

Letter-to-sound rules

Dearest *creature in creation*,
Studying English pronunciation,
I will teach you in my verse
Sounds like *corpse, corps, horse and worse*.

[...]

Finally: which rhymes with “enough”,
Though, through, plough, cough, hough, or tough?
Hiccough has the sound of “cup”...
My advice is – give it up!

These lines are the first and last stanzas of the famous poem “The Chaos”, which was written by Gerard Nolst Trenité with the aim of drawing attention to the irregularities of English letter-to-sound correspondences. While such poems are indeed entertaining, there is a danger that they discourage learners as they create the false impression that English spelling is totally idiosyncratic and therefore unlearnable. However, the truth is that the majority of English words **are** pronounced according to a set of predictable (and learnable) rules, and the regular examples vastly outnumber the irregular ones that poems like “The Chaos” enjoy pinpointing. So, instead of the message conveyed by the poem above, our advice is: **don't** give it up, but rely on your knowledge or intuitions of the regularities, and it is enough to memorise the words that are irregular.

Many will have heard the joke that the English are weird because they spell “Manchester” but pronounce it “livöpúl” (if you are not familiar with the story, you may learn about it [*here*](#)). This funny example has been widely cited among teachers and learners of the English language to make a mockery of the apparent idiosyncrasies of English spelling.

On the one hand, it is understandable why English spelling seems illogical and therefore unpredictable to learners – let us list a few aspects that have contributed to this view, focusing on the most significant differences between English and Hungarian spelling.

- While Hungarian letter-to-sound correspondences are dominated by the phonemic principle (according to which one letter corresponds to one sound, and one sound is denoted by one letter), this is not true of English spelling rules. In English, the same sound may be denoted by several different letters: e.g., [f] may be spelt <f>, <ph> or <gh>, as in *half, graph* and *laugh*, respectively; *meet, meat* and *mete* are all pronounced the same; etc. The reverse is also true: the same letter may represent several different sounds, e.g., <a> is pronounced in four different ways in *cat, car, came*, and *care*; the letters <ough> may be pronounced in eight different ways; etc. (the latter is a quite famous example; sentences have been compiled that include all eight pronunciations of the letter combination: *A rough-coated, dough-faced, ploughman strode through the streets of Scarborough, coughing and hiccoughing thoughtfully*).

- Hungarian and English use consonant doubling for two totally different purposes. In Hungarian, a double consonant letter signals a long consonant sound – e.g., *halom* ‘pile’ and *hallom* ‘I can hear it’ differ in that the former contains a short [l] and the latter contains a long one ([l:]). In English there are no long consonants¹: the words *diner* and *dinner* both contain a short [n]; what is different is the *vowel sound* before the [n], so the role of consonant doubling in English is to indicate what the preceding vowel can be.
- English makes an extensive use of silent letters. There is only one letter in the English alphabet that is never silent (<v>); every other letter may occur in words where it is spelt but not pronounced, e.g.: *climb*, *hour*, *knife*, *talk*, *autumn*, *psychology*, *castle*, *write*, etc. Unpronounced letters are much less frequent in Hungarian, and even so they are either dialectal (e.g., some speakers pronounce the [h] in words like *méh* and *düh*, while others do not) or the deletion of a sound happens in a specific case, namely when three consonants would stand next to one another as a result of suffixation or compounding (in such cases one of the three consonants, usually the middle one gets deleted, e.g., the <d> is pronounced in *mond* ‘say’ and *rend* ‘order’, but not in *mondta* ‘he said it’ and *rendtartás* ‘keeping of order’).

On the other hand, however, the truth is that the majority of English words have a regular pronunciation – the irregularities may seem frightening, especially if they are presented in a collection, but they are not so large in number that it would make the regularities pointless to know/learn.

It is important to point out that a learner of the language does not need to be familiar with all of the regularities because the majority of them can be acquired implicitly – in other words, although it may seem difficult at first, it is perfectly possible to develop the intuitions necessary to guess how English words are pronounced. What is crucial though is that *teachers* should be familiar with the regularities so that they will be able to *anticipate* what irregular words the students are likely to mispronounce, and they should *prevent* the students from memorising such words in an incorrectly pronounced form by drilling the correct pronunciations before the students would make an attempt at pronouncing them based on their intuitions.

With this manual, our purpose is not to present a comprehensive overview of the letter-to-sound rules of English (this is provided by various textbooks – we recommend Chapters 11-

¹ In fact, they do occur, but only in very limited environments (e.g., on certain types of prefix+root / root+suffix boundaries, such as in *unnatural* and *keenness*).

