

A DUAL-STREAM NEUROANATOMY OF SINGING

PSYCHE LOUI
Wesleyan University

SINGING REQUIRES EFFORTLESS AND EFFICIENT USE of auditory and motor systems that center around the perception and production of the human voice. Although perception and production are usually tightly coupled functions, occasional mismatches between the two systems inform us of dissociable pathways in the brain systems that enable singing. Here I review the literature on perception and production in the auditory modality, and propose a dual-stream neuroanatomical model that subserves singing. I will discuss studies surrounding the neural functions of feedforward, feedback, and efference systems that control vocal monitoring, as well as the white matter pathways that connect frontal and temporal regions that are involved in perception and production. I will also consider disruptions of the perception-production network that are evident in tone-deaf individuals and poor pitch singers. Finally, by comparing expert singers against other musicians and nonmusicians, I will evaluate the possibility that singing training might offer rehabilitation from these disruptions through neuroplasticity of the perception-production network. Taken together, the best available evidence supports a model of dorsal and ventral pathways in auditory-motor integration that enables singing and is shared with language, music, speech, and human interactions in the auditory environment.

Received: June 6, 2014, accepted October 1, 2014.

Key words: singing, voice, pitch, neuroanatomy, arcuate fasciculus

SINGING REQUIRES EFFICIENT AND EFFORTLESS control of the vocal apparatus, accurate perceptual and cognitive representations of pitch, rhythm, and structural components of music, and continuous monitoring and transformation between musical intentions and their motoric implementations. In order to accomplish proficiency in singing, the human brain must operate multiple neural circuits in concert, especially auditory-motor circuits that give rise to the feedforward and feedback circuits of vocal control (Zarate,

2013). Research in multiple subfields of neuroscience has offered neural models of perceptual-cognitive and motor behavior including vision, audition, and speech and language (Goodale & Milner, 1992; Hickok & Poeppel, 2004; Rauschecker & Scott, 2009). While these models may be extensible to provide descriptions of singing behavior, newer results from trained singers and tone-deaf individuals – broadly construable as extremes in the broad spectrum of individual differences in singing and musical ability – further refine existing models of the neural substrates of singing. Here I define a possible neuroanatomical basis for a dual-stream model of singing. This model is centered around the superior longitudinal fasciculus, a white matter pathway that includes the arcuate fasciculus at its core, but contains multiple way stations including two streams of anatomical connectivity between superior temporal and inferior frontal regions of the brain. The dual-stream neuroanatomical model of singing explains behavioral results and the neuropsychological literature on pitch perception and production, offering specific neural substrates for the dissociable features of singing behavior, while generating predictions for optimal and impaired behaviors subserved by each of the dorsal and ventral pathways. The possible use of dual-stream models in singing is also posited elsewhere, notably by Zarate’s review on the neural control of singing (Zarate, 2013), which surveys the evidence from training and neuroplasticity, amusia, and dual-stream models of “what” vs. “where” pathways. In the current model I relate the neuroanatomical model for singing more explicitly to white matter, which forms the anatomical structure of dorsal and ventral streams in the present dual-stream model.

A Dual-Stream Neuroanatomical Architecture of Singing

In this special issue on singing, Pfordresher and colleagues propose a theoretical model of singing that posits the existence of multiple pathways involved, one that involves direct sensorimotor translations, and another that passes through categorization and categorical representations of sounds. In the current dual-stream neuroanatomical model of singing proposed in this paper, I link the proposed cognitive model to its proposed neuroanatomical substrates. The neuroanatomical model of

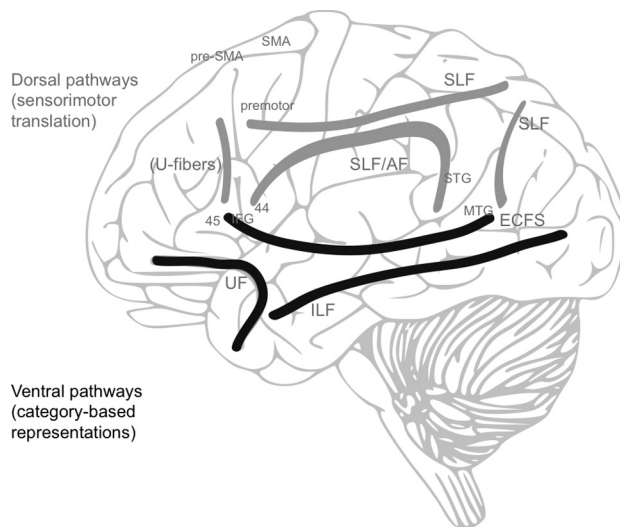


FIGURE 1. Schematic of the approximate neuroanatomical locations of white matter pathways described in the current review. White matter and grey matter structures are labeled in grey text. White matter structures: SLF = superior longitudinal fasciculus; AF = arcuate fasciculus; UF = uncinate fasciculus; ILF = inferior longitudinal fasciculus; ECFS = extreme capsule fiber system. Grey matter structures: STG = superior temporal gyrus; MTG = middle temporal gyrus; IFG = inferior frontal gyrus; 44 and 45 = Brodmann areas 44 and 45; SMA = supplementary motor area; pre-SMA = pre-supplementary motor area.

singing assumes that intact connectivity within a network of white matter association fibers within the brain is crucial for singing. This network of connectivity is illustrated in Figure 1. It centers around the superior longitudinal fasciculus, which includes the arcuate fasciculus but extends to the extreme capsule fiber system, the uncinate fasciculus, and u-fibers that link adjacent branches of white matter pathways. Specifically, the dual streams refer to a dorsal pathway and a ventral pathway. The dorsal pathway connects the posterior superior temporal gyrus with frontal areas, specifically the inferior frontal gyrus (IFG), which includes its two subregions the pars opercularis (Brodmann area 44) and pars triangularis (Brodmann area 45), premotor area, Supplementary Motor Area (SMA), and pre-Supplementary Motor Area (pre-SMA). The ventral pathway extends into the middle temporal gyrus and includes the extreme capsule fiber system, connecting to the pars triangularis of the IFG.

