目 录

Chapter 1	Description of English Speech Sounds	1
	1. Basic concepts of phonetics	2
	1.1 Branches of phonetics	2
	1.2 Main features of phonetics	3
	2. Phonetic transcriptions	1
	3. Phonemes and allophones	5
	4. Physiology of speech sounds	7
Chapter 2	English Vowels	13
	1. Classification of English phonemes	15
	2. Description of vowels	15
	3. Cardinal Vowels	16
	4. Classification of English vowels	17
	4.1 Description of monophthongs	18
	4.2 Description of diphthongs	27
	4.3 Description of triphthongs	31
Chapter 3	English Consonants	49
	1. Description of consonant phonemes	51
	1.1 Place of articulation	51
	1.2 Manner of articulation	52
	1.3 Voicing	52
	2. Description of English consonants	55
	2.1 Plosives	55
	2.2 Nasals	59
	2.3 Fricatives	50
	2.4 Affricates t/J and $d/d_3/$	53
	2.5 Approximants	54
Chapter 4	English Sounds and Spellings	31
	1. The notion of vowel and consonant letters	33
	2. English spelling system	33
	3. Sound-spelling regularities	34
	3.1 Vowel letters	34

	3.2 Consonant letters	88
Chanter 5 S	vllables and Consonant Clusters	99
1.	Svllable	
1	1.1 Svllable structures	
	1.2 Syllabic consonants	
	1.3 Syllable types	
2.	Consonant clusters	
	2.1 Cluster constraints	
	2.2 Characteristics of consonant clusters	
3.	. Syllabification	106
Chapter 6 V	Vord Stress and Vowel Reduction	115
1.	. Stress	117
	1.1 Classification of stress	117
	1.2 Levels of word stress	117
	1.3 Orthographic representation of stress	118
	1.4 Characteristics of stressed and unstressed syllables	118
2.	. Vowel reduction	119
	2.1 Unstressed vowels in open syllables	120
	2.2 Unstressed vowels in closed syllables	120
	2.3 Unreduced vowel	121
3.	. Pitch patterns of a word	121
	3.1 Two-syllable words	122
	3.2 Three-syllable words	122
	3.3 Four-syllable words	123
	3.4 Five-syllable words	123
	3.5 More-than-six-syllable words	123
4.	. Placement of word stress	124
	4.1 Two-syllable words	124
	4.2 Three-syllable words	125
	4.3 More-than-three-syllable words	125
	4.4 Prefixes	125
	4.5 Suffixes	126
5.	. Stress of compounds	128
	5.1 Noun compounds	129
	5.2 Verb compounds	131

5.3 Adjective compounds	
5.4 -ing + noun compounds	
5.5 Adverb + -ing compounds	
Chapter 7 Sentence Stress and Weak Forms	
1. Sentence stress	
1.1 Basic sentence stress	
1.2 Levels of sentence stress	
2. Weak forms	
Chapter 8 Aspects of Connected Speech	
1. Rhythm	
1.1 Stress-timed rhythm vs. syllable-timed rhythm	m166
1.2 Rhythm and music	
1.3 Rhythm and meter in English poetry	
1.4 Rhythm group	
1.5 Eurhythmic structure	
2. Linking	
2.1 Consonant + vowel linking	
2.2 Vowel + vowel linking	
2.3 Linking [r]	
2.4 Intrusive [r]	
3. Assimilation	
3.1 Direction of assimilation	
3.2 Assimilation in voicing	
3.3 Assimilation in place of articulation	
3.4 Assimilation in manner of articulation	
3.5 Coalescent assimilation	
4. Elision	
4.1 Elision of vowels	
4.2 Elision of consonants	
Chapter 9 English Intonation	
1. Tonality	
1.1 Intonation phrases	
1.2 The size of intonation phrases	
1.3 The intonation phrase boundary	

2	2. Tonicity	197
	2.1 Stress and accent	197
	2.2 Transcription conventions	197
	2.3 The nuclear placement	198
	2.4 Broad and narrow focus	201
3	B. Tone	202
	3.1 Types of nuclear tones	202
	3.2 The meanings and applications of nuclear tones	
Chapter 10	Pitch Contour	231
	1. The components of the intonation phrase	233
	2. Pitch range	236
	3. Marking system for intonation contour	237
	4. Pitch contour of prenuclear segment	237
	4.1 The prehead	237
	4.2 The head	
	5. Pitch contour of nuclear segment	241
	5.1 Pitch contour of falls	
	5.2 Pitch contour of rises	244
Answers to	Diagnostic Test	
References		270

Chapter 1

Description of English Speech Sounds

Fundamental Knowledge

Although language is obviously composed of sound, speech sounds came to be the main focus of linguistic investigation only in the 20th century. In the 19th century, linguists were more interested in written rather than spoken language. Only with the work of the Swiss linguist Ferdinand de Saussure in the early 20th century did linguists recognise the primacy of sound in human language and the secondary nature of writing.

