Chapter 2: The Sounds of Language: Phonetics

1 Representing sounds: the International Phonetic Alphabet
The most basic problem that one faces when writing about speech sounds is how to represent them on paper (or, generally speaking, in any visual form). Sounds aren’t visible things, after all. For example, if I want to write about the first speech sound in the E word *cat*, what do I write on the sheet? Well, you may say it is quite easy: the word does have a written form, it begins with a *c*, so why not use the conventional spelling — in this case, the letter *c*?

That’s where the problems start. Think now of the word *Kate*: it starts with the same *sound*, albeit it is written differently. The spelling “pretends”, as it were, that the two words begin with different sounds, which, of course, isn’t the case. The underlined parts of the words *technical, trick, queen, khan* illustrate further occurrences of the same sound — spelt differently in each word. Or, think of *kept* and *peeped*, both words ending in the same consonant cluster, but having a radically different spelling. One problem that immediately faces us when trying to use conventional spelt forms to represent speech sounds is that very often, the same sound is written in a variety of ways. This is especially frequent in the case of English spelling, but we can observe it in Hungarian, too, cf. *Hu gőlya and bőja, or skicc, látz, and ufca.*

The opposite can also happen: two or more different sounds can be spelt identically. English examples include the letter *th*, as in *through vs. them*, or the *e* in *set, scene, and term*, for instance, but the plural ending *(e)s* is a fine example, too, cf. *stones, bats, and boxes*. In Hungarian writing, this happens much less frequently, but it does occur, cf. the *n* in *nagy vs. lengyel* (in the latter, it represents the same sound as the *ny* in *lány*).

To sum up, conventional *orthographies* (= spelling systems) are inadequate for linguistic use, because they fail to represent sounds unambiguously — they often violate the “one sound — one letter” principle (in both directions), a principle we do need as linguists. But this is not all to it. Orthographies are the products of cultural tradition (often centuries or millennia old), differing from place to place. In the western part of Europe, for instance, the Roman alphabet (itself a modified form of the Greek script) has been in use for a long time; in Eastern Europe, several cultures use the Cyrillic script, another modification of the Greek alphabet. Now, Serbs and Croats, who speak basically the same language (with very small differences), use different alphabets. Look at the (Serbian) sentence below (meaning ‘Vera is a student of French language and literature’):

(1) Вера је студентица француског језика и књижевности.

This sentence becomes much less “exotic” for the “Western” eye if we write it down in the Croatian way — i.e., using Roman letters:

(2) Vera je studentica francuskog jezika i književnosti.

How different these two sentences look! Yet, they are in fact the same sentence, pronounced perfectly identically and meaning exactly the same! This leads us to the second important point why we cannot rely on conventional orthographies: they use different scripts.

But even within the same basic tradition there can be significant differences. Consider, for example, the first sound of the Hungarian word *só*. The same consonant occurs in a wide range of languages (and the majority of European ones). Now, even if we disregard the Cyrillic alphabet, and remain strictly within the Roman-based orthographic tradition, we find...
an enormous variety of different spellings for this poor consonant. Here’s a selection, just to give you a taste of it:

(3)

- English: generally sh (ship) but also ti (nation), ch (machine);
- German: mostly sch (Schiff ‘ship’), but often s (stehlen ‘to steal’);
- French: ch (chat ‘cat’);
- Portuguese: mostly ch (chegar ‘to arrive’), but often s (estar ‘to be’), sometimes even x (peixe ‘fish’);
- Italian: sc (nasce ‘(s)he is born’);
- Welsh: si (siarad ‘to speak’);
- Czech (and some other Slavonic orthographies): š (šest ‘six’);
- Polish: sz (Warszawa ‘Warsaw’);
- Romanian: ș (pește ‘fish’).

Imagine now the confusion of the linguist who wishes to denote this single sound in a single manner! Which one is he/she to choose?

But there are still worse things. First, many languages use writing systems which do not represent sounds but syllables or even complete signs (irrespective of sound shape). Chinese is a good example for a word writing system, in which one symbol represents a word (it must be added that Chinese words are generally simplex, i.e., they consist of one morpheme); Japanese uses both this type and a syllabic writing. In either case, there’s no way a single consonant can be written down! Second, most languages are not written in any way! This may seem odd to you, but remember that there are between 4,000 to 6,000 languages in the world, and the majority of them are spoken by communities which either do not use writing, or, more typically, they use another language for the purposes of written communication (this isn’t very unusual in the history of Europe, including Hungary, either — think of the almost exclusive use of Latin in writing for many centuries). If a linguist wants to write about the speech sounds of such a language, there is no way he/she can rely on the conventional orthography (for there is none)!

To sum up, traditional writing systems are inadequate for our purposes. For this reason, British and French phoneticians founded a society called the International Phonetic Association (IPA) — Association Phonétique Internationale (APhI) in French — at the end of the 19th century, with the explicit aim of devising a unique alphabet, where there is a symbol for each speech sound that exists in a human language. This alphabet was named the International Phonetic Alphabet — abbreviated, just for the sake of fun, IPA (and yes, APhI in French). In this book, we will use the abbreviation IPA to refer to the alphabet, rather than the society (which is the usual practice, anyway).

