Neural networks underlying vocal control in experienced classical singers

Dissertation
der Fakultät für Informations- und Kognitionswissenschaften
der Eberhard-Karls-Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Dipl.-Psych. Boris Kleber
aus Stuttgart

Tübingen
2009

Tag der mündlichen Qualifikation: 15.07.2009

Dekan: Prof. Dr.-Ing. Oliver Kohlbacher

1. Berichterstatter: Prof. Dr. Martin Hautzinger

2. Berichterstatter: Prof. Dr. Niels Birbaumer

Meinen Eltern, Heidemarie und Karl Kleber

Acknowledgements

First and foremost, I would like to express my deepest gratitude to Prof. Niels Birbaumer. I will never forget when we first met at the Stuttgart Opera and he asked me to visit him at his institute in Tübingen, because “we may have a project you could be interested in”. It all came differently though. He strongly supported my idea for a new project involving singing and allowed me to change over from the BCI group, when the DFG awarded the proposal. He also entrusted me with the organization of a project the content of which was way beyond the scope of this dissertation. I feel deeply honored by his faith in me! Ever since, Prof. Birbaumer has been a great mentor whose door was always open when his support was needed. He is not only an outstanding scientist but also a great source of inspiration that never seem to run dry.

I am deeply grateful to Prof. Martin Hautzinger, who has always been supportive during all stages of this dissertation. His expertise and understanding was indispensable for accomplishing this work.

Without Prof. John Gruzelier, I would most likely not be where I am now. He stimulated and fostered my interest in neuroscience and showed me that the sciences and the arts can indeed celebrate a successful marriage. I am indebted to you for leading me on this path! Thank you ever so much!

I would like to thank Dr. Ralf Veit and Prof. Martin Lotze for sharing their knowledge on fMRI with me, for giving me their invaluable expertise and friendship, and also for spending so many hours with me at the scanner listening to “caro mio ben”! I am sure this song will never be forgotten!

The Institute of Medical Psychology and Behavioral Neurobiology is a special place and there are so many people who are responsible for that. A big “thanks” in particular goes to PD Dr. Ute Strehl for always being so supportive and to Lydia, Angela, Slavica, Hannelore, and Maike. There are many more, and it was a privilege to work with you guys!

Finally, I would like to thank my family! Although I only had 10 years with my father, this was a very precious time. He was of a rare kind with his warm-heartedness, his loving being, his charisma and his unmatched musicality as a widely respected concert pianist. He placed a seed within me and the plant is still growing. I will thank you forever! There are no words to express my gratitude to my mother. Your faith in me never failed and I cannot say enough how much I thank you for this! Last but definitely not least I would like to express my deeply felt thankfulness to Helga, to Mark and also to Olga!

Eidestattliche Erklärung

Hiermit erkläre ich, dass ich unter Verwendung der im Literaturverzeichnis aufgeführten Quellen und unter fachlicher Betreuung diese Dissertation selbstständig verfasst habe.

(Boris Kleber)

Table of Contents

Abbreviations

A1	primary auditory cortex
AAL	automatic anatomic labeling
ACC	anterior cingulate cortex
ACG	anterior cingulate gyrus
BA	Brodmann area
BOLD	blood oxygenation level dependent
CBF	cerebral blood flow
cM1	contralateral primary motor cortex
df	degrees of freedom
DLPFC	dorsolateral prefrontal cortex
EMG	electromyography
EPI	echo planar imaging
FDR	false discovery rate
fMRI	functional Magnetic Resonance Imaging
FWE	family wise error rate
HRF	hemodynamic response function
HDR	hemodynamic response
Hipp	hippocampus
IPC	inferior parietal cortex
LI	lateralization index
M1	primary motor cortex
MNI	Montreal Neurological Institute
MRI	magnetic resonance imaging
MyHC	myosin heavy chain
NMR	nuclear magnetic resonance
p	probability
PFC	prefrontal cortex
PMC	premotor cortex
RF	radio frequency
ROI	region of interest
S1	primary somatosensory cortex
SD	standard deviation
SFG	superior frontal gyrus
SII	secondary somatosensory cortex

SMA	supplementary motor cortex
SPL	superior parietal cortex
SPM	statistical parametrical mapping
SPS	samples per second
SPSS	statistical package for the social science
STF	slow twitch muscle fibers
STG	superior temporal gyrus
STS	superior temporal sulcus
t	Student's t-test
TE	echo time
TE1.0-2	cytoarchitectonic definition of primary auditory cortex
TR	repetition time
VAS	visual analogue scale

Publications:

Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., & Lotze, M. (2007). Overt and imagined singing of an Italian aria. *Neuroimage, 36*(3), 889-900.

Kleber, B., Veit, R., Birbaumer,N., Gruzelier, J., & Lotze, M. (2009). The brain of opera singers: experience-dependent changes in functional activation. *Cerebral Cortex*. Accepted for publication

Overview

Chapter 1 briefly illustrates the content, background, and motivation for the studies of the present dissertation.

Chapter 2 provides an overview on voice physiology including the structural anatomy of the larynx and the functional systems involved in voice production. Furthermore, it outlines differences between trained and untrained voices.

Chapter 3 describes the concepts of expertise research in the music domain, which basically explain musical proficiency as a function of deliberate practice.

Chapter 4 gives a brief introduction on the basic principles of magnetic resonance imaging, with a particular focus on the methods and procedures of functional magnetic resonance imaging (fMRI) and data processing.

Chapter 5 reports on an fMRI experiment with professionally trained singers. This study focused on the common and distinct neural networks underlying vocal control during both overt and imagined singing of an Italian aria.

Chapter 6 presents the results of a subsequent fMRI study investigating experience dependent changes in functional brain activation during overt singing in a large sample of opera singers, vocal students and amateurs.

Chapter 7 gives a summary on the results and procedures employed in the present dissertation, and provides an outlook to future research.

Chapter 1 Introduction

The human voice is a remarkable system, but one that is not well understood with regard to the neural networks that govern it. During the last 50 years, scientific studies have substantially increased our knowledge of the behavioral and biomechanical principles involved in voice production. Among these, it was particularly the work of Johann Sundberg and Ingo Titze that contributed most to the discovery of the unique properties of the vocal system in general and their complex interaction for singing in particular. Although vocal function in speaking and singing shares the same basic principles, singing requires a much higher level of control over vocal subsystems. This is particularly evident in classically trained singers, whose vocal efficiency can be distinguished from untrained singers as well as from other singing styles. Such highly developed singing skills affect both vocal biomechanics and resulting sound production.

Much less is known about the neural mechanisms underlying vocal motor control in singing. The technological advances in the cognitive neurosciences provide the means to investigate this question. Particularly the developments in imaging technology allow unraveling the neural underpinnings of behavior, and offer new possibilities to study the cognitive processes involved when actions are planned and executed by the brain. To date, several studies investigated these mechanisms with respect to speaking, while only few of them involved simple singing in untrained subjects. Studying singing in trained experts allows investigating the role of sensorimotor integration and neural plasticity in a domain that has not yet received much attention and that is distinct from the well-researched peripheral motor system.

The present work addressed this subject matter for the first time in two neuroimaging experiments with professionally trained (classical) singers. The first experiment aimed at identifying the brain networks involved in overt professional singing, and compared these results with the neural activation during imagined singing. Experienced singers were to sing excerpts of an Italian aria within a magnetic resonance scanner utilizing an advanced functional magnetic resonance imaging (fMRI) design that allows for undisturbed auditory feedback during singing. Subsequently, the same aria was sung entirely in their minds, without any overt production. Motor imagery has often been viewed as a window to cognitive motor processes and particularly to motor control. We were interested in detecting the possible differences or commonalities in neural networks involved in overt and imagined singing.

In the second study, we investigated the effects of expertise and training in classical singing on corresponding neural adaptation of the central nervous system. Three groups were measured in an fMRI scanner during overt singing, differing with respect to their singing skills: professional opera singers, vocal students and amateurs. To date, most of the neuroimaging work on motor control has addressed motor skill acquisition related to peripheral movement. Neuroimaging studies of professional instrumentalists demonstrated extensive experience-dependent neural adaptation (cortical expansion) related to the training of complex motor sequences and sensory feedback/control of the peripheral motor system. Yet, singing has been largely neglected and the neural consequences of vocal training are unknown.

The project of the present research was stimulated by the question of how the brain organizes and controls the complex vocal system during singing. A system that works largely autonomously when we speak, and that has already been trained extensively in verbal communication, but that requires a much higher level of control and additional mechanisms when one sings. The present work aims at identifying (a) the brain networks underlying vocal control during overt and imagined singing in voice experts and (b) at investigating how the brain adapts its functions with the development of higher levels of voice control (i.e. singing skills), which may also have ramifications for understanding disorders of vocal motor control.

Chapter 2 Fundamentals of voice physiology

The qualities of the human vocal system allow the production of highly complex phonological patterns in fluent speech by the age of 3 or 4 but also to create an artful combination of words and music in singing. I would like to begin with a brief overview on vocal physiology. At first I will outline the anatomy of the larynx, followed by a functional description of the vocal subsystems and their respective role in sound production. Finally, I will present some of the characteristics that distinguish singing from speaking and singers from non-singers.

2.1 Anatomy of the larynx

The larynx, the primary source of sound production, is suspended within a complex system of extrinsic laryngeal muscles attached to the oblique line of thyroid cartilage and the cricoid cartilage (Figure 1).

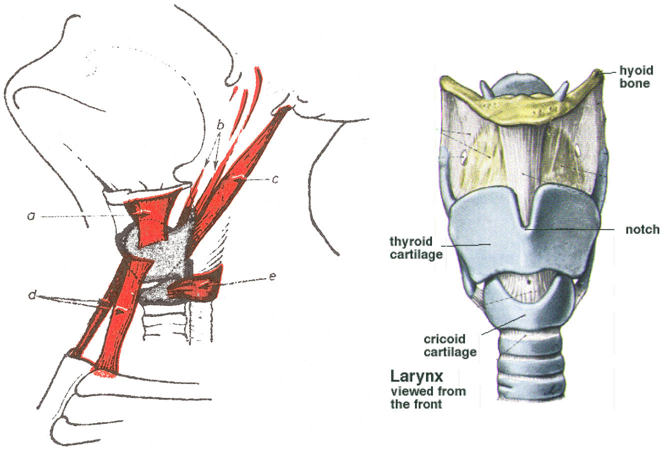

Figure 1 Left: Extrinsic larynx muscles. Muscles raising the thyroid cartilage are (a) m. thyrohyoideus, (b) m. palato-laryngeus, and (c) m. stylo-pharyngeus. Muscles lowering the larynx are (d) m. sterno-thyreoideus and (e) m. crico-pharyngeus. Right: Front view of the larynx. (adapted from Husler & Rodd-Marling, 1976)

One of their functions is to change the position of the larynx by raising and lowering the thyroid cartilage, thereby affecting vocal tract length (measured from the vocal folds to the lips), airway protection during swallowing but also intrinsic laryngeal muscle function (Ludlow, 2005). The intrinsic muscles of the larynx are specialized for vibration and airway protection reflexes (Figure 2). They are located within thyroid and cricoid cartilages. Most of these muscles have multiple compartments that play functionally distinct roles, which is also expressed by their respective muscle fiber type composition. The main functions of the intrinsic laryngeal muscles are airway protection, respiration, and phonation (i.e. the generation of sound). Traditionally, they are grouped analogously to their main actions into adductors, abductors, and tensors.

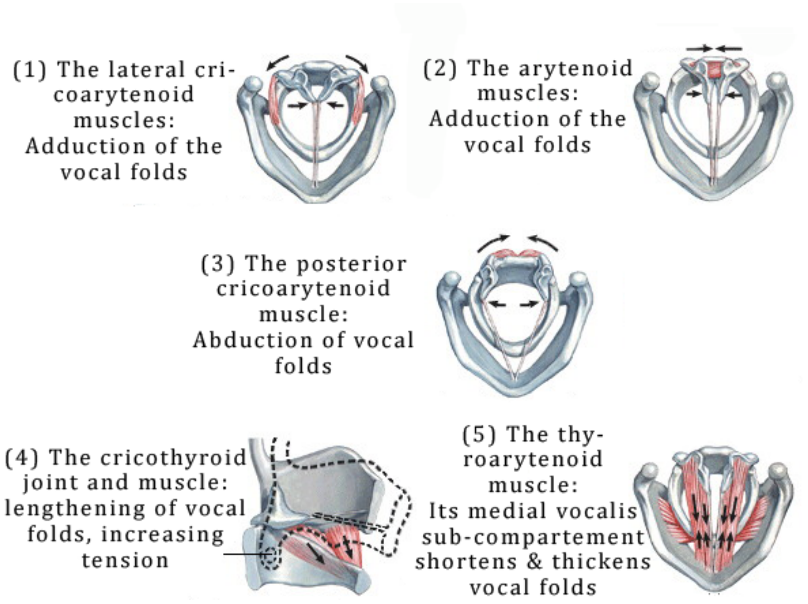

Figure 2: Intrinsic laryngeal muscle functions (adapted from http://webanatomy.net/anatomy/laryngeal_muscles.jpg).

The adducting muscles are comprised of (1) the lateral cricoarytenoid, which increases medial compression of the vocal folds and (2) the arytenoid (consisting of the transverse arytenoid that adducts vocal folds and the oblique arytenoid that pulls the apex medially). The only intrinsic muscle that abducts vocal folds is (3) the posterior cricoarytenoid – the antagonist of the lateral cricoarytenoid. There are two vocal tensors in the larynx consisting of (4) the cricothyroid and (5) the thyroarytenoid. The latter is the principal component of the vocal folds and plays a significant role in controlling voice quality, intensity, and pitch. It consists of two sub-compartments: the lateral muscularis and medial vocalis both having different attachments and different directions.

The vocalis muscle fibers run parallel and lateral to the vocal ligament, which is often referred to as the vocal cords. The vocalis not only affects phonation frequency of the vocal folds but also the natural oscillation of the vocal ligaments that cover the vocal folds (Titze, 1994).

Three pairs of nerves provide the motor and sensory innervation of the larynx: (1) the recurrent laryngeal nerve, (2) the external, and (3) the internal superior laryngeal nerve. All of them are branching off the vagus nerve (Hoh, 2005; Monfared, et al., 2001). As depicted in Figure 3, the superior laryngeal nerve descends into the carotid triangle where it divides into a smaller external and a larger internal branch, the latter of which is considered to carry mainly afferent (sensory) fibers from the mucosa of the larynx and trachea (Merkel, 1902). It may, however, also contain motor efferences to the arytenoid muscle for laryngeal protective reflexes (Sanders & Mu, 1998). The external branch descends along and deep to the superior thyroid artery to reach and supply the cricothyroid muscle. All other laryngeal muscles receive innervation from the terminal branch of the recurrent laryngeal nerve. A nerve plexus of extreme complexity is found in the thyroarytenoid muscle containing the vocalis muscle, which plays a major role in vocal fold vibration (Sanders, et al., 1993).

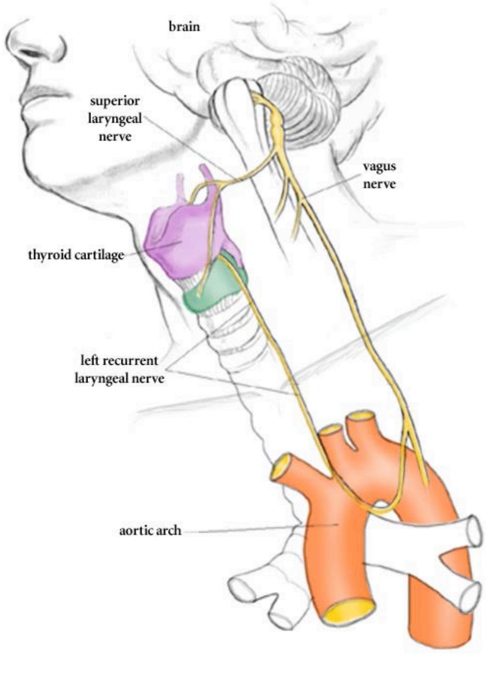

Figure 3: Innervation of the larynx (from Koufman, et al., 2009).

It is probably highly relevant for both voice production and airway protection, that the densest concentration of sensory receptors in the human body is found in the mucosa of the larynx (Sanders & Mu, 1998). The laryngeal mucous receptors are of vital importance for laryngeal reflexes during breathing and in the protection of the airway during coughing, swallowing and vomiting (Storey, 1968; Harding et al., 1978). However, these reflexes also contribute substantially to volitional motor control in voice production (Ludlow, 2004). The mechanoreceptors in the vocal folds mucous play therefore an important role in the motor management of intrinsic laryngeal muscles during singing and speaking (Wyke, 1974a, 1974b), where they act in concert with articular mechanoreceptors located in the fibrous capsules of the intercartilaginous joints of the larynx and with stretch-sensitive myotatic mechanoreceptors (muscle spindles) located in the intrinsic laryngeal muscles (Abo-el-Enein, 1966).

The peculiarities of the vocal mechanisms also demand for particular muscle properties. In fact, intrinsic laryngeal muscles are not only uniquely suited to their specialized role in phonation and articulation; they are also distinct from any other human skeletal and craniofacial muscle. Mammalian skeletal muscle fibers can be classified into functional types by the heavy chain (MyHC) and light chain (MyLC) isoforms of myosin, the primary motor protein that they contain. Myosin controls the kinetics of energy transduction and therefore the kinetic properties of muscle fibers (Hoh, 2005). Sciote et al. (2003) reported that laryngeal muscles possess a highly heterogeneous expression of MyHC that allows for variable shorting speeds that can be up to twice as fast as limb muscles. Fast-type MyHC fibers dominate the adducting twitch-muscles. In contrast, mixed type fast-slow MyHC fibers are largely present in abductor muscles (Shiotani, et al., 1999). Interestingly, functionally adequate differences were observed by Wu et al. (2000) within the sub-compartments of the thyroarythenoid muscle. As mentioned above, the thyroarythenoid muscle is a highly complex muscle that is crucially important in controlling voice quality, intensity, and pitch (Hirano, 1974; Titze, 1988). Its medial sub-compartment, the vocalis muscle, is the vibrating part of the vocal folds and the main influence on vocal fold behavior in speaking and singing. A high proportion of slow-twitch muscle fibers (STF) are found in the vocalis muscle. This allows for fatigue resistant, sustained and precisely controlled contractions (Han, et al., 1999). Only the human vocalis muscle is characterized by STF, with the highest expression of STF in the most mobile, medial part of the vibrating vocal folds (Han, et al., 1999), which indicates a special role of STF in providing appropriate muscular control for singing. Interestingly, the superior medial vocalis also contains a higher amount of muscle spindles (Sanders, et al., 1998). Muscle spindles provide the central nervous system with sensory information about the muscle, resulting in precise motor adjustments only in this part in which the sensory organ is present (English & Weeks, 1984). This co-occurrence of sensory muscle receptors and SFT along with the high density of mucosal mechanoreceptors points to a strong role of proprioception for voice control in singing (Han, et al., 1999).

With respect to the evolution of the human vocal system, only comparative methods are available. When comparing vocal function across all primates, the essentials of voice production are similar. Yet the existing difference between human speech and non-human vocalization can be directly attributed to two anatomical differences that occurred during the course of evolution. The descended larynx and the increased thoracic innervation (Ghazanfar & Rendall, 2008). The deeper position of the human larynx in the vocal tract compared to non-human primates results in an enlarged pharyngeal cavity in addition to the oral cavity. This two-tube configuration coupled with an agile tongue and rapid lip movements allows the precise shaping of dynamic resonances, which enables the latitude of articulatory movements in humans and affects the resonances defining the components of language and the sound properties of the singing voice. The increased innvervation of the thorax, on the other hand, is directly linked to an increased control of air-pressure from the respiratory system. This exquisite control over breathing is required for speaking or singing long sentences without disruptive pauses for inhalation. A third component that reflects evolutionary differences between human and non-human primate vocalization is the neo-cortical control of voice production in humans, with a widely distributed network of speech functions in the cortex that is extended in singing.

2.2 Functional anatomy of the vocal system

Voice occurs through an interaction of three components: a power source, an oscillator, and a resonator (Figure 4): the lungs represent the power source, providing the vocal folds with energy. Air stream control is defined by action of the diaphragm, the chest muscles, and abdominal muscles. The sound source is generated at the larynx, or more precisely, by the vocal folds.

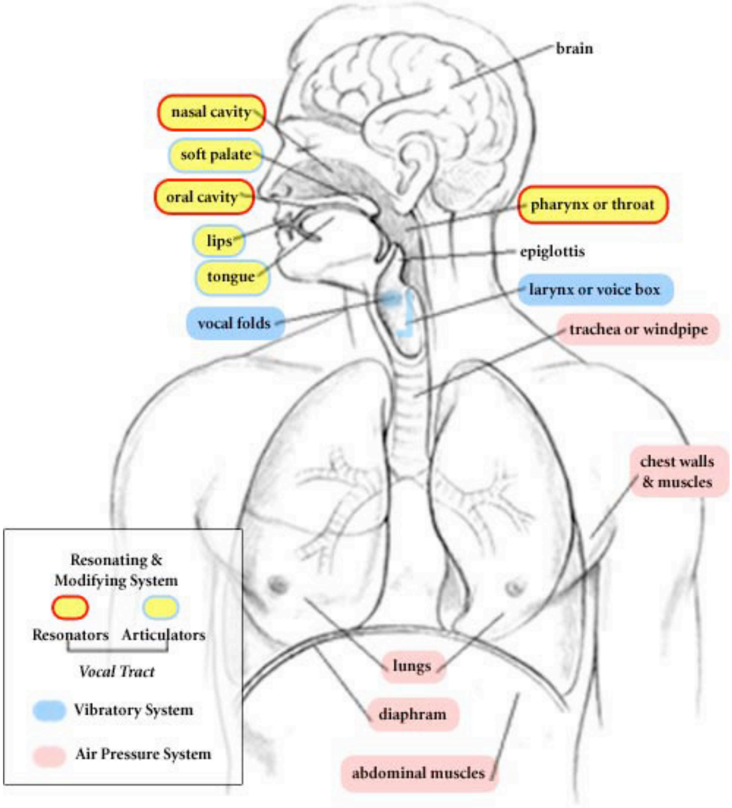

Figure 4: Diagram of the vocal subsystems involved in voice production (from Koufman, et al., 2009)

A complex interaction of vocal fold muscular forces with aerodynamic forces causes the vocal folds to oscillate. The subsequent resonating system modifies the sound as it travels further from its source. It is comprised of the throat (pharynx), the oral cavity and the nasal cavity. This “vocal tract” is constantly changed through activity of the articulators (Titze, 1994). The interaction of aerodynamic forces with laryngeal muscular forces is described by the aerodynamic-myoelastic theory of voice production (Titze, 1980). That is, airflow from the lungs sets the vocal folds into vibration and the mass and tension of the vocal folds, controlled by intrinsic laryngeal muscles, determines the quality and quantity of the sound produced. These factors work in opposite directions; increases in mass lower vibration frequency while increases in tension raise frequency. When the vocal folds vibrate, they create a series of high-speed air puffs that convert air pressure into sound waves. A particular characteristic of vocal fold vibration is that they do not vibrate as a single unit (Figure 5). Instead, each of the inferior and superior masses of the vocalis muscle vibrate slightly out of phase (Adachi & Yu, 2005).

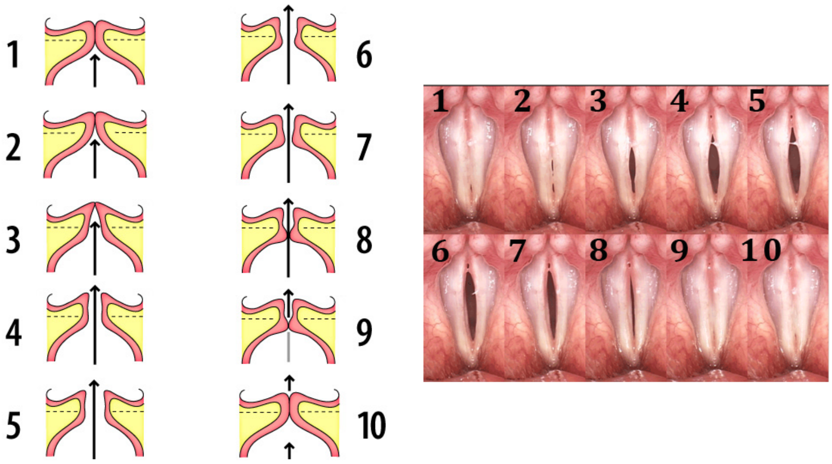

Figure 5: Left image: Diagram of a vocal fold vibration cycle seen from the front. Right image: high-speed stroboscope images of vocal folds in motion during one vibration cycle, viewed from the top. 1. Air pressure moves from the lungs towards the closed vocal folds. 2-3:Air pressure opens bottom of vibrating layers (mucosal wave: orange; muscle: yellow). 4-5: Air pressure opens the top of vocal folds. 6-10: The low pressure behind the fast moving air column causes the bottom to close (Bernoulli effect), followed by the top layer. 10: Closure of the vocal folds cuts off the air column and releases a pulse of air. The inferior and superior vocalis muscles support the two masses seen in vocal fold vibration. (Left image from (Titze, 1980), right image from (Kayelemetrics, 2009))

The mucous membrane, which covers the vocal folds, plays a crucial role during this process. In fact, the voice quality depends critically upon the mucous. For example, alterations in the flexibility of the mucous cause the well-known difficulties in speaking and singing during an infection. The vocal folds are characterized by three layers: epithelium, lamina propria and vocalis muscle (Figure 6).

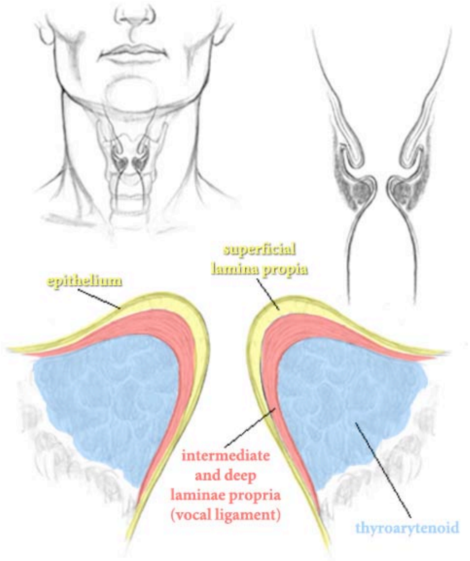

Figure 6: Cross-section of vocal folds layers (from Koufman, et al., 2009).

The mucous consist of the epithelium and the superficial lamina propria. The vocal ligament (i.e. the vocal cord) beneath is composed of the intermediate lamina propria and the deep lamina propria. The superficial mucous has the greatest degree of elasticity and is passively involved in vibration during phonation. This passive involvement is explained by the Bernoulli (1783) effect. That is, when the vocal folds are adducted during phonation, air pressure begins to build up below the vocal folds. When the pressure is high enough, the soft tissues of the vocal folds are forced to separate and the airstream flows through the vocal folds. As the airstream accelerates, a drop in pressure follows it, which forces the vocal folds back together. This airstream-driven wave sustains vibration and takes part in the production of higher frequencies of the sounds spectrum. If several opening and closing periods are sustained at steady time intervals, a sound wave of a certain frequency is generated. This vocal fold oscillation frequency determines the fundamental frequency, which characterizes the pitch level. For example, when a soprano sings the musical note A5’ with a fundamental frequency of 880 Hz, the vocal folds open and close 880 times per second. In addition, a whole spectrum of upper harmonic partials is produced and modulated through the ligament's tensile stress, which rises rapidly with elongation of the vocal folds.

In conclusion, the interaction of air-pressure from the lungs and adjustments of intrinsic laryngeal muscles determine the three main characteristics of the resulting sound: (1) pitch level, (2) loudness, and (3) primary sound spectrum. Loudness is chiefly related to increased sub glottal air pressure while the lengthening and shortening of laryngeal muscles determine the primary control of pitch. Vocal fold stiffness modulates the wave of the mucosal cover. Importantly, these parameters are not independent and must be tightly controlled in singing (Sundberg, 1987).

The third functional component of voice production is the resonating vocal tract, comprised of the throat (pharynx), the oral cavity and the nasal cavity (Sundberg, 1987; Titze, 1994). In general, a resonator amplifies and reinforces the frequencies that the sound source produces. In musical instruments, their size and shape mainly determine the resonance properties. Yet the resonating vocal tract, which acts as a source filter of the primary sound from the larynx, is constantly modulated by activity of the articulators (Figure 7A). As a consequence, changes affecting the vocal tract cavity (e.g. the position of the tongue, the area of opening of the mouth, the position of the larynx) will also retune the resonance characteristics. Resonance phenomena are often described by “formants”; strong peaks in the harmonic sound spectrum that consist of closely grouped energy-rich harmonic partials. The relationship between the first two formants determines the vowel we hear (Figure 7B), whereas the opening of the jaw determines the frequency level of the first formant, the tongue shape the second and the tip of the tongue the third (Hall, 1997).

