by transcranial magnetic stimulation of primary somatosensory cortex. *European Journal of Neuroscience*, 26(4):1007 – 1010.

Mürbe, D., Pabst, F., Hofmann, G., and Sundberg, J. (2002). Significance of auditory and kinethetic feedback to singers' pitch control. *Journal of voice*, 16(1):44 – 51.

Mürbe, D., Pabst, F., Hofmann, G., and Sundberg, J. (2004). Effects of professional solo singer education on auditory and kinesthetic feedback - a longitudinal study of singers' pitch control. *Journal of voice*, 18(2):236 – 41.

Murry, T. (1990). Pitch-matching accuracy in singers and nonsingers. *J. Voice*, 4:317–321.

Mürbe, D., Lindner, P., Zollner, S., Welsch, H., Kuhlisch, E., Huttenbrink, K. B., and Sundberg, J. (2004). Change of voice characteristics during +3 gz acceleration. *Aviation, space, and environmental medicine*, 75(12).

Nasir, S. M. and Ostry, D. J. (2006). Somatosensory precision in speech production. *Current biology*, 16(19):1918–23.

Nasir, S. M. and Ostry, D. J. (2008). Speech motor learning in profoundly deaf adults. *Nature neuroscience*, 11(10):1217–22.

Natke, U., Donath, T. M., and Kalveram, K. T. (2003). Control of voice fundamental frequency in speaking versus singing. *The Journal of the Acoustical Society of America*, 113(3):1587 – 93.

Nikjeh, D. A., Lister, J. J., and Frisch, S. A. (2009). The relationship between pitch discrimination and vocal production: comparison of vocal and instrumental musicians. *J Acoust Soc Am*, 125(1):328–38.

Ogawa, S., Lee, T. M., Kay, A. R., and Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci USA*, 87:9868 – 9872.

Ogawa, S. and Sung, Y. (2007). Functional magnetic resonance imaging. *Scholarpedia*, 2(10):3105.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The edinburgh inventory. *Neuropsychologia*, 9:97 – 113.

Pauling, L. and Coryell, C. D. (1936). The magnetic properties and structure of hemoglobin, oxyhemoglobin and carbonmonoxy hemoglobin. *Proc Nayl Acad Sci*, 22:210 – 216.

Penny, W. D., Friston, K. J., Ashburner, J. T., Kiebel, S. J., and Nichols, T. E. (2011). *Statistical parametric mapping: the analysis of functional brain images: the analysis of functional brain images*. Academic press.

Peretz, I., Cummings, S., and Dube, M. P. (2007). The genetics of congenital amusia (tone deafness): a family-aggregation study. *American journal of human genetics*, 81(3):582–8.

Perkell, J. S., Lane, H., Denny, M., Matthies, M. L., Tiede, M., Zandipour, M., Vick, J., and Burton, E. (2007). Time course of speech changes in response to unanticipated short-term changes in hearing state. *The Journal of the Acoustical Society of America*, 121(4):2296–311.

Premji, A., Ziluk, A., and Nelson, A. (2010). Bilateral somatosensory evoked potentials following intermittent theta-burst repetitive transcranial magnetic stimulation. *BMC Neuroscience*, 11(1):91.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ragert, P., Franzkowiak, S., Schwenkreis, P., Tegenthoff, M., and Dinse, H. R. (2008). Improvement of tactile perception and enhancement of cortical excitability through intermittent theta burst rtms over human primary somatosensory cortex. *Exp Brain Res*, 184:1–11.

Rai, N., Premj, A., Tommerdahl, M., and Nelson, A. J. (2012). Continous theta-burst rtms over primary somatosensory cortex modulates tactile perception on the hand. *Clin Neurophysiol*, 123(6):1226 – 33.

Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., and Grodd, W. (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport*, 11(9):1997–2000.

Riedel, G., Platt, B., and Micheau, J. (2003). Glutamate receptor function in learning and memory. *Behavioural Brain Res.*, page 1 – 47.

Rödel, R. M., Olthoff, A., Tergau, F., Simonyan, K., Kraemer, D., Markus, H., and Kruse, E. (2004). Human cortical motor representation of the larynx as assessed by transcranial magnetic stimulation (tms). *The Laryngoscope*, 114(5):918 922.

Rossini, P. M., Barker, A. T., Beradelli, A., and et al. (1994). Non-invasive eletrical and magnetic stimulation of the brain, spinal cord and roots: basic principles and procedures for routine clinical application. report of an ifcn committee. *Electrocephal Clin Neurophysiol*, 91:79 – 92.

