In the representation of sounds it is important not to get confused by spelling, which is historically determined. In English especially this means, for instance, that what you see is often not what you hear (or pronounce)! The late 19th century saw the first attempts at creating a system of symbols to accurately represent the sounds used in different languages. This culminated in the creation of the **International Phonetic Alphabet (IPA)** by the **International Phonetic Association** (confusingly also abbreviated to IPA), which was founded in 1886 in France by the English teacher Paul Passy (1859-1940). The phonetic alphabet represents **phonemes**, i.e. the smallest units of sounds in a language that serve to distinguish one word from another. An easy way to distinguish between phonemes is through so-called **minimal pairs** – word pairs in which only one sound differs: e.g. we can say that /p/ and /b/ are two **phonemes** in English since they distinguish (both in terms of sound – pronunciation - and meaning) the words *pin* and *bin*. Phoneticians would say that /p/ and /b/ **contrast** within the above minimal pair. Indeed, minimal pairs rely on what is known as **contrastive distribution**, i.e. the substitution of contrastive sounds which results in different words (lexical items). The fact of replacing one sound for another in order to create a minimal pair is known as the **commutation test**. It is important to add that contrastive distribution can reveal contrasts not just at the beginning of words, but also in the middle or at the end, whereas it is used for both vowels and consonants: *seat/soot* (/sɪt/-/sʊt/). While minimal pairs are a wonderful tool in order to prove the **phonemic status** of two sounds, their use is not always simple. The main problem of course involves contrasts for which no minimal pair can be found in the language. For instance, while probably all English speakers agree that the sounds at the beginning of *sh* in (/ʃɪn/) and at the end of *beige* (/beɪdʒ/) are two separate phonemes, it is impossible to find a minimal pair in English to show the **contrast** between these two sounds. In cases such as these, an attempt is made to find what is known as a **near minimal pair** – i.e. a set of words that is not really a minimal pair since there is more than one speech sound that is different, but in which the phonetic **environment** is very similar. For /ʃ/-/ʒ/, we can, for instance, come up with *fission/vision* (/ˈfɪʃən/-/ˈvɪʒən/), in which the identical contexts provide a sufficient contrast between the two sounds.

Although in many languages each letter corresponds to one sound, English spelling is such that the same letter may represent several sounds. Consider, for instance, the word *Morocco*; despite containing only one **vowel letter**, its pronunciation reveals that there are, in
fact, three vowel sounds, i.e. /maˈrʊkəu/ (the superscript mark ‘i’ indicates that the following syllable is stressed, i.e. is pronounced with more prominence).

Undoubtedly the most (in)famous example of the disparity between sound and spelling in English is the name Featherstonehaugh, which is pronounced like Fanshaw (/ˈfænʃəʊ/). However, everyday English is full of spectacular letter-to-sound peculiarities. The <ough>\(^1\) cluster, for instance, can be pronounced in nine (!) different ways: lough (/ˈlɒx/), cough (/kɒf/), drought (/ˈdraʊt/), hiccough (/ˈhɪkəp/), though (/ˈðəʊ/), thought (/ˈθɔːt/), thorough (/ˈθɔːrə/), through (/ˈθruː/), and tough (/tɑːf/). Conversely, one sound can be represented by several spellings: e.g. /f/ as in enough (/ɪnˈaʊf/), ruff (/rɑːf/), soft (/sɒft/), and photography (/faˈtɒgrəfɪ/). Another example of our inefficient use of the alphabet are the so-called homophones (words that sound the same, but have different meanings): e.g. see-sea. English also has a large number of homographs, i.e. words that have identical spellings but different pronunciation (and meaning of course): e.g. row, which can be pronounced as rhyming with know or with now, with /rəʊ/ meaning ‘line’ and /rɔː/, a quarrel.

The fact that we distinguish a number of sounds in a given language does not, of course, mean that these sounds are pronounced exactly the same by everyone or in all circumstances and contexts. For a start, we do not all have identical vocal organs, while pronunciation suffers under stress, excitement, fear, etc. or … when someone has drunk too much! However, phonetic contexts are by far the most determining. For instance, all speakers of English distinguish one k, and so those in keep and cool are recognized by all speakers of English as being the same sound. Yet, a closer investigation reveals that these two /k/ sounds are produced in different ways. If you pronounce them slowly, you will find that the /k/ in keep is produced with the back of the tongue raised more forward in the mouth (and is thus said to be fronted), whereas the back of the tongue in cool is more towards the back (and is said to be retracted). Both are variations of the same sound, or phoneme. So, a phoneme should, in fact, be considered an abstract in that it merely symbolizes the various ways in which a sound unit is realized in various different contexts, and by individual speakers. In other words, a phoneme consists of several distinct articulations, all of which are considered identical by native speakers, since one variety may be substituted for another without any change in meaning. If we apply this to our examples with /k/, this means that the use of a retracted /k/ in keep, or a fronted one in cool does not alter the meaning of the word.

