

Please cite the chapter in your own work as: Cohen, A. J., Levitin, D., & Kleber, B. (2020). Brain mechanisms underlying singing. In F.A. Russo, B. Ilari, & A. J. Cohen (Eds.), *Routledge Companion to Interdisciplinary Studies in Singing: Vol I, Development* (pp. 79-86). Routledge.

Brain mechanisms underlying singing

Annabel J. Cohen, University of Prince Edward Island

Daniel Levitin, Minerva Schools at KGI and McGill University

Boris A. Kleber, Aarhus University and The Royal Academy of Music Aarhus/Aalborg

Introduction

Singing relies on activity of the brain. As discussed historically in Chapter 4 (Graziano, Born, & Johnson), Henschen (1920) had searched for a brain center for singing in his comprehensive investigations of aphasia. He had pointed to the left frontal cortex, near the previously identified center for speech in Broca's Area. Just as in Henschen's time, in the late 20th and early 21st century, neuroscientific studies of singing have been far fewer than those of language. However, due to some shared mechanisms underlying speaking and singing, progress in understanding brain function underlying language often enlightens the neuroscience of singing.

Moreover, in recent years, neuroscientists (e.g., Steven Brown, Boris Kleber, Isabelle Peretz, Séverine Samson, Jean Zarate, and Robert Zatorre) have directed specific attention to "singing in the brain". The present chapter surveys this literature, drawing from a wide range of sources, including previous reviews (Brown, Martinez, Hodges, Fox, & Parsons, 2004; Cohen, 2019; Kleber & Zarate, 2014; Loui, 2015). Taking on a challenging task of integration of information from these sources, this chapter also aims to provide foundational concepts for readers from disciplines outside of the cognitive sciences. Thus, the chapter begins with a brief introduction to

the brain. It then discusses the reliance of singing on feedback from the auditory and motor systems and their co-operation in the singing network. The chapter closes with a brief consideration of the neurochemical aspects of singing and a contemporary look at aphasia, coming full circle with the beginning of the chapter.

The human brain

The human brain is the most complex organ of the body. It is hierarchical, based on phylogenetics — newer structures build in layers onto older ones, outward from a core (Arbib, 2013, p. 24; Luu, Kelly & Levitin, 2001). In mammals, the two roughly symmetrical hemispheres of the brain are covered by the cerebral cortex. The cortex for each hemisphere is divided into four lobes, each with some degree of functional specialization of function within each hemisphere (see Figure 6.1 A1)

<Figure 6.1 here> [see pp. 40-41]

Four decades of neuroimaging work have lent support to the idea that specific brain functions are localized within these lobes, but much of that work was conducted analyzing only gray matter; when we take white matter tracts into consideration, the emerging picture is that most higher cognitive functions (such as reasoning, thinking, speaking, singing, memory) engage networks of activity spatially distributed across different locations (Fox & Friston, 2012; Menon, 2013; Ross, 2010), an idea proposed earlier by Pribram (1982).

The *temporal lobe*, located above the ear, is the most significant for hearing. It contains the part of the brain known as the auditory cortex. The left temporal lobe is associated with the ability to process coarse but fast changes in sound (e.g., transients

that distinguish phonemes), and the right is more specialized for representation of fine but slow frequency relations as those which characterize vowels or musical instrument sounds (Zatorre & Baum, 2012). The *frontal lobe*, located behind the forehead, is associated with executive functions, learning, memory and planning. It contains the motor cortex and is implicated in voluntary muscle motion, which would include altering the pitch or timing of a note. The distribution of muscles is systematically mapped onto a portion of the frontal lobe and is known as the motor homunculus. The *occipital lobe* at the back of the head is associated with vision, and is relevant to singing from the standpoint of processing signals of moving lips, facial expression, and bodily rhythmic motion of fellow choristers or a choral leader, or checking in a mirror one's own posture and motor behavior when practicing. The *parietal lobe*, on the top, separating the frontal and occipital lobes, and also the left from the right temporal lobes, is well-positioned to receive and integrate information from the different senses: auditory, visual, tactile, skin receptors, kinesthetic (motion perception), and proprioceptive (position of muscles and joints). It contains the sensory homunculus, a distorted map of the body representing sensory responsiveness, with disproportionate space dedicated to sensations from the vocal tract (tongue, lips, pharynx), as well as the nose, eye area, and fingers/hand (see Figure 6.1 A1).

Anatomical terms of location that describe the cortical architecture broadly distinguish between regions that are located at the top of the brain (dorsal) and those that are located below (ventral), and those that are forward (anterior) and rearward (posterior). Brodmann's (1909) system of mapping the brain into 52 areas, based on

histological tissue analysis, regained popularity in modern neuroscience with the introduction of novel imaging techniques in the 1980s that allowed scientists to non-invasively visualize the structure of the living brain as well as its neural activity during task performance, and thus to associate brain structural units with their underlying functions. For example, recently the part of the motor cortex responsible for controlling the vocal folds, a major function of the intrinsic musculature of the larynx, was revealed in a brain-imaging study of persons who carried out a variety of tasks such as singing the first five notes of the musical scale, making “glottal stops”, tongue movements and lip protrusions. The revealed area was named the *larynx/phonation area (LPA)* (Brown, Ngan, & Liotti, 2008). (Figure 6.1 A)

Hierarchical brain organization underlying vocalization

Following evidence presented by Simonyan and Horwitz (2011), Kleber and Zarate (2014) distinguish a hierarchically organized brain system with parallel pathways for controlling innate vocalization and voluntary fine-motor control, which includes the intentional control of emotional voice production. (A1). The vocal pattern generator is located in the brainstem’s reticular formation, hosting phonatory (larynx) motoneurons that receive input from both pathways. (A5) The periaqueductal gray (PAG) in the midbrain is closely connected to limbic structures involved in emotional vocalization and voice initiation, including the amygdala, hypothalamus, hippocampus, and the anterior cingulate cortex (ACC) (A4 and A5). Lesions of the ACC lead to a complete loss over volitional emotional prosody (Zarate, 2013, p. 2) and lesions of the PAG to mutism (Simonyan & Horwitz, 2011). The basal ganglia

(putamen and nucleus caudatus) play a role in learning new songs (A2). They contain a limbic portion (ventral striatum, ventral tegmental area), which are known as important components of the dopaminergic reward system (A3). The primary motor cortex (M1) represents the highest level in the hierarchy as a crucial area for acquiring control over vocalizations of learned song and speech, which is facilitated by unique direct connections in humans between the vocalization area in M1 and the brainstem motor neurons that command phonatory (laryngeal) muscles (see Belyk & Brown, 2017). Learned vocalizations, however, can be modulated by lower brain regions (i.e., putamen, globus pallidus, thalamus, pontine gray and cerebellum (Kleber & Zarate, pp. 2-3).

Auditory feedback from vocal productions, as well as somatosensory feedback from receptors of the vocal tract, larynx, and respiratory system engage other parts of the brain (Jürgens, 2009; Zarate, 2013). Somatosensory feedback is transmitted to primary and secondary somatosensory cortex, as well as to the insula, whereas auditory feedback is transmitted to the auditory cortex in the superior temporal gyrus (STG). Potential neural regions for audio-vocal integration for singing include the PAG, posterior superior temporal sulcus (pSTS), inferior parietal cortex, ACC, inferior frontal gyrus (IFG), and the anterior insula (aINS), as hypothesized by Kleber and Zarate (2014, p. 5).

A detailed discussion of brain mechanisms underlying music perception is beyond the scope of the present chapter, but the reader is referred to Levitin (2006), and Tan, Pfordresher, and Harré (2016) for introductions, and to Oxenham (2019) and

Koelsch (2019) for more detail. Regarding the perception of the quality of the voice see McAdams (2019). Belin, Zatorre, Lafaille, Ahad, & Pike (2000) showed that specific regions of the superior temporal sulcus (STS) responded to the human voice rather than to environmental sounds.

