PAUL CARLEY AND INGER M. MEES

AMERICAN ENGLISH PHONETICS AND PRONUNCIATION PRACTICE

ROUTLEDGE

American English Phonetics
and Pronunciation Practice

American English Phonetics and Pronunciation Practice provides an accessible introduction to basic articulatory phonetics for students of American English. Built around an extensive collection of practice materials, this book teaches the pronunciation of modern standard American English to intermediate and advanced learners worldwide.

This book

- provides an up-to-date description of the pronunciation of modern American English;
- demonstrates the use of each English phoneme with a selection of high-frequency words, both alone and in context, in sentences, idiomatic phrases, and dialogues;
- provides examples and practice material on commonly confused sounds, including illustrative pronunciation diagrams;
- is supported by a companion website featuring complete audio recordings of the practice material to check your pronunciation against;
- can be used not only for studying pronunciation in the classroom but also for independent practice.

American English Phonetics and Pronunciation Practice is essential reading for any student studying this topic.

Paul Carley is a lecturer in English at Fakeeh College for Medical Sciences, Jeddah, and has held posts at the University of Leicester, the University of Bedfordshire, and the University of Applied Sciences Utrecht. He is a regular lecturer at the UCL Summer Course in English Phonetics.

Inger M. Mees is an associate professor in the Department of Management, Society, and Communication at the Copenhagen Business School. She has formerly held lectureships at the universities of Leiden and Copenhagen. She is on the academic staff of the UCL Summer Course in English Phonetics.

American English Phonetics and Pronunciation Practice

Paul Carley and Inger M. Mees

Routledge
Taylor & Francis Group

LONDON AND NEW YORK

Contents

English phonemic transcription key

Consonants

Voiceless			Voiced		
Plosives					
p	*pet, lap*	pɛt, læp	b	*bet, lab*	bɛt, læb
t	*town, mat*	taʊn, mæt	d	*down, mad*	daʊn, mæd
k	*cap, luck*	kæp, lək	g	*gap, lug*	gæp, ləg
Affricates					
ʧ	*chin, batch*	ʧɪn, bæʧ	ʤ	*gin, badge*	ʤɪn, bæʤ
Fricatives					
f	*fast, safe*	fæst, seɪf	v	*vast, save*	væst, seɪv
θ	*thigh, breath*	θaɪ, brɛθ	ð	*thy, breathe*	ðaɪ, brið
s	*sink, face*	sɪŋk, feɪs	z	*zinc, phase*	zɪŋk, feɪz
ʃ	*shy, wish*	ʃaɪ, wɪʃ	ʒ	*measure*	ˈmɛʒər
h	*hat*	hæt, –			
Nasals			Voiced		
			m	*meet, team*	mit, tim
			n	*nice, fine*	naɪs, faɪn
			ŋ	*–, sing*	–, sɪŋ
Approximants					
Lateral (approximant)			l	*late, sail*	leɪt, seɪl
(Median)**approximants**			j	*yes, –*	jɛs, –
			w	*wait –*	weɪt, –
			r	*red, dare*	rɛd, dɛr

Notes

1 This phonemic transcription system is, with few exceptions, the same as that in the *Routledge Dictionary of Pronunciation for Current English* (Upton and Kretzschmar 2017).
2 Examples show initial and final position. Note that /ŋ/ does not occur initially; /ʒ/ is virtually restricted to medial position; /h j w/ do not occur finally.
3 The voiceless/voiced contrast is found only in the plosives, affricates, and fricatives.
4 Stress is shown by [ˈ] placed *before* the syllable, e.g., *open* /ˈoʊpən/, *forget* /fərˈgɛt/.
5 Syllabic consonants (Section 8.2) are shown by [ˌ] *below* the symbol, e.g., *apple* /ˈæpl̩/, *mission* /ˈmɪʃn̩/.
6 Tapped /t/ (Section 2.8) is shown as [t̬] e.g., *pretty* /ˈprɪt̬i/, *thirty* /ˈθərt̬i/.

Vowels

The words shown in SMALL CAPITALS are the **keywords** used throughout this book to refer to the vowels. These were first introduced in Wells (1982).

Vowel	Keyword	Additional spellings
Checked		
ɪ	KIT /kɪt/	g**y**m, man**a**ge, b**u**sy, **E**ngland, g**ui**lt
ɛ	DRESS /drɛs/	br**ea**d, fr**ie**nd, s**ai**d, f**ai**r, v**a**ry, th**ei**r
æ	TRAP /træp/	pl**ai**d
ə	schwa /ʃwɑ/	s**o**n, y**ou**ng, bl**oo**d
ər [ɚ]	schwar /ʃwɑr/	g**ir**l, t**er**m, h**ear**d, w**or**d, j**our**ney, c**ur**ry
ʊ	FOOT /fʊt/	p**u**t, w**ou**ld, w**o**man
Free monophthongs		
i	FLEECE /flis/	n**ea**t, th**e**se, techn**i**que, bel**ie**f, b**ee**r, w**ei**rd
ɑ	PALM /pɑm/	n**o**t, sw**a**n, kn**ow**ledge, st**a**rt, h**ea**rt, mem**oi**r
u	GOOSE /gus/	r**u**de, s**ou**p, sh**oe**, d**o**, cr**ew**, t**ou**r, p**oo**r
(ɔ	THOUGHT /θɔt/	c**au**ght, s**aw**, w**a**lk, br**oa**d)
Free diphthongs		
eɪ	FACE /feɪs/	l**ai**d, m**ay**, w**ei**gh, th**ey**, br**ea**k
aɪ	PRICE /praɪs/	tr**y**, l**ie**, b**uy**, g**ui**de
ɔɪ	CHOICE /tʃɔɪs/	b**oy**
oʊ	GOAT /goʊt/	n**o**se, bl**ow**, s**ou**l, t**oe**
[o]	SPORT /sport/	m**ore**, f**our**, **oar**, d**oor**, w**ar**
aʊ	MOUTH /maʊθ/	dr**ow**n
Unstressed vowels		
ə	schwa	comm**a**, **a**bility, usel**e**ss, bon**u**s, fam**ous**
ər [ɚ]	schwar	coll**ar**, und**er**, forg**et**, V**ir**ginia, capt**ure**, mart**yr**
i	FLEECE	happ**y**, mon**ey**, hipp**ie**, med**i**ate

Notes

1 Some American accents have an additional vowel /ɔ/, which is used instead of PALM /ɑ/ in the lexical sets THOUGHT and CLOTH (Section 5.2.1), e.g., *law, loss* /lɔ lɔs/.

2 The allophone of GOAT before /r/ is strikingly different and shown as [o], e.g., *more* [mor]. We have assigned it a separate keyword: SPORT. In accents that have an additional THOUGHT vowel, it is regarded as an allophone of THOUGHT; see Sections 5.2.1 and 5.2.2.

Phonetic symbols and diacritics

This is a list of the phonetic symbols used in this book. We have not included here the symbols used for English phonemic transcription (see previous pages).

[ç]   voiceless palatal fricative, as a realization of /hj/, e.g., *huge*
[ɦ]   voiced glottal fricative, e.g., *ahead*
[ɫ]   voiced velarized alveolar lateral approximant ("dark /l/"), e.g., *well*
[ɹ]   voiced post-alveolar approximant, e.g., *very*
[ʔ]   glottal plosive, e.g., as a realization of /t/ before consonants, *at school* [əʔ ˈskul], or glottal reinforcement of a voiceless plosive, e.g., *background* [ˈbæʔkgraʊnd]
[t̬]   tapped /t/, e.g., *city* [ˈsɪt̬i]
[̥]   devoiced consonant, e.g., *rob* [rɑb̥] (normally below the symbol, but above for descending symbols, e.g., *rag* [ræg̊])
[ʰ]   aspirated, e.g., *pie* [pʰaɪ]
[̩]   syllabic consonant, e.g., *Britain* [ˈbrɪtn̩] (normally below the symbol, but above for descending symbols, e.g., *bacon* [ˈbeɪkn̍])
[ˈ]   primary stress, e.g., *intend* /ɪnˈtɛnd/
[ˌ]   secondary stress, e.g., *kangaroo* /ˌkæŋɡəˈru/
/ /   phonemic transcription
[ ]   phonetic transcription
< >   orthographic form
*   incorrect or unattested form

THE INTERNATIONAL PHONETIC ALPHABET (revised to 2015)

CONSONANTS (PULMONIC)   © 2015 IPA

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	N		
Trill	B			r					R		
Tap or Flap		ⱱ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	L			

Symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations judged impossible.

CONSONANTS (NON-PULMONIC)

Clicks	Voiced implosives	Ejectives
ʘ Bilabial	ɓ Bilabial	ʼ Examples:
ǀ Dental	ɗ Dental/alveolar	pʼ Bilabial
ǃ (Post)alveolar	ʄ Palatal	tʼ Dental/alveolar
ǂ Palatoalveolar	ɠ Velar	kʼ Velar
ǁ Alveolar lateral	ʛ Uvular	sʼ Alveolar fricative

OTHER SYMBOLS

ʍ Voiceless labial-velar fricative

w Voiced labial-velar approximant

ɥ Voiced labial-palatal approximant

ʜ Voiceless epiglottal fricative

ʢ Voiced epiglottal fricative

ʡ Epiglottal plosive

ɕ ʑ Alveolo-palatal fricatives

ɺ Voiced alveolar lateral flap

ɧ Simultaneous ʃ and x

Affricates and double articulations can be represented by two symbols joined by a tie bar if necessary.

t͡s k͡p

VOWELS

Where symbols appear in pairs, the one to the right represents a rounded vowel.

SUPRASEGMENTALS

ˈ Primary stress	ˌfoʊnəˈtɪʃən
ˌ Secondary stress	
ː Long	eː
ˑ Half-long	eˑ
˘ Extra-short	ĕ
ǀ Minor (foot) group	
‖ Major (intonation) group	
. Syllable break	ɹi.ækt
‿ Linking (absence of a break)	

TONES AND WORD ACCENTS

LEVEL			CONTOUR		
e̋ or ˥	Extra high		ě or ˩˥	Rising	
é ˦	High		ê ˥˩	Falling	
ē ˧	Mid		e᷄ ˦˥	High rising	
è ˨	Low		e᷅ ˩˨	Low rising	
ȅ ˩	Extra low		e᷈	Rising-falling	
↓ Downstep			↗ Global rise		
↑ Upstep			↘ Global fall		

DIACRITICS Some diacritics may be placed above a symbol with a descender, e.g. ŋ̊

̥ Voiceless	n̥ d̥	̤ Breathy voiced	b̤ a̤	̪ Dental	t̪ d̪
̬ Voiced	s̬ t̬	̰ Creaky voiced	b̰ a̰	̺ Apical	t̺ d̺
ʰ Aspirated	tʰ dʰ	̼ Linguolabial	t̼ d̼	̻ Laminal	t̻ d̻
̹ More rounded	ɔ̹	ʷ Labialized	tʷ dʷ	̃ Nasalized	ẽ
̜ Less rounded	ɔ̜	ʲ Palatalized	tʲ dʲ	ⁿ Nasal release	dⁿ
̟ Advanced	u̟	ˠ Velarized	tˠ dˠ	ˡ Lateral release	dˡ
̠ Retracted	e̠	ˤ Pharyngealized	tˤ dˤ	̚ No audible release	d̚
̈ Centralized	ë	̴ Velarized or pharyngealized	ɫ		
̽ Mid-centralized	e̽	̝ Raised	e̝ (ɹ̝ = voiced alveolar fricative)		
̩ Syllabic	n̩	̞ Lowered	e̞ (β̞ = voiced bilabial approximant)		
̯ Non-syllabic	e̯	̘ Advanced Tongue Root	e̘		
˞ Rhoticity	ɚ a˞	̙ Retracted Tongue Root	e̙		

Preface and acknowledgments

American English Phonetics and Pronunciation Practice is unique in combining an introduction to English phonetics with extensive material for practicing English pronunciation. It presents a twenty-first-century model of educated American English, "General American" (GA), and is supported by a website containing complete recordings of all exercises. A parallel version for British English is also available (Carley, Mees, Collins 2018).

Modern corpus-based descriptions of colloquial English have been used to create materials that teach pronunciation while simultaneously practicing useful idiomatic language. Each English phoneme is demonstrated in high-frequency words in different phonetic contexts, in common phrases, in sentences, and in dialogues. Moreover, each phoneme is also extensively practiced in contrast with similar, confusable sounds in minimal pairs, phrases, and sentences.

Beyond the segmental level of vowels and consonants, considerable attention is given to the difference between strong and weak syllables, a very important component of English rhythm, and two complete chapters are dedicated to the difficult area of consonant clusters.

Learners who will benefit from the book include

- Learners of English as a foreign language
- Students of English language and linguistics
- Trainee English teachers
- Professionals wishing to speak English with clarity and accuracy

<div align="right">

Paul Carley and Inger M. Mees
Jeddah and Copenhagen, May 2019

</div>

Chapter 1

Basic concepts

1.1 Pronunciation priorities

When learning to pronounce a new language, it's essential to get your priorities right. The most important sounds are the ones that can change the meaning of words. These are called **phonemes** (see Section 1.2). If you say *pin* and it sounds like *bin*, people will misunderstand you. And if you say *I hid them* and it sounds like *I hit them*, there will also be a breakdown in communication. Furthermore, you should be aware that sounds may be pronounced differently in different contexts, e.g., **pre-vocalically** (before vowels), **intervocalically** (between vowels), or **pre-consonantally** (before consonants). They may also be pronounced differently in different positions in the word – at the beginning (**initial**), in the middle (**medial**), and at the end (**final**). For instance, /p/ is more like a /b/ when it occurs after /s/, e.g., *port* vs. *sport*; /r/ sounds different in *red* and *tread*, the two /t/ sounds in *tight* are different, and the quality of <**oa**> is different in *goat* and *goal*. Note that when we refer to the letters in a word – as opposed to the sounds – we show them in **angle brackets**, e.g., <f> or <ie>. Phonemes are shown in **slant brackets**, e.g., /r/ or /ɛ/. The word *spread* would phonemically be shown as /sprɛd/.