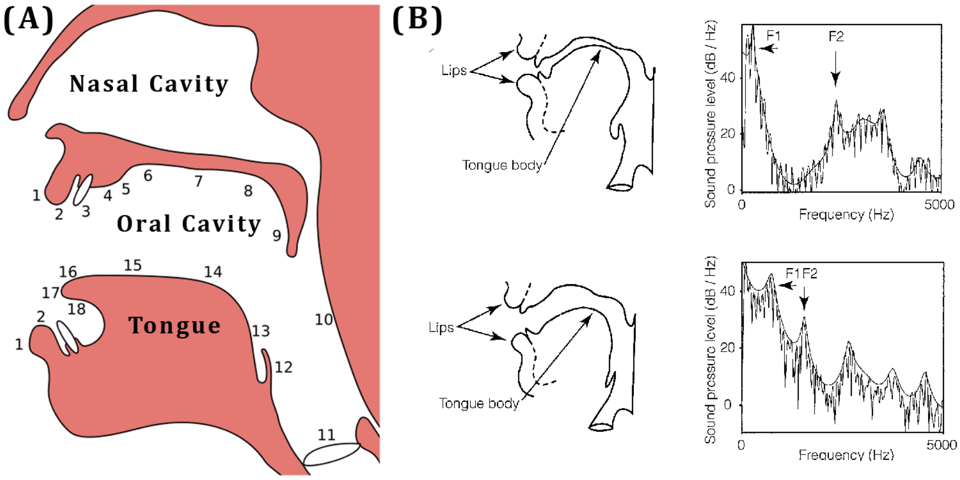

Figure 7: (A) The resonating vocal tract and places of articulation (passive & active): 1. Exo-labial, 2. Endo-labial, 3. Dental, 4. Alveolar, 5. Post-alveolar, 6. Pre-palatal, 7. Palatal, 8. Velar, 9. Uvular, 10. Pharyngeal, 11. Glottal, 12. Epiglottal, 13. Radical, 14. Postero-dorsal, 15. Antero-dorsal, 16. Laminal, 17. Apical, 18. Sub-apical (from http://en.wikipedia.org/ wiki/Places_of_articulation). (B) Examples of vocal tract configurations (left) and resulting frequency spectra (right) for the vowels /i/ (in "beat") and /ae/ (in "bat") in the top and bottom row respectively. Arrows indicate the first two formant peaks in the spectrograms (F1 & F2). The relationship between the first two formants determines which vowel we hear (adapted from Patel, 2008).

Trained singers, however, require a control over vocal tract resonances that reaches far beyond speech intelligibility, affecting voice quality, timbre, as well as loudness. Traditionally, the linear theory of vocal acoustics assumes that the source of sound (larynx) and the resonator of sound (vocal tract) work independently (Sundberg, 1987). However, this poses problems considering the limitations of vocal tract size. A recent non-linear source filter theory proposes that there is an interaction between the sound source and the resonances of the vocal tract (Titze, 2008a). Because the vocal tract is only about 17 cm long, reinforcing a whole cluster of harmonics simultaneously using an energy-feedback process would increase its efficiency dramatically. This effect raises the inert reactance of the vocal tract, which in turn enhances the driving pressures of the vocal folds and the glottal flow, thereby increasing the energy level at the source. When singers create the appropriate conditions in their vocal tract, they ‘provide an extra, precisely timed "kick" to each cyclic opening and closing of the folds that reinforces their vibration to create stronger sound waves. A singer's task is to adjust the shape of the vocal tract (by carefully selecting favorable "singing" vowels) so that the inert reactance is experienced over most of the pitch range’ (Titze, 2008b). This process takes many years before mastery is achieved.

2.3 Trained and untrained voices

Up to now, numerous studies have demonstrated functional and acoustical differences between speaking and singing and between trained and untrained singing voices. Singing compared to speaking requires enhanced control over air-pressure, more efficient vocal fold vibration, and distinct vocal tract resonance characteristics (Sundberg, 1987). Since a complete review of the literature is beyond the scope of this chapter, I will briefly mention two characteristic differences between trained and untrained voices. It is obvious to listeners that classical singers can produce higher sound pressure levels (i.e. a louder voice) than non-singers. However, what we cannot hear is that singers achieve this with less sub-glottal air pressure and lower muscle tension than amateurs, which reflects the increased efficacy of vocal fold vibrations in professionals (Koufman, 1996). This study showed that classical singers had lower muscle tension scores than non-classical singers, with the lowest muscle tension scores being seen in those singing choral music (41%), art song (47%), and opera (57%), and the highest being seen in those singing jazz/pop (65%), musical theater (74%), bluegrass/country and western (86%), and rock/gospel (94%). With respect to vocal tract resonances, skilled classical singing was also found to result in more pronounced higher frequencies in opera singers compared to speaking and to other forms of singing (Sundberg, 1987). In particular, classically trained singers show a prominent amplitude in their sound spectrum in the range of 3 kHz, referred to as the "singer's formant". This formant results from a clustering of formants 3, 4, and 5 because of articulatory fine-tuning (Sundberg, 1974, 2001).

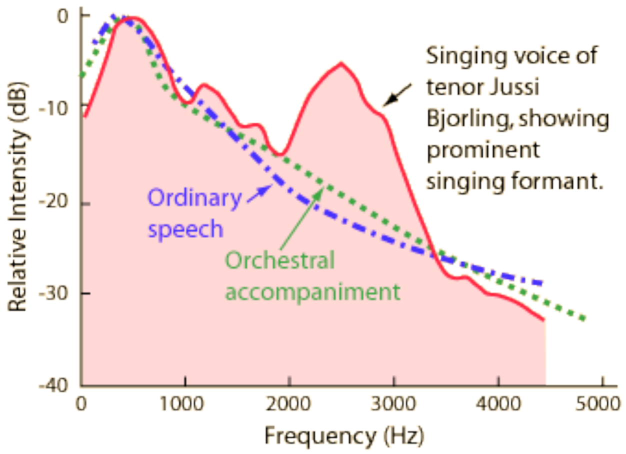

Figure 8: The effect of the “singing formant” on the audibility of the voice despite competing sound from an orchestra (adapted from Sundberg, 1977)

Such resonance phenomena have several advantages (Sundberg, 1994): They contribute to voice quality and have significant effects on the audibility of the singing voice (Figure 8). In Chapter 5 , we will demonstrate for the first time in a neuroimaging experiment the neural correlates associated with both overt and imagined singing of an Italian aria in professionally trained singers.

Chapter 3

Expertise and deliberate practice

Musicians represent a highly select group that differs most apparently from non-musicians in their exceptional skill level and motor dexterity. Within this group, there is only a very small sample of excellent musicians, who passed the entrance qualifications for music conservatories, trained extensively for many years and work eventually as professional musicians.

An often-heard characterization of musicians refers to their allegedly natural giftedness or talent for music (for an overview, see McPherson & Williamon, 2006). The term giftedness corresponds to innate capacities that are distinctly above average in at least one aptitude domain, involving intellectual, creative, socio-affective and sensorimotor abilities. The individual blending of these characteristics is thought to determine a predisposition for musical achievement. Musical talent, on the other hand, refers to superior performance in a specific domain of musical activity (e.g. performing, improvising or composing) and becomes evident through behavior and training (Gagné, 1985, 2000). Thus, a multitude of factors seems to contribute to musical achievement, ranging from genetic components for musical ability (Simonton, 1999), gene-environment interaction (Scarr & McCartney, 1983), and hormonal differences (Hassler, 2000), to intrapersonal characteristics (e.g. motivation, volition, personality and physique), environmental factors (e.g. parental support) and chance (McPherson & Williamon, 2006). However, while the distribution of these aspects may vary between musicians, it is evident that musical talent is more than just being naturally inclined or disposed for learning an instrument. In fact, no amount of musical aptitude will guarantee success as professional musician without intense, systematic learning and practice (Howe, et al., 1998). Therefore, exceptional skill is thought to be predominantly related to the amount of deliberate practice, which refers to a ‘highly structured activity with the explicit goal of improving some aspect of performance’ (Jorgensen, 2001).

Expertise research convincingly demonstrated evidence for a close connection between accumulated musical practice and overall achievement in both young and older musicians (Sloboda, et al., 1996; Williamon & Valentine, 2000). The "monotonic benefit assumption" states that 'the amount of time an individual is engaged in deliberate practice is monotonically related to that individual's acquired performance' (Ericsson, et al., 1993, p.368). Deliberate practice is thus thought as being a direct cause of musical achievement (Howe, et al., 1998). This route to expertise generally involves four distinct stages: (1) a period of play and fun with music, (2) a period of deliberate practice with high assistance of family and teachers, (3) an intensification of deliberate practice with acknowledged experts in a music conservatory, whereby music becomes the core of life planning, and (4) a stage after graduation, where the level of expertise shows through outstanding performance (see Jorgensen, 2001). The amount of training required before a professional career can be started is stunning and involves more than 10 000 hours of dedicated practice over a period of approximately 10 years. Young pianists in a study by Sosniak (1985) even required an average of 17 years of training before they started their careers as concert soloists, having accumulated over 60 000 hours of dedicated practice towards the end of their careers (Krampe, 1997).

According to Jorgensen (2002), the assessment of deliberate practice is based on:

1. The initial starting age when formal lessons began.
2. The accumulated amount of practice from initial starting age with formal training to present situation.
3. The amount of practice at one particular time or during a limited period of time.

The majority of musicians who have reached high levels of expertise also started early, in the order of 4-6 years, and therefore accumulated also a large amount of practice (Jorgensen, 2002). Yet, not all musicians are early starters as Jorgensen (2001) observed. He found in a study on music conservatory students that piano and string players started earliest (7.8 and 7.3 years respectively), followed by brass (11.6 years) and woodwind (12.8 years) players, while vocal students were comparably late beginners (14.4 years on average). However, the generally positive relationship between starting age and achievement was confirmed for the whole music student population under investigation. Similar results were found for instrumentalists with respect to the amount of weekly practice and the number of practice days in a week but not for vocal students, who practiced only half of the time than instrumentalists (Jorgenson, 2002). A nevertheless existing difference in performance levels between singers had thus been attributed to the overall accumulated practice time.

Professional musicians have increasingly become the focus of cognitive neuroscience, mainly because of the large amount of time spent with the training and development of highly specific motor skills. This is what makes musicians a particularly attractive study group for investigating learning dependent plasticity of the brain. Over the last 15 years, several neuroscientific studies have demonstrated that extensive training with a musical instrument can result in corresponding functional and structural adaptation of the brain (for a review, see Munte, et al., 2002). Interestingly, the magnitude of change was usually related to the age at which musical training commenced, which points towards a role of accumulated motor skill practice in neural plasticity. However, these studies have entirely focused on the peripheral motor system. Up till now, evidence for neural adaptation in conjunction with vocal motor training is sparse and mainly derived from clinical rehabilitation studies. Changes in neural correlates with respect to experience and training levels in classical singing have not yet been reported. In Chapter 6 we will describe a functional magnetic resonance imaging experiment with professional opera singers, vocal students, and amateurs during overt singing of an Italian aria. We will provide the first evidence that accumulated training and performance of classical singing skills results in distinct and functionally related neural reorganization of the brain.

Chapter 4

Principles of functional magnetic resonance imaging (fMRI)

4.1 Introduction

In this chapter, I will give a brief introduction into the principles of functional magnetic resonance imaging, which was the method we employed in the experiments of this dissertation. Magnetic resonance imaging (MRI) is a widely used non-invasive tool for clinical diagnosis that produces high quality images of the soft tissues of the human body. Unlike Computer Tomography, it does not involve the use of ionising radiation. Whereas MRI is useful for anatomical and morphometric investigations, technical advances in the 1980s and 1990s resulted in the development of a functional MRI (fMRI) technique that allowed for in-vivo measurements of brain activity.

MRI is based on the nuclear magnetic resonance (NMR) phenomenon. Atomic nuclei with an odd number of protons and neutrons possess specific quantum mechanical magnetic properties called magnetic spin. If placed in an externally applied magnetic field, these nuclei align accordingly and absorb radiofrequency energy, which they re-emit if transferred back to their original state. Described and measured in molecular beams by Isidor Rabi in 1938, the first successful nuclear magnetic resonance experiment in condensed matter was carried out in the laboratories of Felix Bloch and Edward Purcell in 1946, for which they shared the Nobel Prize in physics in 1952. It took another 27 years before Lauterbur (1973), and Mansfield and Granell (1973) independently proposed that the NMR can be used to obtain tomographic images of objects. Paul Lauterbur was the first to illustrate the process of acquiring 2D and 3D images by NMR, known as magnetic resonance imaging (MRI). Both Mansfield and Lauterbur received the Nobel Prize for their discoveries related to MRI in 2003. Another MRI pioneer was Raymon Damadian, who discovered that normal and cancerous tissue can be distinguished based on the different response signals (i.e. relaxation times) they emit (Damadian, et al., 1974), which eventually resulted in the application of MRI in clinical diagnosis. Since the 1980s, MRI technique has developed rapidly and became an indispensible tool for uncovering the involvement of the nervous system in cognitive, emotional, perceptual and motor processes.

4.2 The physics of MRI

The strength of a magnetic spin differs with the type of nucleus. The hydrogen nucleus, which is composed of a single proton, has the strongest magnetic moment. In addition, it has also the highest biological abundance in humans. Fat and water, the major hydrogen containing components, are widely distributed in the body. MRI takes advantage of these facts and uses the NMR signal from the hydrogen nuclei for constructing high-resolution 3D images with excellent soft-tissue contrasts.

Generally, the magnetic spin of a proton has no particular orientation and therefore does not produce an overall magnetization (Figure 9A). Yet, if exposed to an externally applied magnetic field the hydrogen nucleus is restricted to adopt two principle quantum states that differ energetically. That is, the protons are either oriented anti-parallel to the external magnetic field (requiring a higher energetic state) or they are assuming an orientation in parallel with the external magnetic field (requiring a lower energetic state). A majority of protons, however, will align themselves in the direction of the applied magnetic field, thereby producing a net polarization (or magnetization) parallel to the external field (B_0, see Figure 9B). The higher the field strengths are (e.g. 1.5 or 3.0 Tesla) the higher the number of protons aligning with the external field, which enhances the sensitivity of the technique. Because of the net magnetization (M) in the XY plane, each spin will rotate (precess) in a cone-like fashion about the Z-axis of the externally applied magnetic field B_0 (Figure 9C). Described by the Larmor equation, the precession frequency of the individual proton behaves proportional to the strength of the external magnetic field (B_0) according to the gyromagnetic ratio.

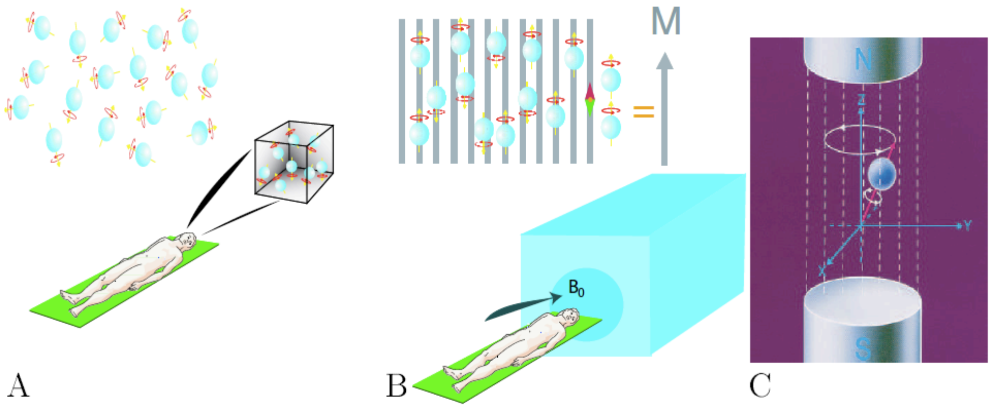

Figure 9: (A) Without an externally applied magnetic field, the magnetic spin has now particular orientation. (B) When a strong external magnetic field is applied, the majority of hydrogen protons align with the external field. While protons are forced to precess in the direction of B_0 to lower their energy content, they produce the net magnetisation 'M' (C) The proton precesses in a cone-like fashion about the Z axis in the magnetic field (adapted from Mathiak 2002).

For a hydrogen nucleus, the gyromagnetic ratio is 4257 Hz/Gauss. Thus at 1.5 Tesla (15,000 Gauss), the precession frequency is 63.855 Megahertz, which is in the radio frequency (RF) range. This relation is of central importance for magnetic resonance imaging. A further step required for detecting a signal involves the introduction of a second magnetic field referred to as B_1. This new field must follow two important rules: (1) it has to be applied perpendicular to B_0, and (2) it has to be at the resonant frequency of the precessing protons. An RF pulse (B_1), applied with an appropriate RF coil, will aggregate the protons to a higher energy state by forcing their spin orientation out of alignment with B_0 and towards the direction of the coil (see Figure 10). That is, the RF pulse tips the orientation of the proton out of the longitudinal Z plane and towards the transverse XY plane of B_1. The length and intensity of this pulse determine the flip angle of the spins. Once the RF pulse is turned off, the protons will return to their original equilibrium (low energy) state in the direction of the z-axis (B_0), thereby releasing the absorbed radio frequency.

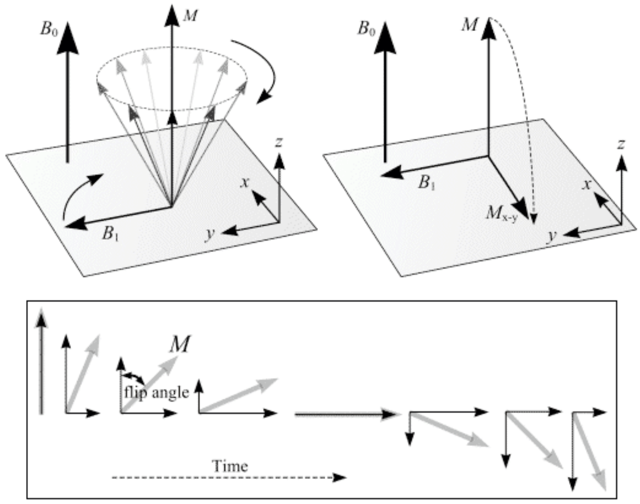

Figure 10: (top) Applying an RF pulse (B_1) on the net magnetisation M creates a second magnetic field Mx-y. M is flipped from its original along the direction of the external magnetic field B_0, into the transverse x-y plane along the direction of the second magnetic field B_1. The protons aggregate to a higher energetic state. The bottom image illustrates the flip angle, which is determined by the strength and duration of the RF pulse (from http:// www.easymeasure.co.uk/principlesmri.aspx).

This process is called ‘relaxation’ and can be detected by the MRI coils. The time constant that describes the return to an equilibrium state is called the spin lattice relaxation time (T1). More exactly, T1 is the time taken for approximately 63% of the initial longitudinal (B_0) magnetization to be restored following a resonant RF pulse. In contrast, T2 relaxation (also known as T2 decay) is the decrease in the transverse x-y (or B_1) component of magnetization, which does not involve the emission of energy. Because the T1 and T2 relaxation times vary with the type of body tissue, they deliver excellent tissue contrasts. For the final spatial encoding step of the position of the spins, magnetic field gradients must be superimposed on the otherwise homogeneous external magnetic field (B_0). The principle role of the gradient coils is to produce linear changes in the magnetic field in the x, y and z directions. Three magnetic field gradients, placed orthogonally to one another inside the bore of the magnet, are required to encode information in three dimensions. When the gradients are switched on, some columns of spins perpendicular to the axis precess at slightly different Larmor frequencies depending on the specific gradient strength. This allows exciting only a narrow plane of protons with the correct resonant RF pulse while all other protons in the sample receive the wrong frequency of excitation for resonance to occur.

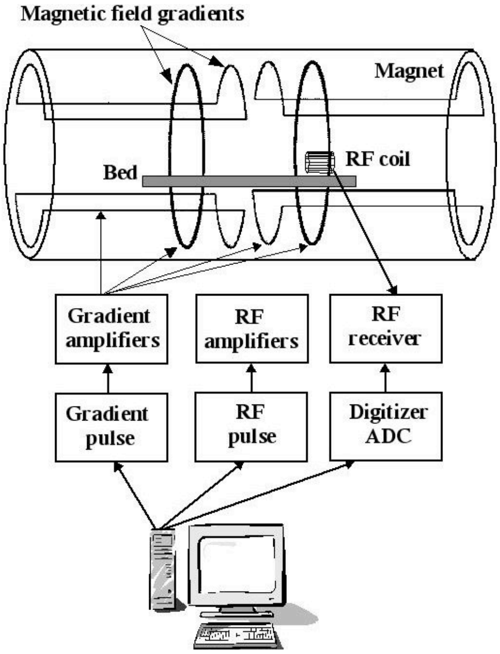

Figure 11: Diagram of a whole-body magnetic resonance tomography.

By employing time constant maps based on T1 or T2 relaxation, encoding of spatial information and reconstruction of anatomic images is made possible. A summary of the processes involved is depicted in Figure 11.

4.3 Functional MRI (fMRI)

Functional MRI is based on the magnetic resonance principles described above. It works by detecting levels of oxygen throughout the brain. Since its development in the early 1990s by Seiji Ogawa (Ogawa, et al., 1990; Ogawa, et al., 1993), it has become one of the leading research tools for mapping functional brain activity. Basically, the fMRI takes advantage of two principles: (1) the interactions of cerebral blood flow (CBF), energy demand, and neural activity, also known as neurovascular coupling; and (2) that oxygenated and deoxygenated blood have different magnetic properties.

The brain constantly demands adequate supply with oxygen and glucose, delivered with the cerebral blood flow through a complex mechanism. If neural activation occurs, an energy-consuming metabolic process is launched that is followed by enhanced substrate delivery (oxygen and glucose) to the activated brain area. This hemodynamic response occurs as a direct cause of regional neural activity (Logothetis & Pfeuffer, 2004).

The hemodynamic activity is mainly correlated with neurotransmitter release and synaptic field potentials, reflecting local signaling, although the exact relation is not yet fully understood (Logothetis, 2003). Following the increased blood flow, different concentrations of oxygenated and deoxygenated venous blood emerge with respect to the baseline state (Figure 12). Deoxygenated blood is slightly paramagnetic relative to brain tissue and causes distortions in the magnetic field around vessels. Oxygenated blood, on the other hand, is isomagnetic and does not cause distortions. fMRI detects these inhomogeneities and measures what is know as the "blood oxygenation level dependent" (BOLD) signal. For capturing the BOLD signal, a fast imaging method is required that allows the collection of whole brain data in a few seconds or less. Echo planar imaging (EPI) is such a method, which enables the formation of an image in 30-100ms at the expense of spatial resolution (Stehling, et al., 1991). It is based on the T2 decay of transverse magnetization but measures inhomogeneities related to changes in the oxygenation of venous blood (T2*). The correlation of the signal time course in each volume element (voxel) with the known time course of the task makes it possible to identify those voxels in the brain that show changes associated with the brain function under consideration.

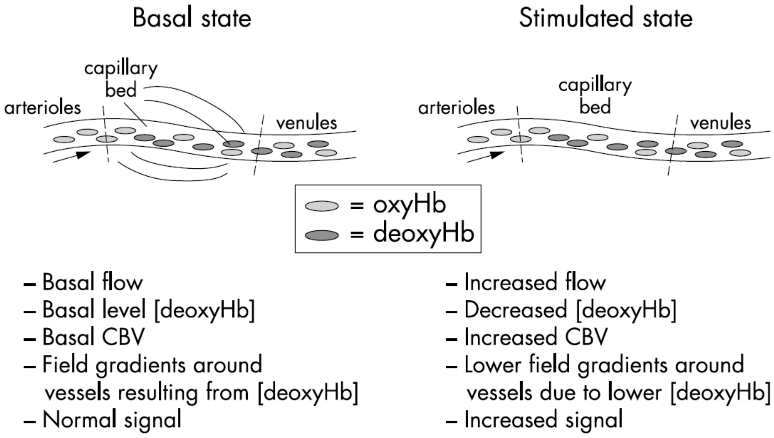

Figure 12: This diagram shows the change in blood desoxygenation during neuronal activity (Matthews & Jezzard, 2004). In the basal state, deoxygenated blood (deoxyHb) causes field inhomogeneities around the blood vessels and reduces BOLD signal. In the activated (or stimulated) state, neurons require increased oxygen and glucose supply causing an increased cerebral blood volume (CBV). The increased cerebral blood flow overcompensates for the increase in oxygen and ultimately delivers an oversupply of oxygenated blood. As a consequence, the relative amount of deoxyHb decreases whereas the BOLD signal increases.

This requires an exact modeling of the hemodynamic time course. The hemodynamic response (HDR) typically reaches its plateau between 6-12 seconds before it falls back to baseline level (Figure 13). BOLD responses are modeled for statistical analysis by means of a transfer function, called the Hemodynamic Response Function (HRF; Friston, 2002). The HRF represents the theoretical signal the BOLD fMRI would detect “in response to a single very short stimulus of unit intensity” (Marrelec, et al., 2003). Importantly, BOLD fMRI represents only a relative measure of signal change compared to a baseline condition. These signal changes are usually very small, in the order of 2-4% for motor tasks, and much smaller for cognitive tasks. Therefore, great efforts must be undertaken to control for possible artifacts, and advanced statistical methods must be employed for reliably localizing functional brain activity.

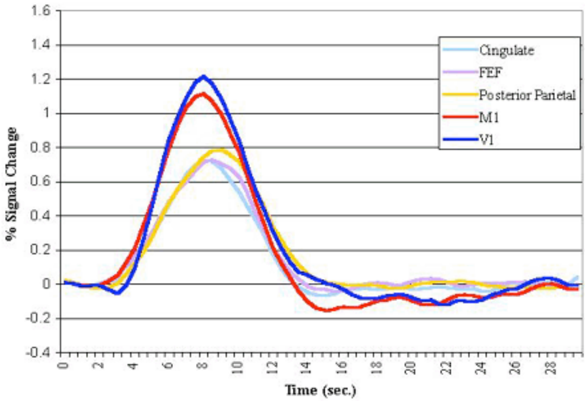

Figure 13: Group average hemodynamic response curves for different brain structures. It has been previously assumed that the vascular response remains constant throughout the human brain and across subjects. However, studies provided evidence for substantial variability in the shape of responses across subjects and between areas. The typical response however reaches its plateau between 6-12 seconds, before falling back to baseline. A precisely modeled HDR is critical for detecting activated brain networks in neuroimaging experiments.

4.4 Experimental designs of fMRI studies

Functional MRI is a powerful method for detecting and localizing brain responses to numerous stimuli and various active tasks. The careful selection of an experimental design that suits the respective stimulus or tasks under investigation is of crucial importance. Considerations must take into account the temporal dynamics of the HDR, the spatial and temporal resolution of the sequence employed, the area of brain coverage, and possible artifacts. Paradigms that require auditory processing must additionally consider interfering effects with the loud scanner noise. Based on the temporal dynamics of the BOLD signal, the two main strategies for acquiring images in fMRI experiments are “block” and “event-related” designs.

An fMRI experiment typically consists of several scanning sessions, during which the subjects’ brain responses to stimuli are repeatedly measured. In block designs, these sessions are segregated in relatively long blocks (30-60 seconds), where stimuli belonging to the same experimental condition are presented in close succession (Figure 14A). In its simplest form, there are only two conditions (AB block-design). One active condition and one baseline (or rest) condition alternate with each other. Each measured block contains the averaged activation of all items belonging to the same category (e.g. A or B). The advantage of this approach is the relative robustness of results, increased statistical power and high BOLD signal change ratios compared to baseline (Amaro & Barker, 2006). However, not all task types may be appropriate for this paradigm, such as oddball experiments for example. In contrast, the event-related approach allows an individual analysis of responses to single stimuli or trials belonging to different experimental conditions (Figure 14B). Each of these randomly presented stimuli are preceded and followed by baseline conditions lasting for 6-12 seconds. The great advantage of this designs is that it avoids habituation as well as item prediction and that it allows distinct analyses related to individual responses (e.g. emotional judgments or behavioral errors) through the detection of transient variations in hemodynamic responses to single short stimuli (Amaro & Barker, 2006). Those voxels consistently activated within a distinct brain response pattern are detected as event related activation. Therefore, each trial onset must be convolved with a modeled HRF. In contrast, block designs assess effects of interest using boxcar regressors convolved with a synthetic HRF. The disadvantage of event-related paradigms is their lower signal to noise ratio, resulting in longer acquisition times.

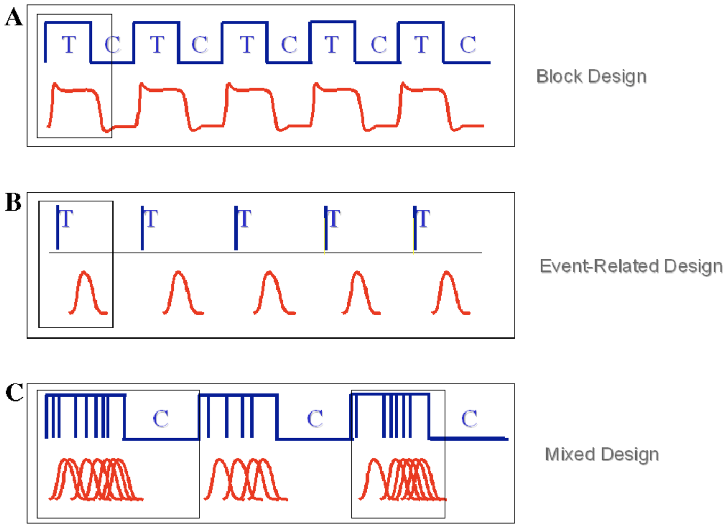

Figure 14: Experimental designs in fMRI studies. (A) Block design: stimulus of the same condition are presented subsequently, the BOLD response is actually composed from individual HRFs from each stimulus, and is generally of higher magnitude. (B) Event-related design: each stimulus' HRF is detected, and may be analyzed in detail; (C) mixed design, a combination of events closely presented alternating with a series of control condition.

