2

Neural Structures Involved in Speech Production

As mentioned in the Introduction, producing even the simplest speech utterance involves a large number of cortical and subcortical regions of the brain. In this chapter we will provide an overview of these regions, with later chapters providing more detailed descriptions of their hypothesized functions within the neural control system for speech. This discussion will utilize standard terminology for relative anatomical locations; these terms and the corresponding anatomical directions are illustrated in figure 2.1.

2.1 The Primate Vocalization Circuit

Control of speech production builds on neural circuits used for other purposes by our evolutionary predecessors, most notably the production of learned voluntary vocalizations. The study of the neural mechanisms of speech thus naturally begins with a treatment of the neural systems underlying these vocalizations in our closest evolutionary relatives, nonhuman primates.

The nonhuman primate vocalization system has been more precisely characterized than the human system because of the availability of invasive techniques that are not suitable for use in humans, including single-unit electrophysiology, focal lesioning, and axonal tracers. Figure 2.2, adapted from Jürgens (2009), schematizes the brain regions and axonal tracts responsible for the production of learned vocalizations in primates, based largely on studies of the squirrel monkey. The remainder of this section follows the Jürgens model presented in this figure.

The *reticular formation* acts as the final convergence zone of projections from higher-level brain areas involved in vocalization and is partially responsible for coordinating movements of the various muscles involved in vocalization, including respiratory, laryngeal, and orofacial muscles. The coordinated commands are sent to the motoneuron pools in the brain stem that control these muscles. Activation of the reticular formation produces full coordinated vocalizations rather than isolated components of vocalization, supporting the assertion of a role for the reticular formation in coordination across muscle systems (Jürgens & Richter, 1986).

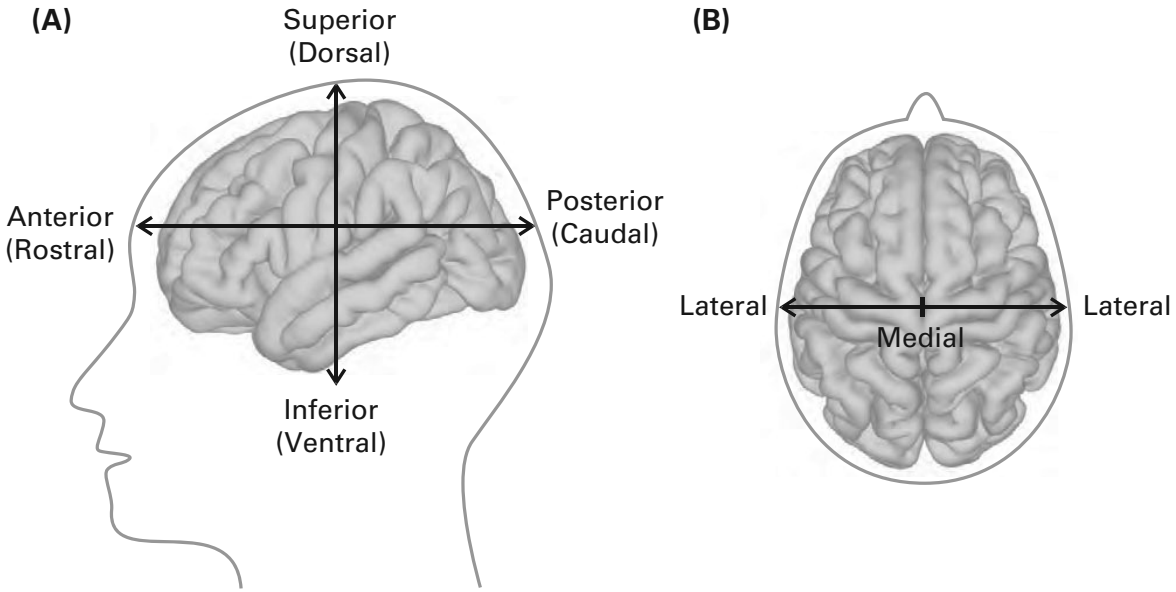

Figure 2.1
Anatomical terms of location. (A) Lateral view. (B) Superior view.

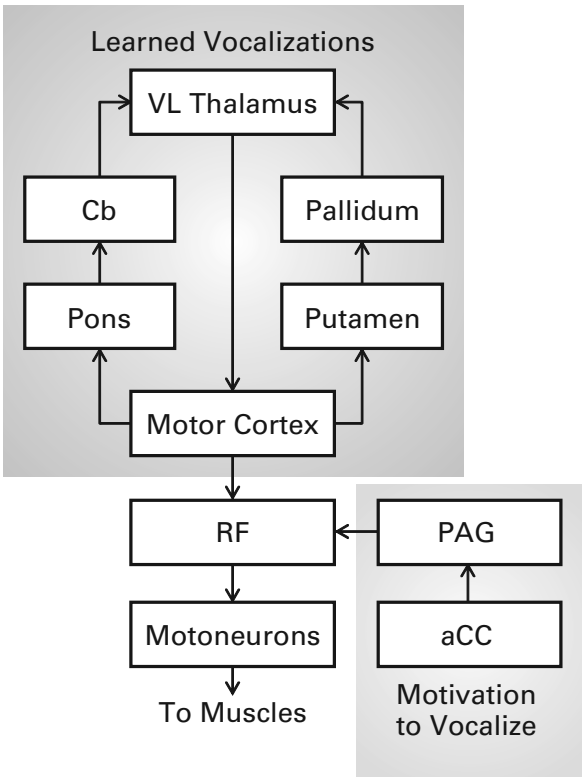

Figure 2.2
Primate vocalization model proposed by Jürgens (2009). aCC, anterior cingulate cortex; Cb, cerebellum; PAG, periaqueductal gray matter; RF, reticular formation; VL, ventral lateral nucleus.