The IPA symbol used to denote the consonant in Hungarian só is ʃ; and, of course, we use the same symbol when talking about the (same) sound in any word in any language. So, here are some words from the list in (3), written down in their traditional spelt form as well as using the appropriate symbols of the IPA. (The symbol ə represents the vowel pronounced at the beginning of the English word ago.) Note that wherever there is a ʃ in the IPA-written forms, the same consonant as in só is pronounced:
When we write down the word in this way, using IPA symbols to show its pronounced form, we transcribe the word, and the result is a transcription. So, for example, the English word *ship* is transcribed as /ʃip/. If two words are pronounced identically, they are also transcribed identically (the E words *scene* and *seen* are both *sin*, for instance), even in two different languages (cf. Hu *szín* = *sin*; note that the symbol : indicates that the preceding symbol represents a long sound).

It is often necessary — for the sake of clarity — to show that a given form is a transcribed one; traditionally, transcriptions are enclosed between square brackets, e.g., *[ʃip]*, *[sin]*, etc., while conventional spelt forms are written between outward pointing arrowheads, e.g., `<ship>`, `<scene>`, if need be. Often, however, these bracketings are omitted, if it is quite clear what one talks about.\(^1\)

### 2 Choice of accent

Before we start our discussion of English speech sounds, one thing must be clarified: what kind of English do we describe? After all, English has spread too far and wide to be uniform: it has many variants, especially in pronunciation. As all advanced students of English will have noticed, a Londoner speaks with a different accent than someone from Chicago or Dublin or Sydney, etc. One type of difference between accents of a language is what speech sounds they use. We can illustrate this with a well-known Hungarian example, the short a sound of the Palóc accent of Northern Hungary, as in the words *alma, Jani*, etc., not found in other accents. What about English?

English has two standard accents used widely in teaching English as a foreign language, standard American and standard British English. Since we are mostly foreign learners of English, rather than native speakers, we have probably been acquainted with one of these accents in our school experience (unless taught by a native Australian or Irish speaker, for example); moreover, those who will teach English will also use one of them. It is useful, therefore, to choose one of them as our reference accent, i.e., the accent we will describe during this course.

The accent described here will be the standard accent used in England, known in everyday life as “British accent”. (In reality, there’s no single British accent: people in Britain use a wide range of accents, often radically different from each other as well as from the standard one.) Let us, therefore, say a few words about it. The standard accent in England is known as Received Pronunciation (abbreviated RP) — this is basically what people mean when they talk about a “British accent”. RP is the accent used in England by educated speakers in schools, in the media, in theatres, and all relatively formal public situations. As a result, it is a socially and geographically neutral accent in the sense that if someone speaks RP, no one can tell his/her social/geographical background: it is just like standard Hungarian in this respect.

\(^1\) In fact, another method is used to show transcribed forms, namely, slashes, e.g., /ʃip/ — this is what you find in the Oxford Advanced Learner’s Dictionary, for instance. In fact, there is an important difference in using brackets or slashes, a problem we will discuss in the next chapter.
As opposed to standard Hungarian, though, RP is quite a learned accent: while standard Hungarian is used by many (if not most) Hungarians as a native accent, acquired in early childhood without formal education, few people in Britain are actually native speakers of RP. The majority of RP speakers have a different native accent, and they learn RP during formal education (= at school). Of the native accents of England, RP is closest to those of the South-East of England, more precisely, the general colloquial accent of London.\(^2\)

RP itself is not homogenous, though, and two varieties are generally distinguished. The first one is characteristic of older or more conservative speakers and is termed Conservative RP; younger, less conservative speakers use a slightly different variant called Advanced RP. This book will concentrate on the latter, since it is by far more widely used by younger generations and is likely to displace ConsRP altogether within one or two generations’ time; as most students of this subject are young people, I find this better than describing ConsRP. Nonetheless, the main differences between Cons and Adv RP will be pointed out, but it should be noted that the differences between the two accents are very few\(^3\) — so few indeed that it makes perfect sense to talk about “RP as such”.

The standard accent used in the United States is known as General American (GA). Like RP in England, GA is the accent used in formal contexts, in education, in the media, etc. There are two important points in which the social position of GA differs from that of RP. First, GA is the native accent of a large number of Americans. Second, it is more heterogenous than RP, having a number of variants itself, most of them connected to a particular area or city.

You may now ask why it is RP that we’re describing in this course. The answers are simple and practical. First, RP is probably the best described English accent, taken as a model by the representatives of the influential British phonetic school since the late 19\(^{th}\) century. Second, most textbooks and dictionaries aimed at foreign learners of English in Europe describe RP: this is due partly to the (now, thankfully, receding) tradition in European practice which regarded RP as the only “correct” way of speaking English, partly to the dominant role that British universities, especially Cambridge and Oxford, play in English language teaching and textbook publishing. Third, the most subjective but quite practical reason is that the author of the present book has always learnt RP and knows this accent in more detail than GA (or other accents).