A feedback system

The system of neural activity that drives vocal production relies heavily on sensory feedback (e.g., Dalla Bella, Berkowska, & Sowinski, 2011; Pfordresher et al., 2015; Tsang, Friendly, & Trainor, 2011). Intentional vocalization as in singing and speaking entails neural activation of muscles that control respiration, vocal fold vibration, and vocal tract configuration. These anatomical areas, mechanics and acoustics have been described by Sundberg (1987) (see also Chapter 5 by Wolfe, Garnier, Henrich Bernardoni, & Smith). The process is complex because singing requires precision in pitch (frequency), timing and timbre (tonal quality) that exceeds requirements for speaking (Zatorre & Baum, 2012). It entails the coordination of three complex processes: (a) forming a representation or mental model of the song one wants to sing, (b) guiding the vocal production system (controlling position and tension of the vocal cords as well as breathing mechanisms) to create the sound of the mentally represented model, and (c) using auditory and motor (i.e., bodily) feedback to determine whether the targeted sounds are reached, and if not, making further adjustments.

The first process entails generating an unfolding mental model of the target melody supported by knowledge of the conventions or grammar of one's musical culture (Cohen, 2000; Lerdahl & Jackendoff, 1983; Trainor, 2018). The melody in mind may be one heard before, or it could be a melody spontaneously composed or improvised. Structural complexity and degree of adherence to cultural conventions can obviously influence precision of the models (Fine, Berry, & Rosner, 2006).

The second process entails translation of the mental model of the melody into a program of motor commands to create the melody. This is sometimes referred to as reverse engineering, as the brain predicts the sensory consequences of certain actions based on experience. As the third process, the intended production is then compared with the auditory and kinesthetic feedback arising from vocalization, which leads to a corrective motor response in case of mismatch and an updating of the associated model to generate better predictions in the future (Guenther, 2016; Hickok, 2017).

The complexity of the processes might explain why some adults feel they cannot sing (although the spirit of this chapter is that almost everyone is born with the ability to sing, which can be facilitated by practice and confidence). Even pre-school children can produce melodies that are recognizable by adults (Gudmundsdottir & Trehub, 2017). Indeed, many adults hold the subjective impression or fear that they can't sing for a variety of reasons that are not necessarily grounded. Some have been told that they don't have a pleasing vocal timbre; some find it difficult to remember songs (a problem with auditory sequence memory, not singing per se); and some are simply afraid to sing (Levitin, 1999).

The ability to hear what one has sung and compare it to what one expected to sing is a key factor in singing, as well as playing any musical instrument, particularly those like the violin family, in which one must produce each tone from a continuum of possibilities. The importance of auditory feedback has been revealed by several researchers. Ward and Burns (1978) denied auditory feedback to trained singers (forcing them to rely solely on muscle memory); the singers erred by as much as a minor third, or three semitones. Murry (1990) examined the first five acoustic waves of vocal production (before auditory feedback could take effect) and found that singers who were otherwise good at pitch matching made average errors of 2.5 semitones, and errors as large as 7.5 semitones; however, trained vocalists performed better than those with less training.

Audio-vocal integration for singing requires interactions between the auditory cortex in STG and the IFG (e.g., Broca's area) via the arcuate fasciculus (one of the white-matter tracts connecting regions of the temporal and frontal lobes), and engages constituents of the dorsal sensorimotor stream, such as the dorsal premotor cortex, ACC, the aINS, and inferior parietal lobes (Rauschecker, 2011; Hickok, 2017). These structures underlie both singing and speech, whereas singing recruits a more distributed bilateral network that may engage more right hemispheric regions than speech (Callan et al., 2006; Herbet et al., 2015; Özdemir et al., 2008). Interestingly, it has recently been demonstrated that the motor cortex encodes auditory vocal information in the form of sensorimotor representations of acoustic features rather than articulatory representations (Dichter, Breshears, Leonard, & Chang, 2018; Cheung, Hamilton, Johnson, & Chang, 2016).

This is in line with evidence suggesting that with more singing experience, hearing one's voice becomes less important than feeling the muscle tensions and positions associated with respiratory, laryngeal, and orofacial systems that control the production of pitch (Kleber, Friberg, Zeitouni, & Zatorre, 2017; Mürbe, Pabst, Hofmann, & Sundberg, 2004; Zarate, Wood, & Zatorre, 2010). The representation of laryngeal sensations in S1 (Grabski et al., 2012; Kleber & Zarate, 2014) follows a path like that of the motor presentations in M1 (Brown et al., 2008; Brown et al., 2009). Importantly, the proprioceptive and tactile information is integrated with motor commands already before vocalizations (Bouchard, Mesgarani, Johnson, & Chang, 2013). With experience, these signals become linked with their corresponding acoustic consequences and can thus contribute to coordinating vocal production even in the absence of auditory feedback (Nasir & Ostry, 2006), which at this point will mainly be used to acquire new vocal patterns and to keep the sensory-motor system calibrated (Guenther, 2016).

Chapter 17 by Yennari and Schraer-Joiner on the singing by children who are deaf, offers further insight into the relation between perception and production, as do several other chapters in Part II of this volume on the relation between perception and production. The example of vocalist Mandy Harvey is also a case in point. A university music major with perfect pitch, at the age of 18, she became profoundly deaf due to illness (connective tissue disorder, Ehlers-Danlos syndrome). Unable to hear her own voice, she used visual and tactile feedback to calibrate her vocal system,

and was a semifinalist in “America’s Got Talent” in 2017 (Freeman, 2018; Harvey & Atteberry, 2017).

A singing network

Case studies

Direct brain stimulation (DES)

A case report of an avid singer with a right fronto-temporo-insular lesion¹ provides some evidence for distinct dedicated singing and speaking networks (Herbet et al., 2015). The awake patient underwent direct electrical stimulation (DES) to localize various functions prior to surgery by activating a small cerebral area for a few seconds. He was asked to perform various verbal tasks that also engaged vision and emotion. Stimulation of the anteroposterior pars opercularis of the right inferior frontal gyrus (IFG_{op}), Brodmann areas 44, homologous to part of Broca’s area in the left hemisphere, which also includes area 45) elicited a switch from a speaking to a singing mode. (Accompanying the publication is a web-link to a video of the actual procedure showing the electrode placement and the patient’s verbal response of four syllables. On three occasions, the patient “sings”, producing a requested word in a melodic manner more similar to singing than speaking). Noting that the IFG_{op} has been “previously identified as a crucial cortical area in the response inhibition and task switching networks” (p. 1404), the authors proposed two independent neural networks relatively specialized for either speech or singing, and “a neurocognitive mechanism allowing an individual to flexibly pass from speaking to a singing mode of speech production” (Herbet et al., 2015, p. 1402).² They concluded that similar to

persons who are bilingual, experienced singers may develop a dedicated neural subnetwork for production of “melodically intoned articulation of words” competing with the neural network devoted to language production, and that an inhibitory mechanism enables appropriate use of one over the other (p. 1404).

In contrast to the DES-disruption of speaking by singing, Katlowitz, Oya, Howard, Greenlee, and Long (2017) reported an opposite pattern in a professional male vocalist who was undergoing surgery in the right hemisphere to combat severe epilepsy. The researchers carried out two kinds of direct stimulation to a portion of the right posterior superior temporal gyrus (pSTG)³, applying first DES, as did Herbet et al. (2015) (though in the former case to the IFG), and then *focal cooling*. In this case, singing rather than speaking was suppressed by the electrical stimulation.⁴ Note that the study was not conducted with the aim of investigating a singing network per se. However, Garcea et al. (2017) did deliberately investigate the role of the right STG in music processing using DES with a musician who had a tumor in the right temporal lobe. In this study, the patient was simply asked to hum 74 novel short melodies that were presented to him, 36 of which were presented while receiving DES in three parts of his brain, including the STG. Only during the stimulation of the STG did large errors in melodic production arise. The authors interpret the finding in the context of melodic processing, rather than singing per se, and more research would be needed to determine whether the stimulation affected melodic perception as well as melodic vocalization, as only vocalization (humming, and not perception) was tested.

Transcranial direct current stimulation (tDCS)

Hohmann, Loui, Li, and Schlaug (2018) used transcranial direct current stimulation (tDCS) to disrupt activity independently in four key brain nuclei, the right and left posterior IFG and posterior STG. On separate days, persons without music training or performance experience underwent tDCS stimulation and were asked to imitate individual pitches presented in a comfortable range. Their performance when compared with a sham control condition was disrupted with stimulation to the left IFG and the right STG, consistent with previous identification of these nuclei as key locations in a singing network. The authors conjecture that the right STG plays a role in representing the target pitch and the left IFG plays a role in organizing the motor sequence.