12 in *The Pronunciation of English* by Katalin Balogné Bérces and Szilárd Szentgyörgyi, downloadable from [*here*](#)). What we will do is pinpoint examples of letter-to-sound rules and exceptions that are either often highlighted in coursebooks and/or that are likely to cause problems for Hungarian learners.

Before we turn to a discussion of concrete examples, it needs to be clarified what exactly makes a letter-to-sound correspondence “regular”, because the explanations we will be providing will contain references to “regular” and “irregular” pronunciations. Let us consider the example *live* as a verb [lɪv] and as an adjective [laɪv]. Obviously, as the spelling of the two words is the same, only one pronunciation can be regular. Which one is it and how do we know?

What the regular pronunciation of a given letter (or combination of letters) will be is determined by two factors: (1) frequency and (2) native speaker intuitions.

(1) If we list all the words having the same spelling pattern as *live* (i.e., having a stressed vowel spelt with an <i>, and followed by a consonant letter and a vowel letter), we will find that the ones pronounced with [aɪ] (e.g., *bite, file, shine, time*, etc.) vastly outnumber the ones pronounced with [ɪ] (e.g., *give*), therefore the former is to be considered the regular one.

(2) Native speaker intuitions best manifest themselves in the case of spelling errors, re-spellings (e.g., in graffiti), nonsense words and lesser-known words such as neologisms (including brand names) or proper nouns (especially surnames and geographical names). If a native speaker is asked how they would pronounce non-existent words having the pattern in question (e.g., *hine, pibe, tive*, etc.), they will certainly say that the vowel of such words is [aɪ].

Consequently, out of the two pronunciations of the word *live*, it is the adjective whose pronunciation is the regular one.

Some issues learners (especially beginners) might be puzzled by

a) Letter-to-sound rules for vowels – Why do *cat, car, came, and care* have different vowels if they are all spelt with an <a>?

Beginner learners of English who are used to more transparent spelling systems (such as Hungarian learners) will encounter confusing issues concerning English spelling already at the very beginning of their language learning. The fact that in English the same letter may denote several different sounds can be illustrated the best by the use of vowel letters, so we will disregard consonants here.

What a Hungarian learner is used to in terms of vowel letters is that wherever they see for example a letter <a> spelt in Hungarian words, it is always pronounced the same (words like

Facebook and *hashtag* of course do not count). This is absolutely not the case in English, as <a> is pronounced in four different ways in the words *cat*, *car*, *came*, and *care* (which are only the regular pronunciations of the letter, and we have not listed the irregular ones). Although learners tend to develop an intuition about the pronunciation of vowel letters and get over such initially confusing issues relatively quickly, it is still beneficial for teachers to have some conscious knowledge of the spelling regularities in order to be able to aid their students if they have queries about the confusing aspects of English letter-to-sound correspondences.

Let us see the main difference between Hungarian and English in terms of the use of vowel letters in more detail. Hungarian uses 14 different letters (distinguished by the use of diacritical marks) to denote its 14 vowel phonemes, with one letter corresponding to one sound, while English only uses <a>, <e>, <i>², <o> and <u> to denote at least³ 17 vowel sounds⁴. How is this possible without using diacritics on top of the vowel letters? The strategy used by English spelling is that it creates four sound values belonging to each vowel letter based on what letters *follow* the vowel letter in question. The four options are created by combining two binary criteria: (1) whether the consonant letter following the vowel letter is a <r> or any other consonant; (2) whether there is a silent letter <e> at the end of the word. (To keep the explanation simple, we only consider words consisting of a consonant sound, a vowel sound, and another consonant sound in this order.) The table below illustrates this with example words.

	(1) <r>: ✗ silent <e>: ✓	(2) <r>: ✓ silent <e>: ✓	(3) <r>: ✗ silent <e>: ✗	(4) <r>: ✓ silent <e>: ✗
<a>	<i>came</i>	<i>care</i>	<i>cat</i>	<i>car</i>
<e>	<i>gene</i>	<i>here</i>	<i>pet</i>	<i>her</i>
<i>	<i>fine</i>	<i>fire</i>	<i>sit</i>	<i>sir</i>
<o>	<i>home</i>	<i>core</i>	<i>dog</i>	<i>nor</i>
<u>	<i>mute</i>	<i>cure</i>	<i>fun</i>	<i>fur</i>

² <y> is an alternative to <i> (words like *gym*, *my*, *rhyme*, *lyre*, *myrrh*, etc. could be spelt with an <i> too), but we ignore this here for the sake of simplicity.

³ There are actually more than 17, but we ignore the schwa here (which can be spelt with any of the vowel letters) as well as the few vowels that can only be spelt with a digraph, and the fact that /u/ and /ju:/ (as in *rule* vs. *mute*) can be analysed as two separate vowels.

