

4 Voicing and consonants

4.1 The larynx

We begin this chapter by studying the **larynx**. The larynx has several very important functions in speech, but before we can look at these functions we must examine its anatomy and physiology – that is, how it is constructed and how it works.

The larynx is in the neck; it has several parts, shown in Fig. 10. Its main structure is made of **cartilage**, a material that is similar to bone but less hard. If you press down on your nose, the hard part that you can feel is cartilage. The larynx's structure is made of two large cartilages. These are hollow and are attached to the top of the **trachea**; when we breathe, the air passes through the trachea and the larynx. The front of the larynx comes to a point and you can feel this point at the front of your neck – particularly if you are a man and/or slim. This point is commonly called the **Adam's Apple**.

Inside the “box” made by these two cartilages are the **vocal folds**, which are two thick flaps of muscle rather like a pair of lips; an older name for these is **vocal cords**. Looking down the throat is difficult to do, and requires special optical equipment, but Fig. 11 shows in diagram form the most important parts. At the front the vocal folds are joined together and fixed to the inside of the **thyroid cartilage**. At the back they are attached to a pair of

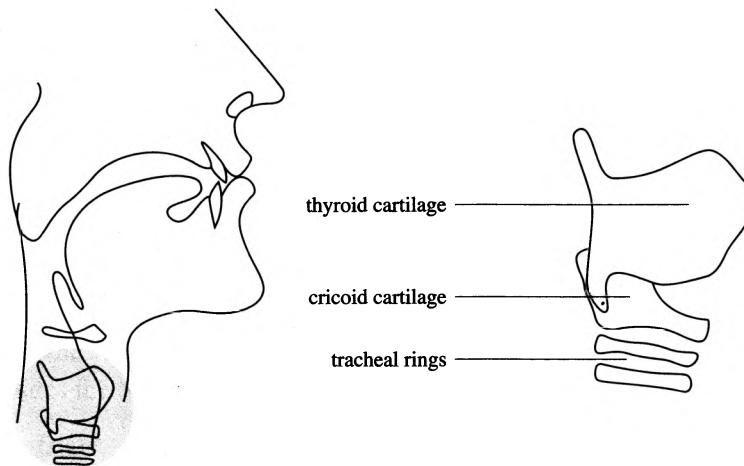


Fig. 10 The larynx

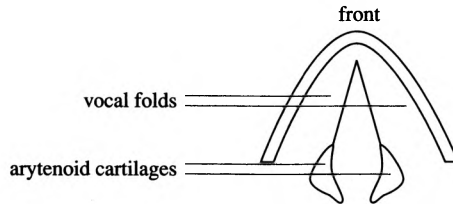


Fig. 11 The inside of the larynx seen from above

small cartilages called the **arytenoid cartilages** so that if the arytenoid cartilages move, the vocal folds move too.

The arytenoid cartilages are attached to the top of the **cricoid cartilage**, but they can move so as to move the vocal folds apart or together (Fig. 12). We use the word **glottis** to refer to the opening between the vocal folds. If the vocal folds are apart we say that the glottis is open; if they are pressed together we say that the glottis is closed. This seems quite simple, but in fact we can produce a very complex range of changes in the vocal folds and their positions.

These changes are often important in speech. Let us first look at four easily recognisable states of the vocal folds; it would be useful to practise moving your vocal folds into these different positions.

- i) **Wide apart:** The vocal folds are wide apart for normal breathing and usually during voiceless consonants like p, f, s (Fig. 13a). Your vocal folds are probably apart now.
- ii) **Narrow glottis:** If air is passed through the glottis when it is narrowed as in Fig. 13b, the result is a fricative sound for which the symbol is h. The sound is not very different from a whispered vowel. It is called a **voiceless glottal fricative**. (Fricatives are discussed in more detail in Chapter 6.) Practise saying hahahaha – alternating between this state of the vocal folds and that described in (iii) below.
- iii) **Position for vocal fold vibration:** When the edges of the vocal folds are touching each other, or nearly touching, air passing through the glottis will usually cause vibration (Fig. 13c). Air is pressed up from the lungs and this air pushes the vocal folds apart so that a little air escapes. As the air flows quickly past the edges of