Taken together, the DES findings of Herbet et al. (2015), Katlowitz et al. (2017), and Garcea et al., and the tDCS findings of Hohmann et al (2018) (i.e., evidence of singing disruption at IFG and STG) are consistent with the idea of independent components if not competing networks for singing and speech. Özdemir, Norton, and Schlaug (2006) in an earlier fMRI study, supported the notion of distinct neural correlates for singing (e.g., right STG and portions of the primary sensorimotor cortex) and speaking (e.g., IFG) as well as overlap (e.g., superior STG, STS bilaterally, and inferior pre-and post-central gyrus). In their study, while in the fMRI scanner, participants were asked to sing and speak two-syllable words as well as simply hum or produce vowels. However, a decade later Brown and colleagues, hypothesized “a single vocal system in the human brain that mediates all the vocal functions of human communication and expression, including speaking, singing, and

the expression of emotions” (Belyk & Brown, 2017, p. 182). The difference in positions may be partially semantic. Conceivably, however, both theories may be compatible if we consider that dynamic changes within the network activity determines how the different vocal tasks are supported based on experience.

Transcranial direct current stimulation (TMS)

In a recent study, Finkel et al. (2019) applied repetitive TMS to right larynx S1 and a non-vocal control area in untrained singers to investigate the underlying neural processes. Before and after stimulation, participants performed a pitch-matching singing task. Results revealed that when auditory feedback was masked with loud noise during singing, larynx S1 stimulation enhanced pitch accuracy and pitch stability. The specific effects on voice production suggest that larynx-S1 stimulation affected the preparation and involuntary regulation of (initial) vocal pitch accuracy in persons with little involvement in singing, a group that may lack accurately developed associations between bodily sensations and auditory pitch, whereas pitch stability was enhanced throughout tone production. Together, these findings support a causal role of somatosensation in vocal pitch regulation.

Effects of vocal training and practice – evidence for a critical period

Comparing trained and untrained vocalists adds to the picture of how neural systems become more differentiated with experience. The brains of musicians have been characterized by both increased gray matter and cortical thickness in selective areas and show an altered white matter organization (Zatorre, Fields, & Johansen-Berg,

2012). Several recent studies have focused on experience-dependent structural plasticity of vocalists.

Halwani, Loui, Rüber and Schlaug (2011) obtained magnetic resonance images of professional singers, professional instrumentalists, and non-musicians. A specific region of interest was the arcuate fasciculus (AF). The images revealed that vocalists had a larger left hemisphere tract volume than instrumentalists. Because singers as compared to instrumentalists produce words at the same time as producing melody, extra language practice might account for the larger AF in the left hemisphere of singers. Singers, however, had lower fractional anisotropy (microstructure) measures of the AF, and the anisotropy decreased with years of vocal training. The reduced anisotropy, generally taken as an adaptation arising from experience, was thought to reflect reliance on increasingly complex integration of feedback and feedforward systems required of virtuoso performance levels.

Whereas Halwani et al. (2011) compared groups of trained vocalists, instrumentalists, and non-musicians, a recent study examined singing and playing the cello in the same instrumentalist (Segado, Hollinger, Thibodeau, Penhune, & Zatorre, 2018). The researchers reported overlap in the fMRI activation patterns that compared 11 highly trained cellists in their production of notes on a (specially designed non-magnetic) cello versus vocalization of the same notes. The earlier the cellists had begun taking lessons before the age of 7, the greater was the overlap, and overlap was also proportional to the extent to which performance was in tune. The singing network is evolutionarily old, and structures that support it are phylogenetically older than those

that support language. Segado et al. suggest that musical performance on an instrument co-opts this system in the same way that evolutionarily new cultural tasks such as arithmetic have co-opted functional brain networks for more basic evolutionarily old tasks like direction processing, in accordance with the Theory of Neuronal Recycling (Dehaene & L. Cohen, 2007).

Using voxel-based morphometry, Kleber et al. (2016) showed that classical singers, as compared to participants without vocal training, have increased right hemisphere gray-matter volume in four areas: ventral primary somatosensory cortex (larynx S1), adjacent rostral supramarginal gyrus (BA40), secondary somatosensory cortex (S2), and primary auditory cortex (A1). In another study, singing experience was also positively correlated with increased functional connectivity between the bilateral aINS and the cortical representations of the larynx and the diaphragm within sensorimotor cortex (M1/S1) during resting state fMRI (Zamorano et al., 2019).

Whereas the fMRI findings of Segado et al. (2018) were suggestive of an early critical period during which musical instrument training has an impact on future pitch production accuracy ability, of importance in the study of Kleber et al. (2016) was that vocalists who began training after the age of 14 years, but not earlier, had increased gray-matter in right S1 and the supramarginal gyrus. The extent of the increase was a function of the amount of training after the age of 14 years. This contrasts with experienced performers of musical instruments who show effects of training at earlier ages. The age of 14 years coincides with a first plateau in speech motor development. One might look at this from the point of view of closing the window on a sensitive period for speech and instrument motor development and

opening a window for singing motor development. An evolutionary explanation for this is difficult to suggest, though the timing coincides with the biologically and socially significant stage for mate selection for which singing can play several roles (Miller, 2000).

In another fMRI brain imaging study, Kleber, Veit, Birbaumer, Gruzelier, and Lotze (2010) found that experienced opera singers, compared to non-singers, showed increased blood-oxygen-level-dependent (BOLD) response in S1 (laryngeal and mouth representation) and inferior parietal cortex, as a function of accumulated practice, reflecting better kinesthetic control of the vocal production mechanisms. (see Figure 6.1 C)

In two neuroimaging studies, the right aINS was identified as the main region for gating somatosensory and auditory feedback integration based on singing experience. When a topical anesthetic was applied to the vocal folds, trained singers (in contrast to laypersons) limited detrimental effects on pitch-matching accuracy through reduced insula activation and sensory feedback integration (Kleber et al., 2013). Conversely, pitch-matching accuracy remained high in singers when auditory self-monitoring was masked with loud noise (Kleber et al., 2017), and brain regions that integrate somatosensory feedback with motor control (IPL, aINS, ACC, premotor cortex) showed enhanced activation. (see Figure 6.1 D) People with little to no formal vocal training showed no such compensation strategy in the brain and thus a greater dependency on auditory feedback for controlling the singing voice. This fits with a role of the right anterior insula (aINS) in the coordination of vocal tract movements

during singing (Riecker, Ackermann, Wildgruber, Dogil, & Grodd, 2000; Ackermann and Riecker, 2004) and prosodic or melodic elements of vocalizations (Oh, Duerden and Pang, 2014).

Zarate and Zatorre (2008) and Zarate et al. (2010) asked trained and untrained singers to retain the pitch of the note they were singing while presented with erroneous auditory feedback. Only trained singers could intentionally ignore the erroneous auditory feedback and maintain the initial pitch-level without any motor adjustments, while there were no significant group differences in the ability to compensate for the pitch-change. The main regions associated with audio–vocal integration were the anterior cingulate cortex, auditory cortex, and the aINS. Experience-dependent differences were found in the posterior STG (auditory feedback monitoring) and its increased connectivity with the inferior parietal sulcus (IPS; presumably encoding the size and direction of the pitch shift), which in turn was functionally connected with the ACC and the aINS (Zatorre & Zarate 2012, p. 280). Note that the ACC and aINS are key components of the Salience Network (Sridharan et al., 2008) shown by Alluri et al. (2017) and others to differentiate musicians and non-musicians. Persons without vocal training, in contrast, showed more activity in the dPMC than did experienced singers, possibly reflecting a less efficient motor planning mechanism.