It must be emphasised that RP is in no way a better, more pleasing, elegant, more educated or “original” accent of English than GA. Non-linguists, including many Americans, tend to believe that GA is a “corrupted” version of RP, the latter being the “English of Shakespeare”, spoiled by American laziness and ignorance. Needless to say, these are all mistaken judgments based on cultural prejudice and ignorance of scientific facts. Both RP and GA are descendants of early 17\(^{th}\) century London English, and both have changed since then, sometimes in different ways, but neither RP nor GA (or any other present-day accent, for that matter) is the way Shakespeare spoke. (In fact, GA is in many ways much closer to Shakespeare’s accent!) Conversely, many Americans regard RP as an affected, posh way of speaking. This view, of course, is not more serious than the opposite: it is equally just a matter of prejudice and individual taste.

\(^2\) This accent, widely used all around London, is not the same as Cockney, which is a socially as well as geographically much more restricted accent, more different from RP. More about this in the chapter on the accents of English.

\(^3\) Where they differ, the general tendency is that AdvRP is closer to colloquial London English.
3 Articulation: how speech sounds are produced
The branch of science that deals with the production of speech sounds — that is, articulation — is called articulatory phonetics. For example, it is an articulatory phonetic statement that the consonant b is produced by the two lips pressed together. Phonetics, the study of speech sounds, has two other branches besides articulatory phonetics: acoustic phonetics and auditory phonetics. The former deals with speech sounds as sound waves spreading in the air. As any sound wave, speech sounds have a characteristic acoustic pattern (frequency, for example). As is clear from this, acoustic phonetics is really a branch of physics. Auditory phonetics studies the way speech sounds are perceived. In this course, apart from some notes, we are not going to deal with acoustic and auditory phonetics in detail, and we’ll use the term phonetics in the sense “articulation”. This is not to say that acoustic or auditory phonetics are uninteresting, but neither is indispensable for an understanding of the essentials of phonology, i.e., how sounds are utilised in language. Of course, from now on, we will use IPA symbols to denote speech sounds.

Please pay special attention to this chapter and the next. All through the rest of the course, terms and concepts introduced and explained here will be used extensively. It will be taken for granted that you are familiar with the basic terms of (articulatory) phonetics, because a solid knowledge of them is indispensable for studying phonology, and — to a lesser extent, though — morphology.

* * * * *

In order to produce sounds, a body of air must be set in motion, it must vibrate — this vibration is what we call a wave of sound. In most cases — and always so in E and Hu — this body of air is provided by a pulmonic egressive airstream mechanism. Don’t be frightened of this scientific-sounding expression, it means something simple. The term airstream mechanism refers to how the air is set into motion in order to produce sounds; pulmonic means ‘of the lungs’, and egressive means ‘outward going, leaving’. Altogether, this whole expression refers to the fact that speech sounds are produced by using the air coming out of the lungs. This first stage in the process of sound production — setting the air in motion — is referred to as the initiation process.4

Of course, the body of air leaving the lungs does not in itself produce any particular effect (this is what happens when you breathe out). In order to produce speech sounds, the air must be modified while leaving the mouth. Those organs that perform this task are called speech organs. The first speech organ in the path of the air leaving the lungs is the larynx, also called Adam’s apple, where the vocal cords (also called vocal folds) are found. When simply breathing, the vocal cords are apart, letting the air pass through the larynx freely. The vocal cords, however, can be pushed close together, in which case the air passes through them, making them vibrate. The effect of this vibration is one of the most important features of speech production: voice. Languages use both voice and the lack of it in the production of speech sounds. If the vocal cords vibrate while the given speech sound is articulated, the sound is voiced; if not, it’s voiceless. This second stage of sound production — whether voice is produced in the larynx or not — is called the phonation process. There is a very simple practical method to see if a sound is voiced or not. Put the tip of your fingers on your larynx, and if you can feel it vibrate, the sound is voiced. Compare, for instance the initial consonants of sip and zip: if you produce s, you cannot feel anything with your fingertips, but the production of z makes the larynx vibrate. Give it a try!

Figure 1.1  The organs of speech
To the right: The glottis seen from above. In the picture at the top, the vocal folds are close together, producing voice. Below, they are apart, letting the air pass through without voice.