The reticular formation receives projections from two distinct pathways from the cerebral cortex: a *limbic pathway* involving the *anterior cingulate cortex* (*aCC*) and *periaqueductal gray matter* (*PAG*) and a *motor cortical pathway* emanating from the primary motor cortex.

The limbic pathway is not heavily involved in motor coordination but instead serves a gating function that allows commands from the cerebral cortex to reach the motor periphery via the reticular formation. This pathway appears to be involved in controlling the intensity, but not the muscle patterning, of the vocalization. The gating signal from the limbic system is believed to represent the *motivation* or *readiness* to vocalize. In keeping with this characterization, neural activity in PAG correlates with global loudness but is not sensitive to specific acoustic patterns within the vocalization (Düsterhöft, Häusler, & Jürgens, 2004).

The motor cortical pathway is responsible for the production of learned vocalizations, including higher-level coordination of the muscles involved in vocalization (in concert with the lower-level coordination provided by the reticular formation). Signals in motor cortex are under the influence of two reentrant loops with subcortical structures: a loop through the *pontine nuclei* (or *pons*), *cerebellum*, and *thalamus* (*cortico-cerebellar loop*) and a loop through the *putamen*, *pallidum* (also called the *globus pallidus*), and *thalamus* (*cortico–basal ganglia loop*). The roles of these loops will be discussed within the context of the speech motor system in later sections.

The remainder of this chapter details the neural structures and pathways that make up the human speech motor system, starting from the motor periphery and moving upward in the nervous system to the cerebral cortex. These circuits include the primate vocalization circuit described by the Jürgens model as well as additional neural structures, particularly in the cerebral cortex.

2.2 Brain Stem Structures

Cranial Nerve Nuclei

The most peripheral neurons of the nervous system that are involved in speech are the *cranial nerve nuclei* located in the *medulla* and *pons*, which, along with the *midbrain*, constitute the *brain stem*. The twelve *cranial nerves* are typically identified by Roman numerals along with a name. Figure 2.3 schematizes the cranial nerves involved in speech, along with their associated cranial nuclei and the projections between these nuclei and the cerebral cortex. The cranial nuclei include sensory nuclei, illustrated on the left half of the brain stem in figure 2.3, and motor nuclei, illustrated on the right half of the brain stem.[1] Some nuclei and nerves have both motor and sensory components while others are primarily motor or primarily sensory. The cranial nerves most important for speech are nerves V and VII–XII; these nerves and their associated nuclei are discussed in the following paragraphs.

