Chapter 4: Identity in Phonology: Contrast

The aim of this chapter is to introduce you to the basic concepts of phonology, that is, sound structure. In Chapter 2, we looked at the phonetic identity of speech sounds. This chapter, too, deals with identity, but from a different point of view: from that of function. This is phonological (that is, linguistic) identity.

You have already seen that phonetic and functional identity is not necessarily the same: remember the case of approximant sounds, such as \( j \), which are phonetically defined as vowels but functionally, they behave like consonants. Now, if something behaves as if it was a consonant, then it is a consonant — in terms of function. The phonologist is interested in this aspect of sounds — their function, hence, for a phonologist, approximants are consonants, while for the phonetician, they are vowels: a phonetician is interested in articulatory and acoustic properties.

There is a famous textbook parallel found in several introductory phonology courses, which may make the difference between phonetic and phonological identity clear and easily understandable. Like speech sounds, coins can be studied from two different aspects: according to their material (physical) properties, or according to their function (that is, value). The science studying coins as physical objects is called numismatics. A numismatician is interested in the coin’s size, what metals it is made of, whose picture is on it, etc.; how much purchasing power the coin has is uninteresting for him. Economics, on the other hand, which also deals with money, is totally uninterested in the coin’s physical properties. In fact, economics is not even interested in whether a given denomination (e.g., $ 1, € 5, etc.) comes in coins, notes, or both, or neither. For a numismatician, € 3 doesn’t exist (there’s no such coin), but it is an existing amount of money for the economist, who is interested in what you can buy for a given amount, not how it is realised. Phonetics is similar to numismatics, while phonology is much like economics: the former is interested in material, the latter in functional identity or value.

In order to see the essence of the “identity” question, we must first clarify what we really mean by the function of speech sounds. Speech sounds (segments) are used to make up words: they are the bricks used to build the form of the linguistic sign. Function in this sense means two, interrelated things. First, what position the sound can occupy in the word? How can it combine with other sounds? For instance, in English, the majority of short vowels cannot occur in word-final position: forms such as *te, *lieq, *sto, etc. are ill-formed (the same is true for short o and ə in Hungarian). The velar nasal ɳ, on the other hand cannot be the first sound in the word (*ŋ1k, cf. k1ŋ <king>). To introduce the appropriate terminology, the first question relating to function is the distribution of the sound: those positions in which it can occur. The second aspect of function derives from distribution. If two sounds can occur in the same position, it is possible for them to distinguish words. All three nasals of English can, for example, appear in word-final position. Take the word sum <some>, and replace the m with one of the other nasals: you’ll get either sun <son> or sün <sung>. These are all actual words in English, and they differ in one segment only: the last one. Two words which differ in one speech sound only are said to constitute a minimal pair, so some — son is a minimal pair, and so are some — sung and son — sung. If there are more than two words differing in one sound, we talk about a minimal set, e.g., some — son — sung (we can, of course, add more words to the set, such as suck, sub, sup). Now, note that sum and sun

---

1 There was a short period in Hungary in the 1990’s, for instance, when there were both coins and notes worth Ft 100 in circulation.

2 This is quite a misnomer, since it isn’t the pair which is minimal but the difference between the members.
represent two words (homonyms) each: the former can mean ‘néhány, valamennyi’ as well as ‘összeg’, the latter ‘fiúgyermek’ or ‘nap’. This happens to be reflected in the spelling, viz. son/sun and some/sum. Nonetheless, son — sun is not a minimal pair, but a pair of homonyms: two words having the same form. Note that we’re dealing with sounds, and they sound identical, even if spelt differently! The same goes for sea — see, road — rode, practise — practise, etc. Also, words of different length cannot be minimal pairs, so the presence vs. absence of a sound doesn’t make a minimal pair, as law — lord (= lɔː — lɔːd) or dine — Dinah (= dæn — dann).

If two sounds can distinguish words — that is, signs — they are said to contrast, and the difference between them is contrastive or distinctive. This is what we get when two sounds have an identical or at least overlapping distribution (identical distribution — when two sounds are always replaceable with each other without loss of well-formedness — is quite rare). Overlapping distribution means that both sounds can appear in some positions, but not in all. The nasals illustrate a case in point: the labial and the coronal one can appear word-initially, but the velar one can’t; word-finally, all can appear. Even though η is not possible initially, it is found word-finally, so it contrasts with both m and n, as shown by the above examples.

It often happens, though, that two sounds cannot appear in identical positions at all. Let us illustrate the situation first using an example from Hungarian. First, let us ask a seemingly simple question: how many nasal consonants does Hungarian have? In the previous chapter, we gave the usual answer (generally given by speakers): three, as in már, nulla, nyak, i.e., m, n, η.¹ Let’s now pronounce the words fánk and leng. Which of these three consonants is pronounced in these words? Careful observation shows that none of them! If you observe the nasal in fánk and leng, you’ll notice that it’s not labial, coronal, or palatal: instead, these words contain a velar nasal, η (fæŋk, lɛŋ). It seems, therefore, that Hungarian has four nasals rather than three!