In a further neuroimaging study of experiential effects with implications for the insula, trained vocalists and non-vocalist/non-musicians produced a vowel under conditions of altered auditory feedback (Wang, Chen, Jones, Gong, & Liu, 2019). Voxel-based morphometry revealed reduced grey matter in the area of the insula in

singers. The size of the reduction was inversely correlated with the extent to which auditory feedback led to involuntary correction (training led to reduced involuntary correction). The results suggested greater efficiency in the insular area with increased vocal training, associated also with increased reliance on motor versus auditory feedback with vocal training,

Covert singing manipulations

Because of possible artifacts from head movements while singing or speaking in an fMRI scanner, researchers have often used covert (imagining rather than carrying out an activity) instead of overt singing and speaking tasks while participants undergo neuroimaging. Zatorre and Halpern (2005) reviewed evidence that covert paradigms engaged neural activity that typically underpins overt musical activity. Similar paradigms have been applied to investigate singing (e.g., Callan et al., 2006; Kleber, Birbaumer, Veit, Trevorrow, & Lotze, 2007). Wilson, Abbott, Lusher, Gentle, and Jackson (2011) tested participants who represented three levels of singing expertise which also coincided with their level of pitch accuracy. In the *singing task*, participants covertly sang the beginning of a familiar folk song. In the *word task*, they covertly generated as many words as possible beginning with a visually presented letter. The fMRI results revealed less overlap with the traditional language areas in expert than in non-expert singers, supporting the idea that singing experience modifies the network for speech and song production. Kleber and colleagues (2007) performed the first fMRI study with professional opera singers during overt and covert production of an Italian aria. The results showed that many of the regions that control overt singing were also active during imagined singing. Moreover, imagery

compared to overt singing revealed a larger fronto-parietal network, including the IFG (e.g., Broca's area) and the IPL, which are involved in motor planning and kinesthetic feedback control. This emphasizes the value of mental imagery for the purpose of song rehearsal.

Neurochemical effects

Singing a familiar song is associated with increased activation of the nucleus accumbens (Jeffries, Fritz, & Braun, 2003), part of the brain's pleasure and reward system that modulates levels of dopamine. Dopamine release in the nucleus accumbens (see Figure 6.1 A4, ventral striatum) and surrounding areas has been associated with increased mood, motivation and a drive toward goal-directed behaviors (Chanda & Levitin, 2013). Dopamine replacement therapy in individuals with Parkinson's Disease can lead to "compulsive singing," further underscoring the connection between dopamine, singing, and reward (Bonvin, Horvath, Christe, Landis, & Burkhard 2007). The connection between dopamine and singing has also been found in birds (Simonyan, Horwitz, & Jarvis, 2012), suggesting an ancient evolutionary origin.

Singing is associated with increased levels of immunoglobulin A (IgA, Chanda & Levitin, 2013), an important antibody that stimulates immune function of the mucosal membranes, and with increased levels of oxytocin (Grape, Sandgren, Hansson, Ericson, & Theorel, 2002), a social saliency hormone associated with feelings of bonding and trust . The connection between oxytocin and singing has been established in members of jazz quartet engaged in improvising (Keeler et al., 2015), and in two species of "singing mice", which display an unusually complex vocal repertoire and

exhibit high oxytocin receptor binding within brain regions associated with social memory (Campbell, Ophir, & Phelps, 2009). See also in Volume 3, Chapter 7 (Fancourt & Warren) and Chapter 12 (Launay & Pearce) as well as the review article by Kang, Scholp, and Jiang (2018) for additional studies which imply the effect of singing on the immune function and other neurochemical effects.

Aphasia

The opening of this chapter drew attention to the work of neurologist Salomon Henschen and the search for a brain center for singing. Almost a century later, with the benefit of brain imaging technologies and behavioral research methodologies, his ideas can be verified and greatly extended. Some of the research reviewed above supports his notion of the significance of left hemisphere components in the vicinity of the speech center, where he located the singing center. Since the 1970's, singing-related therapy has been offered as a means of improving the speech of persons who have aphasia. Melodic intonation therapy (MIT) has been used to assist people without expressive speech to be able to sing their mental and emotional states (Albert, Sparks, & Helm, 1973), and is most widely known to the public through its most famous patient, Congresswoman Gabrielle Giffords, who recovered speech following a gunshot wound and MIT (Giffords & Kelly, 2011)⁵. The application of tDCS has been shown to enhance the effects (Vines, Norton & Schlaug, 2011). Schlaug, Marchina and Norton (2008) demonstrated that melodic intonation therapy yielded significant improvement in propositional speech that generalized to unpracticed words and phrases. The beneficial effects were attributed to engagement of the right hemisphere by music. This classical view was partially upheld in a recent fMRI study that applied MIT for 30 sessions over 6 weeks to subacute (<3 months post onset)

stroke patients with severe non-fluent aphasia. In the same study which included patients with chronic aphasia (>1 year post onset), there was no evidence for right hemisphere recruitment resulting from MIT. Rather the neuromaging data (arising from language listening) suggested that in chronic cases, a “reorganisation of language after MIT occurs in interaction with a dynamic recovery process after stroke” (van de Sandt-Koenderman, et al., 2018, p. 765). In a related study with chronic cases, improvements were unimpressive, mostly restricted to improved repetition of trained items, and required regular maintenance (van der Meulen et al., 2016).

Merritt, Zumbansen, & Peretz (2019, p. 379) note that “it has yet to be fully explained how cognitive systems for music and language that are dissociable in the face of brain injury or congenital abnormalities could at the same time be sufficiently linked to enable music networks to support language function”. Zumbansen and Tremblay (2019) in that same issue suggested that benefits of singing in non-fluent aphasia arise in the motor aspects of speech (i.e., rather than semantic) while others have focused on rhythmic practice rather than melodic being the key (Stahl, Kotz, Henseler, Turner, & Geyer, 2011). See also Vol 3. Chapter 8 by Särkämö.

Concluding remarks

An aim of this chapter was to review and integrate the expanding body of literature on the neuroscience of singing. Within the constraints of the chapter, we hope to have laid a groundwork that may be helpful to others in carrying on with this task.

Regarding the question of the neural mechanisms underlying singing development, we can conclude that research is needed on the short and long term impacts of singing

engagement early in life. A controlled study of the effect of 15 months formal musical instrument training in children of six years of age revealed increased relative voxel size in the musically-significant portion of the right temporal lobe (Hyde et al., 2009). We need to know whether weekly singing lessons and regular practice would have had the same effect, or if formal training in singing has its primary impact only after the age of 14 (Kleber et al., 2017).

List of neuroanatomical acronyms

A1	primary auditory cortex
ACC	anterior cingulate cortex
AF	arcuate fasciculus
BA40	Brodmann area 40 - supramarginal gyrus in the parietal lobe
IFG	inferior frontal gyrus
aINS	anterior insula
IPS	intraparietal sulcus
LMC	larynx motor cortex
M1	primary motor cortex
PAG	periaqueductal gray
dPMC	dorsal premotor cortex
S1	primary somatosensory cortex
S2	secondary somatosensory cortex
STG	superior temporal gyrus
pSTG	posterior superior temporal gyrus
STS	superior temporal sulcus

Glossary

Anterior: towards the front (nose) in a vertebrate.

Aphasia: A brain deficit associated with loss of language function.

Association cortex: Any area of the cortex that receives input from more than one sensory system.*

Basal ganglia: a collection of subcortical nuclei (e.g., striatum—[putamen and caudate]-- and globus pallidus) that have important motor functions.

BOLD signal: A blood-oxygen-level-dependent signal, which is recorded by fMRI and is related to the level of neural firing.

Broca's area: A region of frontal lobe (inferior prefrontal cortex/ frontal operculum) of the dominant hemisphere of the brain concerned with the production of speech. It was discovered by French surgeon Paul Broca. Damage in this area causes Broca's aphasia, characterized by hesitant and fragmented speech with little grammatical structure.

Brodman areas – Brain map of areas created by Korbinian Brodmann (2009) to define structures of the cerebral cortex

Default mode network: A brain network including the posterior cingulate cortex and the ventromedial prefrontal cortex, which is responsible for self-related experiences such as autobiographical processing and self-monitoring

Direct electrical stimulation (DES): DES is an exploratory technique used since the early days of neurosurgery to avoid destruction of speech centers during brain surgery for intractable seizures or otherwise unmanageable critical medical conditions. After temporarily removing a portion of the skull, ultrasound first determines the location of the lesion.

