5 Explain the uses of the notion of markedness in linguistic description.
6 What is naturalness in phonological description? Can some languages be said to have more natural phonological systems than others?
7 What are the principal differences between a conventional phonemic description of a language and a generative description?
8 What are the reasons for using formal notation in phonological description?

6 The Anatomy and Physiology of Speech Production

This chapter provides a comprehensive anatomical background to the book's account of speech sounds. The first two sections set the scene for a technical account using the conventions of anatomical description (6.1 and 6.2).

The bulk of the chapter reviews the various organs of speech in a logical order, moving from the broad underlying structures and functions of the nervous system and respiratory system to the details of specific articulators such as tongue and lips. Given the complex functions of the larynx in speech, the section dealing with the larynx is followed by a separate section on how the larynx functions in phonation. The sections are:

- the nervous system (6.3)
- the respiratory system (6.4)
- the larynx (6.5)
- phonation (6.6)
- the pharynx (6.7)
- the velum and the nasal cavity (6.8)
- the oral cavity (6.9)
- the tongue (6.10)
- the lips (6.11)
- the mandible (6.12).

6.1 Introduction

In chapter 2 we outlined the speech production process from a functional perspective with sufficient detail to allow us to describe the speech sounds of language, but deliberately avoiding much discussion of the underlying technical detail. In this chapter we now provide a more technical examination of the anatomical and physiological processes of speech production. Some readers may choose to skip this and the ensuing chapter on speech acoustics, but for others, and especially those whose interests lie in experimental phonology and phonetics, speech and hearing science, communication disorders, cognitive science, artificial intelligence and speech technology, these two chapters provide an essential foundation. Moreover, these more physical and empirical perspectives on the subject have been basic to the development of some of the more theoretical and abstract units and categories of phonological description and analysis.

6.2 Conventions of anatomical description

We shall confine technical detail to essentials, but an understanding of some basic conventions of anatomical description is necessary. Figure 6.2.1 shows the basic division of the body into three planes (for sections through the body) and the five basic aspects from which anatomical features are viewed.

The following terms will be found in many anatomical descriptions:

CRANIAL	nearest or towards the brain
CAUDAL	away from the head
PROXIMAL	near the source of attachment
DISTAL	away from the source of attachment
AFFERENT	conducting towards the brain or spinal cord
EFFERENT	conducting from a central to a peripheral region
SUPERFICIAL	towards the surface
DEEP	away from the surface
PROCESS	projection or elevation on a bone

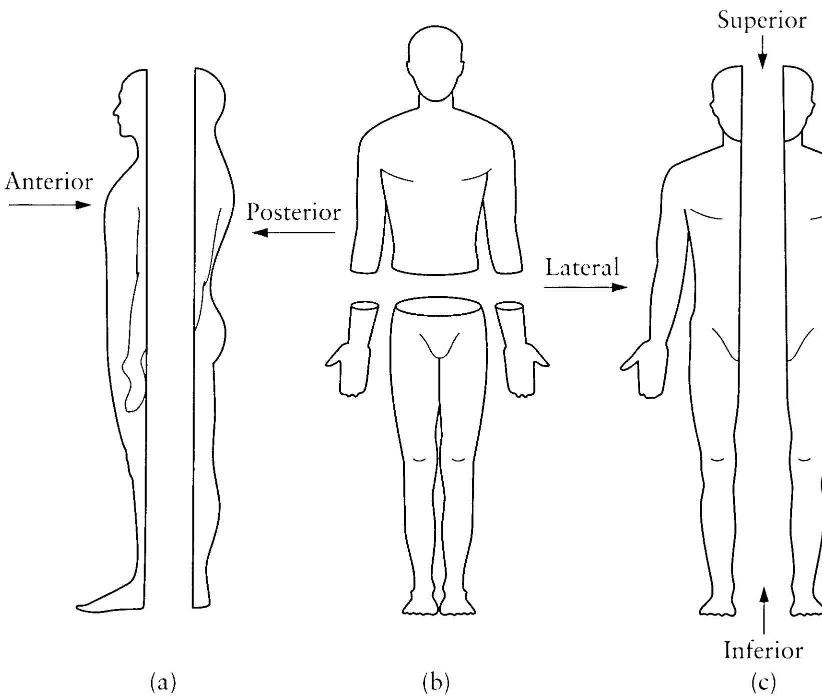


Figure 6.2.1 Sections and viewing aspects in human anatomy: (a) coronal; (b) transverse; (c) sagittal
Adapted from: Zemlin 1968, p. 9.

The organs of speech are all bodily structures composed of a variety of tissue types (such as bone, cartilage and skin) which are specific to their biological (rather than linguistic) function. Such tissue is in turn composed of cells, and the inner structure of these need not concern us further here. Bodily organs are generally grouped into systems which have particular functions in the life of the organism. These include the respiratory system, the digestive system, the reproductive system and so on. While it can be argued that the organs of speech form a system, they do not contribute to life support in the same way as other systems, and they are generally not thought of as performing their primary biological function when they are used in speech production.

In terms of bodily structure, humans are vertebrates – they have a backbone or vertebral column and their bodily structure is a mirror image either side of the backbone. Many parts of the body (such as muscles) are therefore PAIRED.

There are several ways of classifying the various tissues that make up the organs of the body. A common categorization recognizes five basic types, the first of which is of little relevance in the study of speech organs but is included here for completeness:

1. EPITHELIUM is the technical term for the layers of cells that constitute the outer skin of the body and the various membranous linings inside the body, including the so-called mucous membranes that are capable of secretion.
2. SCLEROUS or SKELETAL TISSUE refers to the dense tissues that provide the relatively rigid structure of the body. There are two subtypes: BONE, which is the most dense and rigid of all the tissues, and CARTILAGE, which also constitutes stiff supporting material but is more flexible than bone. Cartilage varies in its elasticity: the most flexible type (generally termed simply ELASTIC CARTILAGE) is not widely distributed through the body but is important in determining the structure and function of the larynx. Most other cartilage, as found in joints for example, is termed HYALINE CARTILAGE.
3. CONNECTIVE TISSUE is a rather loose category (which can be taken in a wider functional sense to include skeletal tissue). It includes adipose or fatty tissue and other tissues which are not particularly relevant in speech. Connective tissues which *are* relevant are the fibrous tissues that constitute TENDONS (which connect muscles to bones and are commonly known as 'sinews'), LIGAMENTS (which connect bones or cartilages to each other) and APONEUROSES (flat sheets of fibrous tissue functioning in the same way as tendons).
4. MUSCULAR TISSUE consists of bundles of fibres. The significant characteristic of muscular tissue is that it can contract, either voluntarily or involuntarily.
5. NERVOUS TISSUE is composed of nerve cells and supporting tissue and is characteristically capable of carrying electrochemical impulses through the nervous system.

The organs of speech, no less than other parts of the body, depend on the nature and function of these tissues. In particular, movement of the speech organs depends on the intricate structure of the body, in which muscles are anchored and interconnected within the skeletal structure, and on the complex operation of these muscles under nervous stimulation. The next section deals with the nervous system and its role in muscular control.

6.3 The nervous system

The nervous system is usually considered to have two parts, the CENTRAL NERVOUS SYSTEM (CNS), consisting of the brain and the spinal medulla, and the PERIPHERAL NERVOUS SYSTEM (PNS), consisting of the nerves distributed through the body.

The CNS begins with the spinal cord in the vertebral column and its extension, the brain stem. Within the brain stem are located the nuclei of the cranial nerves (which are vital to speech production). Posterior to the top of the spinal cord is the CEREBELLUM or 'little brain', whose function is the precisely coordinated muscular control of movement. Above the brain stem are the two cerebral hemispheres, which are fundamentally responsible for complex functions such as speech and vision. Figure 6.3.1 shows this area in sagittal section.

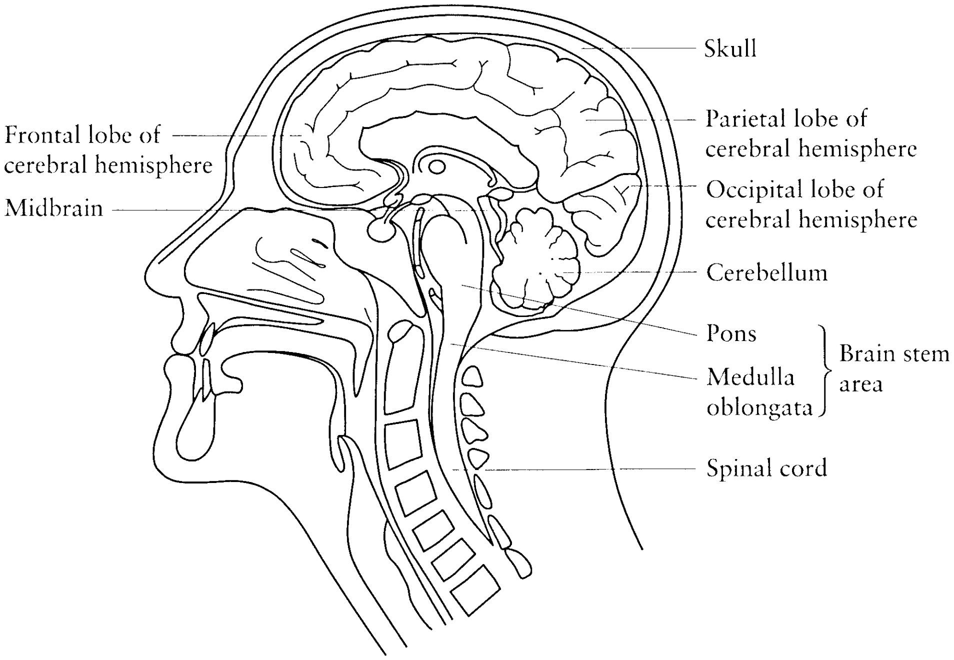


Figure 6.3.1 The central nervous system
Adapted from: Tribe and Eraut 1977, p. 4.

From birth, human development shows increasing evidence that one of the two cerebral hemispheres is dominant in the functioning of spoken language (usually the left, even in left-handed persons). According to Lenneberg (1967) this process of hemispheric lateralization is essentially developmental, and largely complete by the time puberty is reached. More recently, it has been suggested that lateralization is present from birth, and that specific language-related hemispheric function is rather more complex and not primarily developmentally determined (Bryden 1982, Springer and Deutsch 1985). Investigation of the functioning of the brain in language is actually very difficult, and only limited conclusions can be drawn from case studies of persons who have suffered damage to the brain or who undergo brain surgery. In fact little is known about the ways in which the brain and its various neural structures actually initiate and control the complex and integrated activities of speech (see Lenneberg 1967, Kinsbourne 1980, Abbs and Welt 1985, Guenther and Perkell 2004).

The PNS consists of three components: the CRANIAL NERVES, which arise from the brain stem and the head and neck area (much of which is involved in speech production), the SPINAL NERVES, which innervate the trunk and lungs, and the AUTONOMIC NERVOUS SYSTEM, which is responsible for involuntary activities such as blood flow and breathing.

The cranial nerves innervating the vocal tract consist of mixed nerves – that is, they contain both efferent (motor) fibres sending muscle control signals from the CNS, and afferent (sensory) fibres sending information to the CNS from receptors in the skin, mucosa and muscles.

The functional unit of the CNS is the NEURON, which consists of a nerve cell and its nerve fibre extensions (or processes). Figure 6.3.2 shows a motor neuron, consisting of a cell body, the axon (a nerve fibre which conducts impulses to muscles, as shown here, or to other nerves) and dendrites (which are similar to axons but are shorter and may conduct impulses to the neuron). On the cell body and dendrites are connecting points or SYNAPSES, which allow connections with other neurons. Hence the nervous system consists of a complex interconnecting network of neural pathways which can conduct nerve impulses. Muscle commands initiated in the CNS do not travel to their destination via single nerve cells. Rather, they may pass across many synaptic junctions with nearby cells interacting with and modifying the original command impulse before it reaches the muscle.

Neural signals travelling along nerve fibres take the form of short impulses of electrochemical energy caused by the firing of the associated nerve cell. The magnitude of nerve activity is determined not by the strength or amplitude of the pulse but by the number of pulses per unit of time travelling down the axon. A nerve cell can fire only if the impulse energy arriving at one or more of its synapses is above a threshold level. Beyond that level any increase will have no further effect. This 'all-or-none' principle is reflected in the behaviour of muscle fibre.

After a cell has fired there is a short refractory period (typically 0.5 ms) when no further firing can occur. Firing may require several impulses at a

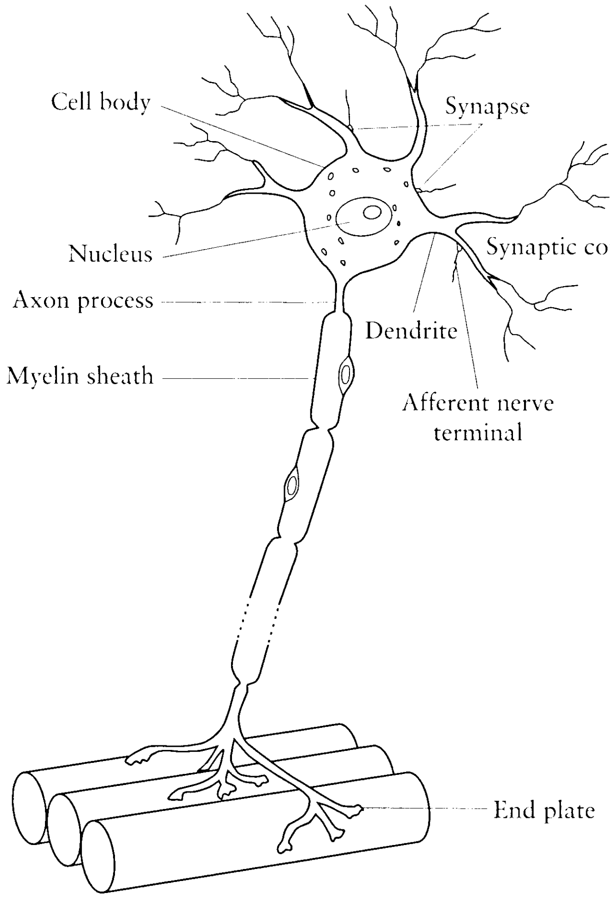


Figure 6.3.2 Motor neuron
Adapted from: Tribe and Eraut 1977, p. 109.

synapse, and there are also inhibitory synapses which can inhibit cell excitation. Thus both summation and inhibition of impulses contribute to the complex and selective control of muscles. The cranial nerves for muscle control in speech are listed in table 6.3.1.

Two types of muscle occur in the body: striated (striped) muscle, which is capable of rapid contraction under voluntary control, and smooth muscle, capable of involuntary contraction (as found in the blood supply system). Only the first of these is of concern in speech production.

Voluntary muscle is made up of bundles of fibres from 10 to 100 microns in diameter and up to 10 cm long. The main fibres are termed EXTRAFUSAL and a smaller group of fibres (which are separately innervated) are termed INTRAFUSAL. Figure 6.3.3 shows typical muscle fibre arrangements. As previously noted, muscle fibre contraction is controlled by impulses from motor neurons, and these are supplied to the muscle fibres by connections at regions called motor end plates. The extrafusal fibres are supplied via large myelinated

Table 6.3.1 Cranial nerves for muscle control in speech

Number and name of nerve		*Function*	*Latency*
V	Trigeminal nerve	Jaw	Short
VII	Facial nerve	Lips	Short
X	Recurrent laryngeal nerve	Larynx	Long
XI	Accessory nerve	Pharynx	Short
XII	Hypoglossal nerve	Tongue	Short

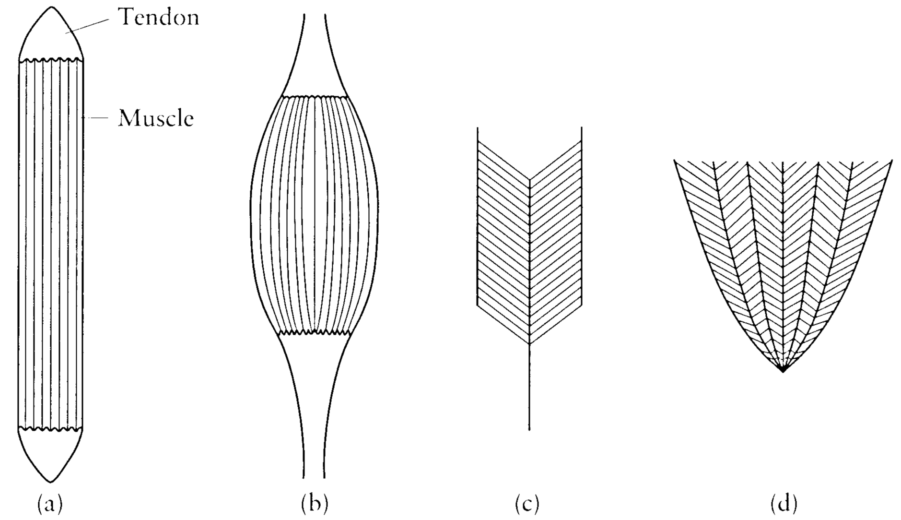


Figure 6.3.3 Muscle fibre patterns: (a) strap; (b) fusiform; (c) bipennate; (d) multipennate
Adapted from: Tribe and Eraut 1977, p. 107.

