

Rules and Ruliness in English Spelling: A Précis

In spite of all the unkind things said over the centuries about its alleged disorder, perversity, and chaos, English spelling is actually ruled and ruly. And thus there is a subject matter there, useful things that can be taught by our English and language arts teachers and learned by their students. What follows summarizes the most basic of those rules and that ruliness, delaying for now complications and details. (For more on these rules see the relevant sections of the [Compendium of English Orthography](#).)

1, Vowels and consonants, V's and C's. Letters are vowel letters when they spell vowel sounds. They are consonant letters when they spell consonant sounds. Thus, <w> and <y> are consonant letters when they spell the consonant sounds [w] and [y], as in *wet* and *yet*. But when they appear in vowel digraphs, they are vowel letters, as in *craw*, *crew*, *crow*, and *crowd*, and in *tray*, *trey*, and *Troy*. And when <y> spells long or short <i>, it is a vowel, as in *typical* and *type*. The letter <u> is usually a vowel, but it is a consonant letter when it spells the consonant sound [w], as in *language*, *suede*, and *pueblo*. It is a consonant letter also whenever it follows the letter <q>, whether it spells [w] or not, as in *queen* and *mosquito*. In the following discussion vowel letters are represented with a “V”, consonant letters with a “C” – as in *cat* = CVC.

2. Tactical Patterns and Rules. Tactical rules describe expected and ruly sequences of V's and C's. The two best known and most extensive are VCV and VCC, each with a stressed first vowel. In a VCV string the first vowel will regularly be long, as in *biter* or *plated*. In a VCC string it will regularly be short, as in *bitter* or *planted*.

Three less extensive tactical rules are the following:

In a CVC# string, where “#” indicates the end of the word, the final vowel, if stressed, is short – as in *bit*, *sedan*, and *thermostat*.

In CVe#, where “e” indicates a silent final <e>, the preceding vowel, if stressed, is long – as in *pie* or *pursue*.

In V.V, where the period indicates a syllable boundary, the first

vowel, whether stressed or not, will be long – as in **duet** and **lion**.

3. Shortening Rules. The VCV rule can be preempted by certain more specific and local rules that account for short first vowels in VCV strings:

French Lemon Rule. Words adapted from French will often have short vowels in VCV strings – as in **lemon** from French vs. *demon* from Latin, and **model** from French vs. *yodel* from German.

Third Syllable Rule. If the vowel is in the third or fourth syllable from the end of the word and is stressed, it will be short even if in a VCV string – as in **national** vs *nation*, **penalty** vs. *penal* – and **holiday** vs. *holy day*. This rule also covers the hundreds of words that contain the suffix *-ity*) – as in *sane* vs. **sanity** and *extreme* vs. **extremity**.

The Suffix *-ic* Rule. Other suffixes than *-ity*) can create VCV strings with short first vowels. The most common is the suffix *-ic*). For instance, compare *athlete* and **athletic**, or *state* and **static**, or *telephone* and **telephonic**.

4. Procedural Rules. Procedural rules describe procedures that are regularly followed when word prefixes, bases, and suffixes combine.

Simple Addition. Far and away the most common and general procedure is Simple Addition, in which word elements combine by simply adding together – as in **unlovely** (un+love+ly) and **reseeded** (re+seed+ed). The Rule of Simple Addition states that “Unless you know of some change that should be made, simply add the word elements together.”

Three Changes. The three changes that can preempt Simple Addition are **insertion**, **deletion**, and **replacement**. In insertion a letter is inserted between the word parts – as in the **twinning** of final consonants – for instance, *swim+ing*) = *swim+m+ing*) = **swimming**. The Twinning Rule states “If you are adding a suffix that starts with a vowel to a word that ends in a CVC# string that has stress on the V both before and after the suffix is added, twin the final consonant letter” – thus, *refer*, **referred**, but *reference* due to the loss of stress on *fer* when the suffix *-ence*) is added.

In **deletion** a letter is deleted when word parts combine – as in the very common **silent final <e> deletion**: *bite* plus the suffix *-ing*) = *bite+ing*) = ***biting***. In most cases, the silent final <e> is deleted before any vowel, but if the <e> is preceded by a soft <c> or <g>, the <e> is deleted only if the suffix starts with the vowel <e>, <i>, or <y>. Thus it is ***manager*** via <e> deletion – *managē+er*), but *manageable* via Simple Addition – *manage+able*), keeping the <e> to mark the soft <g>.