The parts above the larynx where the air can pass through is called the vocal tract. Upon leaving the larynx, the air passes through the pharynx. English and Hungarian do not use the pharynx in speech production, so we can neglect this part here. Above the pharynx, the air can go two different ways: towards the nasal cavity alone or the nasal and oral cavities. The organ responsible for directing the air this or that way is the velum, i.e., the soft, back part of the palate (the very end of the velum, the uvula, can be seen if you open your mouth wide and look into the mirror: it is the little organ which is hanging down in the back of your mouth). During normal breathing, the velum is lowered, letting the air pass through towards the nasal cavity (as well as the mouth). Most speech sounds, however, are produced with a raised velum, which cuts off the passage of the air to the nose and directs all of it towards the oral cavity; we say that most speech sounds are oral. There are some, nevertheless, which are nasal, i.e., during their articulation the velum is lowered and the air is free to pass through the nose. Such a sound is m, as in may, as opposed to, say, b (as in bay) produced with the lips closed tightly, just like m, but with a raised velum — that is, without the air being able to escape through the nose. Lowered velum, and its effect, nasality, is another important articulatory gesture. This third stage of articulation, where the nasal vs. oral distinction is produced, is called the oro-nasal process.
We have now come to the final, fourth, stage of sound production: the **articulation process**. The air passes through the mouth while certain organs, the **articulators** (including the tongue, the lips, the teeth, etc.; we’ll see them in detail later) perform various gestures in order to modify the passage of the air, resulting in various acoustic effects. Depending on what type of gesture the articulators perform, we talk about different **manners of articulation**. The most basic distinctions to be made between speech sounds on the basis of their manner of articulation are as follows; note that these criteria are partly defined on an acoustic, partly on an articulatory basis:

To the left: **States of the velum: raised** (above); **lowered** (below).
1) *vowels vs. consonants;*  
2) *sonorants vs. obstruents;*  
3) *continuants vs. stops.*

Let us now see these divisions in detail.

1) The division between consonants and vowels is one of the most elusive ones in linguistic science. In everyday life, we tend to think of this distinction as a simple and straightforward one, but as it often happens in science, what seems to be simple turns out to be complicated and far from unambiguous.

First of all, it must be borne in mind that in everyday practice (including primary and secondary education), people use the terms “consonant” and “vowel” when they really mean “a letter representing a consonant sound” or “a letter representing a vowel sound”. Of course, whether a letter represents a vowel or a consonant depends on local orthographic tradition, much like spelling in general. The letters <y> and <w> are especially notorious in the Latin-based writing tradition, variably representing vowels or consonants. The letter <w>, for example, represents in German the same sound as E or Hu <v> (G <weit> ‘far’ sounds almost the same as Hu <vájt>), clearly a consonant sound; in Welsh, <w> is the regular spelling for what is spelt in Hu as <u> or <û> (Welsh writing doesn’t indicate the difference in length), so that Welsh <brwd> ‘keen’ sounds like E <brood>, and <bws> ‘bus’ is pronounced just like Hu <busz>. German kids are taught that <w> is a consonant; Welsh children are told that it’s a vowel. In fact, it is neither: it’s but a graphic form to represent a sound. In English school practice, <w> is called a “consonant”, implying that it represents a consonant sound, which it can (as in <wait>), but very often it doesn’t, as in <law> or <few>, yet, English-speaking children are taught that <w> is a “consonant”. In everyday life, such a simplification is not a problem; however, as students of linguistics, we must keep the distinction between sounds and letters in mind all the time.

The linguistic distinction, based on sounds, is rather complex. This is because there are two quite different approaches to how we define the difference between vowels and consonants: a phonetic (articulatory/acoustic) and a functional (phonological) one. The two often do not coincide.

The phonetic difference between **consonants and vowels** is based on **whether the airflow is obstructed or not, and if it is, to what extent.** Let us see this in detail. If you pronounce the word *pie,* observing the first sound of it (represented in the IPA by the symbol p), you will notice that it is produced by the lips which make a complete closure, blocking the way of the airflow, i.e., obstructing it; then the lips burst apart, letting the air out suddenly. Compare this to the last sound of *spa* (the phonetic symbol for it is α: the symbol ı indicates that it’s pronounced long, rather than short): the mouth is wide open, the air escapes freely. In case the articulation involves complete closure, we talk about consonants; if there is no obstruction, we talk about vowels. But what if there is some obstruction, but not a complete colsure, i.e., the airflow is obstructed but not blocked completely? In such cases, the decision in phonetics is made on the basis of the **presence vs. absence of audible friction.** For example, pronounce the initial sound of the word *fine* (transcribed as f in the IPA). During its production, the lower lip touches the upper teeth; but, unlike for p, there is no complete closure. There is a very narrow gap left, through which the air can escape; this is why you can produce a f for a long time (as long as you have enough air in your lungs). Now, pronounce the first sound of the word *wet* (symbolised in the IPA by w): there is again a narrow gap, made by both lips (with the back of the tongue raised at the same time). But there is an important difference between the two sounds. During the articulation of f, the narrowing is close enough to make the airflow turbulent, causing audible friction. As opposed to this, the production of w involves no
friction: the narrowing isn’t close enough for making the airflow turbulent. If there’s audible friction, phoneticians define the sound as a consonant; otherwise, it’s a vowel. Phonetically speaking (i.e., on the basis of articulation and acoustic effects), w is a vowel. This probably strikes you as odd, but try to compare it to the initial sound of the word ooze, which you, I hope, readily identify as a vowel (transcribed as u: in the IPA). It, too, is produced by the two lips approaching each other (and the simultaneous raising of the back of the tongue) — there’s hardly any difference in articulation! Indeed, if you try to produce the w sound for a long time, you will get something like an u:!