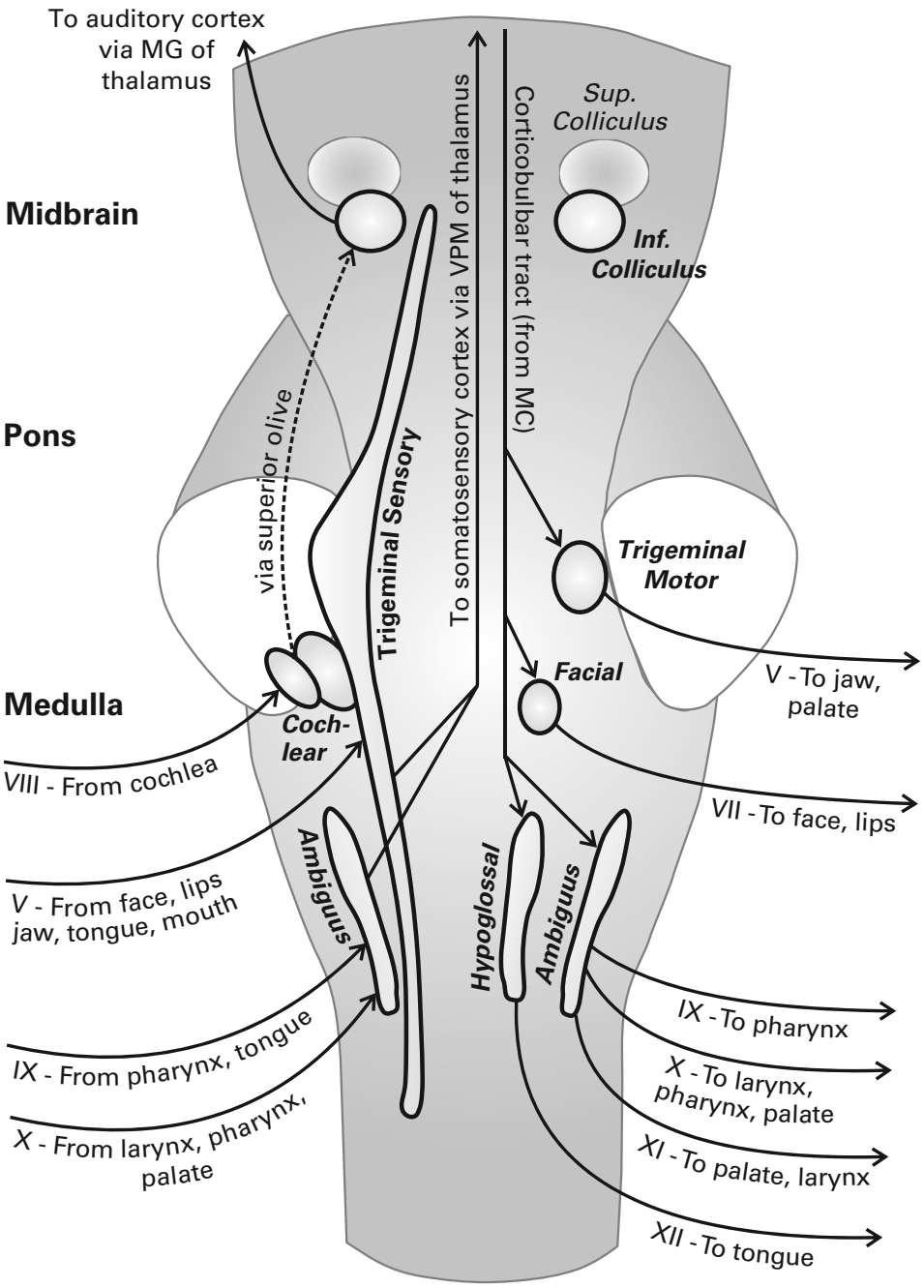

Figure 2.3
Dorsal view of the brain stem (medulla, pons, and midbrain) with cerebellum removed (white patches represent cerebellar peduncles) to illustrate the cranial nerves and associated nuclei involved in speech, including direct connections to cerebral cortex and to the periphery. Connections to the cranial nerve nuclei from the reticular formation and red nucleus are not shown. Sensory pathways are shown on the left; motor projections to muscles on the right. Nerve locations are approximate. Roman numerals indicate cranial nerve number. Inf., inferior; MC, motor cortex; MG, medial geniculate nucleus; Sup., superior; VPM, ventral posterior medial nucleus.

The *Vth nerve* (also called the *trigeminal nerve*) carries tactile and proprioceptive information from the oral cavity (except the posterior third of the tongue), nasal cavity, face, lips, jaw, and pharynx to the *trigeminal sensory nucleus* in the medulla. The motor portion of the trigeminal nerve originates in the *trigeminal motor nucleus* and innervates jaw and soft palate muscles. Somatosensory information from the trigeminal and other somatosensory cranial nerve nuclei projects to the primary somatosensory cortex via the *ventral posterior medial* (*VPM*) *nucleus* of the thalamus. Damage to the trigeminal nerve can cause numbness on the side of the lesion, difficulty chewing, and loss of muscle tone in the floor of the mouth (Zemlin, 1998). Unilateral lesions of the trigeminal nerve have minimal effects on speech, but bilateral damage is devastating for speech since the jaw hangs open (Duffy, 1995).

The *VIIth nerve*, or *facial nerve*, is a combined sensory and motor nerve, but only the motor aspect appears to be heavily involved in speech (Duffy, 1995). This nerve innervates muscles of the face and lips, and unilateral damage can result in paralysis of the facial muscles on the side of the damage, fasciculations (small, local quivering of muscle fibers), and/or atrophy of facial muscles resulting in facial asymmetry (Zemlin, 1998).

The *VIIIth nerve* (also called the *cochleovestibular nerve*, *cochlear nerve*, or *auditory nerve*) is a sensory nerve that carries auditory information from the cochlea to the *dorsal* and *ventral cochlear nuclei* in the medulla. From there, this information projects to the *superior olive* in both hemispheres and continues up to the *inferior colliculus* via a pathway called the *lateral lemniscus*. Auditory information then passes through the *medial geniculate* (*MG*) *nucleus* of the thalamus before arriving at the primary auditory cortex. Unilateral lesions of the VIIIth nerve typically result in partial to full deafness in the ipsilateral ear and/or *tinnitus* (or ringing in the ear), possibly with some facial pain or numbness (Zemlin, 1998). Bilateral lesions can result in total deafness.

The *nucleus ambiguus* in the medulla gives rise to several cranial nerves involved in speech. The *IXth nerve*, or *glossopharyngeal nerve*, innervates the stylopharyngeal muscle of the pharynx and receives somatosensory information from the pharynx and tongue. Damage to this nerve results in difficulty swallowing and loss of sensation and taste in the posterior third of the tongue (Zemlin, 1998). The IXth nerve is also thought to play a role in the gag reflex. Its role in speech motor control appears to be minor. The *Xth nerve*, or *vagus nerve*, innervates muscles in the larynx, pharynx, and soft palate and receives somatosensory information from these same areas. The vagus nerve plays a central role in speech production, and lesions to this nerve can result in severe voicing and swallowing abnormalities as well as weakness in the soft palate, pharynx, and larynx. The *XIth nerve*, or *accessory nerve*, is a motor nerve that innervates muscles of the soft palate and larynx and is intermingled with the vagus nerve. Lesions of the accessory nerve can result in difficulties with head and shoulder movements and cause a variety of voicing problems.