A variant of the event-related design is also known as "sparse-sampling", since scanning is paused during task execution. Image acquisition is only sampled after task completion, taking advantage of the time course of the hemodynamic response. This approach is increasingly used in studies requiring undisturbed auditory processing (Krumhansl, 2003). Mixed designs consisting of block-sampling and sparse-sampling features are sometimes employed to combine the advantages of both (Figure 14C). That is, a series of event-related image acquisitions involve the consecutive presentation of same category tasks, alternating with a series of baseline conditions. Because of the consecutive presentation of similar tasks a complete recovery of the hemodynamic response is avoided, which increases signal to noise ratio and statistical power.

4.5 Analysis of fMRI data

The main purpose of the statistical analysis in functional imaging is to identify signal changes in voxels that vary with the changing brain states of interest across the serially acquired images. A multitude of software packages is available from different labs for carrying out such analyses (e.g., FSL at http://www.fmrib.ox.ac.uk/fsl; XBAM at http://www.brainmap.co.uk; Brain Voyager at http://www.brainvoyager.de; SPM at http://www.Wl. ion.usl.ac.uk/spm; AFNI at http://afni.nimh.nih.gov/afni). Common to most of them are two major steps involved in data analysis: (1) spatial and temporal pre-processing, including movement correction, normalization to a canonical space and spatial smoothing, and (2) statistical inference, performed via parametric or non-parametric statistical approaches (Figure 15).

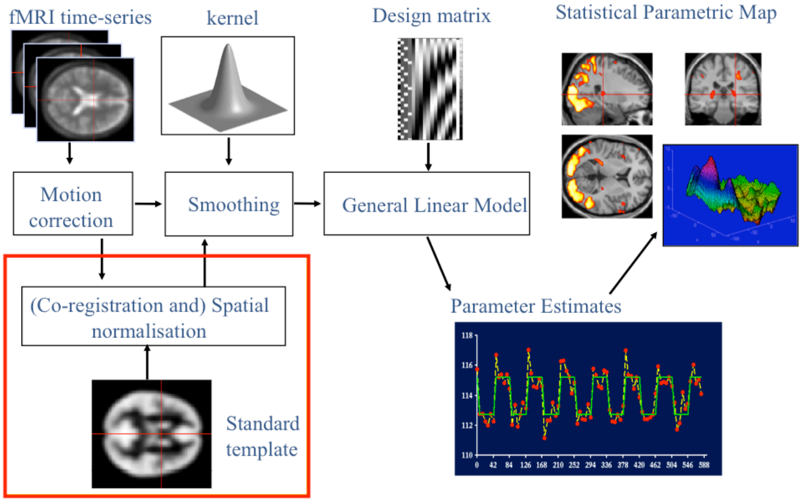

Figure 15: Overview of data pre-processing and statistical analysis with the “Statistical Parametric Mapping” software package (Friston, 2007). The first step requires a series of pre-processing operations, aiming at increasing data quality and statistical inference by reducing motion artifacts, co-registering functional and structural images, normalizing images to a standard brain template, and smoothing data to increase signal to noise ratio. The second step involves statistical modeling of data followed by the calculation of probability levels for each activated voxel and comparisons between subject groups or task modalities.

In this chapter, I will briefly review the steps involved in fMRI data analysis with the “Statistical Parametrical Mapping” (SPM) software package developed at the Wellcome Department of Imaging Neuroscience (London, UK), which was employed in this dissertation.

4.5.1 Spatial and temporal processing

4.5.1.1 Motion correction (Realignment)

Statistical inference depends on the accuracy of locating activated brain voxels. This process is complicated by data from different functional scans or brains of different sizes and shapes that must be properly aligned (or registered). If the images are not aligned, then the statistical analysis at a voxel will be from different locations in the brain across the time series. A significantly confounding factor in fMRI studies is its sensitivity to motion artifacts, originating from whole head movement or from pulsations within the brain due to respiratory or cardiac cycles. Most of the movement happens between consecutive series of images. However, head movement during image acquisition may not only "dislocate" the brain with respect to the previously acquired image in the series, it also alters the spin-magnetization affecting the signal intensity – i.e., some regions become brighter. Although experimenters take measures to restrict head movement during image acquisition, the first step in fMRI analysis always requires the post-hoc realignment of brain volumes using automated algorithms to minimize the difference between subsequent images. In SPM, the realignment algorithms are based on the estimation of six movement parameters (Figure 16) that describe movement along three translations (i.e. differences in mm along the x, y, and z-axis) and three rotations (i.e. the degree of pitch, roll, and yaw). The model uses these parameters to align all of the images in a time series to the first functional image acquired, using a least squares approach. In a second step, these parameters are incorporated as confounds in the statistical model to improve statistical inference. The degree to which motion affects the result of an activation study depends to a certain extent upon the type of experimental design that is employed. In general, event-related related designs are less sensitive to head-movement (Huang, et al., 2002).

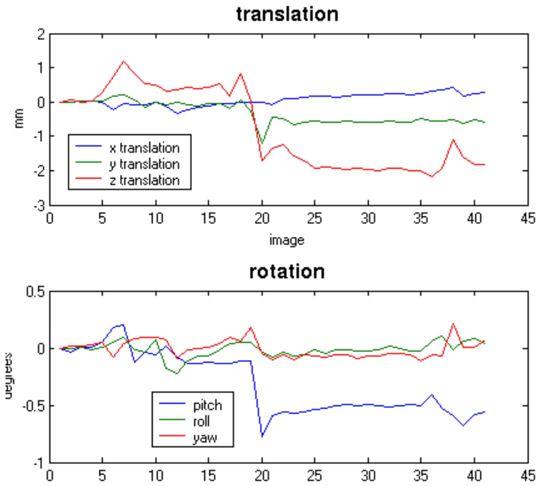

Figure 16: Movement parameter estimates for SPM realignment procedures. Top: translation along the x, y, and z-axis; bottom: the degree of rotation (pitch, roll, and yaw). These images depict motion artifacts due to head nodding.

4.5.1.2 Co-registration

Co-registration is used to register a structural image to a mean functional one. Structural images with higher spatial resolution are usually collected before or after functional experiments. They are used for the superimposition of functional images on the individual's own anatomy or to overlay group-level functional activation onto an average structural image. Because functional and structural scans have different characteristics such as signal intensity, voxel orientation and size, they need to be co-registered by SPM automated image-matching algorithms.

4.5.1.3 Slice-timing correction

This is an optional feature in SPM to correct for differences in acquisition time between slices. Different slices within a scan volume have different acquisition times because the respective images are acquired consecutively throughout the repetition times (TR) and not simultaneously. The order in which slices are acquired may vary, in most cases either ascending (bottom-top) or descending (top-bottom). That implies a delay in acquisition time between the respective images. Consequently, the hemodynamic response for voxels in later slices will appear to be ahead of that for earlier slices. If a single basis function (such as a canonical HRF) were used to model the response, the data of the earlier acquired slices would not match the data of the later acquired slices. This might lead to poor parameter estimates and decreased sensitivity of the statistical analysis. The slice-timing feature in SPM aims to correct for this mismatch. However, if applied before motion correction, rapid movements could cause errors in the slice-timing interpolation, resulting in a blurring across voxels from different brain regions. On the other hand, performing motion correction first might change the order in which the slices were acquired, thus affecting the temporal correction. Therefore, the application of slice timing must be carefully considered.

4.5.1.4 Spatial normalization

Before assigning an observed response to a particular brain structure, the data must conform to a known anatomical space. This is why spatial normalization is carried out. It is employed to warp images from different participants onto a template brain to make functional results comparable within standard stereotactic space. The two widely used brain templates available are the anatomical atlas of Talairach & Tournoux (1988) and the anatomical template of the Montreal Neurological Institute (MNI, Quebec, Canada). The MNI template was created by averaging a high number of fMRI scans from right-handed subjects and is the SPM default template. SPM uses a voxel-intensity-based approach for normalization, adopting a two-step procedure: First, a linear transformation is applied based on translations and rotations, accounting for the major differences in head shape and position. A precise match of the brain shape is achieved with the second step. That is, non-linear transformation (warping) parameters are identified to correct for smaller-scale differences in brain anatomy (Brett, et al., 2002; Talairach & Tournoux, 1988). This method warps the images such that functionally homologous regions from different subjects are as close together as possible. In principal, there are different options for fMRI normalization that either uses the structural or the functional image for warping. For example, one method uses the mean image of the functional series (that was created during co-registration) to estimate the warping parameters for an EPI template. The normalization parameters are then applied to the EPI images and to the (co-registered) anatomical image. The other method, standard in SPM5, first segments the individual's structural image into gray matter, white matter and cerebrospinal fluid to identify normalization parameters to then warp it for matching an estimate for the structural template. The caveats of spatial normalization are that it is impossible to make a perfect structural match between subjects, due to individual differences in anatomy. However, a solution for remaining small-scale differences in anatomical or functional location is provided with the next step in SPM preprocessing.

4.5.1.5 Spatial smoothing

Because of the limitations of normalization, homologous regions can never be precisely registered. Spatial smoothing helps to spread out the different areas and compensates for the discrepancy. The main advantage of smoothing is basically to increase the signal to noise ratio, thus resulting in enhanced sensitivity, and greater validity of statistical assumptions. This is typically achieved by convolving functional data with a Gaussian kernel, specified as full-width at half-maximum (FWHM) of two times the voxel size of the functional data. The effect is a blurring of individual differences through the averaging of neighboring data points, which in turn facilitates inter-subject comparisons. The disadvantages of spatial smoothing are the loss of spatial resolution. This has to be considered with respect to the focus of a given study. Smaller Gaussian kernels are more efficient for detecting individual differences in circumscribed brain regions but less suitable for generalizing the results. A kernel that is three times the size of the original voxel resolution (e.g. 3x3 mm2) typically yields a good balance between spatial resolution and the extension of results to the population.

4.5.2 Statistical analyses

Statistical methods are employed to assess the acquired data for evidence of an experimentally induced effect. There are several different classes of statistics available that include parametric and non-parametric approaches. SPM generally employs parametric statistical methods based on the “General Linear Model” (GLM; Friston, 2007). In order to make inferences about effects of interest, statistical analysis of fMRI data must follow three consecutive steps:

- (1) The modeling of data.
- (2) The estimation of effects.
- (3) The testing of these effects for significance.

Data from fMRI time series must be initially modeled as to decompose the observed variability into effects of interest and effects of no-interest (e.g. confounds). This is expressed by using matrix notation. The emerging design matrix has one row per observation, and one column (or explanatory variable) per model parameter with the remainder being the error measures (i.e. residuals). This quantifies the information about the experimental factors that may explain the expected signal. To further improve the model, it convolves a stimulus function with a canonical hemodynamic response function to predict the time course more accurately. The HRF is included along with other parameters as regressors in the respective model. The next step requires the estimation of effects and error using the least-squares parameter estimate. These parameters define the contribution of each component of the design matrix to the value of Y by calculating a "best-fit" between design matrix and data, thus leading to "fitted values". Differences between the observed data, Y, and that predicted by the model, reflect the error. That is, 'of all linear parameter estimates consisting of linear combinations of the observed data whose expectation is the true value of the parameters, the least squares estimates have the minimum variance' (Friston, 2007). Each voxel in a regressor of the design matrix is then represented by a beta value. Finally, the model parameters are used to look for an effect we are interested in. Therefore, the parameter estimates are converted into Student's t values, by dividing the parameter estimate (beta) through its standard error. In SPM, t-values and corresponding (probability) p-values are calculated for all voxels in the brain separately, which results in a large volume of statistical values. Only voxels that survived a statistical threshold will be presented in a color scale superimposed on a standard brain, representing the "activation map" or "statistical parametrical map" (SPM). For single subjects, these steps are referred to as "first-level" analysis, whereas the "second-level" statistics apply the general linear model to interrogate questions based on group statistics that incorporate combined first-level results (i.e. *t*-contrasts).

In general, performing statistical analysis is testing against the null hypothesis, which states that the condition of interest caused no effect. Therefore one compares the statistical value with a null distribution. The easiest method to decide when the null hypothesis can be rejected is to select a statistical threshold that is applied to every voxel in the statistical map. For example, adopting a conventional threshold of p=0.05 would limit the chance that the experimental results were actually a consequence of the null distribution to 5%. That is, one accepts a 5% chance that the null hypothesis was actually true and the results false positives (i.e. type I error). If statistical values survive this threshold, the null hypothesis is typically rejected. For neuroimaging data, a level of p=0.001 may be acceptable when testing for activated voxels was performed within restricted areas (i.e. small volumes). However, this method is usually problematic in fMRI experiments because of the many voxels in the brain that may result in a large amount of statistical tests to be carried out. This is particularly the case when no ad-hoc hypotheses exist about where in the brain the effect may occur and tests are carried out for the whole brain volume. This increases the chance of type I errors. For example, if 20 000 voxels are tested with a significance threshold of p=0.05, 1000 voxels may be expected to activate by chance, even if no stimulation was applied. One way to solve this problem is to employ statistical adjustments. In SPM, there are two methods available: the Family Wise Error Rate (FWE) and the False Discovery Rate (FDR). The former is based on the "Bonferroni correction", which divides the respective p-values through the number of tests carried out. Applied to the example above, a threshold level of p=0.05 would be divided through 20 000 (voxels), resulting in 0.00000025 falsely activated voxels that survived significance level. This is a very conservative method that increases type II errors (i.e. false negatives) as it likely eliminates also correctly activated voxels. This is because most functional imaging data have also some degree of spatial correlation between neighboring statistic values, so that there may be actually fewer independent values in the statistic volume than voxels. In contrast to the Bonferroni correction, FDR controls for the expected *proportion* of false positive voxels within those declared positive, rather than controlling the *total* number of false positive voxels whether or not the null hypothesis was rejected (Genovese, et al., 2002). The advantage of the FDR method is that it adapts its threshold to the features of the data. It is therefore more sensitive to individual differences across subjects and represents a more reliable method for controlling the multiple testing problems in neuroimaging experiments.

Statistical analysis of fMRI data knows two fundamentally different strategies, depending on the statistical inference that is aimed at. They are known as either "fixed effects" or "random effects" analyses. The ''fixed effects'' analysis does not take into account the subject-to-subject variability and can therefore not be used to make generalizations about the larger population from which the group was drawn. This measure of often employed for first-level analyses of individual subjects or case studies. On the other hand, the random effects model considers the variance in measurements between individuals in addition to the variance in a measurement for an individual subject. It is therefore used for the comparison of group statistics. A deviation from the attempt to detect differences between groups or conditions is the conjunction analysis. This approach is based on the commonalities of conditions. Hence, the conjunction analysis aims at identifying common patterns of BOLD response, which are present in all subjects or conditions analyzed. This method helps to unravel the basic processes involved in the performance of a task. After the statistical analysis is completed, the last step involves the identification of activated areas. This is typically done with automatic labeling software, accessible from within SPM. These labeling packages associate the coordinates of activation loci in stereotactic space with the respective brain area. The most prevalent maps are those based on the cytoarchitectonic areas defined by Brodmann (1909). Recently developed labeling procedures also allow for a combination of probabilistic cytoarchitectonic maps with functional imaging data (Eickhoff, et al., 2005).

Chapter 5

Overt and imagined singing of an Italian aria

5.1 Introduction

Performing music in ones mind is a technique commonly used by professional musicians to rehearse various aspects of a musical piece, e.g. difficult parts of an already executed musical passage (Lotze, et al., 2003). Composers write their music regardless of the presence of a musical instrument, as for example Schubert or Mozart did routinely. Beethoven composed his violin concerto in D-major (op. 61) even though he was already deaf and famous musicians like Vladimir Horowitz and Walter Gieseking reported frequent use of mental practice (Gieseking & Leimer, 1972; Schonberg, 1987). Several experiments have proven the benefits of motor imagery in motor skill learning and its use as a tool to probe the neural basis of overt motor performance is increasingly appreciated (Langheim, et al., 2002; Yue & Cole, 1992). For example, the training of auditory imagery has shown to improve harmony skills in children (Humphreys, 1986). Highben and Palmer (2004) demonstrated that the ability of pianists to learn an unfamiliar piece without auditory feedback depends on their degree of aural skills (e.g. the ability to perceive melodic differences). In a seminal study, Pascual-Leone (2003) demonstrated that mental practice has not only the capacity to improve performance but also to change the size of the cortical representation of the motor cortex over time. Motor imagery can improve the dynamics of motor performance (Yaguez, et al., 1998) and the velocity and smoothness of sequential movements trajectories (Lacourse, et al., 2005). The systematic and repetitive use of mental motor practice is a technique that has been successfully used to improve skill learning in sports (Hall, 2001). Yet mental music rehearsal does not only include motor, kinesthetic and auditory imagery but also emotional aspects of the piece rehearsed, which involves three components: a semantic meaning component, a stimulus component and a response component that defines the imagined emotion

(Lang, 1979). Emotional memory retrieval therefore represents an associative network processing that may reflect a prototype of a propositional network in the brain for overt responding.

Singing in the field of classical music involves both technical-motor and a strong emotional engagement in order to transport the musical intention and to communicate artistic, emotional and semantic aspects of the song or aria. Whilst classical singing shares some aspects with speech: i.e., articulation and movements of the vocal tract, expression of meaning, reading or recalling words and prosodic modulation of the voice, other elements are quite different. These involve the stylistic accuracy demanded in classical singing (e.g. in German Lied, Baroque music or Late Romantic opera), the tight regulation of musical pitch, meter and rhythm as well as an increased sound intensity and vocal range, vibrato and an emphasized dramatic expression of emotions. Such behaviors require a much higher level of vocal control compared to speech (Natke, et al., 2003). Motor aspects of these requirements are reflected in the finer laryngeal motor control (e.g. for accurate intonation), the appropriate adjustment of the resonating vocal tract (affecting for example timbre and vocal quality) and the high involvement of the diaphragm and the thoracic muscles for a precise control of the airflow during singing (Watson & Hixon, 1985). However, studies investigating cerebral correlates of singing are sparse and only focused on non-musicians.

A first fMRI-study that investigated silent singing in musically untrained subjects demonstrated a lateralization of the motor representation of the tongue area to the right during singing in comparison to a left lateralization during speech (Wildgruber, et al., 1996). The first study of overt singing used Positron Emission Tomography (PET) in subjects with musical training less than one year. Cerebral activation during singing of pitch and spoken vowels was compared (Perry, et al., 1999) and activity in motor areas (primary and secondary; cerebellar hemisphere in lobule VI), auditory areas (right gyrus of Heschl), the bilateral insula, the medial cingulate cortex, the parietal and occipital lobe and the brain stem was observed. Overt humming of a melody (W.A. Mozart's "Kleine Nachtmusik") compared to word generation revealed cerebral lateralization to the right in the insula and temporal pole and to the left in the cerebellar hemisphere (Riecker, Ackermann, Wildgruber, Dogil, et al., 2000). Hickok and colleagues (2003) were predominantly interested in audio-motor coupling and compared humming with listening to piano pieces. They reported a left parieto-temporal and frontal network active during humming. One study investigated neural correlates of overt singing and speaking by using the same bisyllabic words/phrases for both conditions in a group of non-professional singers (Özdemir, et al., 2006). They reported a shared network for sensorimotor processes in both singing and speaking, which involved the inferior pre- and post-central gyrus, superior temporal gyrus (STG) and the superior temporal sulcus (STS) on both hemispheres. However, singing in comparison to speech revealed a larger network that involved bilateral activations in the anterior to mid-portions of the STG (enhanced in the right hemisphere) including Heschl's gyrus but also in the anterior portion of BA 22 and BA 38, the most inferior portion of the inferior frontal gyrus (Broca's region), inferior aspects of the central operculum, and the middle and most inferior portions of the sensorimotor cortex, possibly reflecting a greater bihemispheric organization for singing. All of the above-mentioned studies were carried out with non-experienced or amateur singers. Studies with professional singers have not yet been published. The assessment may have been limited by artifacts elicited by movements of the head and jaw and changes of air-space during vocalization and articulation (Soltysik & Hyde, 2006). fMRI-techniques controlling for the artifacts produced by overt singing are now available (Gracco, et al., 2005; Soltysik & Hyde, 2006) and sparse sampling techniques allow for auditory feedback during singing (Gaab, et al., 2003; Özdemir, et al., 2006).

While substantial evidence corroborates the contribution of primary visual cortex during certain visual imagery tasks (Kosslyn & Thompson, 2003), the relative contribution of primary (core) and secondary (belt) auditory areas during auditory imagery is still a matter of discussion (Zatorre & Halpern, 2005). The primary auditory cortex of healthy subjects seems to have "a propensity to spontaneously 'activate' during silence" (Hunter, et al., 2006). Primary auditory cortex activation in the absence of auditory input has been found for short gaps of silence (lasting 2-5s) that were inserted at different points during the soundtrack of familiar music (Kraemer, et al., 2005) and in a sound detection task during the preceding silence of the emerging sounds (Voisin, et al., 2006). A further report of primary auditory activation during mental imagery of a computer generated monotone (Cmaj) comes from Yoo and colleagues (2001) but a recent inspection of the exact locus of activation indicated in that study questioned their results (Halpern, et al., 2004). More compelling evidence exists regarding the contribution of secondary (or association) auditory cortex during mental imagery of complex, meaningful sounds (Bunzeck, et al., 2005; Halpern & Zatorre, 1999; Shergill, et al., 2000; Shergill, et al., 2001; Zatorre, et al., 1996). These results are supported by studies on imagined musical performances in professional instrumentalists (Langheim, et al., 2002; Lotze, et al., 2003). We were therefore particularly interested in the involvement of auditory areas during imagined singing of the Italian aria "Caro mio ben" (by Tommaso Giordani). Given the more complex motor pattern for mental imagery of classical singing compared with other imagery tasks used in experimental situations, we expected strong activation of secondary auditory cortex. Due to the combination of lyrics and melody this activation should be represented bilaterally (Zatorre & Halpern, 2005). Furthermore, we expected left temporo-parietal sylvian cortex activation during imagined singing of a classical aria, supporting auditory-motor integration (Hickok, et al., 2003).

Imagined singing can be regarded as auditory imagery of a familiar tune. Yet mental singing involves not only auditory mental representations but also somatosensory and kinesthetic aspects, which require a declarative knowledge about the different components of the task (Jackson, et al., 2001) – a knowledge almost absent in unskilled subjects. In fact, Driskell and colleagues (1994) found that the behavioral effect of motor imagery in sports depends on the previously acquired knowledge with a physical task. That is, a stronger effect of motor practice on learning and overt performance can be expected for more experienced subjects, whereas when motor representations are lacking, subjects will benefit only in the cognitive domain. Thus, a kinesthetic representation of the motor components involved is essential before motor imagery can efficiently improve the behavioral performance (Fourkas, et al., 2008).

Imaging studies on simple motor paradigms confirmed that imagined movements share certain neural substrates with executed movements in the supplementary motor area (SMA), the dorsal and ventral premotor cortex (PMC) the cerebellar hemispheres and also – to a lesser extent – the primary motor cortex (Lotze, Laubis-Herrmann, et al., 1999; Porro, et al., 1996; Sabbah, et al., 1995). Using fMRI and TMS, Kuhtz-Buschbeck and colleagues (2003) demonstrated that the primary motor cortex contralateral (cM1) to the imagined movements was active during imagery tasks of increasingly complex movements, a result consistent with a previous finding of more prominent involvement of cM1 with performance of complex motor sequences (Gerloff, et al., 1998; Lotze, Erb, et al., 2000). Furthermore, kinesthetic motor imagery (i.e. to try to feel the sensations that are usually felt in the muscle-tendon complexes when actually executing the movement rather than just visualizing the execution) involved the primary motor cortex to the same extent as motor execution conditions, indicating a comparable activation of this brain structure in both conditions (Caldara, et al., 2004). Earlier studies focusing on mental performance in professional instrumentalists however, did not observe significant M1 and S1 activation (Langheim, et al., 2002; Lotze, et al., 2003). It is interesting in this respect that singing and particularly singing of classical music as compared to instrumental music requires more body-core centered motor and anatomic activity involving muscles and internal organs of vital bodily importance such as respiration, salivation, ingestion, speech and social communication. In contrast, string instruments and piano playing investigated in most brain imaging experiments involves rather fine control of peripheral muscle groups. Therefore, we propose that imagery of such vital body-centered activity mobilizes the entire motor system (primary, secondary, basal ganglia, thalamus, cerebellum, brain stem) while mental re-enactment of sensory stimuli frequently used in MR-studies of imagery (see Kosslyn & Thompson, 2003 for an overview) are "capturing" much less vital body-core functions and therefore report less areas activated. In addition, singing is evolutionary older and more genetically primed than instrumental music, constituting the first basic function to express emotive meaning (Brown, 2000; Molino, 2000). Motor and autonomic associative networks should therefore be more widespread. Given that singing an aria is highly associated with emotional expression, and perhaps even more so when free from the demands of actual performance, we also expected activation of areas involved in emotional processing such as the amygdalae, hippocampus, temporal poles, anterior cingulate cortex (ACC), anterior insula, ventral striatum and ventrolateral prefrontal cortex (VLPFC) (Koelsch, et al., 2005).

5.2 Methods

5.2.1 Participants:

16 right-handed classical singers (4 chorus; 4 opera, 8 vocal students; mean age: 31.06 +8.27 years; range: 20-44 years; 5 men) without reported history of neurological or psychiatric disease participated in the study. The average experience of training in singing was 14.06 ±7.59 years; range 7-28 years. Participants took their first professional singing lessons at the average age of 16.91 ±3.14 years and reported an average of 25.73 ±7.26 hours spent with practicing per week during the 3 months before fMRI-scanning. All singers reported experience in mental rehearsal. The study was approved by the Ethics Committee of the Medical Faculty of the University of Tuebingen. Written informed consent was obtained according to the guidelines of the Declaration of Helsinki.

5.2.2 Singing Task:

We selected six phrases from the first line of the bel canto aria “Caro mio ben” by Tommaso Giordani, each phrase sung separately after a visual signal (see Figure 17A). This piece was selected since it is well known by classical singers, it involves strong emotional expression, lyrics and melody are easily memorized, all voice types (tenor, soprano etc) can perform it and the vocal demands are such that it can be performed in a scanner environment without extreme pitches or forti, which might induce additional artifacts. Additionally, the phrase structure of that piece is such that the 3 seconds phrases naturally fit the sparse sampling technique employed, which allowed for natural auditory feedback during singing. Phrases were rehearsed before (by B.K.) with recorded scanner sound and were trained again within the scanner for the first six phrases. The overt singing part was followed by the mental rehearsal of singing in distinct sections. In the latter section, subjects sung the same phrases only in their imagination. Six periods of singing (and imagined singing respectively) alternated with six periods of deep inspiration not followed by singing (or imagined singing).

With respect to the imagined singing task, subjects were required to imagine the physical performance of singing as vividly as possible without performing any movements. Concentration, vividness of imagery of singing, singing related movements, tempo and rhythm during the fMRI-scanning was assessed with Visual Analogue Scales (VAS; from 0 = ”not at all” to 10 = ”very good”) after each block. Auditory control by microphone during the imagery period did not reveal any sound-production. To test for possible movements of throat muscles involved in overt and imagined singing we recorded surface electromyography (EMG) of the musculus cricothyroideus with the Nexus-10 polygraph (MindMedia, NL) and superficial silver chloride electrodes in 10 subjects during baseline (breathing), mental rehearsal (imagined singing) and (overt) singing (see Figure 17B). EMG was measured outside of the scanner since EMG-electrodes resulted in substantial artifacts if placed in a head coil. Signals were acquired at 1024 samples per second, bandpass filtered (20-500Hz; IIR filter, Butterworth 4th order) and sent via Bluetooth to a PC for data visualization, storage and preprocessing using BioTrace Software (MindMedia, NL). EMG amplitudes were calculated (root mean square, epoch size 1/16 second, 32 samples per second (SPS)) and averaged over 30 seconds for baseline, imagined singing and singing for each subject. Differences between the three measures were compared with ANOVA followed by paired t-tests using the Statistical Package for the Social Sciences (SPSS 10.05).