Even if people can understand what you are saying, an off-target pronunciation may still sound *comical*, *irritating*, or *distracting* to listeners. For instance, if you say English /r/ with a back articulation (in your throat) instead of a front articulation (with your tongue-tip), it may sound funny to people who aren't used to it. If listeners are distracted because of a false pronunciation, they may stop concentrating on what you are trying to say. Or if they need to invest a lot of effort in deciphering what you are saying, they may lose track of your message. Furthermore, judgments of your overall ability in English are likely to be based on the impression your pronunciation makes: if you sound like a beginner, you may be treated like a beginner, even if your level is advanced in terms of grammar, vocabulary, reading, and writing.

The best approach is to aim for a pronunciation that (1) can be understood without any difficulty and (2) doesn't irritate or distract your listeners. Note that there's more to learning the pronunciation of a language than mastering the **segments** (vowels and consonants). You have to pay attention to several other points. For instance, correct use of **weak forms** (Section 8.3) helps to get the speech **rhythm** right. **Contractions**, e.g., *don't*, *it's*, *we'll*, improve the fluency. To make your pronunciation more authentic, it's important to have knowledge of **assimilation** (sounds that change under the influence of neighboring sounds, e.g., *when* becomes /wɛm/ in *when my*) and **elision** (disappearing sounds, e.g., /t/ is often lost in *facts*).

1.2 Phonemes and allophones

Some sound differences matter a great deal, whereas others are of little significance. The ones that matter most are those that can change the meaning of otherwise identical words. In English, the words *bit*, *bet*, *boat* are distinguished only by the vowels; in *bit*, *sit*, *wit*, only the initial consonant is different. In *bit*, *bill*, *bin*, it's the final consonant that brings about the change in meaning. Sounds that can distinguish meaning are called **phonemes** (adjective: **phonemic**). A *pair* of words distinguished by a single phoneme is called a **minimal pair**, e.g., *bit – hit*. The variety of English taught in this book (see Section 1.7) has 24 consonant phonemes and 13 vowel phonemes.

Not every sound difference can change the meaning of a word. Listen carefully to *feet* and *feed*. You can hear a distinct difference in the length of the two vowels. But the native English speaker interprets these vowels as two variants of the same phoneme /i/; the different vowel lengths are the result of the influence of the following consonants /t/ and /d/. Similarly, the two /k/ sounds in *keen* and *corn* are different, the first being formed more forward and the second further back in the mouth, but English speakers hear both as variants of the phoneme /k/.

When you say the /d/ in *deal*, your lips are unrounded during the consonant, but when you say /d/ in *door*, they are rounded. In *deal*, the vowel is unrounded, and in *door*, the vowel is rounded. When we say *deal* and *door*, our lips are getting ready for the vowel during the articulation of the consonant. So the lip-shape of the consonant is affected by the lip-shape of the following vowel. Each phoneme is composed of a number of such different variants. These are termed **allophones** (adjective: **allophonic**). Allophones may occur in **complementary distribution** or in **free variation**. Our *deal/door* example is an instance of allophones in complementary distribution. This means that the different allophones complement each other; where one occurs, the other cannot occur. In other words, we can write a rule for the occurrence of the two allophones: /d/ with rounded lips occurs before lip-rounded sounds while /d/ with unrounded lips occurs before all other sounds. Vowels are shortened before voiceless consonants like /s/ while they retain full length before voiced consonants like /z/; for example, the vowel in *face* /feɪs/ is clearly shorter than that in *phase* /feɪz/. Again, the allophones are in complementary distribution. If allophones are in free variation, their occurrence cannot be predicted from the phonetic context. An example of this would be the different possible pronunciations of /t/ in word-final position, as in *hat*. It's possible to pronounce the /t/ with or without glottal reinforcement (see Section 2.7.3). Many speakers vary between these two possibilities, and we cannot predict which of the two they are going to use. The glottally reinforced and non-glottally reinforced variants are therefore said to be in free variation.

Unfortunately for the learner, languages generally don't have the same phoneme system, and they certainly don't have the same range of allophones. So the learner has to work out the phonemic inventory of the new language and all the phonetic variants. Your first task is to make sure you never lose a phoneme contrast. This isn't easy to do in practice. Even though two phonemes may sound very similar, or identical, to the learner, to the native speaker, they are completely different. This is something native speakers and learners are often not aware of. Native speakers are frequently surprised to hear that the vowels in the English words *seat* /siːt/ and *sit* /sɪt/ sound identical to speakers of most other languages, who hear them as the same vowel because they count as allophones of the same phoneme in their languages. Many learners find it difficult to separate the phonemes in *Luke* /luːk/ and *look* /lʊk/. Others find it difficult to distinguish between *cat* /æ/, *cut* /ə/, and *cot* /ɑ/. Yet others can't hear and/or make

the difference between the initial consonants in *three* /θ/ and *tree* /t/, or *three* /θ/ and *free* /f/, or *theme* /θ/ and *seem* /s/. In this book, we have provided exercises for 29 consonant contrasts and 17 vowel contrasts. You'll find that some of these don't pose a problem for speakers of your language while others will take a long time to master. If making a particular contrast isn't difficult for you, you can still use the contrast section as extra material to help you get the two sounds just right. Note that a full command of the contrasts involves being able to say all the different allophones of a phoneme in their appropriate contexts.

Remember that allophones can never change the meaning of words. English /t/ can be said in many different ways (i.e., there are many different allophones or variants), but if we substitute one allophone for another, the meaning remains the same. It will merely sound a bit odd. However, if we replace /t/ in *tight* by /s/, /f/, or /k/, then it turns into *sight*, *fight*, *kite*, and the result is a new word with a different meaning; /t s f k/ are therefore examples of phonemes in English. The English phoneme system is shown in the "English Phonemic Transcription Key" at the start of this book.

1.3 Spelling and sound

English **orthography** (i.e., spelling) is notoriously unreliable. For instance, the vowel /i/ can be spelled in numerous ways. All the letters underlined in the following words represent /i/: *me, see, sea, believe, receive, pizza, people, key, quay, quiche, Portuguese*. Most other phonemes can also be spelled in many different ways, especially vowels. So instead of relying on the orthography, phoneticians use **transcription**. There are two types: (1) **phonemic transcription**, indicating phonemes only; this type, as we have seen, is normally placed inside slant brackets / /, e.g., *part* /pɑrt/. The sign – is used to show **phoneme contrasts**, e.g., *let* /lɛt/ – *met* /mɛt/; (2) **phonetic transcription**, showing more detailed allophonic distinctions, enclosed by **square brackets** [ ], e.g., *part* [pʰɑrt]. To indicate the allophonic distinctions, we often make use of **diacritics**, i.e., marks added to symbols to provide extra information, e.g., [pʰ]. The rounded allophone of /t/ is shown as [tʷ]; as /t/ said with unrounded lips is the default, there's no special symbol to denote it.

Sometimes words with different meanings are spelled completely differently but are pronounced in the same way, as in *key* and *quay* above. Such words are called **homophones** (same pronunciation, different meaning). English has a great many of these. Other examples of homophones are *wait/weight*, *know/no*, *sea/see*, *cite/sight/site*. To confuse matters even more, the opposite also occurs. It's possible for words that are spelled identically to be pronounced differently. The written word *row* can be said with the vowel in GOAT (when it means a "line") or the vowel in MOUTH (when it means a "quarrel"), and it's therefore impossible to tell from the spelling alone which meaning and pronunciation are intended. Words of this type are called **homographs** (same spelling, different pronunciation).

1.4 Phoneme symbols

Unfortunately, at present, there is no consensus among writers on the set of symbols used for transcribing GA. Even those, like us, who use the symbols of the International Phonetic Association's (IPA) International Phonetic Alphabet (see p. xv) don't necessarily use the same symbols in their transcriptions. The main reason for this is that while the IPA provides symbols to represent the range of speech sounds found in language, it doesn't dictate how the sound system of a language should be analyzed. A further reason is that writers have different

approaches depending on whether they are writing for foreign learners, speech and language therapists, professional linguists, actors, dictionary users, and so on. In each case, there may be different traditions of transcription, differences in the linguistic knowledge of readers, different levels of tolerance for unfamiliar symbols, and different assumptions about what needs to be made explicit and what can be taken for granted in transcriptions.

In this work, we use a transcription system which is mostly phonemic but includes a small number of non-phonemic elements. We take a phonemic approach to the schwa /ə/ phoneme, using the same symbol for it in stressed and unstressed syllables (e.g., *above* /əˈbəv/; see Section 5.3) and the same symbol followed by /r/ when it is r-colored in stressed and unstressed syllables (e.g., *murmur* /ˈmərmər/; see Section 5.3). We take a non-phonemic approach to the SPORT [o] vowel (e.g., *four* /for/, *sort* /sort/, *story* /ˈstori/; see Section 5.2.2), t-tapping (e.g., *city* /ˈsɪt̬i/; see Section 2.8), and syllabic consonants (e.g., *kitten* /ˈkɪtn̩/, *rattle* /ˈrætl̩/; see Section 1.5 and Section 8.2). In these three cases we continue to use phonemic slanted brackets in this book in order to avoid the inconvenience of constantly switching between phonemic and phonetic bracketing.

1.5 The syllable

A syllable is a group of sounds that are pronounced together. Words can consist of a single syllable, i.e., a **monosyllable** (*tight, time*) or of two or more syllables (**polysyllabic**), e.g., *waiting* (two syllables – **disyllabic**), *tomato* (three syllables), *participate* (four syllables), *university* (five syllables), and so on. A syllable nearly always contains a vowel (e.g., *eye* /aɪ/); this is called the **syllable nucleus**. The nucleus may be preceded or followed by one or more consonants (e.g., <u>t</u>ea, <u>tr</u>ee, <u>str</u>eam, a<u>t</u>, <u>c</u>a<u>t</u>, <u>c</u>a<u>ts</u>, <u>st</u>am<u>ps</u>). The consonant or consonants preceding the nucleus are known as the **syllable onset**, and the consonants following the nucleus are called the **coda**. A group of consonants in a syllable onset or coda is known as a **cluster**. The English syllable can consist of clusters of up to three consonants in the onset (e.g., *strengths* /<u>str</u>ɛŋθs/), and as many as four in the coda (e.g., *texts* /tɛ<u>ksts</u>/). Note that we are here concerned with *pronunciation*, so even though the word *time* looks as if it consists of two syllables because it has two vowel letters in the orthography, the word consists of only one syllable, as the second vowel letter in the spelling doesn't represent a vowel sound. A syllable that has a coda (i.e., one or more closing consonants) is called a **closed syllable**, while a syllable that ends with a vowel phoneme is called an **open syllable**.

Occasionally, a syllable consists of a consonant only, most frequently /n/ or /l/, e.g., *Britain* /ˈbrɪtn̩/, *hidden* /ˈhɪdn̩/, *mission* /ˈmɪʃn̩/, *middle* /ˈmɪdl̩/, *apple* /ˈæpl̩/. A consonant that forms a syllable without the aid of a vowel is called a **syllabic consonant**. Note that we show a syllabic consonant by means of a small vertical mark beneath the symbol (with descending symbols, a superscript mark is used, e.g., *bacon* /ˈbeɪkŋ̍/). A word like *apple* /ˈæpl̩/ consists of two syllables, but only the first contains a vowel; the second contains a syllabic consonant; see Sections 2.19, 2.23, 2.26, and 8.2.

1.6 Stress

Words consist of more than a set of segments (vowels and consonants) arranged in a certain order. Words of more than one syllable also have a distinctive rhythmic pattern depending on which syllables are pronounced with **stress** and which are not. Stressed syllables are pronounced with greater energy and effort than unstressed syllables, which results in greater

prominence, i.e., they stand out more. The first syllable in *carpet* is **stressed** and the second **unstressed**; the second syllable in *contain* is stressed and the first unstressed. Stress is indicated by means of a vertical mark placed before the stressed syllable, and unstressed syllables are left unmarked, e.g., /ˈkɑrpət/, /kənˈteɪn/. The position of stress in an English word is an important factor in word recognition, and there are even words that are distinguished by stress alone, e.g., the noun *increase* /ˈɪŋkris/ and the verb *increase* /ɪŋˈkris/.

Some words have more than one stressed syllable. In *Alabama*, the first and third syllables are stressed, while in *impossibility*, the second and fourth syllables are stressed. In these examples, as in all cases of multiple stresses, the last stress sounds more prominent than the earlier stress, and this is why the term **primary stress** has been used for the last, more prominent stress and **secondary stress** for any earlier, less prominent stresses. Primary stress is indicated with the usual stress mark and secondary stress with the same symbol at a lower level, e.g., /ˌæləˈbæmə/, /ɪmˌpɑsəˈbɪləʤi/. Although the terminology and transcription seem to suggest that there are three different levels of stress – primary stress, secondary stress, and unstressed – this isn't actually the case. There are only stressed and unstressed syllables, and the difference in prominence between the stresses in words such as *Alabama* and *impossibility* is due to **pitch accent**.

An accented /ˈæksɛntəd/ syllable is one that is accompanied by a change in the **pitch** of the voice. Pitch is related to the speed at which the vocal folds vibrate: faster vibration results in higher pitch and slower vibration lower pitch. When a word is pronounced in isolation, the syllable that takes primary stress is accented, i.e., accompanied by a pitch movement, usually a fall in pitch. When there's a "secondary stress" earlier in the word, this is accompanied by a step up to a relatively high pitch before the pitch movement of the "primary stress." In terms of the English sound system, the pitch movement associated with the "primary stress" is more salient than the step up in pitch associated with the "secondary stress." Thus, the distinction between primary and secondary stress is really a difference between different kinds of pitch accent rather than stress.

