

CHAPTER 1

Introduction

In this chapter ...

In this chapter we explore a system for thinking about, and then describing, English speech sounds. We will see that there are important differences between the usual written system of English and how the system of sounds is structured – so many differences, in fact, that the familiar written system of English could never be used as a transcription of either the structure that lies behind speech or the occurrence of English speech sounds themselves. As we’ll see, in order to work systematically with the sounds of English we need to analyse both *the structure that lies behind speech* (we call this *phonology*) and the nature and occurrence of speech sounds themselves (we call this *phonetics*).

Here, too, we begin to look at some of the principles that govern phonology: the *distribution* of sounds, and how they *contrast*. We draw an analogy between this system and the system, or timetable, of trains, and see that to study phonology is to study part of the ‘timetable of language’.

| | | |
|-----|---|--------|
| 1.1 | Written and spoken English | page 2 |
| 1.2 | More on written and spoken English: the primacy of speech | 3 |
| 1.3 | Speech as a system | 5 |
| 1.4 | <i>Accent</i> and <i>dialect</i> | 7 |
| 1.5 | More on systems and structure | 8 |
| 1.6 | Phonetic observation and phonological generalisation | 10 |
| 1.7 | Transcription types | 12 |
| | Exercises, list of terms, further reading | 14 |

1.1 Written and spoken English

It's critical for our purposes to distinguish between the *written* and the *spoken* systems of English. Although it contains significant clues as to how English was once pronounced, English spelling is unreliable as a guide to recent and present-day pronunciation, so much so that George Bernard Shaw once suggested that the familiar word *fish* should be spelled as <ghoti> – <gh> from *enough*, <o> from *women*, and <ti> from words such as *motion*. Consider also the vowel sound (or sounds) one produces in words such as <oar>. For many speakers of English, particularly those who don't typically pronounce the final *r* of <oar>, the vowel represented by the written symbols <oa> is also found in words such as <auk>, <ought>, <sure> and <ford>, where it's represented by the *written* symbols <au>, <ou>, <ure> and <or>.

The above paragraph introduces a useful convention: when we analyse English, it's convenient to refer to *written* (or *common alphabetic*) *forms* by inserting them within angled brackets, <...>. When we come to analyse the *sounds* of English, we will insert these into different brackets, either /.../ or [...], depending on the kind of transcription of sound we are making (see below, 1.6 and 1.7).

We're usually so familiar with the written form of English that it can mislead us into making wrong assumptions about the sound system. The word <school>, for example, conventionally begins with three common alphabetic symbols, <s+c+h>, but in terms of *sounds*, the word actually begins with two **consonants** (roughly, and just for the moment, an 's' sound and a 'k' sound). Similarly, the word <shore> begins with two symbols, <s+h>, but only one consonant in speech (a kind of 'sh' sound – for the relevant symbol, see chapter 2). And again, for many (though by no means all) speakers of English, the final <r> of words such as <oar>, <ear>, <car> isn't pronounced; for many (though by no means all) speakers of English the final <g> of words like <king>, <song>, <fishing> isn't pronounced. In your studies, as analysts of the English language and its many different varieties, it's always important to distinguish very carefully between the written and the spoken forms of English.

Can you construct other, possibly unusual combinations of letters which 'spell' English words, e.g. <ghoti> = 'fish', <aughturnun> = 'afternoon' (<aught> from <draught>, <ur> from <auburn>, <un> from <lun-atic>)?