(alpha) nerve fibres and some smaller (gamma) fibres. The latter also contribute to innervation of intrafusal muscle fibres.

When a nerve impulse arrives at a motor end plate, a complex electrochemical action occurs at the neuromuscular junction, and an action potential, in the form of a wave of depolarization, moves along the outer sheath (the sarcolemma) of the fibre, causing it to contract or twitch once only, about 2 ms after the arrival of the nerve impulse. The resting potential of the outer sheath is –70 mV with a depolarization spike of +50 mV, as shown in figure 6.3.4. To reach maximum contraction, a fibre will usually require several such stimuli. Thus to remain shortened or tensed, a muscle must receive repeated stimulation from nerve impulses. Such sustained shortening is known as TETANIC contraction.

It is the inner structure of muscle fibres that allows them to contract. A fibre consists of thin threads called MYOFIBRILS which in turn are composed of units called SARCOMERES (which are responsible for the striated appearance of voluntary muscle). Within these are fine MYOFILAMENTS which slide between each

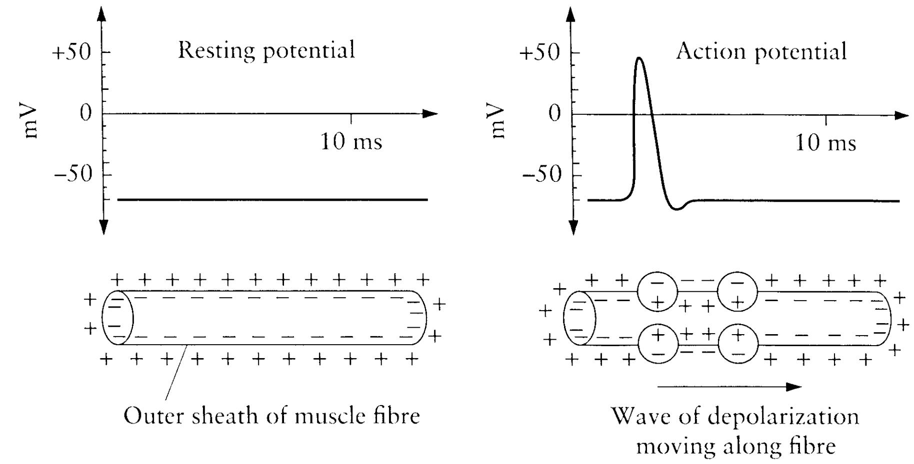


Figure 6.3.4 Potentials across a muscle fibre
Adapted from: Tribe and Eraut 1977, p. 53.

other to cause fibre contraction, their actual length remaining unchanged (Huxley 1958, Tribe and Eraut 1977, Mill 1982).

Muscle activity is of three basic types:

- isotonic contraction, or the dynamic shortening of muscles;
- isometric tension, or increased tension, without shortening; and
- lengthening, where an opposing force is greater than the muscle's active contraction force.

The overall control of muscle can be considered in terms of motor units. Each unit consists of a nerve cell (motor neuron), its axon, and the muscle fibres controlled by the branches of the axon (figure 6.3.2 above). A single impulse in the axon will thus reach all the motor end plates, causing the associated fibres to contract simultaneously. The number of muscle fibres in a motor unit (the innervation ratio) varies from muscle to muscle, depending on their roles. Muscles involved in delicate movement, such as eye muscles or intrinsic tongue muscles, have low innervation ratios (less than 1 to 10). Muscles involved in larger and less precise forms of movement, such as limb muscles, may have much higher innervation ratios (up to 1 to 700). The degree of contraction or tension is determined by the rate of firing of motor units and the number of units firing. The 'all-or-none' principle of contraction is determined by the nerve impulses themselves rather than by the contractile properties of the muscle structure.

So far we have been considering the efferent or motor component of the nervous system. The afferent, or sensory aspect, of the nervous system is also vital to speech production, sending back information from various parts of the vocal tract to the CNS and thereby contributing to appropriate control of the various articulators and their movements. This is one form of the general process known as feedback.

Most of the sensory receptors for speech are found in the oral and respiratory areas. There are also receptors in the muscles and joints which respond to their movement (muscle spindles and joint receptors). Primary afferent endings in the muscle spindle (at the intrafusal fibres) respond to the degree and rate of stretch of the muscle spindle. These afferent neurons make direct synaptic connections with motor neurons in the same muscle, and thus cause the main extrafusal fibres to contract. This is known as stretch reflex. It has been suggested that stretch reflex contributes to muscular control in speech (in conjunction with direct efferent control to both the main muscle fibres and the muscle spindle itself).

The continuous sensory feedback which is needed for coordination of articulatory movement in speech can be divided into three kinds of feedback:

- auditory (by hearing the consequences of an articulation);
- tactile (by the feel or touch of the articulators); and
- proprioceptive (by signals from the muscle spindles and joint receptors which provide information on joint movement and position, and muscle contraction).

Researchers are investigating the question of which feedback is primary, but there appears to be no simple answer. From what happens when a speaker is deprived of feedback, it appears that each of the three types of feedback is related to different aspects of speech control. The relationship is more or less what one would expect from the nature of the feedback. Thus when tactile feedback from the tongue is suppressed, sounds such as [iː] and [s] are inadequately articulated, whereas when auditory feedback is interfered with, it is the overall intensity and pitch of speech that is affected, and so on. Normal feedback mechanisms are clearly important, both in the development of articulatory motor skills in childhood and in the preservation of these skills in adulthood. This is sadly evident in the development of children who are born profoundly deaf: without adequate auditory feedback, these children have great difficulty in acquiring fluent and intelligible speech. Feedback is thus crucial to the acquisition of spoken language (Borden 1980; see also Borden and Harris 1980, MacNeilage 1981, Clark and Palethorpe 1982, Lieberman and Blumstein 1988, Guenther and Perkell 2004 and Kent 2004 for further details of feedback processes in speech).

More extensive general accounts of the anatomical and physiological principles of speech production can be found in Hardcastle (1976), Daniloff et al. (1980), Zemlin (1981), Dickson and Dickson (1982) and Perkins and Kent (1986); and more detailed information on nerve and muscle mechanisms can be found in Mann (1981) and Ottoson (1983). See also Maassen et al. (2004) for a collection of papers on speech motor control, neural behaviour and the acquisition of speech.

6.4 The respiratory system

We begin our detailed examination of the vocal tract with the subglottal respiratory system. The respiratory cycle not only provides the major source of airflow for speech sound sources, but is also important in the sequential organization of speech. For this reason, most accounts of the articulatory processes of speech begin with a categorization of the so-called AIRSTREAM MECHANISMS which provide the sources of energy in the production of speech sounds (section 2.5 above).

The respiratory system (aside from the upper airways in the supraglottal vocal tract) is contained within the chest, or THORAX. It consists of the barrel-shaped rib structure which forms the sides of the thoracic cage itself, the associated muscles, and the lung structure contained within it. There are 12 (paired) ribs, roughly U-shaped, flexibly attached posteriorly to the VERTEBRAL COLUMN, and anteriorly to the breast-bone, or STERNUM, by muscle and connective tissue. (But the two lowermost ribs have no anterior attachments.) The upper limit of the thoracic cage is formed posteriorly by the shoulder blades, or SCAPULAE, and anteriorly by the collar bones, or CLAVICLES. The floor of the cage is formed by the dome-shaped DIAPHRAGM muscle (unpaired). In conjunction with the rib cage, the movement of the diaphragm plays an essential role in the respiratory cycle. Figure 6.4.1 shows the general structure of the thorax.

Within the thoracic cavity are the lungs, which provide the reservoir for airflow in much of speech. In the process of inspiration and expiration in the normal respiratory cycle, they perform the vital function of replenishing oxygen and removing unwanted carbon dioxide from the blood. They consist of soft spongy material which is roughly cone-shaped, with the base resting on the diaphragm and the peak reaching towards the base of the neck. The lungs are connected to the windpipe, or TRACHEA, by two bronchial tubes which join at the base of the trachea. Within each lung the bronchial tubes divide into smaller and smaller tubes, or BRONCHIOLES, which distribute the air supply throughout the lung. These end in tiny air sacs, or ALVEOLI, which make up the bulk of elastic or spongy tissue in the lung structure. Figure 6.4.2 shows the general arrangement of the lungs, bronchial tubes and trachea. The two lungs actually form a single mechanical unit, for they are connected by the PLEURAL LINKAGE, an interface of fluid between the outer lining of the lungs and the inner lining of the thoracic cage. It is through this linkage that changes in the thoracic cavity volume cause changes in lung volume during the respiratory cycle. Since the lungs tend naturally to contract and the thorax to expand (when considered independently of each other), the quiescent volume of the lung–thorax system is the result of a balance of forces.

During inspiration, the thoracic cavity volume is enlarged by two basic means: the rib cage is lifted upwards and outwards, and the floor of the cavity is lowered. The balance between the two is a function of posture, individual habit and respiratory demands. Direct control of rib cage movement during inhalation is principally effected by the EXTERNAL INTERCOSTAL muscles which fill the spaces between the ribs. They are connected between the upper

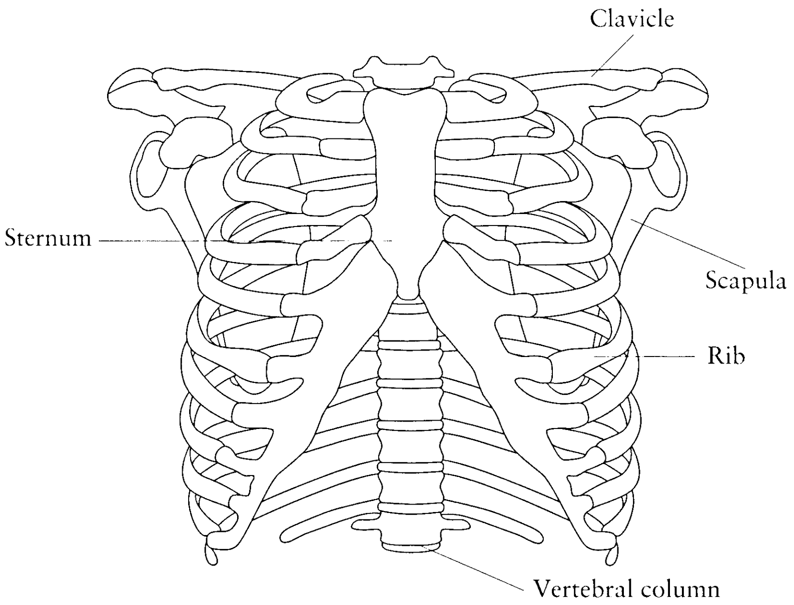


Figure 6.4.1 Structure of the thorax
Adapted from: Zemlin 1968, p. 60.

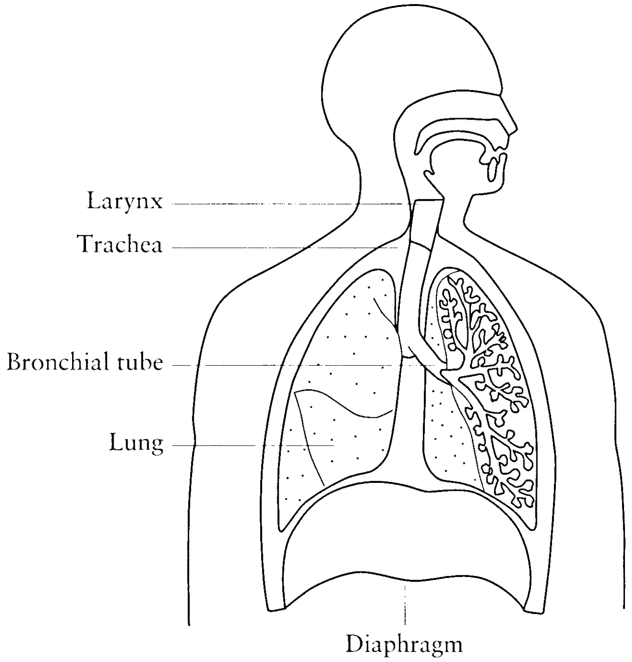


Figure 6.4.2 Lungs, bronchial tubes and trachea: general arrangement
Adapted from: Minifie et al., 1973, p. 78.

and lower edges of adjacent ribs, and as a group run upward and outward relative to the sternum. Their contraction shortens the distance between each rib causing the rib cage structure to be raised. Because of the U-shape of the ribs and the flexible nature of their attachment to the vertebral column and the sternum, raising of the rib cage causes the ribs to rotate relative to their posterior attachment, such that both sides of the rib cage and the sternum move outwards, thereby increasing thoracic cavity volume.

The floor of the thoracic cavity is lowered by contraction of the dome-shaped diaphragm muscle. During quiet breathing it is this diaphragm contraction which is largely responsible for thoracic cavity volume changes during inspiration. In running speech the diaphragm probably retains a major role in increasing the volume of the thoracic cavity during inspiratory phases. Investigations described by Hixon et al. (1977) suggest that the external intercostal muscles may not contribute quite as much as was previously thought, and that the diaphragm may be assisted by tensed abdominal musculature, which optimizes the effect of the diaphragm on thoracic cavity volume by allowing it to work against a taut abdominal wall. These investigations also suggest that under such conditions, the effort of the diaphragm may also contribute to rib cage movement.

In very deep inspiration, as during extreme physical exertion, a number of additional muscles associated with the thorax in the region of the back, shoulders and neck also contribute to the enlargement of the thoracic cavity. These include the serratus posterior superior, the latissimus dorsi, and the levatores costarum in the back, the sternocleidomastoid and the scalenus in the neck, the major and minor pectorals, the anterior serratus, and the subclavius. Figure 6.4.3 shows the rib cage and some of the muscle structure relevant to the inspiratory phase of respiratory activity.

Enlargement of the thoracic cavity volume results, through the pleural linkage, in an increase in lung volume. This in turn lowers the internal air pressure in the lungs relative to external atmospheric air pressure and allows air to flow into the lungs via the nose and mouth to equalize the internal and external pressures. When the lung–thorax system is enlarged, the resultant expansion and movement of muscle and other tissue in the lung–thorax structure also sets up ELASTIC RECOIL FORCES.

For expiration, the lung volume is reduced, causing an increase in internal lung air pressure relative to the external atmosphere, which in turn results in air flowing out of the lungs through the mouth and nose airways to equalize the internal and external pressures. In normal exhalation, the elastic recoil forces set up during inhalation are sufficient to achieve the necessary volume reduction. The resultant air pressure produced in the lungs by the action of these elastic recoil forces is known as RELAXATION PRESSURE. Research into respiratory behaviour during speech described by Ladefoged (1967, although carried out about a decade earlier) has shown that for an appreciable part of the expiratory phase, relaxation pressure produced by the elastic recoil forces is rather greater than required for normal conversational speech. As a result, muscles primarily associated with inspiratory activity, particularly the external intercostals, are used to resist the effects of the elastic recoil forces, until relaxation pressure has lowered to the appropriate level for speech. Hixon (1973) and Hixon et al. (1977) suggest that the effectiveness of the external intercostals in lowering the rib cage to offset excess relaxation pressure is optimized by abdominal forces pulling downwards on the undersurface of the diaphragm. Otherwise, the diaphragm would have a greater tendency to be pulled upwards when relaxed in the expiratory phase. This would decrease thoracic volume and, as a consequence, reduce the effectiveness of rib cage control in offsetting the effects of relaxation pressure.