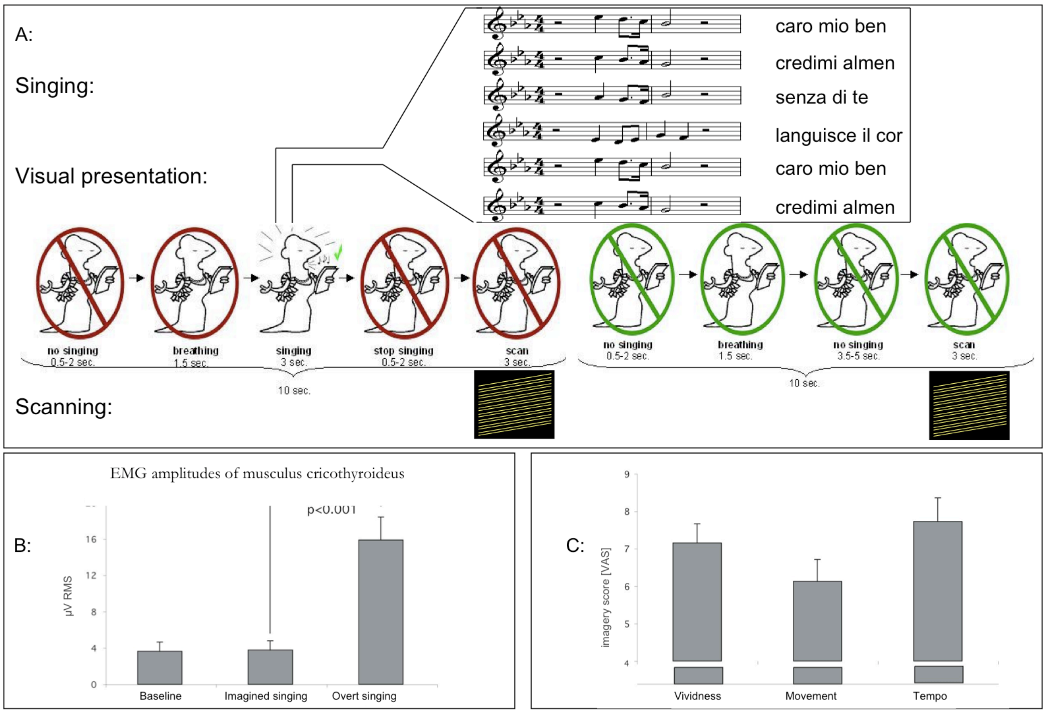

Figure 17: **A**. Design of the procedure. First line: visual signals indicate breathing & singing during fMRI scans. Red circles for singing blocks (overt & imagined) and green circles for breathing only blocks respectively. Below: Scanning during pausing, for each block 66 scans were performed alternating singing and rest. The imagined singing block was indicated with the same symbols but subjects were instructed to use mental rehearsal instead of overt singing. **B**. Superficial EMG amplitudes (µV root mean square (RMS), 32 samples per second) of musculus cricothyroideus, averaged over 30 seconds during baseline (breathing only), imagined singing and overt singing from 10 of the participants recorded outside of the scanner. Baseline and imagined singing did not differ (t(8)=0.59; n.s.) whereas singing versus imagined singing (t(9)=4.90; p<0.001) and singing versus baseline differed significantly (t(9)=4.87; p<0.001). **C**. Rating of imagination intensity (visual analogue score from 0 to 10) by all singers obtained after scanning. Movement rating showed a trend of decreased tempo rating (t(15)=2.11) but this result did not stand correction for multiple comparisons.

5.2.3 fMRI technique:

Whole head scans (66 volumes per block) were performed with a 1.5 Tesla whole body Scanner (Siemens Vision) using echo planar imaging (EPI; TE: 40 ms; TR: 3 sec, 36 transversal slices of 3 mm thickness and 1 mm gap, matrix 64*64). A sparse sampling method was employed to avoid movement artifacts and to allow auditory control during singing. The first 3 EPI data sets of each session were discarded prior to analysis to allow for T1-saturation effects. T-1 weighted images (MPRage; 176 sagittal slices with

1mm effective thickness) were measured for anatomical reference. Data were analyzed with the statistical parametric mapping program (SPM99 (for preprocessing and single subject statistics) and SPM2 (random effects group statistics), Wellcome Department of Imaging Neuroscience) running on Matlab (MathWorks Inc; Natick, MA; USA). SPM2 was used for group statistics to allow for non-sphericity correction in case of unequal variances between levels or groups, an option not available in SPM99. Each individual scan was realigned to the first one of each scanning condition to correct for movement artifacts. The realigned data were spatially normalized to the MNI-template and resliced with 3*3*3 mm. The resulting images were smoothed with a 9 mm (full width at half maximum) Gaussian filter. Individual statistical maps (fixed effect) were calculated for singing and imagined singing (main effect) and for differences between both. Contrast images of each subject were then used for group statistics calculated as random effects analysis at the 2nd level, which takes variance between subjects into account. The statistical threshold used to report group-activations was set as $p<0.05$ corrected for the whole brain (false discovery rate; FDR; Genovese et al., 1999). T-values of significant activations of the highest activated voxels were given for the MNI-coordinates and were assigned to anatomical regions. All regions were detected with the “Automated Anatomical Labeling” software (AAL) (Tzourio-Mazoyer, et al., 2002) or if already cytoarchitectonically mapped (primary motor cortex (BA 4a and p); primary somatosensory cortex (BA 1,2,3), primary auditory cortex (PAC 1.0, 1.1, 1.2), Broca’s area and analogon (BA 44/45) and SII (POP 1,3 and 4; Morosan et al. (2001), Wernickes area) with maps based on cytoarchitectonical data with 50% probability (Eickhoff, et al., 2005). The auditory-association area in the temporo-parietal lobe was selected after the highest activated voxels given by Hickock et al. (2003). The lateralization index (LI) was calculated with the LI-tool (Wilke & Schmithorst, 2006) by the following formula: (left-right)/(left+right). Activation maps were overlayed on a T1-weighted anatomical image, which fits best to the MNI-space by using MRIcro software (http://www.sph.sc.edu/comd/rorden/mricro.html).

5.3 Results

5.3.1 Rating of imagery:

Concentration (7.57 +1.73 of 10), vividness of imagined singing (7.15 +2.04 of 10), tempo (7.72 +2.54 of 10), rhythm (7.56 +2.35 of 10) and singing related movement (6.13 +2.13 of 10) were rated moderate to high and did not differ significantly (Figure 17C).

5.3.2 EMG-measures

ANOVA revealed a significant effect for condition (overt, imagined and baseline; $F(9,1)=48,45$; $p<0.001$) for the superficial EMG of musculus cricothyroideus. Post-hoc paired t-tests (Figure 17B) revealed no difference between baseline and imagined singing ($t(8)=0.52$; n.s.) whereas singing versus imagined singing ($t(9)=5.33$; $p<0.001$) and singing versus baseline differed significantly ($t(9)=5.28$; $p<0.001$).

5.3.3 Results of fMRI analysis

5.3.3.1 Overt singing:

The results of the main effect for singing performance are plotted in Supplementary Table 1 and examples of activation sites are given in Figure 18. Overt singing involved the primary sensorimotor cortex bilaterally (in the somatotopic area of articulatory movements) but also secondary motor areas (SMA, right PMC), basal ganglia, the thalami, the cerebellar hemispheres and the vermis. Interestingly, medulla activation was found around the nuclei of the cerebral nerves V, IX, X, XII, associated with articulation and phonation. The primary auditory cortex (A1) and the surrounding belt area were active bilaterally during singing in comparable magnitude (lateralization index (LI) A1: 0.06). The auditory association area was active bilaterally (left: $t = 4.91$, peak in -51; -42; 21; right $t = 3.03$, 51; -36; 18). Both, Broca's and Wernicke's areas and their homologues were activated. No area (BA 44/45, S1, M1, SII, PMC, parietal) showed relevant lateralization (LI <0.05). Some areas related to emotional processing (anterior insula, ACC) showed significant activation.

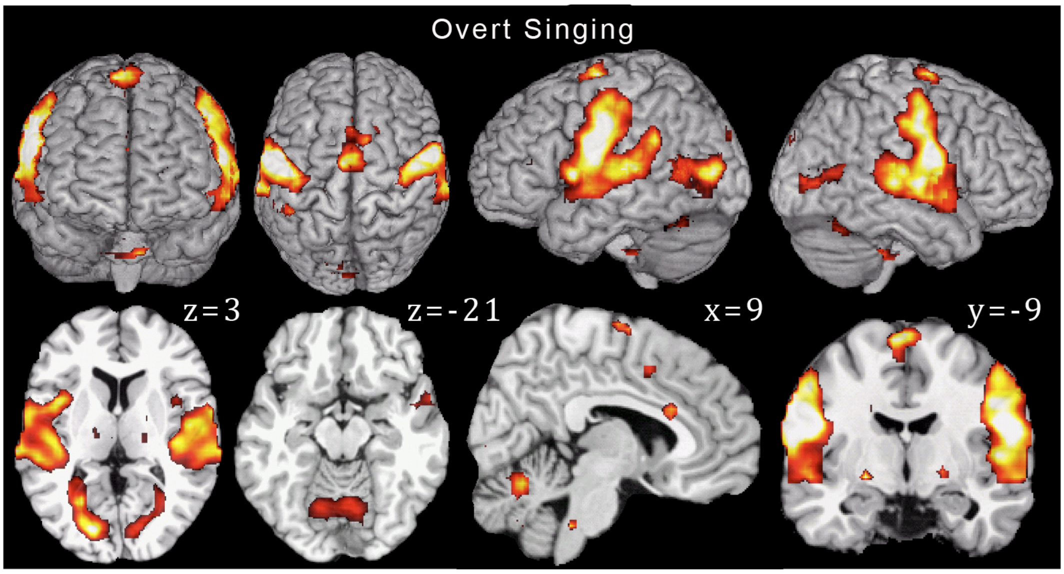

Figure 18: Main effects of overt singing. Activation maps were projected on Collin's reference brain, provided by MRIcro, which is most similar to the MNI-average brain. Bilateral sensorimotor cortex and cerebellum, bilateral auditory cortex, Broca's and Wernicke's area and their homologues, medulla, thalamus and ventral striatum but also ACC and insula were activated. Coordinates of cuts are given above each slice.

5.3.3.2 Imagined singing:

The results of the main effect for imagined singing performance are plotted in Supplementary Table 2 and examples of activation sites are given in Figure 19. Imagined singing activated predominantly the fronto-parietal loop. This included secondary and tertiary motor areas (supplementary motor area (SMA); premotor cortex (PMC)), Broca's area and its homologue (no lateralization; LI < 0.11) and the superior and inferior parietal lobe. Additionally, subcortical motor areas were involved (bilateral cerebellar hemisphere and vermis, bilateral putamen, pallidum and caudate nucleus). Interestingly, the primary motor (M1) and somatosensory cortex (S1) and the thalami were active bilaterally. Furthermore, the medulla was also active during imagery. No activation was found in the primary auditory cortex or in the auditory belt area but in the temporo-parietal lobe bilaterally (left: t= 4.90, peak in –51; -42; 21; right t = 9.73, 48; -39; 15). Several areas associated with emotional processing showed significant activation (anterior insula, bilateral amygdala, ACC, bilateral temporal pole, bilateral anterior hippocampus, mPFC).

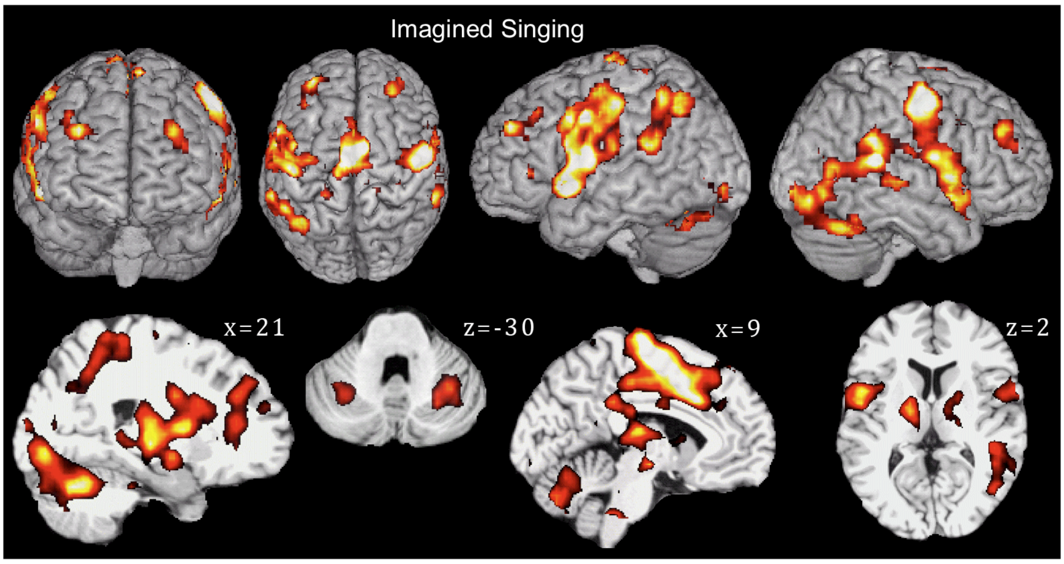

Figure 19: Main effects of imagined singing. Imagined singing involved typical imagery regions (SMA, PMC, parietal lobe, BA 44/45 bilaterally) as well as primary sensorimotor areas, the thalamus, the basal ganglia and more inferior regions of the cerebellum in both hemispheres. Areas processing emotions showed intense activation (ACC and bilateral insula, hippocampus, bilateral amygdala, ventrolateral prefrontal cortex). Bilateral A1 showed no significant activation. Coordinates of cuts are given above each slice.

5.3.3.3 Overt minus imagined singing:

The results of the difference in activation maps for the calculation overt minus imagined singing are plotted in Table 1 and examples of activation sites are given in Figure 20A. Overt minus imagined singing revealed bilateral activation in the primary motor cortex, in the primary and secondary somatosensory cortex, bilateral superior temporal gyrus including primary auditory cortex. Subcortical motor areas (right cerebellar hemisphere, thalamus, medulla) showed also an increase during execution.

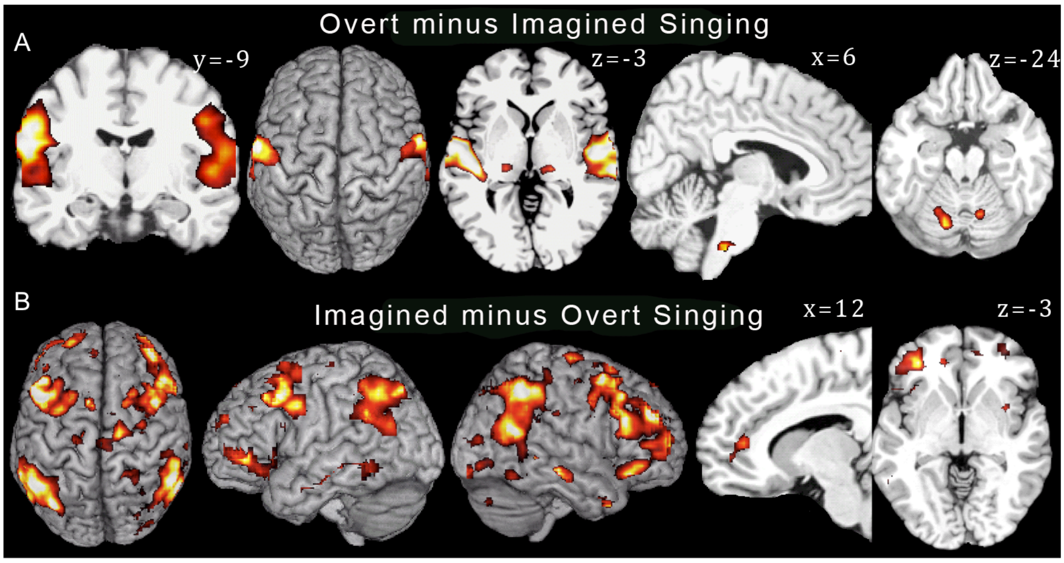

Figure 20: A. Overt singing showed higher activation in bilateral sensorimotor cortex in the somatotopic area of the articulators, bilateral auditory cortex and other areas of the superior temporal gyrus, SII, superior cerebellar hemispheres, thalamus and medulla. **B**. Mental rehearsal of singing showed higher activation in BA 6 (SMA and PMC), dorsal and ventrolateral prefrontal cortex, BA 44 left and 45 right, the parietal lobe and the ACC. Coordinates of cuts are given above each slice.

5.3.3.4 Imagined minus overt singing:

The results of the difference in activation between imagined and overt singing are plotted in

Table 2. Examples of activation sites are given in Figure 20B. Imagined minus overt singing showed increased activity in the SMA, the prefrontal cortex (PFC) and the inferior parietal lobe. Broca's and Wernicke's area and their homologues were stronger activated during the imagination task. Areas associated with emotional processing were also enhanced during imagined singing (ACC, ventrolateral and medial prefrontal cortex).

Table 1: Overt minus imagined singing (FDR corrected: $p < 0.05$ for the whole brain. Voxel extent threshold >5; cytoarchitectural masks are indicated with a *. BA = Brodmann area)

Area	cluster in mask	t-value	x	y	z
left S1* (primary sensorimotor; BA 1,2,3)	34	6,18	-57	-12	33
right S1* (primary sensorimotor; BA 1,2,3)	45	4.85	39	-24	39
left M1* (primary motor cortex; BA 4a+p)	16	5.97	-54	-9	33
right M1* (primary motor cortex; BA 4a+p)	17	5.02	48	-12	36
left A1* (primary auditory cortex; BA 41)	42	6.01	-39	-30	12
right A1* (primary auditory cortex; BA 41)	67	5.44	45	-21	12
left superior temporal gyrus (BA 22)	336	5.66	-45	-21	3
right superior temporal gyrus (BA 22)	333	5.47	66	-3	-6
left occipital lobe (BA 17, 18, 19)	34	5.20	-21	-75	9
left cerebellar hemisphere (H VII)	16	5.02	-9	-42	-24
right cerebellar hemisphere (H VII)	8	4.14	15	-63	-24
left SII* (secondary sensory cortex; OP 1,3,4)	44	7.38	-63	-9	21
right SII* (secondary sensory cortex; OP 1,3,4)	29	5.59	63	-6	12
Medulla	30	5.29	6	-30	-45
Thalamus	5	4.41	18	-24	-3

Table 2: Imagined minus overt singing (FDR corrected: $p < 0.05$ for the whole brain. Voxel extent threshold >5; cytoarchitectural masks are indicated with a *. BA = Brodmann area)

Area	cluster in mask	t-value	x	y	z
SMA* (medial supplementary motor; BA 6)	26	7.25	-9	12	51
right PFC (ant. med front gyrus; BA 9)	5	5.70	42	45	33
left ventrolateral prefrontal (BA 47)	13	4.22	-36	30	-12
right ventrolateral prefrontal (BA 47)	10	4.76	42	39	-9
anterior cingulate cortex (ACC; BA 24)	34	6.18	9	42	9
anterior cingulate gyrus		4.86	-15	45	-6
left parietal inferior (BA 39, 40)	156	6.79	-48	-57	48
right parietal inferior (BA 39, 40)	93	7.23	60	-51	45
left parietal superior (BA 5, 7)	9	4.86	-33	-75	51
right BA 45* (Broca´s analogon, BA 45)	42	5.61	51	27	24
left BA 44* (Broca´s area, BA 44)	6	4.35	-50	9	42
left angular gyrus (BA 40)	110	6.19	-39	-69	45
right angular gyrus (BA 40)	77	5.12	57	-63	24

5.4 Discussion

In the present study, we report cerebral activation sites of professional singers during singing and imagined singing of an Italian aria. Using a sparse sampling technique, cerebral mapping of overt and imagined singing revealed a number of activation sites. These sites were subsumed in the following subgroups: those involved with the motor performance or the motor-concept; those involved in auditory control of the produced sound and audio-motor co-processing, those involved in working-memory processes (or involved in the recall of lyrics and musical aspects of the song respectively) and areas involved in emotional processing. These functional loops do partially overlap. Since our material contains both lyrics and melody this study will not contribute to the extensive discussion about overlapping or differentiated functional maps of language and music (e.g. see Patel, 2003). A striking feature of this study was that most areas involved in motor processing during overt production were also activated during imagined singing. In contrast, most areas involved in auditory control during overt singing were absent during imagery. Yet imagery also showed increased activation in a fronto-parietal working memory network as well as in emotion processing areas.

5.4.1 Motor system:

5.4.1.1 Cortex:

The cortical representation sites in the sensorimotor cortex comprised areas centered in the somatotopic representation of the articulators and lip and tongue movements (Lotze, Erb, et al., 2000; Lotze, Seggewies, et al., 2000). No activity was seen in the supposed representation of the diaphragm and inspiration muscles in the area of the thoracic wall. Probably because deep inspirations were also performed in the baseline period and these activations might have been cancelled out. Activation in the primary motor (M1) and somatosensory cortex (S1) was significantly increased during overt singing in a pattern, similar to one recently reported for non-professional singers (Özdemir, et al., 2006). Interestingly, M1 and S1 were also active during imagined singing. Earlier studies on imagined performance in instrumentalists did not observe significant activation of M1 and S1 (Langheim, et al., 2002; Lotze, et al., 2003). This could be a consequence of the increased statistical power of the present study, which investigated more subjects – it has been reported that imagery involves M1 only about 30% of the intensity than overt performance does (Lotze, Montoya, et al., 1999; Porro, et al., 1996). On the other hand, subtle movements of the articulators during scanning cannot not be fully excluded based on superficial EMG results of baseline and imagined singing outside of the fMRI scanner, or by controlling for sound production during the imagery-task in the scanner – an important difference to the feasible EMG-control of hand movements in instrumentalists during scanning. However, it has been previously reported that imagined singing (without EMG-control) may also involve the precentral gyrus (Wildgruber, et al., 1996). Therefore, we assume that imagined singing might indeed differ from imagined instrumental performance with respect to the involvement of M1 and S1. This inference may be further supported by activation of the anterior insula during both overt and imagined singing. While the anterior insula plays an important role in emotion processing (see discussion below), the left insula also seem to contribute to the actual coordination of muscles engaged in articulation and phonation. (Ackermann and Riecker, 2004).

5.4.1.2 Cerebellum:

The representation sites in the anterior cerebellar hemispheres mirror those of the cortical sites around the central sulcus. The cerebellum is preferentially involved in timing of complex movements with sensory feedback (Thach, et al., 1992) and has been reported to serve additionally as an integrator of multisensory information from the cerebral cortex into a motor frame of reference (Bloedel, et al., 1993). The spinocerebellum is also involved in sensorimotor tasks (Gao, et al., 1996). Activation maps were larger than those reported for simple tongue and lip movements but were predominantly restricted on the area inferior to the fissura prima (Larsell lobule HVIa) (Grodd, et al., 2001). Consistent with previous observations, imaging showed an activation peak inferior and posterior Larsell's lobule HVII to areas involved in motor execution in the cerebellum (Lotze, Montoya, et al., 1999) but extended also anteriorly to HVIb. Cerebellar activity is highly correlated with the temporal sequencing of speech movement (Mathiak, et al., 2002). The functional demands of singing may explain this activity. Vermis activation, not described in previous studies investigating singing, was present during overt singing. The vermis is involved in balancing of the body during gait (lesion-data: (Muley & Bushara, 2004); imaging-data: (Jahn, et al., 2004)). For singing, only speculative interpretations of the function of the vermis can be provided, as it may also be associated with the unnatural lying-position during classical singing in the scanner. It would be interesting in this context to investigate in a further study whether opera singers, who are often required to sing in unusual positions show increased vermis-activation when compared to, for example, concert singers.

5.4.1.3 Basal ganglia:

Parts of the dorsal basal ganglia (putamen and pallidum) were activated during singing. This activation is likely to be associated with the motor loop (Alexander & Crutcher, 1990) and less so with the sensory-attention loop (Liu, et al., 1999). Here, the basal ganglia may be involved in the modulation of the voice, comparable to the modulation of smooth movement performance (Brooks, 1995).

Interestingly, the basal ganglia were also active during imagined singing. The comparison between overt and imagined singing revealed no differences. The basal ganglia activation may indicate an associative connection during overt and imagined singing between the dorsolateral prefrontal and orbito-frontal cortex via the nucleus caudatus pointing to working memory activation (dorsolateral PFC) and perhaps to an anticipatory fine-tuning of text and melody (ventrolateral PFC). The basal ganglia activation also involves the limbic-emotional loop, predominantly connecting the ventral striatum with the ACC (Alexander & Crutcher, 1990). In contrast to other studies on musical perception (Blood & Zatorre, 2001; Menon & Levitin, 2005) we did not observe any activation in the nucleus accumbens located in the ventral striatum. The reason for this is unclear; both the uncomfortable position in the scanner and the strict regime of the task may have decreased the pleasure of this singing aria.

5.4.1.4 Medulla:

We observed circumscribed activation in the superior medulla in a region where cranial nerves VII, IX, X and XII are relayed to movements related to articulation, vocalization and phonation. This activation is most likely not elicited by movement artifacts based on articulation dependent cerebral fluid pulsation, since activation maxima were located distinct from the border of the spinal fluid and the brain stem. Patients with lesions in the superior medulla express drastic impairments in articulation and swallowing (Vuilleumier, et al., 1995). Brainstem activation in our study was significantly increased for overt singing compared to imagined singing, thus pointing to the executive function of this activation site.

5.4.1.5 Secondary motor areas (BA 6):

The supplementary motor area (SMA) and the medial cingulate cortex were comparably active during both overt and imagined singing. Both areas have previously been described to be involved in overt vocalization (Perry, et al., 1999). Together with the premotor cortex (PMC) these areas seem to be critical for the planning, timing and ideation of the movement and the motor imagery (for an overview see: (Lotze & Halsband, 2006). It is possible that the access to motor programs during mental rehearsal is realized through functional activation of BA 6, which has also shown to be involved in kinesthetic imagery tasks (Gerardin, et al., 2000; Stephan, et al., 1995).

5.4.2 Areas involved in auditory control:

During overt singing, we observed bilateral activation of the primary auditory cortex (A1) and the auditory belt area with comparable intensity. Recent functional imaging studies suggest a relative left-hemispheric specialization for rapid temporal and rhythm processing (Boemio, et al., 2005; Overy, et al., 2004; Zatorre & Belin, 2001) whereas right auditory cortex shows a stronger sensitivity for spectral and melody processing (Overy, et al., 2004; Zatorre & Belin, 2001) and for processing of auditory information with slower temporal resolution (Boemio et al., 2005). With regard to lateralization effects of primary auditory cortex, simple humming without language production activates A1 only in the right hemisphere (Perry, et al., 1999) whereas pitch memory has also been associated with increases in the left Heschl's gyrus (Gaab, et al., 2006). The ambiguous role of A1 lateralization in melody perception tasks may be mediated by the listening style (Evers, et al., 1999). For example, listeners who determined pitch information from the spectral composition of a tone showed a pronounced rightward activity in the lateral Heschl's gyrus when compared to listeners who focus on the fundamental frequency only (Schneider, et al., 2005). The Italian aria used in our study was comprised of text, rhythm and melody. Thus, the bilateral activation of A1 with comparable magnitude is not surprising. The activation of the auditory cortex found in our study was similar to the one reported by Ozdemir and colleagues (2006) for non-professional singers, who also found bilateral activation in the middle section of the superior temporal lobe for overt singing as well as for overt speaking when compared to silence, suggesting a bihemispheric activation for both tasks. In their study, only the contrast singing more than speaking revealed A1 activation more pronounced in the right hemisphere. Callan et al. (2006) compared Japanese singing versus speech production and found an increased activation for singing in the right planum temporale. Since speech was highly lateralized to the left in their study, bilateral planum temporiale activation during singing resulted in increased right hemispheric activity after contrasting singing with speech. This result corresponds with our findings of bilateral planum temporale activation during singing.

During imagery, activation of A1 (Morosan, et al., 2001) was completely absent (see Figure 19). This is in line with several observations on sound and music imagery, which found activity the secondary auditory cortex during imagery but not in A1 (Ducreux, et al., 2003; Halpern & Zatorre, 1999; Kraemer, et al., 2005; Yoo, et al., 2001; Zatorre & Halpern, 2005). It is interesting in this respect that our subjects rated the vividness of imagined singing as high and reported no problems with attention during the imagery task. Vividness of imagery has previously been shown to correlate high with psychophysiological responses during imagery (Lang, 1979). Yet we did not assess to what extent our subjects experienced the imagined aria as an auditory or kinesthetic representation. Two earlier studies that investigated imagined performance in instrumentalists did also not observe A1 activation during imagery (Langheim, et al., 2002; Lotze, et al., 2003) although the rating of imagined melody was high (Lotze, et al., 2003). Conversely, when professional and amateur violinist were tapping an imagined and well-known piece of music without auditory feedback, activations in the primary auditory cortex were found, especially in the professional group (Lotze, et al., 2003). This finding may represent a tight interconnection between motor action, sensory perception and auditory control in musicians (Bangert, et al., 2001; D'Ausilio, et al., 2006). Comparable cross-modal interactions in singers could be perturbed by the lack of activity in the somatosensory and motor modality. However, that assumption is contradicted by the activity of M1 and S1 during imagined singing in our study.

An auditory association area in the left hemisphere was reported to be responsible for the control of spoken and listened words (Hickok, et al., 2003). We also observed activity around the peak of this area in the temporo-parietal lobe during both overt and imagined singing, although this was bilaterally expressed. It is possible that auditory association for leading the melody during both singing and imagined singing may be represented in this area.

5.4.3 Speech and song processing:

We observed no relevant lateralization during both overt and imagined singing in Broca's area. Bilateral activation in comparable magnitude may be related to the combination of melody and lyrics in the Italian aria. Compared to simple humming, which showed right lateralized activity (Perry, et al., 1999), the singing of an aria involved also linguistic, tempo and rhythm processing previously shown to be lateralized to the left (Tramo, 2001). In a very recent paper Lahav et al., (2007) compared listening to untrained and trained musical sequences. Interestingly, Brocca's area and its homologue were increasingly active during listening to a trained piece compared to listening to the same piece before training. The authors concluded this area has a multifunctional role in action listening besides its role for sensorimotor integration (Binkofski, et al., 2000). In the light of this study, bilateral activation of the inferior frontal gyrus during both overt and imagined singing might not necessarily be related to the recall and internal representation of lyrics but might point to the recall of trained motor processes or sensorimotor interactions.

