

1 How do you say...?

The pronunciation of English words, including proper names

1.1 sloth

I saw an American TV programme recently based on the theme of the seven deadly sins. One of them is *sloth*, which the presenter pronounced **slɔːθ**, to rhyme with *cloth*. But this is not my pronunciation: I call it **sləʊθ**, to rhyme with *both* and *growth*.



The literal meaning of the name of both the deadly sin and the slow-moving animal named after it is ‘slowness’. The word is cognate with the adjective *slow*: as *warm* is to *warmth*, so *slow* is to *sloth*. There are some abstract nouns in *-th* that involve a vowel change when compared with the adjective they are based on: *deep* – *depth*, *wide* – *width*. But historically *sloth* was not one of them. Indeed, in earlier times it was sometimes spelt *slowth* or *sloath*. So the form rhyming with *cloth* must have originated as a spelling pronunciation based on the now current spelling, which according to the OED was first used in the sixteenth century. British dictionaries generally give only the pronunciation rhyming with *growth*. But in American English (AmE) the pronunciation rhyming with *cloth* appears to predominate.

1.2 nucular snitchel

The pronunciation of *nuclear* as if it were spelt *nucular* is well known and widely condemned as a mispronunciation. President Bush, for example, normally said 'nu:kjələ instead of the standard AmE 'nu:kliə. This (mis)pronunciation presumably arose through the influence of the large number of familiar words ending in -kjələ (British English [BrE] -kjələ or -kjələ), among them *circular*, *particular*, *spectacular*, *molecular*, *secular*, *perpendicular*, *jocular*.

Two other mispronunciations that one hears from time to time are *percolator* as 'pɜ:kjələrtə instead of 'pɜ:kələrtə and *escalator* as 'eskjələrtə instead of 'eskələrtə. In these, a yod has been inserted between **k** and **əl** under the influence of words such as *speculator* and perhaps even our own technical term *articulator*. And then we have *prenuptial* and *nuptials* pronounced -tʃuəl, as if spelt *prenuptual*, *nuptuals*. I have also heard people referring to a *defibrillator* as if it were spelt *defribulator*.

But my all-time favourite of this type of thing has to be the Wiener *schnitzel* masquerading as a 'sni:tʃəl, thereby solving at a stroke what are from the English point of view two phonotactic problems in the German-style pronunciation: initial **ʃn-** and morpheme-internal **-ts-**.

Dictionaries perform various roles. The main users of pronunciation dictionaries such as LPD are teachers or students of English as a foreign language (EFL) who wish to know how to pronounce a word whose written form they perhaps know well but about whose pronunciation they are uncertain. A second group of users are native speakers who want to have their ignorance remedied or their prejudices confirmed or disconfirmed by a lexicographer they consider authoritative.

The lexicographer, though (me, for instance), if he is trained in linguistics, will always be acutely conscious that there is no absolute standard of correctness in language, particularly in a language with such a wide speakership and inadequate orthography as English.

That is why I have always striven to make LPD play a further role: that of documenting the current actual state of the language. That is why I felt it essential to include widespread though 'incorrect' pronunciations.

However, if I had merely listed these variants along with all the other variants that warrant inclusion, I would be failing in my duty, because I would be ignoring the attitude of careful, educated, literary-inclined, speech-conscious users towards them. That's why I use a special warning mark: the exclamation mark enclosed in a triangle.

Not to include such variants would be to fail to document the language properly – a charge that I think could reasonably be levelled against certain other pronunciation dictionaries. To include them but without a distinguishing mark would be to fail to indicate the attitude of other speakers towards them.

1.3 artisanal

How do you pronounce *artisanal*? That was a question a correspondent asked me. At the time I had to reply that I did not know the word: it wasn't part of my vocabulary and I'd never heard anyone say it. However, despite the fact that *artisan* is **ˌɑːtɪˈzæn**, I thought the regular stress effect of the suffix *-al* ought to yield **ɑːˈtɪzənəl**. Adjectives in *-al*, of three or more syllables, have penultimate stress if the penultimate vowel is long (*archetypal*, *primeval*, *universal*) or followed by a consonant cluster (*dialectal*, *incidental*), but otherwise have antepenultimate stress (*personal*, *industrial*, *medicinal*). (There are one or two exceptions in which the penultimate vowel lengthens on adding *-al*, as *adjectival* - **ˈtɑɪvəl**.)