\(^1\) In line with common usage in phonetics, angular brackets will be used to render spellings (as opposed to the transcription of speech sounds).
The variants of a phoneme are known as **allophones**, which will be discussed later in more detail. While for general purposes, we need only one symbol, i.e. /k/, it is sometimes necessary to give a more accurate and detailed description of the actual realization of the sound. In this case, the **broad (phonemic) transcription** is not enough, and special signs – known as **diacritic marks** (usually abbreviated to **diacritics**) – are added. This more detailed representation is referred to as a **narrow (or allophonic) transcription**. In our example, the /k/ in cool would then be represented by means of [k] (the small line underneath is the diacritic indicating that the sound is retracted), and that of keep by [k], in which the plus sign denotes the advanced position of the tongue in the pronunciation of the sound.

You will notice that in the **allophonic transcription** the sound symbols are enclosed in square brackets ([ ]), as opposed to the obliques used in **phonemic transcription**.

In modern phonetic theory, however, a further distinction is made through the term **phone**, which refers to the actual realization, i.e. utterance, of the speech sound; in other words, what we hear when people speak are **phones**, whereas both phonemes and allophones are underlying constructs.²

The first thing that needs to be added is that allophones tend to be **predictable**. Let us consider, for instance, the allophones of /p/, which in English can be an ‘aspirated’ ['pʰ] (pronounced with a soft ‘h’ sound after it, as in Peter), and an unaspirated one (['p'].) However, it is not as if these two variants are used interchangeably. As we shall see, the aspirated variety tends to occur at the beginning of a words, but not after /s/, whereas the unaspirated ['p'] is not used (at least not in Standard British English) at the beginning of words. This means that if there is a p sound at the beginning of a word (not preceded by /s/), then it can be predicted that the aspirated allophone will be used. As a result, in word-initial positions ['pʰ] and ['p'] are said to be in **complementary distribution**. Other examples include the /t/ in tenth ([tɛŋθ]), which is **dental** (rather than alveolar) because of the following dental fricative, and thus only occurs in these contexts.

In some cases, however, it is possible to find both allophones in the same position; this is true for instance for the ['p'] and ['pʰ'] in **word-final** positions (i.e. at the end of words): for instance, in a word like lip, both the aspirated and unaspirated variants are found. As a result, in this case, the allophones are said to be in **free variation**.

² The first to formulate the theory of phonemes and phones was the Polish-Russian scholar Jan Baudouin de Courtenay (1845-1925), whereas it gained currency in Western linguistics through the work of the British phonetician Daniel Jones.
The above situation, of course, only applies to the Received Pronunciation variety of English. Indeed, in South African English, for instance, the de-aspirated /p/ does occur in word-initial positions, whereas most English speakers would be amazed to find that the unaspirated and aspirated /p/ sounds constitute different phonemes (not allophones) in, for instance, Thai or Korean, in which [tʰ] - [t'] and [kʰ] - [k'] are also different phonemes. So, to a Thai speaker, the difference between these two sets of sounds is comparable to the difference to English speakers between, say, /p/ and /b/, or /t/ and /d/. In other words, in Thai [p’]-[pʰ], [tʰ]-[t’] and [kʰ] - [k’] have contrastive distribution, which means that the substitution of the aspirated sound for the unaspirated one results in a difference in meaning, as shown by the following minimal pairs: [pʰɑː]-[pʰɑː] (‘aunt’) - [pʰɑː] (‘cloth’); [tʰɑː]-[tʰɑː] (‘landing place’); [kʰɑː]-[kʰɑː] (‘leg’).

In the above discussion of allophones we used the words keep and cool as an example to indicate that despite being considered the same /k/ sounds by English speakers, there are differences in how it is realized in these two words, with the one in keep being produced with the front of the tongue being raised towards the palate, whereas the one in cool involved a more retracted (i.e. towards the back) pronunciation. Basque or Vietnamese speakers, however, would recognize the difference between these two sounds far more rapidly since in Basque and Vietnamese these two sounds are distinct phonemes rather than allophones of one phoneme (in the IPA the voiceless palatal stop of /kiːp/ is represented by the symbol [ç]).

A final example – and also the most spectacular one from an English speaker’s point of view – involves Japanese. Speakers of this language are notoriously incapable of pronouncing a European /r/, resulting in sentences like ‘I like law (instead of raw) fish.’ However, a closer look at the phonology of Japanese reveals that the language does contain an /r/ sound. So, what is the story? Well, in Japanese, the phoneme /r/ has the allophone /l/ in word-initial positions and the allophone /ɾ/ in intervocalic positions (i.e. between vowels): e.g. nara (‘if’); loku (‘six’). So, when Japanese speakers says lice when in fact meaning rice (as in flaid lice for fried rice) or boring (instead of bowling), they are simply applying a phonological rule of Japanese since to them /ɾ/ and /l/ are simply allophones of the same phoneme!

Although most alphabets started out by being written representations of the sounds used within the linguistic community, there is one alphabet in the world that may be said to be strictly phonemic – Korean. It was created in around 1450 A.D. by a king called Se-
Jong, who must have been quite a phonetician. For instance, in his alphabet he accorded a single letter to an unaspirated lenis /p/ and /b/. Naturally, the king knew these sounds did not sound the same, but since in Korean these occur in complementary distribution (the /b/ between voiced sounds, /p/ in all other contexts), they are considered to be very closely related by Korean speakers. Conversely, since, as we have seen, /p/ and /pʰ/ are separate phonemes in Korean, the king awarded a separate symbol for each!