Increased processing of memory for pitch was reported to be related to activation in BA6, 44, 45, and BA 40 and 7 (Koelsch & Siebel, 2005) – areas that were all active during the imagery task in our study. Together with additional mobilization within BA 9 (Zhang, et al., 2003), this increased activation during mental imagery may suggest that it is less automated than overt singing. If the auditory feedback is missing, increased resources may be needed to keep track, which may lead to an increased working memory demand in the frontal and parietal lobes. The inferior frontal cortex seem to play a particular role in the processing of musical structure and syntax (Tillmann, et al., 2006). It has recently been attributed to emotional intonation processing (Wildgruber, et al., 2005) and motor preparation for vocal production (Özdemir, et al., 2006). Furthermore, auditory imagery of familiar melodies and inner speech involve the inferior frontal gyrus, together with SMA and auditory association areas in the superior temporal gyrus (Halpern & Zatorre, 1999; Shergill, et al., 2001). All these areas where active during imagined singing in our study, which supports the idea that the frontal cortex may be involved in auditory imagery processes (Zatorre & Halpern, 2005).

Parieto-temporal association areas for controlling auditory stimulation are located near or within Wernicke's area and its homologue. Its involvement in auditory imagery has been demonstrated before (Halpern & Zatorre, 1999; Shergill, et al., 2001). We found increased inferior parietal cortex activation bilaterally during both tasks, which could reflect activation of a store for phonetic and intonational acoustic information contributing to the working memory system (Wildgruber, et al., 2005; Zatorre, 2001; Zatorre, et al., 1994).

The superior parietal lobe (BA 7) also showed increased activation during imagined singing. This area has been reported to be activated during imagery of complex movements trajectories (Wolbers, et al., 2003). The superior parietal lobe is closely connected to the posterior SMA but also to the PMC (Rizzolatti & Arbib, 1998). Spatial trajectories of the movement may be transformed in corresponding motor programs, processed and stored in the superior parietal lobe. Patients with parietal lesions were found to have problems with tasks that involve imagined movements (Sirigu, et al., 1996). Thus, the parietal lobe might be crucial for mental rehearsal. This could be especially true for musicians, since more inferior parts of the parieto-temporal lobe found during imagery could be essential for the auditory processing of the piece that was mentally rehearsed.

Activation in the occipital lobe during both overt and imagined singing was more pronounced during overt singing. Arguably, overt singing needs a closer attentive control of the symbolic instructions, which indicated onset and end of a performance period, although Perry et al. (1999) reported occipital activation during humming despite his subjects having their eyes closed. They concluded that visual imagery might account for this activation.

5.4.4 Emotional processing:

The overt singing task activated only the anterior cingulate cortex (ACC; for differentiation see: (Bush, et al., 2000)) and the insula. The ACC contributes strongly to the recall of emotions (Phan, et al., 2004) – a capability which might be important for both overt and imagined performance.

We observed bilateral insula activation during overt and imagined singing, an area that has been termed "limbic sensory cortex" and that is associated with emotional states, emotional recall/imagery and emotional tasks with cognitive demand (Phan, et al., 2002). Insula activation also accompanies the experience of “chills” during the perception of pleasurable music (Blood & Zatorre, 2001). The anterior part seems particularly related to emotional valence, reflecting the intensity dimension of an emotion (Anders, et al., 2004), whereas the dorsal part of this structure is more active during cognitive tasks (for a meta-analysis see Wager, 2004).

Interestingly, imagined singing showed additional activation in areas related to emotion processing involving the medial prefrontal cortex (mPFC), the bilateral ventrolateral prefrontal cortex (VLPFC; also called orbitofrontal lobe; BA 47), the bilateral temporal poles and amygdala. The temporal poles are involved in conscious self-regulation of emotional responses (Beauregard, et al., 2001; Lane, 2000). The amygdala is known to be mainly involved in more passive avoidance or approach tasks (Morris, Ohman, et al., 1999; Morris, Scott, et al., 1999). This is consistent with the fact that the amygdala was not active during overt singing. Although amygdala activation has been reported for the perception of unpleasant musical material (Blood, et al., 1999; Koelsch, 2005; Koelsch, et al., 2005) its activation is not specific for negative emotions.

Other areas such as the mPFC, the VLPFC - active during imagined singing in our study - have also been related to unpleasant musical stimuli (Blood, et al., 1999), but results from meta-analyses on studies investigating visually presented emotional material did not find area specificity for positively or negatively rated stimuli (Murphy, et al., 2003; Phan, et al., 2002). Instead, the mPFC may play a "general" role in emotional processing for appraisal/evaluation, experience, and response and can be related to self-referential processing about one's own emotional experience (Phan, et al., 2004). VLPFC-activation in BA 47 is associated with prosody perception (Wildgruber, et al., 2005) and visually presented expressive gestures (Lotze, et al., 2006). Lesions in this area result in impaired recognition of vocal expressions (Hornak, et al., 2003). Significantly increased activation during imagined compared to overt singing was only observed in the ACC and the VLPFC. Therefore, we believe that enhanced activation of these areas may reflect an increased emotional recall during imagined singing.

5.5 Conclusions

To date, several studies have proven the value of music imagery to improve various aspects of overt musical performance (Highben & Palmer, 2004; Humphreys, 1986; Lacourse, et al., 2005; Pascual-Leone, 2003; Yaguez, et al., 1998). By comparing overt and imagined singing of an Italian aria we showed a broad range of activations in partly overlapping cortical and sub-cortical areas. While overt singing predominantly activated areas processing complex motor sequences and sensory feedback/control, most of the areas related to motor processing were also active during imagery. This is of notable interest, because imagined motor performance commonly activates secondary cortical motor areas (Munzert, et al., 2009), whereas the contribution of primary motor cortex during motor imagery is less clear and still a matter of controversy (Dechent, et al., 2004). Medial and lateral premotor areas, related to the planning and preparation of movement, obviously require adequate motor memory for efficient function and should therefore also be active during imagery. In the primary motor cortex, anticipatory activity based on encoded information for upcoming movement sequences has only recently been suggested and might be less pronounced during imagery (Lu & Ashe, 2005). Yet considering the vital functions of the voice and the extensive training of articulation for verbal communication and singing throughout the life span of an artist, motor sequence encoding in the primary motor area representing voice functions might also be more pronounced. These findings have possible implications for artistic as well as for clinical needs. Since imagined motor rehearsal of skilled peripheral movement has the capacity to enhance overt performance, trained singers might particularly benefit from rehearsing the physiological underpinnings of musical performance without the sensorimotor and auditory confounds of overt production as a means of improving vocal accomplishment (Langheim, et al., 2002). By making imagery a fundamental component of their everyday practice routine, this technique could also help prevent repetitive strain injuries, one of the most common problems of extensive overt practice in professional musicians (Lederman, 2003). Clinical studies increasingly use imagery as a complementary therapeutic tool in motor rehabilitation (Page, et al., 2007). Our results suggest that non-fluent aphasic patients could particularly benefit from this approach.

A major distinction between musical imagery and overt performance, however, is the lack of motor and sensory consequences in the former. Consequently, imagery should also engage areas distinctive from those involved in overt tasks. We found additionally a large fronto-parietal network activated, which points to an increased involvement of working memory processes during mental imagery. This indicates that imagined singing is less automatic than overt singing, although subjects were experienced with imagined rehearsal. A recent finding (Desmurget, et al., 2009) showed that electrical stimulation of inferior parietal cortex triggered a strong intention to move in the contralateral hand, foot, arm or mouth areas, and when the stimulation intensity was increased participants believed they had really performed these movements although no electromyographic activity was detected. Hence, we speculate that the associative networks activated during imagery are more complex, more variable and more adjusted to the individuals' personal training history. The main advantage of imagery is that it "frees" us from the chains of external perceptual cues, providing more resources to be dedicated to other aspects of the task. This is in line with our observations that emotion-processing areas were also increasingly active during imagined singing. Although this may reflect increased emotional recall during imagery, the lack of sensory processing also allows dedicating more attention to emotional aspects of the task.

Chapter 6

Experience-dependent neural adaptation in trained singers – an fMRI study

6.1 Introduction

There is considerable evidence that cortical representation of the body may be continuously changed in response to activity, behavior and skill-acquisition (Adkins, et al., 2006; Buonomano & Merzenich, 1998). Adaptive changes in neural circuitries related to motor skill-training have also been attributed to improved performance (Nielsen & Cohen, 2008). For investigating the latter professional musicians are ideally suited since they begin with training during early childhood and practice extensively throughout their life to achieve the most astonishing levels of motor skill perfection (Jørgensen, 2002). This combination of training specificity (i.e. highly specific motor skills), training intensity, and early training onset results in extensive use-dependent neural plasticity in professional musicians (Munte, et al., 2002).

In string and keyboard players, exceptional skill has been associated with reduced activation of pre-motor areas and a shift towards more focused (or task relevant) activation in regions related to motor execution - namely in the primary sensorimotor cortex (Haslinger, et al., 2004; Lotze, et al., 2003). Adaptations on the level of functional or morphologic characterization appear not only as increased cortical representation of hand and fingers (Amunts, et al., 1997) but involve also the auditory cortex (Schneider, et al., 2002). Corresponding to Hebbian learning rules (Hebb, 1949), cross-modal plasticity has been reported as a consequence of simultaneous integration of auditory and somatosensory signals (Pantev, et al., 2003). Other experience-dependent effects in the musicians' brain involve an increased motor cortical excitability (Rosenkranz, et al., 2007), a larger cerebellum and anterior corpus callosum (Hutchinson, et al., 2003; Schlaug, et al., 1995) and even a more structured fiber tract organization in white matter (Bengtsson, et al., 2005).

With respect to the vocal system, several neuroimaging studies have investigated neural correlates of speech motor control (see Ackermann, et al., 2004). A few studies have examined simple singing (or intoned vocalization) of vowels, syllables, and words (e.g. Riecker, Ackermann, Wildgruber, Dogil, et al., 2000) as well as short musical phrases (Callan, et al., 2006; Gunji, et al., 2007) in untrained subjects, mostly in contrast to speech. Yet little is known about changes in neural activation in response to professional training of the voice (Kleber, et al., 2007; Zarate & Zatorre, 2008). Training induced neural reorganization of the vocal motor system remains poorly understood and seems to be the domain of clinical rehabilitation studies (Crinion & Leff, 2007). However, studying neural adaptation related to professional vocal training, such as classical singing, may provide a window into use-dependent plasticity of the vocal system.

Vocal production in speech and singing requires the involvement of more that 100 muscles, which depicts the functional demands, placed on that system. Although this may be the case for instrumentalists (e.g. wind instrument players) too, the vocal motor system is different in nature from the peripheral motor system. First, all sound is produced and amplified entirely within the body in the absence of visual control over movement. Second, the muscle fibers of the laryngeal and orofacial system are distinct from the skeletal (i.e. peripheral) muscle systems. They are uniquely heterogeneous and designed for fast as well as for variable contraction and fatigue resistance (for review, see Kent, 2004), which enables the vocal folds to open and close several hundred times per second within a precisely coordinated interplay of laryngeal, respiratory and articulatory activity (Dejonckere & Lebacq, 1981). Third, vocal motor activity is routinely performed at a faster discrete rate than any other human behavior, which requires a control system that mainly depends on an intrinsic reflex system (Abo-el-Enein, 1966).

These complex operations are thought to be automatic in speech once development is complete (Smith & Zelaznik, 2004). Thus, singers may build their skills on an already highly evolved system with respect to motor function in singing whereas instrumentalists develop entirely new motor patterns. Nevertheless, singers must train intensively for many years before they are able to adjust their vocal system well enough to achieve the typical sound characteristics and the level of vocal control required in classical singing (Sundberg, 1987). They learn to optimize vocal tract coordination by accessing two major feedback circuits, auditory and kinesthetic, which act as control systems on laryngeal musculature during sound production (i.e. phonation). Auditory feedback seems to be particularly relevant during early stages of vocal training and plays a fundamental role in pitch control. In contrast, the kinesthetic feedback circuit becomes increasingly important with experience and provides information on movement kinematics, subglottal air pressure and the appropriate positioning (i.e. pre-phonatory tuning) of the articulators that are eventually associated with the desired vocal outcome (Mürbe, et al., 2004). Vocal motor control in trained singers compared to laymen might therefore rely more on internal models that involve strong kinesthetic representations of vocal tract behavior in classical singing.

Based on these findings and on our previous study that involved professional singers (Kleber, et al., 2007), the present study aimed at identifying differences in neural activation patterns during overt singing of an Italian aria with respect to the training-level in classical singing. We proposed that experience in classical singing is reflected by increased involvement of cerebral areas related to somatosensory processing and performance monitoring, allowing for fast adjustments of singing related actions during musical expression.

6.2 Methods

6.2.1 Participants:

Forty-nine right-handed subjects without reported history of neurological or psychiatric disease participated in this study. Ten professional opera singers (choristers and soloists) from the Stuttgart State Opera (mean age: 38.07 years; ranging from 30 to 44; 7 female); 21 vocal students (mean age: 25.22 years; ranging from 20 to 30; 14 female) enrolled in the “artist’s diploma” program in singing performance at the State University of Music and Performing Arts Stuttgart; and 18 laymen (medical students; mean age: 23.54; ranging from 23 to 29; 6 female) from the University of Tübingen with minimal singing experience. More than 5 hours of weekly choir singing served as an exclusion criterion for the selection of laymen. Not represented in the demographic data is one amateur subject who was eliminated from data analysis due to excessive head movement (>5mm) and one randomly selected female vocal student, who was excluded to balance for gender. The Ethics Committee of the Medical Faculty of the University of Tübingen approved this study. All subjects gave written informed consent according to the guidelines of the Declaration of Helsinki.

6.2.2 Singing Experience:

Five laymen reported regular choir singing activity between one to five hours the week (mean: 2.6). None of them ever received professional vocal training. Vocal students took their first professional singing lesson on average at the age of 15 (range: 8-25). They had a mean of 9.77 years (range: 4-23) professional singing experience (counted from their first professional singing lesson) and reported an average of 18 hours (range 8-30) weekly singing practice. Opera singers took their first professional singing lesson on average at the age of 17 (range: 12-22). They had a mean of 21.03 years (range 14-28) professional singing experience and reported an average of 27.6 hours weekly singing practice (range: 20-42). An overview of singing experience is given in Figure 21.

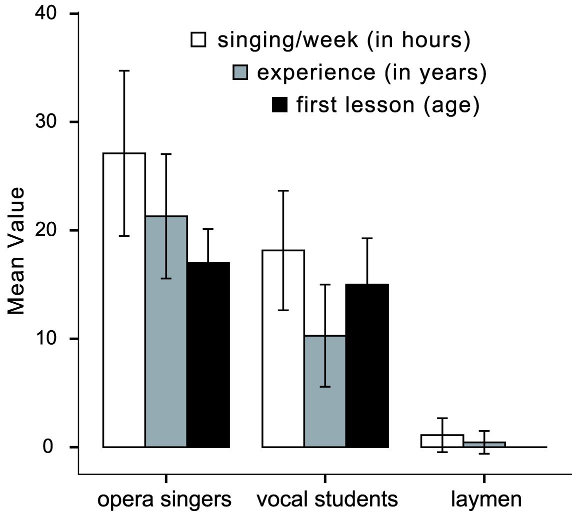

Figure 21: Singing expertise in opera singers, vocal students and laymen. White bars represent the estimated amount of weekly singing practice given in hours; grey bars represent the total time of singing experience given in years, counted from the first professional singing lesson in singers and the number of years with amateur singing experience in laymen; black bars represent the age at which subjects began with professional singing training. Respective mean values are given on the y-scale. Error bars represent one standard deviation.

6.2.3 Singing Task:

Subjects overtly sung six phrases from the first stanza of the Italian aria “Caro mio ben” (by Tommaso Giordani, 1730-1806) in an fMRI scanner, each phrase separately after a visual cue. This aria was selected because lyrics and melody are easy to memorize and the vocal demands are such that all voice types could perform it in a scanner environment. Furthermore, the phrase structure fits naturally the sparse sampling technique employed (see fMRI Data Acquisition). Although this is a popular aria in classical singing, only about 2/3 of all experienced singers reported to have sung the aria before. No subjects had actively sung or rehearsed this aria in the 5 years preceding the experiment. All subjects were provided with sheet music and audio samples of the aria three weeks before scanning. Task execution was rehearsed outside the scanner and then again within the scanner before the actual measurement. Successful task accomplishment was monitored via loud speakers.

6.2.4 fMRI Technique:

The fMRI technique applied corresponds to the one we employed previously (Kleber, et al., 2007). Whole head scans (66 volumes per block) were performed with a 1.5 Tesla whole body Scanner (Siemens Vision) using echo planar imaging (EPI; TE: 40 ms; TR: 10 sec, TA: 3 sec, 36 transversal slices of 3 mm thickness and 1 mm gap, matrix 64*64). Head movement during scanning was minimized by a rubber foam head restraint. A sparse sampling method was employed to delineate singing related movement from volume acquisition and to allow for undisturbed monitoring of the own voice during singing. The design involved six consecutive singing periods (phrases from the Italian aria) of 10 sec TRs respectively alternating with six consecutive baseline periods in which subjects were inhaling/exhaling only (Figure 22). Thirty-three singing conditions and thirty baseline conditions with a total duration of approximately 11 minutes were measured. To control for effects related to excessive inhalation, baseline measurements required subjects to inhale at the same time points as in the singing condition, followed by slow exhalation. Each TR contained a delay of 7 sec in which visual signals were presented, indicating inhalation (1.5 sec) and singing (3 sec) followed by a TA of 3 sec. Within the 7 sec delay, timings of inhalation and task execution were systematically varied by ±500 ms to ensure that we covered the peak amplitude of the hemodynamic response for each task. The first 3 EPI data sets of each session were discarded prior to analysis to allow for T1-saturation effects. T1 weighted images (MPRage; 176 sagittal slices with 1mm effective thickness) were measured for anatomical reference.

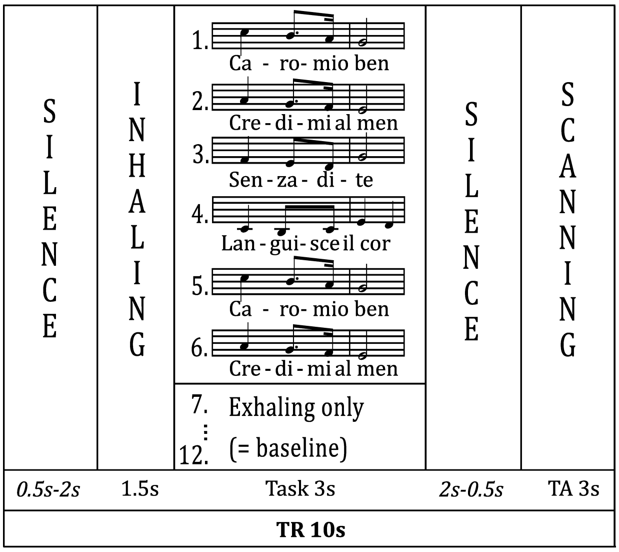

Figure 22: Sparse-sampling fMRI design: six consecutive singing periods alternated with six consecutive baseline periods in which subjects were inhaling/exhaling only. Each TR (10 seconds) contained a delay of 7 seconds in which visual signals were presented, indicating inhalation (1.5 seconds) and singing (3 seconds) followed by a TA of 3 sec. The timings of inhalation and task execution were systematically varied by ±500 ms within a 1.5s time window to ensure to cover the peak amplitude of the hemodynamic response. Silences were adjusted accordingly. A total of thirty-three singing periods and thirty baseline periods were measured.

6.2.5 fMRI Data Analysis:

Data were analyzed with SPM5 (Wellcome Department of Imaging Neuroscience) running on Matlab (MathWorks Inc; Natick, MA; USA). Each individual scan was realigned to the first image of each session to correct for movement artifacts. The realigned data were spatially normalized to the MNI-template and resliced with 3*3*3 mm. The resulting images were smoothed with a 9 mm (full width at half maximum) Gaussian filter. The movement parameters estimated during the realignment procedure were introduced as covariates of no-interest into the general linear model. Individual statistical maps (fixed effect) were calculated for each subject (main effect). First level contrast images of each subject were then used for group statistics calculated as random effects analysis at the 2nd level. A one-way ANOVA with three groups (laymen, vocal students and opera singers) was performed. Non-sphericity correction was applied to account for the possibility of unequal variances due to the different group size of opera singers. Age was added as a nuisance covariate to account for differences between groups. Gender was included as a nuisance covariate to correct for an uneven gender distribution within groups. We measured experience-dependent differences in BOLD response based on expertise-levels in classical singing by comparing cerebral activation during overt singing of an Italian aria in three groups: highly experienced opera singers, music conservatory vocal students who train to become professional singers, and laymen with little singing experience. Main effects of overt singing were calculated for the whole brain in each group followed by a conjunction analysis (conjunction null, Friston, et al., 2005) .

Based on these data as well as on our previously published results (Kleber, et al., 2007), a-priori defined regions of interest (ROIs) were selected for all subsequent analyzes. Therefore, a single anatomical mask was formed from anatomical ROIs to avoid alpha-error inflation due to multiple testing. For regions already cytoarchitectonically mapped we used ANATOMY (Eickhoff, et al., 2005). The following regions were included: primary sensorimotor cortex (Brodmann's area (BA) 1,2,3ab and BA 4ap selected from ANATOMY), premotor cortex and supplementary motor area (BA 6), primary auditory cortex (TE1.0, 1.1, 1.2) and superior temporal gyrus, BA 44 and 45, superior and inferior parietal lobe. For all other regions selected, we used the atlas of Tzourino-Mazoyer (AAL, Tzourio-Mazoyer, et al., 2002): the dorsolateral prefrontal cortex (occupying the posterior superior and middle frontal gyrus), the cerebellar hemispheres, limbic areas (insula, anterior cingulate cortex, amygdala, hippocampus), the thalamus, the basal ganglia (putamen, pallidum, caudate nucleus) and the temporal poles.

Post-hoc contrasts (Student's t-tests) were applied to assess between-group effects for the following comparisons: vocal students versus laymen, opera singers versus laymen and opera singers versus vocal students. Statistical inference for all analyses was based on the resulting t-statistics in the mask using a p-value of 0.001 uncorrected for multiple comparisons. T-values of significant activations of the highest activated voxels were given for MNI-coordinates and were assigned to anatomical regions. Where applicable, probabilities for anatomical localization of activation maxima were calculated (Eickhoff, et al., 2005). While the former analyses addressed only differences between groups as defined by their singing standard (laymen, students, professionals), we performed an additional regression analysis to assess the effect of singing expertise on BOLD response. We entered the estimated total amount of singing practice across all subjects in a linear regression model and controlled for age and gender by adding them as nuisance regressors. The accumulated amount of singing practice was approximated by multiplication of singing experience (based on the age of commencement with formal singing training in experts or years of amateur singing experience in laymen respectively) with the estimated weekly singing practice (years experience*weekly singing practice). The resulting value was subsequently transformed to a logarithmic scale to obtain a more uniform distribution of data along the abscissa, more suitable for linear regression. Activation maps were superimposed on the MRIcron ch2better template, which fits best to the MNI-space (http://www.sph.sc.edu/comd/rorden/mricron/).

6.3 Results

6.3.1 Main effects:

Main effects of each group were calculated for the whole brain. Peak activations included bilateral primary sensorimotor cortex (M1, S1), secondary somatosensory cortex (SII), premotor cortex, primary auditory cortex (A1), the temporal poles, inferior and superior parietal cortex, visual cortex, the insula and the cingulate cortex. At the subcortical level activations were found in the thalamus, the basal ganglia (putamen, pallidum, caudate nucleus) and the cerebellum. Full statistical tables of respective main effects are available in Supplementary Table 3-5.

6.3.2 Conjunction analysis:

Common activation in all groups was found in the superior temporal gyrus including the primary auditory cortex (TE 1.0-3), the bilateral primary sensorimotor cortex (M1 and S1) focused in the cortical representation of the orofacial area, bilateral SII (OP4), the supplementary motor area and right inferior premotor cortex (BA6), as well as in the cerebellum (Larselle's lobule VI) and in left inferior parietal cortex (Figure 23A). Full statistical information of the conjunction analysis is available in Supplementary Table 6.

6.3.3 Vocal students versus laymen:

Vocal students showed increased activation in bilateral S1 (BA 1) in the somatotopic area of articulation and the larynx (see Table 3 and Figure 23B). Further cortical activation maxima could be assigned to the angular gyrus and dorsolateral prefrontal cortex (DLPFC) bilaterally, and left temporal pole. Significant differences in vocal students compared to laymen involved also the right pallidum and the right cerebellum in Larsell's lobule HVI.

6.3.4 Opera singers versus laymen:

Opera singers showed increased activation in left S1 (BA 1) and overlapping S1/M1 (BA 4p and 3a) in the right hemisphere (see Table 4 and Figure 23C). Activation maxima were also found in the left SII (OP4), the left temporal pole and right inferior parietal cortex. Differences in opera singers compared to laymen involved furthermore the basal ganglia (putamen, pallidus and caudate nucleus), the left thalamus and the cerebellum in Larsell's lobules HVI and Crus 1 and 2 bilaterally.

6.3.5 Opera singers versus vocal students:

The comparison opera singers versus vocal students (see Table 4 and Figure 23D) revealed increased activation in right hemispheric overlapping S1/M1 (BA 4p and 3a). Further increased activation in opera singers involved the right putamen, parietal cortex and the cerebellum (right Larsell's lobule HVI, bilateral Crus1, left Crus 2 and left lobule HIV-V).

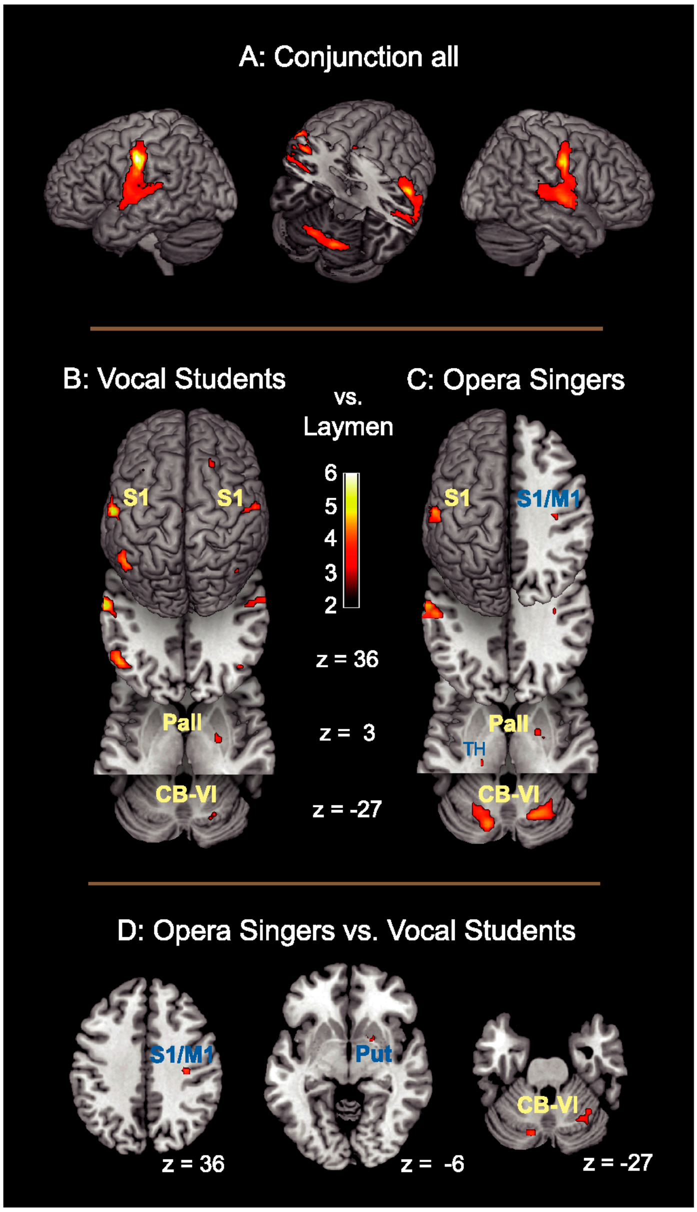

Figure 23: Brain activation associated with overt singing. A: Common activation in opera singers, vocal students and laymen was found in right M1, bilateral S1, SII, A1, SMA, lateral premotor cortex and the cerebellum. B: Increased activation in vocal students compared to laymen involved bilateral S1, right globus pallidus, right cerebellum (Lobule H VI), SMA, IPC, and DLPFC. C: Increased activation in opera singers compared to laymen was found in bilateral S1 and left SII, right M1 and IPC but also in the right putamen, right globus pallidus, and right caudate nucleus, the left thalamus and the bilateral cerebellum. D: Increased activation in opera singers compared to vocal students involved right M1 and S1, right putamen and bilateral cerebellum.

