Central Voice Production and Pathophysiology of Spasmodic Dysphonia

Niv Mor, MD; Kristina Simonyan, MD, PhD; Andrew Blitzer, MD DDS

Objective: Our ability to speak is complex, and the role of the central nervous system in controlling speech production is often overlooked in the field of otolaryngology. In this brief review, we present an integrated overview of speech production with a focus on the role of central nervous system. The role of central control of voice production is then further discussed in relation to the potential pathophysiology of spasmodic dysphonia (SD).

Data Sources: Peer-review articles on central laryngeal control and SD were identified from PUBMED search. Selected articles were augmented with designated relevant publications.

Review Methods: Publications that discussed central and peripheral nervous system control of voice production and the central pathophysiology of laryngeal dystonia were chosen.

Results: Our ability to speak is regulated by specialized complex mechanisms that coordinate high-level cortical processing, brainstem reflexes, peripheral nerves, muscles, and mucosal actions. Recent studies suggest that SD results from a primary central disturbance associated with dysfunction at our highest levels of central voice control. The efficacy of botulinum toxin in treating SD may not be limited solely to its local effect on laryngeal muscles and also may modulate the disorder at the level of the central nervous system.

Conclusion: Future therapeutic options that target the central nervous system may help modulate the underlying disorder in SD and allow clinicians to better understand the principal pathophysiology.

Key Words: Laryngeal motor cortex, phonation, voice, spasmodic dysphonia, laryngeal dystonia, botulinum toxin.

Level of Evidence: N/A

INTRODUCTION

Our ability to speak is regulated by a number of complex specialized mechanisms that coordinate high-level cortical processing, brainstem reflexes, and peripheral nerves. Although vocalization was mapped to the motor cortex by Penfield in 1930s, our current understanding of neural control of voice and speech production is based on studies conducted in the past 5 years. However, in the field of otolaryngology, the role of the central nervous system (CNS) in controlling speech production often has been overlooked. In this review, we present an overview of central control of voice production and discuss the potential neuropathophysiology of spasmodic dysphonia (SD). Although SD was characterized in 1980s and 1990s, much of what we currently know regarding the detailed clinical phenomenology, neural correlates, and genetics of SD is through contemporary studies conducted within the past decade. A better understanding SD has helped us appreciate the importance of CNS regulation in voice production.

Development

Voice production in humans can be voluntary as in speaking and singing, or involuntary as is occasionally observed in response to pain, fright, or emotions. Voluntary and involuntary voice production is coordinated under the control of brain stem, midbrain, and cortical structures. The intricate neural circuitry involved in human voice production gradually develops over time from initial involuntary shrieks and cries in infants to clearly articulated vocal communication later in life. Vocalization is not acquired through explicit instruction; rather, it is implicitly acquired through a gradual process of increased adaptation and development, resulting in more complex behaviors such as speaking and singing. As a child gradually acquires control over orofacial...
and laryngeal muscles, early signs of speech develop that often are heard as babbling. As development continues, speech motor control becomes increasingly more skilled, and voice onset and offset come to be timed to differentiate between different sounds.2

Voice, Speech, and Language

The distinction between voice, speech, and language is important. Voice usually is used for speech, and speech conveys meaning. Language involves the formulation of meaningful phrases in grammatically articular relationships. Examination of voice control without the confounding effects of language would more accurately characterize voice production without meaning. Nonlanguage voice production involves voice changes alone and does not require the use of lips, tongue, and jaw movements for speech or articulation. Although vocalization requires precise control of the larynx and utilizes skilled laryngeal motor patterns necessary for speech production, it does not necessarily convey language or meaning.2 Different brain levels control vocalizations of different degrees of complexity, and the CNS control over voice production can be perceived as somewhat hierarchical.3 As one moves up the hierarchical ladder, increasingly more complex vocalizations begin to incorporate voice with speech and language.