In this book, when individual monosyllabic words are transcribed as examples, we don't use a stress mark, which agrees with the approach taken in most dictionaries and works on English phonetics. Every word must have at least one stressed syllable when pronounced in isolation, and therefore, it's self-evident that the one and only syllable of a word is stressed. When we transcribe an individual polysyllabic word, we only indicate primary stress. When we transcribe utterances longer than a single word, we use the stress mark whenever a syllable is stressed, meaning that monosyllabic words can receive a stress mark but also that some stresses that appear when a word is said in isolation may disappear when the word is spoken in a phrase.

1.7 Pronunciation model

Every language has a number of different **accents**. An accent is a pronunciation variety characteristic of a group of people. Accents can be **regional** or **social**. In the USA, we find many different regional accents; examples are Texas, Kentucky, New York, and Boston, spoken by most of the people who live in these areas. But unless you have reasons for specifically wishing to adopt one of these regional accents, it's best for learners not to use these as a **model** for imitation. The accent of American English we recommend is one heard from educated speakers throughout the USA (and also in Canada). We shall term this social accent **General American** (abbreviated to **GA**). If you listen regularly to the American

media, you're probably already familiar with this accent, since it's the variety used by the majority of American presenters. It's sometimes even called "Network English." It's either completely non-localizable (i.e., it's impossible to tell where speakers come from) or has very few regional traces. Thus, GA can be taken as the common denominator of the speech of educated Americans. When people alter their pronunciation (consciously or unconsciously) to sound less regional, they change in the direction of GA. When there's an accent spectrum within a location, those at the lower end of the social scale speak with the local accent while those toward the other end of the social scale speak with an accent progressively more like GA.

The English we describe in this book is the speech of the average modern General American speaker. Old-fashioned usages have been excluded, as have any "trendy" pronunciations that are too recent to have gained widespread acceptance.

Chapter 2

Consonants

2.1 The vocal tract and tongue

Before we discuss how the 24 English consonant phonemes are made, or **articulated**, let's familiarize ourselves with the anatomy of the **vocal tract** (Figure 2.1) and **tongue** (Figure 2.2).
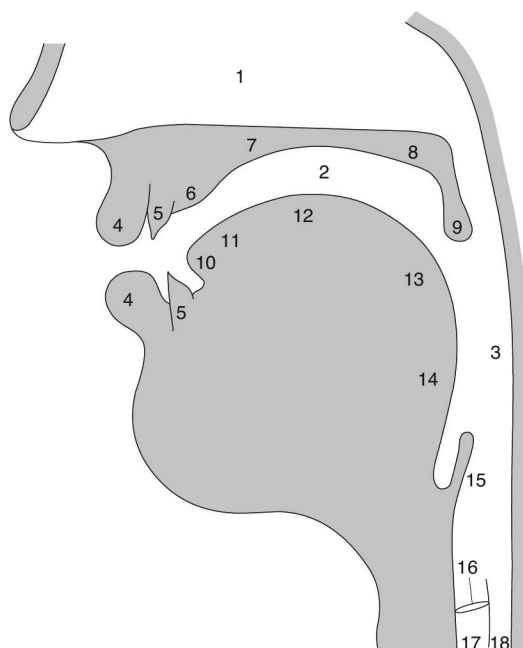
1 Nasal cavity	10 Tip of tongue
2 Oral cavity	11 Blade of tongue
3 Pharynx	12 Front of tongue
4 Lips	13 Back of tongue
5 Teeth	14 Root of tongue
6 Alveolar ridge	15 Epiglottis
7 Hard palate	16 Larynx, containing vocal folds
8 Soft palate (also termed "velum")	17 Trachea
9 Uvula	18 Esophagus

Figure 2.1 The anatomy of the vocal tract

This diagram might appear strange to you at first. The tongue may be larger or smaller than you expected. It's small in the sense that the tip does not extend much further from the place at which it's attached to the lower jaw, and it's large in the sense that it extends deep into the mouth and throat and almost completely fills the **oral cavity** (mouth).

If we start at the bottom of the diagram, we see that there are two passages. The **esophagus** /ɪˈsɑfəgəs/ (food passage) leads to the stomach, and the **trachea** /ˈtreɪkiə/ (windpipe) leads to the lungs. It's the trachea that is of most interest for our purposes. During speech, air flows up from the lungs via the trachea, and the first point of interest that it meets is the **larynx** /ˈlɛrɪŋks/. The larynx joins the trachea to the pharynx (throat), and is a box-like structure made of cartilage. It's larger in men than in women and is what makes the "Adam's apple," the lump at the front of the throat. The larynx contains the **vocal folds**, a pair of lip-like structures that can be brought together to close off the trachea and lungs. If there's a need to expel something from the lungs or trachea, the vocal folds are brought tightly together, the muscles of the chest and abdomen squeeze the lungs strongly, and then the vocal folds are abruptly separated to let the trapped air below escape in an explosion that hopefully clears the blockage. This is a cough. The vocal folds also seal off the lungs to stabilize the chest during lifting or other types of physical exertion. You will notice that before you pick up something heavy, you take a breath and trap it in your lungs by bringing the vocal folds together, and then when you put the load down, you inevitably let out a gasp as you release the air you had trapped in your lungs.

The next feature is the **epiglottis** /ɛpəˈglɑtəs/, a flap of cartilage at the root of the tongue. It isn't involved in making speech sounds in English. Its biological function is to fold over the entrance to the larynx during swallowing in order to guide food and drink into the esophagus.

The space above the larynx and behind the root of the tongue is called the **pharynx** /ˈfɛrɪŋks/. It's smaller when the tongue is pulled back in the mouth and larger when the tongue is pushed forward.

There is then a possible fork in the road for the **airstream**. In our diagram, the **soft palate** /ˈpælət/ (also termed **velum** /ˈviləm/) and the **uvula** /ˈjuvjələ/ at its tip are shown in the lowered position, but it's also possible for the soft palate to form a seal against the back wall of the pharynx and close off the entrance to the **nasal cavity** (nose). This is known as a **velic closure**; see, e.g., Figure 2.3. Thus, the airstream can potentially enter both the oral and nasal cavities (as in Figure 2.1) or only the oral cavity (when the soft palate is raised and a velic closure is formed). There's little to be said about the nasal cavity itself because its dimensions are fixed; it's only the valve-like action of the soft palate opening and closing the entrance to it that is relevant for speech.

The oral cavity is bordered by the tongue at the bottom, the **palate** at the top, and the lips, cheeks, and teeth at the front and sides. By opening and closing the jaw and pulling the tongue back and pushing it forward, the oral cavity can be made larger or smaller. The tongue, lower teeth, and lip move with the lower jaw while the upper teeth and lip are in a fixed position.

Behind the upper front teeth is a lumpy area called the **alveolar ridge** /ælˈviələr/, and to the rear of that is the palate. The palate is divided into the soft palate and the **hard palate**. If you explore your palate with the tip of your tongue, you'll find that it's indeed hard and bony at the front and soft and fleshy at the back. At the very end of the soft palate is the uvula, which you can see hanging down when you look in the mirror.

The position shown in Figure 2.1 is actually a slightly unusual one, but it's useful for demonstration purposes. It shows the position assumed when breathing through the nose and mouth simultaneously. When not speaking, a healthy person would hold the jaws closer
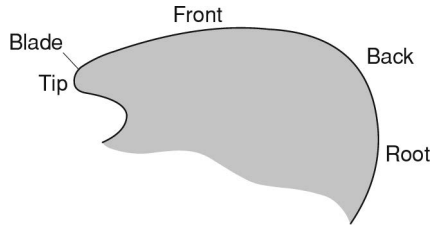
Figure 2.2 Divisions of the tongue

together with the lips and teeth touching; the tongue would fill the oral cavity, touching the roof of the mouth from the alveolar ridge to the soft palate; and the soft palate would be lowered (as in the diagram) to allow for normal breathing in and out via the nose.

The tongue has few obvious natural divisions in the way that the vocal tract does. However, phoneticians find it convenient to divide it into a number of parts when describing sounds and their articulations (see Figure 2.2).

The very point of the tongue is known as the **tip**. The part of the tongue that narrows to the tip and that lies under the alveolar ridge is the **blade** of the tongue. The part that lies under the hard palate and the part under the soft palate are called the **front** and **back** of the tongue respectively. This may seem strange at first, but front and back refer to the part of the tongue used in the articulation of vowels (see Section 5.1.1) – the front is the part of the tongue used to form front vowels and the back is the part used to form back vowels. The tip and blade remain low in the mouth and are not involved in making vowels. Finally, the part of the tongue in the pharynx is the **root**.

2.2 Describing consonants

A consonant is a speech sound that involves an obstruction of the airstream as it passes through the vocal tract. Describing a consonant involves describing the nature of the obstruction, and there are three factors to be taken into consideration: **voicing**, **place of articulation**, and **manner of articulation**.

2.2.1 Voicing

Voicing refers to the actions of the vocal folds during the articulation of a consonant. Different actions of the vocal folds produce **voiced** and **voiceless** sounds.

- For voiced sounds (i.e., vowels and the voiced consonants /b d ɡ dʒ v ð z ʒ m n ŋ l r j w/), the vocal folds are held gently together so that the airflow from the lungs causes them to vibrate.
- For voiceless sounds (i.e., the voiceless consonants /p t k tʃ f θ s ʃ h/), the vocal folds are held apart as in the position for normal breathing.

It's easiest to appreciate the voicing in sounds like /v ð z ʒ m n ŋ l r/ because these can be prolonged. Put your hand on your throat as you say them and feel the vibration. Note how the vibration stops and starts as you stop and start the consonant.

A number of English consonants come in pairs, the only difference between them being that one is voiceless and one is voiced. These pairs are: /f v/, /θ ð/, /s z/, /ʃ ʒ/, /p b/, /t d/, /k g/, and /ʧ ʤ/. Take some of the pairs that can be easily lengthened, such as /f v/ and /s z/, and alternate between the voiceless and the voiced consonants, feeling how the vibration in your larynx stops and starts.

Note that in our diagrams of consonant articulations, we use a plus sign at the larynx to indicate that the consonant is voiced and a minus sign if it's voiceless. If both signs are included (±), both voiced and voiceless articulations are possible.

2.2.2 Place of articulation

The second factor to take into account when describing a consonant is *where* in the vocal tract the obstruction is made. **Place of articulation** is described in terms of an **active articulator** that moves toward a **passive articulator**, which is in a fixed position (Table 2.1).

Table 2.1 English consonants: place of articulation

Place	Active articulator	Passive articulator	Consonants
Bilabial	Lower lip	Upper lip	/p b m/
Labio-dental	Lower lip	Upper incisors	/f v/
Dental	Tongue-tip	Upper incisors	/θ ð/
Alveolar	Tongue-tip	Alveolar ridge	/t d n s z l/
Post-alveolar	Tongue-tip	Rear of alveolar ridge	/r/
Palato-alveolar	Tongue-tip, blade, and front	Alveolar ridge and hard palate	/ʧ ʤ ʃ ʒ/
Palatal	Front of tongue	Hard palate	/j/
Velar	Back of tongue	Soft palate	/k g ŋ/
Glottal	Vocal folds	Vocal folds	/h/
Labial-velar	Back of tongue Lips	Soft palate Lips	/w/

The lip-rounding of labial-velar /w/ and positioning of the vocal folds for glottal /h/ cannot be analyzed in terms of passive and active articulators because they involve two elements moving toward each other (the corners of the mouth for /w/ and the two vocal folds for /h/). The lips and vocal folds are therefore classified as both active and passive in our table.

We refer to /r/ as post-alveolar, but see Section 2.21 for a discussion of an alternative place of articulation for this consonant.

Take some time to silently articulate the consonants listed in the table (or at least those you are confident you can correctly pronounce) in order to identify the different places of articulation.

2.2.3 Manner of articulation

Manner of articulation is the term used to describe the kind of obstruction involved in articulating a consonant. The five manners of articulation found in English are

- **Plosive**: A complete closure is formed in the vocal tract, blocking the airstream, and then released. The GA plosives are /p b t d k g/.
- **Fricative**: A narrowing is formed in the vocal tract, causing turbulence and fricative noise as the airstream is forced through. The GA fricatives are /f v θ ð s z ʃ ʒ h/.

- **Affricate**: A complete closure is formed in the vocal tract, blocking the airstream, and then released slowly, resulting in **homorganic friction** (i.e., fricative noise at the same place of articulation). The GA affricates are /tʃ dʒ/.
- **Nasal**: A complete closure is formed in the oral cavity, the soft palate is in the lowered position, and air exits via the nose. The GA nasals are /m n ŋ/.
- **Approximant**: A narrowing is formed in the vocal tract, but one not narrow enough to cause turbulence and noise as in the case of a fricative. The GA approximants are /r l j w/.

Explore what these terms really mean by articulating some of the consonants you are confident of and feeling the different manners of articulation.

2.2.4 Double and secondary articulations

English /w/ is an example of a **double articulation**, meaning that two articulations of equal magnitude (i.e., two **primary articulations**) take place at the same time – an approximant articulation between the back of the tongue and the soft palate and another approximant articulation consisting of the rounding of the lips. Other double articulations are possible but do not occur in English.

In the case of **secondary articulations**, the primary articulation is accompanied by an articulation of lesser magnitude. Examples of this in GA are the **labialization** (i.e., lip-rounding) accompanying /ʃ ʒ tʃ dʒ r/ (see Sections 2.5, 2.13, and 2.20) and the approximation of the back of the tongue to the soft palate (**velarization**) that accompanies /l/ (see Section 2.25).

2.2.5 Combining voicing, place, and manner

When we bring together the key factors of voicing, place of articulation, and manner of articulation, each English consonant phoneme has its own unique label that phoneticians use when referring to them.