Overview of Fronto-Temporal White Matter Pathways

While the neuropsychological data on tone-deafness provide convincing support for the existence of dual

pathways in singing, the utility of a neuroanatomical model lies in forming specific and testable predictions of how the interconnected components of a theoretically connected model, such as sensorimotor translation, categorization, and even conscious awareness, may map onto the anatomy of the brain. While computations of the human brain rely on neuronal dendrites and cell bodies in grey matter, white matter contains axonal projections that transmit information between cell bodies. In the following paragraphs, based on the white matter atlas by Wakana, Jiang, Nagae-Poetscher, van Zijl, and Mori (2004), I provide a brief outline of the structure of white matter pathways that are known to play a role in the auditory-motor functions involved in singing. These pathways include the superior longitudinal fasciculus, of which the arcuate fasciculus forms a part, as well as the uncinate fasciculus and u-fibers that connect proximal regions of grey matter.

The *superior longitudinal fasciculus* (SLF) is involved in associative tasks, possibly including higher motor behavior motor regulation. The SLF includes three subsets of fiber pathways: SLF1, SLF2, and SLF3. The SLF1 involves medial and dorsal branches from the frontal lobe extending to the superior parietal lobule, as well as the superior precuneus, postcentral gyrus, precentral gyrus, posterior superior frontal gyrus, and SMA. The SLF2 is the main component of SLF. It includes the white matter of the occipital-temporal-parietal transition region (BA 19), angular gyrus (BA 39), supramarginal gyrus (BA 40), postcentral gyrus, precentral gyrus (BA 4), and middle frontal gyrus (BA 6 and 46). This is the bidirectional link between the prefrontal cortex and the parietal lobe, and is involved in the perception of visual space, and the focusing of attention to different parts of space. The SLF3 is ventral and lateral. The fibers of SLF3 course anterior-posteriorly between supramarginal gyrus (BA 40), ventral part of more mesial areas and premotor areas (BA 44). The SLF3 provides the ventral premotor region and the adjacent area 4 with higher order somatosensory input, and may be critical for monitoring orofacial and hand actions, and have a role in the articulatory component of language.

The *arcuate fasciculus* extends from cortex of the posterior ventrolateral frontal lobe, arches around lateral Sylvian fissure, and terminates in posterior superior and middle temporal gyrus. This may also contain a posterior lateral segment connecting temporal and parietal lobes and an anterior lateral segment that connects frontoparietal lobes (indirect pathways). This is the connection between Broca's and Wernicke's language areas. Additionally there is an indirect pathway that projects through the parietal lobe (Catani et al., 2007), which

may be related to semantically based language functions. On the right hemisphere, the dorsal branch of the arcuate fasciculus was not identified using diffusion tensor tractography among tone-deaf individuals (Loui, Alsop, & Schlaug, 2009), suggesting a major disruption in the dorsal branch, possibly corresponding to the direct pathway as identified by Catani et al. (2007). The ventral branch of the arcuate fasciculus is shown to be correlated, in identifiable volume and white matter integrity (as indexed by Fractional Anisotropy), with performance on a music learning task that assesses the acquisition of musical structure via the differentiation of novel grammatical items from ungrammatical items (Loui, Li, & Schlaug, 2011). Together, these two findings suggest that categorical learning (which is a prerequisite of learning musical structure) is a dissociable route from fine-grained pitch perception abilities, and that these two streams rely on the ventral and dorsal arcuate fasciculus respectively.

The *uncinate fasciculus* (UF) is another association tract that connects lateral and medial orbitofrontal cortex with the anterior temporal lobes. It enables interactions between the anterior temporal lobe, which includes areas for voice processing (Capilla, Belin, & Gross, 2013), and the lateral and medial orbitofrontal cortex, which includes areas important for memory association and emotional valence and decision-making (Von Der Heide, Skipper, Klobusicky, & Olson, 2013). In other words, the uncinate fasciculus may enable the linking of voice processing with the formation of emotional associations with autobiographical and well-learned memories, an aspect of cognitive-to-affective function that is especially important for singing.

U-fibers, or small associative fibers, are responsible for cortico-cortical connections within and between adjacent grey matter regions around a sulcus, as well as adjacent white matter tracts. One example of u-fiber pathways is the white matter that connects superior and middle temporal gyri. These regions are implicated in sound perception and categorization respectively, and the volume of u-fibers connecting these regions is found to be larger, especially in the left hemisphere, among people with absolute pitch (Loui, Li, Hohmann, & Schlaug, 2011). These findings suggest that the u-fibers may play a role in enabling the mapping between the perception of pitch and its mapping onto an appropriate categories, an ability that is exceptional in people with absolute pitch and is supported by structural (Loui, Li, et al., 2011) and functional connectivity studies (Loui, Zamm, & Schlaug, 2012). Other examples include fibers that connect Brodmann areas 45 (*pars triangularis*) and 44 (*pars*

opercularis) with areas 6 and 8 in premotor, pre-SMA, and SMA areas. These u-fibers are important in the connection and integration of areas of the brain that are important in sequencing and grammatical knowledge, with areas that are important in the storage and selection of motor plans.