In **replacement**, a letter is deleted and replaced by another – as in the common <y> to <i> replacement: *try+ed*) = *try+i+ed*) = ***tried***. There is also just the reverse: an <i> to <y> replacement, as in *die+ing*) = *diē+y+ing*) = ***dying***.

Another type of replacement involves the final consonants in prefixes. In several prefixes the final consonant sounds **assimilate** – that is, change to be more similar to the following sound to ease pronunciation. And there are changes with the final consonant letter as well. For instance, in the prefix (*com-* “with, together” the [m] assimilates fully to [l] and [r] before [l] and [r], with the letter <m> changing to <l> or <r> – as in ***collect*** (*corh+l+lect*) and ***correct*** (*corh+r+rect*). Before vowels and <h> the [m] assimilates partially by deleting with no replacement in sound or letter – as in ***coerce*** (*corh+erce*) and ***cohere*** (*corh+here2*). Most often the [m] assimilates partially by being replaced with [n], with the <m> being replaced with <n>: ***conclude*** (*corh+n+clude*) – also ***conductor, conform, congestion, conjeture, conquer, consist, contact, and convince***. But because [m], [b], and [p] are very similar sounds, pronounced at the two lips, the <m> remains <m> before <m>, , and <p>, and Simple Addition prevails – as in ***common, combine, composition***.

Other prefixes in which assimilation occurs are (*ad-*, (*ex-*, (*sub-*, (*in-*1 “no, not,” (*in-*2 “in,” (*dis-*, and (*ob-*. A brief sample of words with such assimilations are *affect, allow, announce, appear, assume, attempt; suffer, suggest, suppose; illiterate; illuminate; differ, direct, distant, division; occupy, offer, opponent*.

5. Sound-to-Spelling Correspondences. It is with English sound-to-spelling correspondences that critics and skeptics grow most apoplectic.

But English spelling spells more than sounds: It also spells meanings – as in the consistent spelling of the base word *sign* in ***signed, signal, design,*** and ***designate*** in spite of marked changes in pronunciation. Though essential to the silent speed reader the consistent spelling of meanings admittedly complicates the sound-to-spelling correspondences. Additionally, English spelling indicates historical lineages, which also complicate the correspondences and account for many of the minor spellings listed in the tables below. But even with these complications there are useful patterns and generalizations that can be taught and learned. This précis treatment can only sample the considerable ruliness that is there in the correspondences.

Consider first the short vowels: The following table shows that each short vowel has one or two major spellings that account for the great majority of the spellings and some minor spellings that account for very few. Those minor spellings that occur in three or less words are marked with a superscript exclamation point. Percents are from Hanna et al. *Phoneme-Grapheme as Cues to Spelling Improvement* (Washington D.C.: GPO, 1966):

Short Vowels	Major Spellings	Minor Spellings
[ă]	bat (97%) (97%)	plaid [!] laugh [!]
[ĕ]	bed (93%) bread (4%) (97%)	again [!] says [!] friend [!] heifer [!] leopard [!] bury [!]
[ĭ]	bid (93.8%) syllable (2.5%) (96.3%)	pretty [!] been [!] counterfeit [!] sieve [!] women [!] busy [!] build [!]

Short Vowels	Major Spellings	Minor Spellings
[ɒ]	pot (79.6%) father (9.7%) (89.3%)	autumn crawl sergeant broad' cough knowledge'
[ʊ]	bud (87.6%) come (9.8%) (97.4%)	blood' cousin
[u]	pull (58.3%) wood (31%) (89.3%)	woman' could'

As an example of the rules and ruliness in consonant correspondences, the front stops [p], [b], [t], and [d] have major spellings that follow the tactical and procedural rules outlined above. Their sound-to-spelling correspondences are very simple and straightforward: Each has two major spellings that account for nearly 100% of the occurrences of the sound. The first, and far and away the most common, major spelling is the same letter that is used in square brackets to symbolize the sound: <p> for [p], for [b], <t> for [t], and <d> for [d]. The second major spelling in each case is the doublet of the first: <pp>, <bb>, <tt>, <dd>. Doublets occur at boundaries that

- (i) involve twinning: *ripping, ribbed, wetter, wedding*;
- (ii) involve a full assimilation: *appear, abbreviate, attempt*; or
- (iii) involve a simple addition that brings together two instances of the same letter, as in *dumbbell, outtalk, addict*.