/p/	voiceless bilabial plosive	/s/	voiceless alveolar fricative
/b/	voiced bilabial plosive	/z/	voiced alveolar fricative
/t/	voiceless alveolar plosive	/ʃ/	voiceless palato-alveolar fricative
/d/	voiced alveolar plosive	/ʒ/	voiced palato-alveolar fricative
/k/	voiceless velar plosive	/h/	voiceless glottal fricative
/g/	voiced velar plosive	/m/	voiced bilabial nasal
/tʃ/	voiceless palato-alveolar affricate	/n/	voiced alveolar nasal
/dʒ/	voiced palato-alveolar affricate	/ŋ/	voiced velar nasal
/f/	voiceless labio-dental fricative	/j/	voiced palatal approximant
/v/	voiced labio-dental fricative	/w/	voiced labial-velar approximant
/θ/	voiceless dental fricative	/l/	voiced alveolar lateral approximant
/ð/	voiced dental fricative	/r/	voiced post-alveolar approximant

Strictly speaking, the IPA symbol for a voiced post-alveolar approximant is [ɹ], but when making practical phonemic transcriptions for various languages, phoneticians often replace "exotic" symbols with the nearest "non-exotic" symbol for the sake of ease of printing and writing. In such cases, there's no danger of confusion because the description of the sounds is included with the symbols, as we do here.

Note that in the IPA chart (see p. xv) the term *post-alveolar* is used for /tʃ dʒ ʃ ʒ/, but we prefer the term traditionally used for English, *palato-alveolar*, because it better describes the English articulations.

2.3 The English consonants

Table 2.2 conveniently summarizes the voicing, place, and manner of articulation of the 24 English consonants. Within each cell, the sound on the left is voiceless and the one on the right is voiced.

Table 2.2 English consonant grid

	Bilabial	Labio-dental	Dental	Alveolar	Post-alveolar	Palato-alveolar	Palatal	Velar	Glottal	Labial-Velar
Plosives	p b			t d				k g		
Affricates						tʃ dʒ				
Fricatives		f v	θ ð	s z		ʃ ʒ			h	
Nasals	m			n				ŋ		
Approximants				l	r		j			w

2.4 Obstruents and sonorants

Table 2.2 reveals that consonants fall into two groups: those that typically come in voiceless and voiced pairs (obstruents) and those that do not (sonorants). Plosives, affricates, and fricatives are **obstruents** /ˈɑbstruənts/, the most consonant-like of the consonants, involving the greatest degree of obstruction to the airstream. Nasals and approximants are **sonorants** /ˈsɑnərənts/ (as are vowels) and are the least consonant-like of the consonants, involving a lesser degree of obstruction.

2.4.1 Pre-fortis clipping

Voiceless (also known as *fortis*) obstruents shorten sonorants that precede them in the same syllable. This phenomenon is termed **pre-fortis clipping** and is most often discussed in relation to vowels (see Section 5.6), but it also affects sonorant consonants, for example:

/m/	*lump* /ləmp/, *lymph* /lɪmf/
/n/	*bent* /bɛnt/, *wince* /wɪns/, *bench* /bɛntʃ/
/ŋ/	*bank* /bæŋk/
/l/	*help* /hɛlp/, *belt* /bɛlt/, *sulk* /səlk/, *belch* /bɛltʃ/, *golf* /ɡɑlf/, *health* /hɛlθ/, *else* /ɛls/, *Welsh* /wɛlʃ/
/r/	*sharp* /ʃɑrp/, *short* /ʃort/, *fork* /fork/, *march* /mɑrtʃ/, *scarf* /skɑrf/, *north* /norθ/, *scarce* /skɛrs/, *marsh* /mɑrʃ/

2.4.2 Obstruent devoicing

Although we have so far referred to English consonants as either voiceless or voiced, English voiced obstruents are actually only *potentially* fully voiced. Depending on the phonetic context, they are often partially or even completely **devoiced** /diˈvɔɪst/ (i.e., they partially or completely lose their voicing). For this reason, the terms **fortis** and **lenis** are sometimes used for voiceless and voiced instead. Fortis means *strong*, while lenis means *weak*, which reflects the fact that voiceless obstruents are articulated more forcefully than voiced obstruents. Force of articulation is a difficult feature to perceive, and many learners are confused

by these terms because they feel that voiced sounds are louder and therefore strong, while voiceless sounds are quieter and therefore weak. It's preferable, therefore, to stick with the terms *voiceless* and *voiced*, with the understanding that when we refer to English /b d g dʒ v ð z ʒ/ as voiced obstruents, we actually mean "potentially" fully voiced.

English voiced obstruents are typically fully voiced when they occur between voiced sonorants, i.e., nasals, approximants, and vowels:

/b/	*rabbit* [ˈræbət]	*a boat* [ə ˈboʊt]	*rub it* [ˈrəb ɪt]
/d/	*colder* [ˈkoʊldər]	*a dog* [ə ˈdɑg]	*need it* [ˈnid ɪt]
/g/	*cargo* [ˈkɑrgoʊ]	*my gate* [maɪ ˈgeɪt]	*big oak* [ˈbɪg ˈoʊk]
/dʒ/	*magic* [ˈmædʒɪk]	*a joke* [ə ˈdʒoʊk]	*page eight* [ˈpeɪdʒ ˈeɪt]
/v/	*envy* [ˈɛnvi]	*a view* [ə ˈvju]	*move it* [ˈmuv ɪt]
/ð/	*clothing* [ˈkloʊðɪŋ]	*see this* [ˈsi ˈðɪs]	*loathe it* [ˈloʊð ɪt]
/z/	*easy* [ˈizi]	*the zoo* [ðə ˈzu]	*his own* [hɪz ˈoʊn]
/ʒ/	*vision* [ˈvɪʒn̩]	*a genre* [ə ˈʒɑnrə]	*beige is* [ˈbeɪʒ ˈɪz]

English voiced obstruents are typically partially or fully devoiced when they are preceded by a pause (i.e., silence) or a voiceless consonant or when a pause or voiceless consonant follows:

/b/	*this book* [ˈðɪs ˈb̥ʊk]	*grab some* [ˈgræb̥ ˈsəm]
/d/	*dog* [ˈd̥ɑg]	*load* [loʊd̥]
/g/	*that guy* [ˈðæt ˈg̊aɪ]	*big* [ˈbɪg̊]
/dʒ/	*joke* [ˈd̥ʒoʊk]	*stage fright* [ˈsteɪdʒ̊ fraɪt]
/v/	*eight verbs* [ˈeɪt ˈv̥ərbz]	*save time* [ˈseɪv̥ ˈtaɪm]
/ð/	*not that* [ˈnɑt ˈð̥æt]	*bathe* [beɪð̥]
/z/	*six zones* [ˈsɪks ˈz̥oʊnz̥]	*buzz* [bəz̥]
/ʒ/	*genre* [ˈʒ̊ɑnrə]	*beige top* [ˈbeɪʒ̊ ˈtɑp]

A devoiced consonant is normally shown by means of a subscript circle under the consonant (e.g., [b̥]), but with descending symbols, a superscript circle is used (e.g., [g̊]). In the case of fricatives, the devoicing is greater before a voiceless consonant or pause than after a voiceless consonant or pause. When devoicing occurs, the difference between pairs of voiced and voiceless obstruents is less marked, but the contrast still remains – /b d g dʒ v ð z ʒ/ do not become /p t k tʃ f θ s ʃ/.

2.5 Stops

Plosives and affricates make up the category of **stops**. They have in common a combination of a velic closure and a closure in the oral cavity that results in a complete obstruction to the airstream (hence the term *stop*).

English has three pairs of voiceless and voiced plosives at the bilabial, alveolar, and velar places of articulation.

- For /p/ and /b/, the lips come together and form a complete closure, stopping the airstream (see Figure 2.3).
- For /t/ and /d/, a complete closure is formed by the tip of the tongue against the alveolar ridge and by the sides of the tongue against the upper side teeth (see Figure 2.4).
- For /k/ and /g/, the back of the tongue forms a closure against the soft palate, and the rear of the sides of the tongue form a seal against the rear upper side teeth (see Figure 2.5).
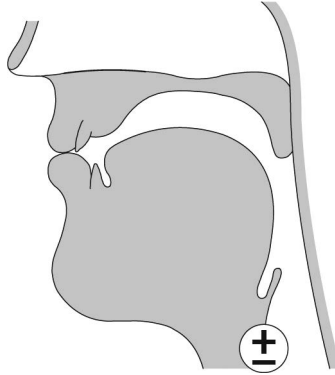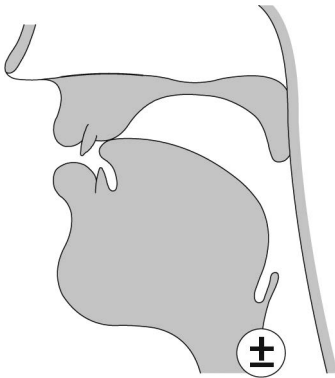
Figure 2.3 English plosives /p/ and /b/

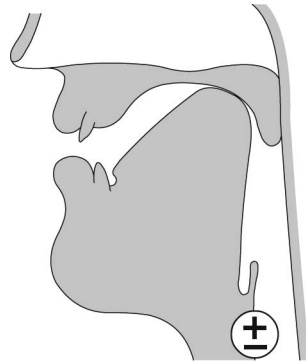*Figure 2.4* English plosives /t/ and /d/

Figure 2.5 English plosives /k/ and /g/

English has a single pair of voiceless and voiced affricates at the palato-alveolar place of articulation.

- For /ʧ/ and /ʤ/, the tip and blade of the tongue form a closure against the rear part of the alveolar ridge, the front of the tongue is raised towards the hard palate, and the sides of the tongue form a seal against the upper side teeth (see Figure 2.6).
- The closure is released slowly, resulting in a brief moment of homorganic friction (Figure 2.7).
- The primary articulation is accompanied by a simultaneous secondary articulation – rounding and protrusion of the lips.

Although the phonemic symbols for the affricates consist of two elements, these phonemic affricates are single sounds, and although the first element of each symbol is the same as that used for the alveolar plosives, the place of articulation is different (as our diagrams demonstrate), the same symbols being used only for the sake of convenience.
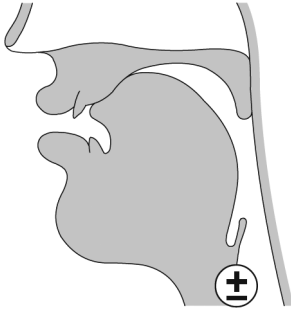
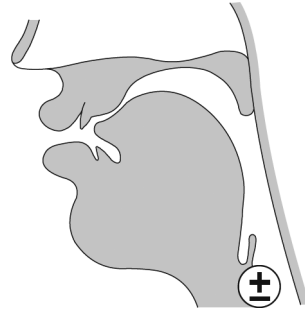*Figure 2.6* English affricates /ʧ/ and /ʤ/ showing palato-alveolar closure

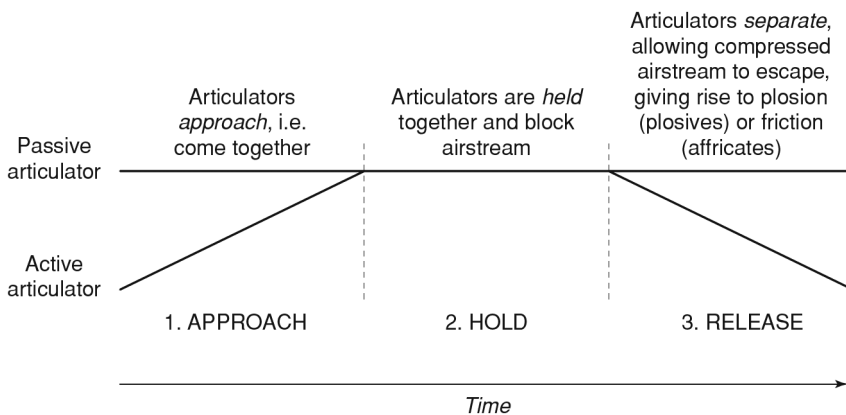*Figure 2.7* English affricates /ʧ/ and /ʤ/ showing release with homorganic friction

Figure 2.8 Articulation timing diagram showing the three stages of a stop

2.5.1 The stages of stops

Stops have three stages (see Figure 2.8):

1 **Approach**: The active articulator moves toward the passive articulator in order to form the closure.
2 **Hold**: The closure is made, the airstream is blocked, and pressure builds up.
3 **Release**: The active articulator moves away from the passive articulator, breaking the closure and releasing the compressed air.

   The difference between plosives and affricates is that the release stage of affricates is slower, and therefore, the articulators spend a brief moment in the position for a fricative at the same place of articulation, resulting in audible friction. Another key difference is that the fricative release stage of affricates is always present while the release stage of plosives is very variable (Section 2.11).
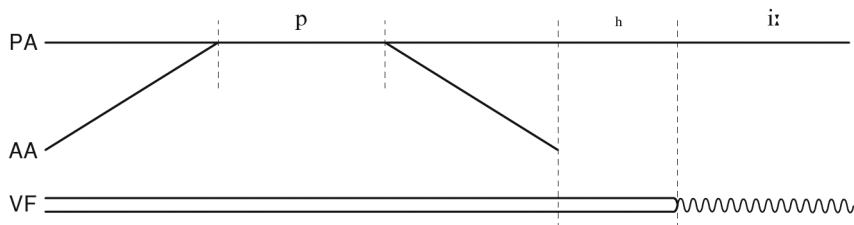
2.6 Aspiration

When the voiceless plosives /p t k/ are at the beginning of a stressed syllable, they are released with **aspiration**. This means that there's a brief period of voicelessness between the release of the plosive and the beginning of voicing for the next sounds (see Figure 2.9).

During this period, the vocal folds remain open, and the air rushing through the vocal tract gives the impression of a short [h]; hence, aspiration is shown with the symbol [ʰ] (e.g., [tʰ]).