Dual-Stream Models in Neuroscience

The idea of dual stream models is by no means unique to singing behavior, but has been pervasive in multiple domains of investigation in neuroscience and cognitive science. In the field of vision, Mishkin, Ungerleider, and Macko (1983) first posited, based on behavioral data and physiological recordings in the rhesus monkey, that multiple visual areas are hierarchically organized into two separate cortical visual pathways that are specialized for object vision and spatial vision respectively. The pathway that is specialized for object vision consists of multi-synaptic projections, from posterior to anterior from the occipital lobe to the temporal lobe, following along the course of the white matter pathway known as the inferior longitudinal fasciculus. In contrast, the pathway that is specialized for spatial vision consists of multi-synaptic occipital-parietal projections that follow the course of the superior longitudinal fasciculus (Mishkin et al., 1983). These dual pathways, commonly termed “what” vs. “where” pathways, have become highly influential in cognitive science and neuroscience. Importantly, these pathways are identified as “streams,” or combinations of specific regions interconnected by multiple synapses that form each pathway, rather than on single regions functioning in isolation. The regions within each stream are anatomically connected to form a brain network, consisting of nodes that are anatomical regions, and edges that are synaptic projections between these regions.

Linking the primate physiological data to human neuropsychology, Goodale and Milner (1992) posit separate anatomical pathways for visual perception and action. Drawing from a combination of neurophysiological data and neuropsychological evidence, they propose that the ventral stream of projections from the striate cortex to the inferior temporal cortex plays a major role in perceptual identification of visual objects, whereas the dorsal stream, which projects from the striate cortex to the posterior parietal lobe, gives rise to the sensorimotor transformations that are necessary for the visual direction of actions towards objects. One piece of evidence in support of these separate pathways comes from rare cases of patients with visual-form agnosia, who have damage in Brodmann areas 18 and 19, which are

part of the ventral projections from the visual cortex. These patients are unable to recognize the size, shape, and orientation of visual objects, but can scale the size of their hand grip when instructed to reach out and pick up the object. In other words, these patients are differentiated with respect to their access to consciousness, in that they have no conscious perception of objects but have unconscious access to perceptual information such as size and orientation. These unique cases suggest that separate pathways can exist to give rise to two sources of information: conscious access to perceptual features of the object, and unconscious sensitivity to features that enable sensorimotor interaction with the same object.

“Where” and “What” Pathways in the Auditory System

While the dual-stream models of vision have received support from neuropsychological and behavioral evidence (Goodale, Milner, Jakobson, & Carey, 1991) as well as neurophysiological data (DeYoe & Van Essen, 1988), researchers have debated the existence and characteristic features of “what” vs. “where” pathways in the auditory domain. What and where processing in the auditory cortex are posited to lie in streams between the auditory cortex and the prefrontal cortex (Kaas & Hackett, 1999, 2000). Based on a combination of behavioral and neurophysiological data from human and non-human primates, Rauschecker and Tian (2000) posited that auditory cortical pathways are also organized in parallel streams. Specifically, a dorsal-parietal pathway, originating from the caudal or caudomedial part of the supratemporal plane and involving the inferior parietal areas, may be an important way station for the processing of auditory movement and spatial information, thus forming the dorsal “where” pathway. In contrast, lateral belt areas of the superior temporal gyrus (STG) are involved in the early processing of species-specific vocalizations as well as human speech, thus forming an auditory “what” pathway (Rauschecker & Tian, 2000). In this model, sensitivity to vocalizations is central to the “what” pathway, as the identification of vocal sounds is necessary in order to derive perceptually invariant categories of sounds, and ultimately to enable communication. However, some later instantiations of this dual-stream model expand the role of the dorsal pathway to encompass some speech and language functions in humans (Rauschecker, 2011), as well as expanding the role of the antero-ventral pathway to include the categorization and identification of non-speech sounds (Rauschecker & Scott, 2009). Furthermore, functions of the dorsal pathway can be construed as a forward

model, in which an efference copy (i.e., an internal copy of the motor commands that happens during motor events such as singing) that is sent from the prefrontal and premotor cortices provides the basis for forward and inverse mapping between motor actions and the production of intentional sounds. By functioning as way stations along the feedforward and feedback loops, these frontal lobe regions couple with inferior parietal lobe and then with sensory areas of the posterior auditory cortex to learn through multiple iterations, eventually arriving at efficient predictions of optimal states at every instance (Rauschecker, 2011).

Feedforward, Feedback, and Efference Copies

In considering how the two streams of a dual-stream model for singing might be integrated, a plausible neural mechanism involves the role of efference copies. The concept of efference copy originates from the motor control literature and describes an identical signal (a copy) of the feedforward motor command that provides internal feedback for one’s own actions to be compared with reafferent input to the sensory input (efference) areas. The role of the efference copy in learning speech is highlighted by work on sensorimotor adaptation during speech production, in which behavioral compensations are observed in response to formant perturbations (Houde & Jordan, 2002; Katseff, Houde, & Johnson, 2012) and pitch perturbations (Burnett, Freedland, Larson, & Hain, 1998; Patel, Niziolek, Reilly, & Guenther, 2011) of one’s own voice during vocalizations. In these situations of perturbed auditory feedback, a typical response of the subjects is to compensate for the perturbation by adjusting one’s voice in the opposite direction to the perturbation. The presence of these compensatory responses are robust evidence for the use of a neural system for feedback processing, that couples with feedforward processes in order to monitor and correct for the errors that are detected in one’s own voice.

