

British English: Received Pronunciation

Peter Roach

School of Linguistics and Applied Language Studies

University of Reading

p.j.roach@reading.ac.uk

The accent described here is the present-day version of the accent that has been used as the standard in phoneticians' description of the pronunciation of British English for centuries. The definition of this accent is a matter of heated debate and frequent controversy: the arguments will not be rehearsed here, but the interested reader is recommended to look at Jones (1917 and subsequent) and Wells (2000). The most important aspects of this accent should, however, be made clear:

- a. The number of native speakers of this accent who originate in Ireland, Scotland and Wales is very small and probably diminishing, and it is therefore a misnomer to call it an accent of BRITISH English. It is an accent spoken by some English people.
- b. The great majority of native speakers of this accent are of middle-class or upper-class origin, educated at private schools and (if of appropriate age) university. This does not mean that the accent cannot be acquired by others: the present author (who attended a state school in the Midlands) originally spoke with an accent with noticeable regional features, but has over many years of teaching the phonetics of English acquired an accent not far from the standard one described here.
- c. The majority of speakers of this accent live in, or originate from, the south-east of England.
- d. The accent is most familiar as that used by most 'official' BBC speakers of English origin (newsreaders and announcers on Radio 4 and Radio 3, and most television channels). It is also frequently heard on the BBC World Service, though that service appears to have adopted the policy of sometimes using newsreaders and announcers with noticeable foreign accents. It is clear that this accent will eventually lose its pre-eminent status in broadcasting as a result of the wish to broaden the social base of broadcast speech, but it will take a long time for this to happen.

The accent has been known for nearly a century as RECEIVED PRONUNCIATION, or by its abbreviation, RP. Early in the 20th century, Daniel Jones, the great exponent of the description of English pronunciation, named it PUBLIC SCHOOL PRONUNCIATION (Jones 1917), but later changed the name to Received Pronunciation. Other names have been proposed, such as GENERAL BRITISH (GB) and EDUCATED SOUTHERN BRITISH ENGLISH. The present author's own preference is for the name BBC PRONUNCIATION or BBC ACCENT (Jones, ed. by Roach et al. 2003), but given the continuing popularity of the name Received Pronunciation, this has been used for the description which follows.

The choice of symbols for the representation of RP is one which has provoked much discussion, but since the 1980s there has existed, largely as a result of pressure from the major ELT publishers, a *de facto* standard set of symbolization conventions which has remained almost unchanged to the present day. It is widely accepted that some modification of these conventions is needed to take account of observable changes in the pronunciation of English, but it is felt to be important that such modifications should not be introduced without general agreement among practitioners of English phonetics so that the benefits of a common system of transcription enjoyed over the last twenty or thirty years should not be lost. An alternative set of transcriptions has, however, been used in the *Oxford Dictionary of Pronunciation* (Upton et al. 2001).

Consonants

	Bilabial	Labiodental	Dental	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosive	p b			t d			k g	
Affricate					tʃ dʒ			
Nasal	m			n			ŋ	
Fricative		f v	θ ð	s z	ʃ ʒ			h
Approximant	(w)				r	j	w	
Lateral approximant				l				

p	<i>pea</i>	b	<i>bee</i>
t	<i>toe</i>	d	<i>doe</i>
k	<i>cap</i>	g	<i>gap</i>
f	<i>fat</i>	v	<i>vat</i>
θ	<i>thing</i>	ð	<i>this</i>
s	<i>sip</i>	z	<i>zip</i>
ʃ	<i>ship</i>	ʒ	<i>measure</i>
h	<i>hat</i>		
tʃ	<i>chin</i>	dʒ	<i>gin</i>
m	<i>map</i>	l	<i>led</i>
n	<i>nap</i>	r	<i>red</i>
ŋ	<i>hang</i>	j	<i>yet</i>
		w	<i>wet</i>