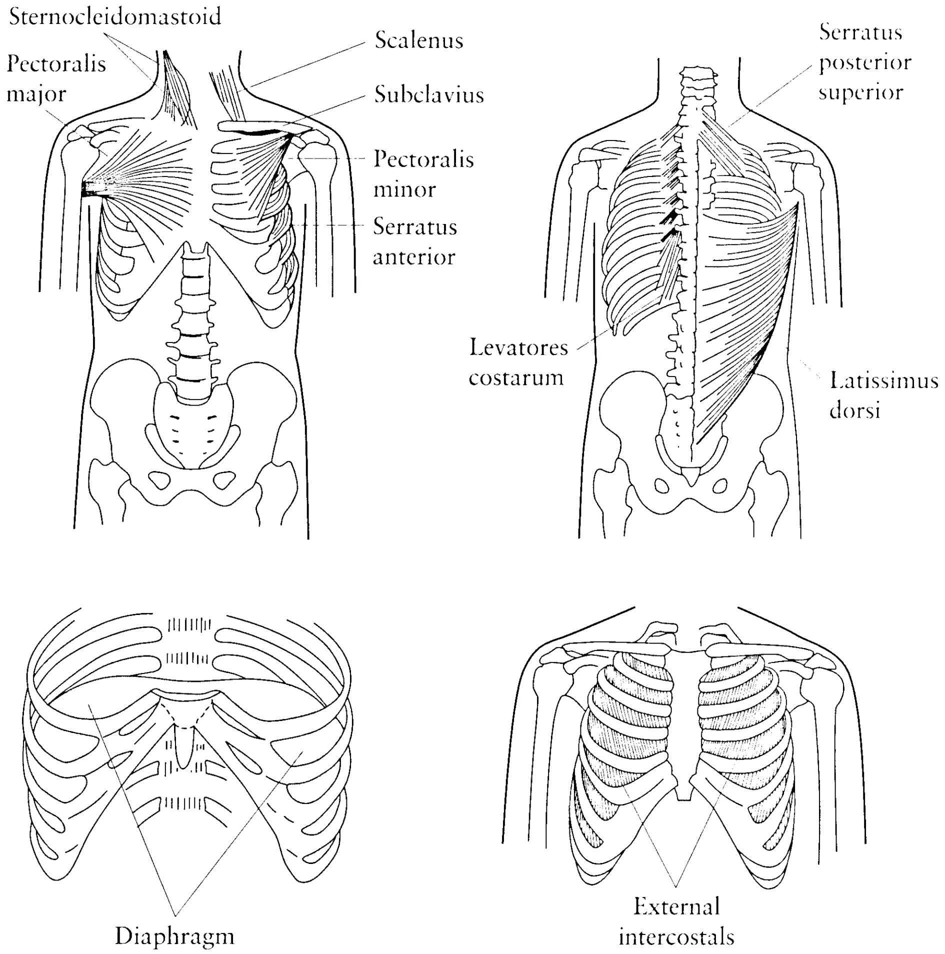


Figure 6.4.3 Muscles used in inspiration
Adapted from: Minifie et al., 1973, pp. 84, 86, 87.

When relaxation pressure can no longer satisfy the aerodynamic demands of speech production, the true muscles of expiration provide the necessary forces to continue reducing lung volume and thereby maintain the required air pressure in the lungs.

The principal muscles which compress the lung–thorax system and therefore reduce its volume are located in the region of the thorax and the abdomen. Certain back muscles also assist under conditions of extreme expiration, such as when shouting, or in producing extremely long continuous utterances, but do not normally make a significant contribution. The principal thoracic muscles involved in exhalation are: the internal intercostals (which lie in the spaces between the ribs approximately at right angles to, and below, the external intercostals); the subcostals; and the transverse thoracic. All of these function to pull the rib cage downwards. The abdominal muscles used in exhalation are: the transverse abdominal; the internal oblique; the external oblique; and the rectus abdominis. All of these function to compress the abdomen, causing upward pressure on the lung–thorax system; in some cases they also assist in compression of the rib cage. Figure 6.4.4 shows some of the muscle structure concerned with exhalation.

The level of respiratory activity required in speech production is greater than in normal quiet breathing, but will vary with the degree of overall vocal effort used. Stetson (1951) proposed that the muscular activity of respiration was to some extent related to the syllabic organization of speech. He claimed that each syllable had an associated 'ballistic chest pulse' initiated by the internal intercostal muscles, but data from physiological experiments (Ladefoged 1967) seem not to support the claim. The general tendency during speech is for the respiratory system musculature to maintain a relatively consistent level of pressure below the glottis, known as SUBGLOTTAL PRESSURE (Psg). Psg is relative to the overall level of vocal effort being employed at the time. Rises in Psg tend to occur on strongly stressed syllables, and falls are often associated with sudden reductions in resistance to airflow at the glottis, as occur in voiceless fricatives. But such changes rarely amount to more than 20 per cent of the average value of Psg. When rapid compensatory activity is required of the respiratory musculature – because of short-term changes in Psg requirements or because

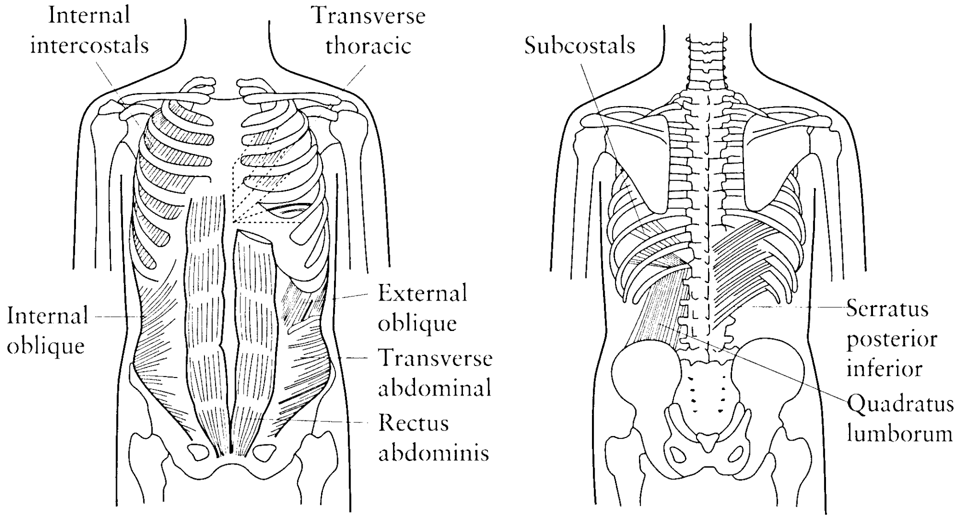


Figure 6.4.4 Muscles used in expiration
Adapted from: Minifie et al., 1973, pp. 91, 92.

of airflow resistance – it appears to be supplied largely by the action of the intercostal muscles, with some contribution from abdominal muscles (Sears and Newsom-Davis 1968, Bouhuys 1974). Fine control of respiratory muscle function in speech is not completely understood, but the available evidence suggests that the role of the diaphragm is not as great as some traditional accounts suggest.

Respiratory capacities vary with posture and bodily size, with a typical TOTAL LUNG CAPACITY in a male adult being 5–7 litres. The maximum volume of air that may be exhaled following maximum inspiration is known as the VITAL CAPACITY; it ranges from about 3.5 to 5 litres. In normal quiet breathing approximately 0.5 litres is inspired and expired. This is known as the QUIET TIDAL VOLUME, representing about 10–15 per cent of vital capacity. During speech, demands on respiratory capacity are somewhat greater, with actual tidal volume depending on the overall degree of vocal effort involved and, to some extent, on the durational demands of the particular utterance. Normal quiet breathing and speech operate in the lower midrange of vital capacity, with minimum respiratory volumes of 30–40 per cent of vital capacity at the end of the exhalation phase. (The figure applies to a speaker standing up and may be slightly different in other postures.) The tidal peak at the end of the inspiratory phase may range from around 45 per cent of vital capacity in quiet breathing to 80 per cent of vital capacity in loud speech. Figure 6.4.5 illustrates typical use of respiratory capacity during quiet breathing and speech.

The aerodynamic demands of speech on the respiratory system vary according to the type of articulation involved, overall vocal effort, and the habits of the individual speaker. Psg provides a measure of the overall articulatory effort being used in a sequence of speech and therefore varies widely between quiet talking and very loud shouting. The perceived loudness of speech is proportional to Psg but in a nonlinear way: it follows a power function law which shows appreciable variability (Isshiki 1964, Ladefoged 1967). In normal conversational speech during phonation (with vocal folds vibrating for voiced sounds) Psg will be in the region of 4–8 cm H_2O, with a minimum differential pressure drop of around 4 cm H_2O needed to initiate vocal fold vibration

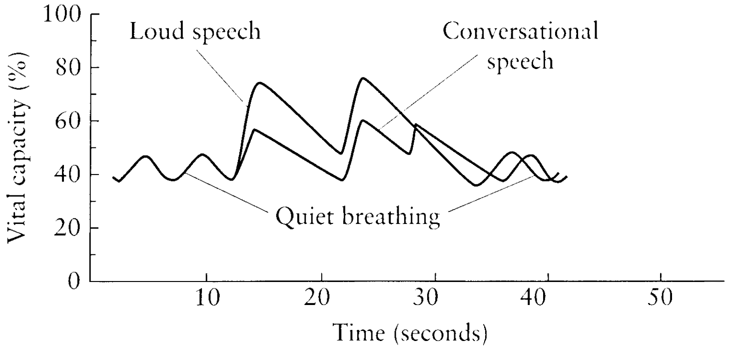


Figure 6.4.5 Respiratory volumes and normal breathing patterns in speech and quiet breathing
Adapted from: Minifie et al., 1973, p. 115.

for voiced sounds (Baken 1987). Airflow rate, while also affected by the general level of vocal effort, varies much more directly according to changes in the resistance to glottal and supraglottal vocal tract airflow caused by individual articulatory configurations. Investigations by Isshiki and Ringel (1964), van Hattum and Worth (1967), Klatt et al. (1968), Gilbert (1973), Clark et al. (1982), Stathopoulos and Weismer (1985) and others suggest that typical flow rates during normal phonation are in the region of 100–200 cm^3 per second for vowels, 200–600 cm^3 per second for voiceless fricatives with the glottis open, and transiently up to 1 litre per second or more at the release of voiceless aspirated stops. Airflow figures of the kind quoted are measured at the lips and nostrils and represent total airflow through the vocal tract (predominantly but not exclusively through the oral cavity).

Some of the basic relationships between respiratory forces and aerodynamic performance are shown in figure 6.4.6, which illustrates the roles of muscular pressure and relaxation pressure in meeting the demands of Psg during a simple sustained vowel sound over most of the range of vital capacity. At point *a* on the graph, which represents the start of expiration after deep inspiration, the positive value of Psg produced by relaxation pressure far exceeds the requirements of the articulation: muscular forces need to work against the excessive relaxation pressure, to bring the net value back to that required for the level of articulatory effort being used. As the available air supply diminishes, the lung volume is reduced, and the relaxation pressure falls; and the level of negative muscular pressure can be correspondingly reduced. At point *b* on the graph, the relaxation pressure curve intersects with the constant Psg value for normal speech and the respiratory muscular pressure value is zero. Relaxation pressure at this point in the expiratory phase is capable of providing the necessary Psg value for normal speech, and no contribution, offsetting or otherwise, is required from the muscular forces. Between points *b* and *c* on the graph, there is a positive contribution from both relaxation pressure and muscular forces since the former is no longer able to supply Psg requirements in this part of the expiratory phase. At point *c* on the graph, the contribution of relaxation pressure has fallen to zero and Psg demands are satisfied by the true inspiratory muscle forces acting to compress the lung–thorax system. Beyond point *c*, the respiratory muscle forces must supply more pressure than is actually required, to offset the now negative effects of relaxation pressure resulting from the compression of the lung–thorax tissue. As can be seen, the negative effects of relaxation pressure become quite large as the expiratory phase approaches zero vital capacity, demanding substantial counterforces. For this reason, speech does not exploit this part of the expiratory phase except under extreme conditions such as prolonged shouting.

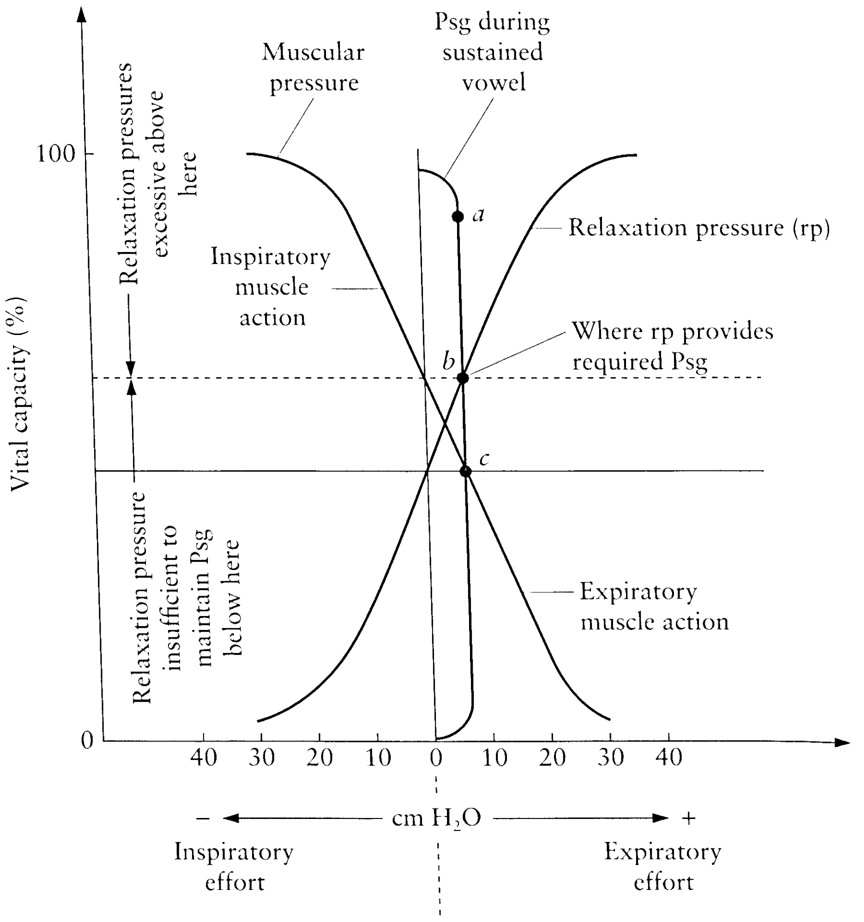


Figure 6.4.6 Respiratory forces producing constant Psg

How the data of figure 6.4.6 relate to the relevant muscle functions can be studied in figure 6.4.7, which is derived from work by Ladefoged and his colleagues. Here, lung volume, Psg, relaxation pressure, and measurements of muscle activity are plotted on a common time scale. It can be seen that at the point where the relaxation pressure curve intersects with the value of Psg, there is a changeover of muscle activity from the inspiratory external intercostals to a number of true expiratory muscles. The data were obtained from a subject counting from 1 to 32 at a reasonably constant level of loudness. Ladefoged and his colleagues are careful to point out that the data come from a single subject and that individuals vary in the way they use their respiratory musculature during speech.

Extensive discussion of respiratory function and the aerodynamics of speech can be found in Hixon (1973, 1987), Warren (1976) and Weismer (1985). Shadle (1997) provides a comprehensive literature review, and a detailed account of methods of measurement.