Central Control of Voice Production

The lowest level in this hierarchical system is under control of the reticular formation and phonatory sensory and motor nuclei within the brainstem, with motoneurons to the intrinsic laryngeal muscles located in the nucleus ambiguous and motoneurorns of extrinsic laryngeal muscles located near the hypoglossal nucleus3–5 (Fig. 1). This level of central control is responsible for production of innate vocalizations, which include nonverbal vocalization such as the cry or laugh of an infant. The structure of innate vocalizations genetically is preprogrammed.6 This means that nonverbal emotional vocalizations are not learned actions and are not under the control of the forebrain. For example, anencephalic infants with intact brain stems and no forebrain still are capable of vocal utterances and verbal reactions to painful stimuli.7

As children develop and become capable of learning and mimicking vocal utterances, innate vocalizations increasingly become more voluntary. At this stage of development, a cry can be produced without the presence of an emotional stimulus or can be suppressed despite the presence of discomfort. Although still part of the innate vocalization system, this level of vocal control is more advanced and requires input from higher brain regions such as the cingulate cortex (CC) and the periaqueductal gray (PAG) (Fig. 1). The CC and PAG are responsible for the control of emotional vocalizations, voice initiation, and modulation of its intensity.8 The PAG appears to act as a gateway between the CC and the brainstem, linking the external stimulus with the motivational vocal reactions. The role of the PAG and CC in voice production further can be understood in their absence. Destruction of the CC does not interfere with voice that is initiated in the PAG. Thus, the ability to speak and vocalize is preserved with the loss of emotional intonation. By contrast, destruction of the PAG abolishes all vocalizations that originate from the CC, resulting in mutism.8,9

The above innate emotional voice system differs from the cortically based system that supports the development of learned voice productions necessary for speech.2,9,10 This highest level of voice production is under the control of the speech motor cortex—including the laryngeal motor cortex (LMC) and orofacial motor cortex—which coordinate more than 100 muscles used in phonation, swallowing, and breathing (Fig. 1). The LMC is responsible for highly skilled learned laryngeal movements, such as speaking and singing. Almost all laryngeal muscles receive bilateral innervation from the left and right LMC; thus, patients with unilateral injury to the LMC still maintain the ability of voluntarily voice control.9 Neuroimaging and electrical stimulation studies have localized the LMC in humans to area 4 of the primary motor cortex (Fig. 2).11–16 Interestingly, the motocortical location of the larynx in nonhuman primates is different than in humans, where it is located far more rostrally and ventrally in area 6 of the premotor cortex. This difference likely represents an

Fig. 1. Hierarchical organization of central voice control depicting different interconnected levels of the voice control. The lowest level represented by the brain stem and spinal cord. Higher level of voice control is represented by the PAG and CC. The highest level is represented by the LMC. The dotted lines represent interconnections between regions. CC = cingulate cortex; LMC = laryngeal motor cortex; PAG = periaqueductal gray. [Color figure can be viewed in the online issue, which is available at www.larygoscope.com.]
evolutionary adaptation toward enhanced verbal communication in humans when compared to nonhuman primates. The LMC in area 4 of the motor cortex is thought to enable direct connection between the LMC and laryngeal motoneurons in the nucleus ambiguous of the brainstem. The importance of the LMC in human voice production is highlighted when juxtaposed with its more limited role in the nonhuman primates, which has markedly limited capacity for complex speech and voice production. Faster more direct coordination of complex laryngeal, orofacial, and respiratory movements in humans likely facilitates learning and voluntary vocal control for the purposes of speech and singing. In nonhuman primates, this connection is made indirectly, which may explain why these species are less capable of learning new vocal tasks. Bilateral lesions to the LMC in nonhuman primates result in a very limited deficit, without a profound effect on vocalizations. By contrast, bilateral lesion to the LMC in humans results in speech loss, preserving only the nonverbal emotional vocalizations such as grunting, crying, and laughing that are controlled through the cingulate-PAG circuitry.

**Somatosensory Feedback**

The exact receptors type and its role in laryngeal somatosensory feedback during speech are less well known and still debated. Some authors have suggested that stretch reflexes in the laryngeal muscles may provide proprioceptive feedback assisting in voice control. Thus far, no studies have demonstrated the physiological effect of stretched human laryngeal muscle. Recent findings suggest that spindle fibers only occur within the interarytenoid muscle and are sparse or absent in the thyroarytenoid, lateral cricoarytenoid, cricothyroid, and posterior crioarytenoid muscles. Furthermore, animal models show that sensory afferent fibers from the internal branch of the superior laryngeal nerve (iSLN) are more sensitive to mucosal deflection than to muscle stretch. Bhabu et al. demonstrated initiation of the laryngeal adductor response in human by an air stimulus to the laryngeal mucosa, supporting the role that mucosal mechanoreceptors provide dynamic sensory feedback to the central nervous system.