Plosives, fricatives (with the exception of /h/) and affricates show a distinction between VOICED and VOICELESS, or, alternatively, between LENIS and FORTIS, e.g. /b/ vs. /p/, /z/ vs. /s/. Voicing of so-called voiced consonants is often very weak or even undetectable, a fact which has led to the idea that force of articulation is what distinguishes pairs of obstruent consonants, /p/, for example, being classed as fortis (strongly articulated) and /b/ as lenis (weakly articulated). Vowels are markedly shortened before fortis consonants in the same syllable. The plosives /p, t, k/ are aspirated before vowels (i.e. part of the following vowel is devoiced), and in an equivalent process /l, /r/, /w/ and /j/ are partially devoiced following these plosives. This is much less noticeable if the following vowel is unstressed. The devoicing is usually blocked if /s/ precedes the plosive within a syllable. Syllable-final /p, t, k, tʃ/ are frequently preceded by a glottal stop [ʔ] unless followed by a vowel, though there are few examples of this in the recording on which our transcription is based. It sometimes happens that /t/ is replaced, rather than preceded by, a glottal stop, particularly before a syllabic nasal,

as in *button* [bʌʔn]. The usual phonetic realization of the glottal stop in English is a few pulses of creaky voice at the end of the preceding vowel: in the transcription of the illustrative passage given below, the word *attempt* is transcribed allophonically as [ə^hˈemʔt], but a truly phonetic transcription might represent it as [ə^hˈemmt]. Among the fricatives, /s, z, ʃ, ʒ/ are of relatively high intensity, while /f, v, θ, ð, h/ are of low intensity. In particular, /ð/ is often a weak dental plosive with no detectable friction noise; the sequence /nð/ often assimilates to [nn̩]. The glottal fricative /h/ is usually voiced [ɦ] in a voiced environment. The lateral /l/ is ‘clear’ ([i]-coloured) before vowels and ‘dark’ ([u]-coloured) elsewhere. The /r/ phoneme is usually realized as a post-alveolar approximant [ɹ] unless devoicing results in a voiceless fricative articulation. Alveolar consonants are frequently assimilated to the place of articulation of a following non-alveolar consonant. The sonorants /m, n, ŋ, l, r/ may be pronounced as syllabic consonants in place of a weak syllable containing a vowel, e.g. *bottle* /bɒtl/, *button* /bʌtɹn/.

Vowels

RP is traditionally described as having a large inventory of vowels that can be grouped into SHORT VOWELS, LONG VOWELS and DIPHTHONGS. The diphthongs may be subdivided into CENTRING diphthongs (ending in /ə/) and CLOSING diphthongs which end in either /ɪ/ or /ʊ/. The number of these vocalic elements is around twenty, but the conventional view that these twenty or so elements are all discrete phonemes of the language is something that is open to counter-argument: a much more parsimonious analysis is possible. It is noticeable in this accent that unstressed vowels are much shorter than stressed ones, and are frequently centralized.

It should be understood that the terms ‘long’ and ‘short’ should be seen in relative terms: the vowels of both classes are subject to the lengthening and shortening effects found in English, with the result that a ‘short’ vowel may, in some contexts, be longer than a ‘long’ vowel in a different context. The length mark : is used to mark the long vowels, though this is actually redundant since the vowel symbols already successfully distinguish each vowel from every other (see Ladefoged 2001, for further discussion). The development of these transcription conventions is usually credited to Gimson (1962) and they were adopted by him for his 14th Edition of the *English Pronouncing Dictionary* (Jones, ed. by Gimson, 1977); it was his belief that this redundancy helped rather than hindered the learner of English phonetics. One convention, however, has become widespread since Gimson’s time, and that is the use of two additional symbols, i and u, to represent unstressed vowels which occur before vowels and in final position, and are qualitatively more similar to /i:/ and /u:/ than to /ɪ/ and /ʊ/. These cannot be classed as true phonemes, since they contrast neither with /i:/, /u:/ nor with /ɪ/, /ʊ/; they represent vowels in contexts where the contrasts between the respective long and short vowels are neutralized. The SCHWA vowel is of great importance: though it has no unique representation in spelling, and is often pronounced so weakly as to be almost undetectable, it is the most frequently-occurring vowel in this accent.