The word *artisanal* was not in the OED until added in 2008. It seems to have come into use quite recently as a kind of antonym of *industrial*, referring to small-scale production methods, agricultural or other.

Some would have us believe that it is **ˈɑːtɪzənəl** or **ˌɑːtɪˈzænəl**. I don't think so. That looks like someone's guess – someone who hasn't fully absorbed the English stress rules.

Early one morning I was jolted out of my half-sleep by hearing someone on my bedside radio, on the farming programme, use the word not once but twice. I can't remember what they were talking about, but my semi-conscious mind did note that the pronunciation used was indeed **ɑːˈtɪzənəl**. Result! And the OED now agrees.

1.4 C. *diff.*

The bacterium *Clostridium difficile* has been in the news recently. It is a cause of diarrhoea, usually acquired in hospital. Like all organisms it has a Latin binomial name, the first part referring to the genus and the second to the species. There's no difficulty about how to say the generic name: clearly, **klɒˈstrɪdiəm**. This can also be abbreviated to *C.*, and in this form is often pronounced just as **siː**.

It's the specific name that raises the interesting question. Apparently this species of *Clostridium* is called *difficile* because when it was first discovered it was difficult to grow in the laboratory. In Latin, *difficile* is the neuter singular nominative of the adjective *difficilis* 'difficult' – neuter because it agrees in gender with the neuter *Clostridium*. (With a masculine or feminine noun it would be *difficilis*.) In accordance with the usual rules for pronouncing Latin words in English, you would expect it to be **dɪˈfɪsɪləɪ** or perhaps **dɪˈfɪkɪləɪ**.

But it so happens that *difficile* is also the French word for 'difficult'. And in French it is pronounced **dɪfɪsil**, Anglicized as **ˌdɪfɪˈsiːl**. Because doctors and

other health care professionals these days are much more likely to know French than to know Latin, it is perhaps not surprising that the bacterium is generally known in English as **klɒˈstrɪdiəm ˌdɪfɪˈsi:l**.

Or *C. diff.* ˌsi: ˈdɪf for short.

1.5 shih tzu

Here's a dreadful pun: *I went to the zoo the other day, and all they had was a dog. It was a shih tzu.* Apologies for the indelicacy.

This breed of dog can be spelt either *shih tzu* or *shi tzu*. Either way, it is an old (Wade–Giles) romanization of the Chinese word for 'lion', which in Hanyu Pinyin would be written *shizi*, or with tones *shīzi*. (See <http://en.wikipedia.org/wiki/Wade-Giles> and http://en.wikipedia.org/wiki/Hanyu_Pinyin for further details.) In Chinese characters it is 狮子 (simplified), 獅子 (traditional). The Chinese pronunciation is **ʃi˥˥ tʂu˥˥**, said on a high level pitch (the first syllable has tone 1, the second is toneless).

This word is in LPD, where I gave its English pronunciation as **ʃɪt ˈtsuː**. On reflection, having heard dog fanciers talking about the breed, I think that the second syllable ought rather to be shown as **ˈzuː** or **ˈsuː**. (In English we can't really manage **ts** in syllable-initial position, so we transfer the plosive into the preceding syllable. The spelling *tz* normally corresponds to the pronunciation **ts**, as in *quartz*; but the presence of *z* in the spelling tends to lead people to use **z** when the plosive and fricative are separated, as here.)

Given that, the pun reduces to **ʃɪt ˈzuː** vs. – what? Well, actually **ʃɪt ˈzuː**. What stops the pun being perfect is intonation: the fact that in the latent version *zoo* has already been mentioned, so that on the second mention it will be deaccented, throwing the nucleus onto *shit*. The dog is a *shih tzu*, but the rubbishy zoo is in this context a *shit zoo*.