[pʰ]	*peas* [pʰiz]	*palm* [pʰɑm]	*pound* [pʰaʊnd]
[tʰ]	*toad* [tʰoʊd]	*tail* [tʰeɪl]	*torn* [tʰorn]
[kʰ]	*kind* [kʰaɪnd]	*care* [kʰɛr]	*curb* [kʰərb]

When /p t k/ are not at the beginning of a stressed syllable, they are weakly aspirated or unaspirated. This is the case at the beginning of unstressed syllables and at the end of syllables but also, quite strikingly, when preceded by /s/ at the beginning of a syllable:

[p]	*perform* [pərˈform]	*pacific* [pəˈsɪfɪk]	*pyjamas* [pəˈdʒæməz]
[t]	*today* [təˈdeɪ]	*taboo* [təˈbu]	*together* [təˈgɛðər]
[k]	*concern* [kənˈsərn]	*kebab* [kəˈbab]	*canal* [kəˈnæl]

[p]	*leap* [lip]	*rope* [roʊp]	*map* [mæp]
[t]	*fight* [faɪt]	*goat* [goʊt]	*rate* [reɪt]
[k]	*like* [laɪk]	*sick* [sɪk]	*cheek* [ʧik]

[pʰ]	*pin* [pʰɪn]	*pot* [pʰat]	*pie* [pʰaɪ]
[p]	*spin* [spɪn]	*spot* [spat]	*spy* [spaɪ]

[tʰ]	*tone* [tʰoʊn]	*top* [tʰap]	*till* [tʰɪl]
[t]	*stone* [stoʊn]	*stop* [stap]	*still* [stɪl]

[kʰ]	*core* [kʰor]	*kill* [kʰɪl]	*cool* [kʰul]
[k]	*score* [skor]	*skill* [skɪl]	*school* [skul]

PA	=	passive articulator
AA	=	active articulator
VF	=	vocal folds
═══	=	vocal folds apart as for voiceless
∿∿∿	=	vocal folds vibrating as for voiced
────	=	vocal folds together as for glottal stop (see Fig. 2.10)

Figure 2.9 Aspiration in English /p/ as in *pea*; the diagram shows the brief period of voicelessness after the release of the plosive and before the voicing for the vowel

When aspirated /p t k/ are followed by approximants /l r w j/, the aspiration takes place during the articulation of the approximant, partially or fully devoicing the approximant [l̥ r̥ j̊ w̥] and causing turbulence and fricative noise at the place of articulation of the approximant, for example:

/p/	*play* [pl̥eɪ]	*print* [pr̥ɪnt]	*pew* [pj̊u]
/t/		*treat* [tr̥it]		*twin* [tw̥ɪn]
/k/	*clay* [kl̥eɪ]	*cream* [kr̥im]	*cue* [kj̊u]	*queasy* [ˈkw̥izi]

When /s/ precedes these clusters, there's no aspiration and therefore the approximants do not become devoiced or fricative:

/sp/	*splay* [spleɪ]	*sprint* [sprɪnt]	*spew* [spju]
/sk/		*scream* [skrim]	*skew* [skju]	*squeezy* [ˈskwizi]

As regards aspiration and devoicing, the /str/ cluster is a little irregular. Unlike other /s/ + voiceless plosive + approximant clusters, the /r/ remains somewhat devoiced and fricative. Consequently, the /tr/ of *strap* is not very different from, or even identical to, the /tr/ of *trap*.

Because the voiced plosives are frequently partially or fully devoiced (see Section 2.4.2), the presence of aspiration is an important cue for distinguishing /p t k/ from /b d ɡ/.

2.7 Glottal plosive [ʔ]

In addition to the bilabial /p b/, alveolar /t d/, and velar /k ɡ/ plosives, the glottal plosive [ʔ] (often referred to as "glottal stop") also occurs in English. The closure for a glottal plosive is made by bringing the vocal folds firmly together in an articulation similar to that of a very weak cough. The vocal folds are unable to vibrate during the production of a glottal plosive and therefore the sound has no voiced equivalent. Although common, the glottal plosive is not an independent phoneme in English (though it may be in other languages). The glottal plosive has a number of uses (see the following sections).

2.7.1 Hard attack

When a word starting with a vowel is preceded by a voiceless consonant or a pause, the usual way to begin the vowel is to gently bring the vocal folds together into the position for voicing or, if the preceding sound is voiced, to continue the voicing as the articulators move into the position for the vowel. An alternative to this is to begin with a glottal plosive: irrespective of whether the vocal folds are apart for a preceding voiceless consonant or a pause or vibrating for a preceding voiced sound, the vocal folds are brought tightly together and on release immediately take up the position for voicing, giving the impression of a very abrupt start to the vowel. This is known as **hard attack** and is used for emphasis in English, although in certain other languages, this may be the most usual treatment of word-initial vowels.

Without hard attack:	*This is an apple.* [ˈðɪs ɪz ən ˈæpl̩]
With hard attack:	*This is an apple (not a pear)!* [ˈðɪs ɪz ən ˈʔæpl̩]

A similar use of the glottal plosive is to separate sequences of vowels within words. In words like *react* /riˈækt/, *cooperate* /koʊˈɑpəreɪt/, and *deodorant* /diˈoʊdərənt/, the transition from one vowel to the next usually consists of a rapid glide of the tongue from the first vowel position to the second. An alternative in emphatic speech is to insert a glottal plosive between the vowels as the tongue moves between the vowel positions. This is only possible when the second vowel is stressed (i.e., not in *serious* /ˈsiriəs/, *fluent* /ˈfluənt/, *leotard* /ˈliətɑrd/).

2.7.2 Glottal replacement

The most important occurrence of the glottal plosive in English is as an allophone of the /t/ phoneme. This is known as **glottal replacement**. Glottal replacement occurs in only a specific set of phonetic contexts. The most important of these is when /t/ is in a syllable coda, preceded by a sonorant (i.e., vowel, nasal, or approximant) and followed by another consonant:

Within words:	*butler* [ˈbəʔlər], *lightning* [ˈlaɪʔnɪŋ], *pitfall* [ˈpɪʔfɑl], *tents* [tɛnʔs]
Between words:	*felt wrong* [ˈfɛlʔ ˈrɑŋ], *sent four* [ˈsɛnʔ ˈfor], *light rain* [ˈlaɪʔ ˈreɪn], *part time* [ˈpɑrʔ ˈtaɪm]

Glottal replacement is also common before a pause, e.g., *wait* [weɪʔ].

Although, in GA, [ʔ] does not occur between vowels in word-medial position, as in *meeting*, it can be heard in word-final position before a vowel in high-frequency words, for example *got a* [ˈgɑʔ ə], *met us* [ˈmɛʔ əs]; note, however, that t-tapping (see Section 2.8) is far more common in this position.

2.7.3 Glottal reinforcement

In addition to [ʔ] acting as an allophone of /t/, it can also occur together with /t/ and the other voiceless stops, /p k ʧ/, in a process known as **glottal reinforcement**. A glottal closure overlaps with the oral closure (see Figure 2.10): first the glottal closure is made; then the bilabial, alveolar, palato-alveolar, or velar closure is made; then the glottal closure is released inaudibly behind the oral closure before finally the oral closure is released; this phenomenon
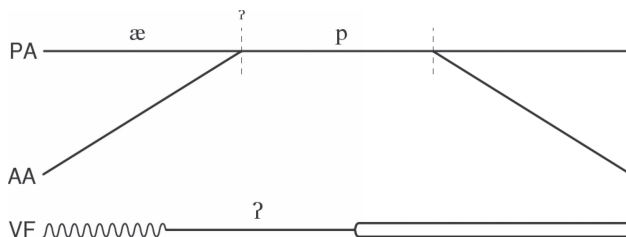

Figure 2.10 Glottal reinforcement in English /p/, as in *captive*. The reinforcing glottal plosive is formed before the hold stage of /p/ and released before the release of the bilabial plosive. (See Figure 2.9 for key to symbols.)

is sometimes also referred to as "pre-glottalization." Glottal reinforcement occurs in the same phonetic contexts as for glottal replacement of [t] (except for /ʧ/, which does not have to be followed by a consonant), and like glottal replacement, although it's common, it isn't obligatory.

captive [ˈkæʔptɪv]	*keep calm* [ˈkiʔp ˈkɑm]	
curtsey [ˈkərʔtsi]	*hot sauce* [ˈhɑʔt ˈsɑs]	
faultless [ˈfɑlʔtləs]	*don't know* [ˈdoʊnʔt ˈnoʊ]	
action [ˈæʔkʃn̩]	*take five* [ˈteɪʔk ˈfaɪv]	
hatchet [ˈhæʔʧət]	*catch me* [ˈkæʔʧ mi]	*catch it* [ˈkæʔʧ ɪt]

2.8 Tapping

In certain phonetic contexts, the alveolar plosives /t d/ and the alveolar nasal /n/ are articulated very rapidly. The tongue tip and blade move to and away from the alveolar ridge very quickly, and the contact between the active and passive articulators is very brief. This manner of articulation is known as a **tap**, and the IPA symbol for a voiced alveolar tap is [ɾ]. A voiced alveolar tap [ɾ] is used as the usual realization of the /r/ phoneme in many languages around the world and can also be heard in a number of English accents. Speakers who have a tap for /r/ in their language or accent of English may feel that the GA tap for /t d/ isn't quite the same as their own.

The most striking effect of tapping in GA is that /t/ becomes voiced and indistinguishable from /d/. They are both realized as a voiced alveolar tap [ɾ], and the difference between them is neutralized in tapping contexts. In contrast, the effect on the realization of /n/ is relatively minor. When tapped, /n/ is realized as a nasalized tap [ɾ̃], the equivalent of tapped /t/ or /d/ with the velum lowered. The result is a sound that isn't very different from /n/ in other contexts and that isn't easily confused with any other sound.

Some works on GA phonetics transcribe **t-tapping** with [ɾ], which emphasizes the particular type of articulation used. Other works transcribe t-tapping with [d], which demonstrates that the tapped variant of /t/ sounds to the ears of English speakers more like a variant of the /d/ phoneme than of any other phoneme. In this book, we prefer a third approach, which is used in two standard reference works: the *Longman Pronunciation Dictionary* and the *Cambridge English Pronouncing Dictionary*. The voiceless alveolar plosive symbol [t] is combined with the IPA "voiced" diacritic [ˌ] (t-tapping is also referred to as **t-voicing**) to give [t̬]. This symbolization has the advantage of making the transcription not too different from the orthography and of using a similar symbol for the same word in different phonetic contexts (e.g., *write* /raɪt/, *writing* /ˈraɪt̬ɪŋ/, *writes* /raɪts/, *write it* /raɪt̬ ɪt/, *write some* /raɪt səm/). This special symbol for t-tapping makes our transcription system a little unphonemic, but for the sake of simplicity, we will continue to use phonemic slanted brackets and not switch to phonetic square brackets for every instance of t-tapping.

T-tapping occurs when /t/ is at the end of a syllable and between vowels. Within a word, the following syllable must be unstressed, but the preceding syllable can be stressed or unstressed:

Stressed:	*city* /ˈsɪt̬i/, *vitamin* /ˈvaɪt̬əmən/, *critic* /ˈkrɪt̬ɪk/, *beautiful* /ˈbjut̬əfl̩/
Unstressed:	*deputy* /ˈdɛpjət̬i/, *quality* /ˈkwɑlət̬i/, *society* /səˈsaɪət̬i/

The following syllable can also be /əl/, realized as syllabic /l/, or /ər/, realized as schwar [ɚ] (the equivalent of syllabic [ɪ], see Section 2.23) but not /ən/, realized as syllabic /n/ (where /t/ is more likely to be realized as a glottal stop):

Syllabic /l̩/:   *bottle* /ˈbɑt̬l̩/, *total* /ˈtoʊt̬l̩/, *little* /ˈlɪt̬l̩/, *fatal* /ˈfeɪt̬l̩/, *rattle* /ˈræt̬l̩/
Schwar /ər/ [ɚ]:   *letter* /ˈlɛt̬ər/, *meter* /ˈmit̬ər/, *butter* /ˈbət̬ər/, *water* /ˈwɑt̬ər/
Syllabic /n̩/:   *button* /ˈbətn̩/, *kitten* /ˈkɪtn̩/, *rotten* /ˈrɑtn̩/, *threaten* /ˈθrɛtn̩/
         button [ˈbəʔn̩], *kitten* [ˈkɪʔn̩], *rotten* [ˈrɑʔn̩], *threaten* [ˈθrɛʔn̩]

The preceding sound doesn't have to be a vowel. It can also be /r/:

After /r/:   *dirty* /ˈdərt̬i/, *poverty* /ˈpavərt̬i/, *artist* /ˈɑrt̬ɪst/, *mortar* /ˈmɔrt̬ər/, *turtle* /ˈtərt̬l̩/

When the preceding sound is /l/, tapping is variable:

After /l/:   *alter* /ˈaltər/ **or** /ˈalt̬ər/, *guilty* /ˈɡɪlti/ **or** /ˈɡɪlt̬i/, *penalty* /ˈpɛnl̩ti/ **or** /ˈpɛnl̩t̬i/

When the preceding sound is /n/, the tendency is for the /t/ to be lost altogether, particularly in familiar, high-frequency words:

After /n/:   *twenty* /ˈtwɛni/, *winter* /ˈwɪnər/, *plenty* /ˈplɛni/, *wanted* /ˈwɑnəd/, *dentist* /ˈdɛnəst/

When t-tapping occurs between words, the following syllable does not have to be unstressed:

Unstressed: *get it* /ˈɡɛt̬ ɪt/, *visit us* /ˈvɪzət̬ əs/, *bet a lot* /ˈbɛt̬ ə ˈlɑt/, *late again* /ˈleɪt̬ ə ˈɡɛn/
Stressed: *get out* /ˈɡɛt̬ ˈaʊt/, *visit others* /ˈvɪzət̬ ˈəðərz/, *bet everything* /ˈbɛt̬ ˈɛvriθɪŋ/, *late evening* /ˈleɪt̬ ˈivnɪŋ/

The neutralization of /t/ and /d/ in these contexts creates a number of homophones: *writer/rider*, *latter/ladder*, *parity/parody*, *atom/Adam*, *metal/medal*, *petal/pedal*.