Table 3: MNI coordinates and peak *t*-score values for region-of-interest (ROI) tests based on the post-hoc contrast (Student's t-test) vocal students versus laymen ($P < 0.001$ uncorrected). For anatomic labeling we used AAL (Tzourio- Mazoyer et al., 2002). Areas already cytoarchitectonically mapped (*) were assigned using the Anatomy Toolbox (Eickhoff et al. 2005) - probabilities are given in brackets. Abbreviations: S1, primary somatosensory cortex; SMA, supplementary motor area; IPC, inferior parietal cortex; PGa, DLPFC, dorsolateral prefrontal cortex; BA = Brodmann's area; L = left; R = right. PGa and PGp correspond to rostral and caudal divisions of the angular gyrus respectively.

		Vocal Students vs. Laymen				
Region		*x*	*y*	*z*	*t*	Assigned to
Sensorimotor						
S1	L	-63	-15	36	5.66	*Area 1 (50%)*
	R	60	-12	39	4.12	*Area 1 (90%)*
SMA	L	-6	-15	57	3.60	*Area 6 (80%)*
Parietal						
IPC	L	-48	-57	33	4.85	*PGa (60%)*
	R	45	-63	36	3.79	*PGp (60%)*
Frontal						
DLPFC	R	21	27	54	3.64	
	L	-27	21	39	3.49	
Temporal						
Temporal Pole	L	-48	6	-24	3.95	
Subcortical						
Globus Pallidus	R	24	-3	3	3.82	
Cerebellum						
Lobule H VI	R	24	-66	-27	3.58	

Table 4: MNI coordinates and peak *t*-score values for region-of-interest (ROI) tests based on the post-hoc contrasts (Student's t-test) opera singers versus laymen and opera singers versus vocal students ($P < 0.001$ uncorrected). For anatomic labeling we used AAL (Tzourio- Mazoyer et al., 2002). Areas already cytoarchitectonically mapped (*) were assigned using the Anatomy Toolbox (Eickhoff et al. 2005) - probabilities are given in brackets. Abbreviations: S1, primary somatosensory cortex; SII, secondary somatosensory cortex; SMA, supplementary motor area; IPC, inferior parietal cortex; SPL, superior parietal lobe; DLPFC, dorsolateral prefrontal cortex; L = left; R = right.

	Opera Singers vs. Laymen					Opera Singers vs. Vocal Students				
Region	*x*	*y*	*z*	*t*	Assigned to	*x*	*y*	*z*	*t*	Assigned to
Sensorimotor										
S1	-60	-15	36	4.84	*Area 1 (60%)*					
	36	-18	39	3.93	*Area 3a (60%)*	33	-21	36	4.10	*Area 3a (50%)*
M1	36	-18	39	3.93	*Area 4p (70%)*	36	-18	39	4.11	*Area 4p (70%)*
SII	-66	-12	24	3.34	*OP 4 (30%)*					
Parietal										
IPC	42	-33	18	3.59	*IPFcm (40%)*					
Precuneus						-12	-54	75	3.84	
Temporal										
Temporal Pole	-48	6	-24	3.35						
Subcortical										
Putamen	21	6	-6	3.86		18	6	-6	3.71	
Globus Pallidus	24	0	3	3.72						
Caudate	12	9	-9	3.42						
Thalamus	-21	-21	6	3.84						
Cerebellum										
Crus 1	-18	-72	-30	4.72		-18	-72	-30	4.22	
Crus 1	42	-69	-30	3.70						
Crus 2	-21	-84	-33	4.14		-18	-84	-33	4.03	
Lobule H VI	21	-66	-24	4.62		33	-57	-27	3.60	
Lobule H VI	27	-63	-30	4.49		27	-63	-30	3.54	
Lobule H VI	-21	-63	-27	3.93						
Lobule H IV-V						-30	-42	-24	3.36	

6.3.6 Singing practice effects:

Shown in Figure 24, regression analysis based on the accumulated amount of singing practice revealed positive correlations with cortical activity in bilateral S1 (left BA 3b and bilateral BA 1), located in the representation area of larynx and articulators. No effect was found for the primary motor cortex. Positive correlations with singing practice were also found in bilateral DLPFC (right 24, 27, 54; t(48)=4.25, r=0.40; left -36, 21, 54; t(48)=3.76, r=0.43), and bilateral angular gyrus (right 42, -63, 39; t(48)=3.95, r=0.40; left -48, -60, 33; t(48)=4.77, r=0.44). Positive correlations with subcortical areas were observed in the right pallidum (24, -3, 3; t(48)=3.82, r=0.40).

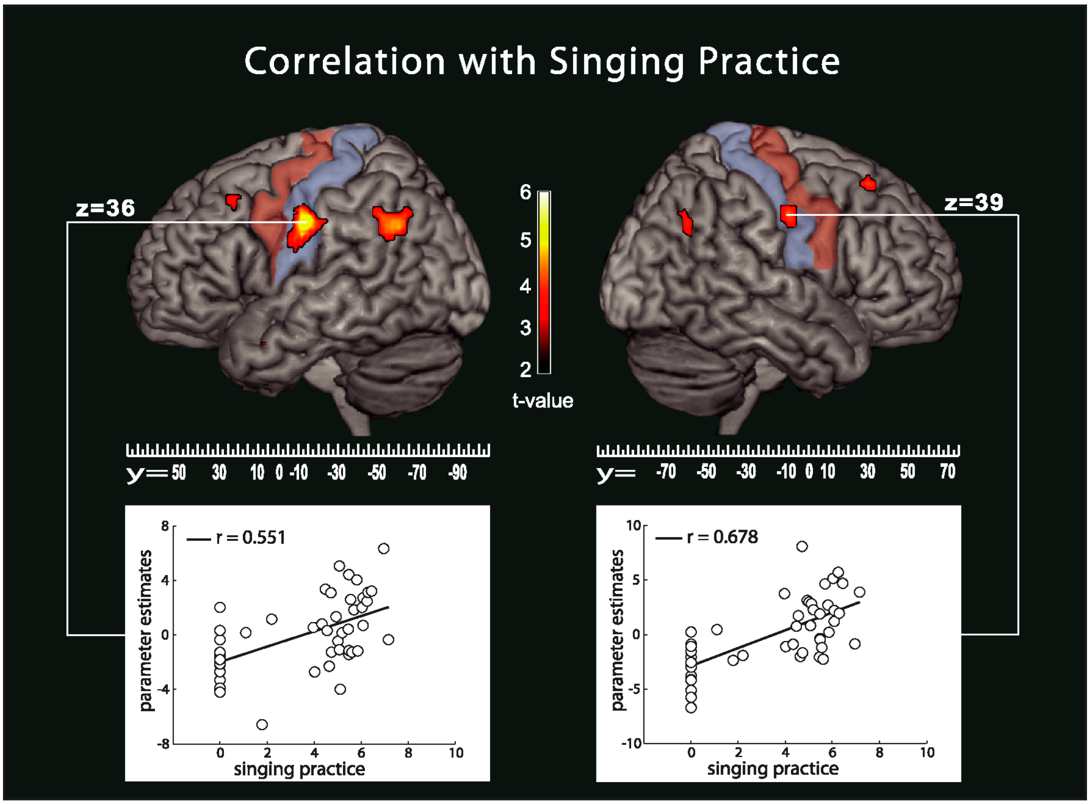

Figure 24: Results from the regression analysis including all subjects (opera singers, vocal students and laymen). The y- scale of the regression plots depicts mean corrected parameter estimates. Statistical inference was based on the resulting t-statistics for each voxel using a p-value of 0.001 uncorrected for multiple comparisons.

6.4 Discussion

We aimed at identifying changes in functional brain activation with regard to expertise level and practice quantity in classical singing. We provided the first evidence that professional training of classical singing gives rise to increased activation of areas associated with kinesthetic information processing from the larynx and the articulators. Additional activations involved cortical regions related to the somatosensory guidance of motor activity and subcortical motor memory areas. We assume that our results reflect experience-dependent changes in functional activation. In contrast to previous findings of experience-dependent neural adaptation in musical instrumentalists, high levels of singing practice did not generally result in increased activation in M1, only in the right hemisphere of the most experienced opera singers.

6.4.1 Sensorimotor cortex

Studies suggest an experience-dependent reorganization of the primary motor cortex associated with the retention and acquisition of motor skills (Karni, et al., 1995). Anatomical and functional reorganization of contralateral sensorimotor cortex in response to manual motor skill training has been repeatedly reported for professional musicians who learned to play an instrument (Amunts, et al., 1997). These changes were also directly linked to the type of musical practice, for example to long-term training of the violin (right M1 reorganization, Lotze, et al., 2003) or piano playing (bilateral M1 reorganization, Hund-Georgiadis & von Cramon, 1999). However, the total practice in classical singing did not reveal changes in M1. Instead, it was strongly correlated with increased activation in ventral primary somatosensory cortex. The reported results suggest that somatosensory feedback might play a particularly important role in the development of classical singing skills. This hypothesis is further supported when considering the physiology of vocal production, which involves two main mechanisms: 1) the generation of sound at the laryngeal level through an interaction of vocal fold muscular forces with air pressure from the lungs and 2) the amplification of sound in the resonating vocal tract by an appropriate shaping of the articulators. Studies on speech articulation have demonstrated that proprioceptive feedback already provides sufficient information on the positions of the articulators to achieve precise articulatory movements, independent from auditory feedback (Nasir & Ostry, 2006). However, articulatory precision in speech is primarily relevant for intelligibility. In singing, additional amplification of certain frequency partials through articulatory shaping of the resonating vocal tract subserves the production of style specific timbre (Ekholm, et al., 1998) that allows the voice to be audible despite the competing sound from an orchestra (Sundberg, 1974). Experienced singers have been extensively trained to focus on timbre, vowel quality, intonation accuracy and sound intensity in the context of increased emotional expression. At the level of sound production, this requires enhanced control over laryngeal motor activity in conjunction with sub-glottal air pressure, whereby the vocal folds oscillate rapidly between 50 and 1500 Hz (e.g. in high sopranos). To account for the speed of vocal fold vibration, the laryngeal motor system is based on an intrinsic reflex system (Abo-el-Enein, 1966). Vocal motor control (i.e. changes in tension and positioning of the vocal folds) during sound production depends largely on kinaesthetic feedback from laryngeal mechanoreceptors (Wyke, 1974a, 1974b). To optimize this vocal motor system, singers utilize a combination of auditory and kinesthetic feedback for vocal control. The significance of auditory feedback for singing has been reported in several studies but the kinesthetic feedback circuit seems to be particularly improved after years of training in classical singing (Mürbe, et al., 2004). This possibly corresponds to the development of more accurate feed-forward commands of vocal tract adjustments based on kinesthetic expectations (or "kinesthetic motor memory"). We therefore believe to have found the first evidence that excellence in classical singing might be accompanied by increased involvement of primary somatosensory cortex, reflecting enhanced kinesthetic motor control for sound production in singing. As the level of peak activation in bilateral S1 roughly coincides with the transversal plane of a larynx specific area in the primary motor cortex as reported by Steven Brown and colleagues (Brown, et al., 2009; Brown, et al., 2007), our results potentially reflect the improved processing of laryngeal feedback within the somatotopic orofacial area.

Increased activation in M1 was only found in the non speech-dominant right hemisphere of opera singers, suggesting that prolonged engagement in classical singing might indeed also result in increased activity patterns in primary motor areas. It has been demonstrated in studies comparing language and singing that there is a significant lateralization in M1 for speech production in the dominant and for melody production in the non-dominant hemisphere (Riecker, Ackermann, Wildgruber, Meyer, et al., 2000; Wildgruber, et al., 1996). Similar lateralization effects have been reported for the auditory cortex (Zatorre & Gandour, 2008). In singing, word production is reduced compared with speaking in favor of an increased focus on vowel production as the carrier of the melodic content. Indeed, the concept of vowel quality plays a very important role in classical singing and is related to timbre and resonance shaping of the vocal tract. The principal motor components in vocal production may have already been exhaustively trained in verbal communication, hence no general increases of M1 activation were found with vocal training. However, the production of words in melody may lead to an increased activation of the right M1 after years of extensive training and performance of classical singing. The musical piece's sensorimotor elements may then change from a sequence oriented processing mode to a more "gestalt" processing mode with prolonged experience and constitute a whole inseparable pattern, as expressed in verbal accounts of highly skilled musicians.

6.4.2 Auditory cortex:

Several studies have demonstrated that training induced improvement of auditory skills corresponds with increased auditory cortical activation (Gaab, et al., 2006; Pantev, et al., 1998). However, we could not observe such differences between groups. A recent study showed increased auditory activation in experienced singers during real-time shifted auditory feedback (Zarate & Zatorre, 2008). The auditory cortex of singers may track unexpected changes in vocal production more efficiently than in untrained subjects as reflected by increased responses of auditory cortex to a mismatch of actual and intended feedback. In our sparse-sampling paradigm, all singers could monitor their voice undisturbed. Thus, increased auditory activation might not be expected. In addition, the perception of self-produced utterances can attenuate auditory activation in A1 in contrast to passive perception of the same sounds (Houde, et al., 2002). In monkeys, approximately 50% of call-responsive neurons in the auditory cortex are inhibited during vocalization (Muller-Preuss & Ploog, 1981). An expected match between predicted and actual sensory consequences in trained singers could thus result in the filtering of self-produced auditory information, which may contribute to our results.

6.4.3 Sensory association cortex

We found increased activation in bilateral inferior parietal cortex in relation to the total singing practice, and in the comparison of vocal students versus laymen. Opera singers versus laymen showed right hemispheric inferior parietal cortex (IPC) activation. The functional roles of the IPC involve auditory-verbal working memory (Paulesu, et al., 1993) and short-term memory for musical pitch, pronounced in the right hemisphere of musically trained subjects (Gaab, et al., 2003). The angular gyrus, correlated with singing practice in our study, has been associated with the processing of semantic features in the left hemisphere (Price, 2000), while its right hemispheric counterpart has shown to be involved in the computation of action awareness (Farrer, et al., 2007). More generally, the IPC has been associated with the integration of sensory and motor signals for the sensory guidance of movements (Fogassi & Luppino, 2005) utilizing also somatosensory information (Jancke, et al., 2001). A model of speech motor control (Guenther, et al., 2006) posits a role of the IPC in a feedforward control mechanism of articulatory motor commands. In this model, the IPC acts as a control system for somatosensory feedback from the vocal tract by comparing the actual kinesthetic feedback with the kinesthetic "expectation" for the produced sound. Accordingly, we assume that experience-related activity of bilateral IPC reflects enhanced processing of a performance module that detects a mismatch between intention, action and consequences and thus allows for more rapid sensorimotor adaptations/corrections in trained singers.

6.4.4 Performance monitoring

We also found increased bilateral DLPFC activation in the comparison of vocal students versus laymen and in the regression analysis based on the total amount of accumulated singing practice in all subjects. The DLPFC is associated with higher cognitive functions such as working memory (Baddeley, 1992). Training induced improvement of working memory has been reported to be accompanied by increased involvement of DLPFC (Olesen, et al., 2004). Studies report its involvement in executive processes such as divided attention between sensory modalities (Johnson, et al., 2007) and goal-directed attention (Weissman, et al., 2004), which includes also the monitoring of self generated performance and preparation of forthcoming sequential actions (Pochon, et al., 2001). Where left DLPFC seems to be preferentially activated during extensive semantic elaboration, right DLPFC is recruited with increasing monitoring demands (Sharp, et al., 2004). In fact, classical singers are not only required to perfectly integrate vocal performance and musical expression but must also simultaneously act while attending to several cues. This is particularly the case in opera singers, who also combine singing with acting on stage. Perhaps the basic motor mechanisms needed to perform the aria are functioning more autonomously in professionals, so that they have more resources for performance monitoring, and perhaps for musical expression, than non-professionals.

6.4.5 Cerebellum, basal ganglia and the thalamus

Cerebellar function is highly correlated with the timing of complex sequential action and speech movements (Braitenberg, et al., 1997). The cerebellum utilizes kinesthetic feedback to monitor and to coordinate movements, thereby acting as a sensorimotor predictor based on a combination of sensory inputs and efference copies of motor commands (Bastian, 2006). Studies on movement control show that 80-90% of the cerebellar signal is attributable to sensory information processing (Jueptner & Weiller, 1998). A recent study could demonstrate that the activation pattern of the cerebellum during conscious monitoring of rhythmic auditory motor synchronization corresponds to the activity pattern found in S1, IPC and DLPFC (Thaut, et al., 2009). We observed increased cerebellar activation in experienced singers compared to laymen, with the largest activation found in the most experienced opera singers. We assume that cerebellar contribution to vocal motor control becomes more pronounced with increased singing experience. This finding is congruent with a morphologic observation that increased cerebellar volume is associated with long-term training with an instrument (Hutchinson, et al., 2003). However, it stands in contrast to functional studies showing decreased cerebellar activation during manual coordination tasks in professional pianists (Haslinger, et al., 2004; Koeneke, et al., 2004). Indeed, there is some uncertainty in the literature as to whether or not more experience leads to a stronger activation or less activation. Regions could also become more efficient in their operation recruiting a smaller number of active neurons to perform a task (Jancke, et al., 2000; Krings, et al., 2000). When considering this conflicting evidence, attention must been drawn towards the tasks employed in these studies. Two studies involved bimanual finger tapping with either one (self-paced, Jancke, et al., 2000) or four fingers (externally paced, Haslinger, et al., 2004). One study investigated the self-paced performance of five finger sequences in the right hand (Krings, et al., 2000) and yet another required two finger coordination to move a cursor along a screen (Koeneke, et al., 2004). Touching a key on a console compared to touching a piano key may result in fundamentally different kinesthetic perception, which is an important aspect in the control of motor sequences in piano playing. The lack of auditory feedback may furthermore contribute to reduced activation since trained musicians show stronger brain responses when processing multimodal information (auditory and somatosensory) compared to the sum of unimodal responses (Schulz, et al., 2003). In contrast, the singing paradigm we employed together with the sparse sampling technique constitutes a realistic demand on motor control in singing, providing both accurate kinesthetic and auditory feedback. The unusual position with restrained head movement in the scanner may furthermore require additional corrective responses. Cerebellar activation was significantly correlated with singing experience but most pronounced in opera singers, who must routinely adapt their vocal system to unusual postures during singing as part of their stage play. It is likely that this group has developed an adaptive system to cope with such demands, requiring increased cerebellar processing.

Activation in the basal ganglia was consistently found in the group comparisons and in the regression analysis. More precisely, we observed increased activation in the right pallidum of vocal students compared to laymen and in the regression analysis. Opera singers showed additionally increased involvement of right putamen and caudate nucleus compared to laymen along with the left thalamus. The putamen was also significantly more active in opera singers than in vocal students. The basal ganglia are part of a distributed network involved in procedural learning, which represents an integral part of musical skill development (Doyon & Benali, 2005). Plasticity within the basal ganglia is related to alterations in the transfer of information and may represent a key neural substrate for adaptive motor control and procedural memory (Kreitzer & Malenka, 2008). Increased activity in the putamen has recently been reported for experienced singers compared to amateurs during simple singing with pitch shifted feedback (Zarate

& Zatorre, 2008) while damage to the basal ganglia gives rise to deficits in laryngeal and articulatory motor control (Merati, et al., 2005). We speculate that extensive training in classical singing leads to alterations in this system, eventually resulting in more efficient information processing and implicit motor control. The ventral thalamus, active in opera singers, integrates somatosensory information from the face and lip area and projects to tertiary cortical areas such as the inferior parietal lobe (Lenz & Dougherty, 1998). A fast projection of somatosensory input and integration with other modalities might be an essential characteristic of professional singing. However, interpretation of thalamic activity is limited by methodological restrictions that do not allow a more precise anatomical differentiation.

6.5 Conclusions

One of the questions in motor learning is the role of sensory mechanisms in movement control. Previous research in this domain has shown that without effective sensorimotor integration, motor learning and the relearning of motor sequences is not possible. Basically, there are two paths for investigating this issue. One is to investigate deficient motor skills compared to average levels of performance, which informs us about the basic principles of motor control and sensorimotor integration, and their relevance for motor learning and relearning. The other one investigates outstanding levels of motor performance compared to standard levels, which allows indentifying the principles of experience dependent neural plasticity in healthy individuals and the role of sensorimotor integration, when higher levels of motor control are required. Professional musicians are perfectly suited for this approach, as they dedicate a whole lifetime to deliberate practice of fine motor skills.

Several neuroimaging studies have demonstrated enhanced sensorimotor integration and experience dependent neural plasticity with respect to the development of peripheral motor control. However, training related neural changes in the vocal system are largely unknown and evidence is limited to the recovery of language function after stroke. We have studied neural activation in association with singing skill levels in healthy subjects. Our results provide the first evidence for adaptive changes in the neural networks underlying vocal control for singing. In particular, we found that training and performance of classical singing give rise to increased activation in circumscribed cortical areas reflecting enhanced somatosensory feedback processing from the larynx and the articulators together with an increased involvement of implicit motor memory areas at the subcortical and cerebellar level. Although the voice may be a special case, because it is only indirectly accessible through auditory and somatosensory feedback, and because its function is based on an intrinsic reflex system, these findings may also have implications for motor skill development in other motor domains. However, in contrast to previous studies of peripheral movements, extensive practice did not generally result in more pronounced activity of primary motor cortex. We speculate that the principal motor component in singing has already been exhaustively trained through verbal communication, which in turn sheds new light on the importance of sensory systems in motor control. That is, increased processing of somatosensory information may be successfully used to fine-tune existing motor patterns in order to match the demands of classical singing, irrespective of functional changes in the primary motor cortex.

In addition to these results, activations attributed to professional singing practice included also regions involved in action monitoring and sensory guidance of motor activity. These activations likely reflect experience-dependent adaptive changes of increased action awareness and appropriate response selection that emerge with extensive skill training. These results complement the concept of enhanced sensorimotor control for development of superior vocal motor control for singing.

Chapter 7 Summary

The present dissertation aimed at investigating the neural networks underlying vocal control for singing. We addressed this issue in two neuroimaging experiments with professionally trained classical singers. In the first study, we studied the neural activation patterns associated with overt and imagined singing of an Italian aria. In the second study, we examined functional changes of the central nervous system associated with the level of vocal skill development and accumulated singing practice. Adopting the methods employed in study one, we investigated overt singing in three groups: experienced professional opera singers, vocal students who train to become professional singers, and laymen with little or no singing experience.

The technique we employed was functional magnetic resonance imaging (fMRI), which allows in-vivo measurements of task related brain activity. The functional MRI takes advantage of two characteristics: (a) the neurovascular coupling in the brain (i.e. the relationship between local neural activity and subsequent changes in cerebral blood flow (CBF)) and (b) the different magnetic properties of oxygenated and deoxygenated venous blood. Deoxygenated blood is slightly paramagnetic and causes inhomogeneities in a magnetic field in contrast to oxygenated blood. During inactivity, the relative amount of deoxygenated blood in the brain is increased, which strengthens the field inhomogeneities. Increased activity triggers increased blood flow to the active area, supplying it with oxygen and glucose, which leads to overcompensation with oxygenated blood. Consequently, the overrepresentation of oxygen results in a reduction of field inhomogeneities. This change in levels of blood oxygenation affects the signal that can be detected during magnetic resonance imaging and enables the determination of task related brain activity. The cognitive neurosciences routinely employ fMRI to determine the neurobiological correlates of behavior. We adopted a sparse-sampling fMRI technique and took advantage of the time lag of the hemodynamic response. That is, brain activity is only scanned after task completion so that subjects can execute their tasks during the silences preceding the scans. This grants unaffected auditory feedback during singing.

In the first study, we assessed 16 professionally trained classical singers during overt and imagined singing of an Italian aria. Over a period of 11 minutes, the participants repeatedly sang six phrases of the Italian aria alternating with six baseline periods in which subjects were instructed to relax in an eyes open position, to breath normally, and to avoid thinking about the next activation block. In a subsequent session, the participants performed the same task only in their minds. Brain activation of overt singing versus baseline revealed a large bilateral network comprising of primary and secondary sensorimotor and auditory cortices. This neural system was symmetrically distributed between the two sides of the brain. In contrast, studies on speech articulation often show additional left lateralized activations in the posterior temporal lobe, the anterior insula and the inferior frontal gyrus. Our results confirmed that singing engages a larger neural network that involves also right hemisphere homologues of classical left hemisphere perisylvian language areas. This may offer an explanation why non-fluent aphasic patients can often sing a song but not speak fluently. Sub-cortical motor areas comprised the cerebellum, the thalamus, the medulla and the basal ganglia and were strongly activated during overt singing. In contrast, areas associated with emotional processing were only vaguely activated. Overt singing directly compared with imagined singing confirmed the association of overt singing with cortical (primary somatosensory cortex, right A1, bilateral primary motor; M1) and sub-cortical (right cerebellum, medulla) sensorimotor activation.

Interestingly, both overt and imagined singing revealed partly overlapping activation. That is, most of the primary and secondary sensorimotor areas that were active during overt singing were active during imagery too, with the exception of primary auditory cortex. We concluded that mental rehearsal might be particularly useful for enhancing overt vocal performance in singers, thus reducing the risk of repetitive strain injuries related to extensive overt practice, and that motor imagery might have clinical implications as a complementary therapeutic tool in the recovery of vocal function after stroke.

Imagined singing also resulted in additional activation within a fronto-parietal network, including several emotion-processing areas, too. The comparison of imagined with overt singing confirmed these results. Interestingly, Wernicke's and Broca's areas and their right hemispheric homologues were active during imagery, too. The additional prefrontal and limbic activation and the larger network of higher order associative functions observed during imagery gave rise to the speculation that the associative networks activated during imagery were more complex and more adjusted to the individuals' personal training history. Because of the lack of external perceptual processing, imagery allows the allocation of more resources to the cognitive and emotional aspects of a given task.

The second experiment addressed the issue of neural plasticity of the vocal system in healthy subjects. Over the last 15 years, several studies demonstrated that extensive training with a musical instrument results in corresponding reorganization of neural networks and brain morphology, reflecting enhanced sensorimotor integration. In contrast, evidence for neural adaptation in conjunction with vocal motor training is sparse and mainly derived from clinical rehabilitation studies. To date, no study has shown adaptive changes in the neural networks governing vocal control in healthy subjects and the role of sensory feedback in the central nervous modulation of vocal motor control was not yet clear. By employing the methods of study one, we investigated neural correlates of overt singing based on the level of singing skills in three groups: opera singers, vocal students, and amateur singers. Data were analyzed by calculating statistical differences between groups and by employing a regression analysis based on the accumulated amount of singing practice including all groups. Our results provided the first evidence for neural plasticity of the vocal system corresponding to the training and development of vocal skills in healthy individuals. We could show that the ability to adapt the vocal system appropriately for classical singing is accompanied by increased functional activation of bilateral primary somatosensory cortex, representing proprioceptive feedback from the articulators and the larynx. We thus demonstrated (a) the principal capacity of neuroplasticity within a highly specialized and over-trained system, and (b) the importance of somatosensory feedback for skilled performance of motor sequences distinct from peripheral movement. Except for the most experience opera singers, we did not find general experience-dependent changes in the primary motor cortex. We speculate that instead increased activity in receptive systems subserve the precise transformation of highly automatic speech motor sequences into appropriately adjusted motor patterns for singing.

Additional skill dependent activations were found in the inferior parietal lobe and bilateral dorsolateral prefrontal cortex, whereas singing experience also increasingly activated the basal ganglia, the thalamus and the cerebellum. A regression analysis based on the accumulated singing practice confirmed that the training and performance of classical singing activates a network for enhanced kinesthetic motor control and sensorimotor guidance at the cortical level together with increased involvement of implicit motor memory areas at the subcortical and cerebellar level.

Considerations with respect to future studies should take into account the limitations of this study. The inferences drawn with respect to neural plasticity of the vocal system do not allow differentiating between the respective roles of verbal and musical components of the singing training. Trained actors might represent an adequate control group for investigating the relative contribution of these two components for adaptive neural processes and sensorimotor integration, as this group intensively trains speech motor skills.

Furthermore, it would be interesting to investigate the relative contribution of auditory and somatosensory feedback in sensorimotor control of the vocal system. To address this issue, we propose to study neural activation maps during singing with and without selectively masked sensory feedback in a pitch reproduction task. In addition, one might examine sensorimotor integration in inaccurate singers who have unimpaired pitch discrimination as a way of understanding how these systems may become decoupled, which might also have implications for understanding voice disorders.

Finally, diffusion tensor imaging (DTI) tractography between professionals and untrained subjects would allow to distinguish between plasticity changes in anatomical connectivity, whereas functional connectivity analysis of overt singing and imagery would provide more detailed information as to how the brain areas under consideration cooperate during these tasks.

Keywords: neural plasticity, imagery, fMRI, singing

Bibliography

Abo-el-Enein, M. A. (1966). Laryngeal myotatic reflexes. *Nature, 209*(5024), 682-686.

Ackermann, H., Ricker, A., & Wildgruber, D. (2004). Cerebral organization of speech motor control: Functional imaging data. In B. Maassen, W. Hultsijn, R. D. Kent, F. M. Peters & P. H. M. M. van Lieshout (Eds.), *Speech motor control in normal and disordered speech* (pp. 85-111). Oxford: Oxford University Press.

Adachi, S., & Yu, J. (2005). Two-dimensional model of vocal fold vibration for sound synthesis of voice and soprano singing. *117*(5), 3213-3224.

Adkins, D. L., Boychuk, J., Remple, M. S., & Kleim, J. A. (2006). Motor training induces experience-specific patterns of plasticity across motor cortex and spinal cord. *J Appl Physiol, 101*(6), 1776-1782.