The *XIIth nerve*, or *hypoglossal nerve*, is primarily a motor nerve[2] (Zemlin, 1998) that originates in the *hypoglossal nucleus* and innervates almost all of the muscles of the tongue.

Damage to the hypoglossal nucleus can cause paralysis, weakness, or fasciculations of the tongue on the side of the lesion.

The cranial motor nuclei contain *lower motor neurons* which connect directly with muscles and thus control the contractile state of the muscle. These neurons are often distinguished from *upper motor neurons*, which are primarily located in the motor cortex[3] and can only affect the muscles via the lower motor neurons. In later chapters we will use the terms *motor neuron* or *motoneuron* to refer to the lower motor neurons, and we will refrain from using the somewhat ambiguous term *upper motor neuron* in favor of more precise terms. Furthermore, we will not distinguish between *alpha* and *gamma motor neuron* types, nor will we explore the considerable complexity of the interactions between interneurons, motoneurons, and sensory afferents at the level of the spinal cord and cranial nerve nuclei (see Brooks, 1986, and Duffy, 1995, for treatments of these topics).

Since there are multiple pathways by which activity in the lower motor neurons can be affected (as described below), these neurons are sometimes referred to as the *final common pathway* for neural signals traveling to the muscles. Because of this organization, damage to the lower motor neurons affects all types of movements (voluntary, reflex, or automatic) of the associated musculature.

Reticular Formation and Red Nucleus

As indicated in figure 2.4, descending projections from the cerebral cortex (especially motor cortex) reach the cranial nerve nuclei via direct projections as well as indirectly via subcortical nuclei in the midbrain and brain stem. The *direct pathway* is a phylogenetically newer pathway[4] that involves the *corticobulbar tract*, which, along with the *corticospinal tract*, forms the descending *pyramidal system*.[5] The corticospinal tract is primarily involved in controlling the limbs and trunk. Its only role in speech concerns respiratory function, which will not be covered in detail here; see Barlow (1999) for a treatment of the neural organization of the respiratory system.

The direct pathway plays an essential role in the control of speech. For the speech articulators, these projections are largely bilateral (i.e., each motor cortex cell projects to lower motor neurons on both sides of the body), with the exception of projections to the facial and hypoglossal nuclei, which are mostly to the contralateral side. As a result, unilateral damage to the corticobulbar tract in humans usually has relatively little effect on speech, with only some weakness of the tongue and lower face on the side contralateral to the damage, whereas bilateral damage can have a devastating effect on speech (Duffy, 1995). The pyramidal projections for most other motor systems are primarily to motor neurons on the contralateral side of the body, and thus unilateral damage to the pyramidal tract causes severe weakness or paralysis of the contralateral side in these systems.

The *indirect pathway* refers to projections from cortex that involve intermediate nuclei, specifically the *red nucleus*, located in the midbrain, and the *reticular formation*, which spans the midbrain and medulla (Darley, Aronson, & Brown, 1975; Duffy, 1995). These