Phonetically speaking, this is true, but then, why is it that Hungarian speakers (unless phonetically trained, of course) are not aware of this, and they can identify three nasals only? You may now say that it is because of Hungarian orthography, which doesn’t have a letter for η, only for the other three (= <m>, <n>, <ny>). However, the same situation holds for illiterate speakers, too, where the influence of spelling is out. Furthermore, why can Hungarian orthography afford not to have a special letter for each nasal consonant? Instead of saying that the spelling is to blame, it seems better to seek another solution, which can answer this question, too.

The orthography uses the letter <n> for η, that is, the usual letter for the coronal nasal. The spelling “pretends” as if the words fán and fánk had the same consonant. Hungarian speakers (including illiterate ones like pre-school children) will readily identify the two consonants as “being the same”; in fact, this is the reason why the spelling doesn’t distinguish them! But why do Hungarians think that fán and fánk have the same consonant?

A possible answer could be that the two sounds are very similar phonetically, and people don’t perceive the small difference, or at least they don’t consider it significant enough to reflect it in writing. But this is not true: n and η differ from each other in place of articulation: the former is coronal (more precisely, dental), the latter is velar. The voiced plosives d and g differ from each other in exactly the same way: the two nasals are no more similar than these two plosives! Yet, the difference between the plosives is obvious for all speakers of Hungarian, no one confuses them, and the spelling reflects this faithfully: it would be extremely confusing if we tried to use the letter <d> for both (imagine spellings like tömed,

---

¹ I will use the symbol η here from now on for the sake of simplicity since there’s no alveolar nasal in Hu to confuse it with, but note that the coronal nasal is dental, rather than alveolar, in this language.
The real reason is suggested by the abovesaid. The plosives d and g have an overlapping distribution, that is, both can occur in most possible positions, such as word-initially or finally, as well as intervocally (= between vowels), so that they can form minimal pairs such as *domb — gomb, éden — égen: they contrast, i.e., they are different not only phonetically but functionally as well. Because the difference between them is distinctive, it identifies signs: éden differs from égen due to it; it follows from this that speakers are aware of the difference. The consequence of this for the orthography is that it must reflect the non-identity of these two sounds.

As opposed to this, n and ŋ never occur in the same position: the former is found word-finally, before vowels, and before coronal consonants (e.g., vén, nagy, tonik, szende, bán, pénz, sansz), while the latter is only possible before velar plosives, as in fánk, inka, leng, harang. Note that n is never found in the same environment. Therefore, the phonetic environment — more precisely, what follows the nasal — determines the choice between n and ŋ. The type of distribution characteristic of these two consonants is called complementary distribution: it means that the two segments are found in complementary, never identical, environments. The most important consequence of this state of affairs is that the two sounds cannot distinguish words, as they aren’t interchangeable. The n of the word *veːn <vén> cannot be replaced by ŋ (*veːŋ), because the velar nasal is impossible word-finally; and while leng <leng> is a possible (and existing) word form, *lEng isn’t, because n is not found before velars. In other words, there are no minimal pairs with these two consonants: if two words differ in this respect, there must be another difference, too — notably, in what follows. These two nasals are, then, phonetically different but functionally, they aren’t. This is the reason why native speakers aren’t aware of the phonetic difference: it doesn’t make a functional — phonological — difference. The two sounds do not contrast.

Note furthermore that ŋ — itself velar — appears precisely before velar consonants. This is surely not a coincidence. Instead, it appears that n and ŋ are each other’s positional variants, the latter occuring before velars. This is shown by the fact that if a word ends in n and gets a velar-initial suffix, the n becomes ŋ, as in öveŋ[k]or. Indeed, except maybe in over-careful theatrical speech, we find the same situation across words, too. The final consonant of nagyon is n if the word is uttered in isolation, but if the next word begins with a velar consonant, it is replaced by ŋ, as in nagyoŋ[k] komoly or nagyoŋ[g] gáz. The n is assimilated to the following velar; since the assimilation involves place of articulation, we talk about place assimilation.

In fact, velar is not the only place of articulation before which n is impossible: it isn’t found before labial or palatal consonants, either, and if a word ending in n receives a suffix or is followed by a word beginning with a labial or a palatal consonant, n becomes a labial or a palatal nasal, respectively. Examples include szé[m]ből, ki[m]pad, vé[n] tyúk, nagyo[p] gyors. To sum up, we can say that in Hungarian, the coronal nasal is only found before homorganic consonants, but not before others; the term homorganic means ‘having the same place of articulation’. Such a behaviour on the part of nasals is not uncommon in the world’s languages. Interestingly, the labial and the palatal nasal in Hungarian can occur before non-homorganic consonants, too, that, is, they never assimilate, cf. szemtől, háromkor, aranyből, lányt, for example.

If in a language two (or more) sounds are each other’s positional variants in complementary distribution, we say that they realise one phoneme. In Hungarian, then, n and

---

4 The terms environment and position are interchangeable.
η realise one phoneme — they are functionally identical, not being able to distinguish words. We can say that fian and fiank contain two different consonant sounds, but the same consonant phoneme. The psychological evidence — i.e., that native speakers are unaware of the phonetic difference and perceive the two sounds to be one — suggests that the phoneme can be thought of as a “mental sound”, as it were. The sounds that act as positional variants of the phoneme — n and η in this case — are referred to as the allophones of the phoneme.