Rothwell, J. C., Hallett, M., Beradelli, A., and et al. (1999). Magnetic stimulation: motor evoked potentials. international federation of clinical neurophysiology. *Electroencephalogr Cin Neurophysiol Suppl*, 52:97 – 103.

Sanders, I., Han, Y., Wang, J., and Biller, H. (1998). Muscle spindles are concentrated in the superior vocalis subcompartment of the human thyroarytenoid muscle. *J Voice*, 12:7 – 16.

Sanders, I., Wu, B.-L., Mu, L., Li, Y., and Biller, H. F. (1993). The innervation of the human larynx. *Archives of Otolaryngology–Head & Neck Surgery*, 119(9):934–939.

Sapir, S., Baker, K. K., Larson, C. R., and Ramig, L. O. (2000). Short-latency changes in voice f0 and neck surface emg induced by mechanical perturbations of the larynx during sustained vowel phonation. *Journal of speech, language, and hearing research : JSLHR*, 43(1):268–76.

Sapir, S., McClean, M. D., and Larson, C. R. (1983). Human laryngeal responses to auditory stimulation. *The Journal of the Acoustical Society of America*, 73(1):315–21.

Siebner, H. R. and Ziemann, R. (2007). Hirnstimulation - physiologische grundlagen. In Siebner, H. R. and Ziemann, R., editors, *Das TMS Buch - Transkranielle Magnetstimulation*. Springer Medizin Verlag, Heidelberg.

Siemens (2003). *Magnete, Spins und Resonanzen - Eine Einführung in die Grundlagen der Magnetresonanztomographie*. Siemens AG Medical Solution.

Simonyan, K. and Horwitz, B. (2011). Laryngeal motor cortex and control of speech in humans. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*, 17(2):197–208.

Simonyan, K., Ostuni, J., Ludlow, C. L., and Horwitz, B. (2009). Functional but not structural networks of the human laryngeal motor cortex show left hemispheric lateralization during syllable but not breathing production. *The Journal of Neuroscience*, 29(47):14912–23.

Smith, A. (2006). Speech motor development: Integrating muscles, movements, and linguistic units. *J Commun Disord*, 39:331 – 349.

Smith, A. and Zelaznik, H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Dev Psychobiol*, 45:22 – 33.

Smotherman, M. S. (2007). Sensory feedback control of mammalian vocalizations. *Behavioural brain research*, 182(2):315–26.

Standring, S., Ellis, H., Healy, J., Johnson, D., Williams, A., Collins, P., and Wigley, C. (2005). Gray's anatomy: the anatomical basis of clinical practice. *American Journal of Neuroradiology*, 26(10):2703.

Stewart, L. (2011). Characterizing congenital amusia. *Quarterly journal of experimental psychology*, 64(4):625 – 638.

Stöcker, H. (2004). *Taschenbuch der Physik*, volume 5. Verlag Harri Deutsch Thun.

Sulica, L. (2004). The superior laryngeal nerve: function and dysfunction. *Otolaryngologic Clinics of North America*, 37(1):183–201.

Sundberg, J. (1974). Articulatory interpretation of the ?singing formant? *The Journal of the Acoustical Society of America*, 55(4):838–844.

Sundberg, J. (1977). The acoustics of the singing voice. *Scientific American*, page 82ff.

Tegenthoff, M., Ragert, P., Pleger, B., Schwenkreis, P., Förster, A., Nicolas, V., and Dinse, H. R. (2005). Improvement of tactile discrimination performance and enlargement of cortical somatosensory maps after 5 hz rtms. *PLoS Biol*, 3(11):e362.

Ternstrom, S., Sundberg, J., and Collden, A. (1988). Articulatory fo perturbations and auditory feedback. *Journal of speech and hearing research*, 31:187 – 192.

Thielscher, A., Opitz, A., and Windhoff, M. (2010). Impact of the gyral geometry on the electric field induced by transcranial magnetic stimulation. *NeuroImage*, 54:234 – 243.

Tian, X. and Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Frontiers in psychology*, 1:166.

Titze, I. R. (1988). The physics of small-amplitude oscillation of the vocal folds. *The Journal of the Acoustical Society of America*, 83(4):1536–1552.

Titze, I. R. (1993). Current topics in voice production mechanisms. *Acta Otolaryngol*, 113:421 – 427.

Titze, I. R. (2008). Nonlinear source–filter coupling in phonation: Theorya). *The Journal of the Acoustical Society of America*, 123(5):2733–2749.