2.9 Nasal release

The bilabial /p b/ and alveolar plosives /t d/ have the same place of articulation as the bilabial /m/ and alveolar /n/ nasals respectively. They differ only in the soft palate being in the raised position (forming a velic closure) for the plosives and in the lowered position for the nasals. Consequently, it's possible to move from /p/ or /b/ to /m/ and from /t/ or /d/ to /n/ by lowering the soft palate only and leaving the oral closure in place. This is called **nasal release** and is usual in English when plosives are followed by their nasal equivalents. It's not usual for velar plosives /k ɡ/ because the velar nasal /ŋ/ does not occur at the beginning of syllables.

/pm/   *chipmunk* /ˈtʃɪpməŋk/   *ripe melon* /ˈraɪp ˈmelən/
/bm/   *submit* /səbˈmɪt/   *grab more* /ˈɡræb ˈmor/
/tn/   *witness* /ˈwɪtnəs/   *got none* /ˈɡɑt ˈnən/
/dn/   *kidney* /ˈkɪdni/   *good news* /ˈɡʊd ˈnuz/

Nasal release often results in syllabic /n/ in unstressed syllables (see Section 8.2.2). Syllabic /m/ and /ŋ/ only occur in such contexts as the result of assimilation (see Section 12.3.1).

[tn̩]	*button* [ˈbətn̩]	*cotton* [ˈkɑtn̩]	*eaten* [ˈitn̩]
[dn̩]	*hidden* [ˈhɪdn̩]	*sudden* [ˈsədn̩]	*wooden* [ˈwʊdn̩]

2.10 Lateral release

A similar process to nasal release occurs when /l/ follows /t/ or /d/. It's possible to go from the /t d/ position to the /l/ position simply by lowering one or both sides of the tongue, this being the only difference in the articulation of the sounds. This is called **lateral release**.

/tl/	*atlas* /ˈætləs/	*at least* /ət ˈlist/
/dl/	*badly* /ˈbædli/	*bad luck* /ˈbæd ˈlək/

Lateral release often results in syllabic /l/ (see Sections 2.26 and 8.2.1).

[tl̩]	*bottle* [ˈbɑtl̩]	*total* [ˈtoʊtl̩]	*crystal* [ˈkrɪstl̩]
[dl̩]	*idol* [ˈaɪdl̩]	*model* [ˈmɑdl̩]	*cradle* [ˈkreɪdl̩]

When the non-alveolar plosives, bilabial /p b/ and velar /k g/, are followed by /l/, **lateral escape** occurs. This means that the tongue-tip contact for /l/ is in place during the hold stage of the bilabial and velar plosives, and when they are released, the pent-up air in the vocal tract travels around the side(s) of this obstruction.

/pl/	*topless* /ˈtɑpləs/	*stop lying* /ˈstɑp ˈlaɪɪŋ/
/bl/	*public* /ˈpəblɪk/	*grab lunch* /ˈgræb ˈlənʧ/
/kl/	*backlash* /ˈbæklæʃ/	*thick legs* /ˈθɪk ˈlɛgz/
/gl/	*burglar* /ˈbərglər/	*big lights* /ˈbɪg ˈlaɪts/

Lateral escape also often results in syllabic /l/ (see Sections 2.26 and 8.2.1).

[pl̩]	*apple* [ˈæpl̩]	*pupil* [ˈpjupl̩]	*simple* [ˈsɪmpl̩]
[bl̩]	*double* [ˈdəbl̩]	*label* [ˈleɪbl̩]	*verbal* [ˈvərbl̩]
[kl̩]	*tackle* [ˈtækl̩]	*local* [ˈloʊkl̩]	*rascal* [ˈræskl̩]
[gl̩]	*eagle* [ˈigl̩]	*legal* [ˈligl̩]	*single* [ˈsɪŋgl̩]

2.11 Stop sequences

When two identical plosives occur in sequence, the first isn't usually released. Instead, the sequence is realized as a single long plosive consisting of an approach stage followed by a long hold stage and then a release stage:

/pp/	*ripe pear* /ˈraɪp ˈpɛr/	**/bb/**	*grab both* /ˈgræb ˈboʊθ/
/tt/	*get two* /ˈgɛt ˈtu/	**/dd/**	*red door* /ˈrɛd ˈdor/
/kk/	*black car* /ˈblæk ˈkɑr/	**/gg/**	*big guy* /ˈbɪg ˈgaɪ/

As in the case of two identical plosives, the first plosive is also usually unreleased when the plosives have the same place of articulation but differ in voicing:

/pb/	stop by /ˈstap ˈbaɪ/	/bp/	rob people /ˈrab ˈpipl̩/
/td/	hot dinner /ˈhat ˈdɪnər/	/dt/	red tie /ˈrɛd ˈtaɪ/
/kg/	black gown /ˈblæk ˈɡaʊn/	/gk/	big cat /ˈbɪɡ ˈkæt/

When plosives occur in sequence at different places of articulation, the closure for the second plosive is made before the closure for the first plosive is released. Consequently, the release of the first plosive is **inaudible** because a closure has already been made further forward in the mouth or because a closure further back in the mouth holds back the compressed air.

Second plosive further forward than first plosive:

/tp/	that part /ˈðæt ˈpart/	/tb/	what beach /ˈwət ˈbitʃ/
/dp/	bad place /ˈbæd ˈpleɪs/	/db/	good boy /ˈɡʊd ˈbɔɪ/
/kp/	black pony /ˈblæk ˈpoʊni/	/kb/	thick book /ˈθɪk ˈbʊk/
/gp/	big picture /ˈbɪɡ ˈpɪktʃər/	/gb/	big bang /ˈbɪɡ ˈbæŋ/
/kt/	lack time /ˈlæk ˈtaɪm/	/kd/	back door /ˈbæk ˈdor/
/gt/	big tip /ˈbɪɡ ˈtɪp/	/gd/	big dog /ˈbɪɡ ˈdɑɡ/

Second plosive further back than first plosive:

/pt/	top team /ˈtap ˈtim/	/pd/	strap down /ˈstræp ˈdaʊn/
/bt/	grab two /ˈɡræb ˈtu/	/bd/	job done /ˈdʒab ˈdən/
/pk/	deep cut /ˈdip ˈkət/	/pg/	cheap gift /ˈtʃip ˈɡɪft/
/bk/	drab colors /ˈdræb ˈkələrz/	/bg/	superb garden /suˈpərb ˈɡardn̩/
/tk/	what car /ˈwət ˈkar/	/tg/	that guy /ˈðæt ˈɡaɪ/
/dk/	sad case /ˈsæd ˈkeɪs/	/dg/	hard game /ˈhard ˈɡeɪm/

When an affricate occurs as the first in a sequence of stops, it is always released. The fricative release stage of affricates is always present, for example:

Sequences of the same affricate:

| /tʃtʃ/ | rich cheese /ˈrɪtʃ ˈtʃiz/ | /dʒdʒ/ | large jaw /ˈlardʒ ˈdʒa/ |

Sequences of affricates differing in voicing:

| /tʃdʒ/ | which job /ˈwɪtʃ ˈdʒab/ | /dʒtʃ/ | large child /ˈlardʒ ˈtʃaɪld/ |

Sequences of an affricate followed by a plosive:

/tʃp/	much pain /ˈmətʃ ˈpeɪn/	/tʃb/	catch both /ˈkætʃ ˈboʊθ/
/tʃt/	each time /ˈitʃ ˈtaɪm/	/tʃd/	each day /ˈitʃ ˈdeɪ/
/tʃk/	each case /ˈitʃ ˈkeɪs/	/tʃg/	teach grammar /ˈtitʃ ˈɡræmər/
/dʒp/	barge past /ˈbardʒ ˈpæst/	/dʒb/	edge back /ˈɛdʒ ˈbæk/
/dʒt/	large team /ˈlardʒ ˈtim/	/dʒd/	charge down /ˈtʃardʒ ˈdaʊn/
/dʒk/	large cuts /ˈlardʒ ˈkəts/	/dʒg/	huge grin /ˈhjudʒ ˈɡrɪn/

A number of alternative realizations for the first member of a stop sequence are possible:

1   When the first stop is voiceless (i.e., /p t k ʧ/), it can undergo glottal reinforcement (see Section 2.7.3).
2   When the first stop is an alveolar plosive (i.e., /t/ or /d/), it can undergo assimilation (see Section 12.3.1), in which case /tp tb dp db/ and /tk tg dk dg/ become /pp pb bp bb/ and /kk kg gk gg/ respectively.
3   When the first stop is /t/, it can undergo glottal replacement (see Section 2.7.2), /tp tb tk tg/ becoming [ʔp ʔb ʔk ʔg].

In all three cases, the first plosive is unreleased or inaudibly released.

2.12 Affricates

So far, as is usual in phonetics textbooks, we have been using the term *affricate* a little imprecisely. There's a distinction between **phonetic affricates** and **phonemic affricates** that we should bear in mind. Phonetically, an affricate is a sound that consists of a complete closure followed by **homorganic** friction. Accordingly, English /ʧ/ and /ʤ/ are phonetic affricates because both the stop element and the fricative element are palato-alveolar. But by our definition, the alveolar sequences /ts dz/ (as in *cats*, *kids*) are also phonetic affricates. The reason we consider /ʧ/ and /ʤ/ to be phonemic as well as phonetic affricates is that they behave as single, indivisible units in the English sound system and are felt to be single sounds by native speakers. /ts dz/, in contrast, behave like sequences and are felt to be sequences by natives.

Another pair of non-phonemic phonetic affricates is formed when /r/ follows /t/ or /d/. The closure for the /t/ or /d/ is post-alveolar in anticipation of the following /r/ (see Figure 2.11), and the release into the position for /r/ results in homorganic friction. The acoustic effect of /tr/ and /dr/ is similar to that of /ʧ/ and /ʤ/, but the two sets of sounds remain distinct: *chain* /ʧeɪn/ vs. *train* /treɪn/, *Jane* /ʤeɪn/ vs. *drain* /dreɪn/.
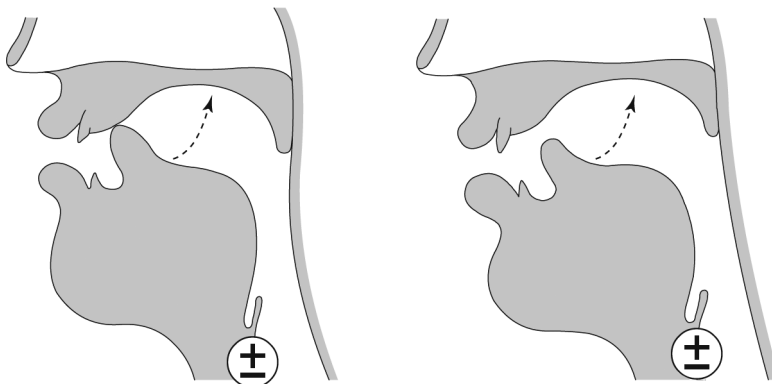

Figure 2.11 English post-alveolar (phonetic) affricates /tr/ and /dr/ [t̠ɹ, d̠ɹ] as in *train*, *drain*; the arrow indicates the raising of the sides of the tongue toward the back teeth

2.13 Fricatives

English has four pairs of voiceless and voiced fricatives at the labio-dental, dental, alveolar, and palato-alveolar places of articulation and a single voiceless glottal fricative.

- For /f/ and /v/, the lower lip lightly touches the upper incisors, and the airstream is forced through the gap (see Figure 2.12).
- For /θ/ and /ð/, the sides of the tongue form a seal against the side teeth, and the airstream is forced through a gap between the tongue tip and the rear of the upper incisors (see Figure 2.13). For some speakers, the articulation of /θ ð/ is interdental, i.e., made with the tongue tip projecting a little between the upper and lower incisors.
- For /s/ and /z/, the airstream is forced through a gap between the tongue tip/blade and the alveolar ridge, while the sides of the tongue form a seal against the upper side teeth (see Figure 2.14).
- For /ʃ/ and /ʒ/, the sides of the tongue form a seal against the upper side teeth, forcing the airstream through a narrow gap between the tongue tip/blade and the alveolar ridge and between the front of the tongue and the hard palate (see Figure 2.15).
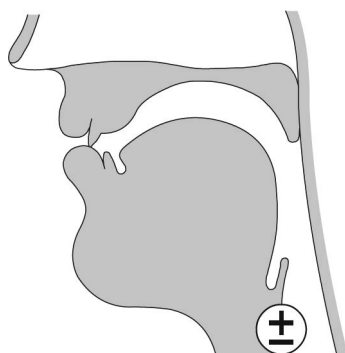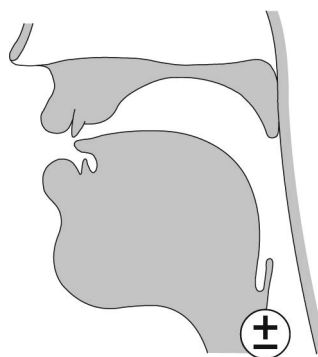

Figure 2.12 English fricatives /f/ and /v/

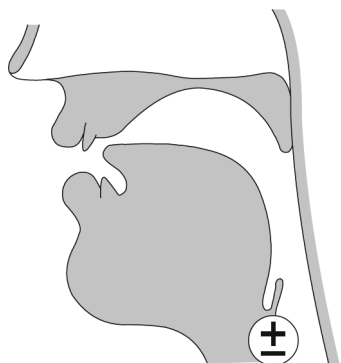*Figure 2.13* English fricatives /θ/ and /ð/

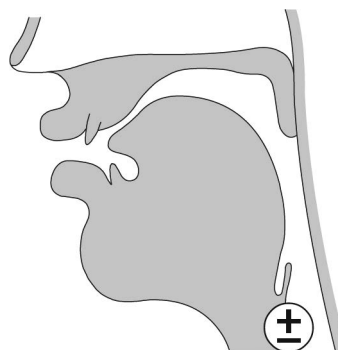*Figure 2.14* English fricatives /s/ and /z/

Figure 2.15 English fricatives /ʃ/ and /ʒ/

The primary articulation is accompanied by a simultaneous secondary articulation: rounding and protrusion of the lips.