To be a successful language learner, it is essential to have a better understanding of how the speech sounds of English are produced. Such knowledge will enable us to understand why we have a 'Chinese accent' when speaking English, since the Chinese accent results partially from an inability to produce the speech sounds of English. In addition, it will enable us to take the necessary steps for self-correction of our pronunciation problems.

1. Basic concepts of phonetics

The study of how speech sounds are produced, how they can be classified and how they differ from each other is known as *phonetics* (语音学).

1.1 Branches of phonetics

Phonetics has three branches: *articulatory phonetics* (发音语音学), *acoustic phonetics* (声学 语音学) and *auditory phonetics* (听觉语音学) (see Fig. 1.1).

- Articulatory phonetics studies the speech sounds from the speaker's point of view (Speaker A in Fig. 1.1); in other words, it studies how a speaker uses his speech organs to articulate sounds.
- Acoustic phonetics studies the speech sounds by looking at the sound waves; in other words, it studies how speech sounds are transmitted through the air from one person to another.
- Auditory phonetics studies the speech sounds from the hearer's point of view (Speaker B in Fig. 1.1); in other words, it studies how the sounds are received by the hearer or recognised by the brain.



Fig. 1.1 Branches of phonetics

The formulation of the speech message in the brain of the speaker and the interpretation of it in the brain of the listener are branches of *psycholinguistics* (心理语言学). In this book, our emphasis will be on articulatory phonetics, this being in many ways the most accessible branch of the subject, and the one with most applications for the learners of English.

1.2 Main features of phonetics

A distinction is often made between the *segments* (音段) and the *suprasegments* (超音段) as shown in Fig. 1.2.



Fig. 1.2 Phonetic features

Segments are units of sounds within a language. In English, the set of phonemes consists of 20 vowel sounds and 24 consonant sounds. *Phonetic features* (语音特征) are more than a matter of vowels and consonants. Indeed, words are represented in speech by vowels and consonants, and by stressed and unstressed syllables, too. Beyond the word level, sentences or utterances display a kind of rhythm and are accompanied by pitch variations, namely, intonation. Thus, the features of speech which involve more than one sound in an utterance are known as suprasegments. The suprasegmental features in English are stress, intonation and how sounds change in connected speech.

A systemic description of English phonetics can be found in the *systemic linguistics* (系统语言 学). In this description, the sound system of English is regarded as a hierarchy of units at four ranks. It moves from the largest or most inclusive unit, the *pitch contour* (音调模式) or *intonation phrase* (语调模式), to the *rhythm group* (节奏群) or *foot* (音步), to the *syllable* (音节), and finally to the smallest unit, the *phonemic unit* (音位单位) or *phoneme* (音位). Every intonation phrase consists of one or more feet, every foot of one or more syllables, every syllable of one or more phonemes. The following is an example of an English sentence with the various ranks in the *phonetic hierarchy* (语音层级) as shown in Fig. 1.3:



Fig. 1.3 Phonetic hierarchy in the sentence I told her to wait in the library

As this simplified example indicates, an intonation phrase consists of one continuous pitch contour with a single dynamic (falling) tone. Each foot consists of one single stress, descending gradually owing to the subglottal pressure and then falling quickly towards the nucleus. The syllables are written with some vowels omitted to indicate that 't' and 'th' do not form syllables themselves. They are better regarded as pseudo-syllables¹ tacked onto real syllables² and containing open transitions. The last rank consists of phonemes and of open transitions between consonants which are represented by $[\bullet]$.

The above account of the phonetic hierarchy of English emphasises an essential point in pronunciation learning. English phonetics is not merely a collection of phonemes whose pronunciation has to be learned and later strung together with superimposed stresses and intonations, but rather a hierarchy of units of different kinds, each with its own learning focus.

2. Phonetic transcriptions

Phonetic transcriptions (音标) are usually written on the basis of the **International Phonetic Alphabet** (IPA for short, 国际音标), in which each English sound has its own symbol. For example, the IPA-based phonetic transcription of *go* is /gəu/, and the transcription of *do* is /du:/. In spelling, they both end in the vowel letter 'o', but their phonetic transcriptions are different, so they are pronounced differently. Dictionaries tell us about the pronunciation of words. In English dictionaries, phonetic transcriptions are necessary because the spelling of a word itself does not tell us how we should pronounce the word.