Dorsal: Toward the surface of the back of a vertebrate or toward the top of the head.*

Efferent nerves: Nerves that carry motor signals from the central nervous system to the skeletal muscles and internal organs.*

Electrocorticography (ECoG): Direct recordings of brain electrical potentials of the cerebral cortex, typically of patients with severe epilepsy who require surgery. Such patients must first undergo craniotomy (removal of part of the skull) leaving a portion of the cortex exposed to allow mapping of the brain.

Exteroceptive stimuli. Stimuli that arise from outside the body (e.g., sound, light).*

Gray matter. Parts of the nervous system that are gray because they are comprised of “neural cell bodies and unmyelinated interneurons” (Pinel, 2014, p. 484).

Functional magnetic resonance imaging (fMRI): A magnetic resonance imaging is a technique for inferring brain activity by measuring increased oxygen flow into particular areas.*

Heschl’s gyri; or transverse temporal gyri found in the primary auditory cortex, occupying Brodmann areas 41 and 42, superior to and separate from the planum temporale; the first cortical structures to process incoming auditory information

Homunculus: The distorted map of the body in the somatosensory cortex (the “sensory homunculus”) and the motor cortex (“the motor homunculus”). Exaggerated portions (e.g., lips, hands) reflect the more extensive innervation of these organs.

Kinesthetic: See proprioceptive.

Neurons: Cells of the nervous system that are specialized for reception, conduction, and transmission of electrochemical signals.*

Planum temporale: an area of the temporal lobe cortex that lies in the posterior region of the lateral fissure and, in the left hemisphere, roughly corresponds to Wernicke's area.

Proprioceptive: The sensation of the location of self-movement and body position, mediated by mechanically sensitive proprioceptive neurons distributed throughout the body, as muscle spindles (embedded in skeletal muscle fibers), Golgi tendon organs (at the interface of muscles and tendons), and joint receptors (embedded in joint capsules)

Salience network: a brain network that includes the anterior cingulate cortex and the anterior insula, which is responsible for identifying salient stimuli and coordinating cognitive resources, such as working memory and attention, between the default mode network and the central executive network.

Somatosensory feedback: Refers to the sense of movement (kinesthesia) and the location of movement (proprioception).

Sulci: Small furrows in the convoluted cortex.*

Transcranial direct current stimulation (tDCS): A non-invasive weak direct current that flows between two cephalic electrodes to modulate levels of regional brain excitability in targeted cortical regions underlying the electrodes, creating a temporary "virtual lesion". Effects last about 30 minutes after 20 – 30 minutes stimulation (cf. Hohmann et al., 2018).

Wernicke's area: The area of the dominant (typically left) temporal cortex (STG) hypothesized by Wernicke to be the center of language comprehension. Broadman area 22.

White matter: Parts of the brain that are white because they are composed of myelinated axons.*

*based on glossary of Pinel (2014, pp. 478 – 497)

References

- Ackermann, H., and Riecker, A. (2004). The contribution of the insula to motor aspects of speech production: a review and a hypothesis. *Brain & Language*, 89, 320-328. doi: 10.1016/S0093-934X(03)00347-X.
- Albert, M. L., Sparks, R. W., & Helm, N. A. (1973). Melodic intonation therapy for aphasia. *Archives of Neurology*, 29(2), 130-131.
- Alluri, V., Toiviainen, P., Burunat, I., Kliuchko, M., Vuust, P., & Brattico, E. (2017). Connectivity patterns during music listening: Evidence for action-based processing in musicians. *Human Brain Mapping*, 38, 2955–2970.
- Arbib, M. A. (Ed.). (2013). *Language, music, and the brain: A mysterious relation*. Cambridge, MA: MIT.
- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, 403(6767), 309–312.
- Belyk, M., & Brown, S. (2017). The origins of the vocal brain in humans. *Neuroscience and Biobehavioral Reviews*, 77, 177–193.

- Bonvin, C., Horvath, J., Christe, B., Landis, T., & Burkhard, P. R. (2007). Compulsive singing: another aspect of punning in Parkinson's disease. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 62(5), 525-528.
- Bouchard, K. E., Mesgarani, N., Johnson, K., & Chang, E. F. (2013). Functional organization of human sensorimotor cortex for speech articulation. *Nature*, 495, 327–332.
- Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde* [Localisation in the cerebral cortex]. Leipzig, Germany: Johann Ambrosius Barth.
- 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 and Cognition*, 70, 31–41.
- Brown, S., Martinez, M. J., Hodges, D. A., Fox, P. T., & Parsons, L. M. (2004). The song system of the human brain. *Cognitive Brain Research*, 20, 363–375.
- Brown, S., Martinez, M. J., & Parsons, L. M. (2006). Music and language side by side in the brain: A PET study of the generation of melodies and sentences. *European Journal of Neuroscience*, 23, 2791–2803.
- Brown, S., Ngan, E., & Liotti, M. (2008). A larynx area in the human motor cortex. *Cerebral Cortex*, 18, 837–845.
- Callan, D.E., Tsytsarev, V., Hanakawa, T., Callan, A.M., Katsuhara, M., Fukuyama, H., & Turner, R. (2006). Song and speech: brain regions involved with perception and

covert production. *Neuroimage* 31, 1327-1342. doi:

10.1016/j.neuroimage.2006.01.036.

Campbell, P., Ophir, A. G., & Phelps, S. M. (2009). Central vasopressin and oxytocin receptor distributions in two species of singing mice. *Journal of Comparative Neurology*, 516(4), 321-333.

Chanda, M. L., & Levitin, D. J. (2013). The neurochemistry of music. *Trends in Cognitive Sciences*, 17(4), 179-193.

Cheung, C., Hamilton, L. S., Johnson, K., & Chang, E. F. (2016). The auditory representation of speech sounds in the human motor cortex, *eLife*,

2016;5:e12577 DOI: [10.7554/eLife.12577](https://doi.org/10.7554/eLife.12577)

Cohen, A. J. (2000). Development of tonality induction: Plasticity, exposure and training. *Music Perception*, 17, 437-459.

Cohen, A. J. (2019). Singing. In P. J. Rentfrow & D. J. Levitin (Eds.), *Foundations in music psychology* (pp. 685 – 750). Cambridge, MA: MIT Press.

Dalla Bella, S., Berkowska, M., & Sowinski, J. (2011). Disorders of pitch production in tone deafness. *Frontiers in Psychology*, 2, 164.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron* 56, 384–398.

doi: 10.1016/j.neuron.2007.10.004

Dichter, B.K., Breshears, J.D., Leonard, M.K., & Chang, E.F. (2018). The control of vocal pitch in human laryngeal motor cortex. *Cell*, 174, 21-31 e29. doi:

10.1016/j.cell.2018.05.016.

- Fine, P., Berry, A., & Rosner, B. (2006). The effect of pattern recognition and tonal predictability on sight-singing ability. *Psychology of Music*, 34, 431-447.
- Finkel, S., Veit, R., Lotze, M., Friberg, A., Vuust, P., Soekadar, A., . . . Kleber, B. (2019). Intermittent theta burst stimulation over right somatosensory larynx cortex enhances vocal pitch-regulation in nonsingers. *Human Brain Mapping*, 40, 2174-2187. <https://doi.org/10.1002/hbm.24515>
- Fox, P. T., & Friston, K. J. (2012). Distributed processing; distributed functions? *Neuroimage*, 61(2), 407-426.
- Freeman, P. (2018). Singer Mandy Harvey: Losing her hearing, finding her path in life. The “America’s Got Talent” sensation inspires audiences. [Forbes interview]. Popcultureclassics.com/mandy_harvey.html].
- Garcea, F.E., Chernoff, B. L., Diamond, . . . Marvin, e., Pilcher, W. H., 7 Mahon, B. Z. (2017). Direct electrical stimulation in the human brain disrupts melody processing. *Current Biology*, 27, 2684-2691.
- Giffords, G., & Kelly, M. (2011). *Gabby: A story of courage and hope*. New York, NY: Scribner.
- Grabski, K., Lamalle, L., Vilain, C., Schwartz, J. L., Vallée, N., Tropres, I., . . . & Sato, M. (2012). Functional MRI assessment of orofacial articulators: Neural correlates of lip, jaw, larynx, and tongue movements. *Human Brain Mapping*, 33, 2306–2321.
- Grape, C., Sandgren, M., Hansson, L. O., Ericson, M., & Theorell, T. (2002). Does singing promote well-being?: An empirical study of professional and amateur

singers during a singing lesson. *Integrative Physiological & Behavioral Science*, 38(1), 65-74.