This coupling between feedback and feedforward systems is crucial for the link between conscious and automatic routes of the dual-stream model is well described by the Directions Into Velocity Articulators (DIVA) model (Guenther, 2006), which is proposed to account for learning of speech. In this model, the left hemisphere is predominantly recruited by the feedforward system whereas the right hemisphere is more involved in the feedback system. The feedforward system includes auditory target cells that code for the intended sound of one’s productions. In contrast, the feedback system involves auditory error detection cells that compare the

feedback with the intended output. While the dual-hemispheric dichotomy between feedforward and feedback systems has received support from fMRI studies on speech perturbation (Golfinopoulos, Tourville, & Guenther, 2010; Tourville, Reilly, & Guenther, 2008), the left-right hemispheric distinction of this model remains to be tested in musical contexts. Nonetheless, a major strength of the DIVA model lies in its formation of specific and testable predictions for each brain region that is involved in the speech motor skill acquisition and speech sound production.

Another proposed neuroanatomy for dual-stream processing entails left-right hemispheric differences and comes from Zatorre, Belin, and Penhune (2002), who posit that the left auditory cortex develops to be relatively specialized for speech whereas the right hemisphere relatively specializes for music in order to optimize a tradeoff in spectral-temporal processing. As speech sounds predominantly emphasize rapid temporal changes whereas other sounds, such as music, feature slower temporal fluctuations but precise spectral fluctuations, this model provides an explanation for the left-lateralization of language processing from speech sounds. A similar view is the Asymmetric Sampling in Time (AST) model (Poeppel, 2003), which also posits that there are different time scales of processing between the two hemispheres. In this model, the left hemisphere specializes in rapid sampling whereas the right hemisphere specializes in slow temporal processing. This conceptual framework is in line with Glasser and Rilling's (2008) model of language, which was offered in light of diffusion tensor imaging data that compared the language pathways (arcuate fasciculus) of humans and other primates (Rilling et al., 2008). Compared to the human arcuate fasciculus, the nonhuman primates' arcuate fasciculus showed less anatomical extensions into the middle temporal gyrus. This finding has profound implications for the evolution of language and led to the authors' model (Glasser & Rilling, 2008), which posits that lexical-semantic access is represented on the left hemisphere and prosodic access on the right hemisphere. It is loosely in line with Hickok and Poeppel's (2004) models of structures and functions in language, which are also widely cited as a conceptual separation between dorsal and ventral streams for language. A later instantiation of the same model expands to include more general hemispheric distinctions between bilateral ventral stream and a left-hemispheric dorsal stream (Hickok & Poeppel, 2007).

A compelling demonstration of the dual-route morphology of language white matter pathways comes from Frey et al's DTI-based neuroanatomical investigation

(Frey, Campbell, Pike, & Petrides, 2008), which shows an anatomical distinction between ECFS and AF, two pathways involved in language processing. The extreme capsule fiber system (ECFS) is a pathway of white matter association fibers that passes from the temporal lobe to the frontal lobe through the extreme capsule, connecting area 45 (pars triangularis of the IFG) to the STG. In contrast, the arcuate fasciculus proper connects to premotor and motor regions in Brodmann areas 6 and 8 (Patterson, Van Petten, Beeson, Rapcsak, & Plante, 2014). Based on these results, as well as combined fMRI and DTI data on differences in grammar learning (Friederici, Bahlmann, Heim, Schubotz, & Anwander, 2006), Friederici and others observed that the dorsal pathway, which projects from the posterior portion of Broca's area (Brodmann area 44) to the superior temporal lobe and presumably encompasses the dorsal pathway of the AF proper, seems to be of particular importance for higher-order language functions including hierarchical grammatical knowledge (Friederici, 2009). In contrast, the ECFS is part of the ventral pathway and may participate in processing local transitions (Friederici et al., 2006).

Dissociating Functional Streams with Feedback Manipulation

The currently proposed dual-stream model of singing posits the anatomical dissociation via frontal-temporal white matter pathways that enable direct sensorimotor translations (the relatively direct mapping of sensory input to motor output) and categorically based representations of sounds. An important prediction of this model is that direct manipulations of sensorimotor information can affect motor output and categorical representations through dorsal and ventral frontotemporal white matter pathways respectively. In this regard, studies that involve the manipulation of feedback (e.g., Hafke, 2008; Zarate & Zatorre, 2008) are particularly informative as they may be able to test directly for the influence of sensorimotor perturbations, without manipulating categorical information, thus manipulating the continuous representation of (i.e., frequency), while keeping its categorical representation of pitch (i.e., pitch class) relatively intact.

There is some evidence to suggest that feedback on pitch accuracy in the visual domain may improve vocal pitch accuracy in singing when it is given in real time. In a study investigating the effects of real-time visual feedback (Welch, Howard, & Rush, 1989), seven-year-old children ($n = 32$) were given a computerized device in school that provided visual feedback in real time

during singing. Results suggest that the inclusion of feedback from an additional (i.e., non-auditory) modality may encourage singing and improved singing development even without adult intervention. In addition to providing a route towards improving the motivational factors that might inhibit children from singing in the classroom, these results also highlight the role of temporal contingency of feedback on vocal pitch accuracy (i.e., near-real-time feedback may be more effective than offline feedback).