6.5 The larynx

The basic function of the larynx is as a valve in the respiratory system. Thus in the process of swallowing, the larynx is automatically shut to ensure that

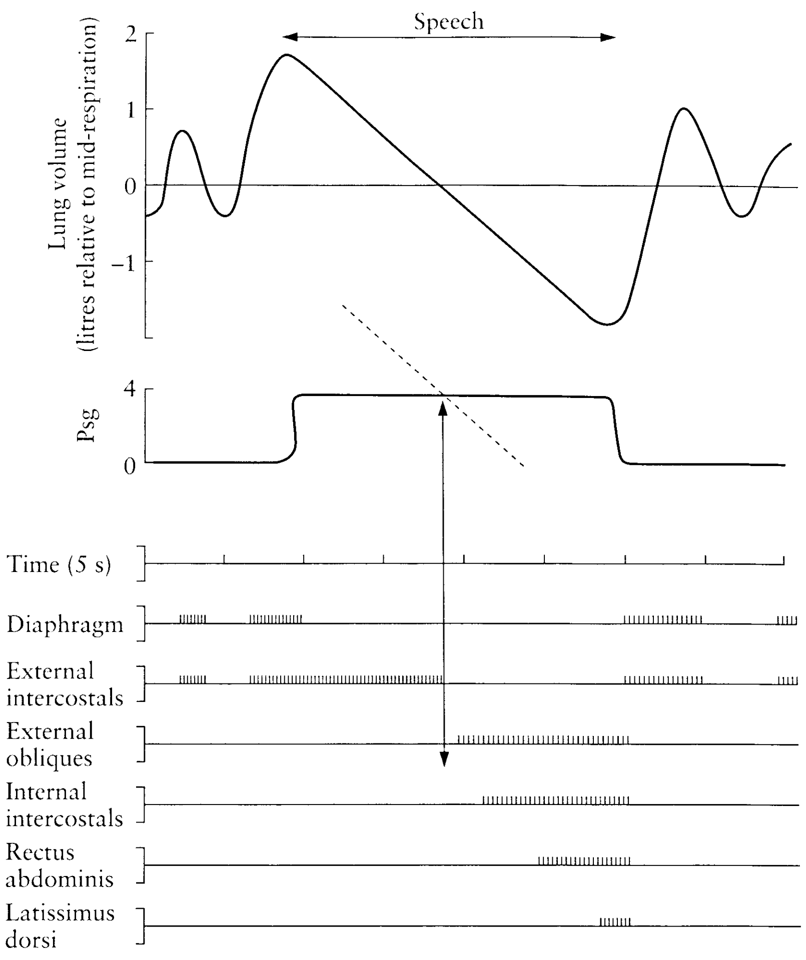


Figure 6.4.7 Respiration and associated muscle activity during speech
Adapted from: Ladefoged 1967, p. 12.

food or drink pass through the pharyngeal cavity into the esophagus and not into the windpipe. We all know the uncomfortable results when this process fails. The valve action of the larynx is also important in short-term physical exertion as a means of stiffening the thorax when we inhale deeply and hold our breath. In speech, the larynx is important as a source of sound and as an articulator.

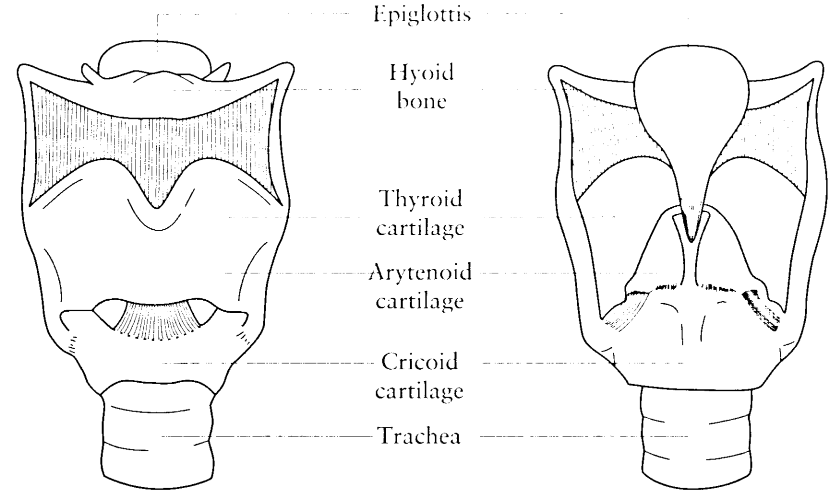


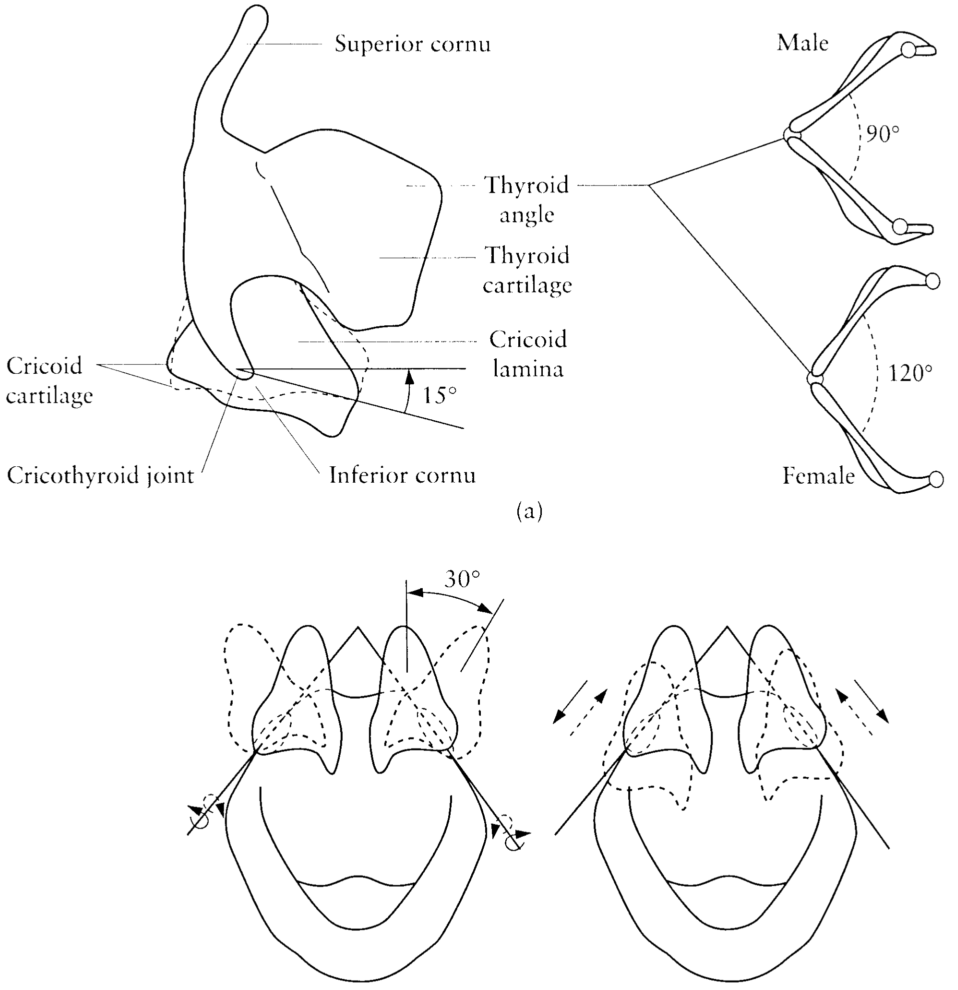
Figure 6.5.1 The larynx: anterior and posterior views

The larynx connects to the lungs via the windpipe or TRACHEA, which consists of a series of roughly horseshoe-shaped cartilaginous sections held together by membranous tissue. It is typically around 11 cm long and 2.5 cm in diameter. The larynx has a skeletal frame formed by a series of cartilages (figure 6.5.1). Some of these cartilages are able to move with respect to each other in ways which affect both the larynx's valving action and its functions in speech production. Figure 6.5.2(a) shows a lateral view of the CRICOID and THYROID cartilages which make up the major part of the cartilaginous laryngeal structure. The cricoid cartilage forms the base of the larynx, and is also the last cartilaginous section of the trachea. It is a complete ring whereas those below it are completed by flexible connective tissue. The cricoid cartilage extends upwards posteriorly to form a plate, or lamina, while anteriorly it is comparable in height to the other tracheal rings. The thyroid cartilage consists of two flat plates forming an angle anteriorly which, among other things, acts as a shield for the vocal folds. The THYROID ANGLE is about 90° in males and about 120° in females. Because the angle is more acute in males, the protrusion can often be seen and felt as the 'Adam's apple'. Posteriorly each plate of the thyroid cartilage has two horns, or CORNUA. The inferior horns form a joint with the cricoid cartilage on its posterior lateral part at matching facets on the two cartilages. This allows the cricoid to tilt over a range of about 15° in an anterior–posterior sense with respect to the thyroid cartilage. The tilting motion plays an important role in controlling vocal fold tension. The superior horns connect to the hyoid bone, which provides the upper suspension of the larynx by muscle connection to the main skeleton structure.

The other important cartilages in the larynx are the small pair of ARYTENOIDS located on the upper posterior lateral part of the cricoid cartilage. The arytenoids move with respect to the cricoid in a rotational and sliding motion which controls positioning of the attached vocal folds, as shown in figure 6.5.2(b). This movement is described in detail by Sonesson (1968), Broad (1973) and Perkins and Kent (1986).

Hinged to the upper anterior part of the thyroid cartilage is the EPIGLOTTIS, commonly described as a leaf- or spoon-shaped cartilage. Its main function appears to be to deflect food from the laryngeal entrance during swallowing.

Extending upward from the superior rim of the cricoid cartilage is a structure of ligamental tissue known, from its shape, as the CONUS ELASTICUS, which



(a)

(b)

Figure 6.5.2 Cartilages of the larynx: (a) cricoid and thyroid; (b) cricoid and arytenoid
Adapted from: (a) Minifie et al., 1973, pp. 130, 133; (b) Sonesson 1968, p. 53.

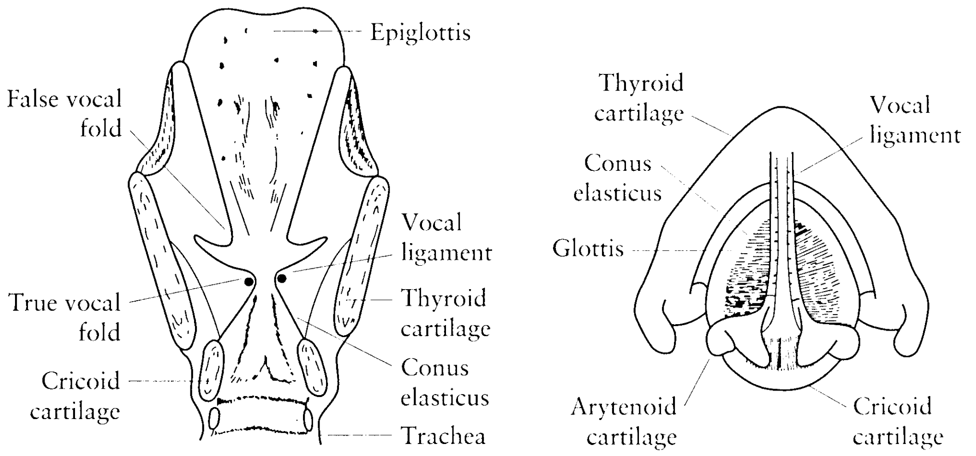
ends in a pair of thickened edges called the VOCAL LIGAMENTS. These form part of the VOCAL FOLDS proper (sometimes a little misleadingly called 'vocal cords'). The folds are roughly triangular in cross-section and they include the upper part of the conus elasticus, the vocalis muscle, and the mucous membrane lining in the laryngeal airway. The vocal folds run from the inferior edge of the thyroid angle to the anterior part of the arytenoid cartilages. The arytenoid cartilages and the vocal folds together form the long slit-like laryngeal valve aperture known as the GLOTTIS. The edges of the glottis (i.e. the length of the vocal folds) are typically about 17 to 22 mm long in males and about 11 to 16 mm long in females. At birth, the length is around 3 mm, and there is no developmental difference between the sexes below the age of 10. Differences in length do appear after 10 years, although there appears to be little evidence that there is any rapid change in the length of the folds accompanying the change in voice pitch in males around puberty. In adults the length of the membranous portion is from four to six times that of the cartilaginous portion; in children the ratio is much lower.

The anatomy of the vocal folds has been well studied, with a view to understanding the behaviour of the folds during speech. The folds are now generally described in terms of cover and body components. These components have distinctive mechanical properties and to some extent move independently of each other; and they may respond differently to the same muscular forces. The general structure of the vocal folds and the glottis is shown in the anterior view and coronal section of the larynx in figure 6.5.3. Hirano et al. (1981) and Kurita et al. (1983) give detailed descriptions of vocal fold structure.

Above the vocal folds is a similar structure known as the FALSE VOCAL CORDS or VENTRICULAR FOLDS (figure 6.5.3(a)). These make no significant contribution to normal vocal fold vibration as such, but they may help lubricate the true folds during phonation, and may assist the glottal valve action when holding high pulmonary (subglottal) pressures.

The various functions of the larynx – valving and phonatory and articulatory activity – depend on a quite complex muscle system. Functionally the muscles fall into two groups, the INTRINSIC laryngeal muscles, which have their attachments within the larynx and are concerned with the control of vocal fold behaviour, and the EXTRINSIC muscles, which are largely concerned with overall movement of the larynx itself.

The intrinsic muscles control the ABDUCTION (opening), ADDUCTION (closing) and TENSIONING of the vocal folds. The posterior CRICOARYTENOID muscle runs from the posterior lamina of the cricoid cartilage to the posterior part of the



(a)

(b)

Figure 6.5.3 Vocal fold structure: (a) anterior view; (b) coronal section
Adapted from: (a) Zemlin 1968, p. 131; (b) Minifie et al., 1973, p. 137.

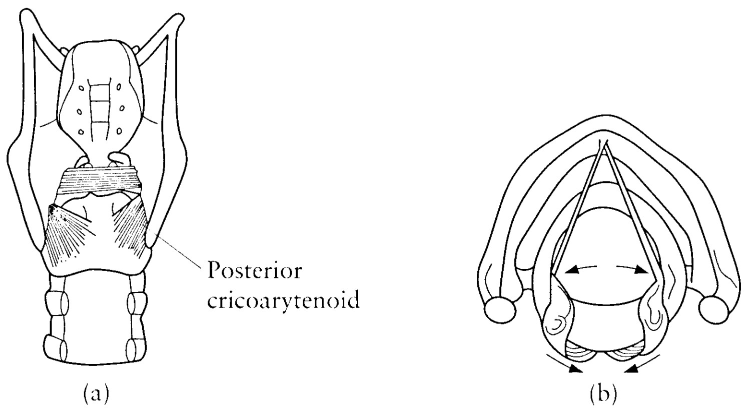


Figure 6.5.4 Posterior cricoarytenoid muscle: (a) posterior view; (b) superior view showing action
Adapted from: Schneiderman 1984, p. 70.

arytenoid cartilages. When contracted, it pulls the arytenoids back and downwards while at the same time causing them to rotate, thus abducting the vocal folds and opening the glottis. This action is illustrated in figure 6.5.4. Under normal conditions of speaking and breathing it is the only muscle responsible for vocal fold abduction, and it is therefore most active during the inspiratory phase of respiration and during the production of voiceless speech sounds.

The LATERAL CRICOARYTENOID and the INTERARYTENOID (or transverse arytenoid) muscles (unpaired) are the principal ones controlling adduction of the vocal folds. The lateral cricoarytenoid muscle runs from the anterior lateral part of the cricoid cartilage to the lateral part of the arytenoid cartilages; it pulls the arytenoids forward and rotates them, thus contributing to vocal fold adduction and overall vocal fold stiffening. The interarytenoid muscle contributes to fold adduction by pulling the arytenoids together, and it tends to reduce vocal fold tension slightly. The actions of both these muscles are shown in figure 6.5.5.

Vocal fold tension, which is important in phonation, is controlled by the THYROARYTENOID, VOCALIS and CRICOTHYROID muscles. The thyroarytenoid muscle runs from the inner part of the thyroid angle to the anterior and lateral surfaces of the arytenoids. Although its exact function in speech production is not fully understood, the muscle appears to shorten and reduce tension in the vocal folds by pulling the arytenoids forward, thus acting as an antagonist to the cricothyroid muscle; but it may also support the action of the vocalis muscle in maintaining tension in the folds. The vocalis muscle, which is sometimes considered simply as a medial component of the thyroarytenoid muscle, runs parallel to the vocal ligaments as part of the vocal fold structure proper. Its function is in general to control tension in the vocal folds, and although the exact nature of this control during speech is still a matter of debate, it appears to stiffen the body while slackening the cover of the folds. The vocalis is generally thought to contribute to quite fine tension control, and possibly to shortening of the vocal folds. Figure 6.5.6 shows a superior view of the muscles.