• For /h/, the airstream is forced through the vocal tract with stronger than usual pressure, resulting in friction throughout the vocal tract and particularly at the narrowest point – the **glottis**, the space between the open vocal folds.

Each fricative is accompanied by the raising of the soft palate, forming a velic closure blocking the entrance to the nasal cavity.

2.14 Sibilants

Fricatives can be divided into the sub-classes of **sibilant** fricatives /s z ʃ ʒ/ and **non-sibilant** fricatives /f v θ ð h/. The tongue assumes a longitudinal **grooved** shape for sibilants (see Figure 2.16) and a flatter shape for non-sibilants. Grooving of the tongue channels the airstream into a jet that becomes turbulent and noisy when it strikes an obstruction further forward in the mouth – the teeth. In the case of non-sibilant fricatives, the turbulence and noise is generated at the stricture itself. The jet-of-air method of producing noisy turbulence is more effective than simple narrowing of the articulators, and hence sibilants are noticeably louder than non-sibilant fricatives. In the case of the English voiced non-sibilants, /v/ and /ð/, there's often very little fricative noise, and the difference between them and their approximant equivalents can be very slight.

2.15 /h/

During the articulation of /h/, the vocal tract assumes the position for the following sound, usually a vowel, which gives /h/ the quality of a voiceless or fricative version of the following sound. Thus, when isolated, the /h/ in *heart* is noticeably different from the /h/ in *hit*, the first having the quality of [ɑ̥] and the second of [ɪ̥]. In the case of a following /j/, as in *huge*, the /h/ is usually [ç], a voiceless palatal fricative. Fricative realizations of /h/ are also sometimes heard before back vowels – pharyngeal before /ɑ/ (e.g., *hot*), uvular before [o] (e.g., *horse*), or velar before /u/ (e.g., *whose*).

Between vowels, as in *ahead* or *a house*, a voiced glottal fricative [ɦ] is a common, but not obligatory, realization. The [ɦ] articulation involves a brief period of **breathy voice**, where the vibrating vocal folds do not fully come together as they vibrate, allowing air to escape between them and resulting in a "breathy" quality.
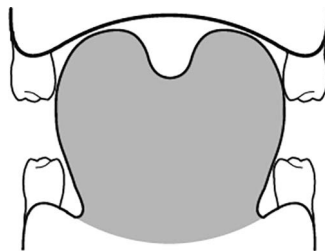

Figure 2.16 Mouth viewed from front showing grooved tongue shape for sibilants

2.16 Distribution of fricatives

Some of the fricatives are restricted in their distribution. /h/ only appears in syllable onsets before vowels (e.g., *hot*, *who*, *home*) or /j/ (e.g., *human*, *hue*), but not in syllable codas. /ʒ/ mainly occurs medially (e.g., *treasure*, *pleasure*, *vision*), being found initially and finally only in relatively recent loanwords (e.g., *genre*, *camouflage*), which often have alternative variants with /dʒ/.

2.17 Nasals

English has three nasals, all voiced, at the bilabial, alveolar, and velar places of articulation:

- For /m/, the lips come together and form a complete closure while the soft palate lowers to allow air to exit via the nose (Figure 2.17).
- For /n/, a complete closure is formed in the oral cavity by the tip of the tongue against the alveolar ridge and by the sides of the tongue against the upper side teeth. Simultaneously, the soft palate is lowered, allowing air to escape via the nose (Figure 2.18).
- For /ŋ/, the back of the tongue forms a closure against the soft palate, and the rear of the sides of the tongue form a seal against the rear upper side teeth. The soft palate is in the lowered position, allowing air to exit via the nose (Figure 2.19).

Note that these three articulations are the same as those for the bilabial /p b/, alveolar /t d/, and velar /k g/ plosives, only differing in the position of the soft palate – raised for plosives, lowered for nasals.
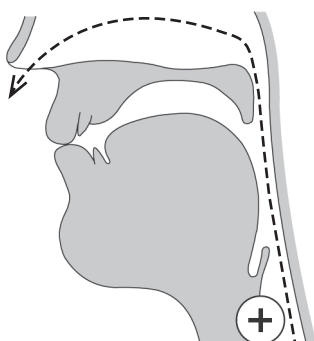
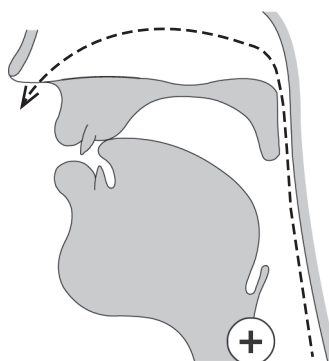*Figure 2.17* English nasal /m/; the arrow indicates the escape of the airstream through the nose

Figure 2.18 English nasal /n/; the arrow indicates the escape of the airstream through the nose
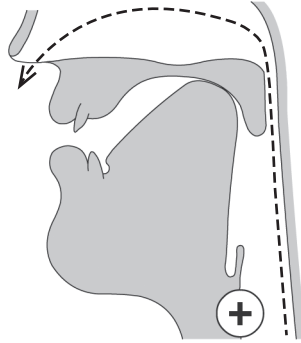
Figure 2.19 English nasal /ŋ/; the arrow indicates the escape of the airstream through the nose

2.18 Distribution of nasals

Historically, the voiced plosives /b/ and /g/ have been lost from the word-final clusters /mb/ and /ŋg/. The only voiced plosive that can occur word-finally after a nasal is /d/. All three *voiceless* plosives /p t k/, however, can occur in this position:

Bilabial /p b/:	*lump* /ləmp/	*comb* /koʊm/
Alveolar /t d/:	*count* /kaʊnt/	*hand* /hænd/
Velar /k g/:	*link* /lɪŋk/	*ring* /rɪŋ/

The preceding examples also demonstrate that within a morpheme, nasal + plosive clusters must be homorganic (i.e., have the same place of articulation). The bilabial nasal /m/ is followed by the bilabial plosive /p/, the alveolar nasal /n/ is followed by the alveolar plosives /t d/, and the velar nasal /ŋ/ is followed by the velar plosive /k/. Across a morpheme boundary, however, sequences of a nasal and a following plosive do not have to be homorganic:

Bilabial /m/ + alveolar /t d/:	*timetable* /ˈtaɪmteɪbl̩/	*someday* /ˈsəmdeɪ/
Bilabial /m/ + velar /k g/:	*tomcat* /ˈtɑmkæt/	*homegrown* /hoʊmˈɡroʊn/
***Alveolar /n/ + bilabial /p b/:**	*pinpoint* /ˈpɪnpɔɪnt/	*sunbed* /ˈsənbɛd/
***Alveolar /n/ + velar /k g/:**	*springclean* /sprɪŋˈklin/	*wineglass* /ˈwaɪnɡlæs/
Velar /ŋ/ + bilabial /p b/:	*gangplank* /ˈɡæŋplæŋk/	*songbird* /ˈsɑŋbərd/
Velar /ŋ/ + alveolar /t d/:	*Washington* /ˈwɑʃɪŋtən/	*kingdom* /ˈkɪŋdəm/

(*Note that /n/ is likely to assimilate when followed by /p b/ or /k g/; see Section 12.3.1.)

In the examples given here, although the nasal + plosive sequences are within a single word (often a compound word), they are separated by a syllable boundary, making them sequences, not true clusters (see Section 10.1). However, when verbs ending in /m/ or /n/ have their regular past/past participle form, they are followed by /d/, forming clusters, /md/ and /ŋd/, within the same syllable:

/md/:	*blamed* /bleɪmd/, *formed* /formd/, *harmed* /hɑrmd/, *named* /neɪmd/
/ŋd/:	*banged* /bæŋd/, *belonged* /bəˈlɑŋd/, *thronged* /θrɑŋd/, *wronged* /rɑŋd/

The velar nasal /ŋ/ is restricted to syllable-coda position (e.g., *sing* /sɪŋ/, *sink* /sɪŋk/) and occurs word-finally through the historical loss of /g/ from /ŋg/ at the end of words. When a suffix is added to a word ending in /ŋ/, the pronunciation remains /g/-less even though it's now in word-medial position: *ring* /rɪŋ/, *ringing* /ˈrɪŋɪŋ/; *hang* /hæŋ/, *hanger* /ˈhæŋər/; and so on. This leads to the generalization that word-medially, /ŋ/ is found at the end of morphemes (e.g., *singer*, *hanging*) and /ŋg/ within morphemes (e.g., *finger*, *angry*, *angle*, *bongo*) because in the latter words, /ŋg/ was never word-final, and therefore, /g/ was never lost. Exceptions to this rule are the words *long*, *young*, and *strong* in their comparative and superlative forms, which retain the historical /g/ lost elsewhere:

long /lɑŋ/	But	*longer* /ˈlɑŋgər/	*longest* /ˈlɑŋgəst/
young /jəŋ/	But	*younger* /ˈjəŋgər/	*youngest* /ˈjəŋgəst/
strong /strɑŋ/	But	*stronger* /ˈstrɑŋgər/	*strongest* /ˈstrɑŋgəst/

2.19 Syllabic nasals

The syllables /əm/, /ən/, and /əŋ/ can in certain circumstances (see Section 8.2) be realized as the syllabic consonants [m̩], [n̩], and [ŋ̩], meaning that the schwa /ə/ isn't articulated, and the nasal becomes the nucleus of the syllable (see Section 1.5). Of the three nasals, however, syllabic /n/ is much more common than the other two (see Section 8.2.2), and syllabic /ŋ/ only occurs as a result of assimilation (see Section 12.3.1). Syllabic consonants are very common in GA but not completely obligatory.

[m̩]	*chasm* /ˈkæzm̩/	*a dozen miles* /ə ˈdəzm̩ ˈmaɪlz/
[n̩]	*reason* /ˈrizn̩/	*a dozen nights* /əˈdəzn̩ ˈnaɪts/
[ŋ̩]	*taken* /ˈteɪkŋ̩/	*a dozen cats* /ə ˈdəzŋ̩ ˈkæts/

2.20 Approximants

English has four approximants, all voiced, at the alveolar, post-alveolar, palatal, and labial-velar places of articulation:

- For /l/, the tongue tip touches the alveolar ridge, while one or both of the sides of the tongue remain lowered, not making a seal with the side teeth and allowing air to flow around the tongue-tip contact (see Figure 2.20).
- For /r/, the tongue blade and front hollow while the tongue-tip curls slightly upward toward the rear part of the alveolar ridge (see Figure 2.21). The lips are often weakly rounded.
- For /j/, the front of the tongue moves toward the hard palate (see Figure 2.22).
- For /w/, the back of the tongue moves toward the soft palate and the lips round (see Figure 2.23).

Each approximant is accompanied by the raising of the soft palate, forming a velic closure blocking the entrance to the nasal cavity.

The double place name **labial-velar** reflects the fact that /w/ is a **double articulation** (i.e., one that involves two simultaneous articulations of equal degree – labial and velar approximants).
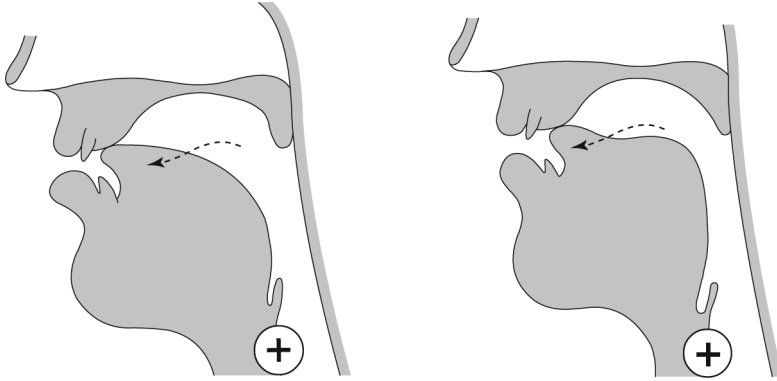
Figure 2.20 Alveolar lateral approximant /l/; left: clear /l/; right: dark /l/ showing velar-ized tongue shape; arrows indicate passage of airstream along lowered sides of tongue
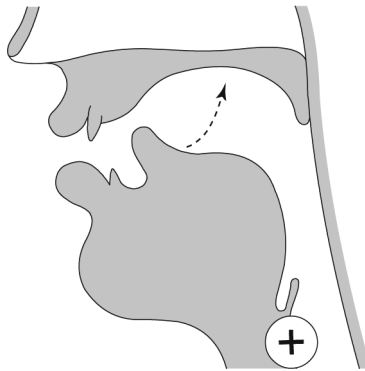

Figure 2.21 English /r/ (post-alveolar approximant); arrow indicates raising of sides of tongue toward back teeth
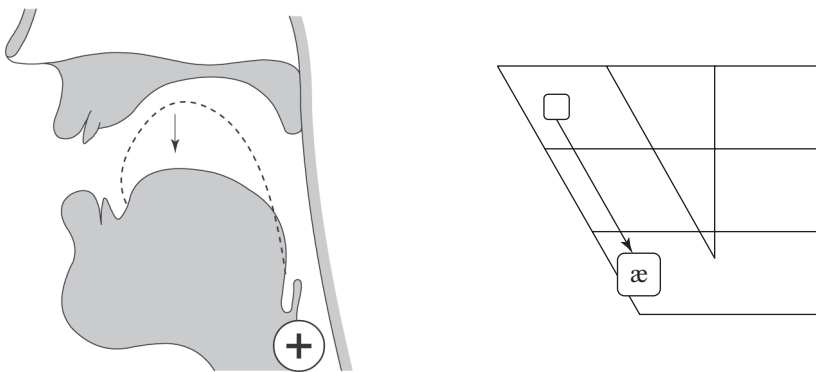