This temporal contingency is also fundamental to feedback in the auditory modality, which represents the bulk of work on vocal feedback. Studies have delayed the auditory feedback of one's own voice (Fairbanks & Guttman, 1958) and observed significant slowing of vocal articulation and linguistic rhythm under conditions of delayed auditory feedback in speaking and singing, as well as delays and disruptions of performance on various musical instruments (Harrington, 1988; Havlicek, 1968). In addition to delaying auditory feedback, studies have also manipulated auditory feedback in other ways, such as by pitch shifting or formant shifting the voice in near-real time. The typical behavioral response to pitch-shifted auditory feedback includes the opposing response, in which subjects compensate for the pitch-shift by adjusting their vocal pitch in the opposite direction. However, in some cases a following response was also observed, in which subjects follow the pitch-shift (Burnett et al., 1998). These following and opposing responses are highly sensitive to task instructions (Hain et al., 2000); furthermore, significant individual variability is also observed in the degree to which subjects show the following and opposing responses. In particular, singers were less perturbed by pitch-shifted feedback, especially when they were told to ignore the feedback and when the size of the pitch shift was large – and thus the singers were better able to de-identify with the feedback as a result of task instructions (Zarate, Wood, & Zatorre, 2010). Effects of voice feedback can elicit early and late pitch-shift responses: the early component (with a latency of ~100 ms) being relatively automatic and the late component (with longer latencies) being more conscious and dependent on task instructions (Burnett & Larson, 2002). Although much work remains to be done on individual differences in following versus opposing responses, the literature on pitch-shifted auditory feedback suggests that reflexive and relatively involuntary processing is dissociable from conscious control of one's own voice due to task instructions.

The ability to identify with one's own voice, in a manner that is independent of categorical perception, is

closely related to the vocal generosity effect (Hutchins, Roquet, & Peretz, 2012), where listeners are more likely to judge a voice as being in tune compared to other instruments. While the original explanations for this effect involve acoustic and cognitive (top-down) factors, it is also conceivable that what subjects might do, when confronted with pitches produced by the voice, is recruit their automatic stream as a mirror neuron mechanism, or motor simulation mechanism, in order to judge intonation, instead of the more canonical, categorically based perceptual route. Because the automatic stream is rapid and coarse-grained, the use of the automatic stream may account for the increased acceptance of fine-grained deviations in pitch as being acceptable matches of one's rapid and automatic estimation of "in-tune" pitch. Viewed in this light, the vocal generosity effect may again provide support for a dual-stream model consisting of a cognitively mediated pathway in contrast to an automatic, rapid, imitation-based stream of processing that is less dependent on conscious access and/or voluntary control. In addition to studies on the vocal generosity effect, other studies do show an advantage for imitating yourself for the general population (Hutchins & Peretz, 2012; Pfordresher & Mantell, 2014), providing further support for a self-imitation route that might be separate from a categorically based stream.

Disruption of Singing: Insights From Congenital Amusia and Poor Pitch Production

Another line of literature providing support for a dual-stream model comes from measuring singing proficiency in persons with musical disorders such as congenital amusia and poor pitch singing (Dalla Bella & Berkowska, 2009; Tremblay-Champoux, Dalla Bella, Phillips-Silver, Lebrun, & Peretz, 2010). Although musical ability is ubiquitous across cultures, an estimated 15% of the normal population (Cuddy, Balkwill, Peretz, & Holden, 2005) self-identifies as tone-deaf. Although true tone deafness is defined as the inability to hear differences in pitch, most people who self-report as tone deaf do not have congenital amusia, but actually present with difficulties in singing (Sloboda, Wise, & Peretz, 2005).

Tone-deafness, also known as congenital amusia, is characterized by poor pitch perception skills, but presents frequently as an inability to sing in tune. A unique dissociation arises when directly comparing pitch discrimination and pitch production in tone-deaf individuals, where the pitch discrimination abilities and pitch production acuity, as determined by psychophysical

staircase procedures as well as the method of fixed stimuli, could be different within the same tone-deaf individual. In striking cases this action-perception mismatch can present as intact direction of imitation of pairs of pitches, without consciously being able to label the direction of intervals between two pitches (Loui, Guenther, Mathys, & Schlaug, 2008). This striking dissociation shows correspondence between tone-deaf subjects and patients with visual agnosia reviewed earlier, with both cases suggesting separate pathways for perception and action (Goodale et al., 1991). Follow-up studies of this original finding (Williamson, Liu, Peryer, Grierson, & Stewart, 2012) suggest that these mismatches are not unidirectional, i.e., amusics could have poor perception thresholds and good production thresholds, or poor production thresholds and good perception thresholds, but the dissociation – or mismatch – between perceptual and production behavior persists in many individuals. This work is also supported by electrophysiological studies in which congenital amusics show earlier potentials elicited by pitch changes, suggesting some sensitivity to small changes in pitch, whereas later potentials show a disconnect with the earlier potentials among tone-deaf individuals, suggesting a lack of awareness along the perceptual-cognitive pathway (Peretz, Brattico, Jarvenpää, & Tervaniemi, 2009).

Voxel-based morphometry and cortical thickness studies on congenital amusia suggest that both frontal and temporal regions can be affected, albeit on both hemispheres of the brain starting as early as the level of the auditory cortex (Albouy et al., 2013; Hyde et al., 2007; Mandell, Schulze, & Schlaug, 2007). Diffusion tensor imaging work on white matter integrity in tone-deafness has shown that the arcuate fasciculus, in both hemispheres but especially on the right hemisphere, has a disrupted pattern of connectivity in tone-deaf individuals (Loui et al., 2009). The superior branch of the arcuate fasciculus, connecting the superior temporal gyrus and the inferior frontal gyrus, was correlated with the pitch discrimination threshold. In contrast, the inferior branch of the arcuate fasciculus, connecting the middle temporal gyrus to the inferior frontal gyrus, was correlated with the size of the mismatch between pitch discrimination and pitch production thresholds. These findings provide support for the proposed dual-stream model of auditory function within the arcuate fasciculus and its co-terminating fibers (possibly including the extreme capsule fiber system and other branches of the superior longitudinal fasciculus) where the superior branch enables fine-grained discrimination, whereas the inferior branch is responsible for matching the sound output to its target.

Both branches are necessary for accurate perception and production.