This is where the question of function comes in. Several sounds — such as w — are frictionless, but are produced with a significant narrowing of the vocal tract. They are called approximants in phonetics (some authors use the term frictionless continuant). Approximants are — because frictionless — vowel-like in the phonetic sense, but they mainly function in the same way as “real” consonants in the phonological sense — i.e., in what part they play in the sound structure of words. For example, unlike vowels, they do not necessarily form a syllable: wet is one syllable, rather than two. In English, there are four approximants, viz. w, j, l, r, as in wet, you, low, ray, resp. Of these, w and j are grouped together as glides (also called semivowels sometimes); l and r are called liquids. We’ll come back to all of them in detail later on; for the time being note that approximants are classified as consonants due to phonological (= functional) reasons, rather than on a phonetic basis. From this point on, we will consider approximants to be consonants, so whenever you meet the term “vowel”, it means a sound without an obstruction to the airflow, a vowel in the classical sense.

2) The distinction between sonorants and obstruents is based on the relative amount of voice in the acoustic content of the speech sound. Remember that speech sounds — as any sound — consist of sound waves. Sound waves are partly produced by the vibration of the vocal cords, that is, by voicing. This is the case with vowels (such as a:) — they are pure, 100% voice, that is the sound waves you hear when hearing a vowel are all produced by the vibration of the vocal cords. Consonants, however, also contain sound waves produced by some obstruction, either complete or partial. The difference between sonorants and obstruents lies in the proportion of sound waves produced by voicing vs. those produced by the obstruction: a sonorant is a sound having a higher proportion of voice, i.e., it predominantly consists of the sound waves produced by voicing. Obstruents, on the other hand, are produced by an obstruction (narrowing or closure) which is significant enough to produce a noise stronger than voicing — an obstruent has a low proportion of voice in it: in other words, it predominantly consists of sound waves produced not by voicing but obstruction, i.e., noise.

Two simple facts follow from this right away:

First, vowels are all obviously sonorants, since their articulation involves no obstruction to the airflow — vowels are 100% voice. (It is no accident that we do not come across voiceless vowels!) Approximants, which phonetically are not consonants, are also sonorants, produced by a weak obstruction, — the noise produced by the obstruction is little. They are not pure voice, but they contain more voice than noise. Second, it is equally obvious that all voiceless consonants are obstruents: they have no amount of voice in them, they are 100% noise. But what about voiced consonants? Are they sonorants or obstruents?

Voiced consonants other than approximants can be classified into two groups according to the oro-nasal process: oral and nasal ones. Nasal ones, as described above, are produced with a lowered velum. They involve a complete closure in the oral cavity (think of m, for instance, pronounced with the lips closed), but, due to the lowered velum, the air passes through the nasal cavity. In the nasal cavity, of course, no obstruction is present. Nasal consonants are sonorants: they do contain some noise, but the voice element is stronger. Non-nasal (= oral) voiced consonants (approximants excepted, of course!) are obstruents; such consonants are
b, v, d, z, etc., as in bay, very, dime, zeal: they do contain some voice, but the noise component is stronger in them.

We can illustrate the proportion of voice in classes of sounds in terms of a scale, known as the sonority scale: sonority means “proportion of voice”; the more sonority a sound contains, the more sonorous it is. You can find a simple sonority scale in (5):

(5) A simple sonority scale, with sonority (proportion of voice) increasing from left to right:

<table>
<thead>
<tr>
<th>OBSTRUENTS:</th>
<th>SONORANTS:</th>
</tr>
</thead>
<tbody>
<tr>
<td>voiceless obstruents</td>
<td>voiced obstruents</td>
</tr>
<tr>
<td>approximants</td>
<td>nasal stops</td>
</tr>
<tr>
<td>vowels</td>
<td></td>
</tr>
</tbody>
</table>

0%

100%

Note that sonorants are typically voiced — there are no voiceless sonorants —, since lack of voice means lack of sonority. Obstruents are either voiced or voiceless.

3) The third important distinction is between stops and continuants. A sound is a stop if its articulation involves a complete closure in the vocal tract; if there is no complete closure, the sound is continuant. The consonant p, for instance, is articulated by the two lips making a complete closure, so p is a stop; the same is true for m, also involving a complete closure made by the lips. Stops can be nasal (such as m), or oral (such as p). Oral stops, called plosives, are produced by a complete closure in the oral cavity with the velum raised so that the air cannot escape either through the mouth or through the nose; instead, a significant body of air accumulates behind the closure, which is released abruptly, letting the air burst out (hence the term plosive). We have already described the production of nasal consonants — note that they are all stops. Continuants include all vowels (logically), approximants, as well as non-stop obstruents — this latter group is called fricatives, such as s, z, f, v. One of the approximants, l, is a strange sound as far as continuancy is concerned. If you pronounce it (as in the word let, for instance), you can feel that the tongue touches the roof of the mouth; indeed, the closure is complete! But then, why is it a continuant — and why an approximant, in the first place? The answer is that the closure here is of a special type: while the tongue tip does block the airflow in the middle of the mouth, the air can escape around the closure, on both sides of the tongue. Such a closure, as well as a sound produced with such a closure, is called lateral (from the Latin word lateralis ‘of/relating to side(s)’). As the air passes through along both sides of the tongue continuously, laterals are continuants; furthermore, the noise produced by the closure is so minimal that there is no audible friction — that’s why l behaves more like an approximant, despite the closure. This is the only lateral consonant of English (and Hungarian), but some languages, e.g., Italian, Russian, Welsh, have more.