I must thank Kwan-hin Cheung for the choice of symbols in the transcription of Chinese here. In 1992 he published the following paper in Chinese: '北京話 '知' '資' 二韻國際音標寫法商榷 [IPA transcription of the so-called 'apical vowels' in Pekinese]', *Research on Chinese Linguistics in Hong Kong* (ed. T. Lee) (Linguistic Society of Hong Kong 1992, pp. 37–46). He comments, 'The collection itself was published bilingually, with a majority of papers in English. This paper predates John Laver's discussion on apical vowels in his *Phonetic Description of Voice Quality* (Cambridge University Press 1994, pp. 319–20).'

In his article Cheung mentions the symbols [ɿ] and [ʅ], which 'linguists within Chinese communities ... often regard ... as the IPA symbols for the two sounds', although in fact 'not only are they not official IPA symbols but they are hardly internationally known'. He would represent *shi* in IPA as [ʃɿ] and *si* as [sʅ]. The slur marks reflect the fact that friction may extend into the vowel.

1.6 omega

‘This food contains valuable omega-3 fatty acids,’ says the nutrition expert on The Food Programme on BBC Radio 4. But he pronounces it **əʊˈmiːgə**.

I’ve got this pronunciation of *omega* in LPD as an AmE variant, along with -‘**meɪgə** and -‘**megə**. But at the time I drafted the first edition of LPD, back in the late 1980s, I clearly assumed that the only British pronunciation was ‘**əʊmiɡə** (or -**mægə**).

Not any more! The traditional BrE pronunciation of the name of the last letter in the Greek alphabet is indeed ‘**əʊmiɡə**. Studying classical Greek from the age of twelve on until I got my BA in Classics at age twenty-one I do not think I ever heard any other way of pronouncing it. Yet given that the classical Greek name was just ω (\bar{o}), this could be seen as a little surprising. Why don’t we simply call the Greek letter **əʊ**? The name *omega* dates from the Byzantine period. By then the distinction between classical short ω and long \bar{o} had been lost, and the two letters \omicron (\bar{o}) and ω (\bar{o}) had to be distinguished by special names: ‘small \omicron ’ *o micron* and ‘big ω ’ *o mega*.

Modern Welsh does something similar with the letters *i* and *u*. In southern Welsh they stand for the same vowel sound (*ci* ‘dog’ rhymes exactly with *tu* ‘side’), so when I was learning Welsh I was taught to call the first one *i dot* and the second *u bedol* (i horseshoe).

OK, the medieval name of the Greek letter was ω $\mu\acute{\epsilon}\gamma\alpha$ (\bar{o} *mega*). That still doesn’t explain why in English we traditionally stress it on the first syllable. Indeed, given that the element *mega* is in implicit contrast with *micron*, you might think that the syllable -*meg*- would bear fossilized contrastive stress. As usual, the reason that it doesn’t is to be found in Latin. Most Greek words in English have passed to us via Latin, and on the way have become subject to the Latin stress rule (see also 1.3 above and 2.12 below). The Latin rule says look at the penultimate syllable: if it is light (= short vowel, not more than one following consonant), then stress the antepenultimate. In the *mega* element the *e* is indeed short (cf. *megabyte*, *mega store*, *megalomaniac*) and it is followed by the single consonant *g*. So stress goes onto the preceding syllable. The *e* itself undergoes vowel reduction. Hey presto: ‘**əʊmiɡə**.

1.7 plethora

How do you pronounce *plethora*?

Not everyone has this word in their vocabulary. The OED defines it as first having a medical meaning – a morbid condition – and then a figurative one, ‘over-fullness in any respect, superabundance; any unhealthy repletion or excess’. *The Longman Dictionary of Contemporary English* (Longman 2009, fifth edition; hereafter LDOCE), more in tune with current usage, defines it as ‘a very large number of something, usually more than you need: *a plethora of suggestions*’.

I pronounce it **'pleθərə**. That's also what's recommended by every dictionary I have to hand, and what I usually hear. I have also, however, heard **plɪ'θɜ:rə**, which you will find in LPD and some other dictionaries as an alternative.

Ha!, I thought to myself, this must be another of those words like *omega*, a Greek/Latin word with a short penultimate vowel followed by a single consonant: the Latin stress rule (see 2.12 below) therefore gives us antepenultimate stress. People ignorant of the quantity of the *o* vowel mistakenly suppose it to be long, and consequently accent it.