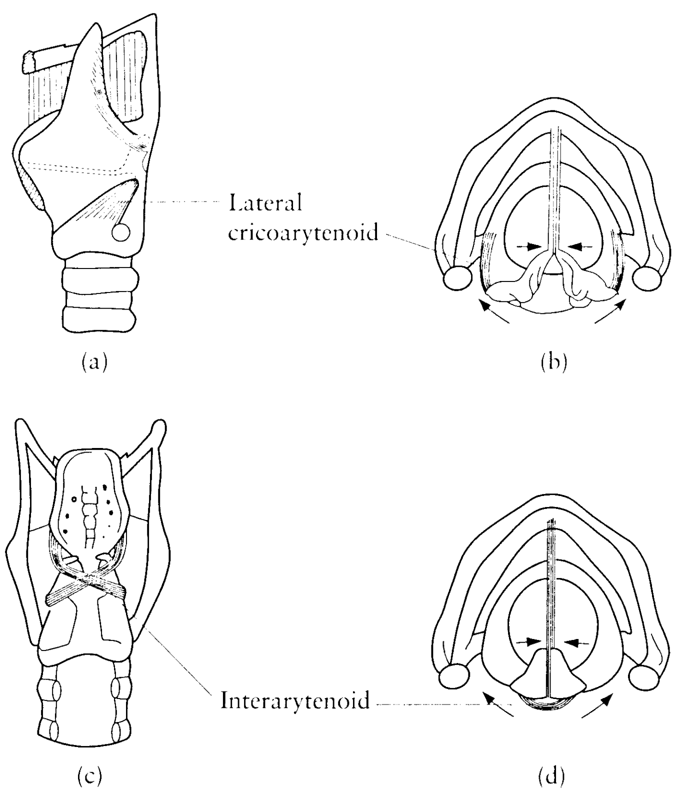


Figure 6.5.5 Lateral cricoarytenoid and interarytenoid muscles: (a) lateral view; (b) superior view showing action; (c) posterior view; (d) superior view showing action
Adapted from: Schneiderman 1984, p. 72.

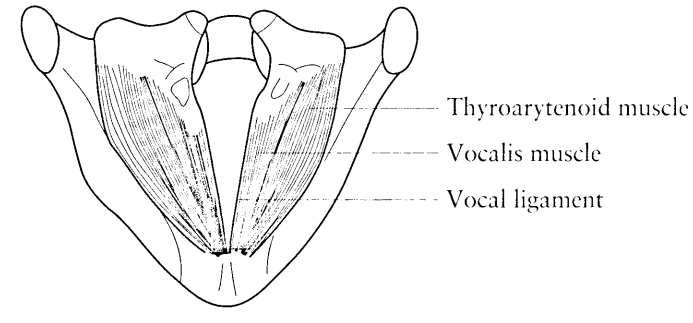


Figure 6.5.6 Vocalis and thyroarytenoid muscles
Adapted from: Zemlin 1968, p. 149.

The cricothyroid muscle runs between the anterior lateral part of the cricoid cartilage and the lower lateral part of the thyroid cartilage. When contracted, it tilts the cricoid cartilage around the pivot formed by the cricothyroid joint, with the result that the arytenoid cartilages move away (backwards) from the thyroid cartilage. This action lowers, stretches, thins and stiffens the vocal fold

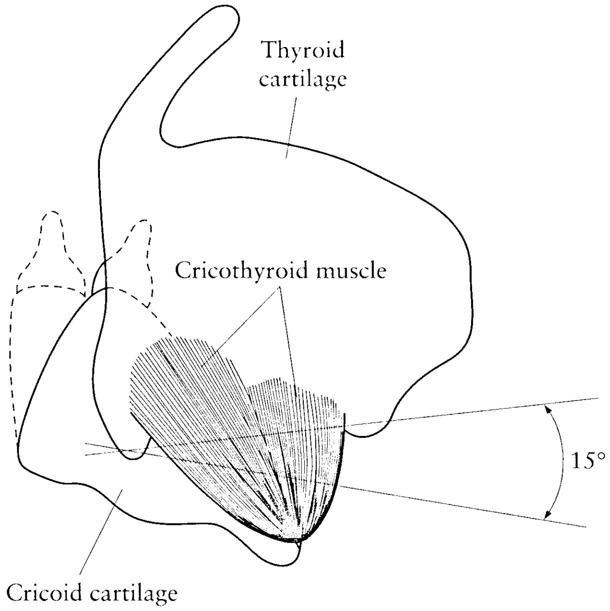


Figure 6.5.7 Cricothyroid muscle: vocal fold tension control
Adapted from: Minifie et al. 1973, p. 158.

structure, increasing both the length and tension of the folds. The tilting action of the cartilages can be seen in figure 6.5.2(b) above; figure 6.5.7 shows the same action in relation to the cricothyroid muscle. For an overview of the role of the intrinsic laryngeal muscles, see Hirano and Kakita (1985).

The extrinsic laryngeal muscles control overall movement and positioning of the larynx, and may be divided into two groups, INFRAHYOIDAL and SUPRAHYOIDAL; the former contribute to lowering the larynx and the latter to raising it. Larynx lowering is largely controlled by the STERNOHYOID, STERNOTHYROID, OMOHYOID and THYROHYOID muscles. The sternohyoid is a long strap muscle which runs from the upper posterior part of the breast-bone to the anterior part of the hyoid bone. Its contraction pulls the hyoid bone downwards and forwards, thus lowering the larynx. The sternothyroid muscle runs from the upper posterior part of the breast-bone to the lateral part of the thyroid cartilage. When contracted it pulls downwards on the thyroid cartilage, contributing to lowering of the larynx. The omohyoid muscle runs from the upper part of the shoulder blade to the lower part of the hyoid bone. When contracted it pulls downward on the hyoid bone and contributes to lowering the larynx. The thyrohyoid muscle runs from the thyroid cartilage to the hyoid bone, and when contracted may help to move both the hyoid bone and the thyroid cartilage, and to lower the larynx, depending on what other extrinsic laryngeal muscles are doing at the same time.

The muscles mainly responsible for raising the larynx are the DIGASTRICUS, GENIOHYOID, MYLOHYOID, STYLOHYOID and HYOGLOSSUS. The digastricus (literally ‘two-bellied’) is a long thin muscle having two components: the anterior component runs from the lower inner face of the jawbone to connective tissue attached to the hyoid bone, and the posterior part runs from the base of the skull to the same connective tissue. When both components are contracted, they will pull the hyoid bone, and hence the larynx, upwards. The geniohyoid muscle runs from the upper anterior part of the inner face of the jawbone to the anterior surface of the hyoid bone, and when contracted will pull the hyoid bone upwards and forwards (provided that the jawbone remains stable). The mylohyoid is a thin sheet of muscle which is part of the structure of the floor of the mouth. It runs from around the inner face of the jawbone via connective tissue to the hyoid bone and (among other functions) aids the action of the geniohyoid and other muscles in raising the larynx. The stylohyoid is a long thin muscle running between the base of the skull and the greater horns of the hyoid bone. When contracted it pulls the hyoid, and thus the larynx, upwards and backwards; its function is therefore similar to that of the posterior component of the digastricus. The hyoglossus (which also functions as a tongue muscle) may contribute to raising the larynx: when the tongue is stabilized by its own extrinsic musculature, the hyoglossus can pull the hyoid bone upwards. The extrinsic laryngeal muscles and the general direction of the laryngeal movements they control are shown in figure 6.5.8.

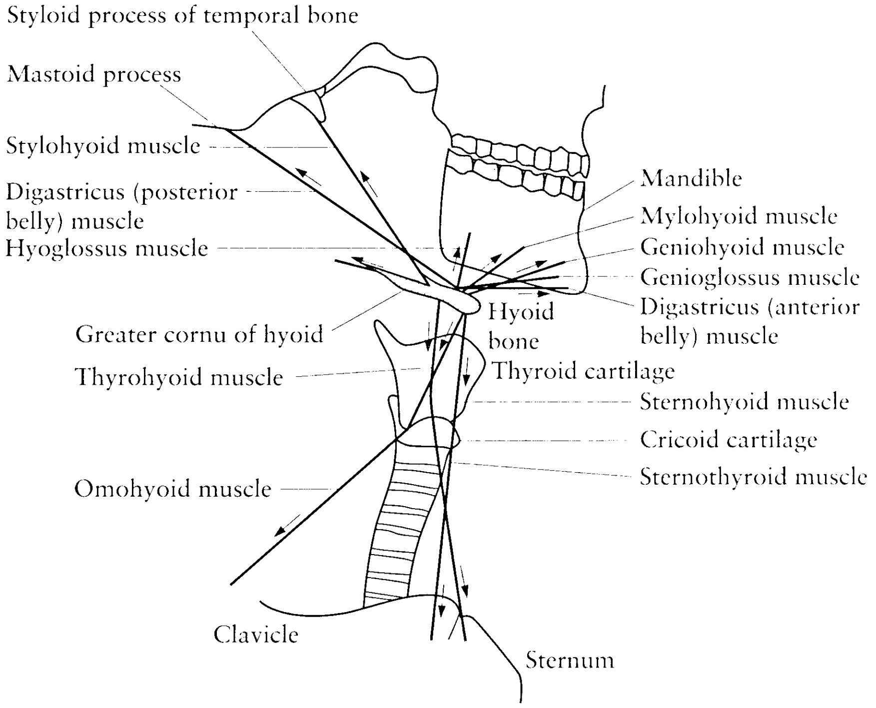


Figure 6.5.8 Laryngeal strap muscles and their actions
Adapted from: Hardcastle 1976, p. 68.

The extrinsic laryngeal musculature is responsible for positioning and stabilizing the larynx, with infrahyoidal muscles acting as antagonists to suprahyoidal. The potential movement of the larynx is mainly vertical, and up–down movement is important in the action of swallowing, as well as in airstream generation for certain sounds that do not use air from the lungs (section 2.5 above). In addition, specific positioning of the larynx can alter the shape and volume of the pharynx, and can indirectly affect the tension of the vocal folds themselves. As a consequence, particular relatively stable settings of the larynx may contribute to voice quality both in speech and in singing.

6.6 Phonation

Phonation (vocal fold vibration, section 2.6 above) is the single most important function of the larynx as a sound source, and the mechanics of phonation have been the object of scientific inquiry for over 250 years. The hypothesis that vocal fold vibration is directly controlled by neural impulses, known as the NEURO-CHRONAXIC theory, is given some credence in older works on phonetics, but the theory finds no support from neurophysiological evidence and in any case presumes a rapidity of muscular control which cannot be substantiated. The explanation of phonation that is now generally accepted is known as the AERODYNAMIC MYOELASTIC theory. This theory takes into account not only the effects of aerodynamic forces, muscle activity and tissue elasticity (van den Berg 1958, 1968), but also the mechanically complex nature of the vocal fold tissue structure (Broad 1979, Hirano and Kakita 1985).

In its simplest form, the aerodynamic myoelastic theory is as follows. When the glottis is closed (i.e. the vocal folds are adducted), expiratory airflow will build up pressure until the vocal folds are forced apart, allowing airflow through the slit of the glottis. Now when a gas or fluid flows through a narrow opening, it accelerates and its pressure drops; the phenomenon is known as the BERNOULLI EFFECT. Thus as air flows through the narrow glottis, the air pressure will be reduced. This in turn will mean that the vocal folds close again, as the pressure reduction sucks them together. The elasticity of the folds assists the entire process, as the folds will part under pressure but will tend to push back once they are apart. The actual opening and closing of the folds has been described as a rippling action: the folds open first at the bottom and the opening moves upward; then the folds close first at the bottom and the closure moves upward. The action is due to the combined effects of the aerodynamic forces and the flexible structure of the folds themselves, in which the cover and body components have some independence of movement. Figure 6.6.1 shows the vibratory cycle (including this effect) as revealed in X-ray studies.

The actual sound produced by the larynx during phonation is created not by the vibration itself, but by the periodic train of puffs of air emitted through the vibrating folds, generating a modulated stream of air. Figure 6.6.2(a) shows the nature of the airstream generated during normal phonation, in the form of what is known as a volume-velocity output waveform. (Details of such waveforms and the techniques of investigation will be found in chapter 7 below, especially section 7.11.)

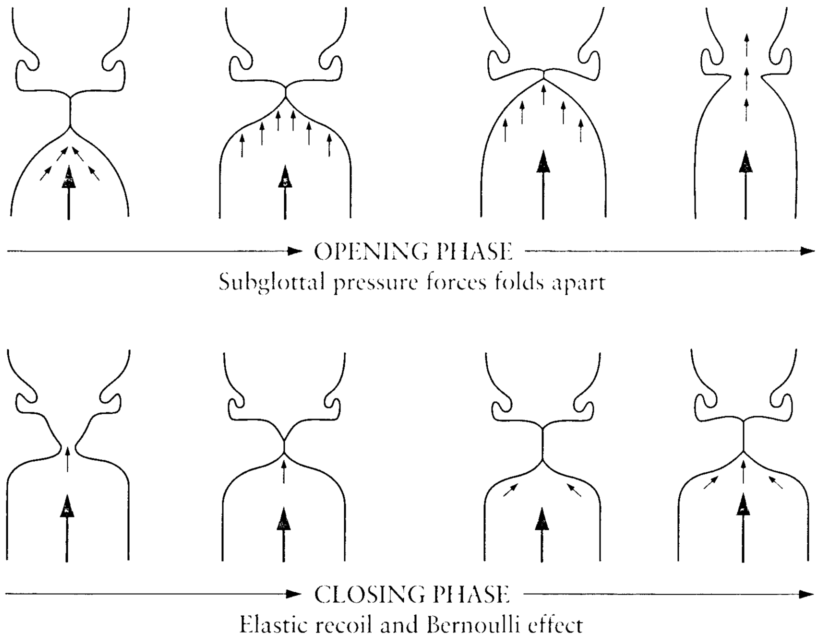


Figure 6.6.1 Vibratory cycle of the vocal folds
Adapted from: Schneiderman 1984, p. 76.

It is possible to distinguish three auditory dimensions or parameters of phonation: loudness, pitch, and a quality of sound that is sometimes called 'timbre'. Normally there is some interaction among the three. Perceived loudness is related to subglottal pressure (section 6.4 above). Pitch is the perceptual correlate of the frequency of vibration of the vocal folds. The frequency is determined by subglottal pressure (Psg) and by laryngeal adjustments governing the length, tension and mass of the vocal folds themselves. Although these control mechanisms do not normally function independently, the logarithm of the frequency of vibration is more or less proportional to Psg for a given laryngeal setting (Ladefoged 1967; see also Ohala 1970 for details of the relationship between frequency and Psg). As already noted, Psg tends to remain relatively constant during a sequence of speech, except for slight rises occurring on strongly stressed syllables. Since pitch contours do not generally follow such a simple pattern, it seems reasonable to suppose that laryngeal adjustments are mainly responsible for controlling the frequency of vocal cord vibration in normal speech (Ohala 1978). On the other hand, Lieberman and Blumstein (1988) maintain that Psg is the primary determinant of frequency, continuing a view long held by Lieberman. It may well be that individual speakers vary in the way that they use laryngeal musculature in relation to Psg to control pitch, but the primacy of laryngeal control is widely accepted.