Later on in this course, we’ll make our sonority scale more elaborate.
We have seen how sounds are characterisable according to the following criteria:

1. Acc. to the initiation process: in E & Hu, all sounds are produced by a pulmonic egressive airstream mechanism;
2. Acc. to the phonation process: voiced vs. voiceless;
3. Acc. to the oro-nasal process: nasal vs. oral;
4. Acc. to the articulation process, i.e., manner of articulation: consonants vs. vowels, sonorants vs. obstruents, stops vs. continuants.

4 Consonants in detail

This is not yet all to it. We know, for instance, that vowels are all sonorants (hence voiced), but how do they differ from each other? Or, for example, p, t, and k, as in the words pie, toe, cow, respectively, are all voiceless plosives: but what’s the difference between them? We will now address this problem, starting with consonants.

Recall that, during the articulation process, the body of air coming out of the lungs is manipulated by organs called the articulators. Let us now see the articulators in detail.

The most important articulator is the tongue — a large body of muscle, capable of performing a variety of delicate gestures. Other articulators include the lips, the upper teeth, and the roof of the mouth. Movable articulators — these are the lips and the tongue — are called active articulators; those which do not move are passive articulators, such as the teeth and the roof of the mouth. During the articulation of speech sounds, an active articulator comes into contact with a passive (or, in one case, with another active) one, creating the obstruction to the airflow, necessary for consonantal articulation. Take, for example, the initial consonant of the English word think. This consonant is a voiceless fricative, produced with the tip of the tongue touching the upper teeth: the active articulator is the tongue, more precisely, the tongue tip, while the role of the passive articulator is played by the upper teeth. (The IPA symbol for this consonant is θ, a typed Greek letter theta.) Compare this consonant with the f in fine: this one, too, is a voiceless fricative as far as its manner of articulation is concerned. The difference is that with f, the active articulator is the lower lip. The two consonants do not differ in phonation, nasality, or manner, but, instead, in place of articulation: both are voiceless oral fricatives but articulated by different articulators. Depending on what articulators produce the obstruction, phoneticians use shorthand terms to refer to the place of articulation. We say that f, for example, is labiodental, because it is articulated by the (lower) lip (hence “labio”) against the (upper) teeth (hence “dental”). The consonant θ, on the other hand, is dental, the tongue tip touches the upper teeth — note here that if the tongue is the active articulator (as with the majority of consonants), we give the place of articulation with reference to the relevant passive articulator only. The following table in (6) gives a summary of places of articulation including the traditional names as well as the relevant articulators. Note: the tongue tip is really part of the tongue blade — its frontest part.

<table>
<thead>
<tr>
<th>Name of place</th>
<th>Passive articulator</th>
<th>Active articulator(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bilabial</td>
<td>—</td>
<td>both lips</td>
</tr>
</tbody>
</table>

It’s obviously not an accident that in many languages, the word for ‘language’ is the same as the one for ‘tongue’ (as Hu nyelv, Czech jazyk, etc.) — in English, too, tongue is sometimes used meaning ‘language’, especially in the expression mother tongue, but sporadically in other, mainly formal or literary, contexts as well, such as English and German are sister tongues.
<table>
<thead>
<tr>
<th>labiodental</th>
<th>upper teeth</th>
<th>lower lip</th>
</tr>
</thead>
<tbody>
<tr>
<td>dental</td>
<td>upper teeth</td>
<td>tongue tip</td>
</tr>
<tr>
<td>alveolar</td>
<td>alveolar ridge</td>
<td>tongue blade</td>
</tr>
<tr>
<td>palato-alveolar</td>
<td>alveolar ridge / hard palate</td>
<td>tongue blade</td>
</tr>
<tr>
<td>retroflex</td>
<td>hard palate</td>
<td>tongue tip</td>
</tr>
<tr>
<td>palatal</td>
<td>hard palate</td>
<td>tongue back</td>
</tr>
<tr>
<td>velar</td>
<td>velum (soft palate)</td>
<td>tongue back</td>
</tr>
</tbody>
</table>

Let us now see each one in detail, providing a description of the consonants of English. It is good to remember that accents of English differ minimally in their consonants, so what is written here is basically valid for GA as well as RP (I will point out if the two differ).

1) **Bilabial** sounds are produced by both lips. Both in E and Hu, bilabials are stops: a nasal one (m) and two plosives: a voiceless and a voiced one (p and b, respectively). Of the better known European languages, Spanish has a voiced bilabial fricative (produced like b but without complete closure), the IPA symbol for which is β, i.e., a Greek letter beta. Since you may come across this consonant during your later linguistics courses (esp. when studying History of English), it is useful to remember it.