While the literature on congenital amusia gives helpful clues to the neural pathways involved in pitch perception and production, another important source of evidence arises from the research on poor-pitch singing (Bradshaw & McHenry, 2005; Dalla Bella, Berkowska, & Sowinski, 2011; Dalla Bella, Giguère, & Peretz, 2007; Pfordresher & Brown, 2007; Wise & Sloboda, 2008). In a study directly comparing pitch discrimination and pitch production among inaccurate singers, no correlation was found between pitch perception and production; however, poor pitch singers divided into two groups: those with intact pitch discrimination but inaccurate pitch production, and those who were inaccurate at both pitch discrimination and production (Bradshaw & McHenry, 2005). Viewed in the light of a possible dual-stream pathway, this single dissociation between perception and production suggests that some poor pitch singers could be lacking both routes while others could have a selective impairment of a single route that gives rise to pitch production via sensorimotor coupling, whereas the pathway that passes through conscious categorization could be preserved.

If tone-deafness and poor pitch singing are characterized by a lack of pitch awareness, reflected neuroanatomically by the more voluntary control-dependent one of the two streams of frontotemporal pathways, then an interesting follow-up question arises as to whether individuals with these musical disorders may be sensitive to perturbations in auditory feedback of their own voice, which can test for involuntary adjustment without the intervention of conscious access. This was directly tested by Hutchins and Peretz (2013), who showed that amusics still responded to pitch-shifted auditory feedback at 25 cents and 200 cents (Hutchins & Peretz, 2013). Furthermore, the size of the pitch shift responses was correlated with pitch matching accuracy rather than with performance on change detection with small pitch differences. In addition, recent data also showed an inverse correlation between note accuracy and perturbability – i.e., poor pitch singers were less affected by altered auditory feedback (Pfordresher & Beasley, 2014), again providing support for an unconscious pathway for pitch modulation in singing that is dissociable from the conscious route that is used for detection of fine-grained pitch changes.

The Role of Neuroplasticity in Expert Singing

If the frontal-temporal connections are critical in learning to sing, then trained musicians, especially

professional singers, who have had a lifetime of singing experience, must have a highly developed arcuate fasciculus. Indeed, when comparing the volume and white matter integrity of the arcuate fasciculus among groups of singers, non-singing instrumentalists, and nonmusicians, Halwani, Loui, Rueber, and Schlaug (2011) observed that singers had larger volume and higher fractional anisotropy (a measure of white matter integrity) in both dorsal and ventral streams of the left hemisphere compared to instrumentalists and non-singers. In contrast, both singers and instrumentalists possessed higher volume in both dorsal and ventral branches in the right hemisphere arcuate fasciculus compared to the nonmusicians. These results suggest that lifelong training in singing refines and enhances both streams in the dual-stream model: the pathway involved in fine-grained control as well as the one involved in coarse-grained, category-based control. Both types of control, involving dorsal and ventral arcuate fasciculus of left and right hemispheres, are required for proficient singing.

Conclusions

Results from singers and from tone-deaf individuals converge with a growing body of evidence suggesting that action and perception arise from dissociated neural pathways that are differentially affected in tone-deafness (Loui et al., 2008; Loui & Schlaug, 2009). These neural pathways of the auditory perception and action network, anatomically distinguishable as dorsal and ventral streams, may be responsible for automatic, category-based sound analysis and conscious access to perceptual information respectively, both of which are important in vocal communication. The pathways and endpoints of this distributed network may specifically parallel existing models positing multiple streams in speech, language, and auditory processing as reviewed above. Alternately, these pathways may be generalizable to provide support for

a more domain-general brain network that subsumes language, music, and auditory scene analysis in the human brain. Such a generalized model of sound perception and action would begin with the analysis of perceptual features such as pitch and rhythm, and subsequently entails the dual processing of dorsal and ventral streams for the linking of action and perception in the human brain. Selected actions can be voluntarily fine-tuned in the dorsal stream based on incoming acoustic input, and/or initiated directly from the ventral stream in a target-matching or category-based manner. Categories selected by the ventral stream may include lexical-semantic information in the linguistic domain (Glasser & Rilling, 2008), pitch class and melodic contour, or object identity in auditory scene analysis (Griffiths et al., 2007).

Taken together, the best available evidence supports an emerging model of dorsal and ventral pathways in auditory-motor integration that subserves singing as a special case of neural functions that include language, music, speech, and human interactions in the sonic perceptual environment. By showing an asymmetry between these dorsal and ventral pathways and linking it to the phenotypes of tone-deafness, absolute pitch, and singing training, the present findings support a dual-stream model for singing and raise interesting questions about the evolution of biological pathways underlying conscious and unconscious perception. Future experiments can compare and contrast different types of singing training and singing impairment, perhaps involving varying degrees of categorical learning, to test the present dual-stream neuroanatomical model of singing.

Author Note

Correspondence concerning this article should be addressed to Psyche Loui, Wesleyan University, 207 High Street, Middletown, CT 06459. E-mail: ploui@wesleyan.edu

References

- ALBOUY, P., MATTOU, J., BOUET, R., MABY, E., SANCHEZ, G., AGUERA, P. E., et al. (2013). Impaired pitch perception and memory in congenital amusia: The deficit starts in the auditory cortex. *Brain*, 136(Pt 5), 1639-1661.
- BRADSHAW, E., & MCHENRY, M. A. (2005). Pitch discrimination and pitch matching abilities of adults who sing inaccurately. *Journal of Voice*, 19(3), 431-439.
- BURNETT, T. A., FREEDLAND, M. B., LARSON, C. R., & HAIN, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *Journal of the Acoustical Society of America*, 103(6), 3153-3161.
- BURNETT, T. A., & LARSON, C. R. (2002). Early pitch-shift response is active in both steady and dynamic voice pitch control. *Journal of the Acoustical Society of America*, 112(3 Pt 1), 1058-1063.