Figure 2.22 English /j/: sequence /jæ/ as in *yak*; diagram shows approximate change in tongue shape; since /j/ is a semi-vowel, it can be indicated on a vowel diagram
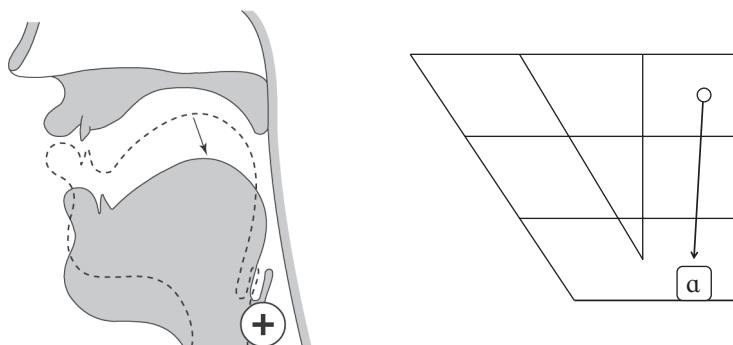
Figure 2.23 English /w/: sequence /wɑ/ as in *wasp*; diagram shows approximate change in tongue shape; since /w/ is a semi-vowel, it can be indicated on a vowel diagram
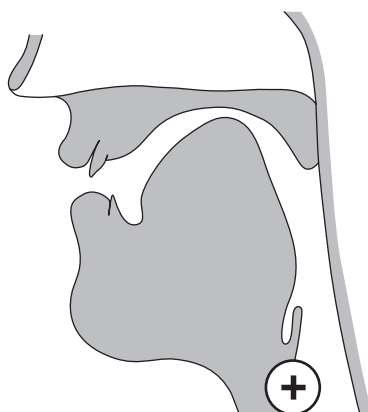

Figure 2.24 English /r/: bunched /r/

The approximants, like the nasals (and vowels), are sonorants and do not undergo the kind of devoicing that obstruents are subject to when adjacent to voiceless consonants or pauses. However, when preceded by an aspirated voiceless plosive, the aspiration takes place during the articulation of the approximant and partially or completely devoices it while the increased airflow through the approximant stricture causes friction (see Section 2.6).

2.21 Bunched /r/

So far, we have described the most common articulation of /r/, with the tongue-tip curling up toward the rear of the alveolar ridge. However, there is another relatively common way of articulating /r/ in GA, referred to as **bunched /r/**. The center of the tongue is raised (i.e., "bunched up") toward the area where the hard palate meets the soft palate (see Figure 2.24). As with the tongue-tip variant of /r/, the lips are usually weakly rounded.

2.22 Distribution of /r/: rhoticity

GA is an example of a **rhotic** accent. This means that it has retained /r/ in positions in which it has been lost in some other accents. GA has /r/ in words like *far* /fɑr/ and *farm* /fɑrm/, while **non-rhotic** accents have no /r/ in these words because historically they have lost /r/ at the ends of words and before consonants. To put it another way, in non-rhotic accents, /r/ only occurs before a vowel (i.e., in words like *red*, *bring*, *carry*, and *attract*). Scottish and Irish accents are usually rhotic, while the General British accent and nearly all the accents of Australia, New Zealand, and South Africa are non-rhotic. General American is rhotic, but the accents of New England and the American South are often non-rhotic.

2.23 Syllabic /r/

To foreign learners and speakers of other English accents, the /r/ phoneme stands out as being very characteristic of the GA accent. This is not only because GA is a rhotic accent or because the GA /r/ is an approximant (unlike the tapped or trilled /r/ of many other languages) but also because of its effect on preceding vowels, which tend to become partially r-colored in anticipation of a following /r/; see Section 5.7.2. When the preceding vowel is schwa /ə/, the whole vowel is r-colored and the phonemic sequence /ər/ is realized phonetically as [ɚ], which is known as "schwar," a very common sound in GA. An r-colored schwa [ɚ] is essentially the same as [ɹ] forming the nucleus of a syllable, and we could, therefore, analyze schwar as syllabic /r/ if we wished and transcribe *murmur*, *nurture*, and *cursor* as /ˈmr̩mr̩/, /ˈnr̩tʃr̩/, and /ˈkr̩sr̩/ instead of /ˈmərmər/, /ˈnərtʃər/, and /ˈkərsər/, but such transcriptions without vowels are difficult to read and using them systematically can be confusing to learners, so we don't use such an approach in this book. Note also that the distribution of schwar is different from the other syllabic consonants (see Section 8.2) because they only occur in unstressed syllables while schwar, as our examples demonstrate, is common in both stressed and unstressed syllables.

2.24 Median and lateral approximants

The approximants can be divided into **lateral** (/l/) and **median** approximants (/r j w/). As the names suggest, the airstream flows to the side of an obstruction for lateral approximants and along the midline of the oral cavity for the median approximants. You can test this by putting your tongue in the position for /l/ and breathing in deeply. You'll notice that one or both sides of your tongue feel cold because of the lateral airflow. Since /l/ is the only lateral approximant, it's often simply called a "lateral," and since all other English approximants are median, the label "median" isn't usually included for them.

2.25 Velarized /l/

Since only tongue-tip contact with the alveolar ridge is necessary for the articulation of /l/, the rest of the tongue, a very flexible articulator, is able to take up a range of different shapes. Different languages and different accents of English have different habits regarding the position of the body of the tongue during the articulation of /l/. Sometimes, the tongue shape is relatively neutral, rather flat (see Figure 2.20, left) or merely anticipates the position of a following vowel. Sometimes the front of the tongue is raised toward the hard palate, creating

a **palatalized** [l] or the back of the tongue is raised toward the soft palate (the velum), creating a **velarized** [l] (see Figure 2.20, right). In some languages, velarized or palatalized [l] contrasts phonemically with non-velarized or non-palatalized [l], the two sounds counting as different phonemes. In other languages or accents, different varieties of [l] occur in different phonetic contexts and count as variants of the same /l/ phoneme. In GA, the tendency is for /l/ to be velarized in all positions, especially at the end of words (e.g., in *hill*, *doll*, and so on) and before consonants (e.g., in *help*, *build*, and so on). Velarization and palatalization are types of secondary articulation (see Section 2.2.4). Velarized /l/ is also known as **dark /l/** and has the IPA symbol [ɫ], while non-velarized /l/ is often referred to as **clear /l/**.

2.26 Syllabic /l/

The syllable /əl/ is very often realized as syllabic /l/ (i.e., [l̩] (see Section 8.2.1)). The articulators move directly from the preceding consonant to /l/, the schwa /ə/ isn't articulated, and the /l/ becomes the nucleus of the syllable. Syllabic /l/ is the most common syllabic consonant in GA, occurring in a greater range of contexts than syllabic /n/.

Syllabic /l/:   *people* [ˈpipl̩], *devil* [ˈdɛvl̩], *satchel* [ˈsætʃl̩], *camel* [ˈkæml̩], *giggle* [ˈɡɪɡl̩]

2.27 Semi-vowels

Another way of grouping the approximants is according to the part of the tongue involved in their articulation, giving the terms **semi-vowel** (/j/ and /w/) and **non-semi-vowel** (/l/ and /r/). The semi-vowels are articulated with the part of the tongue that is used to articulate vowels – the front, center, and back. The position for /j/ and /w/ are the same as for the [i] and [u] vowels respectively. [j] and [i] both consist of an approximation of the front of the tongue to the hard palate, and [w] and [u] both consist of an approximation of the back of the tongue to the soft palate accompanied by simultaneous lip-rounding. Thus, [j] and [w] are glides from these vowel positions to a vowel of longer duration. The non-semi-vowels /l/ and /r/ involve the tip and blade of the tongue, which are not used in the articulation of vowels.

2.28 Distribution of approximants

The semi-vowels /j/ and /w/ only occur in syllable onsets in English (e.g., *young* /jəŋ/, *unit* /ˈjunət/, *few* /fju/, *one* /wən/, *wax* /wæks/, *twin* /twɪn/). Learners from other language backgrounds may feel that /j/ and /w/ can also occur in syllable codas, but in English, such non-onset glides toward the [i] and [u] positions are best analyzed as part of the syllable nucleus, forming diphthongs (see Section 5.5):

FACE:	*cake* /keɪk/	*sail* /seɪl/	*day* /deɪ/
PRICE:	*like* /laɪk/	*ride* /raɪd/	*try* /traɪ/
CHOICE:	*voice* /vɔɪs/	*coin* /kɔɪn/	*boy* /bɔɪ/
GOAT:	*oak* /oʊk/	*rose* /roʊz/	*no* /noʊ/
MOUTH:	*shout* /ʃaʊt/	*clown* /klaʊn/	*plow* /plaʊ/

The lateral approximant /l/ occurs in both syllable onsets and codas, and since GA is a rhotic accent (see Section 5.7), /r/ also occurs freely in these positions.

2.29 Yod-dropping

We've seen that /j/ only occurs in syllable onsets, but even in onsets, /j/ tends to be lost after certain consonants in stressed syllable-initial positions. This phenomenon is known as **yod-dropping**, "yod" being a name for the palatal approximant [j]. Historically, /j/ occurred with /u/ in words, where /u/ has a spelling that includes <u> or <w>, but has been lost after palato-alveolar consonants:

/ʃ tʃ dʒ/:   *chute* /ʃut/, *chew* /tʃu/, *juice* /dʒus/

After approximants:

/r l/:   *rude* /rud/, *blue* /blu/, *crew* /kru/, *lute* /lut/, *allude* /əˈlud/

And after alveolar and dental fricatives:

/s z θ/:   *suit* /sut/, *assume* /əˈsum/, /z/ *presume* /prəˈzum/, *enthusiasm* /ɪnˈθuziæzm̩/

Most GA speakers, in contrast to speakers of some other English accents, have also lost /j/ after alveolar plosives and the alveolar nasal:

/t d n/:   *tune* /tun/, *duke* /duk/, *news* /nuz/

After the remaining consonants, /j/ is retained (i.e., after labio-dentals, bilabials, velars, and /h/):

/f v/:   *few* /fju/, *fuse* /fjuz/, *feud* /fjud/, *view* /vju/
/p b/:   *pew* /pju/, *puny* /pjuni/, *putrid* /ˈpjutrəd/, *beauty* /ˈbjuṭi/, *abuse* /əˈbjus/
/k g/:   *cute* /kjut/, *cue* /kju/, *accuse* /əˈkjuz/, *argue* /ˈɑrgju/, *legume* /ˈlɛgjum/
/h/:   *human* /ˈhjumən/, *huge* /hjudʒ/, *hew* /hju/

In the case of the /hj/ cluster, the /h/ is sometimes dropped by GA speakers, giving *human* /ˈjumən/, *huge* /judʒ/, and so on. It is not necessary for learners to imitate this, however.

Immediately following a stressed syllable, /j/ is retained after /n l/ and variable after /r/, while /tj dj sj zj/ have coalesced over time to yield /tʃ dʒ ʃ ʒ/ respectively. The original /u/ has sometimes been replaced by schwa /ə/, which is typical of unstressed vowels in English.

/n l/:   *venue* /ˈvɛnju/, *annual* /ˈænjuəl/, *value* /ˈvælju/, *volume* /ˈvɑljum/
/r/:   *erudite* /ˈɛrədaɪt/ or /ˈɛrjədaɪt/, *virulent* /ˈvirələnt/ or /ˈvirjələnt/
/tj dj sj zj/:   *statue* /ˈstætʃu/, *module* /ˈmɑdʒul/, *issue* /ˈɪʃu/, *visual* /ˈvɪʒuəl/

2.30 Inflections

The pronunciation of the **-s** suffix (ending), occurring in plurals (e.g., *cats*), third person present tense endings (e.g., *thinks*), and possessives (e.g., *John's*), is determined by the preceding sound. It's pronounced /əz/ following sibilants, /s/ following all other voiceless consonants, and /z/ following all other voiced sounds:

/əz/:	*buses, Alice's, seizes, Rose's, wishes, churches, catches, judges, George's, camouflages*
/s/:	*stops, lips, hats, Pete's, thinks, Mike's, laughs, Jeff's, myths, Beth's*
/z/:	*Bob's, needs, bags, leaves, breathes, seems, Kevin's, things, Jill's, cars, bees, Sue's, bras, ways, Joe's, lies, cows, toys, Hilda's*

The same rule applies in contractions involving auxiliary *has* and *is* (see Sections 8.6.3 and 8.7.2):

has:	*What's happened?* /ˈwɑts ˈhæpənd/, *Mike's left.* /ˈmaɪks ˈlɛft/, *Who's gone.* /ˈhuz ˈgɑn/
is:	*That's good.* /ˈðæts ˈgʊd/, *It's difficult.* /ɪts ˈdɪfəkəlt/, *Where's Pete?* /ˈwɛrz ˈpit/

The pronunciation of the past tense ending -ed (e.g., *talked*) is also governed by the preceding sound. It's pronounced /əd/ following /t/ or /d/, /t/ following all other voiceless consonants, and /d/ following all other voiced sounds, both consonants and vowels:

/əd/:	*waited, needed*
/t/:	*stopped, thanked, watched, laughed, unearthed, kissed, wished*
/d/:	*robbed, begged, judged, saved, breathed, used, seemed, frowned, banged, sailed, starred, played, showed, sighed, allowed, employed*

Note that certain adjectives (e.g., like *crooked, dogged, naked, -legged, wicked*) take the /əd/ ending.