2) **Labiodental** consonants are typically fricatives: voiceless f and voiced v, as in E fee and yase or Hu fa and víza. Labiodental and bilabial consonants are collectively known as **labials**. Because bilabials in E (and Hu) are all stops, while labio-dentals are fricatives, we can simply use the term “labial” when describing them, e.g., we can say that f is a voiceless labial fricative — since there’s no bilabial fricative in E (or Hu), this can only mean “labiodental”.

3) English has two **dental** consonants, a voiceless and a voiced fricative, θ and ð, respectively, as in think and they. Many speakers pronounce these consonants as **interdentals**, with the tongue tip placed between the upper and the lower teeth. Dental fricatives are found in some other European languages as well, the best known one being Spanish, but the Celtic languages, such as Welsh or Irish Gaelic, also have them. Hungarian does not have dental fricatives, but it does have dental stops, as in the words tál, dér, néz, indicated in the IPA with the symbols t, d, n, respectively. Dental stops are articulated with the tongue tip against the area just behind the upper teeth rather than against the upper teeth themselves. English does not have dental stops.
4) **Alveolar** consonants are produced by the tongue blade against the region between the upper teeth and the hard palate known as the **alveolar ridge**. English has three alveolar stops, as in *ten*, *day*, *nice*, denoted by the IPA symbols *t*, *d*, *n*, resp. Note that standard Hungarian does not have alveolar stops; this is why most Hungarian learners of English substitute dental stops for them, which is a sign of a strong Hungarian accent (English learners of Hungarian do just the opposite). Both languages have **alveolar fricatives**, though, as in E *sea* and *sez*, or in Hu *száz*, *zőld*, the former voiceless, the latter voiced (the IPA symbols are *s* and *z*), as well as an **alveolar lateral liquid**, *l*, as in E *let*, Hu *lép*. The initial consonant in Hu *répa* is an **alveolar trill**, produced by the tip of the tongue repeatedly touching the alveolar ridge, which does not exist in RP or GA (but it does so in some other varieties, such as Welsh English or Scottish English). The IPA symbol for the alveolar trill — also found in Italian, Czech, Finnish, Romanian, etc. — is *r*. In GA (and some other accents such as Irish E) there is also an alveolar **flap**, similar to the trill, but the tongue only touches the alveolar ridge once, rather than repeatedly. Its symbol in the IPA is *r*, and it’s found in *waiter*, *riding*, *party*, for instance.

5) **Palatoalveolar** (also called **postalveolar**) consonants are articulated by the tongue blade against the area behind the alveolar region, almost against the palate. In English they include two fricatives, as in *she* and *vision*, also found in Hungarian (e.g., in *só* and *szák*); the IPA symbols for them are *ʃ* and *ʒ*, the first one being voiceless, the second voiced. Here, however, a note is in order. Both English and Hungarian have a type of sound that we have not met yet, illustrated by the initial consonants of E *child*, *gin* and Hu *csupor*, *dzsem*. These are interesting consonants as far as their manner of articulation is concerned: they start out as plosives, viz. *t* and *d*, but the complete closure is not released abruptly; instead, it is released slowly, gradually, making the stop turn into a fricative, i.e., *ʃ* and *ʒ*, respectively. Such “starting-as-a-stop-but-ending-as-a-fricative” consonants are called **affricates**; this double nature of theirs is reflected by the double symbols employed to denote them in the IPA, i.e., *tʃ* and *dʒ*, respectively, for the voiceless and the voiced one. Besides these, Hungarian has two alveolar affricates, too, as in *cérna* and *edz*, the former being voiceless, the latter voiced; not surprisingly, the phonetic symbols for them are *ts* and *dz*.

Due to their common acoustic properties, alveolar and palatoalveolar fricatives and affricates are usually grouped together under the name **sibilants**. In English, there are six: *s*, *z*, *ʃ*, *ʒ*, *tʃ*, *dʒ*; Hungarian has the same six plus *ts* and *dz*.

Finally, RP (and some other accents) has a palatoalveolar non-lateral approximant, as in the word *rain*. The IPA symbol for this sound is *ɹ*, i.e., a turned *r*. This consonant is articulated by the tongue blade making a very loose contact with the palatoalveolar region, hardly touching it. It does not exist in GA (and many other accents such as Scottish English).

6) In GA, words which have a *ɹ* in RP have a **retroflex** approximant instead, symbolised by *ɻ* (a turned *r* with a right tail), produced by the tip of the tongue curled back against the hard palate. Because words with a spelt *ɽ* are pronounced with so many different consonants across the English-speaking world, dictionaries normally follow the phonetically inaccurate
but much more user-friendly practice of using the simple symbol $r$ in the transcription of such words. This practice will be followed in this book, too, except where accents of English are discussed.

Dental, alveolar, palato-alveolar and retroflex consonants — those involving the tongue blade as an active articulator — are collectively known as coronals. (Recall that the tongue tip is part of the tongue blade.)