Let us now see the role of the velar nasal in English. Here, it cannot only occur before velar consonants, but also in word-final position, as we have already seen; accordingly, it contrasts with the labial as well as the coronal nasal (recall sun — sung, i.e., sun — sunη as a minimal pair). In other words, n and η are not only different sounds, but also different phonemes in English. That is, English has three nasal consonant sounds and three corresponding phonemes; Hungarian has four such sounds but three phonemes. This brings us to an important point: how sounds correspond to phonemes is language-specific; furthermore, what characterises languages is not only the inventory of speech sounds they have, but also (more importantly even) how they are organised into a system of phonemes. In other words, two languages may differ not in what sounds they have but what contrasts exist between them.

Recall from Chapter 2 that in phonetics, we use square brackets to indicate that we talk about speech sounds, e.g., [æ], [θ], etc. How can we indicate if we’re talking about phonemes? Phonologists have traditionally used slants (also called slashes) to do this. So, for example, we can say that English has three nasal consonant sounds which are separate phonemes:

(1) Nasal consonant phonemes and sounds in English

<table>
<thead>
<tr>
<th>Phonemes:</th>
<th>/m/</th>
<th>/n/</th>
<th>/η/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sounds:</td>
<td>[m]</td>
<td>[n]</td>
<td>[η]</td>
</tr>
</tbody>
</table>

In Hungarian, there is an extra nasal, the palatal one, so there are four nasal consonant sounds; yet, the number of nasal consonant phonemes is the same, because the coronal and the velar nasal are allophones of one phoneme. We indicate this by using one symbol for it between slants. But what symbol? Should we indicate the phoneme as /n/, /η/, or something else? In such cases, we generally choose the symbol for one of the allophones, specifically, of the one which can be regarded as a basic variant. In our case, the allophone [n] can be regarded as more basic, because it occurs in a wider range of environments, while [η] is restricted to pre-velar position. For this reason, we use /n/ to represent the phoneme. Altogether, here’s the overall picture of Hungarian nasal consonants:
(2) Nasal consonant phonemes and sounds in Hungarian

<table>
<thead>
<tr>
<th>Phonemes:</th>
<th>/m/</th>
<th>/n/</th>
<th>/ɲ/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sounds:</td>
<td>[m]</td>
<td>[n]</td>
<td>[ɲ]</td>
</tr>
</tbody>
</table>

All languages, then, have sounds which act as the realisations (= pronunciations) of a single phoneme. Since the choice of allophones (as n vs. ɲ in Hu) is rule-governed, we can talk about **alophonic rules**. Linguists usually use short formal notation to write down rules instead of using lengthy descriptions “in prose”. For instance, instead of writing “In Hungarian, the phoneme /n/ is realised as a velar nasal sound if immediately followed by k or g”, they normally write something like in (3):

(3) /n/ → [ɲ] / ___ {k,g}

We’ll make this more precise later on, but for the time being note that the arrow is read out as “realised as”; the / is called the **environment bar**, after which the environment of the rule comes. The ___ marks the position of the affected item (here, the phoneme /n/). In our case, the affected item must be followed by a velar consonant in order to be pronounced as a velar nasal, so the ___ itself is followed by {k,g}, where the **braces** (= {...}) indicate an “either — or” choice. All in all, this formula can be read out as “The phoneme /n/ is realised as a velar nasal if followed immediately by either k or g”.

As seen above, it happens that the same sounds in one language act differently than in another one. The nasal consonant system of English has three members on a **phonemic** level, just like Hungarian; but the two systems are not identical. First, Hungarian has a labial, a coronal, and a palatal nasal consonant phoneme; English has a labial, a coronal, and a velar one. Second, one of them, the coronal one has two allophones in Hungarian. Languages then can differ in their **inventory** of phonemes as well as their **alophonic rules**. (Plus, of course, in the inventory of sounds, but this is not what we’re interested in here.)

To illustrate the phenomenon of allophony in English, let us take a look at the lateral liquid /l/. In Hungarian, it is basically always pronounced in the same way (disregarding really small articulatory nuances which need not concern us here): as a lateral alveolar liquid, i.e., [l]. In English, this is but one of the possibilities. In many instances, it is realised quite differently, and this difference is well audible and easily noticeable. Compare the words in (4a) with those in (4b):

(4) (a) let, live, lapse, lock, Luke, alive, melon, palace, valid  
(b) false, cult, milk, help, field, bold

In Southern British E, including RP, the words in (4a) have an alveolar lateral roughly identical to what Hungarian has, i.e., [l]. In (4b), however, we find something different: in these words, the /l/ is pronounced with the back of the tongue raised towards the velum: these words contain a **velarised** alveolar lateral, rather than a plain alveolar one. The term **velarised** means that the primary place of articulation is not velar but some other (in this case, alveolar), but it is supplemented by velar articulation: velarity is a **secondary articulation** here. The IPA symbol for the velarised lateral liquid is [ɻ] (called “L with Tilde”). As the back of the tongue approaches the velum but does not touch it, the [ɻ] sounds quite similar to [w]. In
traditional English grammars, the [ɬ] sound is called “Dark L”, as opposed to [l], called “Clear L”. These two sounds are allophones of the single lateral liquid phoneme /l/. Let us see why.