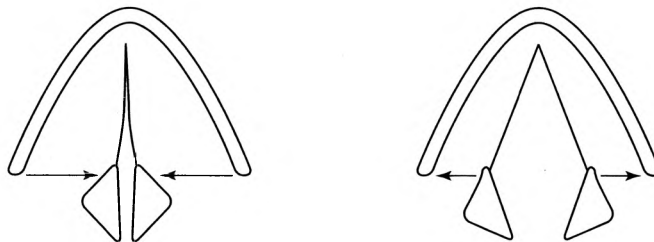


Fig. 12 Arytenoid cartilages causing closing and opening of the glottis

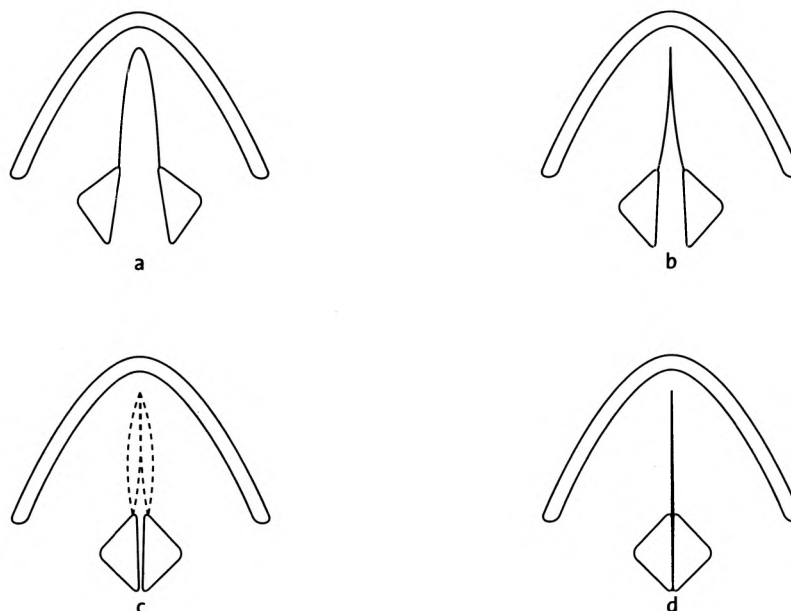


Fig. 13 Four different states of the glottis

the vocal folds, the folds are brought together again. This opening and closing happens very rapidly and is repeated regularly, roughly between two and three hundred times per second in a woman's voice and about half that rate in an adult man's voice.

- iv) Vocal folds tightly closed: The vocal folds can be firmly pressed together so that air cannot pass between them (Fig. 13d). When this happens in speech we call it a **glottal stop** or **glottal plosive**, for which we use the symbol ʔ . You can practise this by coughing gently; then practise the sequence $a\text{ʔ}a\text{ʔ}a\text{ʔ}a\text{ʔ}a$.

4.2 Respiration and voicing

Section 4.1 referred several times to air passing between the vocal folds. The normal way for this airflow to be produced is for some of the air in the lungs to be pushed out; when air is made to move out of the lungs we say that there is an **egressive pulmonic airstream**. All speech sounds are made with some movement of air, and the egressive pulmonic is by far the most commonly found air movement in the languages of the world. There are other ways of making air move in the vocal tract, but they are not usually relevant in the study of English pronunciation, so we will not discuss them here.

How is air moved into and out of the lungs? Knowing about this is important, since it will make it easier to understand many aspects of speech, particularly the nature of stress and intonation. The lungs are like sponges that can fill with air, and they are contained within the rib cage (Fig. 14). If the rib cage is lifted upwards and outwards there

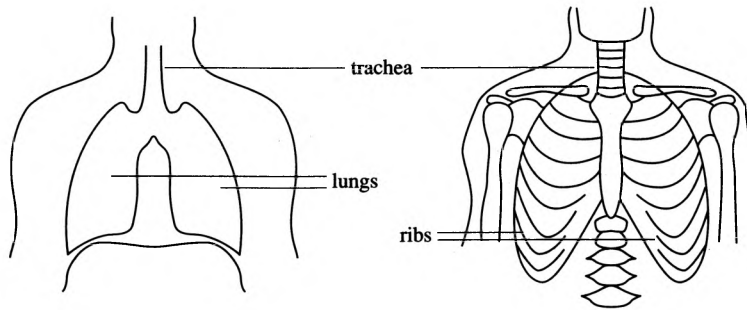


Fig. 14 The lungs and the rib cage

is more space in the chest for the lungs and they expand, with the result that they take in more air. If we allow the rib cage to return to its rest position quite slowly, some of the air is expelled and can be used for producing speech sounds. If we wish to make the egressive pulmonic airstream continue without breathing in again – for example, when saying a long sentence and not wanting to be interrupted – we can make the rib cage press down on the lungs so that more air is expelled.