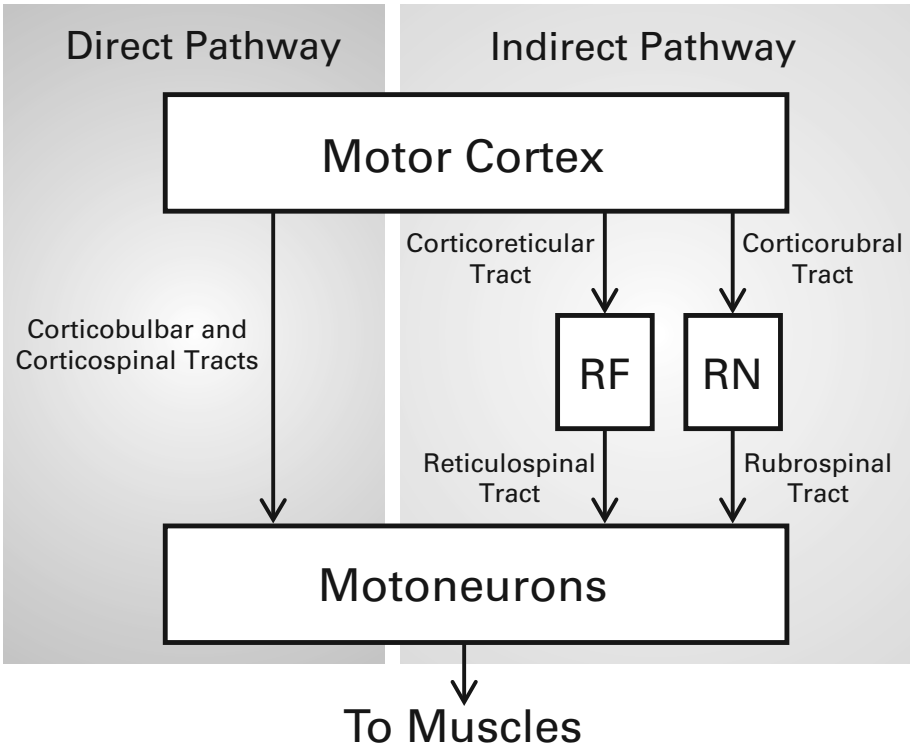

Figure 2.4
Direct and indirect pathways from the cerebral cortex to motoneurons in the cranial nerve nuclei and spinal cord. The direct pathway is formed by the corticobulbar and corticospinal tracts. The indirect pathway has two components: one passing through the reticular formation (RF) via the corticoreticular and reticulospinal tracts and the other passing through the red nucleus (RN) via the corticorubral and rubrospinal tracts.

nuclei receive cortical input via the *corticorubral tract* and *corticoreticular tract*, respectively, as well as input from the cerebellum and thalamus. They project to the motor periphery via the *reticulospinal tract* (which includes projections to the cranial nerve nuclei) and the *rubrospinal tract*, respectively. The indirect pathway's role in speech is poorly understood and appears to be more modulatory than control-oriented, with lesions to the indirect pathway generally affecting muscle tone and reflexes (Duffy, 1995) rather than the patterning of speech movements. However, Jürgens (2009) posits that the reticular formation is responsible for generation of innate vocal patterns, including nonverbal emotional vocal utterances of humans, and it may play a role in coordinating vocal fold vibrations during speech. In chapter 9 we speculate that the reticular formation may also play a role in affective prosody during speech.

It should be noted that the terms *direct* and *indirect pathway* have different meanings when discussed with regard to basal ganglia or cerebellar function, as we will see below. When not obvious from context, we will use care to specify which usage of the terms is intended throughout this book.

Periaqueductal Gray Matter

Less is understood about the PAG in humans than in monkeys. This midbrain structure, which surrounds the cerebral aqueduct, has been known to play a role in vocalization since Brown (1915) demonstrated that stimulation of PAG produced laughter in chimpanzees (for reviews, see Behbehani, 1995; Larson, 2004; Jürgens, 2009). In the Jürgens (2009) primate vocalization model (figure 2.2), PAG plays two key roles: (1) it is responsible for generating motor commands for innate vocalizations such as laughing and crying, and (2) it acts as a gating signal that modulates, rather than generates, the precisely timed motor commands of learned vocalizations which arrive from the motor cortex. This modulation may occur at the reticular formation, where the PAG and motor cortical signals come together. The relatively sparse human literature on PAG involvement in learned vocalizations appears to be compatible with the role it plays in the primate model of Jürgens; that is, as a modulator of descending commands from motor cortical areas rather than a generator of the detailed motor commands needed to coordinate the articulators during speech. However, PAG may coordinate some aspects of learned vocalizations, such as respiratory and laryngeal interactions (Larson, 2004), in concert with the commands generated from motor cortex. The view of PAG as a gating signal is compatible with the observation that bilateral damage to the PAG can lead to complete loss of voluntary vocalization, or *mutism*, in humans (Esposito et al., 1999) as well as animals (Behbehani, 1995; Jürgens, 2009; Larson, 2004) since bilateral damage to PAG would eliminate the gating signal that is needed to read out motor commands for learned vocalization. PAG neural activities also correlate with overall intensity of vocalization but not to acoustic details of the utterance. The notion of PAG as a gating signal for learned vocalizations is not completely settled, however, as studies in squirrel monkeys have shown that bilateral PAG lesions do not significantly affect signals from motor cortex to the laryngeal musculature (Jürgens, 2009), seemingly arguing against a role for PAG in gating of motor cortex commands for learned vocalizations.

2.3 Cerebellum, Basal Ganglia, and Thalamus

Although they are crucial for activating articulatory muscles, the subcortical regions discussed above are not thought to play a major role in the neural computations underlying the finely timed movements that constitute fluent speech. In this section we will treat subcortical structures that have a major influence on the descending motor commands via reentrant loops with the cerebral cortex: the thalamus, cerebellum, and basal ganglia. These structures are schematized in figure 2.5.

Thalamus

The thalamus is situated above the midbrain and below the basal ganglia, which in turn lie below cerebral cortex. In addition to playing roles in arousal and consciousness, the