The vowels may be represented on Cardinal Vowel charts as shown below. It will be noticed that the position of some of the symbols is quite far removed from the Cardinal (or IPA) value of those symbols (see for example /æ/, /ɔ:/, /u:/, /ʌ/, /e/); this is largely for historical reasons, illustrating the need for a modernized transcription. A further symbolization problem concerns diphthongs: in many cases the vowel glide implied by the symbols in such vocalic elements is so small as to be almost undetectable, so that the contrast between, for example, *day* and *dare* may appear to be due to the difference between [e:] and [ɛ:]. A closing diphthong can have a /ə/ vowel attached to it (e.g. *fire* /faɪə/); the resulting complex vowel unit may be classed as a TRIPHTHONG if it is pronounced as a single syllable, but it is often almost impossible to detect any vowel glide in such an element: see for example the vocalic part of the first syllables of *fire-engine*, *Ireland*, which are often [fɪ], [ɹ] despite usually being represented as /aɪə/. Typical formant frequency values for the vowels and diphthongs may be found in Gimson 1962 (ed. by Cruttenden, 2001, pp. 98–102).

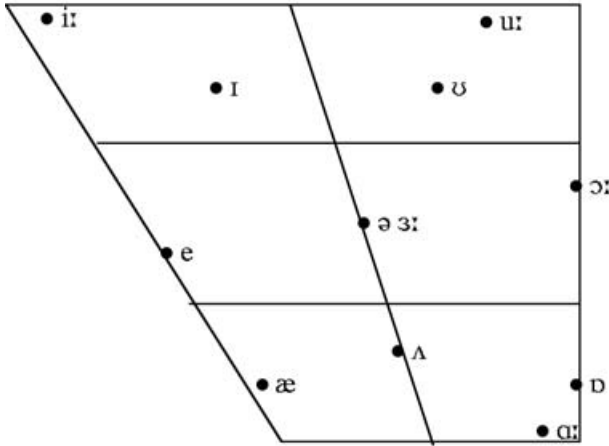


Figure 1 RP pure vowels.

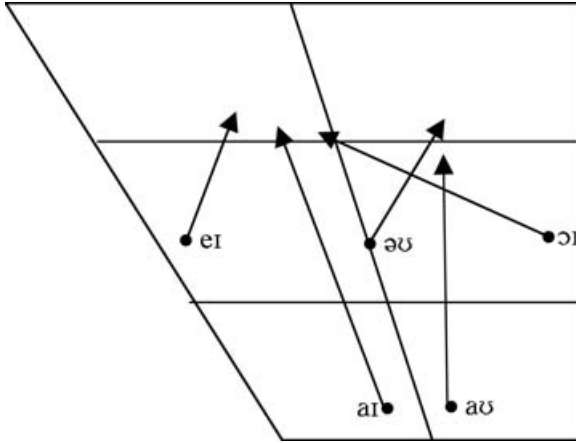


Figure 2 RP closing diphthongs.

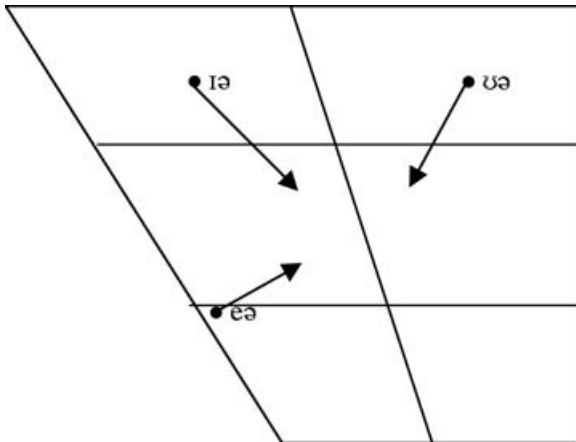


Figure 3 RP centring diphthongs.