- CAPILLA, A., BELIN, P., & GROSS, J. (2013). The early spatio-temporal correlates and task independence of cerebral voice processing studied with MEG. *Cerebral Cortex*, 23(6), 1388-1395.
- CATANI, M., ALLIN, M. P., HUSAIN, M., PUGLIESE, L., MESULAM, M. M., MURRAY, R. M., & JONES, D. K. (2007). Symmetries in human brain language pathways correlate with verbal recall. *Proceedings of the National Academy of Sciences of the United States of America*, 104(43), 17163-17168.
- CUDDY, L. L., BALKWILL, L. L., PERETZ, I., & HOLDEN, R. R. (2005). Musical difficulties are rare: A study of "tone deafness" among university students. *Annals of the New York Academy of Sciences*, 1060, 311-324.
- DALLA BELLA, S., & BERKOWSKA, M. (2009). Singing proficiency in the majority: Normality and "phenotypes" of poor singing. *Annals of the New York Academy of Sciences*, 1169, 99-107.
- DALLA BELLA, S., BERKOWSKA, M., & SOWINSKI, J. (2011). Disorders of pitch production in tone deafness. *Frontiers in Psychology*, 2.
- DALLA BELLA, S., GIGUÈRE, J.-F., & PERETZ, I. (2007). Singing proficiency in the general population. *Journal of the Acoustical Society of America*, 121(2), 1182.
- DEYOE, E. A., & VAN ESSEN, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*, 11(5), 219-226.
- FAIRBANKS, G., & GUTTMAN, N. (1958). Effects of delayed auditory feedback upon articulation. *Journal of Speech and Hearing Research*, 1(1), 12-22.
- FREY, S., CAMPBELL, J. S., PIKE, G. B., & PETRIDES, M. (2008). Dissociating the human language pathways with high angular resolution diffusion fiber tractography. *Journal of Neuroscience*, 28(45), 11435-11444.
- FRIEDERICI, A. D. (2009). Pathways to language: fiber tracts in the human brain. *Trends in Cognitive Sciences*, 13(4), 175-181.
- FRIEDERICI, A. D., BAHLMANN, J. R., HEIM, S., SCHUBOTZ, R. I., & ANWANDER, A. (2006). The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the United States of America*, 103(7), 2458-2463.
- GLASSER, M. F., & RILLING, J. K. (2008). DTI Tractography of the human brain's language pathways. *Cerebral Cortex*, 11, 2471-2482.
- GOLFINOPOULOS, E., TOURVILLE, J. A., & GUENTHER, F. H. (2010). The integration of large-scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage*, 52(3), 862-874.
- GOODALE, M. A., & MILNER, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, 15(1), 20-25.
- GOODALE, M. A., MILNER, A. D., JAKOBSON, L. S., & CAREY, D. P. (1991). Object awareness. *Nature*, 352(6332), 202.
- GRIFFITHS, T. D., KUMAR, S., WARREN, J. D., STEWART, L., STEPHAN, K. E., & FRISTON, K. J. (2007). Approaches to the cortical analysis of auditory objects. *Hearing Research*, 229(1-2), 46-53.
- GUENTHER, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders*, 39(5), 350-365.
- HAFKE, H. Z. (2008). Nonconscious control of fundamental voice frequency. *Journal of the Acoustical Society of America*, 123(1), 273-278.
- HAIN, T. C., BURNETT, T. A., KIRAN, S., LARSON, C. R., SINGH, S., & KENNEY, M. K. (2000). Instructing subjects to make a voluntary response reveals the presence of two components to the audio-vocal reflex. *Experimental Brain Research*, 130(2), 133-141.
- HALWANI, G. F., LOUI, P., RUEBER, T., & SCHLAUG, G. (2011). Effects of practice and experience on the arcuate fasciculus: Comparing singers, instrumentalists, and non-musicians. *Frontiers in Psychology*, 2.
- HARRINGTON, J. (1988). Stuttering, delayed auditory feedback, and linguistic rhythm. *Journal of Speech Hearing Research*, 31(1), 36-47.
- HAVLICEK, L. (1968). Effects of delayed auditory feedback on musical performance. *Journal of Research in Music Education*, 16, 308-318.
- HICKOK, G., & POEPEL, D. (2004). Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2), 67-99.
- HICKOK, G., & POEPEL, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393-402.
- HOUDE, J. F., & JORDAN, M. I. (2002). Sensorimotor adaptation of speech I: Compensation and adaptation. *Journal Speech Language and Hearing Research*, 45(2), 295-310.
- HUTCHINS, S., & PERETZ, I. (2012). Amusics can imitate what they cannot discriminate. *Brain and Language*.
- HUTCHINS, S., & PERETZ, I. (2013). Vocal pitch shift in congenital amusia (pitch deafness). *Brain and Language*, 125(1), 106-117.
- HUTCHINS, S., ROQUET, C., & PERETZ, I. (2012). The vocal generosity effect: How bad can your singing be? *Music Perception*, 30, 147-159.
- HYDE, K. L., LERCH, J. P., ZATORRE, R. J., GRIFFITHS, T. D., EVANS, A. C., & PERETZ, I. (2007). Cortical thickness in congenital amusia: when less is better than more. *Journal of Neuroscience*, 27(47), 13028-13032.
- KAAS, J. H., & HACKETT, T. A. (1999). 'What' and 'where' processing in auditory cortex. *Nature Neuroscience*, 2(12), 1045-1047.
- KAAS, J. H., & HACKETT, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22), 11793-11799.
- KATSEFF, S., HOUE, J., & JOHNSON, K. (2012). Partial compensation for altered auditory feedback: a tradeoff with somatosensory feedback? *Language and Speech*, 55(Pt 2), 295-308.