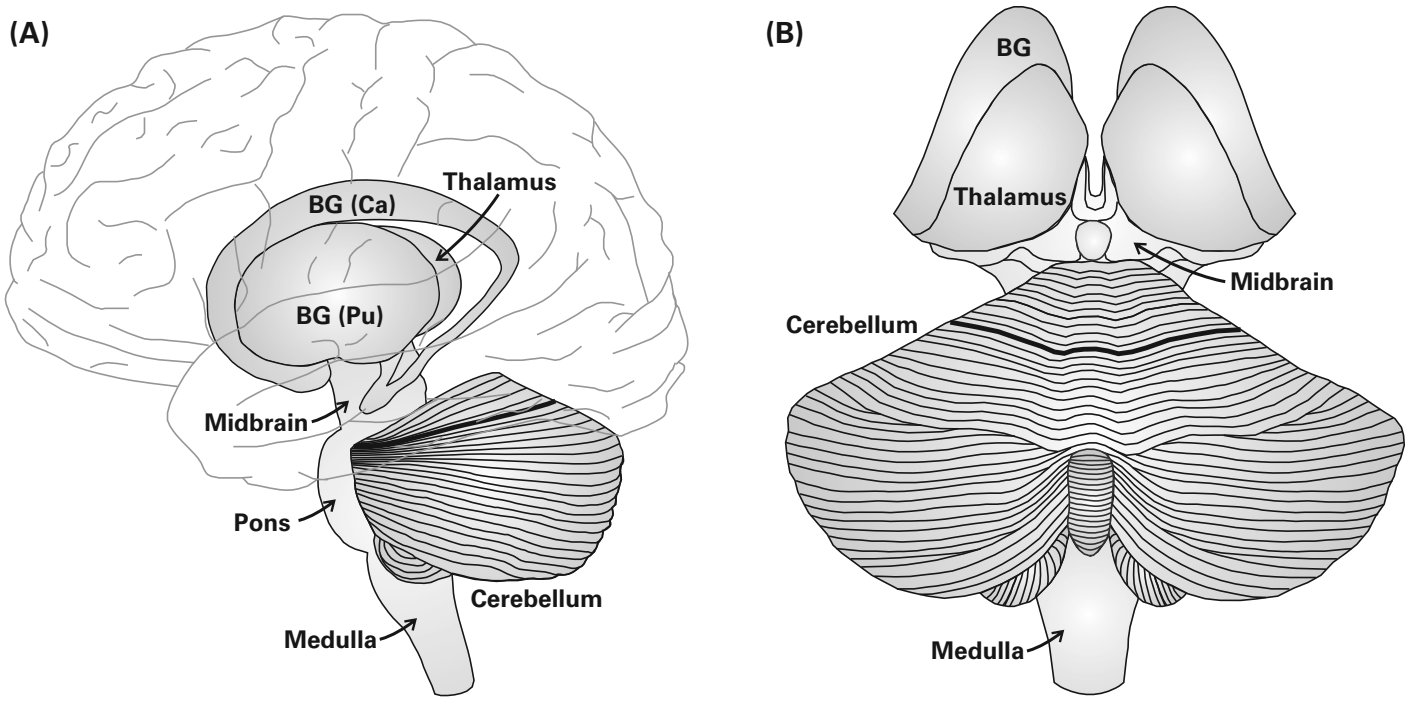

Figure 2.5
(A) Sagittal view showing the location of subcortical structures relative to the cerebral cortex. (B) Dorsal view of subcortical structures with cerebral cortex removed. Bold lines indicate location of the primary fissure of the cerebellum, which separates the anterior and posterior lobes of the cerebellar cortex. BG, basal ganglia; Ca, caudate nucleus; Pu, putamen.

thalamus acts as a massive relay station[6] for information entering or leaving the cerebral cortex. This includes sensory information from the periphery, motor information to the periphery, and outputs of the basal ganglia and cerebellum. The thalamus is divided into a number of distinct nuclei that are connected with different regions of the cerebral cortex. These cortico-thalamic connections are generally bidirectional; that is, each thalamic nucleus sends axons to a particular region of cortex, and that cortical region sends axons back to that same thalamic nucleus.

Electrical stimulation studies have implicated the thalamus, particularly the left *ventral lateral* (*VL*) *nucleus*, in speech motor control (see Johnson & Ojemann, 2000, for a review). Schaltenbrand (1975) reported that stimulation of VL in the language-dominant hemisphere can give rise to compulsory speech (monosyllabic yells and exclamations). Other effects of thalamic stimulation noted by Schaltenbrand (1975) include increasing or decreasing the loudness or rate of speech and *speech arrest*, or sudden loss of ability to speak. Mateer (1978) noted increased utterance duration with slurring and other articulatory distortions when the left VL was stimulated but not the right.

Four thalamic nuclei play important roles in speech motor control; these nuclei and their cortical connection zones are highlighted in figure 2.6. VPM projects to the ventral portion