In order for two sounds to act as allophones of one phoneme, they must be in complementary distribution, i.e., which one is found in a given word is determined by phonetic environment. Let’s see if this is the case. Indeed, if you observe the words in (4a), you’ll find that in each, the /l/ is followed by a vowel sound, whereas in (4b), it’s always followed by a consonant. It seems, then, that [ɬ] and [l] are in complementary distribution, the former found before vowels, the latter before consonants. Using the symbol V to indicate any vowel, and C to indicate any consonant (as usual in linguistics), we can write down two allophonic rules:

\[
\begin{align*}
(5) \quad (a) & \quad /l/ \rightarrow [l] / \_ \_ V \\
(b) & \quad /l/ \rightarrow [ɬ] / \_ \_ C
\end{align*}
\]

The phoneme /l/, however, can also appear in word-final position —is it pronounced dark or clear there? Look at the data in (6):

\[
\begin{align*}
(6) \quad (a) & \quad \text{Yes, I will.} \\
(b) & \quad \text{I will come.} \\
(c) & \quad \text{I will ask.}
\end{align*}
\]

\[
\begin{align*}
(6) \quad (a) & \quad \text{I can’t tell.} \\
(b) & \quad \text{I can’t tell them.} \\
(c) & \quad \text{Tell us about it.}
\end{align*}
\]

\[
\begin{align*}
(6) \quad (a) & \quad \text{I forgot all.} \\
(b) & \quad \text{I know all people here.}
\end{align*}
\]

\[
\begin{align*}
(6) \quad (a) & \quad \text{He likes the girl.} \\
(b) & \quad \text{The girl knows him.}
\end{align*}
\]

\[
\begin{align*}
(6) \quad (a) & \quad \text{I know all of them.} \\
\end{align*}
\]

In (6a), the word-final /l/ is followed by nothing but a pause (= silence): it is not only word-final, but also utterance-final. Linguists use the term absolute word-final position to refer to this situation. In these words, the /l/ is dark. What about (6b) and (6c)? In (6a), too, the /l/ is dark, but in (6c), it is clear. How do we explain the difference? Note please that in both (6b) and (6c), the /l/ is not in absolute word-final position: it is immediately followed by the first sound of the following word. In (6b), this sound is always a consonant; in (6c), it is always a vowel. We can now say that the two rules in (5) are also valid for word-final /l/ if it is not in absolute word-final position. In absolute word-final position, /l/ is realised as [ɬ], which we can formulate as in (7):

\[
(7) \quad /l/ \rightarrow [ɬ] / \_ \_ ##
\]

The symbol # is used to indicate a word boundary; when doubled, it means absolute word boundary, i.e., pause. Here, ## follows the ___ (which, of course, marks the place of /l/), so the rule is read out as “The phoneme /l/ is realised as [ɬ] before an absolute word boundary, i.e., when utterance-final”.

Now we have three rules altogether, collected below in (8):

\[
\begin{align*}
(8) \quad (a) & \quad /l/ \rightarrow [l] / \_ \_ V \\
(b) & \quad /l/ \rightarrow [ɬ] / \_ \_ C \\
(c) & \quad /l/ \rightarrow [ɬ] / \_ \_ ##
\end{align*}
\]

In fact, using braces, it is possible to collapse (8b) and (8c) in one rule, having 2 rules altogether, as in (9):

\[
(9) \quad (a) \quad /l/ \rightarrow [l] / \_ \_ V
\]
Note, however, that these two rules apply in complementary environments, and that we consider the unvelarised variant to be basic. This means that /l/ is not velarised unless followed by a C or ##. We can assume a default rule /l/ → [l]. The term default means “unless specifically indicated otherwise”. In other words, there is no need to tell that /l/ is clear before vowels; instead, we can say that velarisation applies where it can; if it cannot, the /l/ is realised as [l], as shown in (10). Note that (10a) is a more specific rule than (10b), because it is restricted to pre-consonantal and absolute word-final position. The specific rule (10a) applies wherever it can, and the more general rule, (10b), is applied elsewhere (which, logically, means “before vowels”).

(10) Clear vs. Dark L

(a) /l/ → [H] / ___ {C, ##}
(b) /l/ → [l]

In fact, a few notes are necessary. First, this distribution of clear and dark L is characteristic of RP and many other accents, but not of all of them: in Scottish English, and many varieties of GA, /l/ is always dark, i.e., [l], while in Irish E, it is always clear. Second, there are two subcases not described above. First, /l/ is always clear before j, even though j is a consonant, as in the data in (11):

(11) million, value, valiant, will you, pale youth (all with a string [lj])

This, however, is not surprising: j is a palatal approximant, pronounced with the tongue back against the hard palate; this articulation simply inhibits velarisation. We can complete our rule in (10a) to exclude Yod\(^5\) from the darkening consonants to give a final version:

(12) (a) /l/ → [H] / ___ {C, ##}, where C ≠ j
(b) /l/ → [l]

The second case where the /l/ is always dark is when it is syllabic. We’ll not deal with this now, since we lack the necessary background, but we’ll come back to it later.