Phonetic transcriptions may aim to go further and specify the precise phonetic realisation. In all systems of transcription we may distinguish between broad transcription and narrow transcription. *Broad transcription* (宽式标音) uses one symbol to illustrate one phoneme, and

¹ pseudo-syllables (假音节):指句子中应该弱读或省音的虚词,如介词、连词、冠词、人称代词等。

² real syllables (真音节): 指具备句子重音的实义词,如名词、动词、形容词、副词等。详见第7章1.1。

is normally used in dictionaries and textbooks. In broad transcription, a phoneme is illustrated between the slashes (/ /). *Narrow transcription* (严式标音) or *allophonic transcription* (音位变体标音) uses more than one symbol to illustrate a phoneme, and is actually the transcription required by the phoneticians in their study of speech sounds. Conventionally, in narrow transcription, the square brackets ([]) are used to transcribe *allophones* (音位变体) of a phoneme.

The difference between broad and narrow transcription is a continuum. One particular form of the broad transcription is a *phonemic transcription* (音位标音), which disregards all allophonic differences, and is not really a phonetic transcription at all, but a representation of phonemic structure. For example, one particular pronunciation of the English word *little* may be transcribed using the IPA as /'lɪtəl/ or ['lɪtɬ]. The broad, phonemic transcription indicates merely that the word ends with phoneme /l/, but the narrow, allophonic transcription indicates that this final [ɬ] is *dark* (*velarised*) (模糊/软腭化). Therefore, narrow transcription shows more detailed phonetic representation.

The advantage of the narrow transcription is that it helps learners to get exactly the right sounds, and allows linguists to go in depth into slight variations of a phoneme. The disadvantage is that a narrow transcription is rarely representative of all speakers of a language. Most Americans and Australians, for example, would pronounce the /t/ of *little* as a *tap* (闪音) [r]. Many people in England would say /t/ as [?] (a *glottal stop*, 声门塞音) and / or the second /l/ as [u] or something similar. A further disadvantage in less technical contexts is that narrow transcription involves a larger number of symbols which may be unfamiliar to the learners of English.

The advantage of the broad transcription is that it usually allows statements to be made which apply across a more diverse language community. It is thus more appropriate for the pronunciation data in foreign language dictionaries, which may discuss phonetic details in the preface but rarely give them for each entry. A rule of thumb in many linguistic contexts is therefore to use a narrow transcription when it is necessary for the point being made, but a broad transcription whenever possible.

3. Phonemes and allophones

Phonemes and allophones are two of the most basic and important concepts in phonetics and *phonology* (音系学). Phonemes are the smallest contrastive units in the sound system of a language. It is called so because

- it is articulated differently,
- it occurs in the same phonetic context, and
- the change affects the meaning, i.e. any substitution of a sound by another leading to a change in meaning.

In English, for example, /p, b/ and /iɪ, ɪ/ are phonemes; they help differentiate between words that have different meanings, as in *pit* /pɪt/ and *bit* /bɪt/, *sheep* /ʃiɪp/ and *ship* /ʃɪp/.

On the other hand, allophones are phonetic varieties of the same phoneme. They form a set of

sounds that

- do not change the meaning of a word,
- ✤ are all very similar to one another, and
- occur in phonetic contexts different from one another.

For example, $[p^h]$ as in *peak* and [p] as in *speak* are allophones for the same phoneme /p/ in the English language because they cannot distinguish words. In fact, they occur in *complementary distribution* (互补分布). While a native speaker would consider the 'p' sound in *peak* and *speak* to be the same sound, the 'p' is actually pronounced differently in each word: the first is *aspirated* (送气) and the second is *unaspirated* (or *plain*, 不送气). That puff of air is called *aspiration* (送气现象), which is transcribed with a superscripted [h]. Plain [p] also occurs as the 'p' in *cap* [k^hæp], or the second 'p' in *paper* [p^heɪpə]. It is worth noticing that the differences of the allophones are unimportant for the *lexical identity* (词汇辨析) of words though they may be important stylistically and sociolinguistically.

Different languages can have different groupings for their phonemes and allophones. [p] and [p^h] belong to the same phoneme in English, namely, they are two different allophones of the phoneme /p/, but two different phonemes in Chinese. Therefore, in Chinese, switching [p] and [p^h] does change the meaning of the word.