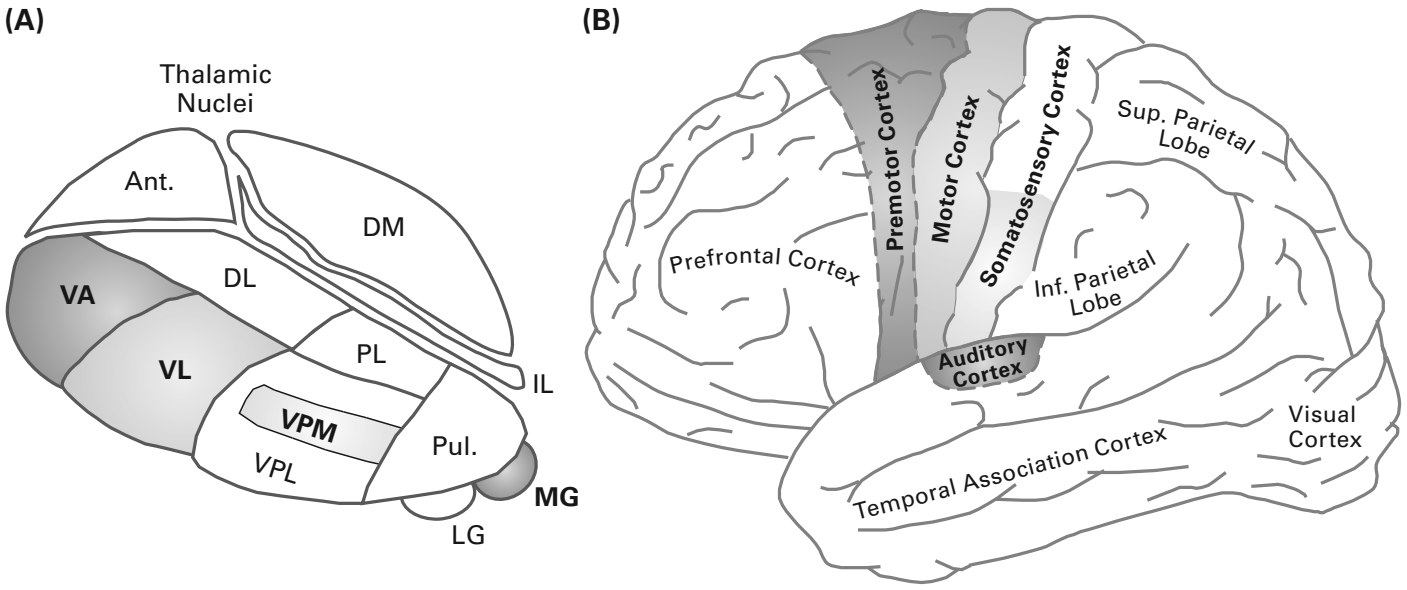

Figure 2.6
(A) Schematic of the thalamus with the nuclei involved in speech motor control highlighted. (B) Cortical projection targets of the thalamic nuclei involved in speech motor control, color coded according to the corresponding thalamic nucleus in panel A. Ant., anterior nucleus; DL, dorsal lateral nucleus; DM, dorsal medial nucleus; IL, intralaminar nucleus; inf., inferior; LG, lateral geniculate nucleus; MG, medial geniculate nucleus; PL, posterior lateral nucleus; Pul., pulvinar; sup., superior; VA, ventral anterior nucleus; VL, ventral lateral nucleus; VPL, ventral posterior lateral nucleus; VPM, ventral posterior medial nucleus.

of the somatosensory cortex, which contains the representation of the speech articulators. MG relays information from auditory brain stem structures to auditory cortex. VL is heavily connected with the primary motor cortex, consistent with the fact that VL is the thalamic nucleus that most frequently affects ongoing articulation when electrically stimulated. The *ventral anterior* (*VA*) *nucleus* is connected with premotor cortex. VL and VA form part of two reentrant loops with the cerebral cortex. The cortico–basal ganglia loop involves projections from the cortex to the basal ganglia, then to the thalamus, and then back to cerebral cortex. The cortico-cerebellar loop involves projections from cortex to the pons, then to the cerebellum, then to the thalamus, then back to cerebral cortex. Within VL and VA, neurons connected with the cerebellum are largely segregated from those connected with the basal ganglia (Sakai, Inase, & Tanji, 2002). These loops will be discussed briefly in the following subsections, and further detail of their involvement in speech production will be provided in later chapters.

Basal Ganglia

The basal ganglia lie beneath the cerebral cortex and are heavily interconnected with the frontal cortex via multiple cortico–basal ganglia loops (Alexander, DeLong, & Strick, 1986; Alexander & Crutcher, 1990; Middleton & Strick, 2000). The architecture of the basal ganglia make them suitable for selectively enabling one output from a set of

competing alternatives (Mink & Thach, 1993; Mink, 1996; Kropotov & Etlinger, 1999), a property evident in several computational models of basal ganglia function (e.g., Redgrave, Prescott, & Gurney, 1999; Brown, Bullock, & Grossberg, 2004; Prescott et al., 2006). Damage or electrical stimulation to the basal ganglia can cause several disturbances in speech. For example, damage to the basal ganglia can result in inaccuracies in articulation (Pickett et al., 1998), and electrical stimulation of the basal ganglia in the language-dominant hemisphere can evoke word production and cause other speech disturbances, including speech arrest (Van Buren, 1963; Gil Robles et al., 2005).

The basal ganglia consist of four distinct nuclei—the *striatum*, *globus pallidus*, *substantia nigra*, and *subthalamic nucleus* (*ST*)—which in turn contain distinct subdivisions. The striatum is the primary input recipient for the basal ganglia. It is typically separated into three subdivisions: the *caudate nucleus* (or simply *caudate*), *putamen*, and *ventral striatum* (which includes the *nucleus accumbens*). The caudate and putamen are similar in terms of microstructure and function; they are considered separate entities anatomically because they are separated by the *internal capsule*, a major white matter projection between the cerebral cortex and brain stem. The striatum receives projections from a vast expanse of the cerebral cortex. Roughly speaking, prefrontal cortical areas project mostly to the caudate and sensorimotor cortical areas to the putamen. The ventral striatum is heavily interconnected with the *limbic system*, which plays important roles in emotion, motivation, and memory. Functions involving the ventral striatum include olfaction and reward processing; it does not appear to play a significant role in speech motor control.