But I was wrong. It turns out that our modern word comes from a medieval Latin *plēthōra*, from a Greek word *πληθώρα* (*plēthōrē*). The *o* in the penultimate syllable was definitely long. The OED reports that the 1731 edition of Nathan Bailey's Dictionary had 'the etymological pronunciation' *ple'thōra*, but that the 1742 and 1755 editions of the same book have *'plethora*. And that's what most of us say today, despite the etymology.

1.8 diocese

The area that an Anglican or Roman Catholic bishop controls is a *diocese*. How do we pronounce this word? More trickily, what is its plural? In England there are a number of *dioceses*. How do we say that? And the derived adjective is *diocesan*, as in *Diocesan Board of Finance*. Where does the stress go?

If you had asked me that when I was aged nine or ten, I think I would have been able to reply confidently: **'daɪsɪs**, **'daɪsɪ:z**, **daɪ'nsɪzn**. But that's because my father was a vicar, and these were everyday words in our house. As far as I know, those were also the pronunciations used by everyone else concerned with ecclesiastical administration, at least in our diocese (Liverpool). It means that *diocese* must have been attracted in its phonetics to the singular-plural alternation we see in such Greek-via-Latin-derived words as *crisis*, *thesis*, *basis*, *oasis*, *emphasis*, *neurosis*, *ellipsis*, *analysis* and *axis*. (Compare also *this* and *these*.)

	singular		plural
<i>crisis</i>	'kraɪsɪs	<i>crises</i>	'kraɪsɪ:z
<i>thesis</i>	'θi:ɪsɪs	<i>theses</i>	'θi:ɪsɪ:z
<i>diocese</i>	'daɪsɪs	<i>dioceses</i>	'daɪsɪ:z

However, this obviously produces a mismatch between sound and spelling. More to the point, most people are not from clergy families and may have no experience of ecclesiastical terminology. Not surprisingly, they tend to pronounce both *diocese* and *dioceses* the way they are spelt, with a regular sibilant-stem plural, making *diocese* perhaps **'daɪsɪ:s** or even **'daɪsɪ:z** and for *dioceses* appending the usual additional-syllable ending **ɪz** (or **əz**).

Jack Windsor Lewis points out that this word was spelt *diocess* from the sixteenth century to the end of the eighteenth; this was the only form

recognized by Dr Johnson and the other eighteenth-century lexicographers, and was retained by some (notably by *The Times* newspaper) into the nineteenth century. During that century, however, *diocese* (as in French) became the established spelling.

The general issue I faced when compiling LPD was: do I prioritize the pronunciation used by those who use the word regularly, day in day out? Or do I heed those speakers for whom the word is a relatively unfamiliar written term? I chose the first, namely **'daɪəsɪs**; other dictionaries, in this case at least, tend to choose the second, **-sɪz** or **-sɪz**.

There's a similar problem with *cervical*, where medical specialists generally use penultimate stress and a long *i* but the general public uses initial stress and a short *i*.

1.9 West Indian islands

Thoughts on returning from abroad: if you fly with Virgin Atlantic or British Airways from London to *Antigua* in the West Indies, you find everyone from passengers to check-in staff to the aircraft captain pronouncing it **æ'n'ti:gə**, just as the locals do. But if you fly there from San Juan or New York on an American airline, as likely as not it will be referred to as **ɑ:n'ti:gwə**, with a Spanish pronunciation (sort of). Probably this is because Americans are much more likely than Brits to have learnt some Spanish at school, and of course the name does look Spanish and was originally Spanish, so that those unfamiliar with the name tend to assume it *is* Spanish, despite the island's three and a half centuries of speaking nothing but English. The Brits, on the other hand, mostly have French, rather than Spanish, as their main foreign language (if indeed they know a foreign language at all), and perhaps anyhow retain some ancestral memory of how to pronounce the names of former colonies.

Why, though, do we pronounce this Spanish-derived name without the **w** it has in Spanish? For the town of Antigua in Spanish-speaking Guatemala we indeed retain it. I don't know the answer. As a Spanish word *antigua* is **an'tiywa** (it is the feminine form of the adjective *antiguo* **an'tiywo**). We preserve this **w** when we anglicize the names *Paraguay*, *Uruguay* and *Guatemala*: so why do we treat *Antigua* differently? I don't know why, but we do.