Gudmundsdottir, H., & Trehub, S. (2018). Adults recognize toddlers' song renditions. *Psychology of Music*, 46, 281-291.

Guenther, F.H. (2016). *Neural control of speech*. Cambridge, MA: The MIT Press.

Halwani, G. F., Loui, P., Rüber, T., & Schlaug, G. (2011) Effects of practice and experience on the arcuate fasciculus: Comparing singers, instrumentalists, and non-musicians. *Frontiers of Psychology*, 2, 1–9.

Harvey, M., & Atteberry, M. (2017). Sensing the rhythm: Finding my voice in a world without sound. Brentwood, TN: Howard Books [Parent company Simon & Schuster]

Henschen, S.E. (1920). Über Aphasie, Amusie und Akalkulie *Klinische und anatomische Beiträge zur Pathologie des Gehirns* [About aphasia, amusia and acalcululia Clinical and anatomical contributions to the pathology of the brain](Vol. 5). Stockholm, Sweden: Nordiska Bokhandeln.

Herbet, G., Lafargue, G., Almairac, F., Moritz-Gasser, S., Bonnetblanc, F., & Duffau, H. (2015). Disrupting the right pars opercularis with electrical stimulation frees the song: C report. *Journal of Neurosurgery*, 123, 1401–1404.

Hickok, G. (2017). A cortical circuit for voluntary laryngeal control: Implications for the evolution language. *Psychonomic Bulletin and Review*, 24, 56-63. doi: 10.3758/s13423-016-1100-z.

Hohmann, A., Loui, P., Li, C.H., & Schlaug, G. (2018). Reverse engineering tone-deafness: Disrupting pitch-matching by creating temporary dysfunctions in the

auditory-motor network. *Frontiers in Human Neuroscience*, 12, 9. Doi: 10.3389/fnhum.2018.00009/full

Hyde, K. L., Lerch, J., Norton, A., Forgeard, M., Winner, E., Evans, A. C., & Schlaug, G. (2009). *The Journal of Neuroscience*, 29, 3019-3025.

Jeffries, K. J., Fritz, J. B., & Braun, A. R. (2003). Words in melody: An H215O PET study of brain activation during singing and speaking. *Neuroreport*, 14(5), 749-754.

Jürgens, U. (2009). The neural control of vocalization in mammals: A review. *Journal of Voice*, 23, 1–10.

Kang, J., Scholp, A., & Jiang, J. J. (2018). A review of the physiological effects and mechanisms of singing. *Journal of Voice*, 32, 390-395.

Katlowitz, K. A., Oya, H., Howard, M. A., Greenlee, J. D. W., & Long, M. A. (2017). Paradoxical vocal changes in a trained singer by focally cooling the right superior temporal gyrus. *Cortex*, 89, 111–119.

Keeler, J.R., Roth, E.A., Neuser, B.L., Spitsbergen, J.M., Waters, D.J., & Vianney, J.M. (2015). The neurochemistry and social flow of singing: bonding and oxytocin. *Frontiers in Human Neuroscience*, 9, 518.

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

Kleber, B., Friberg, A., Zeitouni, A., and Zatorre, R. (2017). Experience-dependent modulation of right anterior insula and sensorimotor regions as a function of noise-

masked auditory feedback in singers and nonsingers. *Neuroimage* 147, 97-110. doi: 10.1016/j.neuroimage.2016.11.059.

Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J., & Lotze, M. (2010). The brain of opera singers: Experience-dependent changes in functional activation. *Cerebral Cortex*, 20, 1144–1152.

Kleber, B., Veit, R., Moll, C. V., Gaser, C., Birmaumer, & Lotze, M. (2016). Voxel-based morphometry in opera singers: Increased gray-matter volume in right somatosensory and auditory cortices. *NeuroImage*, 133, 477–483.

Kleber, B. A., & Zarate, J. M. (2014). The neuroscience of singing. In G. Welch & J. Nix (Eds.), *The Oxford handbook of singing*. Oxford, UK: Oxford University Press. doi:10.1093/oxfordhb/9780199660773.013.015

Kleber, B., Zeitouni, A.G., Friberg, A., and Zatorre, R.J. (2013). Experience-dependent modulation of feedback integration during singing: role of the right anterior insula. *J Neurosci* 33, 6070-6080. doi: 10.1523/JNEUROSCI.4418-12.2013.

Koelsch, S. (2019). Music and the brain. In J. Rentfrew & D. Levitin (Eds.) *Foundations in music psychology* (pp. 407 – 458). Cambridge, MA: MIT Press.

Lerdahl, F. & Jackendoff, R. (1983). *The Generative theory of tonal music*. Cambridge, MA: MIT Press.

Levitin, D. J. (1999). Tone deafness: failures of musical anticipation and self-reference. *International Journal of Computing and Anticipatory Systems*, 4, 243-254.

- Levitin, D. (2006). *This is your brain on music: The science of a human obsession*. New York, NY: Dutton/ Penguin.
- Loui, P. (2015). A dual-stream neuroanatomy of singing. *Music Perception*, 32, 232-241. DOI: 10.1525/MP.2015.32.3.232
- Luu, P., Kelley, J. M., & Levitin, D. J. (2001). Consciousness: A preparatory and comparative process. In P. G. Grossenbacher (Ed.), *Finding consciousness in the brain: A neurocognitive approach* (pp. 243-270). Philadelphia, PA: John Benjamins.
- McAdams, S. & Siedenburg (2019). Perception and cognition of musical timbre. In P. J. Rentfrew & D. Levitin (Eds.) *Foundations in music psychology* (pp. 71 – 120). Cambridge, MA: MIT Press.
- Menon, V. (2013). Developmental pathways to functional brain networks: emerging principles. *Trends in cognitive sciences*, 17(12), 627-640.
- Merrett, D. L., Zumbansen, A., & Peretz, I. (2019). A theoretical and clinical account of music and aphasia. *Aphasiology*, 33, 379-381. doi=10.1080/02687038.2018.1546468
- Miller, G. F. (2000). Evolution of human music through sexual selection. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 329-360). Cambridge, MA: MIT Press.
- Mürbe, D., Pabst, F., Hofmann, G., and Sundberg, J. (2004). Effects of a professional solo singer education on auditory and kinesthetic feedback--a longitudinal study of singers' pitch control. *Journal of voice* 18, 236-241.
- Murry, T. (1990). Pitch-matching accuracy in singers and nonsingers. *Journal of Voice*, 4, 317-321.

- Nasir, S. M., & Ostry, D. J. (2006). Somatosensory precision in speech production. *Current Biology*, 16, 1918–1923.
- Oh, A., Duerden, E.G., and Pang, E.W. (2014). The role of the insula in speech and language processing. *Brain & Language*, 135, 96-103. doi: 10.1016/j.bandl.2014.06.003.
- Ooishi, Y., Mukai, H., Watanabe, K., Kawato, S., & Kashino, M. (2017). Increase in salivary oxytocin and decrease in salivary cortisol after listening to relaxing slow-tempo and exciting fast-tempo music. *PLoS One*, 12(12):e0189075.
- Oxenham, A.J. (2019). Pitch: Perception and neural coding. In *Foundations in music psychology* (pp. 3 – 32). Cambridge, MA: MIT Press.
- Özdemir, E., Norton, A., & Schlaug, G. (2006). Shared and distinct neural correlates of singing and speaking. *NeuroImage*, 33, 628–635.
- Pfordresher, P. Q., Demorest, S. M., Dalla Bella, S., Hutchins, S., Loui, P., Rutkowski, J., & Welch, G. F. (2015). Theoretical perspectives on singing accuracy: An introduction to the special issue on singing accuracy (Part 1). *Music Perception*, 32, 227–231.
- Pinel, J. P. J. (2014). *Biopsychology*. Toronto, Canada: Pearson.
- Pribram, K. H. (1982). Localization and distribution of function in the brain. In J. Orbach (Ed.), *Neuropsychology after Lashley* (pp. 273-296). Hillsdale, NJ: Erlbaum.
- Rauschecker, J.P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing Research* 271, 16-25. doi: 10.1016/j.heares.2010.09.001.