Alexander, G. E., & Crutcher, M. D. (1990). Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci, 13*(7), 266-271.

Amaro, E., Jr., & Barker, G. J. (2006). Study design in fMRI: basic principles. *Brain Cogn, 60*(3), 220-232.

Amunts, K., Schlaug, G., Jäncke, L., Steinmetz, H., Schleicher, A., Dabringhaus, A., et al. (1997). Motor cortex and hand motor skills: Structural compliance in the human brain. *Hum Brain Mapp, 5*(4), 206-215.

Anders, S., Lotze, M., Erb, M., Grodd, W., & Birbaumer, N. (2004). Brain activity underlying emotional valence and arousal: a response-related fMRI study. *Hum Brain Mapp, 23*(4), 200-209.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556-559.

Bangert, M., Haeusler, U., & Altenmuller, E. (2001). On practice: how the brain connects piano keys and piano sounds. *Ann N Y Acad Sci, 930*, 425-428.

Bastian, A. J. (2006). Learning to predict the future: the cerebellum adapts feedforward movement control. *Curr Opin Neurobiol, 16*(6), 645-649.

Beauregard, M., Levesque, J., & Bourgouin, P. (2001). Neural correlates of conscious self-regulation of emotion. *The Journal of neuroscience, 21*(18), 165.

Bengtsson, S. L., Nagy, Z., Skare, S., Forsman, L., Forssberg, H., & Ullen, F. (2005). Extensive piano practicing has regionally specific effects on white matter development. *Nat Neurosci, 8*(9), 1148-1150.

Binkofski, F., Amunts, K., Stephan, K. M., Posse, S., Schormann, T., Freund, H. J., et al. (2000). Broca's region subserves imagery of motion: a combined cytoarchitectonic and fMRI study. *Hum Brain Mapp, 11*(4), 273-285.

Bloedel, J. R., Bracha, V., & Larson, P. S. (1993). Real time operations of the cerebellar cortex. *Can J Neurol Sci, 20 Suppl 3*, 7-18.

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc Natl Acad Sci U S A, 98*(20), 11818-11823.

Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. (1999). Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat Neurosci, 2*(4), 382-387.

Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat Neurosci, 8*(3), 389-395.

Braitenberg, V., Heck, D., & Sultan, F. (1997). The detection and generation of sequences as a key to cerebellar function: experiments and theory. *Behav Brain Sci, 20*(2), 229-245.

Brett, M., Johnsrude, I. S., & Owen, A. M. (2002). The problem of functional localization in the human brain. *Nat Rev Neurosci, 3*(3), 243-249.

Brooks, D. J. (1995). The role of the basal ganglia in motor control: contributions from PET. *J Neurol Sci, 128*(1), 1-13.

Brown, S. (2000). The "musilanguage" model of music evolution. In N. L. Wallin, Merker, B., Brown, S. (Ed.), *The Origins of Music* (pp. 271-300). Cambridge, MA: MIT Press.

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

Brown, S., Ngan, E., & Liotti, M. (2007). A Larynx Area in the Human Motor Cortex. *Cereb Cortex, 18*(4), 837-845.

Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H. J., & Jancke, L. (2005). Scanning silence: mental imagery of complex sounds. *Neuroimage, 26*(4), 1119-1127.

Buonomano, D. V., & Merzenich, M. M. (1998). Cortical plasticity: from synapses to maps. *Annu Rev Neurosci, 21*, 149-186.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn Sci, 4*(6), 215-222.

Caldara, R., Deiber, M. P., Andrey, C., Michel, C. M., Thut, G., & Hauert, C. A. (2004). Actual and mental motor preparation and execution: a spatiotemporal ERP study. *Exp Brain Res, 159*(3), 389-399.

Callan, D. E., Tsytsarev, V., Hanakawa, T., Callan, A. M., Katsuhara, M., Fukuyama, H., et al. (2006). Song and speech: brain regions involved with perception and covert production. *Neuroimage, 31*(3), 1327-1342.

Crinion, J. T., & Leff, A. P. (2007). Recovery and treatment of aphasia after stroke: functional imaging studies. *Curr Opin Neurol, 20*(6), 667-673.

D'Ausilio, A., Altenmuller, E., Olivetti Belardinelli, M., & Lotze, M. (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *Eur J Neurosci, 24*(3), 955-958.

Damadian, R., Zaner, K., Hor, D., & DiMaio, T. (1974). Human tumors detected by nuclear magnetic resonance. *Proc Natl Acad Sci U S A, 71*(4), 1471-1473.

Dechent, P., Merboldt, K. D., & Frahm, J. (2004). Is the human primary motor cortex involved in motor imagery? *Brain Res Cogn Brain Res, 19*(2), 138-144.

Dejonckere, P., & Lebacq, J. (1981). Mechanism of initiation of oscillatory motion in human glottis. *Arch Int Physiol Biochim, 89*(2), 127-136.

Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., & Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science, 324*(5928), 811-813.

Doyon, J., & Benali, H. (2005). Reorganization and plasticity in the adult brain during learning of motor skills. *Curr Opin Neurobiol, 15*(2), 161-167.

Driskell, J. E., Copper, C., & Moran, A. (1994). Does mental practice enhance performance? *J Appl Psychol, 79*(4), 481-492.

Ducreux, D., Marsot-Dupuch, K., Lasjaunias, P., Oppenheim, C., & Fredy, D. (2003). [Lyrical and musical auditive mental imagery in functional MRI]. *J Neuroradiol, 30*(1), 18-24.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage, 25*(4), 1325-1335.

Ekholm, E., Papagiannis, G. C., & Chagnon, C. P. (1998). Relating objective measurements to expert evaluation of voice quality in western classical singing. *J Voice, 12*(2), 182-196.

English, A. W., & Weeks, O. I. (1984). Compartmentalization of single muscle units in cat lateral gastrocnemius. *Exp Brain Res, 56*(2), 361-368.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychol rev, 100*(3), 363-406.

Evers, S., Dannert, J., Rodding, D., Rotter, G., & Ringelstein, E. B. (1999). The cerebral haemodynamics of music perception. A transcranial Doppler sonography study. *Brain, 122*(1), 75-85.

Farrer, C., Frey, S. H., Van Horn, J. D., Tunik, E., Turk, D., Inati, S., et al. (2007). The Angular Gyrus Computes Action Awareness Representations. *Cereb Cortex, 18*(2), 254-261.

Fogassi, L., & Luppino, G. (2005). Motor functions of the parietal lobe. *Curr Opin Neurobiol, 15*(6), 626-631.

Fourkas, A. D., Bonavolonta, V., Avenanti, A., & Aglioti, S. M. (2008). Kinesthetic Imagery and Tool-Specific Modulation of Corticospinal Representations in Expert Tennis Players. *Cereb Cortex, 18*(10), 2382-2390.

Friston, K. J. (2002). Bayesian estimation of dynamical systems: an application to fMRI. *Neuroimage, 16*(2), 513-530.

Friston, K. J. (2007). *Statistical parametric mapping : the analysis of functional brain images* (1st ed.). Amsterdam: Elsevier/Academic Press.

Friston, K. J., Penny, W. D., & Glaser, D. E. (2005). Conjunction revisited. *Neuroimage, 25*(3), 661-667.

Gaab, N., Gaser, C., & Schlaug, G. (2006). Improvement-related functional plasticity following pitch memory training. *Neuroimage, 31*(1), 255-263.

Gaab, N., Gaser, C., Zaehle, T., Jancke, L., & Schlaug, G. (2003). Functional anatomy of pitch memory--an fMRI study with sparse temporal sampling. *Neuroimage, 19*(4), 1417-1426.

Gagné, F. (1985). Giftedness and talent: Reeexamining a reeaxamination of the definitions. *Gifted Child Quarterly, 29*, 103-112.

Gagné, F. (2000). Understanding the complex choreography of talend development through DMGT-based analysis. In K. A. Heller, R. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (2nd ed., pp. 67-79). New York: Elsevier.

Gao, J. H., Parsons, L. M., Bower, J. M., Xiong, J., Li, J., & Fox, P. T. (1996). Cerebellum implicated in sensory acquisition and discrimination rather than motor control. *Science, 272*(5261), 545-547.

Genovese, C. R., Lazar, N. A., & Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage, 15*(4), 870-878.

Gerardin, E., Sirigu, A., Lehericy, S., Poline, J. B., Gaymard, B., Marsault, C., et al. (2000). Partially overlapping neural networks for real and imagined hand movements. *Cereb Cortex, 10*(11), 1093-1104.

Gerloff, C., Corwell, B., Chen, R., Hallett, M., & Cohen, L. G. (1998). The role of the human motor cortex in the control of complex and simple finger movement sequences. *Brain, 121*(9), 1695-1709.

Ghazanfar, A. A., & Rendall, D. (2008). Evolution of human vocal production. *Curr Biol, 18*(11), 457-460.

Gieseking, W., & Leimer, K. (1972). *Piano Technique*: Dover Publications Inc.

Gracco, V. L., Tremblay, P., & Pike, B. (2005). Imaging speech production using fMRI. *Neuroimage, 26*(1), 294-301.

Grodd, W., Hulsmann, E., Lotze, M., Wildgruber, D., & Erb, M. (2001). Sensorimotor mapping of the human cerebellum: fMRI evidence of somatotopic organization. *Hum Brain Mapp, 13*(2), 55-73.

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

Gunji, A., Ishii, R., Chau, W., Kakigi, R., & Pantev, C. (2007). Rhythmic brain activities related to singing in humans. *Neuroimage, 34*(1), 426-434.

Hall, C. (2001). Imagery in sport and exercise. In R. N. Singer, H. A. Hausenblas & C. M. Janelle (Eds.), *Handbook of Sport Psychology* (2 ed., pp. 529-549). New York: John Wiley & Sons.

Hall, D. E. (1997). *Musikalische Akustik. Ein Handbuch*. Mainz: Schott.

Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cereb Cortex, 9*(7), 697-704.

Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia, 42*(9), 1281-1292.

Han, Y., Wang, J., Fischman, D. A., Biller, H. F., & Sanders, I. (1999). Slow tonic muscle fibers in the thyroarytenoid muscles of human vocal folds; a possible specialization for speech. *Anat Rec, 256*(2), 146-157.

Haslinger, B., Erhard, P., Altenmuller, E., Hennenlotter, A., Schwaiger, M., Grafin von Einsiedel, H., et al. (2004). Reduced recruitment of motor association areas during bimanual coordination in concert pianists. *Hum Brain Mapp, 22*(3), 206-215.

Hassler, M. (2000). Music medicine. A neurobiological approach. *Neuro Endocrinol Lett, 21*(2), 101-106.

Hebb, D. O. (1949). *The organization of behavior : a neuropsychological theory*. New York: Wiley.

Hickok, G., Buchsbaum, B., Humphries, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *J Cogn Neurosci, 15*(5), 673-682.

Highben, Z., & Palmer, C. (2004). Effects of auditory and motor mental practice in memorized piano performance. *Bulletin of the Council for Research in Music Education, 159*, 58-65.

Hirano, M. (1974). Morphological structure of the vocal cord as a vibrator and its variations. *Folia Phoniatr (Basel), 26*(2), 89-94.

Hoh, J. F. (2005). Laryngeal muscle fibre types. *Acta Physiol Scand, 183*(2), 133-149.

Hornak, J., Bramham, J., Rolls, E. T., Morris, R. G., O'Doherty, J., Bullock, P. R., et al. (2003). Changes in emotion after circumscribed surgical lesions of the orbitofrontal and cingulate cortices. *Brain, 126*(7), 1691-1712.

Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: an MEG study. *J Cogn Neurosci, 14*(8), 1125-1138.

Howe, M. J., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: reality or myth? *Behav Brain Sci, 21*(3), 399-407; discussion 407-342.

Huang, J., Carr, T. H., & Cao, Y. (2002). Comparing cortical activations for silent and overt speech using event-related fMRI. *Hum Brain Mapp, 15*(1), 39-53.

Humphreys, J. T. (1986). Measurement, prediction, and training of harmonic audiation and performance skills. *Journal of Research in Music Education, 34*, 192-199.

Hund-Georgiadis, M., & von Cramon, D. Y. (1999). Motor-learning-related changes in piano players and non-musicians revealed by functional magnetic-resonance signals. *Exp Brain Res, 125*(4), 417-425.

Hunter, M. D., Eickhoff, S. B., Miller, T. W., Farrow, T. F., Wilkinson, I. D., & Woodruff, P. W. (2006). Neural activity in speech-sensitive auditory cortex during silence. *Proc Natl Acad Sci U S A, 103*(1), 189-194.

Husler, F., & Rodd-Marling, Y. (1976). *Singing. The physical nature of the vocal organ: a guide to the unlocking of the singing voice* (Rev. ed.). London: Hutchinson.

Hutchinson, S., Lee, L. H., Gaab, N., & Schlaug, G. (2003). Cerebellar volume of musicians. *Cereb Cortex, 13*(9), 943-949.

Jackson, P. L., Lafleur, M. F., Malouin, F., Richards, C., & Doyon, J. (2001). Potential role of mental practice using motor imagery in neurologic rehabilitation. *Arch Phys Med Rehabil, 82*(8), 1133-1141.

Jahn, K., Deutschlander, A., Stephan, T., Strupp, M., Wiesmann, M., & Brandt, T. (2004). Brain activation patterns during imagined stance and locomotion in functional magnetic resonance imaging. *Neuroimage, 22*(4), 1722-1731.

Jancke, L., Kleinschmidt, A., Mirzazade, S., Shah, N. J., & Freund, H. J. (2001). The role of the inferior parietal cortex in linking the tactile perception and manual construction of object shapes. *Cereb Cortex, 11*(2), 114-121.

Jancke, L., Shah, N. J., & Peters, M. (2000). Cortical activations in primary and secondary motor areas for complex bimanual movements in professional pianists. *Brain Res Cogn Brain Res, 10*(1-2), 177-183.

Johnson, J. A., Strafella, A. P., & Zatorre, R. J. (2007). The role of the dorsolateral prefrontal cortex in bimodal divided attention: two transcranial magnetic stimulation studies. *J Cogn Neurosci, 19*(6), 907-920.

Jorgensen, H. (2001). Instrumental learning: Is an early start a key to success? *British Journal of Music Education, 18*(3), 227-239.

Jørgensen, H. (2002). Instrumental Performance Expertise and Amount of Practice among Instrumental Students in a Conservatoire. *Music Education Research, 4*(1), 105-119.

Jueptner, M., & Weiller, C. (1998). A review of differences between basal ganglia and cerebellar control of movements as revealed by functional imaging studies. *Brain, 121*, 1437-1449.

Karni, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature, 377*(6545), 155-158.

Kayelemetrics (2009). Retrieved 01.05.2009, from http://www.kayelemetrics.com/ This is a company that provides state-of-the-art technology in laryngoscopy

Kent, R. D. (2004). The uniqueness of speech among motor systems. *Clin Linguist Phon, 18*(6-8), 495-505.

Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., & Lotze, M. (2007). Overt and imagined singing of an Italian aria. *Neuroimage, 36*(3), 889-900.

Koelsch, S. (2005). Investigating emotion with music: neuroscientific approaches. *Ann N Y Acad Sci, 1060*, 412-418.

Koelsch, S., Fritz, T., DY, V. C., Muller, K., & Friederici, A. D. (2005). Investigating emotion with music: An fMRI study. *Hum Brain Mapp, 27*(3), 239-250.

Koelsch, S., & Siebel, W. A. (2005). Towards a neural basis of music perception. *Trends Cogn Sci, 9*(12), 578-584.

Koeneke, S., Lutz, K., Wustenberg, T., & Jancke, L. (2004). Long-term training affects cerebellar processing in skilled keyboard players. *Neuroreport, 15*(8), 1279-1282.

Kosslyn, S. M., & Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychol bull, 129*(5), 723-746.

Koufman, J. A., Sataloff, R. T., Woo, P., Zeitels, S. M., Ford, N. C., & Rosen, C. (2009). Retrieved 25.04.2009, from www.voiceproblem.org This website is edited by six internationally recognized voice surgeons.

Kraemer, D. J., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery: sound of silence activates auditory cortex. *Nature, 434*(7030), 158.

Krampe, R. T. (1997). Age-related changes in practice activities and their relation to musical performance skills. In H. Jørgensen & A. C. Lehman (Eds.), *Does practice make perfect? Current theory and research on instrumental music practice* (pp. 165–178). Oslo: Norwegian Academy of Music.

Kreitzer, A. C., & Malenka, R. C. (2008). Striatal plasticity and basal ganglia circuit function. *Neuron, 60*(4), 543-554.

Krings, T., Topper, R., Foltys, H., Erberich, S., Sparing, R., Willmes, K., et al. (2000). Cortical activation patterns during complex motor tasks in piano players and control subjects. A functional magnetic resonance imaging study. *Neurosci Lett, 278*(3), 189-193.

Krumhansl, C. L. (2003). Experimental strategies for understanding the role of experience in music cognition. *Ann N Y Acad Sci, 999*, 414-428.

Kuhtz-Buschbeck, J. P., Mahnkopf, C., Holzknecht, C., Siebner, H., Ulmer, S., & Jansen, O. (2003). Effector-independent representations of simple and complex imagined finger movements: a combined fMRI and TMS study. *Eur J Neurosci, 18*(12), 3375-3387.

Lacourse, M. G., Orr, E. L., Cramer, S. C., & Cohen, M. J. (2005). Brain activation during execution and motor imagery of novel and skilled sequential hand movements. *Neuroimage, 27*(3), 505-519.

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J Neurosci, 27*(2), 308-314.

Lane, R. D. (2000). Neural correlates of conscious emotional experience. In N. L. Lane RD (Ed.), *Cognitive neuroscience of emotion* (pp. 345-370). New York: Oxford UP.

Lang, P. J. (1979). Presidential address, 1978. A bio-informational theory of emotional imagery. *Psychophysiology, 16*(6), 495-512.

Langheim, F. J., Callicott, J. H., Mattay, V. S., Duyn, J. H., & Weinberger, D. R. (2002). Cortical systems associated with covert music rehearsal. *Neuroimage, 16*(4), 901-908.

Lauterbur, P. C. (1973). Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. *Nature, 242*, 190-191.

Lederman, R. J. (2003). Neuromuscular and musculoskeletal problems in instrumental musicians. *Muscle Nerve, 27*(5), 549-561.

Lenz, F. A., & Dougherty, P. M. (1998). Neurons in the human thalamic somatosensory nucleus (Ventralis caudalis) respond to innocuous cool and mechanical stimuli. *J Neurophysiol, 79*(4), 2227-2230.

Liu, Y., Gao, J. H., Liotti, M., Pu, Y., & Fox, P. T. (1999). Temporal dissociation of parallel processing in the human subcortical outputs. *Nature, 400*(6742), 364-367.

Logothetis, N. K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. *J Neurosci, 23*(10), 3963-3971.

Logothetis, N. K., & Pfeuffer, J. (2004). On the nature of the BOLD fMRI contrast mechanism. *Magn Reson Imaging, 22*(10), 1517-1531.

Lotze, M., Erb, M., Flor, H., Huelsmann, E., Godde, B., & Grodd, W. (2000). fMRI evaluation of somatotopic representation in human primary motor cortex. *Neuroimage, 11*(5 Pt 1), 473-481.

Lotze, M., & Halsband, U. (2006). Motor imagery. *J Physiol Paris, 99*(4-6), 386-395.

Lotze, M., Heymans, U., Birbaumer, N., Veit, R., Erb, M., Flor, H., et al. (2006). Differential cerebral activation during observation of expressive gestures and motor acts. *Neuropsychologia, 44*(10), 1787-1795.

Lotze, M., Laubis-Herrmann, U., Topka, H., Erb, M., & Grodd, W. (1999). Reorganization in the primary motor cortex after spinal cord injury - A functional Magnetic Resonance (fMRI) study. *Restor Neurol Neurosci, 14*(2-3), 183-187.

Lotze, M., Montoya, P., Erb, M., Hulsmann, E., Flor, H., Klose, U., et al. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *J Cogn Neurosci, 11*(5), 491-501.

Lotze, M., Scheler, G., Tan, H. R., Braun, C., & Birbaumer, N. (2003). The musician's brain: functional imaging of amateurs and professionals during performance and imagery. *Neuroimage, 20*(3), 1817-1829.

Lotze, M., Seggewies, G., Erb, M., Grodd, W., & Birbaumer, N. (2000). The representation of articulation in the primary sensorimotor cortex. *Neuroreport, 11*(13), 2985-2989.

Lu, X., & Ashe, J. (2005). Anticipatory activity in primary motor cortex codes memorized movement sequences. *Neuron, 45*(6), 967-973.

Ludlow, C. L. (2004). Recent advances in laryngeal sensorimotor control for voice, speech and swallowing. *Curr Opin Otolaryngol Head Neck Surg, 12*(3), 160-165.

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

Mansfield, P., & Grannell, P. K. (1973). NMR `Diffraction' in solids? *Journal of Physics C: Solid State Physics, 6*, 422-426.

Marrelec, G., Benali, H., Ciuciu, P., Pelegrini-Issac, M., & Poline, J. B. (2003). Robust Bayesian estimation of the hemodynamic response function in event-related BOLD fMRI using basic physiological information. *Hum Brain Mapp, 19*(1), 1-17.

Mathiak, K., Hertrich, I., Grodd, W., & Ackermann, H. (2002). Cerebellum and speech perception: a functional magnetic resonance imaging study. *J Cogn Neurosci, 14*(6), 902-912.

Matthews, P. M., & Jezzard, P. (2004). Functional magnetic resonance imaging. *J Neurol Neurosurg Psychiatry, 75*(1), 6-12.

McPherson, G. E., & Williamon, A. (2006). Giftedness and talent. In G. E. McPherson (Ed.), *The Child as Musician: A Handbook of Musicial Development* (pp. 239-256). Oxford: Oxford University Press.

Menon, V., & Levitin, D. J. (2005). The rewards of music listening: response and physiological connectivity of the mesolimbic system. *Neuroimage, 28*(1), 175-184.

Merati, A. L., Heman-Ackah, Y. D., Abaza, M., Altman, K. W., Sulica, L., & Belamowicz, S. (2005). Common movement disorders affecting the larynx: a report from the neurolaryngology committee of the AAO-HNS. *Otolaryngol Head Neck Surg, 133*(5), 654-665.

Merkel, F. (1902). *Handbuch der Anatomie des Menschen Atmungsorgane*. Jena: Fischer.

Molino, J. (2000). Toward an evolutonary theory of music and language. In N. L. Wallin, Merker, B., Brown, S. (Ed.), *The Origins of Music*. Cambridge, MA: MIT Press.

Monfared, A., Kim, D., Jaikumar, S., Gorti, G., & Kam, A. (2001). Microsurgical anatomy of the superior and recurrent laryngeal nerves. *Neurosurgery, 49*(4), 925-932; discussion 932-923.

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., & Zilles, K. (2001). Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage, 13*(4), 684-701.

Morris, J. S., Ohman, A., & Dolan, R. J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. *Proc Natl Acad Sci U S A, 96*(4), 1680-1685.

Morris, J. S., Scott, S. K., & Dolan, R. J. (1999). Saying it with feeling: neural responses to emotional vocalizations. *Neuropsychologia, 37*(10), 1155-1163.

Muley, S. A., & Bushara, K. O. (2004). Isolated gait ataxia due to cerebellar vermis infarct. *Arch Neurol, 61*(9), 1461.

Muller-Preuss, P., & Ploog, D. (1981). Inhibition of auditory cortical neurons during phonation. *Brain Res, 215*(1-2), 61-76.

Munte, T. F., Altenmuller, E., & Jancke, L. (2002). The musician's brain as a model of neuroplasticity. *Nat Rev Neurosci, 3*(6), 473-478.

Munzert, J., Lorey, B., & Zentgraf, K. (2009). Cognitive motor processes: The role of motor imagery in the study of motor representations. *Brain Res Rev*.

Mürbe, D., Pabst, F., Hofmann, G., & Sundberg, J. (2004). Effects of a professional solo singer education on auditory and kinesthetic feedback--a longitudinal study of singers' pitch control. *J Voice, 18*(2), 236-241.

Murphy, F. C., Nimmo-Smith, I., & Lawrence, A. D. (2003). Functional neuroanatomy of emotions: a meta-analysis. *Cognitive, affective & behavioral neuroscience, 3*(3), 207-233.

Nasir, S. M., & Ostry, D. J. (2006). Somatosensory precision in speech production. *Curr Biol, 16*(19), 1918-1923.

Natke, U., Donath, T. M., & Kalveram, K. T. (2003). Control of voice fundamental frequency in speaking versus singing. *J Acoust Soc Am, 113*(3), 1587-1593.

Nielsen, J. B., & Cohen, L. G. (2008). The Olympic brain. Does corticospinal plasticity play a role in acquisition of skills required for high-performance sports? *J Physiol, 586*(1), 65-70.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci U S A, 87*(24), 9868-9872.

Ogawa, S., Menon, R. S., Tank, D. W., Kim, S. G., Merkle, H., Ellermann, J. M., et al. (1993). Functional brain mapping by blood oxygenation level-dependent contrast magnetic resonance imaging. A comparison of signal characteristics with a biophysical model. *Biophys J, 64*(3), 803-812.

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nat Neurosci, 7*(1), 75-79.

Overy, K., Norton, A. C., Cronin, K. T., Gaab, N., Alsop, D. C., Winner, E., et al. (2004). Imaging melody and rhythm processing in young children. *Neuroreport, 15*(11), 1723-1726.

Özdemir, E., Norton, A., & Schlaug, G. (2006). Shared and distinct neural correlates of singing and speaking. *NeuroImage, 33*(2), 628-635.

Page, S. J., Levine, P., & Leonard, A. (2007). Mental practice in chronic stroke: results of a randomized, placebo-controlled trial. *Stroke, 38*(4), 1293-1297.

Pantev, C., Oostenveld, R., Engelien, A., Ross, B., Roberts, L. E., & Hoke, M. (1998). Increased auditory cortical representation in musicians. *Nature, 392*(6678), 811-814.

Pantev, C., Ross, B., Fujioka, T., Trainor, L. J., Schulte, M., & Schulz, M. (2003). Music and learning-induced cortical plasticity. *Ann N Y Acad Sci, 999*, 438-450.

Pascual-Leone, A. (2003). The brain that makes music and is changed by it. In I. Peretz & R. J. Zatorre (Eds.), *The Cognitive Neuroscience of Music* (pp. 396-412). Oxford: Oxford University Press.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nat Neurosci, 6*(7), 674-681.

Patel, A. D. (2008). *Music, language, and the brain*. Oxford [u.a.]: Oxford University Press.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. (1993). The neural correlates of the verbal component of working memory. *Nature, 362*(6418), 342-345.

Perry, D. W., Zatorre, R. J., Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1999). Localization of cerebral activity during simple singing. *Neuroreport, 10*(18), 3979-3984.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage, 16*(2), 331-348.

Phan, K. L., Wager, T. D., Taylor, S. F., & Liberzon, I. (2004). Functional neuroimaging studies of human emotions. *CNS spectrums, 9*(4), 258-266.

Pochon, J. B., Levy, R., Poline, J. B., Crozier, S., Lehericy, S., Pillon, B., et al. (2001). The role of dorsolateral prefrontal cortex in the preparation of forthcoming actions: an fMRI study. *Cereb Cortex, 11*(3), 260-266.

Porro, C. A., Francescato, M. P., Cettolo, V., Diamond, M. E., Baraldi, P., Zuiani, C., et al. (1996). Primary motor and sensory cortex activation during motor performance and motor imagery: a functional magnetic resonance imaging study. *J Neurosci, 16*(23), 7688-7698.

Price, C. J. (2000). The anatomy of language: contributions from functional neuroimaging. *J Anat, 197*(3), 335-359.

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

Riecker, A., Ackermann, H., Wildgruber, D., Meyer, J., Dogil, G., Haider, H., et al. (2000). Articulatory/phonetic sequencing at the level of the anterior perisylvian cortex: a functional magnetic resonance imaging (fMRI) study. *Brain Lang, 75*(2), 259-276.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends Neurosci, 21*(5), 188-194.

Rosenkranz, K., Williamon, A., & Rothwell, J. C. (2007). Motorcortical excitability and synaptic plasticity is enhanced in professional musicians. *J Neurosci, 27*(19), 5200-5206.

Sabbah, P., Simond, G., Levrier, O., Habib, M., Trabaud, V., Murayama, N., et al. (1995). Functional magnetic resonance imaging at 1.5 T during sensorimotor and cognitive task. *Eur Neurol, 35*(3), 131-136.

Sanders, I., & Mu, L. (1998). Anatomy of the human internal superior laryngeal nerve. *Anat Rec, 252*(4), 646-656.

Sanders, I., Rai, S., Han, Y., & Biller, H. F. (1998). Human vocalis contains distinct superior and inferior subcompartments: possible candidates for the two masses of vocal fold vibration. *Ann Otol Rhinol Laryngol, 107*(10), 826-833.

Sanders, I., Wu, B. L., Mu, L., Li, Y., & Biller, H. F. (1993). The innervation of the human larynx. *Arch otolaryngol, 119*(9), 934-939.

Scarr, S., & McCartney, K. (1983). How people make their own environments: a theory of genotype greater than environment effects. *Child Dev, 54*(2), 424-435.