In talking about making air flow into and out of the lungs, the process has been described as though the air were free to pass with no obstruction. But, as we saw in Chapter 2, to make speech sounds we must obstruct the airflow in some way – breathing by itself makes very little sound. We obstruct the airflow by making one or more obstructions or **strictures** in the vocal tract, and one place where we can make a stricture is in the larynx, by bringing the vocal folds close to each other as described in the previous section. Remember that there will be no vocal fold vibration unless the vocal folds are in the correct position and the air below the vocal folds is under enough pressure to be forced through the glottis.

If the vocal folds vibrate we will hear the sound that we call **voicing** or **phonation**. There are many different sorts of voicing that we can produce – think of the differences in the quality of your voice between singing, shouting and speaking quietly, or think of the different voices you might use reading a story to young children in which you have to read out what is said by characters such as giants, fairies, mice or ducks; many of the differences are made with the larynx. We can make changes in the vocal folds themselves – they can, for example, be made longer or shorter, more tense or more relaxed or be more or less strongly pressed together. The pressure of the air below the vocal folds (the **subglottal pressure**) can also be varied. Three main differences are found:

- i) Variations in **intensity**: We produce voicing with high intensity for shouting, for example, and with low intensity for speaking quietly.
- ii) Variations in **frequency**: If the vocal folds vibrate rapidly, the voicing is at high frequency; if there are fewer vibrations per second, the frequency is lower.
- iii) Variations in **quality**: We can produce different-sounding voice qualities, such as those we might call *harsh*, *breathy*, *murmured* or *creaky*.

4.3 Plosives

A **plosive** is a consonant articulation with the following characteristics:

- a) One articulator is moved against another, or two articulators are moved against each other, so as to form a stricture that allows no air to escape from the vocal tract. The stricture is, then, total.
- b) After this stricture has been formed and air has been compressed behind it, it is **released** – that is, air is allowed to escape.
- c) If the air behind the stricture is still under pressure when the plosive is released, it is probable that the escape of air will produce noise loud enough to be heard. This noise is called **plosion**.
- d) There may be voicing during part or all of the plosive articulation.

To give a complete description of a plosive consonant we must describe what happens at each of the following four phases in its production:

- i) The first phase is when the articulator or articulators move to form the stricture for the plosive. We call this the **closing phase**.
- ii) The second phase is when the compressed air is stopped from escaping. We call this the **compression phase**.
- iii) The third phase is when the articulators used to form the stricture are moved so as to allow air to escape. This is the **release phase**.
- iv) The fourth phase is what happens immediately after (iii), so we will call it the **post-release phase**.

4.4 English plosives

English has six plosive consonants: p, t, k, b, d, g. The glottal plosive ʔ occurs frequently but it is of less importance, since it is usually just an alternative pronunciation of p, t, k in certain contexts. The plosives have different places of articulation. The plosives p, b are bilabial since the lips are pressed together (Fig. 15); t, d are alveolar since the tongue blade is pressed against the alveolar ridge (Fig. 16). Normally the tongue does not touch the front teeth as it does in the dental plosives found in many languages. The plosives k, g are velar; the back of the tongue is pressed against the area where the hard palate ends and the soft palate begins (Fig. 17).

The plosives p, t, k are always voiceless; b, d, g are sometimes fully voiced, sometimes partly voiced and sometimes voiceless. We will consider what b, d, g should be called in Section 4.5 below.

All six plosives can occur at the beginning of a word (**initial position**), between other sounds (**medial position**) and at the end of a word (**final position**). To begin with we will look at plosives preceding vowels (which can be abbreviated as CV, where C stands for a consonant and V stands for a vowel), between vowels (VCV) and following vowels (VC). We will look at more complex environments in later chapters.