- 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*, 1997–2000.
- Ross, E. D. (2010). Cerebral localization of functions and the neurology of language: fact versus fiction or is it something else? *The Neuroscientist*, *16*(3), 222-243.
- Schlaug, G., Marchina, S., & Norton, A. (2008). From singing to speaking: why singing may lead to recovery of expressive language function in patients with Broca's aphasia. *Music perception: An interdisciplinary journal*, *25*, 315-323.
- Segado, M., Hollinger, A., Thibodeau, J., Penhune, V., & Zatorre, R.J. (2018). Partially overlapping brain networks for singing and cello playing. *Frontiers in Neuroscience*, *12*, 351. doi: 10.3389/fnins.2018.00351
- Simonyan, K., & Horwitz, B. (2011). Laryngeal motor cortex and control of speech in humans. *Neuroscientist* *17*(2), 197–208.
- Simonyan, K., Horwitz, B., & Jarvis, E. D. (2012). Dopamine regulation of human speech and bird song: a critical review. *Brain and Language*, *122*(3), 142-150.
- Sridharan, D., Levitin, D. J., & Menon, V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. *Proceedings of the National Academy of Sciences*, *105*, 12569–12574.
- Stahl, B., Kotz, S. A., Henseler, I., Turner, R., and Geyer, S. (2011). Rhythm in disguise: why singing may not hold the key to recovery from aphasia. *Brain* *134*, 3083–3093
- Sundberg, J. (1987). *The science of the singing voice*. DeKalb, IL: University of Northern Illinois Press.

- Tan, S.-L., Pfordresher, P., & Harré, R. (2016). *Psychology of music: From sound to significance*. New York, NY: Routledge.
- Trainor, L. J. (2018). The origins of music: Auditory scene analysis, evolution, and culture in musical creation (pp. 81-112). In H. Honing (Ed.), *The origins of musicality* (pp. 81-112). Cambridge, MA: MIT Press.
- Tsang, C. D., Friendly, R. H., & Trainor, L. J. (2011). Singing development as a sensorimotor interaction problem. *Psychomusicology: Music, Mind, & Brain*, 21, 45–53.
- Van Der Meulen, I., Van De Sandt-Koenderman, M. W. M. E., Heijenbrok, M. H., Visch-Brink, E., & Ribbers, G. M. (2016). Melodic intonation therapy in chronic aphasia: Evidence from a pilot randomized controlled trial. *Frontiers in Human Neuroscience*, 10, 533. doi: [10.3389/fnhum.2016.00533](https://doi.org/10.3389/fnhum.2016.00533)
- Van de Sandt-Koenderman, M. W. E., Mendez Orellana, C. P., van der Meulen, I., Smits, M., & Ribbers, G. M. (2018). Language lateralisation after melodic intonation therapy: an fMRI study in subacute and chronic aphasia. *Aphasiology*, 32(7), 765-783. doi.org/10.1080/02687038.2016.1240353
- Vines, B. W., Norton, A. C., & Schlaug, G. (2011). Non-invasive brain stimulation enhances the effects of melodic intonation therapy. *Frontiers in Psychology*, 2, 230.
- Ward, W.D., & Burns, E. M. (1978). Singing without auditory feedback. *Journal of Research in Singing & Applied Vocal Pedagogy*, 1, 24-44.
- Wang, W., Wei, L., Chen, N., Jones, J. A., Gong, G., & Liu, H. (2019). Decreased gray-matter volume in insular cortex as a correlate of singers' enhanced sensorimotor

control of vocal production. *Frontiers in Neuroscience*, *13*, 815.
doi.org/10.3389/fnins.2019.00815

Wilson, S. J., Abbott, D. F., Lusher, D., Gentle, E. C., & Jackson, G. F. (2011). Finding your voice: A singing lesson from functional imaging. *Human Brain Mapping*, *32*, 2115–2130.

Zamorano, A.M., Zatorre, R.J., Vuust, P., Friberg, A., Birbaumer, N., and Kleber, B. (2019). Enhanced insular connectivity with speech sensorimotor regions in trained singers – a resting-state fMRI study. *bioRxiv*, 793083. doi: 10.1101/793083.

Zarate, J. M. (2013). The neural control of singing. *Frontiers in Human Neuroscience*, *7*, 1–12. doi:10.3389/fnhum.2013.00237.

Zarate, J. M., Wood, S., & Zatorre, R. J. (2010). Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers. *Neuropsychologia*, *48*, 607–618. doi:10.1016/j.neuropsychologia.2009.10.025.

Zarate, J. M., & Zatorre, R. (2008). Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage*, *40*, 1871–1887. doi:10.1016/j.neuroimage.2008.01.026.

Zatorre, R. J., Fields, R. D., & Johansen-Berg, H. (2012). Plasticity in gray and white: Neuroimaging changes in brain structure during learning. *Nature Neuroscience*, *15* (4), 528-536. doi: [10.1038/nn.3045](https://doi.org/10.1038/nn.3045)

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

- Zatorre, R. J., & Zarate, J. M. (2012). Cortical processing of music. In D. Poeppel, T. Overath, A. N. Popper, & R. R. Fay (Eds.), *The human auditory cortex. Springer Handbook of Auditory Research*, 43 (pp. 261– 294). New York: Springer.
- Zatorre, R. J., & Baum, S. R. (2012). Musical melody and speech intonation: Singing a different tune? *PLoS Biology*, 10, e1001372.
- Zumbansen, A., & Tremblay, P. (2019). Music-based interventions for aphasia could act through a motor-speech mechanism: a systematic review and case-control analysis of published individual participant data. *Aphasiology*, 33, 466-497. DOI: 10.1080/02687038.2018.1506089

¹ Recall that Riecker, Ackermann, Wildgruber, Dogil, and Grodd (2000) suggested the aINS coordinates vocal tract movement in singing.

² The pars opercularis is part of Brodmann area 44 (B44), when in the left hemisphere known as Broca's area. Brown, Martinez and Parsons (2006) noted greater activation in the right than left pars opercularis for generation of melodies versus sentences respectively, testing only persons without specialized musical training. The right pars opercularis has been associated with response inhibition and inhibition of speech. A parallel is drawn between the spontaneous activation/suppression of the singing and speech systems and similar evidence of activation/suppression of two languages in bilingual persons.

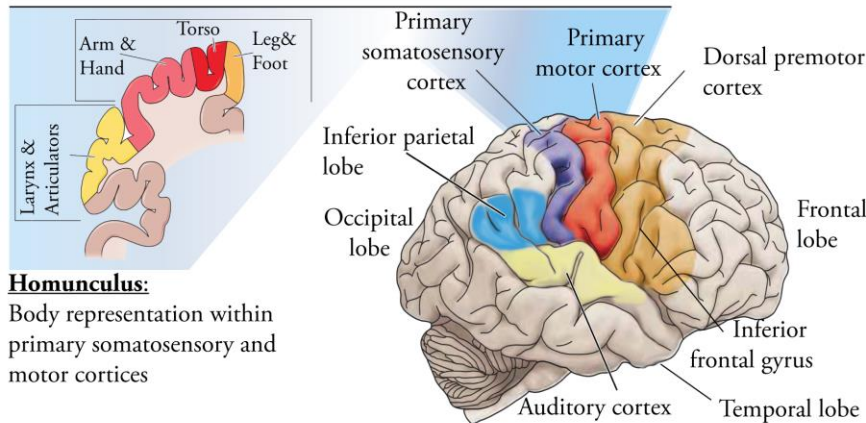
³ Exact borders of Wernicke's area are a matter of debate. The left sided pSTG is commonly assumed to be a part of Wernicke's area. The area uncovered with

electrical stimulation (and thereafter cooled) was within the parallel location on the right side (Kalman Katlowitz & Michael Long, personal communication, 2017).