Within simple words doublets also occur regularly between a preceding short vowel and a succeeding vowel letter—that is, within the VCC pattern: *pepper, cabbage, lettuce, cheddar*. They also occur regularly between short vowels and the string <le>: *ripple, dribble, little, middle*.

Everywhere else the front stops are regularly spelled with the singletons

<p>, , <t>, or <d>. We've seen that the general VCV pattern is often preempted by one of the more local shortening rules: Thus, the Third Vowel Rule leads to singleton spellings after short vowels in words like *property*, *fabulous*, *satellite*, and *federal*. The French *Lemon* Rule leads to *proper*, *cabin*, *atom*, *study*. The Suffix *-ic*) Rule leads to *parasite* vs. *parasitic* and *telescope* vs. *telescopic*.

6. Sets. Another kind of ruliness occurs in sets, of which there are three types:

(1) One type of set is defined by shared features such as bases, prefixes, suffixes, rules, procedures, or even letter strings. For instance, there is the set of words that contain the base *fect*: ***affectionate, confection, defect, effective, infectious, perfect***, and 166 others in the Lexis database. Or the set of words that contain the blend <str>: ***administration, astronaut, construct, demonstrate, illustrate, monstrous, orchestra, restriction, strain*** – and 1,540 others.

(2) Some bases function as partners, having related spellings, senses, and historical sources – for instance, *ceed* in the verb ***exceed***, and *cess* in the related noun ***excess***. The set {*ceed*, *cess*} also occurs in the verb-noun pairs ***succeed, success*** and ***proceed, process***. Closely related pairs are *cede* and *cess*, as in ***accede, access; concede, concession; intercede, intercession; precede, precession; and recede, recession***.

Another partnered pair is *sume* and *sumpt* – as in ***assume, assumption, consume, consumption, presume, presumption, resume, resumption***, etc. Yet another pair are *verse* and *vert* – as in ***avert, averse; convert, conversion; divert, diversion; invert, inversion***; etc. The Lexis database contains 885 bases that belong to such sets, providing the basis of a widespread ruliness and interconnectedness among English words.

(3) A third type of set is defined by all of the forms assumed by a certain word part. For instance, the set of all the forms of the prefix (***ad-*** includes its assimilated forms (***ac-***, (***af-***, (***ag-***, (***al-***, (***an-***, (***ap-***, (***ar-***, (***as-***, (***at-***, which would in turn define the very large set of all words containing those forms – for instance, *accelerator*, *accept*, *accurate*, *acquire*, *adagio*, *addict*, *adhesion*, *afford*, *aggravate*, *adult*, *advance*, *advice*, *ally*, *announce*, *apparatus*, *apply*, *arrange*, *assist*, *attend*, *disappoint*, *exaggerate* – and over 2,000 others.



This précis attempts to show that rather than being a disorderly chaos, our spelling system is ruled and ruly along several dimensions, historical and contemporary. Even this précis demonstrates that in those rules and that ruliness there is much that is useful to teach and learn. A précis must necessarily ignore details and complications, nearly all of which can be explained via more local rules, and when that is impossible, via historical explanations. For more information on the details and the whys and wherefores of the complications, see the introductions to the CommonWords and Lexis databases and other pieces in the Short Articles venue on this site. See also my *American English Spelling: An Informal Description* (Baltimore and London: Johns Hopkins UP, 1988).

Even Further Reading

Christopher Upward and George Davidson. *The History of English Spelling* (Malden, MA and Oxford; Wiley-Blackwell, 2011).

David Crystal. *Spell It Out: The Curious, Enthralling, and Extraordinary Story of English Spelling* (New York: St. Martins, 2012).

Edward Carney, *A Survey of English Spelling*. (London and NY: Routledge, 1994), primarily British English.

G. H. Vallins. *Spelling*, rev. D. G. Scragg (London, Andre Deutsch, 1954, 1973).

Richard Venezky. *The American Way of Spelling: The Structure and Origins of American English Orthography* (NY: Guilford, 1999).

W. A. Craigie. *English Spelling: Its Rules and Reasons* (NY: F.S. Crofts, 1927; repr. Folcroft, PA: Folcroft Library Editions, 1972).