1.2 More on written and spoken English: the primacy of speech

Although it's not the primary object of attention here, the written system of English doesn't lack interest. Studying the physical shapes of the letters, analysing how and why such letter shapes differ from each other, and working out how the alphabet developed, is to study **graphology** and its history. The earliest English alphabets were in fact modified forms of alphabetic shapes used for written Latin, but also incorporated some characters (symbols) inherited from the Germanic **runic** alphabet. (For a brief introduction to runes, see Graddol *et al.* 1996: 42 or Crystal 1995: 9 – though it's worth pointing out that the runic alphabet was itself a special adaptation of Greek and Latin symbols.) It's also the case that many present-day English spellings give us significant clues to the spoken histories of the words in question. It's reasonable to suppose, for example, that written vowel shapes like <ea> were, at some point in the history of English, pronounced differently from vowel shapes written as <ee>. That is, <meat> was once pronounced differently from <meet>, despite the fact that in many present-day varieties of English these words are **homophones**. (Homophones are words that sound identical, despite differences in spelling: other examples in my own variety of spoken English are <sea> and <see>, <site> and <sight>.) So spellings can be and often are used by linguists as important evidence bearing on how a language's sound system has developed, and how its history may be reconstructed.

There's another reason why analysing and transcribing speech is an activity properly distinct from the analysis of written language. Human beings learn to speak long before they can write (even assuming they ever learn to write). Speech is for many of us the primary, and certainly the most overt, mode of human communication, while writing systems usually begin life as an attempt to capture speech sounds, implying that speech is a primary medium, while writing is derived from it.

Writing is usually very much more conservative than speech. The English language is incessantly, though often imperceptibly, changing, and these changes often show up first in speech, rather than in the written system. (Many changes never reach the written system at all.) For example, in the last forty years there has been a definite shift in how the vowel shape represented by <a> is pronounced in some prestige varieties of British English (BrE, and on the abbreviation, see the boxed text below) in words like <cat>, <hand>, or the first – and, in BrE, stressed – syllable of <garage>.

I will be using some abbreviations in this book. ‘British English’ will be abbreviated as ‘BrE’, and ‘General American’ – a variety that typically includes the pronunciation of ‘r’ after vowels and finally in a word (*fourth*, *door*) – as ‘GA’. I will explain abbreviations, and any special symbols used here, in boxed text as we work.

Such a shift in pronunciation isn’t at all represented in changed spellings: the spellings of the words affected have remained constant. This means that often enough, students of language look to speech, not writing, when they are thinking through how languages have changed over time.

How many other pairs of homophones can you find in your own variety of spoken English?

The reason these points are being made now is that many students beginning their study of the sound structure of English are so accustomed to thinking of the written system of the language as in some sense ‘primary’ that they may make faulty generalisations about the sound structure of the language they speak. For example, try the following exercise. Construct a list of ten English words – preferably, words comprising one and only one syllable – that begin with:

- one consonant
- two consonants
- three consonants

This simple exercise contains the word ‘consonant’. The term implies something spoken (‘con+sonant’ = ‘sounding together’). The list of words beginning with one consonant generally presents no problem: **monosyllables** (i.e. words of just one syllable) such as *dog*, *cat*, *house*, *sit*, *pin*, *tar* and *cup* make their appearance. But with the list of words that begin with two consonants, problems arise – and they’re almost invariably problems stemming from the fact that you are still thinking in terms of the *written system* of English. ‘Words that begin with two consonants? Well ... How about *ship*?’ The difficulty there is that <ship> certainly appears to begin with two written consonant shapes, but in terms of the *sound structure* of the language, the word actually begins with just *one* consonant. The following lists make this point clear:

Words only *appearing* to begin with two consonants

| | |
|-------|---|
| ship | (graphic <sh> represents <i>one</i> speech sound) |
| chase | (graphic <ch> ditto) |
| thigh | (graphic <th> ditto) |

- there (graphic <th> ditto)
- phone (graphic <ph> ditto)
- Words only *appearing* to end with two consonants**
- fish (graphic <sh> represents *one* speech sound)
- bath (graphic <th> ditto)
- Bach (proper name: graphic <ch> ditto)
- graph (graphic <ph> ditto)

Things get more complicated if we ask about words that begin and/or end with three consonants. ‘Three consonants at the beginning... Well, what about *school*?’ The problem is that the word *school* appears to begin with *three* written consonant shapes (<s>, <c> and <h>), whereas in terms of the word’s sound structure, *only two* consonants are present. The following lists emphasise this pseudo-problem:

- Words only *appearing* to begin with three consonants**
- school (graphic <sch> represents *two* speech sounds)
- phrase (graphic <phr> ditto)
- shrew (graphic <shr> ditto)
- sphere (graphic <sph> ditto)
- Words only *appearing* to end with three consonants**
- graphs (graphic <pgh> represents *two* speech sounds)
- laughs (graphic <ghs> ditto)
- baths (graphic <ths> ditto)

The point bears repeating: from the beginning of our study of the sound structure of English we need to distinguish carefully between the written and spoken systems of the language. Our familiarity with the written system can sometimes mislead us into making wrong generalisations about the sound structure of the language, or into constructing transcriptions of sound which are inappropriate. Notice that we’re not saying that familiar graphic conventions – the conventions of written English – are ‘wrong’. We’re just saying that the familiar written system of English doesn’t offer us the *symbolic consistency* or the *adequacy* we need in order to describe and transcribe the *system* that underlies the way we speak our varieties of English.

1.3 Speech as a system

In the paragraphs above we’ve begun to use the word *system* – the ‘system of writing’, the ‘sound system of English’. What allows us to make the claim that the sound structure of present-day English is a ‘system’?

Chapter 1, section [1.3]

As we'll see, speech sounds are themselves organised within the overall structure of the English language: certain speech sounds **contrast** with other speech sounds, and such contrasts are meaningful. In many spoken varieties of English, for example, there's a perceptible *spoken* difference between a vowel like that represented by the <i> of <sit>, and one like that represented by the <ea> of <seat>, the <e> of <met> and the <ee> of <meet>. <sit> and <met> contain **short vowels** (we'll define the term 'short' more precisely later, see in particular chapters 9–10), while <seat> and <meet> contain **long vowels**. The difference in length is a *meaningful* contrast.

Speech sounds also tend to behave *predictably*. For example, the speech sounds corresponding to the beginning of the written word <pray> form the beginning of a well-structured syllable (about which you can read more in chapter 6), but the speech sounds corresponding to *<rpai> (see boxed text below) do not.

The asterisk occurring before a particular linguistic form indicates a form that isn't merely non-occurring, but deviant. For instance, the made-up word <bripi> doesn't appear to occur in any variety of English, *even though it is well formed* in terms of its sound structure. Its non-occurrence is merely an accidental gap. On the other hand, *<rpai> is ill-formed: a 'p' simply cannot follow an 'r' in order to begin an English word. Such an ordering would violate the underlying principles of how English speech sounds are ordered.

Similarly, the speech sounds corresponding to <grinds> form a well-structured syllable, but those corresponding to *<rgidns> do not; <blue> is fine, but *<lblue> isn't. If you're asked why the asterisked forms are deviant or otherwise unacceptable, you might reply that they're 'difficult to say' or 'impossible to pronounce'. There's a reason for that difficulty or impossibility: there are *principles* operative within the spoken system of English that determine which speech sounds can co-occur with other speech sounds. Knowing those principles is part of our wider (and usually tacit) knowledge of the structure of the English language. Analysis of *spoken* English can reveal a great deal about what those principles are, and how they might be formulated and studied.

By observing your own variety of spoken English, how much data could you amass to support the claim that your use of that spoken system was largely *systematic*?

1.4 Accent and dialect

Another reason why we might want to study the sounds of English systematically is so that we can analyse the richness of English **accents**. We need to discriminate between the terms *accent* and **dialect**. *Accent* refers to features, patterns and phenomena belonging to variations in *speech*. For example, three speakers of English from different parts of the world may all pronounce the same word – say, the word spelled <path> – rather differently: a speaker of a Northern variety of British English (a speaker from, say, Leeds) may characteristically pronounce the word with a short vowel, a speaker of Southern Standard British English may pronounce it with a long vowel, and a speaker who has learned English as a second language may pronounce the final ‘th’ sound rather like some variety of ‘t’. These variations are variations of *accent*. Professional linguists are interested in precisely these variations, and in answering questions about them. Why do they occur? Where did these variations originate? How historically stable are they? Linguists are *not* interested in making personal judgements about the ‘correctness’ or otherwise of particular English accents. Like it or not, *every* user of English ‘speaks with an accent’. Questioning *why* those accents exist, and asking *how* they are patterned, are the proper concerns of linguists. In this field of study, as in any other science, value judgements are irrelevant.