Another very characteristic allophonic rule of English is Aspiration, affecting voiceless plosive phonemes, i.e., /p, t, k/. In certain positions, they are accompanied by a following audible breath, as if one pronounced a short h after them. This is found, for example, in words like pin, time, cow. In the IPA, a superscript h is used to indicate aspiration, so that these words are transcribable as [pʰin, tʰaim, kʰau]. In fact, the parallel with the glottal fricative h (as well as the use of the symbol h ) is not accidental: aspiration is caused by the vocal cords being spread, similarly to the glottal fricative. The spread state of the glottis makes voicing impossible. When the plosive is pronounced (= the closure is released), the articulators move on to produce the following sound (a vowel in the words above, such as i in pin), which is voiced. It takes time, however, for the vocal cords to assume the position in which voicing can be produced: the result is that between the release of the closure and the start of voicing there is a short period when all one hears is the breath caused by the spread glottis. In other words,

---

5 Remember that Yod is the name of the sound j.
aspiration is none other than the pronunciation of voiceless plosives with spread vocal cords, the result of which is a delay in the voicing of the following sound. In fact, voiceless plosives can also be followed by approximants, i.e., /l, r, j, w/, as in play, proud, twin, cute, queen. If you observe how these words are pronounced, you will find that the h-like breath found in pin, time, etc. is missing, but the approximants are pronounced without voice! The reason for this is precisely the delay in voicing caused by aspiration: by the time voicing starts, the articulators have already finished the production of the approximant, starting to produce the vowel after it. As a result, the approximants are devoiced. Devoiced sounds are indicated by a little subscript circle diacritic, so that twin, cute, play, proud are pronounced as [twɪn, kjuːt, p]eɪ, prəʊd], resp.

Aspiration, however, is not always present in voiceless plosives; in the following words, for example, there’s no aspiration at all:

(13) (a) stay, store, speed, spin, scared, sky
(b) stream, stew (! = stʃuː), splash, square (! = skaʊd)

These words are pronounced without any delay in voicing: the ones in (13a) contain a plain (unaspirated) plosive immediately followed by a vowel, e.g., spin, pron. [spɪn] — compare pin = [pʰɪn]; and the words in (13b) have a voiced approximant, e.g., square, pron. [skwεː], cf. queen = [kwɛːn]. Why is there no aspiration here?

You may have noticed yourself that the voiceless plosive in these words is preceded by /s/. Indeed, this is always the case: voiceless plosives are pronounced without aspiration after /s/. In other positions, they are aspirated, although the aspiration is not equally strong in all positions (it is stronger before a stressed vowel than before an unstressed one, so the first /t/ of tattoo is not as heavily aspirated as the second one, the stress being on the second syllable), and there is some variation amongst speakers, too. We may now formulate two rules for each voiceless plosive: one to account for the unaspirated pronunciation after /s/, and another one to take care of the default cases (i.e., everywhere else):

(14) (a) 1. /p/ → [p] / s ___ 2. /p/ → [pʰ]
(b) 1. /t/ → [t] / s ___ 2. /t/ → [tʰ]
(c) 1. /k/ → [k] / s ___ 2. /k/ → [kʰ]

Each voiceless plosive phoneme, then, has two allophones in English: an aspirated and an unaspirated one. Let us make a diagram (similar to the one illustrating the behaviour of nasal consonants) to visualise the situation:

(15) Voiceless plosives in English

Phonemes: /p/ /t/ /k/

Sounds: [pʰ] [p] [tʰ] [t] [kʰ] [k]

The members of each pair are in complementary distribution, never contrasting with each other: while [pʰɪn] is okay and existent, *[pɪn] isn’t, because unaspirated [p] is only possible in English after s. Conversely, *[spʰɪn] is ill-formed because aspiration never occurs after s.

It is now inevitable to address a very important question: What do we transcribe when we
transcribe English words? Or, more importantly, what do dictionaries transcribe? For instance, take a few words with the phoneme /l/ and check some good dictionaries. You will find that they do not show the difference between [l] and [ɬ]: in the Longman Pronunciation Dictionary (abbreviated LPD), for example, the words *lot* and *call* are transcribed as *lɔt* and *kɔɬ*, respectively. Similarly, the difference between aspirated and unaspirated plosives isn’t shown, either: the LPD transcribes the words *stop*, *twin*, and *time* as *stɒp*, *twɪn*, *tɑɪm*, resp. If the purpose of transcription is to show how a word is pronounced, why do dictionaries not show these differences?

As you can suspect by now, dictionaries neglect allophonic differences, such as the one between Clear and Dark L, or between aspirated and unaspirated plosives. They never fail to distinguish different phonemes! The words *spy* and *sky*, for instance, are transcribed in the LPD as *spaɪ* and *skaɪ*, respectively. These words constitute a minimal pair: they only differ in their second segment. As a result, the difference between /p/ and /k/ is phonemic: the two sounds contrast, hence they are different phonemes. The words *pin* and *spin*, on the other hand, are transcribed as *pɪn* and *spɪn*, resp., because the difference in aspiration is not contrastive: it doesn’t distinguish words, i.e., the difference between /p/ and /p/ is not phonemic. Because the choice between these two sounds in a given word is rule-governed, the speaker doesn’t need to learn separately for each English word whether it contains a /p/ or a /p/: instead, he/she automatically pronounces an unaspirated /p/ after /s/, and an aspirated /p/ elsewhere. This is beautifully reflected by the fact that native speakers are not even aware of the difference: the two sounds are mentally speaking one for them. By contrast, the difference between /p/ and /k/ is contrastive: it identifies words. If you mispronounce the word *spy* as *skɑɪ* by accident, people will not understand you, because *sky* means something other than *spy*; if, on the other hand, you pronounce *spy* as *spɑɪ*, they will feel that it sounds unusual, but they will have no problem in identifying the word.