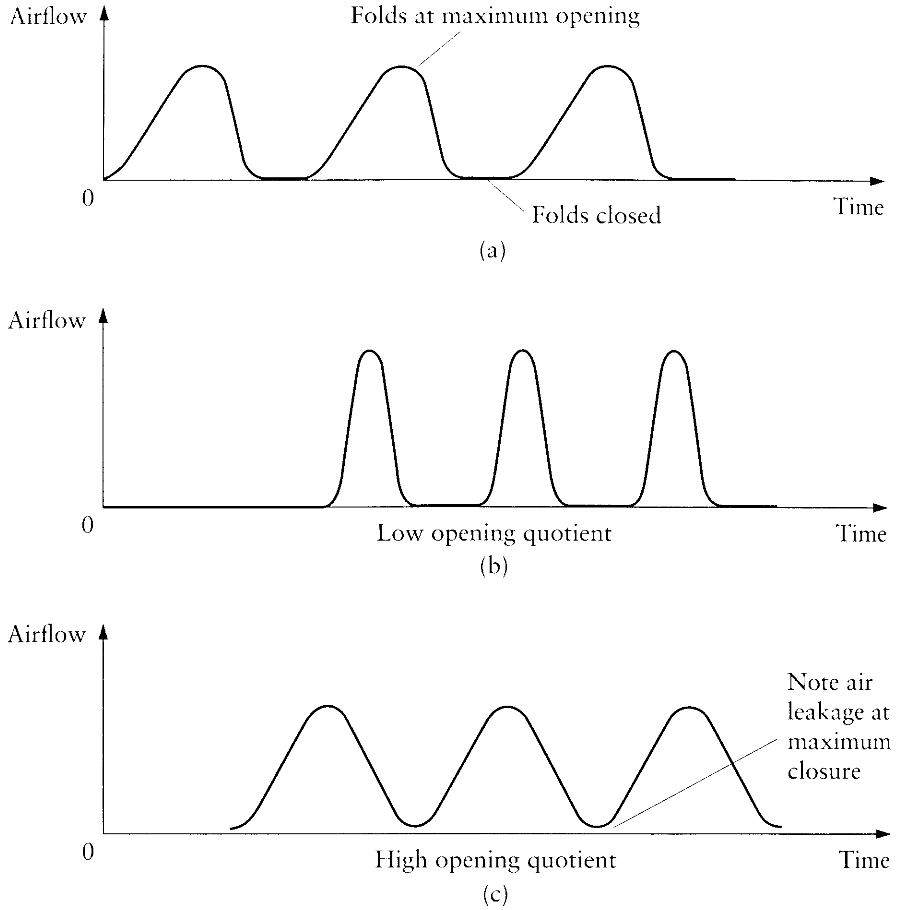


Figure 6.6.2 Phonation airflow waveform: (a) normal (modal) voice; (b) loud (bright) voice; (c) breathy voice

What we have called 'timbre', a quality of sound sometimes reflected in such impressions as the 'mellowness' or 'sharpness' of the voice, is largely determined by the mode of vocal fold vibration during phonation. A simple measure is the OPENING QUOTIENT, the proportion of each vibratory cycle during which the vocal folds are open. The opening quotient is thus the duration of glottal opening during one cycle, divided by the duration of the entire cycle. In normal speech the quotient is typically around 0.5 (see figures 6.6.1 and 6.6.2 above). With increasing loudness, the folds will be closed for longer periods and the opening quotient will fall below 0.5. The explanation for this is that an increase in Psg and in the force of the Bernoulli effect will cause the vocal folds to be forced further apart and pulled together again more rapidly. As the increased kinetic energy must be dissipated, the folds will remain closed longer during the vibratory cycle. The consequent 'sharpness' or 'brightness' of the tonal quality of the voice is due to the additional upper frequency acoustic energy generated by the very rapid changes in the volume-velocity of the airflow. The reverse occurs in soft speech with low vocal effort. Here the opening quotient may become greater than 0.5, giving a 'mellow' tonal quality. With even less vocal effort, the folds do not completely close, and there is some continuous airflow, causing the voice to become 'breathy'. While variation in voice almost always accompanies significant changes in loudness, the reverse is not necessarily true. It is possible to alter the quality without greatly altering the loudness of the voice, by trading Psg and the various muscle forces affecting laryngeal adjustments against each other. Figure 6.6.2 includes volume-velocity waveforms for high and low opening quotients. Finer phonatory distinctions are dealt with under the heading of MODES OF PHONATION (section 2.6 above); see also Lindqvist (1970) and Monsen (1981) for more examples of laryngeal waveforms under varying conditions of vocal effort, showing a variety of opening quotients.

Phonation is never perfectly regular in its periodicity, but (in normal voice) shows a small degree of random variation in both frequency and amplitude from cycle to cycle. Variation in frequency is known as JITTER and variation in amplitude as SHIMMER. These variations are not generally noticeable in the healthy adult voice, but both do increase with ageing, and if the variation is considerable, because of age or vocal or neurophysiological disorder, it may be perceived as a significant component of voice quality. For details, see Heiberger and Horii (1982); for methods of extracting, measuring and analysing laryngeal waveforms, see Baken (1987).

The complex nature of laryngeal control of phonation makes it hard to offer any brief and simple summary. As van den Berg puts it, 'the mean adjustment of the larynx depends mainly on the mean adjustment of the laryngeal muscles' and 'the number of adjustments is infinite' (1968, p. 296). Reviewing the research literature on laryngeal function and control, Ohala notes: 'there is thus much redundancy in the muscular system regulating pitch such that if one or two muscles are lost, the others can take over and pitch regulation is not thereby completely lost, although it may be drastically reduced in range' (1970, p. 19). It must also be recognized that information about laryngeal muscle function during speech is quite limited, because the electromyographic techniques of investigation are both invasive and technically demanding.

Most researchers agree that cricothyroid muscle activity correlates well with pitch control, predominantly in the raising of pitch (by tilting the cricoid cartilage backwards and thereby increasing tension on the vocal folds). At the same time, there may be some decrease in the effective mass of the vibrating part of the folds (due to tissue tension) which will augment this action. There is also an accompanying increase in the length of the folds which may offset the rise in pitch, but tension and its consequences are predominant. The vocalis muscle is also active during pitch rises, although the exact nature of its function is more controversial. It appears to contribute to tensioning and stiffening of the body of the folds; some researchers are of the opinion that it may be responsible for fine incremental control of pitch. The thyroarytenoid and lateral cricoarytenoid muscles also appear to contribute to pitch rises by medial compression of the folds, which may add to their stiffness; when the lateral cricoarytenoid functions as an adductor, it may reduce the effective length of the folds (i.e. reduce the length which is free to vibrate). Such shortening, with a concomitant reduction in effective vibrating mass, will also contribute to a rise in pitch. Other adductor muscles may also play a lesser role in this process. (See van den Berg 1968, Ohala 1970, 1978, Hirose and Gay 1972, Hardcastle 1976, Zemlin 1981 and Honda 1983.)

Although less well understood, pitch lowering is associated with reduced tension in the vocal folds; the vocalis and thyroarytenoid muscles may act as antagonists to the cricothyroid muscle to shorten and reduce vocal fold tension, thereby lowering the rate of vibration. The extrinsic laryngeal muscles also contribute to pitch control, but the relationships between their activity and pitch changes appear to be rather indirect (Sawashima 1974, Honda 1983). There is, however, evidence from Erikson et al. (1983) that infrahyoidal strap muscle activity may complement relaxation of the cricothyroid in pitch lowering. The positioning of the larynx (by extrinsic muscle forces acting on the external cartilaginous structure of the larynx) also has some indirect effect on vocal fold tension: in general, raising of the larynx is associated with raising of pitch, and larynx lowering with lowering of pitch (see Honda et al. 1999 for magnetic resonance imaging (MRI) results showing this effect).

The intrinsic laryngeal muscles also control the timing of laryngeal action relative to supraglottal articulatory activity. The posterior cricoarytenoid muscle, as already noted, is active in abduction of the vocal folds during voiceless sounds. During a voiced sound, it may anticipate the end of voicing by 20–30 ms, cutting off phonation by abduction of the vocal folds. It may also show some activity during voiced fricatives, indicating that the vocal folds are slightly abducted to create a mode of phonation with a higher rate of airflow. The transverse arytenoid muscle appears to function reciprocally with the posterior cricoarytenoid, and may prepare for phonation by adducting the vocal folds 40–50 ms before the start of voicing (Hirose and Gay 1972, Flanagan et al. 1976).

Beyond the normal or modal phonation we have considered so far, and particularly during singing, it is possible to extend pitch range by switching among distinct modes of control known as REGISTERS. Normal speech and the normal range of pitch control operate in CHEST REGISTER. Above the highest pitch which can be attained within chest register, it is possible to enter a higher pitch range by switching abruptly to FALSETTO REGISTER. One of the objectives of voice training for singing is to overcome the abruptness of this change and to control a transitional mode of pitch control known as MIDDLE VOICE REGISTER. According to van den Berg (1968), the pitch control and phonation characteristics of the three registers differ chiefly in tension adjustments in the vocal folds: he describes chest register as having relatively short, thick vocal folds and large amplitude vibrations, and falsetto register as having long, thin vocal folds and much smaller amplitude vibrations. It is difficult to make a smooth transition between registers because the vocal folds have to be considerably elongated and tensed for falsetto, and because the pitch ranges of the two registers may also overlap. To make the transition it is thus necessary to reset the laryngeal musculature. Broad (1973) likens this to shifting gears in a car to suit the speed required. Figure 6.6.3 illustrates the general state of the glottis during the inspiratory phase of respiration, whispered speech, and several phonation modes, including falsetto. Details of various aspects of pitch control and phonation mode can be found in van den Berg (1968), Sawashima (1974), Ohala (1978), Hollien (1983), Lieberman and Blumstein (1988) and Ni Chasaide and Gobl (1997). See Esling and Harris (2005) for a modified classification of phonation modes based on laryngoscopic investigation.

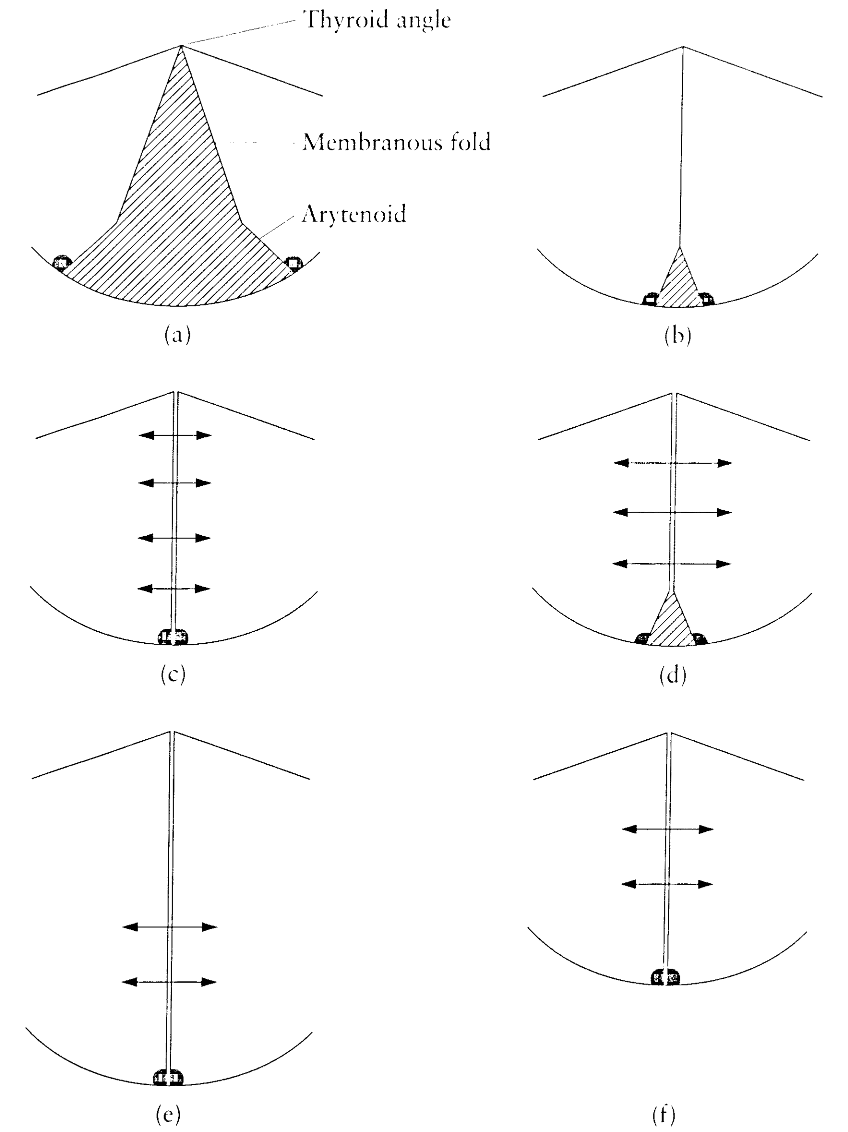


Figure 6.6.3 Glottal state in six common modes: (a) inspiration; (b) whisper; (c) normal phonation; (d) breathy voice; (e) falsetto; (f) creaky voice

6.7 The pharynx

The pharynx is a tube of muscle shaped rather like an inverted cone. Typically around 12 cm long, it lies between the glottis and the base of the skull. It acts as an air passage for respiration, aids in the ingestion of food, and provides drainage for the nasal passages. It makes a passive contribution to speech production by forming part of the length of the supraglottal vocal tract, but its geometry and volume can also be adjusted to vary this contribution or for other articulatory effects. For descriptive purposes it is commonly divided into three functional areas, as shown in figure 6.7.1.

The lowest section is the LARYNGO-PHARYNX, bounded inferiorly by the glottis and superiorly by the hyoid bone (but some writers take the laryngo-pharynx to extend to the tip of the epiglottis). Because of the muscular linkages between the hyoid bone and the body of the tongue, tongue movement can change the diameter of the laryngo-pharynx quite considerably, particularly in a lateral direction. Up and down movement of the larynx also substantially alters the length (and hence volume) of the laryngo-pharynx.

The mid section is the ORO-PHARYNX, bounded inferiorly by the hyoid bone and superiorly by the soft palate. Since the anterior face of this section is formed by the back of the tongue and the upper part of the epiglottis, it also undergoes considerable changes in volume and geometry as the tongue moves. The diameter of the pharynx at the tip of the epiglottis, which has membranous links to the back of the tongue, may vary from 20 mm or more during the articulation of front vowels (as in *heed* or *head*) to 5 or 6 mm during back vowels (as in *board*). Further details can be found in Zemlin (1981).

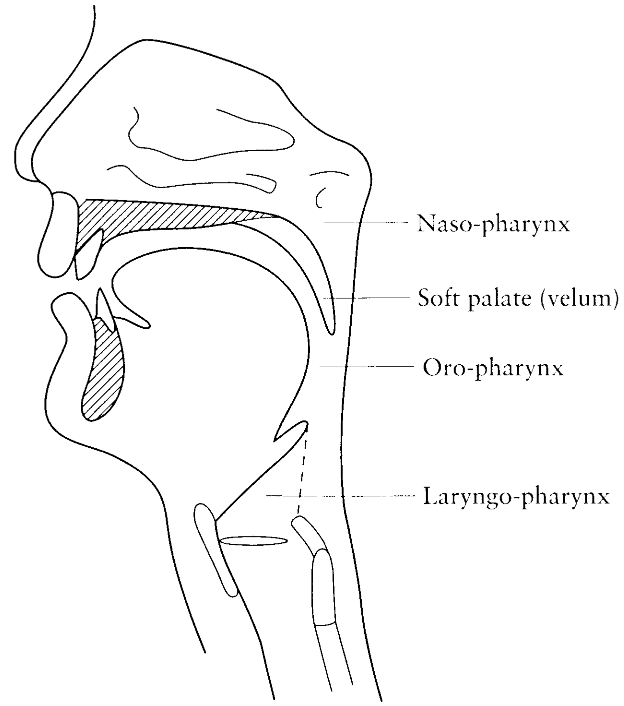


Figure 6.7.1 The pharynx

The upper section is the NASO-PHARYNX, bounded inferiorly by the soft palate and extending to the nasal passages. It can be sealed off from the lower sections of the pharynx by raising the soft palate (figure 6.7.1 above; and section 6.8 below).

6.8 The velum and the nasal cavity

The soft palate or VELUM is a continuation of the roof of the mouth, posterior to the bony structure of the hard palate. It consists of a flexible sheet of muscular tissue covered in mucous membrane ending at the UVULA, a small tip of muscle and flexible tissue. When raised, the velum serves to seal off the nasal cavity by closing the entrance to it, known as the VELOPHARYNGEAL PORT. Three classes of muscle are relevant to the velum and its functions: those which raise the velum, those which enhance velopharyngeal closure, and those which lower the velum. Figure 6.8.1 shows an anterior view of the velum and uvula through the oral cavity.

The most important muscle for raising the velum is the LEVATOR PALATINE, which runs from the temporal bone at the base of the skull (about ear level) down into the medial part of the aponeurosis of the velum. When contracted it pulls the velum upward, tending to close off the velopharyngeal opening.