If the term *accent* refers to spoken features of English, then *dialect* refers to variations that include accent, but also include features of syntax and vocabulary. (In linguistics the word for ‘vocabulary’, or our ‘mental dictionary’ of meaningful words, word parts and phrases, is **lexicon**.)

To make this clearer, consider the following sentence (in linguistics, such a sentence is called a **substitution frame**) and fill in the indicated gap with a demonstrative pronoun – a word such as ‘those’ or ‘them’:

He caught the pike between_____weeds

(A pike is a predatory freshwater fish.) Clearly, you could insert the word *those* into the frame. But for many speakers of English, you could also insert *them* (‘them weeds’). For other speakers, you could insert the form *dey* (and such speakers would also tend to use the form *de* for the definite article – *de pike*). Such variations do not just involve pronunciation, they also involve **grammar** – in this instance, the system of pronoun forms. As such, the variations (including accent, but also embracing other syntactic features of English) belong to the study of *dialect*. They are *dialectal variations*. (Note: please distinguish between the term *dialectal* and the term *dialectical*. This last term belongs properly to philosophy, rather than to linguistics.)

Other examples of dialectal variation: for many speakers of English, *I need this plug mending* is a perfectly usual structure – but not for speakers of some varieties of Scots English, for whom *I need this plug mended* would be normative. This difference, a syntactic difference involving the inflectional **morphology** (roughly, the word-building) of verb forms, is dialectal. Or again, I could refer to an acquaintance *raising her little finger*, while you might normatively refer to her *raising her pinkie*. The difference, between *little finger* and *pinkie*, is a variation that is said to be *lexical* (involving the lexicon, the ‘mental dictionary’ of a speaker).

Every English speaker uses some form of dialect. By historical accident, political choice, or societal pressure (or perhaps all three), the particular dialect used may have become some kind of *standard* form of English, a prestige form, a form taught and transmitted (‘Don’t say *them* weeds, Christopher! Say *those* ...’). But – and uncomfortably for self-appointed guardians of the ‘purity of the English language’ – ‘standard’ forms of English are themselves dialects, and for dialect speakers, whether they be from Somerset, Scotland or Singapore, their native dialect is a perfect communicative medium, neither better nor worse than other dialects. Just as they attempt to study accents with scientific detachment and impartiality, so linguists bring the same analytical detachment to the study of dialect. The questions that interest the linguist are: How did this dialect originate? How has it changed over time? What factors have caused it to change? What is the relationship between spoken and written forms of this particular dialect?

What *accent* of English do you think you use? Would your immediate circle of friends and family agree that you use that form of accent? (Try asking them.) What *dialectal features* can you find in your own variety of English?

1.5 More on systems and structure

I’ve talked about structures and systems, and about how the spoken system of English is rather different from the written. But what sort of object *is* the sound structure of English? How can we study it? What does it mean, ‘making generalisations about’ the behaviour of certain items within that system?

To help understand the word ‘structure’, and what it entails in this kind of linguistic study, I’m going to introduce an analogy. The analogy is between the behaviour of sounds, and the behaviour of trains. The analogy isn’t my own; it’s a reworking of an analogy constructed by the Swiss linguist Ferdinand de Saussure in the early years of the twentieth century (Saussure 1983: 107). Here goes.

For many years I took a morning train to work. The train was the 07.52 from Greenfield to Manchester. Sometimes this train arrived with two yellow carriages, sometimes with four blue ones. Sometimes the train arrived, and subsequently departed, late. Sometimes it didn’t arrive at all.