It follows from this that dictionaries do not really show you the exact pronunciation of words: instead, they only give information that is necessary to identify them and tell them from each other. Non-contrastive differences are not shown, because you can predict them if you know the language. Quite obviously, dictionaries are made for language users — they assume that you are familiar with the language in general terms. To include all details of pronunciation would be very disturbing. To sum up, dictionaries transcribe phonemes, not sounds: they tell you what phonemes a word contains and in what order, but not how exactly a phoneme is pronounced in the given word. Such transcriptions are called phonemic or broad transcriptions, and, therefore, they are enclosed between slants, e.g., *lime, told, pine, spy* = /lɑɪm, ˈtɑʊld, pɑɪn, spɑɪ/.

If we transcribe how a word is actually pronounced, we use square brackets; such a transcription is called a phonetic or narrow transcription, e.g., *like, told, pine, spy* = [lɑɪm, tʰɑʊld, pʰɑɪn, spɑɪ]. Phonetic (narrow) transcriptions are, in fact, very rarely used: only if we are interested in fine phonetic detail. In the majority of cases, and always so in practical (non-scientific) language use, we use phonemic (broad) transcriptions. So please beware! In everyday language, we talk about phonetic transcriptions when we actually mean phonemic ones! This is permissible in everyday discourse, but in linguistics, the two things must be clearly distinguished! Note that because narrow transcriptions are very rarely used, I will usually omit the slants for broad ones, so broad transcriptions will generally be given without any bracketing, but I will always use the square brackets if I give a narrow transcription.

Let us now go back to the formulation of the aspiration rule given in (14). While it describes the situation all right, it is not quite satisfactory: we set up three rules, but in reality, there’s but one single rule at work. To state it “in prose”: “Voiceless plosives are unaspirated
when preceded by /s/”. This statement applies to all voiceless plosives, not only to one or two of them: in other words, this rule affects a group of phonemes rather than a single one (as opposed to L-Darkening, for instance). In any natural language, if a rule affects a group, we always find that the given group is not a random collection of phonemes but rather, it is a class. A class — as opposed to a random grouping of individuals — is based on some common property that all members share; assigning individuals to classes, based on common properties, is called, of course, classification. We can, for instance, classify people based on age, profession, sex, marital status, etc. In a city, for example, “university students” is a class. Members of this class share an important property, something that defines them as a class (i.e., that they attend a university course); moreover, we expect them to show many similarities in their social life, behaviour, attitudes, etc. which may not characterise non-students: university students constitute a socially relevant class, whose members differ from each other but still they behave similarly. (The same is true, of course, for teachers.) As opposed to this, imagine that I make a group of people by selecting two hundred inhabitants of the given city at random. Socially speaking, this is not a relevant class, but a random group: it will probably contain very different people, young and old, male and female, of a range of professions and social standing, etc. In fact, they are unlikely to have anything socially important in common: they are not a class.

To return to phonology, if a rule affects several phonemes, they are members of a single class, sharing some common property or properties that other phonemes lack. The question is, what are these properties like? What are they based on?

If you think about which phonemes of English are subject to the Aspiration rule, and what they have in common (apart from Aspiration itself, of course), you’ll find that they are all voiceless plosives. This is a property that they all share, and no other phoneme of English shares this property: in other words, English has no other voiceless plosives. Both “voiceless” and “plosive” are, of course, phonetic (articulatory) properties. Indeed, we find that segments with common phonetic properties behave similarly. Phonetic properties which are used to define classes of segments are called phonetic features; “voiceless”, for instance, is a phonetic feature. In phonological theory, the names of features are placed between square brackets, so, for example, [voiceless]. Segments which share a common feature constitute a class, known as a natural class in linguistics. All segments of English characterised by lack of voicing are members of a natural class, for example; the complementary class is, of course, that of voiced segments. More often, however, a natural class is not defined on the basis of a single feature, but, rather, on a combination of two or more. The class of voiceless plosives is a good example. All members of this class are voiceless and they are all plosives. So, it seems that this class is defined by two phonetic features. But hang on! The term “plosive” is a cover term, remember! It does not correspond to a single articulatory gesture, but rather, to two. Plosives are, on the one hand, characterised by complete closure; on the other hand, they are also characterised by oral articulation (i.e., lowered velum). Indeed, this latter gesture is what distinguishes them from nasal stops, which are also characterised by complete closure! Voiceless plosives, then, are characterised by three common articulatory gestures which set them apart from other sounds:

1. lack of voice;
2. closure (lack of continuous airflow through the oral cavity);
3. raised velum (lack of nasality).