Titze, I. R. and Hunter, E. J. (2004). Normal vibration frequencies of the vocal ligament. *The Journal of the Acoustical Society of America*, 115(5 Pt 1):2264–9.

Traunmüller, H. and Eriksson, A. (1994). The frequency range of the voice fundamental in the speech of male and female adults. Technical report, Department of Linguistics, University of Stockholm.

Tremblay, S., Shiller, D. M., and Ostry, D. J. (2003). Somatosensory basis of speech production. *Nature*, 423(6942):866–9.

Tremblay-Champoux, A., Dalla Bella, S., Phillips-Silver, J., Lebrun, M. A., and Peretz, I. (2010). Singing proficiency in congenital amusia: imitation helps. *Cognitive neuropsychology*, 27(6):463 – 476.

Trepel, M. (2008). *Neuroanatomie*, volume 4. Urban & Fischer.

Tsang, P., Jacobs, M. F., Lee, K. G., Asmussen, M. J., Zapallow, C. M., and Nelson, A. J. (2014). Continuous theta-burst stimulation over primary somatosensory cortex modulates short-latency afferent inhibition. *Clinical Neurophysiology*, 125(11):2253–2259.

Ueno, S., Tashiro, T., and Harada, K. (1988). Localized stimulation of neural tissue in the brain by means of a paired conuration of time-varying magnetic lds. *Journal of Applied Physics*, 64(10):5862 – 5864.

Ward, L. M. and Greenwood, P. E. (2007). 1/f noise. *Scholarpedia*, 2(12):1537.

Wassermann, E. M. (1998). Risk and safety of repetitive transcranial magnetic stimulation: report and suggested guidelines from the international workshop on the safety of repetitive transcranial magnetic stimulation.

Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 108(1):1 – 16.

Watts, C., Moore, R., and McCaghren, K. (2005). The relationship between vocal pitch-matching skills and pitch discrimination skills in untrained accurate and inaccurate singers. *Journal of Voice*, 19(4):534–43.

Watts, C., Murphy, J., and Barnes-Burroughs, K. (2003). Pitch matching accuracy of trained singers, untrained subjects with talented singing voices, and untrained subjects with nontalented singing voices in conditions of varying feedback. *Journal of Voice*, 17(2):185–94.

Weißhaupt, D., Köchli, V., and Marincek, B. (2014). *Wie funktioniert MRI?: Eine Einfhrung in Physik und Funktionsweise der Magnetresonanzbildgebung*. Springer Science & Business Media.

Weyh, T. and Siebner, R. S. (2007). Hirnstimulation - technische grundlagen. In Siebner, H. R. and Ziemann, R., editors, *Das TMS Buch - Transkranielle Magnetstimulation*. Springer Medizin Verlag, Heidelberg.

Wyke, B. D. (1974a). Laryngeal myotatic reflexes and phonation. *Folia Phoniatrica*, 26:249 – 264.

Wyke, B. D. (1974b). Laryngeal neuromuscular control systems in singing. a review of current concepts. *Folia Phoniatrica*, 26:295 – 306.

Zarate, J. M. (2013). Cortical mechanisms of integrating auditory feedback with vocal pitch control. *The Journal of the Acoustical Society of America*, 133(5):3563.

Zarate, J. M., Delhommeau, K., Wood, S., and Zatorre, R. J. (2010). Vocal accuracy and neural plasticity following micromelody-discrimination training. *PLoS One*, 5.

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

Zatorre, R. J. and Baum, S. R. (2012). Musical melody and speech intonation: singing a different tune. *PLoS Biol*, 10.

Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences*, 6(1):37 – 46.

Zatorre, R. J., Delhommeau, K., and Zarate, J. M. (2012). Modulation of auditory cortex response to pitch variation following training with microtonal melodies. *Frontiers in Psychology*, 3:544.

Acknowledgement

During the period the work of this thesis was conducted several people gave me great support. Therefore I would like to thank first of all Dr. Boris Kleber for all the fruitful discussions about music & neuroscience and keeping me on track. Prof. Birbaumer for giving me the opportunity to work in his lab and Prof. Braun for his invaluable feedback on the previous manuscript. Dr. Ralf Veit for support in acquiring the MRI data. Dr. Anne Schmid for the help with the TMS System. Dr. Anders Friberg for his assistance in analysing the audio data.
Finally, I like to say thank you to my family, Maike and Mira, for supporting me with their sweet smiles that make my heart sing, and to my parents for helping me to come this far.