Antigua's twin island is Barbuda, pronounced – by those who are familiar with it – as **ba:'bjudə**. It tends to crop up mainly in the official name of the country, *Antigua and Barbuda*. Non-locals therefore often need to pronounce this name, and they do tend to say it without its **j** as **ba:'budə**, which is what it would be if it were Spanish-speaking and not just Spanish-derived.

You see something similar to *Antigua* with *jaguar*. In the UK we anglicize it very thoroughly as **'dʒæɡjuə**, while the Americans are more likely to go for a Spanish-style **'dʒæɡwaɪr**. In *Nicaragua*, however, do I detect a slight

movement in Britain away from the anglicized **ˌnɪkəˈræɡjuə** in the direction of **-ˈrɑːgwə**, as in AmE?

Another West Indian island with a tricky pronunciation is *Dominica*. Locally, and locals would insist correctly, it is stressed on the penultimate: **ˌdɒmɪˈniːkə**. Those not familiar with the island tend to assume that it has antepenultimate stress, **dəˈmɪnɪkə**. Fans of Chomsky and Halle's *Sound Pattern of English* (Harper & Row 1968) will see that the disagreement relates to differing views about the length of the penultimate vowel: if long, then stress-attracting, but if short – as in *Monica*, *Veronica*, *America*, *angelica* and many others – then stress-repelling. In the case of *Dominica* matters are made worse by the existence of a similarly named but quite distinct West Indian country, the Dominican Republic. Its name is normally pronounced with the expected antepenultimate stress. (*Dominica*, also known as the Commonwealth of *Dominica*, is the island located between Guadeloupe and Martinique in the eastern Caribbean. Its official language is English, the unofficial language French Creole. The Dominican Republic, aka *Santo Domingo* or *Dominicana*, occupies the eastern half of the island of Hispaniola, the western half being Haiti. Its language is Spanish.)

1.10 place names

Driving to Gatwick Airport a few days ago to meet an arriving passenger, I passed through the village of *Burgh Heath*. As on previous occasions when I have travelled that route, I wondered idly how it's pronounced. Is the first word **bɜː** or **ˈbʌrə**?

When I got back home I looked it up in the *BBC Pronouncing Dictionary of British Names* (Oxford University Press 1990, second edition), which says it can be either. Just not **bɜːg**.

I further learnt that *Burgh* in Norfolk is **ˈbʌrə**, but *Burgh* in adjoining Suffolk is **bɜːg**. Things are different in the north of England: *Burgh-by-Sands* in Cumbria is metathesized to **brʌf**, which must mean that for many locals it's more like **[brʊf]**.

It's worse than *-ough*.

One day I wanted to go to a funeral on the outskirts of Birmingham. To reach my destination the map said I had to look for the road leading to *Alcester*. Er... what was that? I checked with my brother, who lives not too far away, and he said it was **ˈɔːlstə**. Then I looked in LPD and found that I agree.

And there's no call for Americans to feel superior to the wacky British. In the States you never know what will happen with Spanish names. I remember passing through *Salida*, Colorado. That's the Spanish for 'exit', and it was at the mouth of a canyon, so I thought that in English it would be **səˈliːdə**. But the local radio station announcers, who should know, pronounced it **səˈlaɪdə**.

Even English-derived names can be surprising. I remember driving through *Placerville*, California, and discovering to my surprise that it was not **'plɛɪsəvɪl** but **'plæsəvɪl**.

My correspondent John Cowan points out that **'plæsə** is the standard AmE pronunciation of *placer*, a deposit of gravel which contains particles of gold or other precious minerals. *Placer mining* is the kind you do with a pan in which you wash the gravel, letting it flow over the lip of the pan along with most of the water, leaving the heavier gold behind. Despite appearances, the word is derived from Spanish *placer*, which may be of Portuguese or Catalan origin. Another correspondent, William Pitfield, adds that Burgh Heath gave its name to the local telephone exchange, covering a significantly larger area, including the place where he grew up. He doesn't recall any of the locals calling it anything other than **'bʌrə** *Heath*, at least among themselves. The exception was in the days before Subscriber Trunk Dialling, when it was wise to ask the operator for **bɜ:** *Heath*, as otherwise they were liable to look for it under *Borough* and then assure you that the exchange didn't exist.