- LOUI, P., ALSOB, D., & SCHLAUG, G. (2009). Tone deafness: a new disconnection syndrome? *Journal of Neuroscience*, 29(33), 10215-10220.
- LOUI, P., GUENTHER, F. H., MATHYS, C., & SCHLAUG, G. (2008). Action-perception mismatch in tone-deafness. *Current Biology*, 18(8), R331-332.
- LOUI, P., LI, H. C., HOHMANN, A., & SCHLAUG, G. (2011a). Enhanced connectivity in absolute pitch musicians: A model of hyperconnectivity. *Journal of Cognitive Neuroscience* 23(4), 1015-1026.
- LOUI, P., LI, H. C., & SCHLAUG, G. (2011b). White matter integrity in right hemisphere predicts pitch-related grammar learning. *NeuroImage*, 55(2), 500-507.
- LOUI, P., & SCHLAUG, G. (2009). Investigating musical disorders with diffusion tensor imaging: A comparison of imaging parameters. *Annals of the New York Annual Academy of Sciences*, 1169, 121-125.
- LOUI, P., ZAMM, A., & SCHLAUG, G. (2012). Enhanced functional networks in absolute pitch. *NeuroImage*, 63(2), 632-640.
- MANDELL, J., SCHULZE, K., & SCHLAUG, G. (2007). Congenital amusia: An auditory-motor feedback disorder? *Restorative Neurology And Neuroscience*, 25(3-4), 323-334.
- MISHKIN, M., UNGERLEIDER, L. G., & MACKO, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences*, 6, 414-417.
- PATEL, R., NIZIOLEK, C., REILLY, K., & GUENTHER, F. H. (2011). Prosodic adaptations to pitch perturbation in running speech. *Journal of Speech Language and Hearing Research*, 54(4), 1051-1059.
- PATTERSON, D. K., VAN PETTEN, C., BEESON, P. M., RAPCSAK, S. Z., & PLANTE, E. (2014). Bidirectional iterative parcellation of diffusion weighted imaging data: Separating cortical regions connected by the arcuate fasciculus and extreme capsule. *NeuroImage*, 102, Part 2(0), 704-716.
- PERETZ, I., BRATTICO, E., JARVENPAA, M., & TERVANIEMI, M. (2009). The amusic brain: In tune, out of key, and unaware. *Brain*, 132(5), 1277-1286.
- PFORDRESHER, P., & BEASLEY, R. (2014). Making and monitoring errors based on altered auditory feedback. *Frontiers in Psychology*, 5.
- PFORDRESHER, P. Q., & BROWN, S. (2007). Poor-pitch singing in the absence of "tone deafness." *Music Perception*, 25, 95-115.
- PFORDRESHER, P. Q., & MANTELL, J. T. (2014). Singing with yourself: Evidence for an inverse modeling account of poor-pitch singing. *Cognitive Psychology*, 70, 31-57.
- POEPEL, D. (2003). The analysis of speech in different temporal integration windows: Cerebral lateralization as "asymmetric sampling in time." *Speech Communication*, 41, 245-255.
- RAUSCHECKER, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hearing Research*, 271(1-2), 16-25.
- RAUSCHECKER, J. P., & SCOTT, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12(6), 718-724.
- RAUSCHECKER, J. P., & TIAN, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 97(22), 11800-11806.
- RILLING, J. K., GLASSER, M. F., PREUSS, T. M., MA, X., ZHAO, T., HU, X., & BEHRENS, T. E. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience*, 11(4), 426-428.
- SLOBODA, J. A., WISE, K. J., & PERETZ, I. (2005). Quantifying tone deafness in the general population. *Annals of the New York Academy of Sciences*, 1060, 255-261.
- TOURVILLE, J. A., REILLY, K. J., & GUENTHER, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*, 39(3), 1429-1443.
- TREMBLAY-CHAMPOUX, A., DALLA BELLA, S., PHILLIPS-SILVER, J., LEBRUN, M. A., & PERETZ, I. (2010). Singing proficiency in congenital amusia: Imitation helps. *Cognitive Neuropsychology*, 27(6), 463-476.
- VON DER HEIDE, R. J., SKIPPER, L. M., KLOBUSICKY, E., & OLSON, I. R. (2013). Dissecting the uncinate fasciculus: Disorders, controversies and a hypothesis. *Brain*, 136(Pt 6), 1692-1707.
- WAKANA, S., JIANG, H., NAGAE-POETSCHER, L. M., VAN ZIJL, P. C., & MORI, S. (2004). Fiber tract-based atlas of human white matter anatomy. *Radiology*, 230(1), 77-87.
- WELCH, G. F., HOWARD, D. M., & RUSH, C. (1989). Real-time visual feedback in the development of vocal pitch accuracy in singing. *Psychology of Music*, 17(2), 146-157.
- WILLIAMSON, V. J., LIU, F., PERYER, G., GRIERSON, M., & STEWART, L. (2012). Perception and action de-coupling in congenital amusia: Sensitivity to task demands. *Neuropsychologia*, 50(1), 172-180.
- WISE, K. J., & SLOBODA, J. A. (2008). Establishing an empirical profile of self-defined "tone deafness": Perception, singing performance and self-assessment. *Musicae Scientiae*, 12(1), 3-26.
- ZARATE, J. M. (2013). The neural control of singing. *Frontiers in Human Neuroscience*, 7, 237.
- ZARATE, J. M., WOOD, S., & ZATORRE, R. J. (2010). Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers. *Neuropsychologia*, 48(2), 607-618.
- 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., BELIN, P., & PENHUNE, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends in Cognitive Science*, 6(1), 37-46.