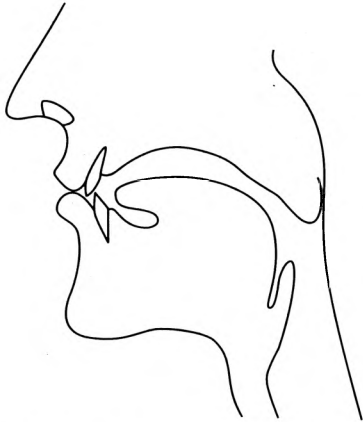


Fig. 15 Bilabial articulation

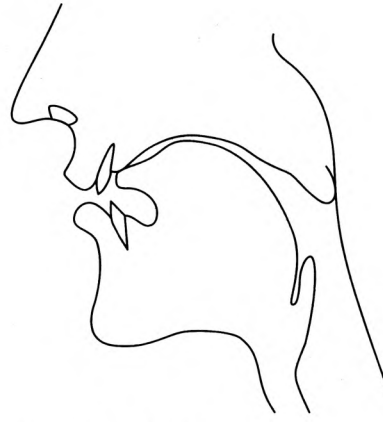


Fig. 16 Alveolar articulation

Ⓞ AU4 (CD 1), Ex 1

- i) Initial position (CV): The closing phase for p, t, k and b, d, g takes place silently. During the compression phase there is no voicing in p, t, k; in b, d, g there is normally very little voicing – it begins only just before the release. If the speaker pronounces an initial b, d, g very slowly and carefully there may be voicing during the entire compression phase (the plosive is then fully voiced), while in rapid speech there may be no voicing at all.

The release of p, t, k is followed by audible plosion – that is, a burst of noise. There is then, in the post-release phase, a period during which air escapes through the vocal folds, making a sound like h. This is called **aspiration**. Then the vocal folds come together and voicing begins. The release of b, d, g is followed by weak plosion, and this happens at about the same time as, or shortly after, the beginning of voicing. The most noticeable and important difference, then, between initial p, t, k and b, d, g is the aspiration of the voiceless plosives p, t, k. The different phases of the plosive all happen very

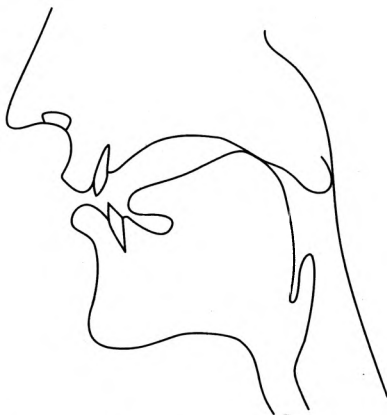


Fig. 17 Velar articulation

rapidly, but the ear distinguishes clearly between p, t, k and b, d, g. If English speakers hear a fully voiced initial plosive, they will hear it as one of b, d, g but will notice that it does not sound quite natural. If they hear a voiceless unaspirated plosive they will also hear that as one of b, d, g, because it is aspiration, not voicing which distinguishes initial p, t, k from b, d, g. Only when they hear a voiceless aspirated plosive will they hear it as one of p, t, k; experiments have shown that we perceive aspiration when there is a delay between the sound of plosion and the beginning (or **onset**) of voicing.

In initial position, b, d, g cannot be preceded by any consonant, but p, t, k may be preceded by s. When one of p, t, k is preceded by s it is unaspirated. From what was said above it should be clear that the unaspirated p, t, k of the initial combinations sp, st, sk have the sound quality that makes English speakers perceive a plosive as one of b, d, g; if a recording of a word beginning with one of sp, st, sk is heard with the s removed, an initial b, d or g is perceived by English speakers.

- ii) **Medial position (VCV):** The pronunciation of p, t, k and b, d, g in medial position depends to some extent on whether the syllables preceding and following the plosive are stressed. In general we can say that a medial plosive may have the characteristics either of final or of initial plosives.

⌚ AU4 (CD 1), Exs 2 & 3

- iii) **Final position (VC):** Final b, d, g normally have little voicing; if there is voicing, it is at the beginning of the compression phase; p, t, k are always voiceless. The plosion following the release of p, t, k and b, d, g is very weak and often not audible. The difference between p, t, k and b, d, g is primarily the fact that vowels preceding p, t, k are much shorter. The shortening effect of p, t, k is most noticeable when the vowel is one of the long vowels or diphthongs. This effect is sometimes known as **pre-fortis clipping**.