Occasionally, allophone selection is not conditioned but may vary from person to person and occasion to occasion. For example, the final consonant of *cap* may not be released by some speakers so there is no audible sound at the end of this word. In this case, it is the same word pronounced in two different ways: [kʰæpʰ] or [kʰæp¬] with the *diacritic* (变音符号) '¬' indicating *no audible release* (听不到爆破) in IPA symbols; in other words, the sound is not actually heard. The difference may be caused by dialect, habit or individual preference, instead of by any distribution rule. Thus, when two sounds appear in the same phonetic context, but do not make a difference in meaning, such a phenomenon is called *free variation* (自由变体). Free variation is often found in regional differences. For example, most Americans pronounce the word *either* as /'iːðə/, whereas most British people say /'aɪðə/. Individual preferences may also determine the use of /dɪ'rekʃn/ vs /daɪ'rekʃn/ for the word *direction*. /i!/ and /e/ in the respective pronunciations of *economics* /_iiːkə'npmīks/ vs /_iekə'npmīks/; or /e/ and /eɪ/ in the respective pronunciations of *again* /ə'gen/ vs /ə'geɪn/. In dictionaries, regional free variants are often listed side by side.

However, when two different forms are identical in every way except for one sound segment which occurs in the same place in the strings, the two words are said to form a *minimal pair* (最 小对立对). For example, in English, *pill* and *bill* are a minimal pair and so are *pill* and *till, till* and *kill, kill* and *dill, dill* and *gill*. Accordingly, we can conclude that /p, b, t, d, k, g/ are phonemes in English. Then all these sound combinations together constitute a minimal set; they are identical in form except for the initial consonant. Also, *beat, bit, bet, bat, boot, but, bait, bite, boat* are a

minimal set. They are identical in form except for the vowel /i1, I, e, æ, u1, Λ , eI, aI, ϑ u/. Minimal pairs are useful because they help us to figure out whether a sound is a phoneme or an allophone.

4. Physiology of speech sounds _____

Human beings all use the same speech organs to produce the sounds we have become accustomed to producing. The set of sounds we acquire, however, may vary: a child brought up in an English-speaking environment will develop the phonemes of English; a Chinese-speaking child will develop a different set; and so on. We also learn to use our voices in different ways: an English-speaking child will learn to use appropriate stress and intonation patterns, and a Chinese-speaking child will learn to use tones to give distinctive meanings to the same set of sounds.

It seems, however, that after childhood our ability to adopt an unfamiliar set of sounds diminishes somewhat. Thus, to a certain extent we need to learn to use our speech organs in a new way in order to produce sounds in the English language or to lose sounds from our own language which are not appropriate in the English language. In other words, we need to adapt our speech organs to a new habit when we try to speak English.

The organs, which take part in the production of speech sounds, are called *speech organs* (发音器官). When we speak, air comes out through the lungs and it is interfered at various places for the production of sounds. Sounds cannot be produced without air. A variety of muscles interact to produce changes in the configuration of the vocal tract so as to allow parts of the speech organs to come into contact with other parts, i.e. to articulate. Phoneticians term these anatomical bits and pieces the *articulators* (发音器官)—hence the term for the branch of science is known as articulatory phonetics.

The organs of speech fall into three groupings: *nasal cavity* (鼻腔), *oral cavity* (口腔) and *pharyngeal cavity* (咽腔).

- Nasal cavity—the nose: the passage through which the air passes when we breathe normally. If the oral cavity is blocked, sounds produced are nasalised, and called nasal sounds.
- ◆ Oral cavity—the mouth: the part of the mouth from the lips to the *uvula* (小舌), including the tongue, the lips and the *lower jaw* (下腭).
- ◆ Pharyngeal cavity—the throat: the top of the *larynx* (喉) and the back of the *soft palate* (软腭).

Figure 1.4 shows the location of the main areas of the head and neck associated with the production of sounds.



Fig. 1.4 Diagram of the organs of speech

Articulation happens when the airstream is interrupted, shaped, restricted or diverted. Producing different speech sounds depends on the movement of speech organs. It is essential to know the placement, the movement and the functions of some important articulators.

The lips

The upper and lower lips can be pressed together to produce *bilabial* sounds (双唇音) /p, b, m, w/, brought into contact with the teeth to produce *labiodental* sounds (唇齿音) /f, v/, or rounded to produce the lip-shape for vowels like /uː/.

The teeth

The upper teeth only take part in the production of consonant sounds. The lower teeth do not participate in producing the sounds. Sounds produced with the tongue touching the upper front teeth are called *dentals* (齿间音, such as / θ , ð/).