The globus pallidus receives input from the striatum and subthalamic nucleus. There are two distinct subregions of the globus pallidus: an *internal segment* (*GPi*) and an *external segment* (*GPe*). GPi contains many of the output neurons of the basal ganglia, which project to the thalamus. In chapter 7 we will discuss the different functional roles played by GPi and GPe in the cortico–basal ganglia loop.

The substantia nigra contains two functionally distinct subregions: the *pars compacta* (*SNc*) and the *pars reticulata* (*SNr*). SNr plays a functionally similar role to GPi whereas SNc consists of *dopaminergic neurons* that supply the neurotransmitter dopamine to the striatum via the *nigrostriatal pathway*. Among other roles, these dopaminergic projections are important for the learning of motor behaviors, acting as a sort of "teaching signal" that strengthens rewarding actions. This topic will be addressed in more detail in chapter 7.

ST is connected to both GPi and GPe, and it is the source of the only excitatory projections of the basal ganglia, which impinge on the globus pallidus. ST also receives input from the cerebral cortex, particularly the motor and premotor cortices (Zemlin, 1998).

The cortico–basal ganglia loop involves a substantial "funneling" or fan-in of information from cortical sources to basal ganglia outputs. For the rat, Zheng and Wilson (2002) estimate the number of neurons in cortex that project to the striatum to be about 10 times the number of striatal neurons that receive these projections. Furthermore, there are

approximately 100 times as many striatal neurons as basal ganglia output neurons in the GPi and SNr (Oorschot, 1996), indicating a ratio of basal ganglia afferents to efferents of roughly 1,000:1, with tens of millions of cortical neurons channeled to tens of thousands of basal ganglia output neurons. Despite this massive funneling, distinct information channels are maintained (Middleton & Strick, 2000). Brown, Bullock, and Grossberg (2004) hypothesize that there is a substantial fan-out of projections from the basal ganglia output channels back to cerebral cortex via the thalamus. Given the small number of basal ganglia output channels, it is highly unlikely that the basal ganglia are responsible for generating the precise motor commands needed for skilled movement. Instead the structure of the cortico–basal ganglia loops seems suited to choosing between alternative motor programs and sending *gating signals* to activate the cortical neurons responsible for carrying out the chosen motor program and inhibit cortical neurons related to competing motor programs (e.g., Mink, 1996; Brown, Bullock, & Grossberg, 2004).

The most important cortico–basal ganglia loop for speech motor control is the *motor circuit*. This circuit, illustrated in figure 2.7, includes several cortical areas that will be described in later sections, including primary motor cortex, premotor cortex, primary somatosensory cortex, and the supplementary motor area, as well as the putamen, globus pallidus, and VL nucleus of the thalamus (Alexander, DeLong, & Strick, 1986). Neuroimaging reveals that the components of the motor circuit are active for speech tasks as simple as production of a single syllable (Ghosh, Tourville, & Guenther, 2008). As discussed further with regard to the supplementary motor area below, this circuit is likely involved in the initiation of speech motor programs but less involved in choosing the precise muscle patterns that make up the motor programs. The motor circuit will be discussed in further detail in chapter 7.

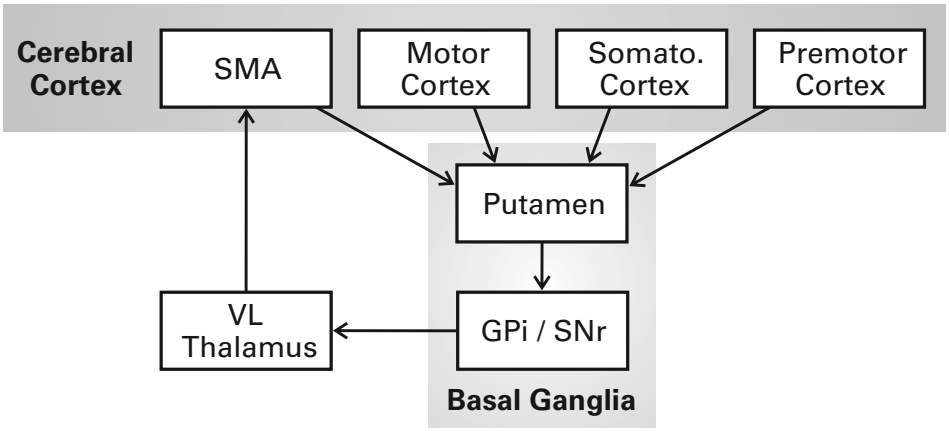

Figure 2.7
The cortico-basal ganglia motor circuit as originally proposed by Alexander, DeLong, and Strick (1986). GPi, internal segment of the globus pallidus; SMA, supplementary motor area; SNr, substantia nigra pars reticulata; Somato., somatosensory; VL, ventral lateral nucleus.