These are the three phonetic gestures that define “voiceless plosives” as a natural class. Each of them defines two natural classes:

1. voiced vs. voiceless sounds

---

This isn’t a feature, in fact, but let’s use it for the moment; we’ll come back to it right away. The use of square brackets may be confusing at first since they are also used to show narrow transcriptions, but the two things are easily distinguished: for one thing, transcriptions are written in IPA symbols!
2. continuant vs. stop sounds  
3. nasal vs. oral sounds

We may now say, for instance, that aspiration is a property of “voiceless” oral stops; “voiced” oral stops are never aspirated, and neither are “nasal” stops, etc. Note, however, that voiceless” and “voiced” are really complementary properties: if something is not voiced, it’s voiceless, and vice versa. This is because “voiceless” is not a gesture: it’s the lack of voice. Therefore, we have a single phonetic feature (= voice) here, as opposed to the lack of it, rather than two independent features.

Therefore, instead of talking about the feature [voiceless] as opposed to [voiced], we’ll use a single feature [±voiced]. The symbol “±” indicates that “voiced” is a binary feature. The term binary means ‘two-valued’: either “plus” or “minus”. To put it simply, a segment is either [+voiced] or [-voiced], there’s no other possibility. In a similar vein, continuants and stops are distinguished by the feature [±continuant], and nasal and oral sounds by the feature [±nasal]: continuant sounds are, of course, [+continuant], stops are [-continuant]; and, needless to say, nasal sounds are [+nasal] while oral ones are [-nasal].

As said above, each such feature defines two natural classes: for instance, [±voiced] defines the division of sounds into the natural classes of voiced (= [+voiced]) and voiceless (= [-voiced]) ones. The class of voiceless plosives can be defined using three features: they are

1. [-voiced],  
2. [-continuant],  
3. [-nasal].

As opposed to this, voiced stops are [+voiced], but the other two features are identical. The feature specifications [-continuant] and [-nasal] together define a bigger natural class: that of plosives. Within the class of plosives, the feature [±voiced] plays a distinctive role: two plosives can differ from each other in voicing only. For example, p and b are identical except for voicing. If a feature can distinguish phonemes, we call it a distinctive feature. Since p and b are distinct phonemes (as testified by minimal pairs such as pin — bin, etc.), and they are distinguished by voice, voice is a distinctive feature in English.

As we said, features correspond to phonetic gestures (like the vibration of the vocal cords or the lowering of the velum, etc.). Sounds produced by different gestures must, therefore, be characterisable with different features. Let us now consider aspirated vs. unaspirated voiceless plosives. What do they differ in? Well, we have already discussed this question: they differ in that aspirated plosives are pronounced with the vocal cords spread apart. This gesture, which is absent in unaspirated sounds, is expressed with the feature [±spread glottis]: aspirated sounds are, of course, [+spread glottis], unaspirated ones are [-spread glottis]. In fact, linguists generally abbreviate long feature names for convenience, a practice I’ll adopt here, so this feature, for instance, will appear in this course as [±sp.gl.] from now on.

The features [±voiced] and [±sp.gl.] both mark differences in articulation, moreover, both are features relating to the state of the glottis (for this reason, they are called laryngeal gestures): [±voiced] means “narrowed vocal cords”, while [±sp.gl.] means “spread vocal cords”. Let us see how the plosive sounds of English can be characterised by these two features:

(16) [b, d, g] as in bee, day, go: [+voiced, -sp.gl.] (= voiced unaspirated)  
[p, t, k] as in spin, stay, sky: [-voiced, -sp.gl.] (= voiceless unaspirated)  
[pʰ, tʰ, kʰ] as in pin, tie, cow: [-voiced, +sp.gl.] (= voiceless aspirated)

Note that there are nine plosive sounds in English, but only six plosive phonemes, cf. (15)
The feature [±sp.gl.] distinguishes sounds in English, but not phonemes: it is not a distinctive feature. Features which do not distinguish phonemes are called redundant features. We say that voiceless stops in English are, in the default situation (= unless preceded by s) redundantly aspirated: there are no two phonemes in English which only differ in aspiration and nothing else.

We are now in a situation to be able to write down the final version of the allophonic rule of Aspiration:

\[(17)\ \text{Aspiration in English 1}\]

\[
\begin{array}{c}
\text{C} \\
-\text{voice} \\
-\text{cont} \\
-\text{nas} \\
\rightarrow \left[ \text{sp.gl.} \right] /s/ \\
\end{array}
\]

This reads: “Voiceless oral stops are unaspirated when preceded by s.” Of course, when this specific condition does not hold (= the plosive isn’t preceded by s), the default rule comes in, which says:

\[(18)\ \text{Aspiration in English 2}\]

\[
\begin{array}{c}
\text{C} \\
-\text{voice} \\
-\text{cont} \\
-\text{nas} \\
\rightarrow \left[ \text{sp.gl.} \right] \\
\end{array}
\]

To sum up: Aspiration and L-Darkening exemplify allophonic rules. Features introduced by them — like [±sp.gl.] are redundant.