4.5 Fortis and lenis

Are b, d, g voiced plosives? The description of them makes it clear that it is not very accurate to call them “voiced”; in initial and final position they are scarcely voiced at all, and any voicing they may have seems to have no perceptual importance. Some phoneticians say that p, t, k are produced with more force than b, d, g, and that it would therefore be better to give the two sets of plosives (and some other consonants) names that indicate that fact; so the voiceless plosives p, t, k are sometimes called **fortis** (meaning ‘strong’) and b, d, g are then called **lenis** (meaning ‘weak’). It may well be true that p, t, k are produced with more force, though nobody has really proved it – force of articulation is very difficult to define and measure. On the other hand, the terms fortis and lenis are difficult to remember. Despite this, we shall follow the practice of many books and use these terms.

The plosive phonemes of English can be presented in the form of a table as shown here:

	PLACE OF ARTICULATION		
	Bilabial	Alveolar	Velar
Fortis (“voiceless”)	p	t	k
Lenis (“voiced”)	b	d	g

Tables like this can be produced for all the different consonants. Each major type of consonant (such as plosives like p, t, k, fricatives like s, z, and nasals like m, n) obstructs the airflow in a different way, and these are classed as different **manners of articulation**.

Notes on problems and further reading

4.1, 4.2 For more information about the larynx and about respiration in relation to speech, see Raphael *et al.*, (2006); Laver (1994: Chapters 6 and 7); Ashby and Maidment (2005: Chapter 2).

4.3 The outline of the stages in the production of plosives is based on Cruttenden (2008: 158). In classifying consonants it is possible to go to a very high level of complexity if one wishes to account for all the possibilities; see, for example, Pike (1943: 85–156).

4.4 It has been pointed out that the transcription sb, sd, sg could be used quite appropriately instead of sp, st, sk in syllable-initial position; see Davidsen-Nielsen (1969). The vowel length difference before final voiceless consonants is apparently found in many (possibly all) languages, but in English this difference – which is very slight in most languages – has become exaggerated so that it has become the most important factor in distinguishing between final p, t, k and b, d, g; see Chen (1970). Some phonetics books wrongly state that b, d, g lengthen preceding vowels, rather than that p, t, k shorten them. The conclusive evidence on this point is that if we take the pair ‘right’ rait and ‘ride’ raid, and then compare ‘rye’ rai, the length of the ai diphthong when no consonant follows is practically the same as in ‘ride’; the ai in ‘right’ is much shorter than the ai in ‘ride’ and ‘rye’.

4.5 The *fortis/lenis* distinction is a very complicated matter. It is necessary to consider how one could measure “force of articulation”; many different laboratory techniques have been tried to see if the articulators are moved more energetically for fortis consonants, but all have proved inconclusive. The only difference that seems reasonably reliable is that fortis consonants have higher air pressure in the vocal tract, but Lisker (1970) has argued convincingly that this is not conclusive evidence for a “force of articulation” difference. It is possible to ask phonetically untrained speakers whether they feel that more energy is used in pronouncing p, t, k than in b, d, g, but there are many difficulties in doing this. A useful review of the “force of articulation” question is in Catford (1977: 199–208). I feel the best conclusion is that any term one uses to deal with this distinction (whether *fortis/lenis* or *voiceless/voiced*) is to be looked on as a **cover term** – a term which

has no simple physical meaning but which may stand for a large and complex set of phonetic characteristics.

Written exercises

- 1 Write brief descriptions of the actions of the articulators and the respiratory system in the words given below. Your description should start and finish with the position for normal breathing. Here is a description of the pronunciation of the word 'bee' bi: as an example:

Starting from the position for normal breathing, the lips are closed and the lungs are compressed to create air pressure in the vocal tract. The tongue moves to the position for a close front vowel, with the front of the tongue raised close to the hard palate. The vocal folds are brought close together and voicing begins; the lips then open, releasing the compressed air. Voicing continues for the duration of an i: vowel. Then the lung pressure is lowered, voicing ceases and the articulators return to the normal breathing position.

Words to describe: (a) goat; (b) ape.

- 2 Transcribe the following words:

a) bake	d) bought	g) bored
b) goat	e) tick	h) guard
c) doubt	f) bough	i) pea