Schlaug, G., Jancke, L., Huang, Y., Staiger, J. F., & Steinmetz, H. (1995). Increased corpus callosum size in musicians. *Neuropsychologia, 33*(8), 1047-1055.

Schneider, P., Scherg, M., Dosch, H. G., Specht, H. J., Gutschalk, A., & Rupp, A. (2002). Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat Neurosci, 5*(7), 688-694.

Schneider, P., Sluming, V., Roberts, N., Scherg, M., Goebel, R., Specht, H. J., et al. (2005). Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nat Neurosci, 8*(9), 1241-1247.

Schonberg, H. C. (1987). *Great Pianists*. New York: Simon & Schuster.

Schulz, M., Ross, B., & Pantev, C. (2003). Evidence for training-induced crossmodal reorganization of cortical functions in trumpet players. *Neuroreport, 14*(1), 157-161.

Sciote, J. J., Horton, M. J., Rowlerson, A. M., & Link, J. (2003). Specialized cranial muscles: how different are they from limb and abdominal muscles? *Cells Tissues Organs, 174*(1-2), 73-86.

Sharp, D. J., Scott, S. K., & Wise, R. J. (2004). Monitoring and the controlled processing of meaning: distinct prefrontal systems. *Cereb Cortex, 14*(1), 1-10.

Shergill, S. S., Bullmore, E., Simmons, A., Murray, R., & McGuire, P. (2000). Functional anatomy of auditory verbal imagery in schizophrenic patients with auditory hallucinations. *Am J Psychiatry, 157*(10), 1691-1693.

Shergill, S. S., Bullmore, E. T., Brammer, M. J., Williams, S. C., Murray, R. M., & McGuire, P. K. (2001). A functional study of auditory verbal imagery. *Psychol med, 31*(2), 241-253.

Shiotani, A., Westra, W. H., & Flint, P. W. (1999). Myosin heavy chain composition in human laryngeal muscles. *Laryngoscope, 109*(9), 1521-1524.

Simonton, D. K. (1999). Talent and its development: An emergenic and epigenetic model. *Psychol rev, 106*, 435-457.

Sirigu, A., Duhamel, J. R., Cohen, L., Pillon, B., Dubois, B., & Agid, Y. (1996). The mental representation of hand movements after parietal cortex damage. *Science, 273*(5281), 1564-1568.

Sloboda, J. A., Davidson, J. W., Howe, M. J. A., & Moore, D. G. (1996). The role of practice in the development of performing musicians. *B J Psych, 87*(2), 287-309.

Smith, A., & Zelaznik, H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Dev Psychobiol, 45*(1), 22-33.

Soltysik, D. A., & Hyde, J. S. (2006). Strategies for block-design fMRI experiments during task-related motion of structures of the oral cavity. *Neuroimage, 29*(4), 1260-1271.

Sosniak, L. A. (1985). Learning to be a concert pianist. In B. S. Bloom (Ed.), *Developing talent in young people* (1. ed. ed., pp. 19–67). New York: Balantine Books.

Stehling, M. K., Turner, R., & Mansfield, P. (1991). Echo-planar imaging: magnetic resonance imaging in a fraction of a second. *Science, 254*(5028), 43-50.

Stephan, K. M., Fink, G. R., Passingham, R. E., Silbersweig, D., Ceballos-Baumann, A. O., Frith, C. D., et al. (1995). Functional anatomy of the mental representation of upper extremity movements in healthy subjects. *J Neurophysiol, 73*(1), 373-386.

Sundberg, J. (1974). Articulatory interpretation of the "singing formant". *J Acoust Soc Am, 5*, 838-844.

Sundberg, J. (1977). The acoustics of the singing voice. *Sci Am, 236*(3), 82-84, 86, 88-91.

Sundberg, J. (1987). *The science of the singing voice*. DeKalb, Ill.: Northern Illinois University Press.

Sundberg, J. (1994). Perceptual aspects of singing. *J Voice, 8*(2), 106-122.

Sundberg, J. (2001). Level and center frequency of the singer's formant. *J Voice, 15*(2), 176-186.

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system; an approach to medical cerebral imaging*. New York: Thieme Medical Publishers.

Thach, W. T., Goodkin, H. P., & Keating, J. G. (1992). The cerebellum and the adaptive coordination of movement. *Annu Rev Neurosci, 15*, 403-442.

Thaut, M. H., Stephan, K. M., Wunderlich, G., Schicks, W., Tellmann, L., Herzog, H., et al. (2009). Distinct cortico-cerebellar activations in rhythmic auditory motor synchronization. *Cortex, 45*(1), 44-53.

Tillmann, B., Koelsch, S., Escoffier, N., Bigand, E., Lalitte, P., Friederici, A. D., et al. (2006). Cognitive priming in sung and instrumental music: activation of inferior frontal cortex. *NeuroImage, 31*(4), 1771-1782.

Titze, I. R. (1980). Comments on the myoelastic - aerodynamic theory of phonation. *J Speech Hear Res, 23*(3), 495-510.

Titze, I. R. (1988). The physics of small-amplitude oscillation of the vocal folds. *J Acoust Soc Am, 83*(4), 1536-1552.

Titze, I. R. (1994). *Principles of voice production*. Englewood Cliffs, N.J.: Prentice Hall.

Titze, I. R. (2008a). Nonlinear source-filter coupling in phonation: theory. *J Acoust Soc Am, 123*(5), 2733-2749.

Titze, I. R. (2008b). The human instrument. *Sci Am, 298*(1), 94-101.

Tramo, M. J. (2001). Biology and music. Music of the hemispheres. *Science, 291*(5501), 54-56.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage, 15*(1), 273-289.

Voisin, J., Bidet-Caulet, A., Bertrand, O., & Fonlupt, P. (2006). Listening in silence activates auditory areas: a functional magnetic resonance imaging study. *J Neurosci, 26*(1), 273-278.

Vuilleumier, P., Bogousslavsky, J., & Regli, F. (1995). Infarction of the lower brainstem. Clinical, aetiological and MRI-topographical correlations. *Brain, 118 (Pt 4)*, 1013-1025.

Wager, T. D., & Feldman-Barrett, L. (2004). *From affect to control: Functional specialization of the insula in motivation and regulation*: Published online at PsycExtra.

Watson, P. J., & Hixon, T. J. (1985). Respiratory kinematics in classical (opera) singers. *J Speech Hear Res, 28*(1), 104-122.

Weissman, D. H., Warner, L. M., & Woldorff, M. G. (2004). The neural mechanisms for minimizing cross-modal distraction. *J Neurosci, 24*(48), 10941-10949.

Wildgruber, D., Ackermann, H., Klose, U., Kardatzki, B., & Grodd, W. (1996). Functional lateralization of speech production at primary motor cortex: a fMRI study. *Neuroreport, 7*(15-17), 2791-2795.

Wildgruber, D., Riecker, A., Hertrich, I., Erb, M., Grodd, W., Ethofer, T., et al. (2005). Identification of emotional intonation evaluated by fMRI. *Neuroimage, 24*(4), 1233-1241.

Wilke, M., & Schmithorst, V. J. (2006). A combined bootstrap/histogram analysis approach for computing a lateralization index from neuroimaging data. *Neuroimage, 33*(2), 522-530.

Williamon, A., & Valentine, E. (2000). Quantity and quality of musical practice as predictors of performance quality. *Br J Psychol, 91*(3), 353-376.

Wolbers, T., Weiller, C., & Buchel, C. (2003). Contralateral coding of imagined body parts in the superior parietal lobe. *Cereb Cortex, 13*(4), 392-399.

Wu, Y. Z., Crumley, R. L., Armstrong, W. B., & Caiozzo, V. J. (2000). New perspectives about human laryngeal muscle: single-fiber analyses and interspecies comparisons. *Arch Otolaryngol Head Neck Surg, 126*(7), 857-864.

Wyke, B. D. (1974a). Laryngeal neuromuscular control systems in singing. A review of current concepts. *Folia Phoniatr (Basel), 26*(4), 295-306.

Wyke, B. D. (1974b). Proceedings: Laryngeal myotatic reflexes and phonation. *Folia Phoniatr (Basel), 26*(4), 249-264.

Yaguez, L., Nagel, D., Hoffman, H., Canavan, A. G., Wist, E., & Homberg, V. (1998). A mental route to motor learning: improving trajectorial kinematics through imagery training. *Behav Brain Res, 90*(1), 95-106.

Yoo, S. S., Lee, C. U., & Choi, B. G. (2001). Human brain mapping of auditory imagery: event-related functional MRI study. *Neuroreport, 12*(14), 3045-3049.

Yue, G., & Cole, K. J. (1992). Strength increases from the motor program: comparison of training with maximal voluntary and imagined muscle contractions. *J Neurophysiol, 67*(5), 1114-1123.

Zarate, J. M., & Zatorre, R. J. (2008). Experience-dependent neural substrates involved in vocal pitch regulation during singing. *Neuroimage, 40*(4), 1871-1887.

Zatorre, R. J. (2001). Neural specializations for tonal processing. *Ann N Y Acad Sci, 930*, 193-210.

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb Cortex, 11*(10), 946-953.

Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *J Neurosci, 14*(4), 1908-1919.

Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations for speech and pitch: moving beyond the dichotomies. *Philos Trans R Soc Lond B Biol Sci, 363*(1493), 1087-1104.

Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: musical imagery and auditory cortex. *Neuron, 47*(1), 9-12.

Zatorre, R. J., Meyer, E., Gjedde, A., & Evans, A. C. (1996). PET studies of phonetic processing of speech: review, replication, and reanalysis. *Cereb cortex, 6*(1), 21-30.

Zhang, J. X., Leung, H. C., & Johnson, M. K. (2003). Frontal activations associated with accessing and evaluating information in working memory: an fMRI study. *Neuroimage, 20*(3), 1531-1539.

Appendix 1

Supplementary Table 1: Main effect of overt singing (FDR corrected: $p < 0.05$ for the whole brain. Voxel extent threshold >5; cytoarchitectonic masks are indicated with a *. BA = Brodmann area).

Area	cluster size	t-value	x	y	z
left S1* (primary sensorimotor; BA 1,2,3)	130	8.39	-48	-15	36
right S1* (primary sensorimotor; BA 1,2,3)	127	6.32	36	-21	42
left M1* (primary motor cortex; BA 4a+p)	39	8.23	-51	-12	39
right M1* (primary motor cortex; BA 4a+p)	49	7.26	48	-12	39
SMA* (medial supplementary motor; BA 6)	26	7.20	-6	-9	69
left PMC* (premotor cortex; lateral BA 6)	66	7.05	-54	-9	42
right PMC* (premotor cortex; lateral BA 6)	71	6.53	51	-12	54
left A1* (primary auditory cortex; BA 41)	45	6.00	-48	-21	12
right A1* (primary auditory cortex; BA 41)	67	6.76	57	-9	3
left anterior insula	192	7.72	-39	6	6
left posterior insula		4.92	-42	-15	6
right anterior insula	130	4.76	48	6	0
right posterior insula		6.50	48	-9	6
left superior temporal gyrus (BA 22)	542	7.42	-69	-24	9
right superior temporal gyrus (BA 22)	676	7.83	54	-12	3
anterior cingulate gyrus (ACC, BA24)	109	6.33	3	18	21
medial cingulate gyrus (BA 24, 32)	125	4.60	-3	15	39
left SII* (secondary somatosensory cortex; OP 1,3,4)	99	7.35	-57	-9	18
right SII* (secondary somatosensory cortex; OP 1,3,4)	100	6.78	60	-9	15
left supramarginal gyrus (BA 40)	80	6.62	-45	-45	47
right supramarginal gyrus (BA 40)	20	3.94	66	-27	18
right cerebellar hemisphere (Larsell lobule H VI)	228	5.48	9	-63	-21
left cerebellar hemisphere (Larsell lobule H VI)	297	5.15	-12	-63	-27

cerebellar vermis	155	5.44	6	-63	-18
left occipital lobe (BA 17, 18, 19)	390	7.81	-15	-81	9
right occipital lobe (BA 17, 18, 19)	277	5.22	21	-75	3
left BA 44* (Broca´s area, BA 44)	11	4.50	-57	3	15
right BA 44* (Broca´s analogon, BA 44)	11	4.70	60	6	18
left pallidum	16	6.71	-21	-6	-6
right pallidum	21	5.19	21	-6	-3
right putamen	9	3.18	27	9	9
medulla	69	5.86	9	-36	-42
left thalamus	83	5.24	-21	-15	-3
right thalamus	104	5.53	18	-21	0

Supplementary Table 2: Main effect of imagined singing (FDR corrected: $p < 0.05$ for the whole brain. Voxel extent threshold >5; cytoarchitectural masks are indicated with a *. BA = Brodmann area).

Area	Cluster size	t-value	x	y	z
SMA* (medial supplementary motor; BA 6)	601	9.95	-3	-9	66
left PMC* (premotor cortex; lateral BA 6)	107	5.94	-51	-9	54
right PMC* (premotor cortex; lateral BA 6)	89	9.56	54	-6	54
left S1* (primary sensorimotor; BA 1,2,3)	119	5.30	-60	-21	48
right S1* (primary sensorimotor; BA 1,2,3)	157	5.86	57	-15	48
left M1* (primary motor cortex; BA 4a+p)	42	6.83	-48	-18	45
right M1* (primary motor cortex; BA 4a+p)	101	7.20	51	-15	48
left anterior insula (BA 24, 32)	233	6.37	-39	12	3
right anterior insula (BA 24, 32)	210	6,47	39	3	-3
left frontomedial anterior (BA 9)	160	6.93	-30	48	38
right frontomedial anterior (BA 9)	114	5.95	30	42	33
left temporal pole (BA 38)	70	9.07	-51	12	-6
right temporal pole (BA 38)	117	6.44	60	9	-9
left sup. temporal gyrus (BA 22)	210	7.67	-51	6	-6
right sup. temporal gyrus (BA 22)	442	9.73	48	39	15
left BA 44* (Broca´s area, BA 44)	58	6.77	-57	9	36
right BA 44* (Broca´s analogon, BA 44)	81	5.65	60	6	12
left BA 45 (Broca´s area, BA 45)	23	5.43	-54	15	33
right BA 45* (Broca´s analogon, BA 44)	81	4.38	60	15	9
left ventrolateral prefrontal (BA 47)	40	6.99	-42	15	-3
right ventrolateral prefrontal (BA 47)	27	5.32	54	18	-6
left parietal inferior (BA 39, 40)	412	6.16	-39	-57	45
right parietal inferior (BA 39, 40)	76	4.04	30	-54	48
left angular gyrus (BA 40)	68	6.15	-36	-60	42
right angular gyrus (BA 40)	59	3.83	27	-54	45
left SII* (secondary sensory cortex; OP 1,3,4)	85	5.04	-60	-15	21
right SII* (secondary sensory cortex; OP 1,3,4)	85	4.77	60	-15	18
left parietal superior (BA 5, 7)	47	4.84	-39	-66	51
right parietal superior (BA 5, 7)	43	4.20	27	-54	48

left occipital lobe (BA 17, 18, 19)	177	5.56	-24	-87	-6
right occipital lobe (BA 17, 18, 19)	307	6.59	36	-81	-12
dorsal medial prefrontal cortex (mPFC)	66	4.81	0	15	42
anterior cingulate cortex (BA 24, 32)	29	3.69	-3	21	30
medial cingulate cortex (BA 24, 32)	391	5.91	-3	9	42
posterior cingulate cortex (BA 23)	19	3.43	-3	-36	30
left thalamus	159	6.89	-15	-12	9
right thalamus	162	5.17	21	-18	0
left cerebellar hemisphere (HVII)	535	7.63	-30	-60	-27
right cerebellar hemisphere (HVII)	725	7.40	36	-57	-30
cerebellar vermis	63	3.25	3	-78	-30
left putamen	48	4.83	-24	-6	3
right putamen	105	4.90	21	0	9
medulla	22	2.53	9	-36	-42
left pallidum	49	5.55	-18	-3	6
right pallidum	20	4.11	18	0	-3
left caudate nucleus	108	4.78	-15	-6	15
right caudate nucleus	49	4.32	15	-9	15
left amygdala	32	4.62	-21	-6	-12
right amygdala	8	3.72	21	-6	-12
left anterior hippocampus	38	4.29	-18	-6	-12
right anterior hippocampus	11	3.25	18	-6	-12

Initialen: _______ Alter: ______ Datum: ____________ ♂ ♀

FMRI: Vorstellung

Wir möchten Sie bitten, auf den folgenden Skalen anzugeben, wie Sie sich während der Aufgabe gefühlt haben. Machen Sie hierzu bitte ein Kreuz an der Stelle auf der Skala, die Ihrem Gefühl entspricht:

1. Wie gut konnten Sie sich auf die Aufgabe konzentrieren:

gar nicht sehr gut

2. Wie gut konnten Sie sich den Gesang vorstellen:

gar nicht sehr gut

3. Wie gut konnten Sie sich Bewegungen vorstellen (z.B. Atmung, Stütze):

gar nicht sehr gut

4. Wie gut konnten Sie sich das Tempo vorstellen:

gar nicht sehr gut

5. Wie gut konnten Sie sich den Rhythmus vorstellen:

gar nicht sehr gut

6. Wie aufgeregt waren Sie während der Kerspinuntersuchung:

gar nicht sehr

Vielen Dank für Ihre Antworten!

Appendix 2

Supplementary Tables 3-5 give MNI coordinates and peak t-score values based on main effects of singing in opera singers, vocal students and opera singers respectively ($P < 0.001$ uncorrected, calculated for the whole brain). Supplementary Table 6 gives MNI coordinates and peak t-score values of a corresponding conjunction analysis ($P < 0.001$ uncorrected, calculated for the whole brain). For anatomic labeling we used AAL (Tzourio-Mazoyer et al., 2002). Areas already cytoarchitectonically mapped (*) were assigned using the Anatomy Toolbox (Eickhoff et al. 2005) - probabilities are given in brackets. Abbreviations: M1, primary motor cortex; S1, primary somatosensory cortex; SII, secondary somatosensory cortex; A1, primary auditory cortex; SMA, supplementary motor area; IPC, inferior parietal cortex; SPL, superior parietal lobe; 7A corresponds to a division of Brodmann's area 7 ; PFcm corresponds to the inferior part of the supramarginal gyrus; Hipp, hippocampus; Hipp (SUB), parahippocampal gyrus; L = left; R = right.

Supplementary Table 3: Main effect of overt singing in opera singers.

Region		*x*	*y*	*z*	*t*	Assigned to
Sensorimotor						
M1*	R	54	-9	33	6.97	*Area 4p (60%)*
	L	-51	-12	33	5.45	*Area 4p (70%)*
S1*	L	-60	-12	39	8.55	*Area 1 (50%)*
	L	-48	-18	33	7.42	*Area 3b (50%)*
	R	54	-12	33	6.55	*Area 3b (80%)*
	R	60	-12	39	6.16	*Area 1 (90%)*
SII*	R	63	-12	27	5.81	*OP 4 (40%)*
	L	-66	-21	15	3.59	*OP1 (60%)*
Premotor						
SMA*		-3	-9	60	4.68	*Area 6 (100%)*
		51	-9	48	4.43	*Area 6 (60%)*
Auditory						
A1*						
	R	60	-12	6	5.11	*TE 1.0 (40%)*
	R	57	-9	3	4.96	*TE 1.2 (50%)*
	L	-36	-33	12	5.01	*TE 1.1 (50%)*
	L	-51	-15	3	4.77	*TE 1.0 (60%)*
	L	-51	-9	3	4.37	*TE 1.2 (40%)*
Parietal						
SPL*	L	-12	-54	75	4.12	*SPL (7A)*
IPC*	R	48	-33	24	4.00	*IPC (PFcm)*
	R	-48	-42	24	3.54	*IPC (PFcm)*
Occipital						
Middle Occipital Gyrus	L	-30	-87	18	5.24	
	R	33	-72	21	3.99	
Cuneus*	L	-3	-84	24	4.98	*Area 18 (60%)*
	R	18	-78	33	4.41	
Calcarine Gyrus*	L	-24	-60	9	4.89	*Area 17 (40%)*
	R	18	-75	6	3.83	*Area 17 (100%)*
Superior Occipital Gyrus	R	21	-81	21	4.61	
	L	-18	-84	21	4.42	
Middle Cingulate Cortex	L	-9	9	36	4.67	
Insula	R	48	-9	3	4.09	

	L	-45	0	3	3.97	
Temporal pole	R	57	6	-3	4.51	
	L	-48	6	-18	4.21	
Subcortical						
Thalamus	L	-18	-9	6	5.20	
	R	21	-21	9	3.68	
Pallidum	L	-21	-6	3	3.66	
	R	21	0	6	4.42	
Putamen	R	24	0	6	3.73	
	L	-27	-21	6	3.29	
Hipp	L	-18	-21	-21	3.29	*SUB (90%)*
Cerebellum						
Lobule H VI	R	21	-66	-24	7.12	
	L	-27	-57	-27	6.14	
Lobule H VII	R	15	-63	-42	4.30	
Lobule H IV-V	R	30	-42	-21	3.86	
	L	-12	-36	-18	3.69	
Crus 1	R	42	-69	-30	4.59	
	L	-21	-84	-30	4.05	
Vermis (8)	R	6	-66	-39	5.19	
Vermis (10)	R	6	-48	-30	3.92	
	L	-3	-45	-33	3.91	

Supplementary Table 4: Main effects of overt singing in vocal students

Region		*x*	*y*	*z*	*t*	Assigned to
Sensorimotor						
M1*	R	48	-12	36	11.22	*Area 4p (70%)*
	L	-54	-6	27	10.44	*Area 4p (40%)*
S1*	L	-57	-12	36	14.17	*Area 1 (60%)*
	L	-57	-6	27	11.59	*Area 3b (50%)*
	R	57	-9	30	9.71	*Area 3b (60%)*
	R	63	-9	36	9.12	*Area 1 (70%)*
SII*	L	63	-12	27	8.92	*OP 4 (40%)*
	R	66	-9	15	6.93	*OP 4 (60%)*
*Premotor**						
	R	48	-15	60	4.59	*Area 6 (60%)*
	R	45	-18	63	4.40	*Area 6 (70%)*
	L	-57	3	15	5.52	*Area 44 (30%)*
	R	60	6	18	4.43	*Area 44 (40%)*
SMA	L	0	-9	63	5.79	*Area 6 (100%)*
Auditory						
A1*	R	54	-12	3	6.60	*TE 1.0 (70%)*
	R	42	-30	12	5.22	*TE 1.1 (30%)*
	R	60	3	0	4.55	*TE 1.2 (30%)*
	L	-66	-27	6	6.49	*TE 3 (90%)*
	L	-54	-9	0	5.99	*TE 1.2 (70%)*
	L	-48	-18	6	5.81	*TE 1.0 (80%)*
Parietal						
IPC*	L	-42	-33	18	6.05	*PFcm*
	R	66	-27	15	5.36	*PF*
Occipital						
Lingual Gyrus	R	24	-54	0	3.80	*Area 17 (80%)*
Superior Occipital Gyrus	L	-15	-84	9	5.17	*Area 17 (40%)*

Calcarine Gyrus	L	-21	-72	6	4.27	*Area 17 (20%)*
Cuneus	R	6	-84	24	3.71	*Area 18 (40%)*
Insula	L	-42	3	3	5.05	
	R	39	-12	15	3.89	
Middle Cingulate Cortex	R	3	15	30	4.82	
Anterior Cingulate Cortex	R	3	15	27	4.42	
	L	-6	12	30	3.72	
Temporal pole	R	60	3	0	4.55	
	L	-57	6	-3	4.48	
Subcortical						
Pallidum	L	-21	-9	-3	4.23	
	R	21	-3	6	3.82	
Putamen	L	-27	-15	12	3.41	
Hipp*	R	27	-36	3	3.99	*CA (80%)*
	L	-27	-39	3	3.78	*CA (90%)*
Thalamus	L	-18	-18	9	3.51	
	R	15	-12	12	3.98	
Cerebellum						
Lobule H VI	L	-15	-63	-24	7.85	
	R	21	-66	-24	6.55	
Lobule H VIII	R	15	-63	-42	4.80	

Supplementary Table 5: Main effects of overt singing in laymen

Region		*x*	*y*	*z*	*t*	Assigned to
Sensorimotor						
M1*	L	-54	-9	36	9.63	*Area 4p (60%)*
	L	-48	-18	54	4.67	*Area 1 (60%)*
	R	45	-12	39	7.76	*Area 4p (50%)*
	R	48	-18	57	4.11	*Area 1 (50%)*
S1*	L	-57	-12	39	8.01	*Area 1 (60%)*
	L	-54	-9	27	8.06	*Area 3b (60%)*
	R	57	-9	30	5.54	*Area 3b (60%)*
	R	57	-15	48	4.19	*Area 1 (90%)*
SII*	L	-51	-15	12	6.28	*OP 1 (50%)*
	R	63	-9	15	5.50	*OP 4 (50%)*
*Premotor**						
	L	-54	-9	48	6.42	*Area 6 (70%)*
	R	48	-15	60	5.17	*Area 6 (60%)*
	R	48	9	3	3.96	*Area 44 (40%)*
SMA	L	-9	0	48	5.24	*Area 6 (60%)*
SMA	R	6	-6	66	4.98	*Area 6 (100%)*
Auditory						
A1*	R	54	-9	0	6.42	*TE 1.2 (60%)*
	R	54	-15	3	6.17	*TE 1.0 (70%)*
	R	45	-24	6	5.51	*TE 1.1 (80%)*
	L	-51	-9	3	5.51	*TE 1.2 (40%)*
	L	-39	-30	9	5.30	*TE 1.1 (70%)*
Parietal						
IPC*	L	-48	-36	18	5.97	*PFcm (70%)*
	L	-57	-24	21	3.58	*Pfop (50%)*
	R	63	-30	18	4.08	*PF*
Occipital						

Precuneus	R	15	-42	9	3.72	
Insula	R	51	-3	0	5.26	
	L	-42	-15	6	4.69	
Temporal pole	L	-51	9	-9	4.80	
	R	51	6	-15	4.53	
Middle temporal gyrus	R	51	-48	3	4.10	
Posterior Cingulate Cortex	R	12	-39	12	3.44	
Subcortical						
Caudate	L	-15	6	18	4.72	
	R	18	21	9	4.22	
Hipp*	R	30	-33	-3	4.18	*Hipp CA (40%)*
	R	15	-36	9	3.96	*Hipp CA (60%)*
	L	-12	-36	9	3.36	*Hipp CA (40%)*
	L	-12	-33	-9	3.78	*Hipp SUB (80%)*
Thalamus	R	3	-24	6	4.28	
		-3	-24	9	3.58	
Cerebellum						
Lobule H VI	R	9	-63	-18	5.64	
	L	-15	-63	-18	4.96	
Lobule H IV-V	L	-3	-48	-18	3.62	
	R	12	-45	-9	4.07	
Vermis (6)	R	6	-63	-15	5.68	
Vermis (4/5)	R	0	-51	-3	3.80	
Vermis (8)	R	0	-66	-39	3.71	

Supplementary Table 6: Conjunction analysis of overt singing (all groups)

Region		*x*	*y*	*z*	*t*	Assigned to
Sensorimotor						
M1*	R	45	-12	36	6.31	*Area 4p (70%)*
S1*	L	-57	12	39	8.01	*Area 1 (60%)*
	R	57	-9	36	6.01	*Area 3b (40%)*
SII*	L	-60	-12	24	5.26	*OP 4 (40%)*
	R	63	-15	9	4.84	*OP 4 (20%)*
*Premotor**						
	R	51	-9	48	4.43	*Area 6 (50%)*
SMA	L	0	-6	60	4.03	*Area 6 (80%)*
Auditory						
A1*	R	60	-9	0	5.16	*TE 1.0 (40%)*
	R	57	-12	3	5.07	*TE 1.0 (40%)*
	R	45	-27	9	4.91	*TE 1.1 (600%)*
	L	-48	-15	3	5.16	*TE 1.0 (60%)*
	L	-66	-21	9	4.38	*TE 3 (80%)*
Parietal						
IPC*	L	-42	-36	15	4.89	*PFcm (50%)*
Cerebellum						
Lobule H VI	R	9	-66	-18	5.35	
	L	-15	-63	-21	4.93	

Fragen zu Erfahrungen mit dem Singen

Name:

1. Händigkeit (bitte ein x in das zutreffende Feld)

Mit welchem Ihrer Hände	schreiben,	essen (Löffel),	Besen führen,	malen
Rechte Hand				
Linke Hand				
Beide Hände werden gleich gebraucht				

2. Grundsätzliche Ausbildung und Beruf

Schulabschluss:

Ausbildung/Beruf: z.B. Medizinstudent:

3. Musikalische Tätigkeiten

A: Gesang

Singen Sie, allein oder in einem Chor? ____

Seit wie vielen Jahren singen Sie bereits? ____

Wie viele Stunden (**h**) singen Sie pro Woche? ____

Hatten Sie schon einmal Gesangsunterricht? Wenn ja, wie lange? __________________________

B. Instrument

Spielen Sie ein Instrument?____

Welche Instrumente spielen Sie?____

In welchem Alter haben Sie begonnen, ein Instrument zu erlernen?__________________________

Von wann bis wann haben Sie Instrumentalunterricht erhalten?____________________________

Haben Sie danach regelmäßig Ihr Instrument gespielt?___________________________________

Wie viele Stunden (**h**) spielen Sie pro Woche Ihr Instrument?_____________________________