Now, whatever the physical appearance of the train, and however late it was, this didn’t alter the fact that the train itself was still the 07.52 from Greenfield to Manchester.

The point is this. The identity of the train I took to work depended on its place in the *timetable*, and that timetable is a *structure*. Even when the 07.52 was late, cancelled, or varied in colour it was still, always, the 07.52, whose identity was guaranteed by the timetable of trains – specifically, by the fact that the 07.52 behaved in a certain way (it travelled from Greenfield to Manchester, not to Blackpool, Bolton or Paris), and by the fact that this train *wasn’t*, and could never be, the 08.05 or the 08.15.

When we start to think about how the sounds of English or any other language ‘work’, we have to understand that these speech sounds operate in terms of a *structure*. Whatever the physical, or acoustic, properties of a sound (for example, whether the sound represented by the symbol ‘g’ is pronounced loudly or softly, spoken, whispered, or sung), this doesn’t alter the fact that in English we still understand it as that particular sound.

How can we prove, or infer, the existence of a linguistic structure? We can infer the timetable (or structure) of the running of trains by looking at their physical arrival and departure, and similarly, when we start thinking about the sound structure of English, we can infer a great deal from the physical nature and distribution of the speech sounds themselves – that is, whether a particular sound can begin a syllable, or end a syllable, or both, or whether it can occur after ‘s’ in the beginning of a syllable, or not... and so on. However, while the railway timetable represents the underlying structure of the running of trains, it doesn’t tell us whether the trains are red or yellow. These are part of the physical characteristics of the trains themselves, and not part of the underlying timetable or structure. And when a linguist thinks about structure, he or she is thinking *primarily about the system, rather than the actual physical implementation of that system*.

Because it’s useful to have a term for that kind of thinking, let’s use one: the sound structure of a language is the **phonology** of that language, and the physical manifestation of the actual sounds is the **phonetics** of that language.

1.6 Phonetic observation and phonological generalisation

From 1.5 it follows that there are different kinds of way in which we could study the sound structure of English. We could focus on the physical characteristics, and acoustic properties, of various sounds (this would be a largely phonetic focus). Or we could study the relatedness (or lack of relatedness) of particular sounds, as these occurred as part of the structure of English, and we could also study how these same sounds, or classes of sounds, were distributed within the syllable, i.e. which consonants, for example, could function as **pre-vocalic** (occurring before vowels), or **post-vocalic** (occurring after vowels); which consonants could occur after the speech sound ‘s’; which consonants were of apparently restricted distribution, and so on. If we studied the sound structure of English in this way, we would be thinking *phonologically*.

The prime focus of this book is on (English) phonology. But there’s a problem here: to get a secure phonological generalisation, we must also take into account some phonetic, that is, acoustic, detail. Why? Precisely because we can make phonological *inferences* from that detail.

To give further definition to the material we’ve started to think about, try an experiment. This involves the speech sound which is invariably written as <p> – the ‘p’ you get in English monosyllables such as *pin*, *spin* or *nip*.

First, hold the open palm of your right/left hand about 5cm from your mouth. Now pronounce the monosyllable *pin*, clearly and distinctly. No need to shout or whisper, just say the word clearly, and without undue emphasis. As you utter the ‘p’ of *pin*, do you feel anything on your open palm?

You should feel a definite puff of air as you pronounce the syllable-initial consonant. This is because the English speech sound represented (so far) by ‘p’ is produced, when it occurs at the beginning of words like *pin* or *path*, with a rapid explosive release of air (on the precise mechanism involved, see chapter 2). Such a puff of air, occurring in this environment (initially, in a stressed syllable), is known as *plosion* or *aspiration*.

Next, and while holding your open palm in the same position, utter the monosyllable *nip*, and notice what happens when you pronounce the ‘p’. This time, was there the same puff of air on your open palm? I doubt it. In fact, in many varieties of English, particularly in quick or casual speech, the ‘p’ sound that occurs finally in a syllable is accompanied by no explosive release of air at all.