⁴ Here we did not count the use of vowel digraphs such as <ee> in *see* or *keen* – the majority of the digraphs denote the same vowels that can be spelt with single vowel letters, anyway.

The system is of course more complex than how it is presented here, but this simplified version is suitable for our purposes (which are to shed light on the regularities governing how each vowel letter has multiple regular pronunciations).

b) The “strange” names of vowel letters of the alphabet – Why is the letter <e> pronounced [i:] and <i> as [aɪ]? Why is [i:] not <i>?

The alphabet is among the first things that a beginner learner of English (or in fact any language) will learn in the new language. The English alphabet may be of a surprise to learners, especially in terms of the names of the vowel letters (<a>, <e>, <i>, <o>, <u> = “aye”, “ee”, “eye”, “oh”, “you”) – in many other languages, “ee” is the name of <i>, but in English, this is the name of the letter <e>. These “strange” names come from the fact that the alphabetical names of the vowel letters of English are the same as the vowels they represent in column (1) in the table above⁵. This is again something that learners tend to get used to fairly soon, but the type of learner who questions everything might want to hear this explanation.

c) Doubling consonants when adding suffixes – Why do we have to double the root-final consonant in *stopped* and *omitted*, but not in *developed* and *vomited*?

When introducing the Present Continuous and Past Simple tenses, coursebooks warn students about some spelling peculiarities concerning the *-ing* and the past tense form of (regular) verbs, such as the fact that in words like *stop*, the word-final letter <p> needs to be doubled if the *-ing* or the *-ed* ending is attached to the root.

There are two problematic issues regarding this. First, as a more accurate presentation of the rule would be too complex for the learners (considering the fact that Present Continuous and Past Simple are both taught at elementary level), coursebooks usually say that consonant doubling needs to be applied in the case of one-syllable verbs whose last two letters are a vowel and a consonant. An elementary learner obviously needs no more information at this level, but for teachers it is worth bearing in mind that this is a simplified version of the rule, because what determines whether the last consonant is to be doubled or not is not the number of syllables the verb consists of, but *stress*: the final consonant is doubled if the last syllable of the verb is stressed. This is why *omitted* (which is stressed on the second syllable) is spelt with <tt>, but *vomited* (stressed on the first syllable) is with <t>.

⁵ The name of the letter <u> contains a [j], which can be analysed as being part of the vowel of words like *mute* and *cure*.

Second, coursebooks do not tend to point out *why* the doubling is necessary, though this might facilitate the learners' remembering when to double consonants. Teachers are therefore advised to point out when teaching this spelling regularity that *without* doubling the word-final consonants, the vowel of the verbs would be different. E.g., *rapping* spelt with a single <p> happens to be an existing word, so it is easy to explain why a difference needs to be made (there is a HUGE difference between a *rapping* clown and a *raping* clown, isn't there? 😊 – cf. [*this story*](#)), but even if the wrongly spelt form of a word does not exist as a separate word, it *would* be pronounced differently, e.g. **stoped* spelt with a single <p> would rhyme with *hoped* (and not *hopped*), **omited* would be pronounced [ə'maɪtɪd], etc.

Some famous Hunglish pronunciation problems related to letter-to-sound rules

a) “Vircsuöl rieliti” – The lack of application of R-influence

As we have already seen above, whether a vowel letter is followed by <r> or some other consonant is one of the two factors based on which each vowel letter has four sound values. A letter <r> (even if it is not pronounced in certain pronunciation varieties!) is thus able to change how the preceding vowel is to be pronounced (*cab*, *can*, *cap*, and *cat* will all have the same vowel, but *car* will have a different one). This is called R-influence on vowels.

The application of R-influence is not usually problematic in the case of frequent words – a beginner learner may be initially puzzled by the fact that words like *cat* and *car*, *sit* and *sir*, etc. have different vowels although they are spelt with the same vowel letter, but experience shows that learners quickly get used to this spelling regularity. The problem rather affects words that also exist in Hungarian as loanwords (e.g., *virtual*), proper nouns (e.g., *Mercury* and *Sherlock*), and less frequent words (e.g., *stir*). Hungarian speakers (even higher level learners!) often pronounce such examples as “vircsuöl”, “merkjuri”, “serlokk” and “sztir”, although the stressed vowel of *all* of these words is [ɜ:].

b) “Hó máccs iz it?” – The pronunciation of the digraph <ow> (and <ou>)

<ow> (as well as its variant <ou>) is regularly pronounced [aʊ]. This means that words like *round* and *cow* are regular, while *soul* and *know* are irregular. This may be problematic for non-native speakers for at least two reasons: first, some words in which <ou>/<ow> is pronounced irregularly are quite frequent words (such as *know*), therefore it might not be obvious which pronunciation is the regular one, which may lead to the development of faulty intuitions. Second, the presence of a letter <o> as part of the digraph may also influence how a non-native

speaker will pronounce a word containing the digraph: they are likely to have the faulty intuition that the regular pronunciation of <ow> is [əʊ], which is one of the regular sound values of the letter <o>.