Often, however, we can observe that a feature which is distinctive for one class of sounds is non-distinctive for another. One such feature is [±voiced], clearly distinctive for obstruents, since two obstruent phonemes can differ only in voice, as \( f \) and \( v \), for example. On the other hand, no two nasals can differ in voice: nasal consonants in English are all voiced. This can be formulated as in (19):

\[(19) \ [\text{+nas}] \rightarrow [\text{+voice}] \quad \text{i.e., “All nasals are voiced.”}\]

Such statements are called redundancy rules. This redundancy rule captures an important fact about the relationship between nasality and voicing: nasals cannot differ from each other in voicing, because they cannot be but voiced. As opposed to this, there’s no such redundancy rule for obstruent consonants: they are either voiced or voiceless., and two obstruents can differ in voicing only.

Take approximants now. In connection with aspiration, we mentioned that they are pronounced voiceless when preceded by an aspirated plosive, remember \( \text{twin, cute, play, proud} \), pronounced as \( [\text{twɪn, kjuːt, pləɪ, prɔud}] \), resp. But this is the only environment in which they are voiceless: normally, they are voiced. This, as we have seen, is also true for nasals; moreover, it is valid for vowels as well — that is, all sonorants are generally voiced. Introducing a feature [±sonorant] to distinguish sonorants from obstruents, we can formulate a redundancy rule:
This means that we never find minimal pairs with voiceless vs. voiced sonorants: while *low* is an English word, low isn’t and couldn’t be. Of course, it must be emphasised that voiceless sonorants (notably, approximants) do occur, as just mentioned; but they are only found after a voiceless aspirated plosive, where voiced approximants do not occur. Voiceless approximants are but allophonic variants which occur only when the following allophonic rule can apply:


That is, a non-nasal sonorant consonant (= approximant) is devoiced after an aspirated consonant. Since voiced approximants never occur in this environment, and voiceless ones only occur here, voiced and voiceless approximants are in complementary distribution: voicing is, therefore, redundant for approximants. In sum, voice is a redundant property of all sonorant phonemes.

There are several other phonetic features which characterise important natural classes of English. It is not the aim of this course, however, to give an exhaustive description of all distinctive (or redundant) features of English. We will introduce a few more while discussing important phonological phenomena. Suffice it to say that features express phonetic properties which may or may not serve to distinguish phonemes from each other. In case a feature distinguishes phonemes, it is distinctive; if it doesn’t, it is redundant.

To sum up the idea of contrast, words can be distinguished by speech sounds. Speech sounds that distinguish words realise different phonemes. Phonemes often have two or more pronunciations, that is, allophones, depending on environment. Allophones are phonetically different from each other, but this difference is not contrastive: it doesn’t distinguish words. So, words are distinguished by phonemes. Phonemes, in turn, are distinguished from each other by distinctive features. Allophones of a phoneme may differ in a non-distinctive (redundant) feature, such as [+sp.gl]: redundant features never distinguish phonemes. Finally, a given natural class may be redundantly specified as “plus” or “minus” for a feature which is distinctive for another class of phonemes. Rules specifying such cases are called redundancy rules, such as the one stating that voice is redundant for sonorants.

We mentioned earlier that two sounds are said to represent different phonemes if they contrast, i.e., they distinguish words: we can find minimal pairs with them, e.g., *pin — bin*. In order to contrast, the two sounds must be able two occur in identical environments (in the example above, in word-initial position before a vowel). As we mentioned, it is nevertheless very rare for two sounds to have an identical distribution, that is, to be able to replace each other in all environments. Typically, there are some environments where the contrast isn’t possible.

Let us take a look at /p/ and /b/. They are clearly different phonemes, as testified by a large number of minimal pairs, cf. *pin — bin* as well as *cap — cab, staple — stable*, etc. If you take /p/, you can replace it with a /b/ in word-initial, word-final and word-medial position. There is, however, one instance where this replacement is not possible: after /s/. Take words like *spair, spin, grasps* <spy, spin, grasp>: the /p/ cannot be replaced by /b/, because within a word, /b/ is not possible in English after /s/. (Be careful that we’re not dealing with letters: written <sb>, of course, occurs, as in <Crosby, lesbian>, etc., but it represents *zb*, not *sb!*?)

7 “Within a word” is very important: across words, this restriction does not hold, cf. *nice boy* with *sb*. Compound
Because /b/ cannot occur in this environment, the two phonemes fail to contrast here: they cannot distinguish words in these cases. We say that the contrast between /p/ and /b/ is neutralised after /s/. The same is true for all plosives: only voiceless plosives can occur after /s/, cf. words like sky, stay = [skai, stei]: the voiced [d, g] could occur in place of the voiceless ones. (Note that the voiceless plosives themselves are realised without aspiration after /s/) We’ll have ample occasion later on to look at instances of neutralisation, a really frequent phenomenon in natural language. For the time being, let me present an example from Hungarian. You will recall that in Hungarian, short mid rounded vowels are not possible in word-final position: while jó, kö: <só, kő> are okay, *jó, *kö are not. Note that the long vowels contrast with the short ones, i.e., they distinguish words, cf. minimal pairs like por — pör and örölt — örült. This contrast is neutralised in word-final position, as the short ones are not allowed in this environment.

words count as two words, so crossbones is pronounced with sb, too.