1.11 Madejski

Any visitor to Reading, on the way from London to Bristol, will be struck by the road signs everywhere pointing to the Madejski stadium. Why does the stadium have a Polish name? And how do people pronounce it?

Reading Football Club's stadium is named after John Madejski, one of the wealthiest people in the UK, the club's benefactor. He was born in Britain, with an English mother, but his father was Polish, an airman who was stationed here during the Second World War. As a Polish name, *Madejski* is pronounced exactly as spelt. The Polish diphthong written *ej* is similar to the English vowel in *face*, so the name would anglicize very easily as **mə'dɛɪski**. However, English people do not know much about the pronunciation of foreign languages, except to some extent French. So they do not expect the letter *j* to stand for anything other than **dʒ** or (in foreign words) **ʒ**. And the resultant ***mə'dɛʒski** would violate English phonotactics and thus be impossible. In fact, I am told that the usual pronunciation of the name in Reading is **mə'dʒɛski** or **mə'dʒɛdski**, as if spelt *Madjeski* or *Majedski*. The metathesis makes it pronounceable.

1.12 Charon

A few years ago astronomers were discussing just what a planet is and exactly which heavenly bodies qualify as pukka planets. One of the candidates for planetary status is Charon, Pluto's moon. One TV news presenter called it **'ʃɛərən**. It is, however, traditionally pronounced in English with **k**, just as are

other Greek words and names spelt *ch-*: *Charybdis*, *chemistry*, *chiropractic*, *chlamydia*, *Chloë*, *chlorine*, *chlorophyll*, *cholesterol*, *choreography*, *Christ*, *chromosome*, *chronograph* and also (usually) *chimera*, *chiropody*. Pronunciation dictionaries are unanimous for initial **k** in *Charon*: **'kæɹən**, **-ɒn**.

Long before his name was given to the moon of Pluto, Charon was a figure in classical mythology: the grim ferryman who ferried dead souls across the river Acheron (or Styx) in Hades. I well remember Charon from Virgil's *Aeneid*, book six, line 299. As a teenager I was required to learn off by heart 20–30 lines of it:

*portitor has horrendus aquas et flumina seruat
terribili squalore Charon, cui plurima mento
canities inculta iacet, stant lumina flamma,
sordidus ex umeris nodo dependet amictus.*

'A frightful ferryman serves these waters and streams, in terrible filth, Charon. On his chin is an unkempt white beard, his eyes stand out with flame, a dirty tunic hangs knotted from his shoulders.'

Virgil's hexameters make it clear that the first vowel in *Charon* was actually short in Greek and Latin:

terrībīlī squāllōrē Chālrōn, cūī | plūrīmā | mentō

The Greek form, Χάρων, makes it clear that his second vowel was long. So perhaps we ought to be saying **'kæɹən** (or even **'kæɹəɒn**) rather than **'kæɹən**.

I found most of the Latin literature I studied at school and university pretty uninspiring. But the sixth book of the *Aeneid* was an exception. Virgil's lines about the lost souls yearning to be transported across the river really have that tingle factor:

*huc omnis turba ad ripas effusa ruebat,
matres atque uiri defunctaque corpora uita
magnanimum heroum, pueri inuuptaeque puellae,
impositique rogis iuuenes ante ora parentum:
quam multa in siluis autumnī frigore primo
lapsa cadunt folia, aut ad terram gurgite ab alto
quam multae glomerantur aues, ubi frigidus annus
trans pontum fugat et terris immittit apricis.
stabant orantes primi transmittere cursum
tendebantque manus ripae ulterioris amore.*

'To this place there rushed a whole crowd, pouring out onto the river bank: mothers and men and the lifeless bodies of great-hearted heroes, boys and unwed girls, young men placed on the funeral pyre before their parents: as many as the leaves that fall in the woods at the first frost of autumn, as many as the birds that flock together onshore, leaving the deep abyss, when the cold weather drives them across the sea to warmer lands. They were standing there, praying to be carried across first, and holding out their hands in yearning for the opposite bank.'