Notice how native speaker intuitions work completely differently from what Hungarians may falsely deduce about English spelling regularities: native speakers of English who do not know how J. K. Rowling pronounces her surname very often pronounce *Rowling* as ['raʊlɪŋ] (instead of ['rəʊlɪŋ], which is an irregular pronunciation), as they will rely on their intuitions, according to which <ow> is to be pronounced [aʊ]. In other words, native speakers might mispronounce irregular proper nouns and apply the regular pronunciation, while Hungarians will do the exact opposite: influenced by the letter <o> in the digraph, they are likely to pronounce even regular words irregularly.

c) “Kattints a ‘launcs míting’-re”, “nyomd meg a ‘pauzé’-t” – The pronunciation of the digraph <au> (and <aw>)

Stressed vowels denoted by the digraph <au> (a variant of which is <aw>), such as in *launch* and *pause*, constitute a potential problem for Hungarians, who are likely to falsely think that this digraph denotes two separate vowels. The truth is that <au> and <aw> denote one vowel (namely [ɔ:]). Words containing this digraph may thus be pronounced by Hungarians one syllable longer than they actually are. Words also existing in Hungarian (names such as *Laura*, or the prefix *audio-*) have a higher chance of being mispronounced with an extra syllable due to the misleading effect of the pronunciation of the Hungarian equivalents. The Hungarian intuition manifests itself in how certain proper nouns are pronounced in dubbed films – e.g., Jack Dawson from the movie *Titanic*, whose surname is pronounced ['dɔ:sən], is called “dzsek dauszon” in the dubbed version.

Some notes on the “find the odd one out” exercise

As it has been pointed out, it would be impossible (nor is it our aim) to provide an exhaustive overview of all the letter-to-sound rules of English, therefore in what follows we only present a miscellaneous collection of regularities that are often pointed out in language coursebooks. The ten rules presented below explain the spelling rules that the “find the odd one out” exercise on the website calls the learners’ attention to, in the same order as the issues appear in the exercise.

1. A letter <h> at the beginning of a word is silent in four English words: *hour*, *honest*, *honour*, *heir* (and their derivatives, e.g., *dishonest*, *heirloom*, etc.).
2. In the standard accents of English, <wh> is pronounced [w]; however, in a few words such as *whole*, *who*, *whom*, etc., it is pronounced [h].
3. The letter <j> is always pronounced [dʒ]. If we disregard some words of foreign origin such as *hallelujah* (in which the <j> is pronounced [j]), we can basically say that this rule has no exceptions.
4. <ch> is most often pronounced [tʃ], but in quite a few words there might be two further alternatives: it may be pronounced [ʃ] in words of French origin (*machine*, *moustache*, *champagne*, *chauffeur*, etc.), and [k] in words of Latin and Greek origin (*chaos*, *chameleon*, *character*, *chemical*, *choir*, *echo*, etc.).
5. At the beginning of words, <s> is regularly pronounced [s]. There are two words that are exceptions: *sure* and *sugar* (and their derivatives such as *assurance* and *sugary*).
6. <ow> is regularly pronounced [aʊ], but it is irregularly pronounced [əʊ] in quite a few words. This is described above in more detail.
7. If a letter <i> is followed by a consonant and a vowel letter, it is regularly pronounced [aɪ] (as opposed to [ɪ] when it is followed by two consonants or a consonant and nothing else), though there are some exceptions (e.g., *city*, *give*, *river*, *wizard*, etc.). This is a regularity that learners do not usually have problems with – even without explicitly learning about the spelling regularity, they will develop the intuition fairly soon that the vowel of words like *bite*, *fine*, *like*, *time*, etc. is [aɪ], while that of *big*, *sit*, *fidget*, *kitty* is [ɪ].
8. <au> is pronounced [ɔː]. This spelling rule has very few exceptions, e.g., *sausage* (pronounced with [ɒ] or [ɑ], depending on the pronunciation variety) or *laugh* ([ɑː] or [æ], again depending on the pronunciation variety).
9. An <o> followed by an <r> is regularly pronounced [ɔː], though there are some exceptions, such as words like *work*, *world*, *worm*, in which it is pronounced [ɜː].
10. An <i> followed by an <r> and another consonant letter (as in *girl*) is pronounced [ɜː]. (This also applies if there is no letter after the <r>, as in *sir*.) There are no exceptions to this spelling rule.