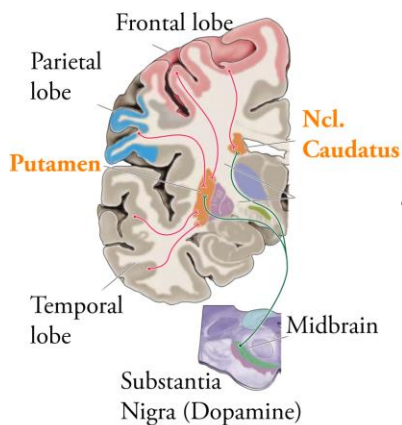
⁴ Focal cooling caused the fundamental (f_0 , pitch) of vowels for speech to increase by a small audible amount. For both singing and speech, the first and second formants increased slightly in frequency. When the cortex returned to its original state of warmth, these formant changes returned to baseline. Because the vocal tract shape creates the resonances that influence the formant frequencies, it appears that the stimulated area of the brain slightly influenced the muscles of the vocal tract.

⁵ ABC News (2011). Gabby Giffords: Finding words through song. https://abcnews.go.com/Health/w_MindBodyNews/gabby-giffords-finding-voice-music-therapy/story?id=14903987

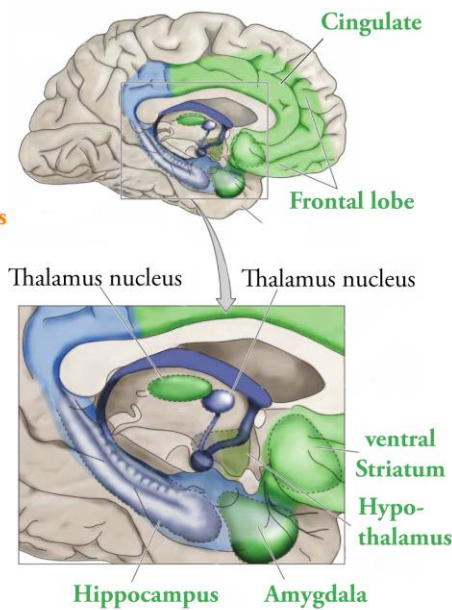
(A1) CORTICAL AREAS INVOLVED IN SINGING



**(A2) BASAL GANGLIA
(BASAL GANGLIA)**



(A3) LIMBIC SYSTEM



(A4) BRAINSTEM

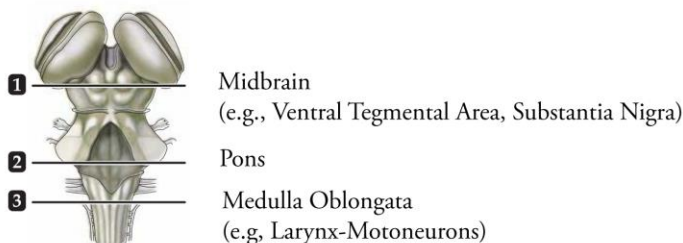


Figure 6.1 Caption

(A) Brain areas involved in human song production: A1 cerebral cortex; A2 basal ganglia; A3 limbic system; A4 brainstem. Images adapted by B. Kleber from “Neuroscience – Fifth edition” (edited by Purves et al., 2012). A1: Figure 17.5 p. 381; A2 Figure 18.2 p. 400; A3 Figure 29.4 p. 652, and A4 17.12 p. 391. Used with kind permission of Oxford University Press.

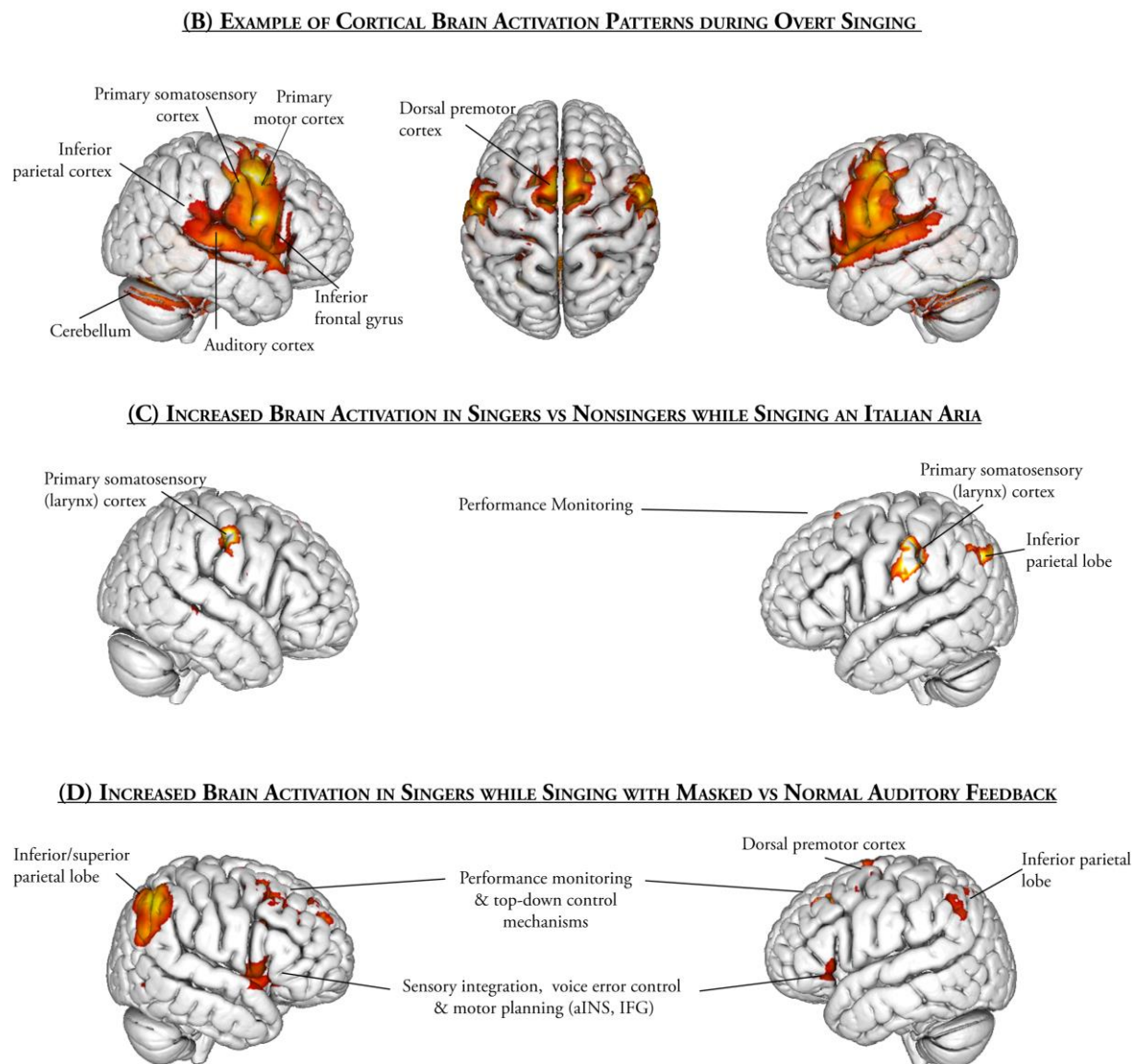


Figure 6.1 Caption

- (B) Brain areas involved in human song production: A1 cerebral cortex; A2 basal ganglia; A3 limbic system; A4 brainstem. Images adapted by B. Kleber from “Neuroscience – Fifth edition” (edited by Purves et al., 2012). A1: Figure 17.5 p. 381; A2 Figure 18.2 p. 400; A3 Figure 29.4 p. 652, and A4 17.12 p. 391. Used with kind permission of Oxford University Press.
- (C) Cortical activation patterns during singing in an fMRI scanner for 42 persons (15 classical singers, 13 rock/jazz singers, and 14 non-singers) comparing overt song production to rest. Involvement of the cerebellum bilaterally is also shown.
- (D) Cortical activation patterns during singing an Italian aria related to accumulated singing practice (i.e., the number of years x the average weekly singing hours) including 10 opera singers, 21 vocal students, and 18 medical students
- (E) Data from 11 highly trained singers who imitated (sang) two-note sequences under two conditions (i) with loud noise masking auditory feedback from their own voice and (ii) without noise and normal auditory feedback.

NOTE: Images from 6.1 B, C, and D are provided by Boris Kleber: B – from his original unpublished data; C new graphical presentations based on original data discussed in Kleber, Veit, Birbaumer, Gruzelier and Lotze (2010); D -recreated images from data previously presented in another format as Figure 4B (Kleber, Friberg, Zeitouni, & Zatorre, 2017)