---

8 This is, remember, the IPA symbol for the alveolar trill, as in Hungarian or Italian, this is why the use of this letter is phonetically not quite justified (except when transcribing English accents that do have such a sound, such as Scottish English).
7) **Palatal** articulation means the back of the tongue against the hard palate. English has but one palatal consonant, *j*, as in *yet* — it’s a palatal approximant (glide). This sound is never spelt as *<j>* in English, so pay attention not to call it “jay” — in fact, phoneticians use the name *yod* to refer to this glide. It is also found in Hungarian, as well as in many other languages. Hungarian, however, has three more palatal consonants, all stops. Two are plosives: a voiceless one, as in *nyíl* (its IPA symbol is *ç* — do not confuse it with the letter *<c>*), and a voiced one, as in *gyere*, denoted by the IPA symbol *ɟ* (a dotless barred *j*). The third stop is a nasal one, as in *nyak*, denoted by the symbol *ɲ* of the IPA.

8) **Velar** consonants are produced by the back of the tongue against the velum. English has three velar consonants, as illustrated by the words *cow, goal, sing*, transcribed as *k, g, ɡ*, respectively. All three are stops: the first and the second are oral ones (plosives), voiceless and voiced; the third one is a nasal. (Please pay attention not to confuse its symbol with that of the palatal nasal, i.e., *ɲ*!) Hungarian has *k* and *ɡ* as velars. In several languages, there is also a voiceless velar fricative, as in German *mach-en* ‘make’, Spanish *Juan* ‘John’, Czech *chleb* ‘bread’, denoted by the IPA symbol *χ* (don’t confuse it with the letter *<x>* which generally indicates clusters such as *ks* or *gz* in E and Hu writing). Its voiced counterpart, indicated by the symbol *γ*, is found in Spanish *lago* ‘lake’ or Greek *gala* ‘milk’. English at present does not have velar fricatives but centuries ago, it did, which is why you’ll come across these consonants when studying History of English, so it’s useful to recognise the symbols.

There is a velar approximant, too, in English, as in *wet*, produced by the back of the tongue raised to the velum. At the same time, the lips are rounded like for the vowel *u* (as in E *moon*, Hu *ű*). This is why *w* is termed *labiovelar*, rather than simply velar; in fact, it has two simultaneous places of articulation.

Palatal and velar (= tongue-back) consonants are grouped together as *dorsal* (from the Latin word *dorsum* ‘back’).

Many languages have uvular and pharyngeal consonants, too. Uvulars are produced with the tongue back against the uvula; pharyngeals are articulated by the tongue root obstructing the pharynx. Uvular consonants are found in French, for example, but in some Northumbrian accents of English, there is a voiced uvular fricative, too, denoted in the IPA by the symbol *ʃ*: it corresponds to RP *j*, GA *ɻ*, as in *rat* (RP *ɹet*, GA *ɻet*, Northumbrian *ɻat*).

We have almost finished our discussion of consonantal articulations, but the observant reader will have noticed that a consonant, found in English as well as in Hungarian, is missing: this is the one found in E *hay* or Hu *hat*, indicated in the IPA by *h*. If you observe how this sound is produced, you will find that it originates from the larynx: there is no obstruction in the vocal tract. In this sense, *h* is like a vowel. In the glottis, however, the vocal cords are spread apart, producing a “heavy breathing” effect, similar to friction. We can, therefore, regard *h* as a **glottal fricative** — a fricative produced in the glottis — adding a further place of articulation for consonants, **glottal**. In fact, English also has a **glottal stop**, more precisely a glottal plosive, produced by the vocal cords being closed completely then released suddenly. The glottal stop’s IPA symbol is *ʔ*. It is not an obligatory sound of English, but it is used to
varying degrees of frequency by all speakers under certain circumstances, for example, as a replacement for an oral plosive, especially $t$. The glottal stop is especially widely used in colloquial London speech, and it can be illustrated by words like *Scotland, battle*, etc., usually pronounced by southern British speakers in all but the most careful types of speech with a glottal stop rather than a $t$, i.e., as 'sknɔlnd, 'bæl. Note please that the articulation of both the glottal stop and the glottal fricative excludes the possibility of the narrowing of the vocal cords responsible for voicing, so that glottal consonants are voiceless only.

We can make a complete table of English consonants, as in (7); the ones in shaded boxes are sibilants. When there are two symbols in a box separated by a slash, the one on the left is voiceless; glottals are all voiceless. I use the symbol $r$ for the palatoalveolar approximant, but remember that the phonetically accurate one for it is $ɹ$. Articulatory categories not found in RP are not included.

Finally, some notes on what we call the symbols if talking about them. Most of them are identical to lowercase Roman letters, and are called accordingly; e.g., the symbol $f$ is called “lowercase F”, $w$ = “lowercase W”, etc. The following ones have special names:

<table>
<thead>
<tr>
<th>IPA symbol</th>
<th>Name</th>
<th>IPA symbol</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>?</td>
<td>glottal stop</td>
<td>$ʃ$</td>
<td>esh</td>
</tr>
<tr>
<td>$θ$</td>
<td>theta</td>
<td>$ʒ$</td>
<td>yogh</td>
</tr>
<tr>
<td>$ð$</td>
<td>eth</td>
<td>$ɲ$</td>
<td>eng</td>
</tr>
</tbody>
</table>