**Spasmodic Dysphonia**

The importance of the role of the CNS in controlling speech production can be further understood within the...
context of centrally derived speech disorders such as spasmodic dysphonia. Laryngeal dystonia is a clinical syndrome characterized by task-specific involuntary contractions of the internal laryngeal muscles. Respiratory laryngeal dystonia results in laryngeal contraction of varied degrees during breathing that usually disappears with sleep and spares patients' fluency during speech. Patients with singers dystonia have laryngeal hyperkinesias only while singing—not during conversational speech. Spasmodic dysphonia (SD) is a focal dystonia affecting fluency of voice during speech. It is the most commonly affected task of the laryngeal dystonias, which were believed to result from dysfunction at the level of the basal ganglia, although this view has recently been expanded to include the pathophysiological network to the cerebellum and sensorimotor cortical regions.

Although the exact pathophysiology of SD is unknown, recently structural alterations in brain organization were demonstrated in patients with SD, including focal reduction of axonal density and myelin along the corticobulbar/corticospinal tract. Functional magnetic resonance imaging (fMRI) identified brain abnormalities in patients with SD and demonstrated a greater extent of brain activation in the cortical brain regions responsible for the control of voice production during both symptomatic driven and asymptomatic tasks. The extent of activation within the subcortical structures (basal ganglia, thalamus, and cerebellum) also increased, but only during symptomatic speech; it was decreased during asymptomatic laryngeal tasks. These changes were noted both in patients with adductor and abductor SD and suggest that the primary disturbance in SD is associated with dysfunction of the sensorimotor cortex as well as the basal ganglia-thalamocortical circuitry (Fig. 3).

A recent study also found neurochemical alterations in the basal ganglia in patients with SD. Positron emission tomography with the radioligand $[^{11}C] \text{raclopride (RAC)}$ was used to explore striatal dopaminergic neurotransmission during symptomatic speech and was compared to healthy controls. Finger tapping was used as an internally controlled task and was designated as an asymptomatic task. Patients with SD had fewer available striatal dopamine D2/D3 receptors, as well as decreased levels of dopamine release during symptomatic speech. Interestingly, patients with SD demonstrated increased dopamine release during the asymptomatic tasks (finger tapping) when compared to healthy controls. It is possible that decreased dopaminergic transmission is responsible for the generation of symptoms in patients with SD, and the observed increased striatal dopaminergic function compensatory adaptation of the nigrostriatal dopaminergic system to decreased dopamine D2/D3 receptor availability. Also patients who were more symptomatic had greater RAC displacement, and those with longer duration of spasmodic dysphonia had decreased task-induced RAC displacement.

Neural abnormalities described in spasmodic dysphonia explain which brain regions and connections or neurochemical makeup are implicated in the
pathophysiology of this disorder. Further identification of these alterations in spasmodic dysphonia and other voice disorders not only can help the physician better understand the disorder itself but could simultaneously help define the contribution of specific brain regions in normal voice control.

**Treatment of Spasmodic Dysphonia**

The role of the CNS in voice production can be further elucidated in the context of the evolution of the treatment of SD. Although SD is now recognized as a CNS disorder, initial attempts to restore voice focused at alterations to the laryngeal framework, muscles or peripheral nerves.\(^{47}\) Initial studies reported an 85% to 90% success rate following recurrent laryngeal nerve (RLN) transection; however, a follow-up study showed that 64% of patients had return of pathologic voice quality.\(^{48-50}\) Previous explanations for the return of the disorder have been proposed, including reinnervation by proximal RLN axons.\(^{51}\) In an attempt to provide permanent selective denervation, Berke et al. performed selective laryngeal adduction denervation and reinnervation (SLAD/R).\(^{52}\) This procedure incorporates bilateral RLN transection to branches responsible for innervating only the adductor laryngeal muscles. To prevent unwanted reinnervation from the RLN and to maintain muscle tone, the cut branches were anastomosed to a branch of the ansa cervicalis. Initial results with SLAD/R were promising and showed improved success when compared to RLN transection alone. However, long-term results demonstrated the return of voice breaks in 26% and a persistent breathy voice quality in 30%.\(^{53}\)