Short vowels

ɪ	<i>pit</i>
e	<i>pet</i>
æ	<i>pat</i>
ʌ	<i>cut</i>
ɒ	<i>pot</i>
ʊ	<i>put</i>
ə	<i>about, India</i>

Long vowels

i:	<i>key</i>
ɑ:	<i>half</i>
ɔ:	<i>paw</i>
u:	<i>coo</i>
ɜ:	<i>cur</i>

Diphthongs

eɪ	<i>bay</i>	əʊ	<i>go</i>
aɪ	<i>buy</i>	aʊ	<i>cow</i>
ɔɪ	<i>boy</i>		

ɪə	<i>peer</i>
eə	<i>pear</i>
ʊə	<i>poor</i>

Weak close vowel symbols

ɪ	<i>happy</i>	ʊ	<i>throughout</i>
---	--------------	---	-------------------

Prosodic features

It is generally considered sufficient to transcribe three levels of stress on syllables: primary (marked '), secondary (marked ,) and unstressed (no mark). The use of these marks in the example text should be taken to record the transcriber's impression of the degree of prominence heard. It has often been remarked that English stress is both FREE (in that any syllable is capable in principle of receiving stress) and FIXED (since it only rarely happens in a particular context that more than one stress placement is acceptable). Many attempts have been made to produce rules for the placement of stress, either within the word or in higher-level units, but such rules have frequent exceptions.

Many transcription systems for English intonation have been proposed, but no accepted standard exists. There is general agreement that English speech may be divided into intonational units of one or more syllables, each of which contains one major accent, and that different pitch-accents or tones correspond to various attitudinal or syntactic functions, but it is difficult to identify any intonational features that are unique to English, and the nature and the number of pitch-patterns or pitch-accents is still a matter of dispute.

English RHYTHM is said to be STRESS-TIMED, i.e. the intervals between stressed syllables tend to be constant and unstressed syllables are compressed to preserve the isochrony of the inter-stress intervals. While the evidence for this is not completely conclusive, it is clear that in RP there is a very marked difference between weak, unstressed syllables which in some contexts may be almost undetectable and strong syllables (stressed or unstressed) which are fully pronounced.

Transcriptions

The recording used for this illustration was made by a female speaker who was born in 1953. She was educated first at a private preparatory school, then at a 'traditional girls' grammar

school', and finally at Oxford University. She teaches in the English Department of a foreign university.

The original recording is on Minidisk, and has been made available to the IPA along with this analysis. The acoustic analysis of the recording was made using the SIL Speech Analyzer program: the file was divided into nine chunks and phonetic labelling of acoustic segments was carried out. The phonemic transcription is idealized, and various connected speech phenomena that are recorded in the allophonic transcription are ignored. The allophonic transcription is different from a purely phonetic transcription as described by Pike (1943) in that it adds phonetic detail to the basic phonological structure given in the phonemic transcription. It should be noted that many of the phonologically voiced consonants marked as devoiced are at times within the consonant less than fully devoiced, but transcribing both voiced and devoiced sections of them would have made the transcription too unwieldy. Punctuation is included because intonation is not transcribed.