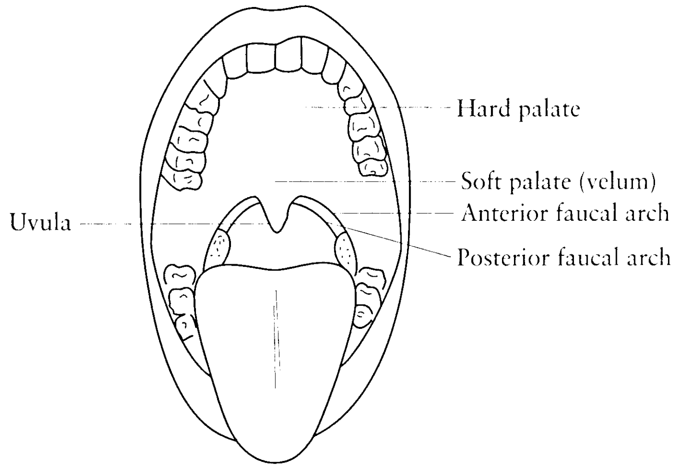


Figure 6.8.1 Oral vocal tract: anterior view
Adapted from: Zemlin 1968, p. 254.

This action may be assisted by the UVULAR MUSCLE which runs from the posterior part of the aponeurosis of the velum to the mucous membrane of the uvula. When contracted, the uvular muscle will shorten and raise the uvula.

Velopharyngeal closure can be enhanced by the PALATAL TENSOR. This muscle runs from the sphenoid bone at the skull base to connective tissue which passes around a projection of the same bone to insert into the aponeurosis at the sides of the velum. When contracted, the muscle serves to stretch the velum laterally, helping to press it against the oropharyngeal wall. In addition, the PALATOPHARYNGEAL SPHINCTER, a single muscle which runs around the pharyngeal wall, may further improve sealing of the velopharyngeal port by forming a bulge, known as 'Passavant's ridge', which presses against the velum. But not many speakers seem to exploit this mechanism in normal speaking, although it may contribute to the pharyngeal wall movement that occurs in some vowels.

The velum is lowered by the action of the PALATOGLOSSUS and PALATOPHARYNGEAL muscles. The palatoglossus, which also counts as an extrinsic tongue muscle, runs from the aponeurosis of the velum to the posterior lateral edges of the tongue. When contracted, it will pull the velum downwards if the tongue is steady. Conversely, it raises the tongue if the velum is stabilized. The palatopharyngeal muscle, which is posterior to the palatoglossus, runs from the inferior aspect of the aponeurosis of the velum to the pharyngeal walls in the area of the posterior part of the thyroid cartilage. When contracted with the larynx stabilized, it may contribute to pulling the velum downwards, and conversely may assist in raising the larynx if the velum is stable. These two muscles and their associated tissue form the ANTERIOR and POSTERIOR FAUCAL ARCHES, shown in figure 6.8.1 above. Figure 6.8.2 shows a sagittal section view of the velum and its associated muscle structure.

In the production of vowels, the velum may be raised to direct airflow through the oral cavity, in which case the vowel is said to be oral; or it may be lowered to allow air to flow out through the nasal cavity as well as through the oral cavity, in which case the vowel is said to be nasalized. In some languages, such as French and Portuguese, the difference is systematically exploited, so that some of the vowels are oral and some are nasalized. In other languages, such as English and German, there is no such distinction and all vowels are oral. But in oral vowels it is quite common for some flow to occur through the nasal cavity as well, because of incomplete velopharyngeal closure. This nasal flow may be due to persistent or anticipatory velum movement (because of neighbouring nasal consonants), or to the habitual articulatory patterns of the individual, or to the linkage between the velum and the tongue such as that formed by the palatoglossus muscle. According to Moll and Shriner (1967), this linkage or 'mechanical coupling' could account for the fact that the velopharyngeal port tends to be more open during a vowel with low tongue position (as in *hard*) than during a vowel with a high tongue position (as in *heed*). But Lubker (1975) disagrees with Moll and Shriner, and the question of muscle function in velum control (particularly the roles of the palatopharyngeal sphincter and the palatopharyngeal muscle) is not fully resolved. Part of the problem is that there appear, once again, to be marked differences among individual speakers in the way they use the musculature to control the velopharyngeal port. Specific muscular activity also appears to be influenced by phonetic context. Whether there is in fact a greater degree of inherent nasality in low vowels (as Ladefoged 1971 suggests) seems to depend in part on the way in which nasality itself is defined and measured. (See Bell-Berti 1980 for an extensive discussion of velopharyngeal function, and Baken 1987 for a review of acoustic and aerodynamic techniques for estimating port size, velum height and movement, and nasality.)

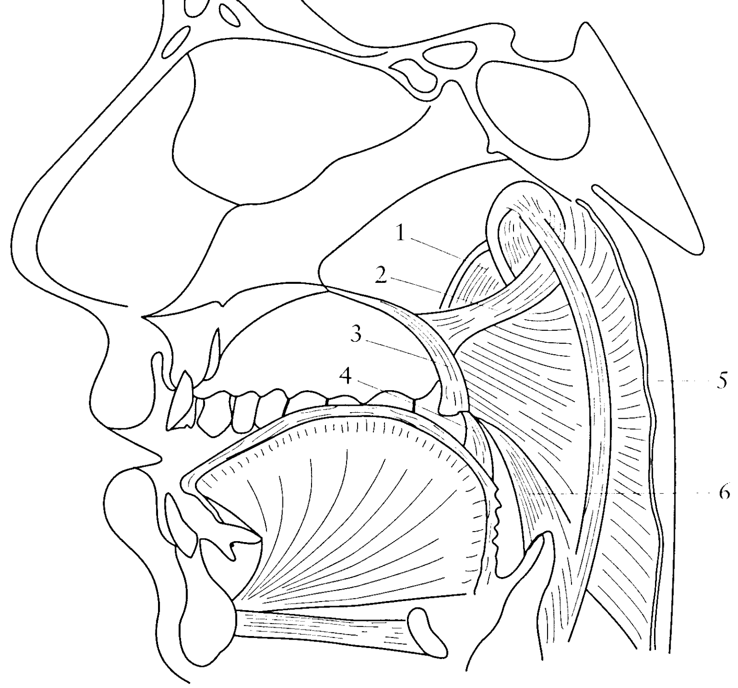


Figure 6.8.2 Supraglottal vocal tract showing soft palate musculature: 1 palatal tensor muscle; 2 levator palatine muscle; 3 uvular muscle; 4 palatoglossus muscle; 5 palatopharyngeal sphincter; 6 palatopharyngeal muscle
Adapted from: Zemlin 1968, p. 299.

In the articulation of stops, the velum must be fully raised to allow adequate buildup of intra-oral air pressure during the stoppage. In fricatives, velum raising is also important, though to a lesser extent. The speech of those who have a cleft palate or related structural deficiency reveals the consequences of inability to create a reasonable degree of velopharyngeal closure when required.

In some languages, the uvula also has a function as an articulator, in conjunction with the body of the tongue, in producing certain trill and fricative sounds.

The naso-pharynx leads into the nasal cavity, which has a rather complex cross-sectional shape, giving it a large surface area. The size of the area and its covering of mucous membrane mean that incoming air is warmed and humidified during normal respiration through the nose. The cavity is typically about 10 cm long from pharynx to nostrils, and is divided into two passages by the SEPTUM, which is cartilaginous at the nostril end and joined to bone structure in the skull. Three bony protrusions, or CONCHAE, extend from the lateral walls of the nasal cavity, partially dividing it into three passages on each side and contributing to the large surface area. Coupled to them is a series of auxiliary cavities, or SINUSES.

The nasal cavity system has a complex shape, but lacks muscular structure to vary this shape. External factors do affect the size and shape of the nasal cavity, however: as Fant (1960) points out, the mucous content varies, and tissue may swell, causing considerable variation in the volume and geometry of the cavities. These variations are obviously involuntary but certainly affect the nasal cavity's resonant properties and its contribution to the acoustic and perceptual characteristics of speech. Voluntary control of the cavity's contribution to sound quality can be achieved only indirectly, by muscular tensions which affect the nature of the nasopharyngeal coupling to the oropharyngeal part of the vocal tract. It is possible, for example, that the nasalized vowels of French gain a particular quality from the kind of nasal coupling which occurs when the velum is lowered and tension is maintained on associated muscles.

6.9 The oral cavity

The oral cavity is the single most important part of the vocal tract in determining the phonetic qualities of speech sounds. Its importance rests on our ability to control the geometry and volume of the cavity, by shaping and positioning the tongue and by moving the lips, jaw and soft palate (section 2.10 above).

The limits of the oral cavity are defined anteriorly by the lips, and posteriorly by the arch-shaped entry formed by the palatoglossus muscle (the ANTERIOR FAUCAL PILLARS). Inferiorly, the floor of the mouth is formed by the tongue, flexible connective tissue, and the (extrinsic laryngeal) mylohyoidal muscle. Superiorly, the oral cavity is divided from the nasal cavity by the roof of the mouth, the front of which is defined by the edges and inner surfaces of the upper teeth. Just behind the upper teeth is the ALVEOLAR RIDGE, the thick membranous covering on the bone structure which joins the tooth-bearing bone of the upper jaw and the vaulted or arched bone structure of the hard palate. At the meeting point of the tooth ridge and the hard palate, the membranous cover has a series of distinctive ridges or corrugations across it. These and the shape of the palatal arch itself vary widely from individual to individual. The hard palate ends approximately level with the rearmost molars and the partition between the nasal and oral cavities is continued by the soft palate or velum (section 6.8 above). Laterally, the oral cavity is delimited by the teeth and associated bone structure in the jaws, extending to the flexible tissue and muscle structure of the cheeks when the mandible is lowered. For reference, the figure used in section 2.10 above (figure 2.10.1) to illustrate the principal areas of articulation is repeated here as figure 6.9.1.

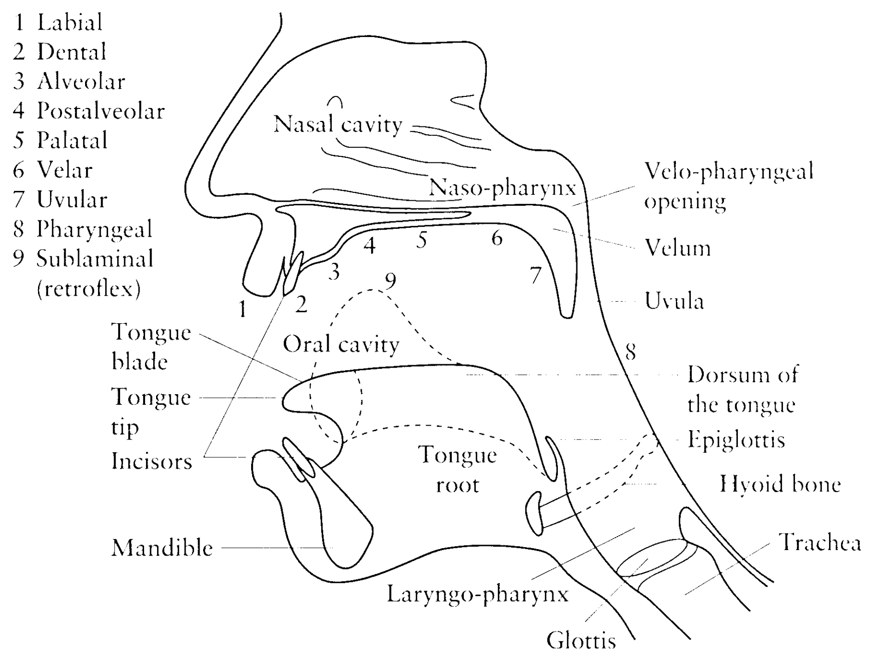


Figure 6.9.1 Mid-sagittal oral vocal tract showing major areas of articulation
Adapted from: Minifie et al. 1973, p. 173.

6.10 The tongue

Within the oral cavity is the tongue, which makes the greatest contribution to changes in the volume and geometry of the cavity. The tongue consists largely of muscle, with an outer covering of mucous membrane and a fibrous septum dividing it longitudinally. It is anchored anteriorly by some of its extrinsic muscles to the hyoid bone. For purposes of phonetic description, the upper surface of the tongue is usually divided into functional areas (section 2.11 above). Actually, as Hardcastle (1976) points out, there is no anatomical basis for such subdivision. Probably for this very reason, writers on articulatory phonetics differ in their use of terms (see Heffner 1964, p. 32; Abercrombie 1967, p. 53; Daniloff 1973, p. 175; Zemlin 1981, p. 318; a detailed account of tongue anatomy is found in Fucci and Petrosino 1981).

The extrinsic muscles of the tongue make for highly versatile positioning of the tongue, while the intrinsic muscles work with the extrinsic to give control over tongue shape. The consequent mobility and plasticity of the tongue are fundamental to speech production.

The extrinsic muscles of the tongue are the PALATOGLOSSUS, STYLOGLOSSUS, GENIOGLOSSUS and HYOGLOSSUS (figure 6.10.1). When contracted, the palatoglossus (see section 6.8 above) assists in raising the back part of the tongue. The styloglossus runs from the base of the skull down and forward to

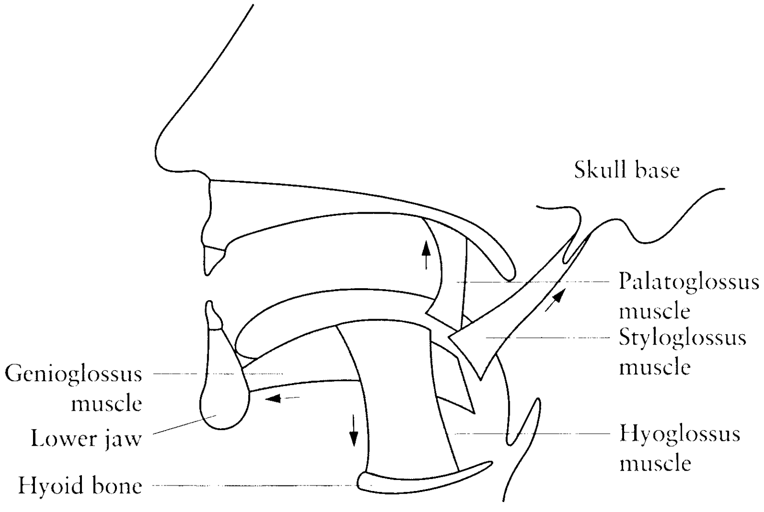


Figure 6.10.1 Extrinsic muscles of the tongue
Adapted from: Sonesson 1968, p. 67.

the back edges of the tongue, dividing and distributing into the hyoglossus and (intrinsic) inferior longitudinal muscle running towards the tongue tip. Its contraction will pull the tongue upwards and backwards. The genioglossus is a bulky muscle which runs from the medial part of the posterior surface of the jawbone, fanning out upward into the tongue from the tip to the root and extending downward towards the hyoid bone. The fibres in the anterior and posterior parts of this muscle are capable of independent contraction, which enables the muscle to perform a variety of functions in the control of the tongue. When the anterior part is contracted, it may pull the tip back and down within the jawbone trough. When the posterior part is contracted, it may pull the tongue forward, causing the tip to protrude. The hyoglossus (mentioned in section 6.5 above) runs upward and forward from the greater horns and anterior lateral part of the hyoid bone to the root of the tongue, blending with other tongue muscle fibres from the back to the tip of the tongue. When contracted, it may pull the tongue downwards. With its anterior fibres it aids the action of the genioglossus in depressing and pulling back the tongue tip, and also functions as an antagonist to the palatoglossus and styloglossus.

The intrinsic muscles of the tongue mostly lie above the extrinsic, and are configured in three planes at right angles to each other (Sonesson 1968, Zemlin 1981). The SUPERIOR and INFERIOR LONGITUDINAL muscles run along the tongue. The superior longitudinal is directly under the surface of the dorsum, and runs from the tongue root to the tip and lateral edges of the tongue. Its contraction can shorten the tongue and contribute to raising the tip and edges. The inferior longitudinal runs from the root of the tongue, blending with the genioglossus and hyoglossus muscle fibres, to the lower surface of the tongue tip. When contracted, it may lower the tip and contribute to shortening the tongue. The TRANSVERSE muscle, which forms a significant part of tongue body bulk, runs from the fibrous septum towards the lateral edges of the tongue, blending with other muscles in this region. When contracted, it may narrow and elongate the tongue and contribute to grooving the tip and blade. The VERTICAL muscle runs from the mucous membrane of the dorsum downwards towards the lower side of the tongue, blending with the inferior longitudinal and transverse muscle fibres. When contracted, it may flatten and widen the tongue. Figure 6.10.2 shows the location of the intrinsic muscles.