Disappointing results with any surgical alterations, whether to the end-organ or the peripheral nerves, to provide long-term symptomatic relief likely are due to the failure of surgery to address the CNS. Surgery is permanent, and fixed alterations do not account for the possibility of CNS plasticity and adaption.\(^{54}\) In addition, RLN transection does not address the numerous interconnections within the larynx and interconnections between the superior laryngeal nerve SLN and RLN (Fig. 4\(^{55}\)). Theoretically, these interconnections could allow persistent CNS access to alter normal laryngeal muscles physiology.\(^{56-61}\)

**Botulinum Toxin**

In 1985, Blitzer found dramatic improvement of voice quality following direct injection of botulinum toxin (BoNT) to the affected laryngeal muscles.\(^{62}\) Botulinum toxin is a 150-kilodalton exotoxin produced from clostridium botulinum, the action of which is mediated through the cleavage of docking proteins, responsible for membrane fusion of presynaptic vesicles, and now is the gold standard for treatment for laryngeal dystonia. Cleavage of these docking proteins inhibits release of acetylcholine (Ach) at the neuromuscular junction and results in muscle weakness. Unlike surgery, BoNT continuously is metabolized, and the ever-changing effect does not allow for central adaptation. In addition to weakening laryngeal muscle activity, BoNT decreases the activation of muscle fibers directly through its effect on the intrafusal sensory fibers. However, local chemodenervation does not fully explain the clinical effects of BoNT. If BoNT interfered solely with muscle action and sensory fiber tone, then BoNT injections should have no effect on central efferent pathways. However, it is well known that unilateral BoNT injections reduce involuntary aberrant contractions to the contralateral untreated laryngeal muscle groups.\(^{63,64}\) A possible explanation for this finding is that modulation of the laryngeal muscles has an effect on the sensory feedback loop.

It has also been shown that the aberrant central activity at the primary sensorimotor cortex (areas 3,1,2) is normalized after peripheral BoNT injections, demonstrating BoNT’s effect on modulating the central nervous system.\(^{65}\) Although the exact mechanism is unknown, it is feasible that elements of BoTN are transmitted through the peripheral nerves in a retrograde fashion and modulate the central interneurons directly.\(^{66,67}\)
Whatever the cause, BoNT is effective and its activity does not appear to be limited to local alterations of the laryngeal muscles.

Future therapeutic options targeting the central nervous system in SD currently are being investigated. A recent survey showed that 55.9% of patients with SD had improvement in voice quality following ingestion of alcohol. This observation is likely related to alcohol’s effect on the CNS system (via GABA function). This finding led researchers to investigate the effects of drugs that could potentially improve SD voice symptoms by acting like alcohol on GABAergic transmission. Thus far, a metabolite of sodium oxybate, which behaves as a GABA receptor agonist, has found to improve SD symptoms in 82.2% of SD patients who had previously had symptomatic improvement following alcohol ingestion.

CONCLUSION

Our ability to produce purposeful vocalizations and speak fluently is regulated by a complex network of mechanisms originating at the level of the CNS. Central regulation in voice production and speech is crucial. Spasmodic dysphonia is a disorder of the CNS and an example of selectively disordered central voice regulation. Recent studies suggest that SD results from a primary central disturbance in the LMC and its circuitry. Botulinum toxin has shown great efficacy in treating patients’ symptoms. Injection of BoNT into the laryngeal muscles is not limited to its effect locally and also conveys an effect to the CNS. Nevertheless, it is important to highlight that BoNT is effective at only temporarily treating the symptoms related to SD and does not ultimately treat the underlying central disturbance. Future therapeutic options that target the central nervous system may help modulate the disorder and allow clinicians to better understand the pathophysiology of SD. Lastly, a better understanding of the types of receptors and the role they play in laryngeal somatosensory feedback during voice production and speech may provide us with additional understandings of the exact therapeutic role of BoNT both locally and centrally in treating the vocal symptoms related to SD.

BIBLIOGRAPHY