Phonemic

ðə 'nɔ:θ 'wind ən ðə 'sʌn wə dɪ'spju:tɪŋ 'wɪtʃ wəz ðə 'strɒŋgə, wen ə 'trævlə keɪm ə ɪnŋ 'ræptɪn ə 'wɔ:m 'kləʊk. ðeɪ ə'gri:d ðæt ðə 'wʌn hu 'fɜ:s sək'sɪ:dɪdɪn 'meɪkɪŋ ðə 'trævlə teɪk hɪz 'kləʊk ɒf ʃʊd bi kən'sɪdəd 'strɒŋgə ðən ði 'ʌðə. 'ðen ðə 'nɔ:θ wind 'blu: əz 'hɑ:d əz i 'kʊd, bət ðə 'mɔ: hi 'blu: ðə 'mɔ: 'kləʊsli dɪd ðə 'trævlə 'fəʊld hɪz 'kləʊk ə'raʊnd hɪm, ænd ət 'lɑ:st ðə 'nɔ:θ wind ɡeɪv 'ʌp ði ə'tempt. 'ðen ðə 'sʌn ʃʊn aʊt 'wɔ:mli, ænd ə'mɪ:diətli ðə 'trævlə 'tʊk ɒf ɪz 'kləʊk. ɪ ˌsəʊ ðə 'nɔ:θ 'wɪn wəz ə'blaɪdɪd tʊ kən'fes ðæt ðə 'sʌn wəz ðə 'strɒŋgɪ əv ðə 'tu:.

Allophonic

ðə 'nɔ:θ 'wɪnd ənŋə 'sʌn wə dɪ'spju:tɪŋ 'wɪtʃ wəz ðə 'strɒŋgə, wen ə 'trævlə k'heɪm ə ɪnŋ 'ræptɪn ə 'wɔ:m 'kləʊk^x. ðeɪ ə'gri:d ðæt ðə 'wʌn hu 'fɜ:s sək'sɪ:dɪdɪn 'meɪkɪŋ ðə 'trævlə t'heɪkɪs 'kləʊk ɒf ʃʊbbi k'hən'sɪdəd 'strɒŋgə ðənŋi 'ʌðə. 'ðenŋə 'nɔ:θ wɪnd 'blu: əz 'hɑ:d əs i 'kʊd, bət ðə 'mɔ: hi 'blu: ðə 'mɔ: 'kləʊsli dɪd ðə 'trævlə 'fəʊld hɪz 'kləʊk^h ə'raʊnd hɪm, ænd ət 'lɑ:st ðə 'nɔ:θ wɪnd ɡeɪv 'ʌp ði ə't'emɪt. 'ðenŋə 'sʌn ʃʊn aʊt 'wɔ:mli, ænd ə'mɪ:diətli ðə 'trævlə 'tʰʊk ɒf ɪz 'kləʊk^x. ɪ ˌsəʊ ðə 'nɔ:θ 'wɪn wəz ə'blaɪdɪz t'hə k'hən'fes ðæt ðə 'sʌn wəzəzə 'strɒŋgɪ əv ðə 't'hɪ:.

Orthographic

The North Wind and the Sun were disputing which was the stronger, when a traveller came along wrapped in a warm cloak. They agreed that the one who first succeeded in making the traveller take his cloak off should be considered stronger than the other. Then the North Wind blew as hard as he could, but the more he blew the more closely did the traveller fold his cloak around him, and at last the North Wind gave up the attempt. Then the Sun shone out warmly, and immediately the traveller took off his cloak. And so the North Wind was obliged to confess that the Sun was the stronger of the two.

References

- GIMSON, A. C. (1962). *An Introduction to the Pronunciation of English*. London: Arnold. [6th edn. in 2001, edited by A. Cruttenden.]

- JONES, D. (1917). *English Pronouncing Dictionary*. London: Dent. [14th edn. in 1977, edited by A. C. Gimson; 16th edn. in 2003, edited by P. Roach, J. Hartman & J. Setter (Cambridge University Press).]
- LADEFOGED, P. (2001). *A Course in Phonetics*. New York: Harcourt.
- PIKE, K. L. (1943). *Phonetics*. Ann Arbor, MI: University of Michigan Press.
- UPTON, C., KRETZSCHMAR, W. & KONOPKA, R. (2001). *Oxford Dictionary of Pronunciation for Current English*. Oxford: Oxford University Press.
- WELLS, J. C. (2000). *Longman Pronunciation Dictionary*. London: Longman.