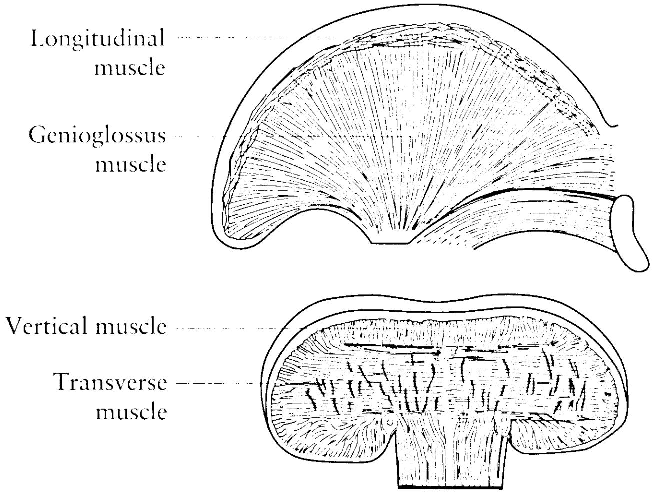


Figure 6.10.2 Intrinsic muscles of the tongue
Adapted from: Sonesson 1968, p. 68.

In completing this summary of the role of individual muscles in tongue control, we should not lose sight of the way in which the muscles work together to create an enormous diversity of tongue shape and position during speech. Cooperation between extrinsic and intrinsic muscles is made clear, for example, in a study by MacNeilage and Sholes (1964) of tongue muscle activity in vowels. Studies designed to capture the complex positioning and posture of the tongue include the famous and early example of Daniel Jones, who used lateral X-ray photographs as evidence in his theory of vowel description (section 2.7 above). Johnson et al. (1993), Perkell (1996) and others have undertaken similar analyses in a more modern context using the respective techniques of X-ray microbeam and electromagnetic midsagittal articulography (EMMA) to track tongue movement during vowel production.

In a now classic study, Perkell (1969) analysed the dynamic behaviour of the tongue during articulation by using ciné X-ray film measurements. He suggests that the extrinsic tongue muscles are responsible for the relatively slow positional adjustments required mainly in vowel production, and that the intrinsic muscles control the more rapid localized movement and shaping required mainly in consonant production. The slower positioning includes, for example, raising and retracting the tongue, using muscles such as the styloglossus, for a back vowel as in *hoard*. The faster and more localized kind of activity includes blockage of the vocal tract by placing the tongue tip just behind the upper teeth, followed by rapid release of the blockage (as in the stop [d] at the beginning of *deal*). This action uses muscles such as the superior and inferior longitudinals to control the requisite rapid movement of the tongue tip. As a result of this earlier work and later research using EMMA, Perkell (1996) proposes a two-dimensional model of the tongue that incorporates aspects of tongue muscular anatomy and biomechanical properties. A number of researchers have also described tongue activity or devised models of the tongue using X-ray microbeam data and ultrasound (e.g. Stone 1990), MRI (Baer et al. 1991, Stone et al. 2001) or MRI and electromyography (e.g. Honda 1996), and ultrasound (Stone 2005). Modern articulatory tracking and monitoring devices described above are enabling less invasive recording of tongue *movement* but, as with most of the articulatory system, the relationship between muscular activity and articulatory function is not simple. As Gracco (1993) and others have observed, individual muscle activity in the vocal tract is meaningful only when understood in terms of the goals of the articulatory gesture and context.

6.11 The lips

The lips are the anterior termination of the oral cavity and thus also of the entire vocal tract. They consist of two fleshy folds which are richly supplied with muscles and are formed externally of skin and internally of mucous membrane. The muscle arrangements are such that the lips and mouth show considerable plasticity and mobility and therefore contribute significantly to the range of vocal tract configurations possible in the articulation of speech.

Muscles associated with the lips allow control over opening and closure, raising and lowering of the upper and lower lips, rounding and protrusion of the lips, and vertical or lateral movement of the angles or corners of the mouth (figure 6.11.1).

The major muscular component of the lips proper is the ORBICULARIS ORIS, a sphincter muscle consisting of an oval band of fibres, some of which are shared with other facial muscles which pass into the lips. The muscle is thus capable of providing a range of different movements associated with lip control. When the muscle is contracted, movements include lip closure, pursing, rounding and protrusion, drawing the upper lip down and the lower lip up, and pressing the lips against the teeth. Lip protrusion is also assisted by the MENTALIS muscle, which runs from the anterior part of the mandible below the lower incisors down to the lower part of the chin. When contracted, it may contribute to raising and protruding the lower lip.

Raising of the upper lip is controlled by a series of levator muscles, which have insertions in the region of the upper lip. These are the ZYGOMATIC MINOR and the two LEVATOR LABII SUPERIOR muscles. The zygomatic minor runs from the cheekbone of the skull to the upper lips and orbicularis oris fibres. The levator labii superior muscles run from the maxilla bone to the medial part of the upper lip around the nasolabial groove. All these muscles contribute to lip raising when contracted.

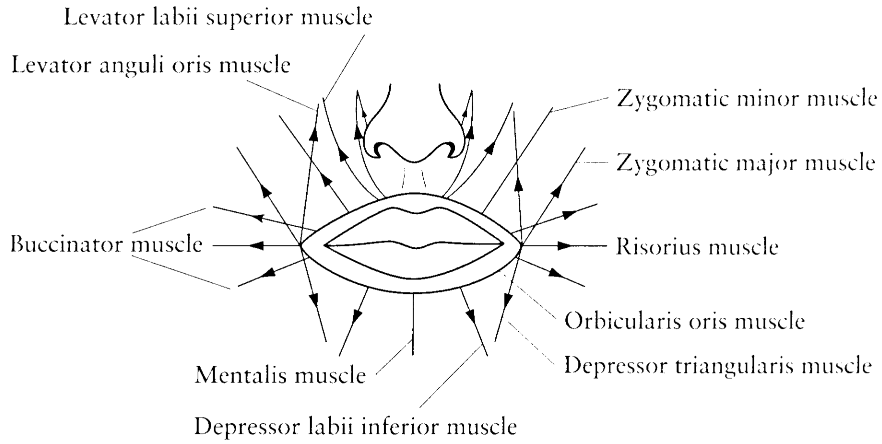


Figure 6.11.1 Labial muscles
Adapted from: Hardcastle 1976, p. 113.

Lowering the lower lip is controlled by the DEPRESSOR LABII INFERIOR muscle which runs from the anterior face of the jawbone to the lower lip, blending with the orbicularis oris.

Lateral movement of the mouth angle is controlled by the BUCCINATOR, RISORIUS and ZYGOMATIC MAJOR muscles. The buccinator, or bugler's muscle, is a thin sheet of muscle which runs via connective tissue from the lateral parts of the maxilla, jawbone and skull in the region of the back molars to blend with the orbicularis oris at the mouth angle. When contracted, it will draw the mouth angle back, spreading the lips. It also has an important function in maintaining tension in the cheeks during oral activity, including speech production. The name 'bugler's muscle' points to its role as an antagonist to distension of the cheeks during blowing or bugling. The risorius, which runs from the region of the lateral part of the jawbone to the lateral part of the lips and mouth angle, also contributes to the action of spreading the lips. The zygomatic major runs from the outer cheekbone to blend with the orbicularis oris at the mouth angle. When contracted, it contributes to drawing the mouth angle back and upwards.

Longitudinal mouth angle movement is controlled by the LEVATOR ANGULI ORIS, the ZYGOMATIC MAJOR (described in the preceding paragraph) and the DEPRESSOR TRIANGULARIS. The levator anguli oris runs from the lateral part of the maxilla and blends with the orbicularis oris at the mouth angle. When contracted, it will raise the mouth angle, as in laughing; the zygomatic major also contributes to this function. The depressor triangularis, so called because of its triangular shape, runs from the anterior lateral part of the mandible and blends with the orbicularis oris at the mouth angle.

Like the tongue muscles, the lip muscles operate in various combinations to yield a considerable range of lip configurations. Precise, rapid closure and release of the lips, as required in the articulation of labial stop sounds such as [p] and [b], may involve the action of the orbicularis oris to close and hold the lips together, and the levator and depressor muscles to open the lips rapidly at the release of the stop. For a sound such as the fricative [f], it is necessary to draw the lower lip against the upper teeth and to spread the lips. This may require the orbicularis oris to pull the lower lip inwards, and the buccinator, risorius and zygomatic major muscles to spread the lips by retracting the mouth angle. Certain vowels, as in English *board* or *talk*, have lip rounding and protrusion as part of their articulatory configuration, which may involve the orbicularis oris and the mentalis. Other vowels, as in *heed*, require the lips to be spread, which may involve the buccinator, risorius and zygomatic major muscles for lip spreading and the triangularis to maintain lip opening. These and other functions of the lip muscles during articulation are discussed in Hardcastle (1976) and Zemlin (1981). Kennedy and Abbs (1979) give a detailed account of labial musculature, and Abbs et al. (1984) offer some evidence that sections of the orbicularis oris may be activated independently. This independence would certainly contribute to the extent and versatility of labial movement control.

6.12 The mandible

The jaw or MANDIBLE does not play the same kind of role in speech production as the lips and tongue. If the lips or tongue are immobilized, speech is seriously impaired; by contrast, it is possible to produce quite intelligible speech with an object such as a pencil clenched between the teeth. The mandible does nevertheless function both as a moving articulator and as an important anchor point for a number of muscles which affect and are affected by its movement. It is approximately U-shaped with vertical extensions known as RAMI at each end of the U; these are heavy bone structures at the ends of which the mandible has joints with the skull base. The mandible is capable of movement in vertical, longitudinal and lateral directions. It may be lowered to produce a typical aperture of around 40 mm, protruded about 10 mm and moved laterally about 20 mm to either side (Heffner 1964, Zemlin 1981). Of these adjustments, vertical movement is the most important in speech; forward longitudinal movement plays a minor part; and lateral movement seems to make no significant contribution to normal articulatory processes.

In vertical movement, the mandible is lowered by the action of the mylohyoid, geniohyoid and digastricus muscles (section 6.5 above) and the genioglossus muscle (section 6.10 above). All of these muscles have attachments in the posterior face of the anterior part of the mandible and, if the hyoid bone is stable, will pull the mandible downwards when contracted. In addition, gravity will also contribute a downward force, as can be seen when the muscles relax and the jaw drops. Raising of the mandible is controlled by the INTERNAL PTERYGOID, TEMPORAL and MASSETER muscles. The internal pterygoid runs from the lateral part of the skull to the posterior of the ramus of the mandible.

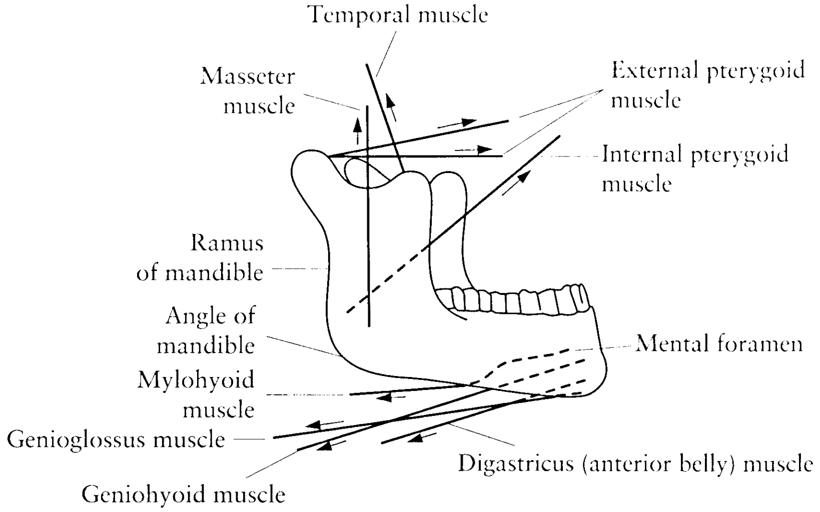


Figure 6.12.1 Muscles of the mandible
Adapted from: Hardcastle 1976, p. 107.

The temporal muscle runs from a wide area of the upper lateral part of the skull to the front of the upper end of the ramus of the mandible, while the masseter muscle runs from the lateral part of the cheekbone to most of the outer surface of the ramus of the mandible. Contraction of these muscles will raise the mandible, with the masseter being the most powerful of the three. Longitudinal movement or protrusion of the mandible is effected by the EXTERNAL PTERYGOID muscle, with some contribution from the internal pterygoid and masseter muscles just described. The external pterygoid runs from the area of the cheekbone to the posterior part of the extremity of the ramus of the mandible. Contraction of these muscles will pull the jaw forward, although in the case of the masseter and internal pterygoid muscles, forward movement is combined with vertical movement. Figure 6.12.1 shows the mandible and the muscle structure related to its movement during speech production.

Given an abnormal condition in which mandible movement is resisted or fixed, most speakers are able to make rapid and adequate compensatory manoeuvres in lip and tongue articulatory movements to produce highly intelligible speech, as Folkins and Abbs (1975) have shown. Under normal conditions, however, mandible movement and positioning are an important adjunct to certain classes of lip and tongue articulatory activity. Thus in vowel articulation, an increase in jaw aperture normally accompanies the lowering of the tongue for production of relatively low or open vowels (as in *ban* or *barn*). The role of jaw aperture in speech production is quantified in models of articulatory processes outlined by Lindblom and Sundberg (1971) and Coker (1973).

There is a connection between mandible movement and lip movement, particularly in such articulatory gestures as lip closure or constriction for stops and fricatives. Photographic studies by Fujimura (1961) have shown that mandible movement normally accompanies the opening of the lips at the release of a stop such as [p] (as in *pay*). Lindblom (1967) has also made indirect photographic measurements of the relationships between lip and jaw movement. Investigations of this sort show that mandible movement does not correspond directly with lip movement, but may lead or lag depending on the phonetic context. This is partly a matter of inertia, as the mandible has a greater mass than the lips. Since mandible movement is generally involved in setting the tongue position for low vowels, the mass of the mandible may help to explain why low vowels tend to be longer than high or close vowels.

Labial and mandibular movement have been extensively studied, not least because the lips and jaw can be fairly easily observed. Studies of lip and jaw motion have formed the basis of general models of articulatory kinematics (e.g. Kelso et al. 1985, Ostry et al. 1997), and a whole suite of studies in the area of articulatory prosody have examined lip and jaw data to uncover the articulatory correlates of linguistic stress and phrasing (e.g. Edwards et al. 1991, Harrington et al. 1995, Byrd 2000). Useful reviews of many of the techniques used in speech production research in general are provided by Bell-Berti and Raphael (1995), Stone (1997) and Hardcastle and Hewlett (1999, chs 10–15). Kent (2004) provides a concise summary of what is special about speech-related muscles. Readers wishing to consult more technical works on the anatomy and physiology of speech production are reminded of works already cited, notably Hardcastle (1976), Daniloff et al. (1980), Zemlin (1981), Dickson and Dickson (1982), Perkins and Kent (1986) and Maassen et al. (2004). Perkell (1997) provides a good discussion of articulatory processes in general, and Löfqvist (1997) includes a concise summary of gestural-based speech production models. The colour photographs of relevant anatomical sections and specimens in McMinn and Hutchings (1988) may also be helpful.

Exercises

1 Briefly describe the following and their function(s) in speech production.

a. the tongue and its muscles
b. the velum
c. the vocal folds
d. the lungs

2 Check that you understand the meaning of the following and their relevance to speech.

a. the innervation ratio of a muscle
b. the Bernoulli effect
c. subglottal pressure
d. the aerodynamic myoelastic theory of phonation
e. the opening quotient of the vocal folds

3 What are the cartilages of the larynx, how do they move, and how do they contribute to speech?

4 What are the CNS, PNS and autonomic nervous system?

5 Most of the various organs used in producing speech have other functions, sometimes considered their 'primary' functions. Identify these primary functions for as many organs of speech as you can. Can you justify the use of the word 'primary' here?

6 How would you explain or demonstrate the following to an introductory phonetics class, without using complex equipment?

a. that air flows out through the nose during the production of nasal consonants
b. that jaw movement is less important in speech than tongue and lip movement
c. that we normally need feedback to speak fluently and intelligibly
d. that most English speech sounds require airflow from the lungs