The alveolar ridge

The *alveolar ridge* (齿龈) is the part between the upper teeth and the *hard palate* (硬腭). The sound produced with the tongue touching the alveolar ridge is called *alveolars* (齿龈音), such as /t, d, s, z, n, l/.

The velum (the soft palate)

The *velum* (软腭) or soft palate is seen in Fig. 1.4. It is the roof of the mouth, which separates the oral and nasal cavity. The end part of the soft palate is called uvula. When the velum is lowered, the nasal sounds (/m, n, η /) are produced. When it is raised, the air passes out through the oral cavity and the oral sounds are produced.

The tongue

The tongue is an important organ of speech. It has the greatest variety of movement. It is usually divided into five parts: the *tip* (舌尖), the *blade* (舌端), the *front* (舌前), the *back* (舌后)

and the *root* (舌根). A number of vowels are produced with the help of the tongue. Vowels differ from each other because of the position of the tongue. The tip of the tongue helps to produce such consonants as /t, d, s, z/. The blade of the tongue helps to produce /ʃ, ʒ, tʃ, dʒ/. The front of the tongue helps to produce palatal sound /j/ and the back of the tongue helps to produce /k, g, ŋ/.



Fig.1.5 Diagram of the tongue

The vocal cords

The larynx contains the *vocal cords* (声带), two flaps of flexible, connective tissues which can open and close. When we breathe in or out, the vocal cords are open. This is the position of the production of *unvoiced* (*voiceless*) sounds (清音), such as /p, t, k/. When the edges of the vocal cords come close together, the air which passes between them makes them vibrate, resulting in *voiced* sounds (浊音) as in /b, d, g/. Therefore, the main function of the vocal cords is to produce unvoiced and voiced sounds.

Finally, the movement of the lower jaw is important in speaking, but the lower jaw is not an articulator in the same way as others because it cannot itself make contact with other articulators. The above descriptions and functions of the organs of speech may help us to produce the consonants and vowels in a proper way. More detailed explanation of the articulation of individual phonemes will be discussed in Chapters 2 and 3.



1. Phonetic transcription

First write the broad (or phonemic) transcription of the five words in the table. Then listen to these words. Concentrate on the actual realisation of 't' sounds, and write down the narrow (or allophonic) transcription of 't' sounds in each word. The first one has been done for you as a reference.

	Word	Broad transcription	Narrow transcription
	top	/tpp/	[thop]
Dhoneme /t /	stop		
r noneme / t/	eighth		
	tree		
	button		

2. Minimal pairs

The following minimal pairs are identical except for one sound segment which may appear in word-initial, word-mid or word-final positions. Put them into the proper column according to the three positions. The first three pairs have been done for you as a reference.

mat — mad	late — rate	tell — tail	bade — bird	peace — Pete
Patty — batty	fleece — fleet	Pete — beat	chick — cheque	bail — pale
plate — played	fist — fished	collect — correct	pen — Ben	sinner — singer
line — Rhine	leak — league	fill — fell	worth — worse	middle — medal
loyal — royal				

Word-initial position	Word-mid position	Word-final position
late — rate	tell — tail	mat — mad

3. The speech organs

Fill in the blanks with the name of each particular part of the speech organs by consulting the text.





- 14)
- 16)



Fig. 1.6 Exercise on speech organs

4. Basic knowledge

Complete the following sentences.

- 1) The soft palate is also called _____
- 2) The back of the tongue helps to produce velar sounds /k/, /g/ and ______.
- 3) /p, b, m, w/ are produced with the help of _____
- 4) The sounds produced with the help of the upper teeth are called ______ sounds.
- 5) The organs that take part in the production of speech sounds are called ______.
- 6) The study of speech organs is important to know the nature of ______.
- 7) The positional variants of the same phoneme are known as ______.
- 8) In English, [1] and [1] are allophones of the phoneme /l/ because they occur in _____; in other words, they never appear in the same sound contexts. [1] always occurs before vowels (as in *leaf* and *look*), while [1] comes after vowels (as in *feel*) and before consonants (as in *cold*).
- 9) Using the diacritics provided by the IPA, ______ captures as many aspects of a specific pronunciation as possible and makes very subtle distinctions between sounds.
- 10) /p/ and /b/ can appear in initial position (as in *pin* and *bin*) and also in final position (as in *rope* and *robe*). They are in phonemic contrast and said to form a(n) ______ because they occur in the same environment and distinguish meaning.